Lecture Notes for Linear Algebra

James S. Cook Liberty University Department of Mathematics

Spring 2015

preface

Before we begin, I should warn you that I assume a few things from the reader. These notes are intended for someone who has already grappled with the problem of constructing proofs. I assume you know the difference between \Rightarrow and \Leftrightarrow . I assume the phrase "iff" is known to you. I assume you are ready and willing to do a proof by induction, strong or weak. I assume you know what \mathbb{R} , \mathbb{C} , \mathbb{Q} , \mathbb{N} and \mathbb{Z} denote. I assume you know what a subset of a set is. I assume you know how to prove two sets are equal. I assume you are familar with basic set operations such as union and intersection. More importantly, I assume you have started to appreciate that mathematics is more than just calculations. Calculations without context, without theory, are doomed to failure. At a minimum theory and proper mathematics allows you to communicate analytical concepts to other like-educated individuals.

Some of the most seemingly basic objects in mathematics are insidiously complex. We've been taught they're simple since our childhood, but as adults, mathematical adults, we find the actual definitions of such objects as \mathbb{R} or \mathbb{C} are rather involved. I will not attempt to provide foundational arguments to build numbers from basic set theory. I believe it is possible, I think it's well-thought-out mathematics, but we take the existence of the real numbers as a given truth for these notes. We assume that \mathbb{R} exists and that the real numbers possess all their usual properties. In fact, I assume \mathbb{R} , \mathbb{C} , \mathbb{Q} , \mathbb{N} and \mathbb{Z} all exist complete with their standard properties. In short, I assume we have numbers to work with. We leave the rigorization of numbers to a different course.

These notes are offered for the Spring 2015 semester at Liberty University. These are a major revision of my older linear algebra notes. They reflect the restructuring of the course which I intend for this semester. In particular, there are three main parts to this course:

- (I.) matrix theory
- (II.) abstract linear algebra
- (III.) applications (actually, we'll mostly follow Damiano and Little Chapters 4,5 and 6, we just use Chapter 8 on determinants and §11.7 on the real Jordan form in the Spring 2015 semester)

Each part is paired with a test. Each part is used to bring depth to the part which follows. Just a bit more advice before I get to the good part. How to study? I have a few points:

- spend several days on the homework. Try it by yourself to begin. Later, compare with your study group. Leave yourself time to ask questions.
- come to class, take notes, think about what you need to know to solve problems.
- assemble a list of definitions, try to gain an inuitive picture of each concept, be able to give examples and counter-examples
- learn the notation, a significant part of this course is learning to deal with new notation.
- methods of proof, how do we prove things in linear algebra? There are a few standard proofs, know them.
- method of computation, I show you tools, learn to use them.

• it's not impossible. You can do it. Moreover, doing it the right way will make the courses which follow this easier. Mathematical thinking is something that takes time for most of us to master. You began the process in Math 200 or 250, now we continue that process.

style guide

I use a few standard conventions throughout these notes. They were prepared with IAT_EX which automatically numbers sections and the hyperref package provides links within the pdf copy from the Table of Contents as well as other references made within the body of the text.

I use color and some boxes to set apart some points for convenient reference. In particular,

- 1. definitions are in green.
- 2. remarks are in red.
- 3. theorems, propositions, lemmas and corollaries are in blue.
- 4. proofs start with a **Proof:** and are concluded with a \Box .

However, I do make some definitions within the body of the text. As a rule, I try to put what I am defining in **bold**. Doubtless, I have failed to live up to my legalism somewhere. If you keep a list of these transgressions to give me at the end of the course it would be worthwhile for all involved.

The symbol \Box indicates that a proof is complete. The symbol ∇ indicates part of a proof is done, but it continues.

reading guide

A number of excellent texts have helped me gain deeper insight into linear algebra. Let me discuss a few of them here.

- 1. Damiano and Little's A Course in Linear Algebra published by Dover. I chose this as the required text in Spring 2015 as it is a well-written book, inexpensive and has solutions in the back to many exercises. The notation is fairly close to the notation used in these notes. One noted exception would be my $[T]_{\alpha,\beta}$ is replaced with $[T]^{\beta}_{\alpha}$. In fact, the notation of Damiano and Little is common in other literature I've read in higher math. I also liked the appearance of some diagrammatics for understanding Jordan forms. The section on minimal and characteristic polynomials is lucid. I think we will enjoy this book in the last third of the course.
- 2. Berberian's *Linear Algebra* published by Dover. This book is a joy. The exercises are challenging for this level and there were no solutions in the back of the text. This book is full of things I would like to cover, but, don't quite have time to do.
- 3. Takahashi and Inoue's *The Manga Guide to Linear Algebra*. Hillarious. Fun. Probably a better algorithm for Gaussian elimnation than is given in my notes.

- 4. Axler *Linear Algebra Done Right.* If our course was a bit more pure, I might use this. Very nicely written. This is an honest to goodness linear algebra text, it is actually just about the study of linear transformations on vector spaces. Many texts called "linear algebra" are really about half-matrix theory. Admittedly, such is the state of our course. But, I have no regrets, it's not as if I'm teaching matrix techinques that the students already know before this course. Ideally, I will openly admit, it would be better to have two courses. First, a course on matrices and applications. Second, a course like that outlined in this book.
- 5. Hefferon's *Linear Algebra*: this text has nice gentle introductions to many topics as well as an appendix on proof techniques. The emphasis is linear algebra and the matrix topics are delayed to a later part of the text. Furthermore, the term linear transformation as supplanted by homomorphism and there are a few other, in my view, non-standard terminologies. All in all, very strong, but we treat matrix topics much earlier in these notes. Many theorems in this set of notes were inspired from Hefferon's excellent text. Also, it should be noted the solution manual to Hefferon, like the text, is freely available as a pdf.
- 6. Anton and Rorres' *Linear Algebra: Applications Version* or Lay's Linear Algebra, or Larson and Edwards Linear Algebra, or... standard linear algebra text. Written with non-math majors in mind. Many theorems in my notes borrowed from these texts.
- 7. Insel, Spence and Friedberg's *Elementary Linear Algebra*. This text is a little light on applications in comparison to similar texts, however, the theory of Gaussian elimination and other basic algorithms are extremely clear. This text focus on column vectors for the most part.
- 8. Insel, Spence and Friedberg's *Linear Algebra*. It begins with the definition of a vector space essentially. Then all the basic and important theorems are given. Theory is well presented in this text and it has been invaluable to me as I've studied the theory of adjoints, the problem of simultaneous diagonalization and of course the Jordan and rational cannonical forms.
- 9. Strang's *Linear Algebra*. If geometric intuition is what you seek and/or are energized by then you should read this in paralell to these notes. This text introduces the dot product early on and gives geometric proofs where most others use an algebraic approach. We'll take the algebraic approach whenever possible in this course. We relegate geometry to the place of motivational side comments. This is due to the lack of prerequisite geometry on the part of a significant portion of the students who use these notes.
- 10. my advanced calculus notes. I review linear algebra and discuss multilinear algebra in some depth. I've heard from some students that they understood linear in much greater depth after the experience of my notes. Ask if interested, I'm always editing these.
- 11. Olver and Shakiban *Applied Linear Algebra*. For serious applications and an introduction to modeling this text is excellent for an engineering, science or applied math student. This book is somewhat advanced, but not as sophisticated as those further down this list.
- 12. Sadun's Applied Linear Algebra: The Decoupling Principle this is a second book in linear algebra. It presents much of the theory in terms of a unifying theme; decoupling. Probably this book is very useful to the student who wishes deeper understanding of linear system theory. Includes some Fourier analysis as well as a Chapter on Green's functions.
- 13. Curtis' Abstract Linear Algebra. Great supplement for a clean presentation of theorems. Written for math students without apology. His treatment of the wedge product as an abstract algebraic system is .

- 14. Roman's Advanced Linear Algebra. Treats all the usual topics as well as the generalization to modules. Some infinite dimensional topics are discussed. This has excellent insight into topics beyond this course.
- 15. Dummit and Foote Abstract Algebra. Part III contains a good introduction to the theory of modules. A module is roughly speaking a vector space over a ring. I believe many graduate programs include this material in their core algebra sequence. If you are interested in going to math graduate school, studying this book puts you ahead of the game a bit. Understanding Dummit and Foote by graduation is a nontrivial, but worthwhile, goal.

And now, a picture of Hannah in a shark,



I once told linear algebra that Hannah was them and my test was the shark. A wise student prayed that they all be *shark killers*. I pray the same for you this semester. I've heard from a certain student this picture and comment is unsettling. Therefore, I add this to ease the mood:



As you can see, Hannah survived to fight new monsters.

Contents

Ι	mat	trix calculation	11
1	four	ndations	13
	1.1	sets and multisets	13
	1.2	functions	16
	1.3	finite sums	19
	1.4	matrix notation	22
	1.5	vectors	24
		1.5.1 geometric preliminaries	24
		1.5.2 n-dimensional space	26
		1.5.3 concerning notation for vectors	30
2	Gau	ss-Jordan elimination	31
	2.1	systems of linear equations	31
	2.2	Gauss-Jordan algorithm	33
	2.3	classification of solutions	39
	2.4	applications to curve fitting and circuits	43
	2.5	conclusions	46
3	alge	bra of matrices	49
	3.1	addition and multiplication by scalars	49
	3.2	matrix algebra	51
	3.3	all your base are belong to us $(e_i \text{ and } E_{ij} \text{ that is}) \ldots \ldots \ldots \ldots \ldots \ldots \ldots$	57
		3.3.1 diagonal and triangular matrices have no chance survive	60
	3.4	elementary matrices	61
	3.5	invertible matrices	62
	3.6	matrix multiplication, again !	66
	3.7	how to calculate the inverse of a matrix \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots	67
		3.7.1 concatenation for solving many systems at once	68
		3.7.2 the inverse-finding algorithm	68
		3.7.3 solving systems by inverse matrix	69
	3.8	symmetric and antisymmetric matrices	70
	3.9	block matrices	72
	3.10	applications	75
	3.11	conclusions	78

4	linear independence and spanning										
	4.1	matrix notation for systems	80								
	4.2	linear combinations and spanning	82								
		4.2.1 solving several spanning questions simultaneously	85								
	4.3	linear independence	86								
	4.4	The Column Correspondence Property (CCP)	91								
	4.5	theoretical summary	93								
5	line	ar transformations of column vectors	95								
	5.1	a gallery of linear transformations	95								
	5.2	properties of linear transformations	104								
	5.3	new linear transformations from old $\ldots \ldots \ldots$	107								
		5.3.1 composition and matrix multiplication	108								
	5.4	applications	109								

II abstract linear algebra

6	vect	tor space	115
	6.1	definition and examples	116
	6.2	subspaces	120
	6.3	spanning sets and subspaces	124
	6.4	linear independence	128
	6.5	bases and dimension	132
		6.5.1 how to calculate a basis for a span of row or column vectors	135
		6.5.2 calculating basis of a solution set	138
	6.6	theory of dimensions $\ldots \ldots \ldots$	141
		6.6.1 application to fundamental matrix subspaces	147
	6.7	general theory of linear systems	148
		6.7.1 linear algebra in DEqns	150
7	abst	tract linear transformations	151
	7.1	basic terminology	152
	7.2	theory of linear transformations	153
	7.3	matrix of linear transformation	160
	7.4	coordinate change	166
		7.4.1 coordinate change of abstract vectors	166
		7.4.2 coordinate change for column vectors	167
		7.4.3 coordinate change of abstract linear transformations	169
		7.4.4 coordinate change of linear transformations of column vectors	172
	7.5	theory of dimensions for maps	173
	7.6	quotient space	177
		7.6.1 the first isomorphism theorem	180
	7.7	structure of subspaces	183
	7.8	examples of isomorphisms	187

113

III applications

8	dete	erminants	193
	8.1	a criteria for invertibility	193
	8.2	determinants and geometry	197
	8.3	cofactor expansion for the determinant	199
	8.4	properties of determinants	203
	8.5	examples of determinants	208
	8.6	Cramer's Rule	210
	8.7	adjoint matrix	214
	8.8	applications	216
	8.9	similarity and determinants for linear transformations	219
	8.10	conclusions	221
a	ouel	idean geometry	9 93
9	0 1	Fuclidean geometry of \mathbb{R}^n	225
	0.2	\mathbb{D} arthogonality in \mathbb{R}^n	220
	9.2 9.3	orthogonal complements and projections	223
	9.9 Q /	orthogonal transformations and geometry	$\frac{200}{245}$
	9.5	least squares analysis	$\frac{240}{247}$
	5.0	9.5.1 the closest vector problem	211
		9.5.2 inconsistent equations	240
		9.5.2 the least squares problem	251
	96	inner products	251 257
	5.0	9.6.1 examples of inner-products	259
		9.6.2 Fourier analysis	$\frac{263}{262}$
	9.7	orthogonal matrices and the QR factorization	264
10			0.0 5
10	com	piex vectorspaces	267
	10.1	10.0.1 concerning matrices and vectors with complex entries	207
	10.1		208
11	eige	nvalues and eigenvectors	271
	11.1	why eigenvectors?	272
		11.1.1 quantum mechanics	272
		11.1.2 stochastic matrices	272
		11.1.3 motion of points under linear transformations	274
	11.2	basic theory of eigenvectors	278
	11.3	complex eigenvalues and vectors	284
	11.4	examples of real and complex eigenvectors	288
		11.4.1 characteristic equations	288
		11.4.2 real eigenvector examples	289
		11.4.3 complex eigenvector examples	292
	11.5	eigenbases and eigenspaces	295
	11.6	generalized eigenvectors	301
	11.7	real Jordan form	305
	11.8	eigenvectors and orthogonality	309
	11.9	select applications	310

11.9.1 linear differential equations and e-vectors: diagonalizable case
11.9.2 linear differential equations and e-vectors: non-diagonalizable case $\ldots \ldots 314$
adratic forms 315
1 conic sections and quadric surfaces
2 quadratic forms and their matrix
12.2.1 summary of quadratic form analysis
3 Taylor series for functions of two or more variables
12.3.1 deriving the two-dimensional Taylor formula
12.3.2 examples
4 intertia tensor, an application of quadratic forms
stems of differential equations 333
1 calculus of matrices
2 introduction to systems of linear differential equations
3 the matrix exponential
13.3.1 analysis for matrices
13.3.2 formulas for the matrix exponential
4 solutions for systems of DEqns with real eigenvalues
5 solutions for systems of DEqns with complex eigenvalues
6 geometry and difference equations revisited
L

Part I

matrix calculation

Chapter 1

foundations

In this chapter we settle some basic notational issues. There are not many examples in this chapter and the main task the reader is assigned here is to read and learn the definitions and notations.

1.1 sets and multisets

A set is a collection of objects. The set with no elements is called the **empty-set** and is denoted \emptyset . If we write $x \in A$ then this is read "x is an element of A". In your previous course you learned that $\{a, a, b\} = \{a, b\}$. In other words, there is no allowance for repeats of the same object. In linear algebra, we often find it more convenient to use what is known as a **multiset**. In other instances we'll make use of an **ordered set** or even an **ordered multiset**. To summarize:

- 1. a set is a collection of objects with no repeated elements in the collection.
- 2. a **multiset** is a collection of objects. Repeats are possible.
- 3. an **ordered set** is a collection of objects with no repeated elements in which the collection has a specific ordering.
- 4. an **ordered multiset** is a collection of objects which has an ordering and possibly has repeated elements.

Notice, every set is a multiset and every ordered set is an ordered multiset. In the remainder of this course, we make the slight abuse of langauge and agree to call an ordinary set a **set with no repeated elements** and a multiset will simply be called in sequel a **set**. This simplifies our langauge and will help us to think better¹.

Let us denote sets by capital letters in as much as is possible. Often the lower-case letter of the same symbol will denote an element; $a \in A$ is to mean that the object a is in the set A. We can abbreviate $a_1 \in A$ and $a_2 \in A$ by simply writing $a_1, a_2 \in A$, this is a standard notation. The **union** of two sets A and B is denoted² $A \cup B = \{x | x \in A \text{ or } x \in B\}$. The **intersection** of two sets is

¹there is some substructure to describe here, multisets and ordered sets can be constructed from sets. However, that adds little to our discussion and so I choose to describe multisets, ordered sets and soon Cartesian products formally. Formally, means I describe there structure without regard to its explicit concrete realization.

²note that $S = \{x \in R : x \text{ meets condition } P\} = \{x \in R \mid x \text{ meets condition } P\}$. Some authors use : whereas I prefer to use | in the set-builder notation.

denoted $A \cap B = \{x | x \in A \text{ and } x \in B\}$. It sometimes convenient to use unions or intersections of several sets:

$$\bigcup_{\alpha \in \Lambda} U_{\alpha} = \{x \mid \text{there exists } \alpha \in \Lambda \text{ with } x \in U_{\alpha}\}$$
$$\bigcap_{\alpha \in \Lambda} U_{\alpha} = \{x \mid \text{for all } \alpha \in \Lambda \text{ we have } x \in U_{\alpha}\}$$

we say Λ is the **index set** in the definitions above. If Λ is a finite set then the union/intersection is said to be a finite union/interection. If Λ is a countable set then the union/intersection is said to be a countable union/interection³.

Suppose A and B are both sets then we say A is a **subset** of B and write $A \subseteq B$ iff $a \in A$ implies $a \in B$ for all $a \in A$. If $A \subseteq B$ then we also say B is a **superset** of A. If $A \subseteq B$ then we say $A \subset B$ iff $A \neq B$ and $A \neq \emptyset$. Recall, for sets A, B we define A = B iff $a \in A$ implies $a \in B$ for all $a \in A$ and conversely $b \in B$ implies $b \in A$ for all $b \in B$. This is equivalent to insisting A = B iff $A \subseteq B$ and $B \subseteq A$. Note, if we deal with ordered sets equality is measured by checking that both sets contain the same elements in the same order. The **difference** of two sets A and B is denoted A - B and is defined by $A - B = \{a \in A \mid \text{ such that } a \notin B\}^4$.

A Cartesian product of two sets A, B is the set of ordered pairs (a, b) where $a \in A$ and $b \in B$. We denote,

$$A\times B=\{(a,b)\ |\ a\in A,\ b\in B\}$$

Likewise, we define

$$A \times B \times C = \{(a, b, c) \mid a \in A, b \in B, c \in C\}$$

We make no distinction between $A \times (B \times C)$ and $(A \times B) \times C$. This means we are using the obvious one-one correspondence $(a, (b, c)) \leftrightarrow ((a, b), c)$. If A_1, A_2, \ldots, A_n are sets then we define $A_1 \times A_2 \times \cdots \times A_n$ to be the set of ordered *n*-tuples:

$$\prod_{i=1}^{n} A_i = A_1 \times \dots \times A_n = \{(a_1, \dots, a_n) \mid a_i \in A_i \text{ for all } i \in \mathbb{N}_n\}$$

Notice, I define $\mathbb{N} = \{1, 2, ...\}$ as the set of **natural numbers** whereas \mathbb{N}_n is the set of natural numbers up to and including $n \in \mathbb{N}$; $\mathbb{N}_n = \{1, ..., n\}$. If we take the Cartesian product of a set A with itself *n*-times then it is customary to denote the set of all *n*-tuples from A as A^n :

$$\underbrace{A \times \cdots \times A}_{n-copies} = A^n.$$

Real numbers can be constructed from set theory and about a semester of mathematics. We will accept the following as \mathbf{axioms}^5

7

³recall the term **countable** simply means there exists a bijection to the natural numbers. The cardinality of such a set is said to be \aleph_o

⁴other texts somtimes use $A - B = A \setminus B$

⁵an axiom is a basic belief which cannot be further reduced in the conversation at hand. If you'd like to see a construction of the real numbers from other math, see Ramanujan and Thomas' *Intermediate Analysis* which has the construction both from the so-called Dedekind cut technique and the Cauchy-class construction. Also, I've been informed, Terry Tao's Analysis I text has a very readable exposition of the construction from the Cauchy viewpoint.

Definition 1.1.1. real numbers

The set of real numbers is denoted \mathbb{R} and is defined by the following axioms:

- (A1) addition commutes; a + b = b + a for all $a, b \in \mathbb{R}$.
- (A2) addition is associative; (a + b) + c = a + (b + c) for all $a, b, c \in \mathbb{R}$.
- (A3) zero is additive identity; a + 0 = 0 + a = a for all $a \in \mathbb{R}$.
- (A4) additive inverses; for each $a \in \mathbb{R}$ there exists $-a \in \mathbb{R}$ and a + (-a) = 0.
- (A5) multiplication commutes; ab = ba for all $a, b \in \mathbb{R}$.
- (A6) multiplication is associative; (ab)c = a(bc) for all $a, b, c \in \mathbb{R}$.
- (A7) one is multiplicative identity; a1 = a for all $a \in \mathbb{R}$.
- (A8) multiplicative inverses for nonzero elements; for each $a \neq 0 \in \mathbb{R}$ there exists $\frac{1}{a} \in \mathbb{R}$ and $a\frac{1}{a} = 1$.
- (A9) distributive properties; a(b+c) = ab + ac and (a+b)c = ac + bc for all $a, b, c \in \mathbb{R}$.
- (A10) totally ordered field; for $a, b \in \mathbb{R}$:
 - (i) antisymmetry; if $a \leq b$ and $b \leq a$ then a = b.
 - (ii) transitivity; if $a \leq b$ and $b \leq c$ then $a \leq c$.
 - (iii) totality; $a \leq b$ or $b \leq a$
- (A11) least upper bound property: every nonempty subset of \mathbb{R} that has an upper bound, has a least upper bound. This makes the real numbers **complete**.

Modulo A11 and some math jargon this should all be old news. An **upper bound** for a set $S \subseteq \mathbb{R}$ is a number $M \in \mathbb{R}$ such that M > s for all $s \in S$. Similarly a lower bound on S is a number $m \in \mathbb{R}$ such that m < s for all $s \in S$. If a set S is bounded above and below then the set is said to be **bounded**. For example, the open set (a, b) is bounded above by b and it is bounded below by a. In contrast, rays such as $(0, \infty)$ are not bounded above. Closed intervals contain their least upper bound and greatest lower bound. The bounds for an open interval are outside the set.

We often make use of the following standard sets:

- natural numbers (positive integers); $\mathbb{N} = \{1, 2, 3, \dots\}$.
- natural numbers up to the number n; $\mathbb{N}_n = \{1, 2, 3, \dots, n-1, n\}$.
- integers; $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$. Note, $\mathbb{Z}_{>0} = \mathbb{N}$.
- non-negative integers; $\mathbb{Z}_{\geq 0} = \{0, 1, 2, \dots\} = \mathbb{N} \cup \{0\}.$
- negative integers; $\mathbb{Z}_{<0} = \{-1, -2, -3, \dots\} = -\mathbb{N}.$
- rational numbers; $\mathbb{Q} = \{ \frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0 \}.$

- irrational numbers; $\mathbb{J} = \{x \in \mathbb{R} \mid x \notin \mathbb{Q}\}.$
- open interval from a to b; $(a, b) = \{x | a < x < b\}.$
- half-open interval; $(a, b] = \{x \mid a < x \le b\}$ or $[a, b) = \{x \mid a \le x < b\}$.
- closed interval; $[a, b] = \{x \mid a \le x \le b\}.$

We define $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$. I refer to \mathbb{R}^2 as "R-two" in conversational mathematics. Likewise, "R-three" is defined by $\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$. We are ultimately interested in studying "R-n" where $\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R} \text{ for } i = 1, 2, \ldots, n\}$. In this course if we consider \mathbb{R}^m it is assumed from the context that $m \in \mathbb{N}$.

In terms of cartesian products you can imagine the x-axis as the number line then if we paste another numberline at each x value the union of all such lines constructs the plane; this is the picture behind $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. Another interesting cartesian product is the **unit-square**; $[0,1]^2 =$ $[0,1] \times [0,1] = \{(x,y) \mid 0 \le x \le 1, 0 \le y \le 1\}$. Sometimes a rectangle in the plane with it's edges included can be written as $[x_1, x_2] \times [y_1, y_2]$. If we want to remove the edges use $(x_1, x_2) \times (y_1, y_2)$.

Moving to three dimensions we can construct the **unit-cube** as $[0,1]^3$. A generic rectangular solid can sometimes be represented as $[x_1, x_2] \times [y_1, y_2] \times [z_1, z_2]$ or if we delete the edges: $(x_1, x_2) \times (y_1, y_2) \times (z_1, z_2)$.

1.2 functions

Suppose A and B are sets, we say $f : A \to B$ is a **function** if for each $a \in A$ the function f assigns a single element $f(a) \in B$. Moreover, if $f : A \to B$ is a function we say it is a *B*-valued function of an *A*-variable and we say A = dom(f) whereas B = codomain(f). For example, if $f : \mathbb{R}^2 \to [0, 1]$ then f is real-valued function of \mathbb{R}^2 . On the other hand, if $f : \mathbb{C} \to \mathbb{R}^2$ then we'd say f is a vector-valued function of a complex variable. The term **mapping** will be used interchangeably with function in these notes. Suppose $f : U \to V$ and $U \subseteq S$ and $V \subseteq T$ then we may consisely express the same data via the notation $f : U \subseteq S \to V \subseteq T$.

Definition 1.2.1.

Suppose $f: U \to V$. We define the **image of** U_1 **under** f as follows:

 $f(U_1) = \{ y \in V \mid \text{there exists } x \in U_1 \text{ with } f(x) = y \}.$

The range of f is f(U). The inverse image of V_1 under f is defined as follows:

$$f^{-1}(V_1) = \{ x \in U \mid f(x) \in V_1 \}.$$

The inverse image of a single point in the codomain is called a **fiber**. Suppose $f: U \to V$. We say f is surjective or onto V_1 iff there exists $U_1 \subseteq U$ such that $f(U_1) = V_1$. If a function is onto its codomain then the **function is surjective**. If $f(x_1) = f(x_2)$ implies $x_1 = x_2$ for all $x_1, x_2 \in U_1 \subseteq U$ then we say **f** is injective on U_1 or 1 - 1 on U_1 . If a function is injective on its domain then we say the **function is injective**. If a function is both injective and surjective then the function is called a **bijection** or a **1-1 correspondance**. **Example 1.2.2.** Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ and f(x, y) = x for each $(x, y) \in \mathbb{R}^2$. The function is not injective since f(1,2) = 1 and f(1,3) = 1 and yet $(1,2) \neq (1,3)$. Notice that the fibers of f are simply vertical lines:

$$f^{-1}(x_o) = \{(x, y) \in dom(f) \mid f(x, y) = x_o\} = \{(x_o, y) \mid y \in \mathbb{R}\} = \{x_o\} \times \mathbb{R}$$

Example 1.2.3. Suppose $f : \mathbb{R} \to \mathbb{R}$ and $f(x) = \sqrt{x^2} + 1$ for each $x \in \mathbb{R}$. This function is not surjective because $0 \notin f(\mathbb{R})$. In contrast, if we construct $g : \mathbb{R} \to [1, \infty)$ with g(x) = f(x) for each $x \in \mathbb{R}$ then can argue that g is surjective. Neither f nor g is injective, the fiber of x_0 is $\{-x_0, x_0\}$ for each $x_0 \neq 0$. At all points except zero these maps are said to be **two-to-one**. This is an abbreviation of the observation that two points in the domain map to the same point in the range.

Example 1.2.4. Suppose $f : \mathbb{R}^3 \to \mathbb{R}^2$ and $f(x, y, z) = (x^2 + y^2, z)$ for each $(x, y, z) \in \mathbb{R}^3$. You can easily see that $range(f) = [0, \infty] \times \mathbb{R}$. Suppose $R^2 \in [0, \infty)$ and $z_o \in \mathbb{R}$ then

$$f^{-1}(\{(R^2, z_o)\}) = S_1(R) \times \{z_o\}$$

where $S_1(R)$ denotes a circle of radius R. This result is a simple consequence of the observation that $f(x, y, z) = (R^2, z_o)$ implies $x^2 + y^2 = R^2$ and $z = z_o$.

Function composition is one important way to construct new functions. If $f: U \to V$ and $g: V \to W$ then $g \circ f: U \to W$ is the composite of g with f. We also create new functions by extending or restricting domains of given functions. In particular:

Definition 1.2.5.

Let $f: U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping. If $R \subset U$ then we define the **restriction of** f **to** R to be the mapping $f|_R: R \to V$ where $f|_R(x) = f(x)$ for all $x \in R$. If $U \subseteq S$ and $V \subset T$ then we say a mapping $g: S \to T$ is an **extension of** f iff $g|_{dom(f)} = f$.

When I say $g|_{dom(f)} = f$ this means that these functions have matching domains and they agree at each point in that domain; $g|_{dom(f)}(x) = f(x)$ for all $x \in dom(f)$. Once a particular subset is chosen the restriction to that subset is a unique function. Of course there are usually many subbets of dom(f) so you can imagine many different restictions of a given function. The concept of extension is more vague, once you pick the enlarged domain and codomain it is not even necessarily the case that another extension to that same pair of sets will be the same mapping. To obtain uniqueness for extensions one needs to add more stucture. This is one reason that complex variables are interesting, there are cases where the structure of the complex theory forces the extension of a complex-valued function on a one-dimensional subset of \mathbb{C} of a complex variable to be unique. This is very surprising. An even stronger result is available for a special type of function called a linear transformation. We'll see that a linear transformation is uniquely defined by its values on a basis. This means that a linear transformation is uniquely extended from a zero-dimensional subset of a vector space⁶.

Definition 1.2.6.

Let $f: U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping, if there exists a mapping $g: f(U) \to U$ such that $f \circ g = Id_{f(U)}$ and $g \circ f = Id_U$ then g is the inverse mapping of f and we denote $g = f^{-1}$.

⁶technically, we don't know what this word "dimension" means just yet. Or linear transformation, or vector space, all in good time...

If a mapping is injective then it can be shown that the inverse mapping is well defined. We define $f^{-1}(y) = x$ iff f(x) = y and the value x must be a single value if the function is one-one. When a function is not one-one then there may be more than one point which maps to a particular point in the range.

Notice that the inverse image of a set is well-defined even if there is no inverse mapping. Moreover, it can be shown that the fibers of a mapping are disjoint and their union covers the domain of the mapping:

$$f(y) \neq f(z) \quad \Rightarrow \quad f^{-1}\{y\} \cap f^{-1}\{z\} = \emptyset \qquad \qquad \bigcup_{y \in range(f)} f^{-1}\{y\} = dom(f).$$

This means that the fibers of a mapping *partition* the domain.

Example 1.2.7. Consider $f(x, y) = x^2 + y^2$ this describes a mapping from \mathbb{R}^2 to \mathbb{R} . Observe that $f^{-1}\{\mathbb{R}^2\} = \{x^2 + y^2 = \mathbb{R}^2 \mid (x, y) \in \mathbb{R}^2\}$. In words, the nonempty fibers of f are concentric circles about the origin and the origin itself.

Technically, the emptyset is always a fiber. It is the fiber over points in the codomain which are not found in the range. In the example above, $f^{-1}(-\infty, 0) = \emptyset$. Perhaps, even from our limited array of examples, you can begin to appreciate there is a unending array of possible shapes, curves, volumes and higher-dimensional objects which can appear as fibers. In contrast, as we will prove later in this course, the inverse image of any linear transformation is essentially⁷ a line, plane or *n*-volume containing the origin.

Definition 1.2.8.

Let $f: U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping. A **cross section** of the fiber partition is a subset $S \subseteq U$ for which $S \cap f^{-1}\{v\}$ contains a single element for every $v \in f(U)$.

How do we construct a cross section for a particular mapping? For particular examples the details of the formula for the mapping usually suggests some obvious choice. However, in general if you accept the **axiom of choice** then you can be comforted in the existence of a cross section even in the case that there are infinitely many fibers for the mapping. In this course, we'll see later that the problem of constructing a cross-section for a linear mapping is connected to the problem of finding a representative for each point in the quotient space of the mapping.

Example 1.2.9. An easy cross-section for $f(x, y) = x^2 + y^2$ is given by any ray eminating from the origin. Notice that, if $ab \neq 0$ then $S = \{t(a, b) \mid t \in [0, \infty)\}$ interects the *a* circle of radius $R^2 = t^2(a^2 + b^2)$ at the point (ta, tb)

Proposition 1.2.10.

Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping. The restriction of f to a cross section S of U is an injective function. The mapping $\tilde{f} : U \to f(U)$ is a surjection. The mapping $\tilde{f}|_S : S \to f(U)$ is a bijection.

The proposition above tells us that we can take any mapping and cut down the domain and/or codomain to give the modified function the structure of an injection, surjection or even a bijection.

⁷up to an isomorphism which is roughly speaking a change of notation

Example 1.2.11. Continuing with our example, $f : \mathbb{R}^2 \to \mathbb{R}$ with $f(x, y) = x^2 + y^2$ is neither surjective or injective. However, just to make a choice, $S = \{(t,0) \mid t \in [0,\infty)\}$ then clearly $\tilde{f}: S \to [0,\infty)$ defined by $\tilde{f}(x,y) = f(x,y)$ for all $(x,y) \in S$ is a bijection.

Definition 1.2.12.

Let $f: U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping then we say a mapping g is a **local inverse** of f iff there exits $S \subseteq U$ such that $g = (f|_S)^{-1}$.

Usually we can find local inverses for functions in calculus. For example, $f(x) = \sin(x)$ is not 1-1 therefore it is not invertible. However, it does have a local inverse $g(y) = \sin^{-1}(y)$. If we were more pedantic we wouldn't write $\sin^{-1}(y)$. Instead we would write $g(y) = \left(\sin|_{[-\pi/2, \pi/2]}\right)^{-1}(y)$ since the inverse sine is actually just a local inverse. To construct a local inverse for some mapping we must locate some subset of the domain upon which the mapping is injective. Then relative to that subset we can reverse the mapping. I mention this concept in passing so you may appreciate its absense from this course. In linear algebra, the existence of a local inverse for a linear transformation will imply the existence of a global inverse. The case we study in this course is very special. We provide the bedrock on which other courses form arguments. Calculus linearizes problems locally, so, to understand local problems we must first understand linear problems. That is our task this semester, to unravel the structure of linear transformations as deeply as we dare.

1.3 finite sums

In this section we introduce a nice notation for finite sums⁸ of arbitrary size. Most of these statements are "for all $n \in \mathbb{N}$ " thus proof by mathematical induction is the appropriate proof tool. I offer a few sample arguments and leave the rest to the reader. Let's begin by giving a precise definition for the finite sum $A_1 + A_2 + \cdots + A_n$:

Definition 1.3.1.

Let
$$A_i$$
 for $i = 1, 2, ..., n$ be objects which allow addition. We recursively define:

$$\sum_{i=1}^{n+1} A_i = A_{n+1} + \sum_{i=1}^n A_i$$
for each $n \ge 1$ and $\sum_{i=1}^1 A_i = A_1$.

The "summation notation" or "sigma" notation allows us to write sums precisely. In $\sum_{i=1}^{n} A_i$ the index *i* is called the **dummy index of summation**. One dummy is just a good as the next, it follows that $\sum_{i=1}^{n} A_i = \sum_{i=j}^{n} A_j$. This relabeling is sometimes called *switching dummy variables*, or *switching the index of summation from i to j*. The terms which are summed in the sum are called **summands**. For the sake of specificity I will assume real summands for the remainder of this section. It should be noted the arguments given here generalize with little to no work for a wide variety of other spaces where addition and multiplication by numbers is well-defined⁹.

⁸the results of this section apply to objects which allow addition and multiplication by numbers, it is quite general ⁹in the middle part of this course we learn such spaces are called vector spaces

Proposition 1.3.2.

Let $A_i, B_i \in \mathbb{R}$ for each $i \in \mathbb{N}$ and suppose $c \in \mathbb{R}$ then for each $n \in \mathbb{N}$,

(1.)
$$\sum_{i=1}^{n} (A_i + B_i) = \sum_{i=1}^{n} A_i + \sum_{i=1}^{n} B_i$$

(2.) $\sum_{i=1}^{n} cA_i = c \sum_{i=1}^{n} A_i.$

Proof: Let's begin with (1.). Notice the claim is trivially true for n = 1. Inductively assume that (1.) is true for $n \in \mathbb{N}$. Consider, the following calculations are justified either from the recursive definition of the finite sum or the induction hypothesis:

$$\sum_{i=1}^{n+1} (A_i + B_i) = \sum_{i=1}^n (A_i + B_i) + A_{n+1} + B_{n+1}$$
$$= \left(\sum_{i=1}^n A_i + \sum_{i=1}^n B_i\right) + A_{n+1} + B_{n+1}$$
$$= \left(\sum_{i=1}^n A_i\right) + A_{n+1} + \left(\sum_{i=1}^n B_i\right) + B_{n+1}$$
$$= \sum_{i=1}^{n+1} A_i + \sum_{i=1}^{n+1} B_i.$$

Thus (1.) is true for n + 1 and hence by proof by mathematical induction (PMI) we find (1.) is true for all $n \in \mathbb{N}$. The proof of (2.) is similar. \Box

Proposition 1.3.3.

Let
$$A_i, B_{ij} \in \mathbb{R}$$
 for $i, j \in \mathbb{N}$ and suppose $c \in \mathbb{R}$ then for each $n \in \mathbb{N}$,

$$(1.) \sum_{i=1}^n \left(\sum_{j=1}^n B_{ij}\right) = \sum_{j=1}^n \left(\sum_{i=1}^n B_{ij}\right).$$

$$(2.) \sum_{i=1}^n \sum_{j=1}^n A_i B_{ij} = \sum_{i=1}^n A_i \sum_{j=1}^n B_{ij}$$

Proof: The proof of (1.) proceeds by induction on n. If n = 1 then there is only one possible term, namely B_{11} and the sums trivially agree. Consider the n = 2 case as we prepare for the induction step,

$$\sum_{i=1}^{2} \sum_{j=1}^{2} B_{ij} = \sum_{i=1}^{2} [B_{i1} + B_{i2}] = [B_{11} + B_{12}] + [B_{21} + B_{22}]$$

On the other hand,

$$\sum_{j=1}^{2} \sum_{i=1}^{2} B_{ij} = \sum_{j=1}^{2} [B_{1j} + B_{2j}] = [B_{11} + B_{21}] + [B_{11} + B_{21}].$$

The sums in opposite order produce the same terms overall, however the ordering of the terms may differ¹⁰. Fortunately, real number-addition commutes.

Assume inductively that (1.) is true for some n > 1. Using the definition of sum throughout and the induction hypothesis in transitioning from the 3-rd to the 4-th line:

$$\sum_{i=1}^{n+1} \sum_{j=1}^{n+1} B_{ij} = \sum_{i=1}^{n+1} \left[B_{i,n+1} + \sum_{j=1}^{n} B_{ij} \right]$$

$$= \sum_{i=1}^{n+1} B_{i,n+1} + \sum_{i=1}^{n+1} \sum_{j=1}^{n} B_{ij}$$

$$= \sum_{i=1}^{n+1} B_{i,n+1} + \sum_{j=1}^{n} B_{n+1,j} + \sum_{i=1}^{n} \sum_{j=1}^{n} B_{ij}$$

$$= \sum_{i=1}^{n+1} B_{i,n+1} + \sum_{j=1}^{n} B_{n+1,j} + \sum_{j=1}^{n} \sum_{i=1}^{n} B_{ij}$$

$$= \sum_{i=1}^{n+1} B_{i,n+1} + \sum_{j=1}^{n} \left[B_{n+1,j} + \sum_{i=1}^{n} B_{ij} \right]$$

$$= \sum_{i=1}^{n+1} B_{i,n+1} + \sum_{j=1}^{n} \sum_{i=1}^{n+1} B_{ij}$$

$$= \sum_{i=1}^{n+1} \sum_{i=1}^{n+1} B_{ij}$$

Thus n implies n + 1 for (1.) therefore by proof by mathematical induction we find (1.) is true for all $n \in \mathbb{N}$. In short, we can swap the order of finite sums. The proof of (2.) involves similar induction arguments. \Box

From (1.) of the above proposition we find that multiple summations may be listed in any order. Moreover, a notation which indicates multiple sums is unambiguous:

$$\sum_{i,j=1}^{n} A_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij}.$$

If we have more than two summations nested the same result holds. Therefore, define:

$$\sum_{i_1,\dots,i_k=1}^n A_{i_1\dots i_k} = \sum_{i_1=1}^n \dots \sum_{i_k=1}^n A_{i_1\dots i_k}.$$

Remark 1.3.4.

The purpose of this section is primarily notational. I want you to realize what is behind the notation and it is likely I assign some homework based on utilizing the recursive definition given here. I usually refer to the results of this section as " properties of finite sums".

¹⁰reordering terms in the infinite series case can get you into trouble if you don't have absolute convergence. Riemann showed a conditionally convergent series can be reordered to force it to converge to any value you might choose.

1.4 matrix notation

Matrices can be constructed from set-theoretic arguments in much the same way as Cartesian Products. I will not pursue those matters in these notes. We will assume that everyone understands how to construct an array of numbers.

Definition 1.4.1.

An $m \times n$ matrix is an array of objects with m rows and n columns. The elements in the array are called entries or components. If A is an $m \times n$ matrix then A_{ij} denotes the object in the *i*-th row and the *j*-th column. The label *i* is a row index and the index *j* is a column index in the preceding sentence. We usually denote $A = [A_{ij}]$. The set $m \times n$ of matrices with real number entries is denoted $\mathbb{R}^{m \times n}$. The set of $m \times n$ matrix with complex entries is $\mathbb{C}^{m \times n}$. Generally, is *S* is a set then $S^{m \times n}$ is the set of $m \times n$ arrays of objects from *S*. If a matrix has the same number of rows and columns then it is called a **square matrix**.

Example 1.4.2. Suppose

$$A = \left[\begin{array}{rrr} 1 & 2 & 3 \\ 4 & 5 & 6 \end{array} \right].$$

We see that A has 2 rows and 3 columns thus $A \in \mathbb{R}^{2\times 3}$. Moreover, $A_{11} = 1$, $A_{12} = 2$, $A_{13} = 3$, $A_{21} = 4$, $A_{22} = 5$, and $A_{23} = 6$. It's not usually possible to find a formula for a generic element in the matrix, but this matrix satisfies $A_{ij} = 3(i-1) + j$ for all i, j.

In the statement "for all i, j" it is to be understood that those indices range over their allowed values. In the preceding example $1 \le i \le 2$ and $1 \le j \le 3$.

Example 1.4.3. Let S be a set of cats. If $A \in S^{2 \times 2}$ then A_{ij} is a cat for all i, j.

Definition 1.4.4.

Two matrices A and B are equal if and only if they have the same size and $A_{ij} = B_{ij}$ for all i, j.

If you studied vectors before you should identify this is precisely the same rule we used in calculus III¹¹ Two vectors were equal iff all the components matched. Vectors are just specific cases of matrices so the similarity is not surprising.

Example 1.4.5. Solve A = B where $A = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$ and $B = \begin{bmatrix} x^2 & 3 \\ 3y & w \end{bmatrix}$. Observe, A = B iff the following four equations are true:

$$x = x^2, y = 3, z = 3y, w = w$$

We can solve these by algebra. Of course, $x^2 = x$ implies x(x-1) = 0 hence x = 0 or x = 1. The y equation is easy to solve and thus z = 3(3) = 9. Finally, the only equation for w is w = w hence there is no restriction on w, it is a **free variable**. The solution as a set is given by

$$\{(x,3,9,w) \mid x=0,1 \ w \in \mathbb{R}\}.$$

¹¹I wrote a special subsection to help you see the geometry of vectors if you didn't get a chance to see it already in another course.

1.4. MATRIX NOTATION

Definition 1.4.6.

Let $A \in \mathbb{R}^{m \times n}$ then a submatrix of A is a matrix which is made of some rectangle of elements in A. Rows and columns are submatrices. In particular,

1. An $m \times 1$ submatrix of A is called a column vector of A. The *j*-th column vector is denoted $col_j(A)$ and $(col_j(A))_i = A_{ij}$ for $1 \le i \le m$. In other words,

$$col_{k}(A) = \begin{bmatrix} A_{1k} \\ A_{2k} \\ \vdots \\ A_{mk} \end{bmatrix} \Rightarrow A = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix} = [col_{1}(A)|col_{2}(A)|\cdots|col_{n}(A)]$$

2. An $1 \times n$ submatrix of A is called a row vector of A. The *i*-th **row vector** is denoted $row_i(A)$ and $(row_i(A))_j = A_{ij}$ for $1 \le j \le n$. In other words,

$$row_{k}(A) = \begin{bmatrix} A_{k1} & A_{k2} & \cdots & A_{kn} \end{bmatrix} \Rightarrow A = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix} = \begin{bmatrix} row_{1}(A) \\ \hline row_{2}(A) \\ \hline \vdots \\ \hline row_{m}(A) \end{bmatrix}$$

Suppose $A \in \mathbb{R}^{m \times n}$, note for $1 \leq j \leq n$ we have $col_j(A) \in \mathbb{R}^{m \times 1}$ whereas for $1 \leq i \leq m$ we find $row_i(A) \in \mathbb{R}^{1 \times n}$. In other words, an $m \times n$ matrix has n columns of length m and n rows of length m.

Example 1.4.7. Suppose $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$. The columns of A are,

$$col_1(A) = \begin{bmatrix} 1\\4 \end{bmatrix}$$
, $col_2(A) = \begin{bmatrix} 2\\5 \end{bmatrix}$, $col_3(A) = \begin{bmatrix} 3\\6 \end{bmatrix}$.

The rows of A are

$$row_1(A) = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$
, $row_2(A) = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}$

Definition 1.4.8.

Let $A \in \mathbb{R}^{m \times n}$ then $A^T \in \mathbb{R}^{n \times m}$ is called the **transpose** of A and is defined by $(A^T)_{ji} = A_{ij}$ for all $1 \le i \le m$ and $1 \le j \le n$.

Example 1.4.9. Suppose $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ then $A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$. Notice that $row_1(A) = col_1(A^T), \ row_2(A) = col_2(A^T)$

and

$$col_1(A) = row_1(A^T), \ col_2(A) = row_2(A^T), \ col_3(A) = row_3(A^T)$$

Notice $(A^T)_{ij} = A_{ji} = 3(j-1) + i$ for all i, j; at the level of index calculations we just switch the indices to create the transpose.

The preceding example shows us that we can quickly create the transpose of a given matrix by switching rows to columns. The transpose of a row vector is a column vector and vice-versa.

1.5 vectors

The first subsection in this section is intended to introduce the reader to the concept of geometric vectors. I show that both vector addition and scalar multiplication naturally flow from intuitive geometry. Then we abstract those concepts in the second subsection to give concrete definitions of vector addition and scalar multiplication in \mathbb{R}^n .

1.5.1 geometric preliminaries

The concept of a vector is almost implicit with the advent of Cartesian geometry. Rene Descartes' great contribution was the realization that geometry had an algebraic description if we make an identification of points in the plane with pairs of real numbers. This identification is so ubiquitious it is hard to imagine the plane without imagining pairs of numbers. Euclid had no idea of x or y coordinates, instead just lines, circles and constructive axioms. Analytic geometry is the study of geometry as formulated by Descartes. Because numbers are identified with points we are able to state equations expressing relations between points. For example, if $h, k, R \in \mathbb{R}$ then the set of all points $(x, y) \in \mathbb{R}^2$ which satisfy

$$(x-h)^2 + (y-k)^2 = R^2$$

is a circle of radius R centered at (h, k). We can analyze the circle by studying the algebra of the equation above. In calculus we even saw how implicit differentiation reveals the behaviour of the tangent lines to the circle.

Very well, what about the points themselves ? What relations if any do arbitrary points in the plane admit? For one, you probably already know about how to get directed line segments from points. A common notation in highschool geometry¹² is that the line from point $P = (Q_1, Q_2)$ to another point $Q = (Q_1, Q_2)$ is \overrightarrow{PQ} where we define:

$$\overrightarrow{PQ} = Q - P = (Q_1 - P_1, Q_2 - P_2).$$

A directed line-segment is also called a **vector**¹³.



¹²a dying subject apparently

¹³however, not every vector in this course is a directed line segment.

1.5. VECTORS

Consider a second line segment going from Q to $R = (R_1, R_2)$ this gives us the directed line segment of $\overrightarrow{QR} = R - Q = (R_1 - Q_1, R_2 - Q_2)$. What then about the directed line segment from the original point P to the final point R? How is $\overrightarrow{PR} = R - P = (R_1 - P_1, R_2 - P_2)$ related to \overrightarrow{PQ} and \overrightarrow{QR} ? Suppose we define addition of points in the same way we defined the subtraction of points:

$$(V_1, V_2) + (W_1, W_2) = (V_1 + W_1, V_2 + W_2).$$

Will this definition be consistent with the geometrically suggested result $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$? Consider,

$$\overrightarrow{PQ} + \overrightarrow{QR} = (Q_1 - P_1, Q_2 - P_2) + (R_1 - Q_1, R_2 - Q_2)$$
$$= (Q_1 - P_1 + R_1 - Q_1, Q_2 - P_2 + R_2 - Q_2)$$
$$= (R_1 - P_1, R_2 - P_2)$$
$$= \overrightarrow{PR}.$$

We find the addition and subtraction of directed line segments is consistent with the usual tip-tail addition of vectors in the plane.



What else can we do ? It seems natural to assume that $\overrightarrow{PQ} + \overrightarrow{PQ} = 2\overrightarrow{PQ}$ but what does multiplication by a number mean for a vector? What definition should we propose? Note if $\overrightarrow{PQ} = (Q_1 - P_1, Q_2 - P_2)$ then $\overrightarrow{PQ} + \overrightarrow{PQ} = 2\overrightarrow{PQ}$ implies $2(\overrightarrow{PQ}) = (2(Q_1 - P_1), 2(Q_2 - P_2))$. Therefore, we define for $c \in \mathbb{R}$,

$$c(V_1, V_2) = (cV_1, cV_2).$$

This definition is naturally consistent with the definition we made for addition. We can understand multiplication of a vector by a number as an operation which **scales** the vector. In other words, multiplying a vector by a number will change the length of the vector. Multiplication of a vector by a number is often called **scalar multiplication**. Scalars are numbers.



Vectors based at the origin are naturally identified with points: the directed line segment from Q = (0,0) to P is naturally identified with the point P.

$$\overrightarrow{QP} = (P_1, P_2) - (0, 0) = (P_1, P_2).$$

In other words we can identify the point $P = (P_1, P_2)$ with the directed line segment from the origin $\vec{P} = (P_1, P_2)$. Unless context suggests otherwise vectors in this course are assumed to be based at the origin.

1.5.2 n-dimensional space

Two dimensional space is $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. To obtain *n*-dimensional space we just take the Cartesian product of *n*-copies of \mathbb{R} .

Definition 1.5.1.

Let $n \in \mathbb{N}$, we define $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) \mid x_j \in \mathbb{R} \text{ for } j = 1, 2, \dots, n\}$. If $v \in \mathbb{R}^n$ then we say v is an **n-vector**. The numbers in the vector are called the **components**; $v = (v_1, v_2, \dots, v_n)$ has *j*-th component v_j .

Notice, a consequence of the definition above and the construction of the Cartesian product¹⁴ is that two vectors v and w are equal iff $v_j = w_j$ for all j. Equality of two vectors is only true if all components are found to match. Addition and scalar multiplication are naturally generalized from the n = 2 case. I use $e_1 = (1, 0)$ and $e_2 = (0, 1)$ for illustration below:



¹⁴see my Math 200 notes or ask me if interested, it's not entirely trivial

Definition 1.5.2.

Define functions $+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and $\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ by the following rules: for each $v, w \in \mathbb{R}^n$ and $c \in \mathbb{R}$:

(1.) $(v+w)_j = v_j + w_j$ (2.) $(cv)_j = cv_j$

for all $j \in \{1, 2, ..., n\}$. The operation + is called **vector addition** and it takes two vectors $v, w \in \mathbb{R}^n$ and produces another vector $v + w \in \mathbb{R}^n$. The operation \cdot is called **scalar multiplication** and it takes a number $c \in \mathbb{R}$ and a vector $v \in \mathbb{R}^n$ and produces another vector $c \cdot v \in \mathbb{R}^n$. Often we simply denote $c \cdot v$ by juxtaposition cv.

If you are a gifted at visualization then perhaps you can add three-dimensional vectors in your mind. If you're mind is really unhinged maybe you can even add 4 or 5 dimensional vectors. The beauty of the definition above is that we have no need of pictures. Instead, algebra will do just fine. That said, let's draw another picture, I already showed how we can write a two dimensional vector as a sum of $e_1 = (1, 0)$ and $e_2 = (0, 1)$ on the previous page.



Notice these pictures go to show how you can break-down vectors into component vectors which point in the direction of the coordinate axis. In \mathbb{R}^3 we have $e_1 = (1,0,0)$, $e_2 = (0,1,0)$ and $e_3 = (0,0,1)$. Vectors of length¹⁵ one which point in the coordinate directions make up what is called the **standard basis**¹⁶ It is convenient to define special notation to describe the standard basis in arbitrary dimension. First I define a useful shorthand,

Definition 1.5.3.

The symbol
$$\delta_{ij} = \begin{cases} 1 & , i = j \\ 0 & , i \neq j \end{cases}$$
 is called the **Kronecker delta**.

For example, $\delta_{22} = 1$ while $\delta_{12} = 0$.

¹⁵the length of vectors is an important concept which we mine in depth later in the course

¹⁶for now we use the term "basis" without meaning, in Chapter 5 we make a great effort to refine the concept.

Definition 1.5.4.

Let $e_i \in \mathbb{R}^{n \times 1}$ be defined by $(e_i)_j = \delta_{ij}$. The size of the vector e_i is determined by context. We call e_i the *i*-th standard basis vector.

Example 1.5.5. Let me expand on what I mean by "context" in the definition above: In \mathbb{R} we have $e_1 = (1) = 1$ (by convention we drop the brackets in this case) In \mathbb{R}^2 we have $e_1 = (1,0)$ and $e_2 = (0,1)$. In \mathbb{R}^3 we have $e_1 = (1,0,0)$ and $e_2 = (0,1,0)$ and $e_3 = (0,0,1)$. In \mathbb{R}^4 we have $e_1 = (1,0,0,0)$ and $e_2 = (0,1,0,0)$ and $e_3 = (0,0,1,0)$ and $e_4 = (0,0,0,1)$.

Example 1.5.6. Any vector in \mathbb{R}^n can be written as a sum of these basic vectors. For example,

$$v = (1, 2, 3) = (1, 0, 0) + (0, 2, 0) + (0, 0, 3)$$

= 1(1, 0, 0) + 2(0, 1, 0) + 3(0, 0, 1)
= $e_1 + 2e_2 + 3e_3$.

We say that v is a linear combination of e_1, e_2 and e_3 .

The concept of a linear combination is very important.

Definition 1.5.7.

A linear combination of objects
$$A_1, A_2, \ldots, A_k$$
 is a sum
 $c_1A_1 + c_2A_2 + \cdots + c_kA_k = \sum_{i=1}^k c_iA_i$
where $c_i \in \mathbb{R}$ for each i .

We will look at linear combinations of vectors, matrices and even functions in this course. If $c_i \in \mathbb{C}$ then we call it a *complex linear combination*. The proposition below generalizes the calculation from Example 1.5.6.

Proposition 1.5.8.

Every vector in \mathbb{R}^n is a linear combination of e_1, e_2, \ldots, e_n .

Proof: Let $v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$. By the definition of vector addition:

$$v = (v_1, v_2, \dots, v_n)$$

= $(v_1, 0, \dots, 0) + (0, v_2, \dots, v_n)$
= $(v_1, 0, \dots, 0) + (0, v_2, \dots, 0) + \dots + (0, 0, \dots, v_n)$
= $(v_1, 0 \cdot v_1, \dots, 0 \cdot v_1) + (0 \cdot v_2, v_2, \dots, 0 \cdot v_2) + \dots + (0 \cdot v_n, 0 \cdot v_n, \dots, v_n)$

In the last step I rewrote each zero to emphasize that the each entry of the k-th summand has a v_k factor. Continue by applying the definition of scalar multiplication to each vector in the sum above we find,

$$v = v_1(1, 0, \dots, 0) + v_2(0, 1, \dots, 0) + \dots + v_n(0, 0, \dots, 1)$$

= $v_1e_1 + v_2e_2 + \dots + v_ne_n$.

Therefore, every vector in \mathbb{R}^n is a linear combination of e_1, e_2, \ldots, e_n . For each $v \in \mathbb{R}^n$ we have $v = \sum_{i=1}^n v_n e_n$. \Box

Proposition 1.5.9. the vector properties of \mathbb{R}^n .

Suppose $n \in \mathbb{N}$. For all $x, y, z \in \mathbb{R}^n$ and $a, b \in \mathbb{R}$,

1. (P1) x + y = y + x for all $x, y \in \mathbb{R}^n$,

- 2. (P2) (x+y) + z = x + (y+z) for all $x, y, z \in \mathbb{R}^n$,
- 3. (P3) there exists $0 \in \mathbb{R}^n$ such that x + 0 = x for all $x \in \mathbb{R}^n$,
- 4. (P4) for each $x \in \mathbb{R}^n$ there exists $-x \in \mathbb{R}^n$ such that x + (-x) = 0,
- 5. (P5) $1 \cdot x = x$ for all $x \in \mathbb{R}^n$,
- 6. (P6) $(ab) \cdot x = a \cdot (b \cdot x)$ for all $x \in \mathbb{R}^n$ and $a, b \in \mathbb{R}$,
- 7. (P7) $a \cdot (x+y) = a \cdot x + a \cdot y$ for all $x, y \in \mathbb{R}^n$ and $a \in \mathbb{R}$,
- 8. (P8) $(a+b) \cdot x = a \cdot x + b \cdot x$ for all $x \in \mathbb{R}^n$ and $a, b \in \mathbb{R}$,
- 9. (P9) If $x, y \in \mathbb{R}^n$ then x + y is a single element in \mathbb{R}^n , (we say \mathbb{R}^n is closed with respect to addition)
- 10. (P10) If $x \in \mathbb{R}^n$ and $c \in \mathbb{R}$ then $c \cdot x$ is a single element in \mathbb{R}^n . (we say \mathbb{R}^n is closed with respect to scalar multiplication)

We call 0 in P3 the **zero vector** and the vector -x is called the **additive inverse** of x. We will usually omit the \cdot and instead denote scalar multiplication by juxtaposition; $a \cdot x = ax$.

Proof: all the properties follow immediately from the definitions of addition and scalar multiplication in \mathbb{R}^n as well as properties of real numbers. Consider,

$$(x+y)_j = x_j + y_j = \underbrace{y_j + x_j}_{\star} = (y+x)_j$$

where \star follows because real number addition commutes. Since the calculation above holds for each j = 1, 2, ..., n it follows that x + y = y + x for all $x, y \in \mathbb{R}^n$ hence P1 is true. Very similarly P2 follows from associativity of real number addition. To prove P3 simply define, as usual, $0_j = 0$; The zero vector is the vector with all zero components. Note

$$(x+0)_j = x_j + 0_j = x_j + 0 = x_j$$

which holds for all j = 1, 2, ..., n hence x + 0 = x for all $x \in \mathbb{R}^n$. I leave the remainder of the properties for the reader. \Box

The preceding proposition will be mirrored in an abstract context later in the course. So, it is important. On the other hand, we will prove it again in the next chapter in the context of a subcase of the matrix algebra. I include it here to complete the logic of this chapter.

1.5.3 concerning notation for vectors

Definition 1.5.10. points are viewed as column vectors in this course.

In principle one can use column vectors for everything or row vectors for everything. I choose a subtle notation that allows us to use both. On the one hand it is nice to write vectors as rows since the typesetting is easier. However, once you start talking about matrix multiplication then it is natural to write the vector to the right of the matrix and we will soon see that the vector should be written as a column vector for that to be reasonable. Therefore, we adopt the following **convention**

$$(v_1, v_2, \dots, v_n) = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

If I want to denote a real row vector then we will just write $[v_1, v_2, \ldots, v_n]$.

The convention above is by no means universal. Various linear algebra books deal with this notational dilemma and number of different ways. In the first version of my linear algebra notes I used $\mathbb{R}^{n\times 1}$ everywhere just to be relentlessly explicit that we were using column vectors for \mathbb{R}^n . The set of all $n \times 1$ matrices is the set of all column vectors which I denote by $\mathbb{R}^{n\times 1}$ whereas the set of all $1 \times n$ matrices is the set of all row vectors which we denote by $\mathbb{R}^{1\times n}$. We discuss these matters in general in next chapter. The following example is merely included to expand on the notation.

Example 1.5.11. Suppose x + y + z = 3, x + y = 2 and x - y - z = -1. This system can be written as a single vector equation by simply stacking these equations into a column vector:

$$\begin{bmatrix} x+y+z\\x+y\\x-y-z \end{bmatrix} = \begin{bmatrix} 3\\2\\-1 \end{bmatrix}$$

Furthermore, we can break up the vector of variables into linear combination where the coefficients in the sum are the variables x, y, z:

$$x \begin{bmatrix} 1\\1\\1 \end{bmatrix} + y \begin{bmatrix} 1\\1\\-1 \end{bmatrix} + z \begin{bmatrix} 1\\0\\-1 \end{bmatrix} = \begin{bmatrix} 3\\2\\-1 \end{bmatrix}$$

Note that the solution to the system is x = 1, y = 1, z = 1.

Chapter 2

Gauss-Jordan elimination

Gauss-Jordan elimination is an optimal method for solving a system of linear equations. Logically it may be equivalent to methods you are already familar with but the matrix notation is by far the most efficient method. This is important since throughout this course we will be faced with the problem of solving linear equations. Additionally, the Gauss-Jordan produces the *reduced row echelon form*(rref) of the matrix. Given a particular matrix the rref is unique. This is of particular use in theoretical applications.

2.1 systems of linear equations

Let me begin with a few examples before I state the general definition.

Example 2.1.1. Consider the following system of 2 equations and 2 unknowns,

$$x + y = 2$$
$$x - y = 0$$

Adding equations reveals 2x = 2 hence x = 1. Then substitute that into either equation to deduce y = 1. Hence the solution (1,1) is unique

Example 2.1.2. Consider the following system of 2 equations and 2 unknowns,

$$x + y = 2$$
$$3x + 3y = 6$$

We can multiply the second equation by 1/3 to see that it is equivalent to x + y = 2 thus our two equations are in fact the same equation. There are infinitely many equations of the form (x, y) where x + y = 2. In other words, the solutions are (x, 2 - x) for all $x \in \mathbb{R}$.

Both of the examples thus far were **consistent**.

Example 2.1.3. Consider the following system of 2 equations and 2 unknowns,

$$x + y = 2$$
$$x + y = 3$$

These equations are inconsistent. Notice substracting the second equation yields that 0 = 1. This system has no solutions, it is inconsistent

It is remarkable that these three simple examples reveal the general structure of solutions to linear systems. Either we get a unique solution, infinitely many solutions or no solution at all. For our examples above, these cases correspond to the possible graphs for a pair of lines in the plane. A pair of lines may intersect at a point (unique solution), be the same line (infinitely many solutions) or be paralell (inconsistent).¹



Remark 2.1.4.

It is understood in this course that i, j, k, l, m, n, p, q, r, s are in N. I will not belabor this point. Please ask if in doubt.

Definition 2.1.5. system of *m*-linear equations in *n*-unknowns

Let x_1, x_2, \ldots, x_n be n variables and suppose $b_i, A_{ij} \in \mathbb{R}$ for $1 \le i \le m$ and $1 \le j \le n$ then $A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n = b_1$ $A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n = b_2$ $\vdots \quad \vdots \quad \vdots \quad \vdots$ $A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n = b_m$ in called a system of linear equations. If $b_1 = 0$ for $1 \le i \le m$ there exists the system of the system of

is called a system of linear equations. If $b_i = 0$ for $1 \le i \le m$ then we say the system is **homogeneous**. The solution set is the set of all $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ which satisfy all the equations in the system simultaneously.

Remark 2.1.6.

We use variables x_1, x_2, \ldots, x_n mainly for general theoretical statements. In particular problems and especially for applications we tend to defer to the notation x, y, z etc...

Definition 2.1.7.

The augmented coefficient matrix is an array of numbers which provides an abbreviated notation for a system of linear equations.

$\begin{bmatrix} A_{11}x_1 + A_{12}x_2 + \dots + A_{1n}x_n = b_1 \\ A_{21}x_1 + A_{22}x_2 + \dots + A_{2n}x_n = b_2 \end{bmatrix}$	11 • 4 11	$\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}$	$\begin{array}{c} A_{12} \\ A_{22} \end{array}$	 	$\begin{array}{c} A_{1n} \\ A_{2n} \end{array}$	$egin{array}{c} b_1 \ b_2 \end{array}$	
$\begin{bmatrix} \vdots \vdots \vdots \vdots \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \dots + A_{mn}x_n = b_m \end{bmatrix}$	abbreviated by	\vdots A_{m1}	\vdots A_{m2}	:	\vdots A_{mn}	$\vdots \\ b_m$	

¹I used the *Graph* program to generate these graphs. It makes nice pictures, these are ugly due to user error.

The vertical bar is optional, I include it to draw attention to the distinction between the matrix of coefficients A_{ij} and the nonhomogeneous terms b_i . Let's revisit my three simple examples in this new notation. I illustrate the Gauss-Jordan method for each.

Example 2.1.8. The system x + y = 2 and x - y = 0 has augmented coefficient matrix:

$$\begin{bmatrix} 1 & 1 & | & 2 \\ 1 & -1 & | & 0 \end{bmatrix} \xrightarrow{r_2 - r_1 \to r_2} \begin{bmatrix} 1 & 1 & | & 2 \\ 0 & -2 & | & -2 \end{bmatrix}$$

$$\xrightarrow{r_2/-2 \to r_2} \begin{bmatrix} 1 & 1 & | & 2 \\ 0 & 1 & | & 1 \end{bmatrix} \xrightarrow{r_1 - r_2 \to r_1} \begin{bmatrix} 1 & 0 & | & 1 \\ 0 & 1 & | & 1 \end{bmatrix}$$

The last augmented matrix represents the equations x = 1 and y = 1. Rather than adding and subtracting equations we added and subtracted rows in the matrix. Incidentally, the last step is called the **backward pass** whereas the first couple steps are called the **forward pass**. Gauss is credited with figuring out the forward pass then Jordan added the backward pass. Calculators can accomplish these via the commands ref (Gauss' row echelon form) and rref (Jordan's reduced row echelon form). In particular,

$$ref\begin{bmatrix} 1 & 1 & | & 2\\ 1 & -1 & | & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & | & 2\\ 0 & 1 & | & 1 \end{bmatrix} \qquad rref\begin{bmatrix} 1 & 1 & | & 2\\ 1 & -1 & | & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & | & 1\\ 0 & 1 & | & 1 \end{bmatrix}$$

Example 2.1.9. The system x + y = 2 and 3x + 3y = 6 has augmented coefficient matrix:

$$\begin{bmatrix} 1 & 1 & 2 \\ 3 & 3 & 6 \end{bmatrix} \xrightarrow{r_2 - 3r_1 \to r_2} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

The nonzero row in the last augmented matrix represents the equation x + y = 2. In this case we cannot make a backwards pass so the ref and rref are the same.

Example 2.1.10. The system x + y = 3 and x + y = 2 has augmented coefficient matrix:

$$\begin{bmatrix} 1 & 1 & | & 3 \\ 1 & 1 & | & 2 \end{bmatrix} \xrightarrow{r_2 - 3r_1 \to r_2} \begin{bmatrix} 1 & 1 & | & 1 \\ 0 & 0 & | & 1 \end{bmatrix}$$

The last row indicates that 0x+0y = 1 which means that there is no solution since $0 \neq 1$. Generally, when the bottom row of the rref(A|b) is zeros with a 1 in the far right column then the system Ax = b is inconsistent because there is no solution to the equation.

2.2 Gauss-Jordan algorithm

To begin we need to identify three basic operations we do when solving systems of equations. I'll define them for system of 3 equations and 3 unknowns, but it should be obvious this generalizes to m equations and n unknowns without much thought.

The following operations are called **Elementary Row Operations**.

(1.)	scaling	row	1 by	v nonzero	$\operatorname{constant}$	0
------	---------	-----	------	-----------	---------------------------	---

$\begin{bmatrix} A_{11} & A_{12} & A_{13} & b_1 \\ A_{21} & A_{22} & A_{23} & b_2 \\ A_{31} & A_{32} & A_{33} & b_3 \end{bmatrix} \xrightarrow{cr_1 \to r_1} \begin{bmatrix} A_{11} & A_{12} & A_{13} & b_1 \\ A_{21} & A_{22} & A_{23} & b_2 \\ A_{31} & A_{32} & A_{33} & b_3 \end{bmatrix}$
--

(2.) replace row 1 with the sum of row 1 and row 2

$\int A_{11}$	A_{12}	A_{13}	b_1		$A_{11} + A_{21}$	$A_{12} + A_{22}$	$A_{13} + A_{23}$	$b_1 + b_2$
A_{21}	A_{22}	A_{23}	b_2	$r_1 + r_2 \to r_1$	A_{21}	A_{22}	A_{23}	b_2
A_{31}	A_{32}	A_{33}	b_3	/	A_{31}	A_{32}	A_{33}	b_3

(3.) swap rows 1 and 2

$\begin{bmatrix} A_{11} \end{bmatrix}$	A_{12}	A_{13}	b_1		A_{21}	A_{22}	A_{23}	b_2
A_{21}	A_{22}	A_{23}	b_2	$r_1 \longleftrightarrow r_2$	A_{11}	A_{12}	A_{13}	b_1
A_{31}	A_{32}	A_{33}	b_3		A_{31}	A_{32}	A_{33}	b_3

I illustrate how to use these elementary row operations to simplify a given matrix in the example below. The matrix in the example corresponds to equations x + 2y - 3z = 1, 2x + 4y = 7 and -x + 3y + 2z = 0.

Example 2.2.1. Given
$$A = \begin{bmatrix} 1 & 2 & -3 & 1 \\ 2 & 4 & 0 & 7 \\ -1 & 3 & 2 & 0 \end{bmatrix}$$
 calculate $rref(A)$.

$$A = \begin{bmatrix} 1 & 2 & -3 & | & 1 \\ 2 & 4 & 0 & | & 7 \\ -1 & 3 & 2 & | & 0 \end{bmatrix} \xrightarrow{r_2 - 2r_1 \to r_2} \begin{bmatrix} 1 & 2 & -3 & | & 1 \\ 0 & 0 & 6 & | & 5 \\ -1 & 3 & 2 & | & 0 \end{bmatrix} \xrightarrow{r_1 + r_3 \to r_3} \xrightarrow{r_1 + r_3 \to r_3} \begin{bmatrix} 1 & 2 & -3 & | & 1 \\ 0 & 0 & 6 & | & 5 \\ 0 & 5 & -1 & | & 1 \end{bmatrix} \xrightarrow{r_2 \leftrightarrow r_3} \begin{bmatrix} 1 & 2 & -3 & | & 1 \\ 0 & 5 & -1 & | & 1 \\ 0 & 0 & 6 & | & 5 \end{bmatrix} = ref(A)$$

that completes the forward pass. We begin the backwards pass,

$$ref(A) = \begin{bmatrix} 1 & 2 & -3 & | & 1 \\ 0 & 5 & -1 & | & 1 \\ 0 & 0 & 6 & | & 5 \end{bmatrix} \xrightarrow{r_3 \leftarrow \frac{1}{6}r_3} \begin{bmatrix} 1 & 2 & -3 & | & 1 \\ 0 & 5 & -1 & | & 1 \\ 0 & 0 & 1 & | & 5/6 \end{bmatrix} \xrightarrow{r_2 + r_3 \leftarrow r_2}$$
$$\begin{bmatrix} 1 & 2 & -3 & | & 1 \\ 0 & 5 & 0 & | & 11/6 \\ 0 & 0 & 1 & | & 5/6 \end{bmatrix} \xrightarrow{r_1 + 3r_3 \leftarrow r_1} \begin{bmatrix} 1 & 2 & 0 & | & 21/6 \\ 0 & 5 & 0 & | & 11/6 \\ 0 & 0 & 1 & | & 5/6 \end{bmatrix} \xrightarrow{r_1 + 3r_3 \leftarrow r_1} \begin{bmatrix} 1 & 2 & 0 & | & 21/6 \\ 0 & 5 & 0 & | & 11/6 \\ 0 & 0 & 1 & | & 5/6 \end{bmatrix} \xrightarrow{r_1 - 2r_2 \leftarrow r_1} \begin{bmatrix} 1 & 0 & 0 & | & 83/30 \\ 0 & 1 & 0 & | & 11/30 \\ 0 & 0 & 1 & | & 5/6 \end{bmatrix} \xrightarrow{r_1 - 2r_2 \leftarrow r_1} \begin{bmatrix} 1 & 0 & 0 & | & 83/30 \\ 0 & 1 & 0 & | & 11/30 \\ 0 & 0 & 1 & | & 5/6 \end{bmatrix} = rref(A)$$

Thus, we've found the system of equations x + 2y - 3z = 1, 2x + 4y = 7 and -x + 3y + 2z = 0 has solution x = 83/30, y = 11/30 and z = 5/6.

The method used in the example above generalizes to matrices of any size. It turns out that by making a finite number of the operations (1.),(2.) and (3.) we can reduce the matrix to the particularly simple format called the "reduced row echelon form" (I abbreviate this rref most places). The Gauss-Jordan algorithm tells us which order to make these operations. The following definition is borrowed from the text *Elementary Linear Algebra: A Matrix Approach*, 2nd ed. by Spence, Insel and Friedberg, however you can probably find nearly the same algorithm in dozens of other texts.

Definition 2.2.2. Gauss-Jordan Algorithm.

Given an m by n matrix A the following sequence of steps is called the Gauss-Jordan algorithm or Gaussian elimination. I define terms such as **pivot column** and **pivot position** as they arise in the algorithm below.

- **Step 1:** Determine the leftmost nonzero column. This is a **pivot column** and the topmost position in this column is a **pivot position**.
- Step 2: Perform a row swap to bring a nonzero entry of the pivot column below the pivot row to the top position in the pivot column (in the first step there are no rows above the pivot position, however in future iterations there may be rows above the pivot position, see 4).
- **Step 3:** Add multiples of the pivot row to create zeros **below** the pivot position. This is called "clearing out the entries below the pivot position".
- **Step 4:** If there is a nonzero row below the pivot row from (3.) then find the next pivot postion by looking for the next nonzero column to the right of the previous pivot column. Then perform steps 1-3 on the new pivot column. When no more nonzero rows below the pivot row are found then go on to step 5.
- **Step 5:** the leftmost entry in each nonzero row is called the **leading entry**. Scale the bottommost nonzero row to make the leading entry 1 and use row additions to clear out any remaining nonzero entries **above** the leading entries.
- **Step 6:** If step 5 was performed on the top row then stop, otherwise apply Step 5 to the next row up the matrix.

Steps (1.)-(4.) are called the **forward pass**. A matrix produced by a foward pass is called the reduced echelon form of the matrix and it is denoted ref(A). Steps (5.) and (6.) are called the **backwards pass**. The matrix produced by completing Steps (1.)-(6.) is called the reduced row echelon form of A and it is denoted rref(A).

The ref(A) is not unique because there may be multiple choices for how Step 2 is executed. On the other hand, it turns out that rref(A) is unique. The proof of uniqueness can be found in Appendix E of Insel Spence and Friedberg's elementary linear algebra text. The backwards pass takes the ambiguity out of the algorithm. Notice the forward pass goes down the matrix while the backwards pass goes up the matrix. **Example 2.2.3.** Given $A = \begin{bmatrix} 1 & -1 & 1 \\ 3 & -3 & 0 \\ 2 & -2 & -3 \end{bmatrix}$ calculate rref(A).

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 3 & -3 & 0 \\ 2 & -2 & -3 \end{bmatrix} \xrightarrow{r_2 - 3r_1 \to r_2} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & -3 \\ 2 & -2 & -3 \end{bmatrix} \xrightarrow{r_3 - 2r_1 \to r_3} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & -3 \\ 0 & 0 & -5 \end{bmatrix} \xrightarrow{\frac{3r_3 \to r_3}{5r_2 \to r_2}} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & -15 \\ 0 & 0 & -15 \end{bmatrix} \xrightarrow{\frac{r_3 - r_2 \to r_3}{\frac{-1}{15}r_2 \to r_2}} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & -15 \\ \frac{1}{15}r_2 \to r_2 \to r_2 \to r_2 \to r_3 \to r_$$

Note it is customary to read multiple row operations from top to bottom if more than one is listed between two of the matrices. The multiple arrow notation should be used with caution as it has great potential to confuse. Also, you might notice that I did not strictly-speaking follow Gauss-Jordan in the operations $3r_3 \rightarrow r_3$ and $5r_2 \rightarrow r_2$. It is sometimes convenient to modify the algorithm slightly in order to avoid fractions.

Example 2.2.4. Find the rref of the matrix A given below:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 1 & 1 \end{bmatrix} \underbrace{r_2 - r_1 \to r_2}_{2} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & -2 & 0 & -1 & 0 \\ -1 & 0 & 1 & 1 & 1 \end{bmatrix} \underbrace{r_3 + r_1 \to r_3}_{1 \to r_1 \to r_2}$$
$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 2 & 2 \\ 0 & -2 & 0 & -1 & 0 \end{bmatrix} \underbrace{r_3 + 2r_2 \to r_3}_{2 \to r_2 \to r_2}$$
$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 2 & 2 \\ 0 & -2 & 0 & -1 & 0 \end{bmatrix} \underbrace{r_3 + 2r_2 \to r_3}_{1 \to r_1 \to r_1}$$
$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 2 & 2 \\ 0 & 0 & 4 & 3 & 4 \end{bmatrix} \underbrace{r_1 \to r_1}_{2r_2 \to r_2} \begin{bmatrix} 4 & 4 & 4 & 4 & 4 \\ 0 & 2 & 4 & 4 & 4 \\ 0 & 0 & 4 & 3 & 4 \end{bmatrix} \underbrace{r_2 - r_3 \to r_2}_{r_1 - r_3 \to r_1}$$
$$\begin{bmatrix} 4 & 4 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 4 & 3 & 4 \end{bmatrix} \underbrace{r_1 - 2r_2 \to r_1}_{0 \to 1} \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 4 & 3 & 4 \end{bmatrix} \underbrace{r_1 - 2r_2 \to r_1}_{r_2 \to r_2} \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 4 & 3 & 4 \end{bmatrix} \underbrace{r_1 - 2r_2 \to r_1}_{r_2 \to r_2} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 4 & 3 & 4 \end{bmatrix} \underbrace{r_1 - 2r_2 \to r_1}_{r_2 \to r_2} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 4 & 3 & 4 \end{bmatrix} \underbrace{r_1 - 2r_2 \to r_1}_{r_2 \to r_2} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 4 & 3 & 4 \end{bmatrix} \underbrace{r_1 - 2r_2 \to r_1}_{r_2 \to r_2} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 3 & 4 \end{bmatrix} \underbrace{r_2 - r_3 \to r_2}_{r_3 \to r_2} \underbrace{r_3 \to r_2}_{r_3 \to r_3}$$
Example 2.2.5.

$$\begin{bmatrix} A|I \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & | & 1 & 0 & 0 \\ 2 & 2 & 0 & | & 0 & 1 & 0 \\ 4 & 4 & 4 & | & 0 & 0 & 1 \end{bmatrix} \xrightarrow{r_2 - 2r_1 \to r_2}_{r_3 - 4r_1 \to r_3}$$

$$\begin{bmatrix} 1 & 0 & 0 & | & 1 & 0 & 0 \\ 0 & 2 & 0 & | & -2 & 1 & 0 \\ 0 & 4 & 4 & | & -4 & 0 & 1 \end{bmatrix} \xrightarrow{r_3 - 2r_2 \to r_3}_{r_3 - 2r_2 \to r_3} \begin{bmatrix} 1 & 0 & 0 & | & 1 & 0 & 0 \\ 0 & 2 & 0 & | & -2 & 1 & 0 \\ 0 & 0 & 4 & | & 0 & -2 & 1 \end{bmatrix} \xrightarrow{r_2/2 \to r_2}_{r_3/4 \to r_3}$$

$$\begin{bmatrix} 1 & 0 & 0 & | & 1 & 0 & 0 \\ 0 & 1 & 0 & | & -1 & 1/2 & 0 \\ 0 & 0 & 1 & | & 0 & -1/2 & 1/4 \end{bmatrix} = rref[A|I]$$

Example 2.2.6. *easy examples are sometimes disquieting, let* $r \in \mathbb{R}$ *,*

$$v = \begin{bmatrix} 2 & -4 & 2r \end{bmatrix} \xrightarrow{\frac{1}{2}r_1 \to r_1} \begin{bmatrix} 1 & -2 & r \end{bmatrix} = rref(v)$$

Example 2.2.7. here's another next to useless example,

$$v = \begin{bmatrix} 0\\1\\3 \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{bmatrix} 1\\0\\3 \end{bmatrix} \xrightarrow{r_3 - 3r_1 \to r_3} \begin{bmatrix} 1\\0\\0 \end{bmatrix} = rref(v)$$

Example 2.2.8.

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 3 & 2 & 0 & 0 \end{bmatrix} \xrightarrow{r_4 - 3r_1 \to r_4} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 2 & -3 & 0 \end{bmatrix} \xrightarrow{r_4 - r_2 \to r_4} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & -3 & 0 \end{bmatrix} \xrightarrow{r_4 - r_2 \to r_4} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & -3 & 0 \end{bmatrix} \xrightarrow{r_4 + r_3 \to r_4} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{r_2/2 \to r_2} \xrightarrow[r_3/3 \to r_3]{r_3 \to r_3} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{r_1 - r_3 \to r_1} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = rref(A)$$

I should remind you that there are numerous online resources to help you become efficient in your row reduction. I provide links in Blackboard and on my website.

Proposition 2.2.9.

If a particular column of a matrix is all zeros then it will be unchanged by the Gaussian elimination. Additionally, if we know rref(A) = B then rref[A|0] = [B|0] where 0 denotes one or more columns of zeros.

Proof: adding nonzero multiples of one row to another will result in adding zero to zero in the column. Likewise, if we multiply a row by a nonzero scalar then the zero column is uneffected. Finally, if we swap rows then this just interchanges two zeros. Gauss-Jordan elimination is just a finite sequence of these three basic row operations thus the column of zeros will remain zero as claimed. \Box

Example 2.2.10. Use Example 2.2.3 and Proposition 2.2.9 to calculate,

	1	0	1	0	0 -	1	0	0	0	0]	
nn c f	0	2	0	0	0	0	1	0	0	0	
rrej	0	0	3	1	0	0	0	1	0	0	
	3	2	0	0	0	0	0	0	1	0	

Similarly, use Example 2.2.7 and Proposition 2.2.9 to calculate:

	1	0	0	0		1	0	0	0]
rref	0	0	0	0	=	0	0	0	0
	3	0	0	0		0	0	0	0

I hope these examples suffice. One last advice, you should think of the Gauss-Jordan algorithm as a sort of road-map. It's ok to take detours to avoid fractions and such but the end goal should remain in sight. If you lose sight of that it's easy to go in circles. Incidentally, I would strongly recommend you find a way to check your calculations with technology. Mathematica will do any matrix calculation we learn. TI-85 and higher will do much of what we do with a few exceptions here and there. There are even websites which will do row operations, I provide a link on the course website. All of this said, I would remind you that I expect you be able perform Gaussian elimination correctly and quickly on the test and quizzes without the aid of a graphing calculator for the remainder of the course. The arithmetic matters. Unless I state otherwise it is expected you show the details of the Gauss-Jordan elimination in any system you solve in this course.

Theorem 2.2.11.

Let $A \in \mathbb{R}^{m \times n}$ then if R_1 and R_2 are both Gauss-Jordan eliminations of A then $R_1 = R_2$. In other words, the reduced row echelon form of a matrix of real numbers is unique.

Proof: The proof of uniqueness can be found in Appendix E of Insel Spence and Friedberg's elementary linear algebra text. It is straightforward, but a bit tedious. \Box

2.3 classification of solutions

Surprisingly Examples 2.1.8,2.1.9 and 2.1.10 illustrate all the possible types of solutions for a linear system. In this section I interpret the calculations of the last section as they correspond to solving systems of equations.

Example 2.3.1. Solve the following system of linear equations if possible,

$$x + 2y - 3z = 1$$

$$2x + 4y = 7$$

$$-x + 3y + 2z = 0$$

We solve by doing Gaussian elimination on the augmented coefficient matrix (see Example 2.2.1 for details of the Gaussian elimination),

	1	2	-3	1		1	0	0	83/30		x = 83/30
rref	2	4	0	7	=	0	1	0	11/30	\Rightarrow	y = 11/30
		3	2	0		0	0	1	5/6		z = 5/6

(We used the results of Example 2.2.1).

Remark 2.3.2.

The geometric interpretation of the last example is interesting. The equation of a plane with normal vector $\langle a, b, c \rangle$ is ax + by + cz = d. Each of the equations in the system of Example 2.2.1 has a solution set which is in one-one correspondence with a particular plane in \mathbb{R}^3 . The intersection of those three planes is the single point (83/30, 11/30, 5/6).

Example 2.3.3. Solve the following system of linear equations if possible,

$$x - y = 1$$

$$3x - 3y = 0$$

$$2x - 2y = -3$$

Gaussian elimination on the augmented coefficient matrix reveals (see Example 2.2.3 for details of the Gaussian elimination)

$$rref \begin{bmatrix} 1 & -1 & | & 1 \\ 3 & -3 & | & 0 \\ 2 & -2 & | & -3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

which shows the system has no solutions. The given equations are inconsistent.

Remark 2.3.4.

The geometric interpretation of the last example is also interesting. The equation of a line in the xy-plane is is ax + by = c, hence the solution set of a particular equation corresponds to a line. To have a solution to all three equations at once that would mean that there is an intersection point which lies on all three lines. In the preceding example there is no such point. **Example 2.3.5.** Solve the following system of linear equations if possible,

$$x - y + z = 0$$

$$3x - 3y = 0$$

$$2x - 2y - 3z = 0$$

Gaussian elimination on the augmented coefficient matrix reveals (see Example 2.2.10 for details of the Gaussian elimination)

$$rref \begin{bmatrix} 1 & -1 & 1 & | & 0 \\ 3 & -3 & 0 & | & 0 \\ 2 & -2 & -3 & | & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \implies \begin{bmatrix} x - y = 0 \\ z = 0 \end{bmatrix}$$

The row of zeros indicates that we will not find a unique solution. We have a choice to make, either x or y can be stated as a function of the other. Typically in linear algebra we will solve for the variables that correspond to the pivot columns in terms of the non-pivot column variables. In this problem the pivot columns are the first column which corresponds to the variable x and the third column which corresponds the variable z. The variables x, z are called **basic variables** while y is called a **free** variable. The solution set is $\{(y, y, 0) \mid y \in \mathbb{R}\}$; in other words, x = y, y = y and z = 0 for all $y \in \mathbb{R}$.

You might object to the last example. You might ask why is y the free variable and not x. This is roughly equivalent to asking the question why is y the dependent variable and x the independent variable in the usual calculus. However, the roles are reversed. In the preceding example the variable x depends on y. Physically there may be a reason to distinguish the roles of one variable over another. There may be a clear cause-effect relationship which the mathematics fails to capture. For example, velocity of a ball in flight depends on time, but does time depend on the ball's velocty ? I'm guessing no. So time would seem to play the role of independent variable. However, when we write equations such as $v = v_o - gt$ we can just as well write $t = \frac{v-v_o}{-g}$; the algebra alone does not reveal which variable should be taken as "independent". Hence, a choice must be made. In the case of infinitely many solutions, we customarily **choose** the pivot variables as the "dependent" or "basic" variables and the non-pivot variables as the "free" variables. Sometimes the word *parameter* is used instead of variable, it is synonomous.

Example 2.3.6. Solve the following (silly) system of linear equations if possible,

$$x = 0$$

$$0x + 0y + 0z = 0$$

$$3x = 0$$

Gaussian elimination on the augmented coefficient matrix reveals (see Example 2.2.10 for details of the Gaussian elimination)

we find the solution set is $\lfloor \{(0, y, z) \mid y, z \in \mathbb{R}\} \rfloor$. No restriction is placed on the free variables y and z.

Example 2.3.7. Solve the following system of linear equations if possible,

$$x_1 + x_2 + x_3 + x_4 = 1$$

$$x_1 - x_2 + x_3 = 1$$

$$-x_1 + x_3 + x_4 = 1$$

Gaussian elimination on the augmented coefficient matrix reveals (see Example 2.2.4 for details of the Gaussian elimination)

$$rref\left[\begin{array}{rrr}1 & 1 & 1 & 1 & 1\\1 & -1 & 1 & 0 & 1\\-1 & 0 & 1 & 1 & 1\end{array}\right] = \left[\begin{array}{rrr}1 & 0 & 0 & 0 & 0\\0 & 1 & 0 & 1/2 & 0\\0 & 0 & 1 & 3/4 & 1\end{array}\right]$$

We find solutions of the form $x_1 = 0$, $x_2 = -x_4/2$, $x_3 = 1 - 3x_4/4$ where $x_4 \in \mathbb{R}$ is free. The solution set is a subset of \mathbb{R}^4 , namely $[(0, -2s, 1 - 3s, 4s) | s \in \mathbb{R}]]$ (I used $s = 4x_4$ to get rid of the annoying fractions).

Remark 2.3.8.

The geometric interpretation of the last example is difficult to visualize. Equations of the form $a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = b$ represent volumes in \mathbb{R}^4 , they're called hyperplanes. The solution is parametrized by a single free variable, this means it is a line. We deduce that the three hyperplanes corresponding to the given system intersect along a line. Geometrically solving two equations and two unknowns isn't too hard with some graph paper and a little patience you can find the solution from the intersection of the two lines. When we have more equations and unknowns the geometric solutions are harder to grasp. Analytic geometry plays a secondary role in this course so if you have not had calculus III then don't worry too much. I should tell you what you need to know in these notes.

Example 2.3.9. Solve the following system of linear equations if possible,

$$x_1 + x_4 = 0$$

$$2x_1 + 2x_2 + x_5 = 0$$

$$4x_1 + 4x_2 + 4x_3 = 1$$

Gaussian elimination on the augmented coefficient matrix reveals (see Example 2.2.5 for details of the Gaussian elimination)

$$rref \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 & 1 & 0 \\ 4 & 4 & 4 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & -1/2 & 1/4 \end{bmatrix}$$

Consequently, x_4, x_5 are free and solutions are of the form

$$x_1 = -x_4 x_2 = x_4 - \frac{1}{2}x_5 x_3 = \frac{1}{4} + \frac{1}{2}x_5$$

for all $x_4, x_5 \in \mathbb{R}$.

Example 2.3.10. Solve the following system of linear equations if possible,

```
x_1 + x_3 = 0

2x_2 = 0

3x_3 = 1

3x_1 + 2x_2 = 0
```

Gaussian elimination on the augmented coefficient matrix reveals (see Example 2.2.8 for details of the Gaussian elimination)

$$rref\left[\begin{array}{rrr}1 & 0 & 1 & 0\\0 & 2 & 0 & 0\\0 & 0 & 3 & 1\\3 & 2 & 0 & 0\end{array}\right] = \left[\begin{array}{rrr}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & 1\end{array}\right]$$

Therefore, there are no solutions.

Example 2.3.11. Solve the following system of linear equations if possible,

$$x_1 + x_3 = 0$$

$$2x_2 = 0$$

$$3x_3 + x_4 = 0$$

$$3x_1 + 2x_2 = 0$$

Gaussian elimination on the augmented coefficient matrix reveals (see Example 2.2.10 for details of the Gaussian elimination)

$$rref \begin{bmatrix} 1 & 0 & 1 & 0 & | & 0 \\ 0 & 2 & 0 & 0 & | & 0 \\ 0 & 0 & 3 & 1 & | & 0 \\ 3 & 2 & 0 & 0 & | & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & 0 & | & 0 \\ 0 & 0 & 1 & 0 & | & 0 \end{bmatrix}$$

Therefore, the unique solution is $x_1 = x_2 = x_3 = x_4 = 0$. The solution set here is rather small, it's $\{(0,0,0,0)\}$.

2.4 applications to curve fitting and circuits

We explore a few fun simple examples in this section. I don't intend for you to master the outs and in's of circuit analysis, those examples are for site-seeing purposes.².

Example 2.4.1. Find a polynomial P(x) whose graph y = P(x) fits through the points (0, -2.7), (2, -4.5) and (1, 0.97). We expect a quadratic polynomial will do nicely here: let A, B, C be the coefficients so $P(x) = Ax^2 + Bx + C$. Plug in the data,

P(0) = C = -2.7		A	B	C	
P(0) = C = -2.1	ς	0	0	1	-2.7
P(2) = 4A + 2D + C = -4.3 P(1) = 4A + D + C = 0.07	\Rightarrow	4	2	1	-4.5
P(1) = A + B + C = 0.97		1	1	1	0.97

I put in the A, B, C labels just to emphasize the form of the augmented matrix. We can then perform Gaussian elimination on the matrix (I omit the details) to solve the system,

rref	$\begin{bmatrix} 0 \\ 4 \\ 1 \end{bmatrix}$	${0 \\ 2 \\ 1}$	1 1 1	-2.7 -4.5 0.97	=	$\begin{bmatrix} 1\\0\\0 \end{bmatrix}$	$\begin{array}{c} 0 \\ 1 \\ 0 \end{array}$	$egin{array}{c} 0 \\ 0 \\ 1 \end{array}$	$\left \begin{array}{c}-4.52\\8.14\\-2.7\end{array}\right]$	\Rightarrow	A = -4.52 $B = 8.14$ $C = -2.7$
The requested poly	Inom	nial	is	P(x) =	= -4.	52x	$^{2} +$	8.1	4x - 2.7		

Example 2.4.2. Find which cubic polynomial Q(x) have a graph y = Q(x) which fits through the points (0, -2.7), (2, -4.5) and (1, 0.97). Let A, B, C, D be the coefficients of $Q(x) = Ax^3 + Bx^2 + Cx + D$. Plug in the data,

O(0)	-D - 27	A	B	C	D]
Q(0)	-D2.1	0	0	0	1	-2.7
Q(2)	$= 6A + 4D + 2C + D = -4.3 \qquad \Rightarrow \qquad \qquad$	8	4	2	1	-4.5
Q(1)	= A + B + C + D = 0.97	1	1	1	1	0.97

I put in the A, B, C, D labels just to emphasize the form of the augmented matrix. We can then perform Gaussian elimination on the matrix (I omit the details) to solve the system,

 $rref \begin{bmatrix} 0 & 0 & 0 & 1 & | & -2.7 \\ 8 & 4 & 2 & 1 & | & -4.5 \\ 1 & 1 & 1 & 1 & | & 0.97 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -0.5 & 0 & | & -4.07 \\ 0 & 1 & 1.5 & 0 & | & 7.69 \\ 0 & 0 & 0 & 1 & | & -2.7 \end{bmatrix} \Rightarrow \begin{array}{c} A = -4.07 + 0.5C \\ B = 7.69 - 1.5C \\ C = C \\ D = -2.7 \end{array}$

It turns out there is a whole family of cubic polynomials which will do nicely. For each $C \in \mathbb{R}$ the polynomial is $Q_C(x) = (c - 4.07)x^3 + (7.69 - 1.5C)x^2 + Cx - 2.7$ fits the given points. We asked a question and found that it had infinitely many answers. Notice the choice C = 4.07 gets us back to the last example, in that case $Q_C(x)$ is not really a cubic polynomial.

Example 2.4.3. Consider the following traffic-flow pattern. The diagram indicates the flow of cars between the intersections A, B, C, D. Our goal is to analyze the flow and determine the missing pieces of the puzzle, what are the flow-rates x_1, x_2, x_3 . We assume all the given numbers are cars per hour, but we omit the units to reduce clutter in the equations.

 $^{^{2}}$...well, modulo that homework I asked you to do, but it's not that hard, even a Sagittarian could do it.



We model this by one simple principle: conservation of vehicles

 $A: \quad x_1 - x_2 - 400 = 0$ $B: \quad -x_1 + 600 - 100 + x_3 = 0$ $C: \quad -300 + 100 + 100 + x_2 = 0$ $D: \quad -100 + 100 + x_3 = 0$

This gives us the augmented-coefficient matrix and Gaussian elimination that follows (we have to rearrange the equations to put the constants on the right and the variables on the left before we translate to matrix form)

$$rref \begin{bmatrix} 1 & -1 & 0 & | & 400 \\ -1 & 0 & 1 & | & -500 \\ 0 & 1 & 0 & | & 100 \\ 0 & 0 & 1 & | & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & | & 500 \\ 0 & 1 & 0 & | & 100 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

From this we conclude, $x_3 = 0, x_2 = 100, x_1 = 500$. By the way, this sort of system is called **overdetermined** because we have more equations than unknowns. If such a system is consistent they're often easy to solve. In truth, the rref business is completely uncessary here. I'm just trying to illustrate what can happen.

Example 2.4.4. (taken from Lay's homework, $\S1.6\#7$) Alka Seltzer makes fizzy soothing bubbles through a chemical reaction of the following type:

$$\underbrace{NaHCO_3}_{\text{odium bicarbonate}} + \underbrace{H_3C_6H_5O_7}_{\text{citric acid}} \rightarrow \underbrace{Na_3C_6H_5O_7}_{\text{sodium citrate}} + \underbrace{H_2O+CO_2}_{\text{water and carbon dioxide}}$$

The reaction above is **unbalanced** because it lacks weights to describe the relative numbers of the various molecules involved in a particular reaction. To **balance** the equation we seek integers x_1, x_2, x_3, x_4, x_5 such that the following reaction is balanced.

$$x_1(NaHCO_3) + x_2(H_3C_6H_5O_7) \rightarrow x_3(Na_3C_6H_5O_7) + x_4(H_2O) + x_5(CO_2)$$

In a chemical reaction the atoms the enter the reaction must also leave the reaction. Atoms are neither created nor destroyed in chemical reactions³. It follows that the number of sodium(Na),

s

³chemistry is based on electronic interactions which do not possess the mechanisms needed for alchemy, transmutation is in fact accomplished in nuclear physics. Ironically, alchemy, while known, is not economical

hydrogen(H), carbon(C) and oxygen(O) atoms must be conserved in the reaction. Each element can be represented by a component in a 4-dimensional vector; (Na, H, C, O). Using this notation the equation to balance the reaction is simply:

$$x_{1}\begin{bmatrix}1\\1\\1\\3\end{bmatrix} + x_{2}\begin{bmatrix}0\\8\\6\\7\end{bmatrix} = x_{3}\begin{bmatrix}3\\5\\6\\7\end{bmatrix} + x_{4}\begin{bmatrix}0\\2\\0\\1\end{bmatrix} + x_{5}\begin{bmatrix}0\\0\\1\\2\end{bmatrix}$$

In other words, solve

$$\begin{array}{c} x_1 = 3x_3 \\ x_1 + 8x_2 = 5x_3 + 2x_4 \\ x_1 + 6x_2 = 6x_3 + x_5 \\ 3x_1 + 7x_2 = 6x_3 + x_5 \end{array} \Rightarrow \begin{bmatrix} 1 & 0 & -3 & 0 & 0 & 0 \\ 1 & 8 & -5 & -2 & 0 & 0 \\ 1 & 6 & -6 & 0 & -1 & 0 \\ 3 & 7 & -6 & 0 & -1 & 0 \end{bmatrix}$$

After a few row operations we will deduce,

$$rref \begin{bmatrix} 1 & 0 & -3 & 0 & 0 & 0 \\ 1 & 8 & -5 & -2 & 0 & 0 \\ 1 & 6 & -6 & 0 & -1 & 0 \\ 3 & 7 & -6 & 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & \frac{-1}{3} & 0 \\ 0 & 0 & 1 & 0 & \frac{-1}{3} & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{bmatrix}$$

Therefore, $x_1 = x_5, x_2 = x_5/3, x_3 = x_5/3$ and $x_4 = x_5$. Atoms are indivisible (in this context) hence we need to choose $x_5 = 3k$ for $k \in \mathbb{N}$ to assure integer solutions. The basic reaction follows from $x_5 = 3$,

$$3NaHCO_3 + H_3C_6H_5O_7 \rightarrow Na_3C_6H_5O_7 + 3H_2O + 3CO_2$$

Finding integer solutions to chemical reactions is more easily solved by the method I used as an undergraduate. You guess and check and adjust. Because the solutions are integers it's not too hard to work out. That said, if you don't want to guess then we have a method via Gaussian elimination. Chemists have more to worry about than just this algebra. If you study reactions carefully then there are a host of other considerations involving energy transfer and ultimately quantum mechanics.

Example 2.4.5. Let $R = 1\Omega$ and $V_1 = 8V$. Determine the voltage V_A and currents I_1, I_2, I_3 flowing in the circuit as pictured below:



Conservation of charge implies the sum of currents into a node must equal the sum of the currents flowing out of the node. We use Ohm's Law V = IR to set-up the currents, here V should be the

voltage dropped across the resistor R.

Substitute the first three equations into the fourth to obtain

$$\frac{V_A}{R} = \frac{2V_1 - V_A}{4R} + \frac{V_1 - V_A}{4R}$$

Multiply by 4R and we find

$$4V_A = 2V_1 - V_A + V_1 - V_A \Rightarrow 6V_A = 3V_1 \Rightarrow V_A = V_1/2 = 4V.$$

Substituting back into the Ohm's Law equations we determine $I_1 = \frac{16V - 4V}{4\Omega} = 3A$, $I_2 = \frac{4V}{1\Omega} = 4A$ and $I_3 = \frac{8V - 4V}{4\Omega} = 1A$. This obvious checks with $I_2 = I_1 + I_3$. In practice it's not always best to use the full-power of the rref.

2.5 conclusions

We concluded the last section with a rather believable (but tedious to prove) Theorem. We do the same here,

Theorem 2.5.1.

Given a system of m linear equations and n unknowns the solution set falls into one of the following cases:

- 1. the solution set is empty.
- 2. the solution set has only one element.
- 3. the solution set is infinite.

Proof: Consider the augmented coefficient matrix $[A|b] \in \mathbb{R}^{m \times (n+1)}$ for the system (Theorem 2.2.11 assures us it exists and is unique). Calculate rref[A|b]. If rref[A|b] contains a row of zeros with a 1 in the last column then the system is inconsistent and we find no solutions thus the solution set is empty.

Suppose rref[A|b] does not contain a row of zeros with a 1 in the far right position. Then there are less than n + 1 pivot columns. Suppose there are n pivot columns, let c_i for i = 1, 2, ..., m be the entries in the rightmost column. We find $x_1 = c_1, x_2 = c_2, ..., x_n = c_m$. Consequently the solution set is $\{(c_1, c_2, ..., c_m)\}$.

If rref[A|b] has k < n pivot columns then there are (n + 1 - k)-non-pivot positions. Since the last column corresponds to b it follows there are (n - k) free variables. But, k < n implies 0 < n - k hence there is at least one free variable. Therefore there are infinitely many solutions. \Box

2.5. CONCLUSIONS

Theorem 2.5.2.

Suppose that $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{m \times p}$ then the first *n* columns of rref[A] and rref[A|B] are identical.

Proof: The forward pass of the elimination proceeds from the leftmost-column to the rightmostcolumn. The matrices A and [A|B] have the same *n*-leftmost columns thus the *n*-leftmost columns are identical after the forward pass is complete. The backwards pass acts on column at a time just clearing out above the pivots. Since the ref(A) and ref[A|B] have identical *n*-leftmost columns the backwards pass modifies those columns in the same way. Thus the *n*-leftmost columns of Aand [A|B] will be identical. \Box

The theorem below is continued as we work through this course. Eventually, it has about a dozen seemingly disconnected parts.

Theorem 2.5.3.

Given *n*-linear equations in *n*-unknowns Ax = b, a unique solution x exists iff rref[A|b] = [I|x]. Moreover, if $rref[A] \neq I$ then there is no unique solution to the system of equations.

Proof: If a unique solution $x_1 = c_1, x_2 = c_2, \ldots, x_n = c_n$ exists for a given system of equations Ax = b then we know

$$A_{i1}c_1 + A_{i2}c_2 + \dots + A_{in}c_n = b_i$$

for each i = 1, 2, ..., n and this is the only ordered set of constants which provides such a solution. Suppose that $rref[A|b] \neq [I|c]$. If rref[A|b] = [I|d] and $d \neq c$ then d is a new solution thus the solution is not unique, this contradicts the given assumption. Consider, on the other hand, the case rref[A|b] = [J|f] where $J \neq I$. If there is a row where f is nonzero and yet J is zero then the system is inconsistent. Otherwise, there are infinitely many solutions since J has at least one non-pivot column as $J \neq I$. Again, we find contradictions in every case except the claimed result. It follows if x = c is the unique solution then rref[A|b] = [I|c]. The converse follows essentially the same argument, if rref[A|b] = [I|c] then clearly Ax = b has solution x = c and if that solution were not unique then we be able to find a different rref for [A|b] but that contradicts the uniqueness of rref. \Box

There is much more to say about the meaning of particular patterns in the reduced row echelon form of the matrix. We will continue to mull over these matters in later portions of the course. Theorem 2.5.1 provides us the big picture. It is remarkable that two equations and two unknowns already revealed these patterns.

Incidentally, you might notice that the Gauss-Jordan algorithm did not assume all the structure of the real numbers. For example, we never needed to use the ordering relations < or >. All we needed was addition, subtraction and the ability to multiply by the inverse of a nonzero number. Any field of numbers will likewise work. Theorems 2.5.1 and 2.2.11 also hold for matrices of rational (\mathbb{Q}) or complex (\mathbb{C}) numbers. We will encounter problems which require calculation in \mathbb{C} . If you are interested in encryption then calculations over a finite field \mathbb{Z}_p are necessary. In contrast, Gausssian elimination does not work for matrices of integers since we do not have fractions to work with in that context. For a much deeper look at linear algebra, see the Part III of Dummit and Foote's third edition of *Algebra*. In that text, the concept of a module is detailed and an analog for Gaussian elimination is given where the field is replaced with a ring (good examples of rings are \mathbb{Z}_n , \mathbb{Z} or the set of \mathbb{R} -valued functions on some space. Every good math major should leave their undergraduate with a command of basic ring theory.

Chapter 3

algebra of matrices

I decided to devote a chapter of these notes to matrices. Our goal here is to appreciate the richness of the matrix construction. These arrays of numbers were at first merely a book-keeping device to manage solutions of many equations and many unknowns. Matrix multiplication, probably first discovered as it relates to solutions by substitution, now is used without reference to any system of equations. Such is the life of matrices, these were born from equations, but now they are often used a langauge of their own. Matrix notation allows us to group many equations into a single elegant matrix equation. The algebraic identities which matrices can encode are boundless. We can use matrices to construct \mathbb{C}, \mathbb{Z}_n and a host of things I ought not name here. For example, any finite dimensional Lie algebra can be realized as a commutator algebra on a set of matrices of sufficiently large size (Ado's Theorem). Later in this course, we'll see how matrices are intimately connected with linear transformations, a single matrix captures the essence of the action of a linear transformation on all of space. My point is just this, matrices are interesting on their own. They're much more than a box of numbers.

3.1 addition and multiplication by scalars

Definition 3.1.1.

Let $A, B \in \mathbb{R}^{m \times n}$ then $A + B \in \mathbb{R}^{m \times n}$ is defined by $(A + B)_{ij} = A_{ij} + B_{ij}$ for all $1 \le i \le m$, $1 \le j \le n$. If two matrices A, B are not of the same size then there sum is not defined.

Example 3.1.2. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$. We calculate

$$A + B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}.$$

Definition 3.1.3.

Let $A, B \in \mathbb{R}^{m \times n}, c \in \mathbb{R}$ then $cA \in \mathbb{R}^{m \times n}$ is defined by $(cA)_{ij} = cA_{ij}$ for all $1 \le i \le m$, $1 \le j \le n$. We call the process of multiplying A by a number c **multiplication by a scalar**. We define $A - B \in \mathbb{R}^{m \times n}$ by A - B = A + (-1)B which is equivalent to $(A - B)_{ij} = A_{ij} - B_{ij}$ for all i, j.

Example 3.1.4. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$. We calculate

$$A - B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} -4 & -4 \\ -4 & -4 \end{bmatrix}.$$

Now multiply A by the scalar 5,

$$5A = 5\begin{bmatrix} 1 & 2\\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 10\\ 15 & 20 \end{bmatrix}$$

Example 3.1.5. Let $A, B \in \mathbb{R}^{m \times n}$ be defined by $A_{ij} = 3i + 5j$ and $B_{ij} = i^2$ for all i, j. Then we can calculate $(A + B)_{ij} = 3i + 5j + i^2$ for all i, j.

Definition 3.1.6.

The **zero matrix** in $\mathbb{R}^{m \times n}$ is denoted 0 and defined by $0_{ij} = 0$ for all i, j. The additive inverse of $A \in \mathbb{R}^{m \times n}$ is the matrix -A such that A + (-A) = 0. The components of the additive inverse matrix are given by $(-A)_{ij} = -A_{ij}$ for all i, j.

The zero matrix joins a long list of other objects which are all denoted by 0. Usually the meaning of 0 is clear from the context, the size of the zero matrix is chosen as to be consistent with the equation in which it is found.

Example 3.1.7. Solve the following matrix equation,

$$0 = \begin{bmatrix} x & y \\ z & w \end{bmatrix} + \begin{bmatrix} -1 & -2 \\ -3 & -4 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} x-1 & y-2 \\ z-3 & w-4 \end{bmatrix}$$

The definition of matrix equality means this single matrix equation reduces to 4 scalar equations: 0 = x - 1, 0 = y - 2, 0 = z - 3, 0 = w - 4. The solution is x = 1, y = 2, z = 3, w = 4.

Theorem 3.1.8.

If $A \in \mathbb{R}^{m \times n}$ then 1. $0 \cdot A = 0$, (where 0 on the L.H.S. is the number zero) 2. 0A = 0, 3. A + 0 = 0 + A = A.

Proof: I'll prove (2.). Let $A \in \mathbb{R}^{m \times n}$ and consider

$$(0A)_{ij} = \sum_{k=1}^{m} 0_{ik} A_{kj} = \sum_{k=1}^{m} 0A_{kj} = \sum_{k=1}^{m} 0 = 0$$

for all i, j. Thus 0A = 0. I leave the other parts to the reader, the proofs are similar. \Box

Matrix addition and scalar multiplication is very natural in general. Let us collect the important facts for future reference.

Theorem 3.1.9.

If $A, B, C \in \mathbb{R}^{m \times n}$ and $c_1, c_2 \in \mathbb{R}$ then 1. (A + B) + C = A + (B + C),2. A + B = B + A,3. $c_1(A + B) = c_1A + c_2B,$ 4. $(c_1 + c_2)A = c_1A + c_2A,$ 5. $(c_1c_2)A = c_1(c_2A),$ 6. 1A = A,

Proof: Nearly all of these properties are proved by breaking the statement down to components then appealing to a property of real numbers. I supply proofs of (1.) and (5.) and leave (2.),(3.),(4.) and (6.) to the reader.

Proof of (1.): assume A, B, C are given as in the statement of the Theorem. Observe that

$((A+B)+C)_{ij}$	$= (A+B)_{ij} + C_{ij}$	defn. of matrix add.
	$= (A_{ij} + B_{ij}) + C_{ij}$	defn. of matrix add.
	$= A_{ij} + (B_{ij} + C_{ij})$	assoc. of real numbers
	$= A_{ij} + (B+C)_{ij}$	defn. of matrix add.
	$= (A + (B + C))_{ij}$	defn. of matrix add.

for all i, j. Therefore (A + B) + C = A + (B + C). \Box

Proof of (5.): assume c_1, c_2, A are given as in the statement of the Theorem. Observe that

 $\begin{array}{ll} ((c_1c_2)A)_{ij} &= (c_1c_2)A_{ij} & \text{defn. scalar multiplication.} \\ &= c_1(c_2A_{ij}) & \text{assoc. of real numbers} \\ &= (c_1(c_2A))_{ij} & \text{defn. scalar multiplication.} \end{array}$

for all i, j. Therefore $(c_1c_2)A = c_1(c_2A)$. \Box

The proofs of the other items are similar, we consider the i, j-th component of the identity and then apply the definition of the appropriate matrix operation's definition. This reduces the problem to a statement about real numbers so we can use the properties of real numbers at the level of components. After applying the crucial fact about real numbers, we then reverse the steps. Since the calculation works for arbitrary i, j it follows the the matrix equation holds true. This Theorem provides a foundation for later work where we may find it convenient to prove a statement without resorting to a proof by components. Which method of proof is best depends on the question. However, I can't see another way of proving most of 3.1.9.

3.2 matrix algebra

This may be the most important section in this chapter. Here we learn how to multiply matrices, what their basic algebraic properties are and we begin study of matrix inversion.

Definition 3.2.1.

Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$ then the product of A and B is denoted by juxtaposition AB and $AB \in \mathbb{R}^{m \times p}$ is defined by:

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$$

for each $1 \le i \le m$ and $1 \le j \le p$. In the case m = p = 1 the indices i, j are omitted in the equation since the matrix product is simply a number which needs no index.

This definition is very nice for general proofs and we will need to know it for proofs. However, for explicit numerical examples, I usually think of matrix multiplication in terms of *dot-products*.

Definition 3.2.2.

Let $v = \langle v_1, \ldots, v_n \rangle$ and $w = \langle w_1, \ldots, w_n \rangle$ be *n*-vectors then the **dot-product** of v and w is the number defined below:

$$v \cdot w = v_1 w_1 + v_2 w_2 + \dots + v_n w_n = \sum_{k=1}^n v_k w_k$$

There are many things to say about dot-products. The geometric content of this formula is hard to overstate. We should return to that task in the third part of this course.

Proposition 3.2.3.

Let $v, w \in \mathbb{R}^n$ then $v \cdot w = v^T w$.

Proof: Since v^T is an $1 \times n$ matrix and w is an $n \times 1$ matrix the definition of matrix multiplication indicates $v^T w$ should be a 1×1 matrix which is a number. Note in this case the outside indices ij are absent in the boxed equation so the equation reduces to

$$v^T w = v^T_1 w_1 + v^T_2 w_2 + \dots + v^T_n w_n = v_1 w_1 + v_2 w_2 + \dots + v_n w_n = v \cdot w.\Box$$

Proposition 3.2.4.

The formula given below is equivalent to the Definition 3.2.1. Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$ then

$$AB = \begin{bmatrix} row_1(A) \cdot col_1(B) & row_1(A) \cdot col_2(B) & \cdots & row_1(A) \cdot col_p(B) \\ row_2(A) \cdot col_1(B) & row_2(A) \cdot col_2(B) & \cdots & row_2(A) \cdot col_p(B) \\ \vdots & \vdots & \ddots & \vdots \\ row_m(A) \cdot col_1(B) & row_m(A) \cdot col_2(B) & \cdots & row_m(A) \cdot col_p(B) \end{bmatrix}$$

Proof: The formula above claims $(AB)_{ij} = row_i(A) \cdot col_j(B)$ for all i, j. Recall that $(row_i(A))_k = A_{ik}$ and $(col_j(B))_k = B_{kj}$ thus

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} = \sum_{k=1}^{n} (row_i(A))_k (col_j(B))_k$$

Hence, using definition of the dot-product, $(AB)_{ij} = row_i(A) \cdot col_j(B)$. This argument holds for all i, j therefore the Proposition is true. \Box

Example 3.2.5. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $v = \begin{bmatrix} x \\ y \end{bmatrix}$ then we may calculate the product Av as follows: $Av = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x \begin{bmatrix} 1 \\ 3 \end{bmatrix} + y \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} x + 2y \\ 3x + 4y \end{bmatrix}.$

Notice that the product of an $n \times k$ matrix with a $k \times 1$ vector yields another vector of size $k \times 1$. In the example above we observed the pattern $(2 \times 2)(2 \times 1) \rightarrow (2 \times 1)$.

Example 3.2.6. The product of a 3×2 and 2×3 is a 3×3

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} [1,0][4,7]^T & [1,0][5,8]^T & [1,0][6,9]^T \\ [0,1][4,7]^T & [0,1][5,8]^T & [0,1][6,9]^T \\ [0,0][4,7]^T & [0,0][5,8]^T & [0,0][6,9]^T \end{bmatrix} = \begin{bmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \\ 0 & 0 & 0 \end{bmatrix}$$

Example 3.2.7. The product of a 3×1 and 1×3 is a 3×3

$$\begin{bmatrix} 1\\2\\3 \end{bmatrix} \begin{bmatrix} 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 4 \cdot 1 & 5 \cdot 1 & 6 \cdot 1\\4 \cdot 2 & 5 \cdot 2 & 6 \cdot 2\\4 \cdot 3 & 5 \cdot 3 & 6 \cdot 3 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 6\\8 & 10 & 12\\12 & 15 & 18 \end{bmatrix}$$

Example 3.2.8. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ and $v = \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}$ calculate Av.

$$Av = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix} = \begin{bmatrix} (1,2,3) \cdot (1,0,-3) \\ (4,5,6) \cdot (1,0,-3) \\ (7,8,9) \cdot (1,0,-3) \end{bmatrix} = \begin{bmatrix} -2 \\ -14 \\ -20 \end{bmatrix}.$$

Example 3.2.9. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$. We calculate

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$
$$= \begin{bmatrix} [1,2][5,7]^T & [1,2][6,8]^T \\ [3,4][5,7]^T & [3,4][6,8]^T \end{bmatrix}$$
$$= \begin{bmatrix} 5+14 & 6+16 \\ 15+28 & 18+32 \end{bmatrix}$$
$$= \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

Notice the product of square matrices is square. For numbers $a, b \in \mathbb{R}$ it we know the product of a

and b is commutative (ab = ba). Let's calculate the product of A and B in the opposite order,

$$BA = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$
$$= \begin{bmatrix} [5,6][1,3]^T & [5,6][2,4]^T \\ [7,8][1,3]^T & [7,8][2,4]^T \end{bmatrix}$$
$$= \begin{bmatrix} 5+18 & 10+24 \\ 7+24 & 14+32 \end{bmatrix}$$
$$= \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix}$$

Clearly $AB \neq BA$ thus matrix multiplication is noncommutative or nonabelian.

If the **commutator** of two square matrices A, B is given by [A, B] = AB - BA. If $[A, B] \neq 0$ then clearly $AB \neq BA$. There are many interesting properties of the commutator. It has deep physical significance in quantum mechanics. It is also the quintessential example of a **Lie Bracket**.

When we say that matrix multiplication is noncommutative that indicates that the product of two matrices does not *generally* commute. However, there are special matrices which commute with other matrices.

Example 3.2.10. Let $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. We calculate

Likewise calculate,

$$AI = \left[\begin{array}{cc} a & b \\ c & d \end{array} \right] \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right] = \left[\begin{array}{cc} a & b \\ c & d \end{array} \right]$$

Since the matrix A was arbitrary we conclude that IA = AI for all $A \in \mathbb{R}^{2 \times 2}$.

Definition 3.2.11.

The identity matrix in $\mathbb{R}^{n \times n}$ is the $n \times n$ square matrix I which has components $I_{ij} = \delta_{ij}$. The notation I_n is sometimes used if the size of the identity matrix needs emphasis, otherwise the size of the matrix I is to be understood from the context.

	г	1	0	0 7		1	0	0	0	
$I = \begin{bmatrix} 1 & 0 \end{bmatrix}$	τ_	1	1		τ_	0	1	0	0	
$I_2 = \begin{bmatrix} 0 & 1 \end{bmatrix}$	$I_3 =$	0	1		$I_4 =$	0	0	1	0	
	L	0	0	ŢŢ		0	0	0	1	

Proposition 3.2.12.

If $X \in \mathbb{R}^{n \times p}$ then $XI_p = X$ and $I_n X = X$.

Proof: I omit the p in I_p to reduce clutter below. Consider the i, j component of XI,

$$(XI)_{ij} = \sum_{k=1}^{p} X_{ik} I_{kj} \qquad \text{defn. matrix multiplication}$$
$$= \sum_{k=1}^{p} X_{ik} \delta_{kj} \qquad \text{defn. of } I$$
$$= X_{ij}$$

The last step follows from the fact that all other terms in the sum are made zero by the Kronecker delta. Finally, observe the calculation above holds for all i, j hence XI = X. The proof of IX = X is left to the reader. \Box

Definition 3.2.13.

Let $A \in \mathbb{R}^{n \times n}$. If there exists $B \in \mathbb{R}^{n \times n}$ such that AB = I and BA = I then we say that A is **invertible** and $A^{-1} = B$. Invertible matrices are also called **nonsingular**. If a matrix has no inverse then it is called a **noninvertible** or **singular** matrix.

The power of a matrix is defined in the natural way. Notice we need for A to be square in order for the product AA to be defined.

Definition 3.2.14.

Let $A \in \mathbb{R}^{n \times n}$. We define $A^0 = I$, $A^1 = A$ and $A^m = AA^{m-1}$ for all $m \ge 1$. If A is invertible then $A^{-p} = (A^{-1})^p$.

As you would expect, $A^3 = AA^2 = AAA$.

Proposition 3.2.15.

Let $A, B \in \mathbb{R}^{n \times n}$ and $p, q \in \mathbb{N} \cup \{0\}$ 1. $(A^p)^q = A^{pq}$.

- 2. $A^p A^q = A^{p+q}$.
- 3. If *A* is invertible, $(A^{-1})^{-1} = A$.

Proof: left to reader. \Box

You should notice that $(AB)^p \neq A^p B^p$ for matrices. Instead,

$$(AB)^2 = ABAB, \qquad (AB)^3 = ABABAB, etc...$$

This means the binomial theorem will not hold for matrices. For example,

$$(A+B)^{2} = (A+B)(A+B) = A(A+B) + B(A+B) = AA + AB + BA + BB$$

hence $(A+B)^2 \neq A^2 + 2AB + B^2$ as the matrix product is not generally commutative. If we have A and B commute then AB = BA and we can prove that $(AB)^p = A^p B^p$ and the binomial theorem holds true.

Example 3.2.16. Consider A, v, w from Example 3.6.1.

$$v + w = \begin{bmatrix} 5\\7 \end{bmatrix} + \begin{bmatrix} 6\\8 \end{bmatrix} = \begin{bmatrix} 11\\15 \end{bmatrix}$$

Using the above we calculate,

$$A(v+w) = \begin{bmatrix} 1 & 2\\ 3 & 4 \end{bmatrix} \begin{bmatrix} 11\\ 15 \end{bmatrix} = \begin{bmatrix} 11+30\\ 33+60 \end{bmatrix} = \begin{bmatrix} 41\\ 93 \end{bmatrix}.$$

In constrast, we can add Av and Aw,

$$Av + Aw = \begin{bmatrix} 19\\43 \end{bmatrix} + \begin{bmatrix} 22\\50 \end{bmatrix} = \begin{bmatrix} 41\\93 \end{bmatrix}.$$

Behold, A(v+w) = Av + Aw for this example. It turns out this is true in general.

Properties of matrix multiplication are given in the theorem below. To summarize, matrix math works as you would expect with the exception that matrix multiplication is not commutative. We must be careful about the order of letters in matrix expressions.

Theorem 3.2.17.

If
$$A, B, C \in \mathbb{R}^{m \times n}$$
, $X, Y \in \mathbb{R}^{n \times p}$, $Z \in \mathbb{R}^{p \times q}$ and $c_1, c_2 \in \mathbb{R}$ then
1. $(AX)Z = A(XZ)$,
2. $(c_1A)X = c_1(AX) = A(c_1X) = (AX)c_1$,
3. $A(X + Y) = AX + AY$,
4. $A(c_1X + c_2Y) = c_1AX + c_2AY$,
5. $(A + B)X = AX + BX$,

Proof: I leave the proofs of (1.), (2.), (4.) and (5.) to the reader. Proof of (3.): assume A, X, Y are given as in the statement of the Theorem. Observe that

$$\begin{array}{ll} ((A(X+Y))_{ij} &= \sum_k A_{ik}(X+Y)_{kj} & \text{defn. matrix multiplication,} \\ &= \sum_k A_{ik}(X_{kj}+Y_{kj}) & \text{defn. matrix addition,} \\ &= \sum_k (A_{ik}X_{kj}+A_{ik}Y_{kj}) & \text{dist. of real numbers,} \\ &= \sum_k A_{ik}X_{kj} + \sum_k A_{ik}Y_{kj}) & \text{prop. of finite sum,} \\ &= (AX)_{ij} + (AY)_{ij} & \text{defn. matrix multiplication}(\times 2), \\ &= (AX + AY)_{ij} & \text{defn. matrix addition,} \end{array}$$

for all i, j. Therefore A(X + Y) = AX + AY. \Box

The proofs of the other items are similar, I invite the reader to try to prove them in a style much like the proof I offer above.

3.3 all your base are belong to us (e_i and E_{ij} that is)

Recall that we defined $e_i \in \mathbb{R}^n$ by $(e_i)_j = \delta_{ij}$. We call e_i the *i*-th standard basis vector. We proved in Proposition 1.5.8 that every vector in \mathbb{R}^n is a linear combination of e_1, e_2, \ldots, e_n . We can define a standard basis for matrices of arbitrary size in much the same manner.

Definition 3.3.1.

The *ij*-th standard basis matrix for $\mathbb{R}^{m \times n}$ is denoted E_{ij} for $1 \le i \le m$ and $1 \le j \le n$. The matrix E_{ij} is zero in all entries except for the (i, j)-th slot where it has a 1. In other words, we define $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$.

Proposition 3.3.2.

Every matrix in $\mathbb{R}^{m \times n}$ is a linear combination of the E_{ij} where $1 \leq i \leq m$ and $1 \leq j \leq n$.

Proof: Let $A \in \mathbb{R}^{m \times n}$ then

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}$$
$$= A_{11} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix} + A_{12} \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix} + \cdots + A_{mn} \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$= A_{11}E_{11} + A_{12}E_{12} + \dots + A_{mn}E_{mn}$$

The calculation above follows from repeated mn-applications of the definition of matrix addition and another mn-applications of the definition of scalar multiplication of a matrix. We can restate the final result in a more precise language,

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} E_{ij}$$

As we claimed, any matrix can be written as a linear combination of the E_{ij} . \Box

The term "basis" has a technical meaning which we will discuss at length in due time. For now, just think of it as part of the names of e_i and E_{ij} . These are the basic building blocks for matrix theory.

Example 3.3.3. Suppose $A \in \mathbb{R}^{m \times n}$ and $e_i \in \mathbb{R}^n$ is a standard basis vector,

$$(Ae_i)_j = \sum_{k=1}^n A_{jk}(e_i)k = \sum_{k=1}^n A_{jk}\delta_{ik} = A_{ji}$$

Thus, $|[Ae_i] = col_i(A)|$. We find that multiplication of a matrix A by the standard basis e_i yields the i - th column of A.

Example 3.3.4. Suppose $A \in \mathbb{R}^{m \times n}$ and $e_i \in \mathbb{R}^{m \times 1}$ is a standard basis vector,

$$(e_i^T A)_j = \sum_{k=1}^n (e_i)_k A_{kj} = \sum_{k=1}^n \delta_{ik} A_{kj} = A_{ij}$$

Thus, $[e_i^T A] = row_i(A)$. We find multiplication of a matrix A by the transpose of standard basis e_i yields the i - th row of A.

Example 3.3.5. Again, suppose $e_i, e_j \in \mathbb{R}^n$ are standard basis vectors. The product $e_i^T e_j$ of the $1 \times n$ and $n \times 1$ matrices is just a 1×1 matrix which is just a number. In particular consider,

$$e_i^T e_j = \sum_{k=1}^n (e_i^T)_k (e_j)_k = \sum_{k=1}^n \delta_{ik} \delta_{jk} = \delta_{ij}$$

The product is zero unless the vectors are identical.

Example 3.3.6. Suppose $e_i \in \mathbb{R}^{m \times 1}$ and $e_j \in \mathbb{R}^n$. The product of the $m \times 1$ matrix e_i and the $1 \times n$ matrix e_j^T is an $m \times n$ matrix. In particular,

$$(e_i e_j^T)_{kl} = (e_i^T)_k (e_j)_k = \delta_{ik} \delta_{jk} = (E_{ij})_{kl}$$

Thus we can construct the standard basis matrices by multiplying the standard basis vectors; $E_{ij} = e_i e_j^T$.

Example 3.3.7. What about the matrix E_{ij} ? What can we say about multiplication by E_{ij} on the right of an arbitrary matrix? Let $A \in \mathbb{R}^{m \times n}$ and consider,

$$(AE_{ij})_{kl} = \sum_{p=1}^{n} A_{kp} (E_{ij})_{pl} = \sum_{p=1}^{n} A_{kp} \delta_{ip} \delta_{jl} = A_{ki} \delta_{jl}$$

Notice the matrix above has zero entries unless j = l which means that the matrix is mostly zero except for the j-th column. We can select the j-th column by multiplying the above by e_j , using Examples 3.3.5 and 3.3.3,

$$(AE_{ij}e_j)_k = (Ae_ie_j^T e_j)_k = (Ae_i\delta_{jj})_k = (Ae_i)_k = (col_i(A))_k$$

This means,

$$AE_{ij} = \begin{bmatrix} & & column \ j & & \\ 0 & 0 & \cdots & A_{1i} & \cdots & 0 \\ 0 & 0 & \cdots & A_{2i} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_{mi} & \cdots & 0 \end{bmatrix}$$

Right multiplication of matrix A by E_{ij} moves the *i*-th column of A to the *j*-th column of AE_{ij} and all other entries are zero. It turns out that left multiplication by E_{ij} moves the *j*-th row of A to the *i*-th row and sets all other entries to zero.

Example 3.3.8. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ consider multiplication by E_{12} ,

$$AE_{12} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 \mid col_1(A) \end{bmatrix}$$

Which agrees with our general abstract calculation in the previous example. Next consider,

$$E_{12}A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} row_2(A) \\ 0 \end{bmatrix}.$$

Example 3.3.9. Calculate the product of E_{ij} and E_{kl} .

$$(E_{ij}E_{kl})_{mn} = \sum_{p} (E_{ij})_{mp} (E_{kl})_{pn} = \sum_{p} \delta_{im} \delta_{jp} \delta_{kp} \delta_{ln} = \delta_{im} \delta_{jk} \delta_{ln}$$

For example,

 $(E_{12}E_{34})_{mn} = \delta_{1m}\delta_{23}\delta_{4n} = 0.$

In order for the product to be nontrivial we must have j = k,

$$(E_{12}E_{24})_{mn} = \delta_{1m}\delta_{22}\delta_{4n} = \delta_{1m}\delta_{4n} = (E_{14})_{mn}$$

We can make the same identification in the general calculation,

 $(E_{ij}E_{kl})_{mn} = \delta_{jk}(E_{il})_{mn}.$

Since the above holds for all m, n,

$$E_{ij}E_{kl} = \delta_{jk}E_{il}$$

this is at times a very nice formula to know about.

Remark 3.3.10.

You may find the general examples in this portion of the notes a bit too much to follow. If that is the case then don't despair. Focus on mastering the numerical examples to begin with then come back to this section later. These examples are actually not that hard, you just have to get used to index calculations. The proofs in these examples are much longer if written without the benefit of index notation.

Example 3.3.11. Let $A \in \mathbb{R}^{m \times n}$ and suppose $e_i \in \mathbb{R}^{m \times 1}$ and $e_j \in \mathbb{R}^n$. Consider,

$$(e_i)^T A e_j = \sum_{k=1}^m ((e_i)^T)_k (A e_j)_k = \sum_{k=1}^m \delta_{ik} (A e_j)_k = (A e_j)_i = A_{ij}$$

This is a useful observation. If we wish to select the (i, j)-entry of the matrix A then we can use the following simple formula,

$$A_{ij} = (e_i)^T A e_j$$

This is analogus to the idea of using dot-products to select particular components of vectors in analytic geometry; (reverting to calculus III notation for a moment) recall that to find v_1 of \vec{v} we learned that the dot product by $\hat{i} = <1, 0, 0>$ selects the first components $v_1 = \vec{v} \cdot \hat{i}$. The following theorem is simply a summary of our results for this section.

Theorem 3.3.12.

Assume $A \in \mathbb{R}^{m \times n}$ and $v \in \mathbb{R}^n$ and define $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$ and $(e_i)_j = \delta_{ij}$ as we previously discussed, $v = \sum_{i=1}^n v_n e_n \qquad A = \sum_{i=1}^m \sum_{j=1}^n A_{ij} E_{ij}.$ $[e_i^T A] = row_i(A) \qquad [Ae_i] = col_i(A) \qquad A_{ij} = (e_i)^T Ae_j$ $E_{ij} E_{kl} = \delta_{jk} E_{il} \qquad E_{ij} = e_i e_j^T \qquad e_i^T e_j = \delta_{ij}$

3.3.1 diagonal and triangular matrices have no chance survive

Definition 3.3.13.

Let $A \in \mathbb{R}^{m \times n}$. If $A_{ij} = 0$ for all i, j such that $i \neq j$ then A is called a **diagonal** matrix. If A has components $A_{ij} = 0$ for all i, j such that $i \leq j$ then we call A a **upper triangular** matrix. If A has components $A_{ij} = 0$ for all i, j such that $i \geq j$ then we call A a **lower triangular** matrix.

Example 3.3.14. Let me illustrate a generic example of each case for 3×3 matrices:

ſ	A_{11}	0	0 -	A_{11}	A_{12}	A_{13}	A_{11}	0	0]
	0	A_{22}	0	0	A_{22}	A_{23}	A_{21}	A_{22}	0
	0	0	A_{33}	0	0	A_{33}	A_{31}	A_{32}	A_{33}

As you can see the diagonal matrix only has nontrivial entries on the diagonal, and the names lower triangular and upper triangular are likewise natural.

If an upper triangular matrix has zeros on the diagonal then it is said to be **strictly upper triangular**. Likewise, if a lower triangular matrix has zeros on the diagonal then it is said to be **strictly lower triangular**. Obviously and matrix can be written as a sum of a diagonal and strictly upper and strictly lower matrix,

$$A = \sum_{i,j} A_{ij} E_{ij}$$
$$= \sum_{i} A_{ii} E_{ii} + \sum_{i < j} A_{ij} E_{ij} + \sum_{i > j} A_{ij} E_{ij}$$

There is an algorithm called LU-factorization which for many matrices A finds a lower triangular matrix L and an upper triangular matrix U such that A = LU. It is one of several factorization schemes which is calculationally advantageous for large systems. There are many many ways to solve a system, but some are faster methods. Algorithmics is the study of which method is optimal.

Proposition 3.3.15.

- Let $A, B \in \mathbb{R}^{n \times n}$.
 - 1. If A, B are upper diagonal then AB is diagonal.
 - 2. If A, B are upper triangular then AB is upper triangular.
 - 3. If A, B are lower triangular then AB is lower triangular.

Proof of (1.): Suppose A and B are diagonal. It follows there exist a_i, b_j such that $A = \sum_i a_i E_{ii}$ and $B = \sum_j b_j E_{jj}$. Calculate,

$$AB = \sum_{i} a_{i}E_{ii}\sum_{j} b_{j}E_{jj}$$
$$= \sum_{i}\sum_{j} a_{i}b_{j}E_{ii}E_{jj}$$
$$= \sum_{i}\sum_{j} a_{i}b_{j}\delta_{ij}E_{ij}$$
$$= \sum_{i} a_{i}b_{i}E_{ii}$$

thus the product matrix AB is also diagonal and we find that the diagonal of the product AB is just the product of the corresponding diagonals of A and B.

Proof of (2.): Suppose A and B are upper diagonal. It follows there exist A_{ij}, B_{ij} such that $A = \sum_{i \leq j} A_{ij} E_{ij}$ and $B = \sum_{k \leq l} B_{kl} E_{kl}$. Calculate,

$$AB = \sum_{i \le j} A_{ij} E_{ij} \sum_{k \le l} B_{kl} E_{kl}$$
$$= \sum_{i \le j} \sum_{k \le l} A_{ij} B_{kl} E_{ij} E_{kl}$$
$$= \sum_{i \le j} \sum_{k \le l} A_{ij} B_{kl} \delta_{jk} E_{il}$$
$$= \sum_{i \le j} \sum_{j \le l} A_{ij} B_{jl} E_{il}$$

Notice that every term in the sum above has $i \leq j$ and $j \leq l$ hence $i \leq l$. It follows the product is upper triangular since it is a sum of upper triangular matrices. The proof of (3.) is similar. \Box .

I hope you can appreciate these arguments are superior to component level calculations with explicit listing of components and \cdots . The notations e_i and E_{ij} are extremely helpful on many such questions. Furthermore, a proof captured in the notation of this section will more clearly show the root cause for the truth of the identity in question. What is easily lost in several pages of brute-force can be elegantly seen in a couple lines of carefully crafted index calculation.

3.4 elementary matrices

Gauss Jordan elimination consists of three *elementary row operations*:

(1.)
$$r_i + ar_j \rightarrow r_i$$
, (2.) $br_i \rightarrow r_i$, (3.) $r_i \leftrightarrow r_j$

Left multiplication by **elementary matrices** will accomplish the same operation on a matrix.

Definition 3.4.1.

Let $[A : r_i + ar_j \to r_i]$ denote the matrix produced by replacing row *i* of matrix *A* with $row_i(A) + arow_j(A)$. Also define $[A : cr_i \to r_i]$ and $[A : r_i \leftrightarrow r_j]$ in the same way. Let $a, b \in \mathbb{R}$ and $b \neq 0$. The following matrices are called **elementary matrices**:

$$E_{r_i + ar_j \to r_i} = [I : r_i + ar_j \to r_i]$$
$$E_{br_i \to r_i} = [I : br_i \to r_i]$$
$$E_{r_i \leftrightarrow r_j} = [I : r_i \leftrightarrow r_j]$$

Example 3.4.2. Let $A = \begin{bmatrix} a & b & c \\ 1 & 2 & 3 \\ u & m & e \end{bmatrix}$

$$E_{r_{2}+3r_{1}\rightarrow r_{2}}A = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ 1 & 2 & 3 \\ u & m & e \end{bmatrix} = \begin{bmatrix} a & b & c \\ 3a+1 & 3b+2 & 3c+3 \\ u & m & e \end{bmatrix}$$
$$E_{7r_{2}\rightarrow r_{2}}A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ 1 & 2 & 3 \\ u & m & e \end{bmatrix} = \begin{bmatrix} a & b & c \\ 7 & 14 & 21 \\ u & m & e \end{bmatrix}$$
$$E_{r_{2}\rightarrow r_{3}}A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a & b & c \\ 1 & 2 & 3 \\ u & m & e \end{bmatrix} = \begin{bmatrix} a & b & c \\ 7 & 14 & 21 \\ u & m & e \end{bmatrix}$$

Proposition 3.4.3.

Let $A \in \mathbb{R}^{m \times n}$ then there exist elementary matrices E_1, E_2, \ldots, E_k such that $rref(A) = E_1 E_2 \cdots E_k A$.

Proof: Gauss Jordan elimination consists of a sequence of k elementary row operations. Each row operation can be implemented by multiply the corresponding elementary matrix on the left. The Theorem follows. \Box

Example 3.4.4. Just for fun let's see what happens if we multiply the elementary matrices on the right instead.

$$AE_{r_{2}+3r_{1}\rightarrow r_{2}} = \begin{bmatrix} a & b & c \\ 1 & 2 & 3 \\ u & m & e \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a+3b & b & c \\ 1+6 & 2 & 3 \\ u+3m & m & e \end{bmatrix}$$
$$AE_{7r_{2}\rightarrow r_{2}} = \begin{bmatrix} a & b & c \\ 1 & 2 & 3 \\ u & m & e \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a & 7b & c \\ 1 & 14 & 3 \\ u & 7m & e \end{bmatrix}$$
$$AE_{r_{2}\rightarrow r_{3}} = \begin{bmatrix} a & b & c \\ 1 & 2 & 3 \\ u & m & e \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} a & c & b \\ 1 & 3 & 2 \\ u & e & m \end{bmatrix}$$

Curious, they generate column operations, we might call these elementary column operations. In our notation the row operations are more important.

3.5 invertible matrices

Proposition 3.5.1.

Elementary matrices are invertible.

Proof: I list the inverse matrix for each below:

$$(E_{r_i+ar_j\to r_i})^{-1} = [I: r_i - ar_j \to r_i]$$

$$(E_{br_i \to r_i})^{-1} = [I : \frac{1}{b}r_i \to r_i]$$
$$(E_{r_i \leftrightarrow r_i})^{-1} = [I : r_j \leftrightarrow r_i]$$

I leave it to the reader to convince themselves that these are indeed inverse matrices. \Box

Example 3.5.2. Let me illustrate the mechanics of the proof above, $E_{r_1+3r_2\to r_1} = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $E_{r_1-3r_2\to r_1} = \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ satisfy,

$$E_{r_1+3r_2\to r_1}E_{r_1-3r_2\to r_1} = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Likewise,

$$E_{r_1-3r_2\to r_1}E_{r_1+3r_2\to r_1} = \begin{bmatrix} 1 & -3 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1 \end{bmatrix}$$

Thus, $(E_{r_1+3r_2\to r_1})^{-1} = E_{r_1-3r_2\to r_1}$ just as we expected.

Theorem 3.5.3.

Let $A \in \mathbb{R}^{n \times n}$. The solution of Ax = 0 is unique iff A^{-1} exists.

Proof:(\Rightarrow) Suppose Ax = 0 has a unique solution. Observe A0 = 0 thus the only solution is the zero solution. Consequently, rref[A|0] = [I|0]. Moreover, by Proposition 3.4.3 there exist elementary matrices E_1, E_2, \cdots, E_k such that $rref[A|0] = E_1E_2 \cdots E_k[A|0] = [I|0]$. Applying the concatenation Proposition 3.6.2 we find that $[E_1E_2 \cdots E_kA|E_1E_2 \cdots E_k0] = [I|0]$ thus $E_1E_2 \cdots E_kA = I$.

It remains to show that $AE_1E_2\cdots E_k = I$. Multiply $E_1E_2\cdots E_kA = I$ on the left by E_1^{-1} followed by E_2^{-1} and so forth to obtain

$$E_k^{-1} \cdots E_2^{-1} E_1^{-1} E_1 E_2 \cdots E_k A = E_k^{-1} \cdots E_2^{-1} E_1^{-1} I$$

this simplifies to

$$A = E_k^{-1} \cdots E_2^{-1} E_1^{-1}.$$

Observe that

$$AE_1E_2\cdots E_k = E_k^{-1}\cdots E_2^{-1}E_1^{-1}E_1E_2\cdots E_k = I_1$$

We identify that $A^{-1} = E_1 E_2 \cdots E_k$ thus A^{-1} exists.

(\Leftarrow) The converse proof is much easier. Suppose A^{-1} exists. If Ax = 0 then multiply by A^{-1} on the left, $A^{-1}Ax = A^{-1}0 \Rightarrow Ix = 0$ thus x = 0. \Box

Proposition 3.5.4.

Let $A \in \mathbb{R}^{n \times n}$. 1. If BA = I then AB = I. 2. If AB = I then BA = I.

Proof of (1.): Suppose BA = I. If Ax = 0 then BAx = B0 hence Ix = 0. We have shown that Ax = 0 only has the trivial solution. Therefore, Theorem 3.5.3 shows us that A^{-1} exists. Multiply

BA = I on the left by A^{-1} to find $BAA^{-1} = IA^{-1}$ hence $B = A^{-1}$ and by definition it follows AB = I.

Proof of (2.): Suppose AB = I. If Bx = 0 then ABx = A0 hence Ix = 0. We have shown that Bx = 0 only has the trivial solution. Therefore, Theorem 3.5.3 shows us that B^{-1} exists. Multiply AB = I on the right by B^{-1} to find $ABB^{-1} = IB^{-1}$ hence $A = B^{-1}$ and by definition it follows BA = I. \Box

Proposition 3.5.4 shows that we don't need to check both conditions AB = I and BA = I. If either holds the other condition automatically follows.

Proposition 3.5.5.

If $A \in \mathbb{R}^{n \times n}$ is invertible then its inverse matrix is unique.

Proof: Suppose B, C are inverse matrices of A. It follows that AB = BA = I and AC = CA = I thus AB = AC. Multiply B on the left of AB = AC to obtain BAB = BAC hence $IB = IC \Rightarrow B = C$. \Box

Example 3.5.6. In the case of a 2×2 matrix a nice formula to find the inverse is known:

$$\left[\begin{array}{cc} a & b \\ c & d \end{array}\right]^{-1} = \frac{1}{ad - bc} \left[\begin{array}{cc} d & -b \\ -c & a \end{array}\right]$$

It's not hard to show this formula works,

$$\frac{1}{ad-bc} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad-bc} \begin{bmatrix} ad-bc & -ab+ab \\ cd-dc & -bc+da \end{bmatrix}$$
$$= \frac{1}{ad-bc} \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

How did we know this formula? Can you derive it? To find the formula from first principles you could suppose there exists a matrix $B = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$ such that AB = I. The resulting algebra would lead you to conclude x = d/t, y = -b/t, z = -c/t, w = a/t where t = ad - bc. I leave this as an exercise for the reader.

There is a giant assumption made throughout the last example. What is it?

Example 3.5.7. Recall that a counterclockwise **rotation** by angle θ in the plane can be represented by a matrix $R(\theta) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$. The inverse matrix corresponds to a rotation by angle $-\theta$ and (using the even/odd properties for cosine and sine) $R(-\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} = R(\theta)^{-1}$. Notice that $R(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ thus $R(\theta)R(-\theta) = R(0) = I$. Rotations are very special invertible matrices, we shall see them again.

Theorem 3.5.8.

If $A, B \in \mathbb{R}^{n \times n}$ are invertible, $X, Y \in \mathbb{R}^{m \times n}$, $Z, W \in \mathbb{R}^{n \times m}$ and nonzero $c \in \mathbb{R}$ then 1. $(AB)^{-1} = B^{-1}A^{-1}$, 2. $(cA)^{-1} = \frac{1}{c}A^{-1}$, 3. XA = YA implies X = Y, 4. AZ = AW implies Z = W,

Proof: To prove (1.) simply notice that

$$(AB)B^{-1}A^{-1} = A(BB^{-1})A^{-1} = A(I)A^{-1} = AA^{-1} = I.$$

The proof of (2.) follows from the calculation below,

$$(\frac{1}{c}A^{-1})cA = \frac{1}{c}cA^{-1}A = A^{-1}A = I.$$

To prove (3.) assume that XA = YA and multiply both sides by A^{-1} on the right to obtain $XAA^{-1} = YAA^{-1}$ which reveals XI = YI or simply X = Y. To prove (4.) multiply by A^{-1} on the left. \Box

Remark 3.5.9.

The proofs just given were all matrix arguments. These contrast the *c*omponent level proofs needed for 3.1.9. We could give component level proofs for the Theorem above but that is not necessary and those arguments would only obscure the point. I hope you gain your own sense of which type of argument is most appropriate as the course progresses.

We have a simple formula to calculate the inverse of a 2×2 matrix, but sadly no such simple formula exists for bigger matrices. There is a nice method to calculate A^{-1} (if it exists), but we do not have all the theory in place to discuss it at this juncture.

Proposition 3.5.10.

If
$$A_1, A_2, \dots, A_k \in \mathbb{R}^{n \times n}$$
 are invertible then

$$(A_1 A_2 \cdots A_k)^{-1} = A_k^{-1} A_{k-1}^{-1} \cdots A_1^{-1}$$

Proof: follows from induction on k. In particular, k = 1 is trivial. Assume inductively the proposition is true for some k with $k \ge 2$,

$$(\underbrace{A_1A_2\cdots A_k}_B A_{k+1})^{-1} = (BA_{k+1})^{-1} = A_{k+1}^{-1}B^{-1}$$

by Theorem 3.5.10. Applying the induction hypothesis to B yields

$$(A_1 A_2 \cdots A_{k+1})^{-1} = A_{k+1}^{-1} A_k^{-1} \cdots A_1^{-1} \qquad \Box.$$

3.6 matrix multiplication, again !

In a previous section we proved Proposition 3.2.4 and calculated a number of explicit products. There are cases where a specific matrix is not given and we need to see patterns at the level of rows or columns. In this section we find several new ways to decompose a product which are ideal to reveal such row or column patterns. In some sense, this section is just a special case of the later section on block-multiplication. However, you could probably just as well say block multiplication is a simple outgrowth of what we study here. In any event, we need this material to properly understand the method to calculate A^{-1} and the final proposition of this section is absolutely critical to properly understand the structure of the solution set for Ax = b.

Example 3.6.1. The product of a 2×2 and 2×1 is a 2×1 . Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and let $v = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$,

$$Av = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} = \begin{bmatrix} [1,2][5,7]^T \\ [3,4][5,7]^T \end{bmatrix} = \begin{bmatrix} 19 \\ 43 \end{bmatrix}$$

Likewise, define $w = \begin{bmatrix} 6\\8 \end{bmatrix}$ and calculate

$$Aw = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} = \begin{bmatrix} [1,2][6,8]^T \\ [3,4][6,8]^T \end{bmatrix} = \begin{bmatrix} 22 \\ 50 \end{bmatrix}$$

Something interesting to observe here, recall that in Example 3.2.9 we calculated $AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$. But these are the same numbers we just found from the two matrix-vector products calculated above. We identify that B is just the **concatenation** of the vectors v and w; $B = [v|w] = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$. Observe that:

$$AB = A[v|w] = [Av|Aw].$$

The term **concatenate** is sometimes replaced with the word **adjoin**. I think of the process as gluing matrices together. This is an important operation since it allows us to lump together many solutions into a single matrix of solutions. (I will elaborate on that in detail in a future section)

Proposition 3.6.2.

Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$ then we can understand the matrix multiplication of A and B as the concatenation of several matrix-vector products,

$$AB = A[col_1(B)|col_2(B)|\cdots|col_p(B)] = [Acol_1(B)|Acol_2(B)|\cdots|Acol_p(B)]$$

Proof: see the Problem Set. You should be able to follow the same general strategy as the Proof of Proposition 3.2.4. Show that the i, j-th entry of the L.H.S. is equal to the matching entry on the R.H.S. Good hunting. \Box

There are actually many many different ways to perform the calculation of matrix multiplication. Proposition 3.6.2 essentially parses the problem into a bunch of (matrix)(column vector) calculations. You could go the other direction and view AB as a bunch of (row vector)(matrix) products glued together. In particular,

Proposition 3.6.3.

Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$ then we can understand the matrix multiplication of A and B as the concatenation of several matrix-vector products,

$$AB = \begin{bmatrix} row_1(A) \\ row_2(A) \\ \vdots \\ row_m(A) \end{bmatrix} B = \begin{bmatrix} row_1(A)B \\ row_2(A)B \\ \vdots \\ row_m(A)B \end{bmatrix}$$

Proof: left to reader, but if you ask I'll show you. \Box

There are stranger ways to calculate the product. You can also assemble the product by adding together a bunch of outer-products of the rows of A with the columns of B. The dot-product of two vectors is an example of an inner product and we saw $v \cdot w = v^T w$. The outer-product of two vectors goes the other direction: given $v \in \mathbb{R}^n$ and $w \in \mathbb{R}^m$ we find $vw^T \in \mathbb{R}^{n \times m}$.

Proposition 3.6.4. matrix multiplication as sum of outer products.

Let
$$A \in \mathbb{R}^{m \times n}$$
 and $B \in \mathbb{R}^{n \times p}$ then
 $AB = col_1(A)row_1(B) + col_2(A)row_2(B) + \dots + col_n(A)row_n(B).$

Proof: consider the i, j-th component of AB, by definition we have

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} = A_{i1} B_{1j} + A_{i2} B_{2j} + \dots + A_{in} B_{nj}$$

but note that $(col_k(A)row_k(B))_{ij} = col_k(A)_i row_k(B)_j = A_{ik}B_{kj}$ for each k = 1, 2, ..., n and the proposition follows. \Box

A corollary is a result which falls immediately from a given result. Take the case $B = v \in \mathbb{R}^{n \times 1}$ to prove the following:

Corollary 3.6.5. matrix-column product is linear combination of columns.

Let $A \in \mathbb{R}^{m \times n}$ and $v \in \mathbb{R}^n$ then $Av = v_1 col_1(A) + v_2 col_2(A) + \dots + v_n col_n(A).$

Some texts use the result above as the foundational definition for matrix multiplication. We took a different approach in these notes, largely because I wish for students to gain better grasp of index calculation. If you'd like to know more about the other approach, I can recommend some reading.

3.7 how to calculate the inverse of a matrix

We have not needed to solve more than one problem at a time before, however the problem of calculating an inverse amounts to precisely the problem of simultaneously solving several systems of equations at once. We thus begin with a bit of theory before attacking the inverse problem head-on.

3.7.1 concatenation for solving many systems at once

If we wish to solve $Ax = b_1$ and $Ax = b_2$ we use a concatenation trick to do both at once. In fact, we can do it for $k \in \mathbb{N}$ problems which share the same coefficient matrix but possibly differing inhomogeneous terms.

Proposition 3.7.1.

Let $A \in \mathbb{R}^{m \times n}$. Vectors v_1, v_2, \ldots, v_k are solutions of $Av = b_i$ for $i = 1, 2, \ldots k$ iff $V = [v_1|v_2|\cdots|v_k]$ solves AV = B where $B = [b_1|b_2|\cdots|b_k]$.

Proof: Let $A \in \mathbb{R}^{m \times n}$ and suppose $Av_i = b_i$ for i = 1, 2, ..., k. Let $V = [v_1|v_2|\cdots|v_k]$ and use the concatenation Proposition 3.6.2,

$$AV = A[v_1|v_2|\cdots|v_k] = [Av_1|Av_2|\cdots|Av_k] = [b_1|b_2|\cdots|b_k] = B.$$

Conversely, suppose AV = B where $V = [v_1|v_2|\cdots|v_k]$ and $B = [b_1|b_2|\cdots|b_k]$ then by Proposition 3.6.2 AV = B implies $Av_i = b_i$ for each i = 1, 2, ..., k. \Box

Example 3.7.2. Solve the systems given below,

x + y + z = 1		x + y + z = 1
x - y + z = 0	and	x - y + z = 1
-x + z = 1		-x + z = 1

The systems above share the same coefficient matrix, however $b_1 = [1, 0, 1]^T$ whereas $b_2 = [1, 1, 1]^T$. We can solve both at once by making an extended augmented coefficient matrix $[A|b_1|b_2]$

	[1	1	1	1	1	Γ 1	0	0	-1/4	0	
$[A b_1 b_2] =$	1	-1	1	0	1	$rref[A b_1 b_2] = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$	1	0	1/2	0	
	$\lfloor -1$	0	1	1	1		0	1	3/4	1	

We use Proposition 3.7.1 to conclude that

 $\begin{array}{l} x+y+z=1\\ x-y+z=0\\ -x+z=1 \end{array} \qquad has \ solution \ x=-1/4, y=1/2, z=3/4\\ x+y+z=1\\ x-y+z=1\\ -x+z=1 \end{array} \qquad has \ solution \ x=0, y=0, z=1. \end{array}$

3.7.2 the inverse-finding algorithm

PROBLEM: how should we calculate A^{-1} for a 3×3 matrix ?

Consider that the Proposition 3.7.1 gives us another way to look at the problem,

$$AA^{-1} = I \iff A[v_1|v_2|v_3] = I_3 = [e_1|e_2|e_3]$$

Where $v_i = col_i(A^{-1})$ and $e_1 = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$, $e_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$, $e_3 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$. We observe that the problem of finding A^{-1} for a 3 × 3 matrix amounts to solving three separate systems:

$$Av_1 = e_1, Av_2 = e_2, Av_3 = e_3$$

when we find the solutions then we can construct $A^{-1} = [v_1|v_2|v_3]$. Think about this, if A^{-1} exists then it is unique thus the solutions v_1, v_2, v_3 are likewise unique. Consequently, by Theorem 2.5.3,

$$rref[A|e_1] = [I|v_1], \ rref[A|e_2] = [I|v_2], \ rref[A|e_3] = [I|v_3].$$

Each of the systems above required the same sequence of elementary row operations to cause $A \mapsto I$. We can just as well do them at the same time in one big matrix calculation:

$$rref[A|e_1|e_2|e_3] = [I|v_1|v_2|v_3]$$

While this discuss was done for n = 3 we can just as well do the same for n > 3. This provides the proof for the first sentence of the theorem below. Theorem 2.5.3 together with the discussion above proves the second sentence.

Theorem 3.7.3.

If $A \in \mathbb{R}^{n \times n}$ is invertible then $rref[A|I] = [I|A^{-1}]$. Otherwise, A^{-1} not invertible iff $rref(A) \neq I$ iff $rref[A|I] \neq [I|B]$.

This is perhaps the most pragmatic theorem so far stated in these notes. This theorem tells us how and when we can find an inverse for a square matrix.

Example 3.7.4. Recall that in Example 2.2.5 we worked out the details of

$$rref \begin{bmatrix} 1 & 0 & 0 & | & 1 & 0 & 0 \\ 2 & 2 & 0 & | & 0 & 1 & 0 \\ 4 & 4 & 4 & | & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & | & 1 & 0 & 0 \\ 0 & 1 & 0 & | & -1 & 1/2 & 0 \\ 0 & 0 & 1 & | & 0 & -1/2 & 1/4 \end{bmatrix}$$

Thus,

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 4 & 4 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1/2 & 0 \\ 0 & -1/2 & 1/4 \end{bmatrix}.$$

Example 3.7.5. I omit the details of the Gaussian elimination,

$$rref \begin{bmatrix} 1 & -1 & 0 & | & 1 & 0 & 0 \\ 1 & 0 & -1 & | & 0 & 1 & 0 \\ 6 & 2 & 3 & | & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & | & -2 & -3 & -1 \\ 0 & 1 & 0 & | & -3 & -3 & -1 \\ 0 & 0 & 1 & | & -2 & -4 & -1 \end{bmatrix}$$

Thus,

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 6 & 2 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} -2 & -3 & -1 \\ -3 & -3 & -1 \\ -2 & -4 & -1 \end{bmatrix}.$$

3.7.3 solving systems by inverse matrix

Let us return to the problem we solved via Gauss-Jordan elimination in the previous chapter. If we wish to solve Av = b for $A \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$ then we can calculate rref(A|I) to compute A^{-1} then the solution is obtained simply by multiplying Av = b on the left by A^{-1} ; $A^{-1}Av = v$ thus $v = A^{-1}b$. This is a very silly technique from an efficiency perspective. It is much faster to simply calculate rref(A|b) to find the unique solution. Moreover, when infinitely many solutions exist, we can still find the solution set from rref(A|b). Thus, the technique we discovered in this section is not the best method for solving an explicit, given, system. On the other hand, I often use multiplication by inverse to solve problems which are symbolic.

3.8 symmetric and antisymmetric matrices

Definition 3.8.1.

Let $A \in \mathbb{R}^{n \times n}$. We say A is symmetric iff $A^T = A$. We say A is antisymmetric iff $A^T = -A$.

At the level of components, $A^T = A$ gives $A_{ij} = A_{ji}$ for all i, j. Whereas, $A^T = -A$ gives $A_{ij} = -A_{ji}$ for all i, j. I should mention **skew-symmetric** is another word for antisymmetric. In physics, second rank (anti)symmetric tensors correspond to (anti)symmetric matrices. In electromagnetism, the electromagnetic field tensor has components which can be written as an antisymmetric 4×4 matrix. In classical mechanics, a solid's propensity to spin in various directions is described by the intertia tensor which is symmetric. The energy-momentum tensor from electrodynamics is also symmetric. Matrices are everywhere if look for them.

Example 3.8.2. Some matrices are symmetric:

$$I, O, E_{ii}, \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix}$$

Some matrices are antisymmetric:

$$O, \left[\begin{array}{rrr} 0 & 2\\ -2 & 0 \end{array}\right]$$

Only 0 is both symmetric and antisymmetric (can you prove it?). Many other matrices are neither symmetric nor antisymmetric:

$$e_i, E_{i,i+1}, \begin{bmatrix} 1 & 2\\ 3 & 4 \end{bmatrix}$$

I assumed n > 1 so that e_i is a column vector which is not square.

Proposition 3.8.3.

Let
$$A, B \in \mathbb{R}^{n \times n}$$
 and $c \in \mathbb{R}$ then
1. $(A^T)^T = A$
2. $(AB)^T = B^T A^T$ socks-shoes property for transpose of product
3. $(cA)^T = cA^T$
4. $(A + B)^T = A^T + B^T$
5. $(A^T) = b = (A + b)^T$

5. $(A^T)^{-1} = (A^{-1})^T$.

Proof: To prove (1.) simply note that $((A^T)^T)_{ij} = (A^T)_{ji} = A_{ij}$ for all i, j. Proof of (2.) is left to the reader. Proof of (3.) and (4.) is simple enough,

$$((A + cB)^T)_{ij} = (A + cB)_{ji} = A_{ji} + cB_{ji} = (A^T)_{ij} + ((cB)^T)_{ij}$$

for all i, j. Proof of (5.) is again left to the reader¹. \Box

 $^{^{1}}$ I wouldn't be surprised if I was asked to prove (2.) or (5.) on a quiz or test.

Proposition 3.8.4.

All square matrices are formed by the sum of a symmetric and antisymmetric matrix.

Proof: Let $A \in \mathbb{R}^{n \times n}$. Utilizing Proposition 3.8.3 we find

$$\frac{1}{2}(A + A^T))^T = \frac{1}{2}(A^T + (A^T)^T) = \frac{1}{2}(A^T + A) = \frac{1}{2}(A + A^T)$$

thus $\frac{1}{2}(A + A^T)$ is a symmetric matrix. Likewise,

$$\left(\frac{1}{2}(A-A^T)\right)^T = \frac{1}{2}(A^T-(A^T)^T) = \frac{1}{2}(A^T-A) = -\frac{1}{2}(A-A^T)$$

thus $\frac{1}{2}(A - A^T)$ is an antisymmetric matrix. Finally, note the identity below:

$$A = \frac{1}{2}(A + A^{T}) + \frac{1}{2}(A - A^{T})$$

The theorem follows. \Box

The proof that any function on \mathbb{R} is the sum of an even and odd function uses the same trick.

Example 3.8.5. The proof of the Proposition above shows us how to break up the matrix into its symmetric and antisymmetric pieces:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \frac{1}{2} \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \right) + \frac{1}{2} \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \right)$$
$$= \begin{bmatrix} 1 & 5/2 \\ 5/2 & 4 \end{bmatrix} + \begin{bmatrix} 0 & -1/2 \\ 1/2 & 0 \end{bmatrix}.$$

Example 3.8.6. What are the symmetric and antisymmetric parts of the standard basis E_{ij} in $\mathbb{R}^{n \times n}$? Here the answer depends on the choice of i, j. Note that $(E_{ij})^T = E_{ji}$ for all i, j. Suppose i = j then $E_{ij} = E_{ii}$ is clearly symmetric, thus there is no antisymmetric part. If $i \neq j$ we use the standard trick,

$$E_{ij} = \frac{1}{2}(E_{ij} + E_{ji}) + \frac{1}{2}(E_{ij} - E_{ji})$$

where $\frac{1}{2}(E_{ij}+E_{ji})$ is the symmetric part of E_{ij} and $\frac{1}{2}(E_{ij}-E_{ji})$ is the antisymmetric part of E_{ij} .

Proposition 3.8.7.

Let $A \in \mathbb{R}^{m \times n}$ then $A^T A$ is symmetric.

Proof: Proposition 3.8.3 yields $(A^T A)^T = A^T (A^T)^T = A^T A$. Thus $A^T A$ is symmetric. \Box

Proposition 3.8.8.

If A is symmetric then A^k is symmetric for all $k \in \mathbb{N}$.

Proof: Suppose $A^T = A$. Proceed inductively. Clearly k = 1 holds true since $A^1 = A$. Assume inductively that A^k is symmetric.

 $\begin{aligned} (A^{k+1})^T &= (AA^k)^T & \text{defn. of matrix exponents,} \\ &= (A^k)^T A^T & \text{socks-shoes prop. of transpose,} \\ &= A^k A & \text{using inducition hypothesis.} \\ &= A^{k+1} & \text{defn. of matrix exponents,} \end{aligned}$

thus by proof by mathematical induction A^k is symmetric for all $k \in \mathbb{N}$. \Box

3.9 block matrices

If you look at most undergraduate linear algbera texts they will not bother to even attempt much of a proof that block-multiplication holds in general. I will foolishly attempt it here. However, I'm going to cheat a little and employ uber-sneaky physics notation. The Einstein summation convention states that if an index is repeated then it is assumed to be summed over it's values. This means that the letters used for particular indices are reserved. If i, j, k are used to denote components of a spatial vector then you cannot use them for a spacetime vector at the same time. A typical notation in physics would be that v^j is a vector in xyz-space whereas v^{μ} is a vector in txyz-spacetime. A spacetime vector could be written as a sum of space components and a time component; $v = v^{\mu}e_{\mu} = v^{0}e_{0} + v^{1}e_{1} + v^{2}e_{2} + v^{3}e_{3} = v^{0}e_{0} + v^{j}e_{j}$. This is not the sort of langauge we use in mathematics. For us notation is usually not reserved. Anyway, cultural commentary aside, if we were to use Einstein-type notation in linear algebra then we would likely omit sums as follows:

$$v = \sum_{i} v_i e_i \longrightarrow v = v_i e_i$$

 $A = \sum_{ij} A_{ij} E_{ij} \longrightarrow A = A_{ij} E_{ij}$

We wish to partition a matrices A and B into 4 parts, use indices M, N which split into subindices m, μ and n, ν respectively. In this notation there are 4 different types of pairs possible:

$$A = [A_{MN}] = \begin{bmatrix} A_{mn} & A_{m\nu} \\ \hline A_{\mu n} & A_{\mu\nu} \end{bmatrix} \qquad B = [B_{NJ}] = \begin{bmatrix} B_{nj} & B_{n\gamma} \\ \hline B_{\mu j} & B_{\mu\gamma} \end{bmatrix}$$

Then the sum over M, N breaks into 2 cases,

$$A_{MN}B_{NJ} = A_{Mn}B_{nJ} + A_{M\nu}B_{\nu J}$$

But, then there are 4 different types of M, J pairs,

$$[AB]_{mj} = A_{mN}B_{Nj} = A_{mn}B_{nj} + A_{m\nu}B_{\nu j}$$
$$[AB]_{m\gamma} = A_{mN}B_{N\gamma} = A_{mn}B_{n\gamma} + A_{m\nu}B_{\nu \gamma}$$
$$[AB]_{\mu j} = A_{\mu N}B_{Nj} = A_{\mu n}B_{nj} + A_{\mu\nu}B_{\nu j}$$
$$[AB]_{\mu \gamma} = A_{\mu N}B_{N\gamma} = A_{\mu n}B_{n\gamma} + A_{\mu\nu}B_{\nu \gamma}$$

Let me summarize,

$$\begin{bmatrix} A_{mn} & A_{m\nu} \\ \hline A_{\mu n} & A_{\mu\nu} \end{bmatrix} \begin{bmatrix} B_{nj} & B_{n\gamma} \\ \hline B_{\mu j} & B_{\mu\gamma} \end{bmatrix} = \begin{bmatrix} [A_{mn}][B_{nj}] + [A_{m\nu}][B_{\nu j}] & [A_{mn}][B_{n\gamma}] + [A_{m\nu}][B_{\nu\gamma}] \\ \hline [A_{\mu n}][B_{nj}] + [A_{\mu\nu}][B_{\nu j}] & [A_{\mu n}][B_{n\gamma}] + [A_{\mu\nu}][B_{\nu\gamma}] \end{bmatrix}$$

Let me again summarize, but this time I'll drop the annoying indices:
Theorem 3.9.1. block multiplication.

Suppose $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$ such that both A and B are partitioned as follows:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{bmatrix}$$

where A_{11} is an $m_1 \times n_1$ block, A_{12} is an $m_1 \times n_2$ block, A_{21} is an $m_2 \times n_1$ block and A_{22} is an $m_2 \times n_2$ block. Likewise, $B_{n_k p_k}$ is an $n_k \times p_k$ block for k = 1, 2. We insist that $m_1 + m_2 = m$ and $n_1 + n_2 = n$. If the partitions are compatible as decribed above then we may multiply A and B by multiplying the blocks as if they were scalars and we were computing the product of 2×2 matrices:

$$\begin{bmatrix} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \hline A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$$

To give a careful proof we'd just need to write out many sums and define the partition with care from the outset of the proof. In any event, notice that once you have this partition you can apply it twice to build block-multiplication rules for matrices with more blocks. The basic idea remains the same: you can parse two matrices into matching partitions then the matrix multiplication follows a pattern which is as if the blocks were scalars. However, the blocks are not scalars so the multiplication of the blocks is nonabelian. For example,

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \\ \hline A_{31} & A_{32} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \hline A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \\ \hline A_{31}B_{11} + A_{32}B_{21} & A_{31}B_{12} + A_{32}B_{22} \end{bmatrix}$$

where if the partitions of A and B are compatible it follows that the block-multiplications on the RHS are all well-defined.

Example 3.9.2. Let $R(\theta) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$ and $B(\gamma) = \begin{bmatrix} \cosh(\gamma) & \sinh(\gamma) \\ \sinh(\gamma) & \cosh(\gamma) \end{bmatrix}$. Furthermore construct 4×4 matrices Λ_1 and Λ_2 as follows:

$$\Lambda_1 = \begin{bmatrix} B(\gamma_1) & 0\\ 0 & R(\theta_1) \end{bmatrix} \qquad \qquad \Lambda_2 = \begin{bmatrix} B(\gamma_2) & 0\\ 0 & R(\theta_2) \end{bmatrix}$$

Multiply Λ_1 and Λ_2 via block multiplication:

$$\Lambda_1 \Lambda_2 = \begin{bmatrix} B(\gamma_1) & 0 \\ 0 & R(\theta_1) \end{bmatrix} \begin{bmatrix} B(\gamma_2) & 0 \\ 0 & R(\theta_2) \end{bmatrix}$$
$$= \begin{bmatrix} B(\gamma_1)B(\gamma_2) + 0 & 0 + 0 \\ 0 + 0 & 0 + R(\theta_1)R(\theta_2) \end{bmatrix}$$
$$= \begin{bmatrix} B(\gamma_1 + \gamma_2) & 0 \\ 0 & R(\theta_1 + \theta_2) \end{bmatrix}.$$

The last calculation is actually a few lines in detail, if you know the adding angles formulas for cosine, sine, cosh and sinh it's easy. If $\theta = 0$ and $\gamma \neq 0$ then Λ would represent a velocity boost on spacetime. Since it mixes time and the first coordinate the velocity is along the x-coordinate. On the other hand, if $\theta \neq 0$ and $\gamma = 0$ then Λ gives a rotation in the yz spatial coordinates in space

time. If both parameters are nonzero then we can say that Λ is a Lorentz transformation on spacetime. Of course there is more to say here, perhaps we could offer a course in special relativity if enough students were interested in concert.

Example 3.9.3. Problem: Suppose M is a square matrix with submatrices A, B, C, 0. What conditions should we insist on for $M = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$ to be invertible.

Solution: I propose we partition the potential inverse matrix $M^{-1} = \begin{bmatrix} D & E \\ F & G \end{bmatrix}$. We seek to find conditions on A, B, C such that there exist D, E, F, G and $MM^{-1} = I$. Each block of the equation $MM^{-1} = I$ gives us a separate submatrix equation:

$$MM^{-1} = \begin{bmatrix} A & B \\ \hline 0 & C \end{bmatrix} \begin{bmatrix} D & E \\ \hline F & G \end{bmatrix} = \begin{bmatrix} AD + BF & AE + BG \\ \hline 0D + CF & 0E + CG \end{bmatrix} = \begin{bmatrix} I & 0 \\ \hline 0 & I \end{bmatrix}$$

We must solve simultaneously the following:

(1.) AD + BF = I, (2.) AE + BG = 0, (3.) CF = 0, (4.) CG = I

If C^{-1} exists then $G = C^{-1}$ from (4.). Moreover, (3.) then yields $F = C^{-1}0 = 0$. Our problem thus reduces to (1.) and (2.) which after substituting F = 0 and $G = C^{-1}$ yield

(1.) AD = I, (2.) $AE + BC^{-1} = 0$.

Equation (1.) says $D = A^{-1}$. Finally, let's solve (2.) for E,

$$E = -A^{-1}BC^{-1}.$$

Let's summarize the calculation we just worked through. IF A, C are invertible then the matrix $M = \begin{bmatrix} A & B \\ \hline 0 & C \end{bmatrix}$ is invertible with inverse

$$M^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}BC^{-1} \\ \hline 0 & C^{-1} \end{bmatrix}.$$

Consider the case that M is a 2×2 matrix and $A, B, C \in \mathbb{R}$. Then the condition of invertibility reduces to the simple conditions $A, C \neq 0$ and $-A^{-1}BC^{-1} = \frac{-B}{AC}$ we find the formula:

$$M^{-1} = \begin{bmatrix} \frac{1}{A} & \frac{-B}{AC} \\ 0 & \frac{1}{C} \end{bmatrix} = \frac{1}{AC} \begin{bmatrix} C & -B \\ 0 & A \end{bmatrix}.$$

This is of course the formula for the 2×2 matrix in this special case where $M_{21} = 0$.

Of course the real utility of formulas like those in the last example is that they work for partitions of arbitrary size. If we can find a block of zeros somewhere in the matrix then we may reduce the size of the problem. The time for a computer calculation is largely based on some power of the size of the matrix. For example, if the calculation in question takes n^2 steps then parsing the matrix into 3 nonzero blocks which are $n/2 \times n/2$ would result in something like $[n/2]^2 + [n/2]^2 + [n/2]^2 = \frac{3}{4}n^2$ steps. If the calculation took on order n^3 computer operations (flops) then my toy example of 3 blocks would reduce to something like $[n/2]^3 + [n/2]^3 = \frac{3}{8}n^2$ flops. A savings of more than 60% of computer time. If the calculation was typically order n^4 for an $n \times n$ matrix then the saving

is even more dramatic. If the calculation is a determinant then the cofactor formula depends on the factorial of the size of the matrix. Try to compare 10!+10! verses say 20!. Hope your calculator has a big display:

 $10! = 3628800 \Rightarrow 10! + 10! = 7257600$ or 20! = 2432902008176640000.

Perhaps you can start to appreciate why numerical linear algebra software packages often use algorithms which make use of block matrices to streamline large matrix calculations. If you are very interested in this sort of topic you might strike up a conversation with Dr. Van Voorhis. I suspect he knows useful things about this type of mathematical inquiry.

Finally, I would comment that breaking a matrix into blocks is basically the bread and butter of quantum mechanics. One attempts to find a basis of state vectors which makes the Hamiltonian into a block-diagonal matrix. Each block corresponds to a certain set of statevectors sharing a common energy. The goal of representation theory in physics is basically to break down matrices into blocks with nice physical meanings. On the other hand, abstract algebraists also use blocks to rip apart a matrix into it's most basic form. For linear algebraists², the so-called Jordan form is full of blocks. Wherever reduction of a linear system into smaller subsystems is of interest there will be blocks.

3.10 applications

Definition 3.10.1.

Let $P \in \mathbb{R}^{n \times n}$ with $P_{ij} \ge 0$ for all i, j. If the sum of the entries in any column of P is one then we say P is a stochastic matrix.

Example 3.10.2. Stochastic Matrix: A medical researcher³ is studying the spread of a virus in 1000 lab. mice. During any given week it's estimated that there is an 80% probability that a mouse will overcome the virus, and during the same week there is an 10% likelyhood a healthy mouse will become infected. Suppose 100 mice are infected to start, (a.) how many sick next week? (b.) how many sick in 2 weeks? (c.) after many many weeks what is the steady state solution?

$I_k = infected mice at beginning of week k$, _ [0.2	0.1	
$N_k = noninfected mice at beginning of week k$	_	0.8	0.9	

We can study the evolution of the system through successive weeks by multiply the state-vector $X_k = [I_k, N_k]$ by the probability transition matrix P given above. Notice we are given that $X_1 = [100, 900]^T$. Calculate then,

$$X_2 = \left[\begin{array}{cc} 0.2 & 0.1\\ 0.8 & 0.9 \end{array}\right] \left[\begin{array}{c} 100\\ 900 \end{array}\right] = \left[\begin{array}{c} 110\\ 890 \end{array}\right]$$

After one week there are 110 infected mice Continuing to the next week,

$X_3 =$	0.2	0.1	[110]	_ [· 111]
	0.8	0.9	890	=	889

²mostly dead by now sad to say.

³this example and most of the other applied examples in these notes are borrowed from my undergraduate linear algebra course taught from Larson's text by Dr. Terry Anderson of Appalachian State University

After two weeks we have 111 mice infected. What happens as $k \to \infty$? Generally we have $X_k = PX_{k-1}$. Note that as k gets large there is little difference between k and k-1, in the limit they both tend to infinity. We define the steady-state solution to be $X^* = \lim_{k\to\infty} X_k$. Taking the limit of $X_k = PX_{k-1}$ as $k \to \infty$ we obtain the requirement $X^* = PX^*$. In other words, the steady state solution is found from solving $(P-I)X^* = 0$. For the example considered here we find,

$$(P-I)X^* = \begin{bmatrix} -0.8 & 0.1\\ 0.8 & -0.1 \end{bmatrix} \begin{bmatrix} u\\ v \end{bmatrix} = 0 \qquad v = 8u \qquad X^* = \begin{bmatrix} u\\ 8u \end{bmatrix}$$

However, by conservation of mice, u + v = 1000 hence 9u = 1000 and $u = 111.\overline{11}$ thus the steady state can be shown to be $X^* = [111.\overline{11}, 888.\overline{88}]$

Example 3.10.3. Diagonal matrices are nice: Suppose that demand for doorknobs halves every week while the demand for yo-yos it cut to 1/3 of the previous week's demand every week due to an amazingly bad advertising campaign⁴. At the beginning there is demand for 2 doorknobs and 5 yo-yos.

$$D_k = demand for doorknobs at beginning of week k$$

$$Y_k = demand for yo-yos at beginning of week k$$

$$P = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/3 \end{bmatrix}$$

We can study the evolution of the system through successive weeks by multiply the state-vector $X_k = [D_k, Y_k]$ by the transition matrix P given above. Notice we are given that $X_1 = [2, 5]^T$. Calculate then,

$$X_2 = \begin{bmatrix} 1/2 & 0\\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} 2\\ 5 \end{bmatrix} = \begin{bmatrix} 1\\ 5/3 \end{bmatrix}$$

Notice that we can actually calculate the k-th state vector as follows:

$$X_{k} = P^{k}X_{1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/3 \end{bmatrix}^{k} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 2^{-k} & 0 \\ 0 & 3^{-k} \end{bmatrix}^{k} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 2^{-k+1} \\ 5(3^{-k}) \end{bmatrix}$$

Therefore, assuming this silly model holds for 100 weeks, we can calculate the 100-the step in the process easily,

$$X_{100} = P^{100} X_1 = \begin{bmatrix} 2^{-101} \\ 5(3^{-100}) \end{bmatrix}$$

Notice that for this example the analogue of X^* is the zero vector since as $k \to \infty$ we find X_k has components which both go to zero.

Example 3.10.4. Naive encryption: in Example 3.7.5 we found observed that the matrix A has inverse matrix A^{-1} where:

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 6 & 2 & 3 \end{bmatrix} \qquad A^{-1} = \begin{bmatrix} -2 & -3 & -1 \\ -3 & -3 & -1 \\ -2 & -4 & -1 \end{bmatrix}.$$

We use the alphabet code

$$A = 1, B = 2, C = 3, \dots, Y = 25, Z = 26$$

 $^{^{4}}$ insert your own more interesting set of quantities that doubles/halves or triples during some regular interval of time

and a space is encoded by 0. The words are parsed into row vectors of length 3 then we multiply them by A on the right; [decoded]A = [coded]. Suppose we are given the string, already encoded by A

[9, -1, -9], [38, -19, -19], [28, -9, -19], [-80, 25, 41], [-64, 21, 31], [-7, 4, 7].

Find the hidden message by undoing the multiplication by A. Simply multiply by A^{-1} on the right,

$$[9, -1, -9]A^{-1}, [38, -19, -19]A^{-1}, [28, -9, -19]A^{-1}$$

 $[-80, 25, 41]A^{-1}, [-64, 21, 31]A^{-1}, [-7, 4, 7]A^{-1}$

This yields,

[19, 19, 0], [9, 19, 0], [3, 1, 14], [3, 5, 12], [12, 5, 4]

which reads CLASS IS CANCELLED ⁵.

If you enjoy this feel free to peruse my Math 121 notes, I have additional examples of this naive encryption. I say it's naive since real encryption has much greater sophistication by this time.

Example 3.10.5. Complex Numbers: matrices of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ multiply like complex numbers. For example, consider $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ observe

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -I$$

This matrix plays the role of $i = \sqrt{-1}$ where $i^2 = -1$. Consider,

$$\left[\begin{array}{cc} a & -b \\ b & a \end{array}\right] \left[\begin{array}{cc} x & -y \\ y & x \end{array}\right] = \left[\begin{array}{cc} ax - by & -(ay + bx) \\ ay + bx & ax - by \end{array}\right]$$

Recall, (a+ib)(x+iy) = ax - by + i(ay + bx). These 2×2 matrices form a model of the complex number system.

Many algebraic systems permit a representation via some matrix model.⁶

Example 3.10.6. Jacobian matrix of advanced calculus: Matrix multiplication and the composition of linear operators is the heart of the chain rule in multivariate calculus. The derivative of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ at a point $p \in \mathbb{R}^n$ gives the best linear approximation to f in the sense that

$$L_f(p+h) = f(p) + D_p f(h) \cong f(p+h)$$

if $h \in \mathbb{R}^n$ is close to the zero vector; the graph of L_f gives the tangent line or plane or hypersurface depending on the values of m, n. The so-called Frechet derivative is $D_p f$, it is a linear transformation from \mathbb{R}^n to \mathbb{R}^m . The simplest case is $f : \mathbb{R} \to \mathbb{R}$ where $D_p f(h) = f'(p)h$ and you should recognize $L_f(p+h) = f(p) + f'(p)h$ as the function whose graph is the tangent line, perhaps $L_f(x) = f(p) + f'(p)(x-p)$ is easier to see but it's the same just set p+h = x. Given two functions, say $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^p$ then it can be shown that $D(g \circ f) = Dg \circ Df$. In turn, the

 $^{^5 \}mathrm{Larson's}$ pg. 100-102 # 22

⁶Minh Nguyen, Bailu Zhang and Spencer Leslie worked with me to study the calculus over semisimple algebras. In that work, one important concept is the matrix formulation of the given algebra. I may have an open project which extends that work, ask if interested

matrix of $D(g \circ f)$ is simply obtain by multiplying the matrices of Dg and Df. The matrix of the Frechet derivative is called the Jacobian matrix. The determinant of the Jacobian matrix plays an important role in changing variables for multiple integrals. It is likely we would cover this discussion in some depth in the Advanced Calculus course, while linear algebra is not a pre-req, it sure would be nice if you had it. Linear is truly foundational for most interesting math.

3.11 conclusions

The theorem that follows here collects the various ideas we have discussed concerning an $n \times n$ matrix and invertibility and solutions of Ax = b.

Theorem 3.11.1.

Let A be a real $n \times n$ matrix then the following are equivalent:

- (a.) A is invertible,
- (b.) rref[A|0] = [I|0] where $0 \in \mathbb{R}^n$,
- (c.) Ax = 0 iff x = 0,
- (d.) A is the product of elementary matrices,
- (e.) there exists $B \in \mathbb{R}^{n \times n}$ such that AB = I,
- (f.) there exists $B \in \mathbb{R}^{n \times n}$ such that BA = I,
- (g.) rref[A] = I,
- (h.) rref[A|b] = [I|x] for an $x \in \mathbb{R}^n$,
- (i.) Ax = b is consistent for every $b \in \mathbb{R}^n$,
- (j.) Ax = b has exactly one solution for every $b \in \mathbb{R}^n$,
- (k.) A^T is invertible.

These are in no particular order. If you examine the arguments in this chapter you'll find we've proved most of this theorem. What did I miss? 7

⁷teaching moment or me trying to get you to do my job, you be the judge.

Chapter 4

linear independence and spanning

Spanning and Linear Independence (LI) are arguably the most important topics in linear algebra. In this chapter we discuss spanning and linear independence in the context of \mathbb{R}^n . We begin by developing the necessary matrix result. Then the idea of spanning is explained and a number of explicit examples are given. We also see how to solve several spanning questions simultaneously. Then we turn to the question of minimality. How can we reduce the size of the spanning set while maintaining the span? This requires us to introduce the concept of LI. A fundamental proposition is proved and we again see how to solve the typical problem with a matrix technique. Next, we learn how the Column Correspondance Property (CCP) gives efficient solutions to all the questions we faced in this chapter (and much more later). This current context is however very special, later in the course we will not be quite as free to use the CCP directly. The problems we face here are particularly simple. Finally, we draw together results about spanning, LI and matrix invertibility. This continues a series of theorems we saw in previous chapters.

4.1 matrix notation for systems

Let us begin with a simple example.

Example 4.1.1. Consider the following generic system of two equations and three unknowns,

$$ax + by + cz = d$$
$$ex + fy + gz = h$$

in matrix form this system of equations is Av = b where

$$Av = \underbrace{\begin{bmatrix} a & b & c \\ e & f & g \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}_{v} = \begin{bmatrix} (a, b, c) \cdot (x, y, z) \\ (e, f, g) \cdot (x, y, z) \end{bmatrix} = \begin{bmatrix} ax + by + cz \\ ex + fy + gz \end{bmatrix} = \underbrace{\begin{bmatrix} d \\ h \end{bmatrix}}_{b}$$

Definition 4.1.2.

Let x_1, x_2, \ldots, x_k be k variables and suppose $b_i, A_{ij} \in \mathbb{R}$ for $1 \le i \le k$ and $1 \le j \le n$. The system of linear equations

$$A_{11}x_1 + A_{12}x_2 + \dots + A_{1k}x_k = b_1$$

$$A_{21}x_1 + A_{22}x_2 + \dots + A_{2k}x_k = b_2$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots$$

$$A_{n1}x_1 + A_{n2}x_2 + \dots + A_{nk}x_k = b_n$$

has coefficient matrix A, the inhomogeneous term b and augmented coefficient matrix [A|b] defined below:

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nk} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \quad [A|b] = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} & b_1 \\ A_{21} & A_{22} & \cdots & A_{2k} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nk} & b_n \end{bmatrix}.$$

A vector $x \in \mathbb{R}^k$ for which Ax = b is called a **vector solution** to the matrix form of the system. Also, the solution set is $Sol_{[A|b]} = \{x \in \mathbb{R}^k \mid Ax = b\}.$

Naturally, solutions x_1, x_2, \ldots, x_k to the original system are in 1-1 correspondence with the vector solutions of the corresponding matrix form of the equation. Moreover, from Chapter 2 we know Gauss-Jordan elimination on the augmented coefficient matrix is a reliable algorithm to solve any such system.

Example 4.1.3. We found that the system in Example 2.3.1,

$$x + 2y - 3z = 1$$

$$2x + 4y = 7$$

$$-x + 3y + 2z = 0$$

has the unique solution x = 83/30, y = 11/30 and z = 5/6. This means the matrix equation Av = b where

$$Av = \underbrace{\begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & 0 \\ -1 & 3 & 2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{v} = \underbrace{\begin{bmatrix} 1 \\ 7 \\ 0 \end{bmatrix}}_{b} \quad has \ vector \ solution \quad \begin{bmatrix} v = \begin{bmatrix} 83/30 \\ 11/30 \\ 5/6 \end{bmatrix}.$$

Example 4.1.4. We can rewrite the following system of linear equations

$$x_1 + x_4 = 0$$

$$2x_1 + 2x_2 + x_5 = 0$$

$$4x_1 + 4x_2 + 4x_3 = 1$$

in matrix form this system of equations is Av = b where

$$Av = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 2 & 2 & 0 & 0 & 1 \\ 4 & 4 & 4 & 0 & 0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}}_{v} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{b}.$$

Gaussian elimination on the augmented coefficient matrix reveals (see Example 2.2.5 for details of the Gaussian elimination)

Consequently, x_4, x_5 are free and solutions are of the form

$$x_1 = -x_4 x_2 = x_4 - \frac{1}{2}x_5 x_3 = \frac{1}{4} + \frac{1}{2}x_5$$

for all $x_4, x_5 \in \mathbb{R}$. The vector form of the solution is as follows:

Remark 4.1.5.

You might ask the question: what is the geometry of the solution set above ? Let $S = Sol_{[A|b]} \subset \mathbb{R}^5$, we see S is formed by tracing out all possible linear combinations of the vectors $v_1 = (-1, 1, 0, 1, 0)$ and $v_2 = (0, -\frac{1}{2}, \frac{1}{2}, 0, 1)$ based from the point $p_o = (0, 0, \frac{1}{4}, 0, 0)$. In other words, this is a two-dimensional plane containing the vectors v_1, v_2 and the point p_o . This plane is placed in a 5-dimensional space, this means that at any point on the plane you could go in three different directions away from the plane.

We saw in Section 3.6 there are a number of interesting ways to look at matrix multiplication. One important view was given in Corollary 3.6.5. For example:

Example 4.1.6. Let $A = \begin{bmatrix} 1 & 1 & 1 \\ a & b & c \end{bmatrix}$ and $v = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ then we may calculate the product Av as

follows:

$$Av = \begin{bmatrix} 1 & 1 & 1 \\ a & b & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = x \begin{bmatrix} 1 \\ a \end{bmatrix} + y \begin{bmatrix} 1 \\ b \end{bmatrix} + z \begin{bmatrix} 1 \\ c \end{bmatrix} = \begin{bmatrix} x+y+z \\ ax+by+cz \end{bmatrix}.$$

In general, a Corollary to Corollary 3.6.5 is simply:

Proposition 4.1.7.

If $A = [A_1|A_2|\cdots|A_k] \in \mathbb{R}^{n \times k}$ and $b \in \mathbb{R}^n$ then the matrix equation Ax = b has the same set of solutions as the vector equation

 $x_1A_1 + x_2A_2 + \dots + x_kA_k = b.$

Moreover, the solution set is given by Gauss-Jordan reduction of $[A_1|A_2|\cdots|A_k|b]$.

4.2 linear combinations and spanning

Proposition 1.5.8 showed that linear combinations of the standard basis will generate any vector in \mathbb{R}^n . A natural generalization of that question is given below:

PROBLEM: Given vectors v_1, v_2, \ldots, v_k and a vector b do there exist constants c_1, c_2, \ldots, c_k such that $c_1v_1 + c_2v_2 + \cdots + c_kv_k = b$? If so, how to find c_1, \ldots, c_k ?

We have all the tools we need to solve such problems. Ultimately, the CCP gives us the most efficient solution, However, I think it is best for us to work our way with less optimal methods before we learn the fastest method. For now, we just use Proposition 4.1.7 or common sense.

Example 4.2.1. Problem: given that v = (2, -1, 3), w = (1, 1, 1) and b = (4, 1, 5) find values for x, y such that xv + yw = b (if possible).

Solution: using our column notation we find xv + yw = b gives

	2		1		4		$\begin{bmatrix} 2x+y \end{bmatrix}$		4
x	-1	+y	1	=	1	\Rightarrow	-x+y	=	1
	3		1		5		3x+y		5

We are faced with solving the system of equations 2x + y = 4, -x + y = 1 and 3x + y = 5. As we discussed in depth last chapter we can efficiently solve this type of problem in general by Gaussian elimination on the corresponding augmented coefficient matrix. In this problem, you can calculate that

$$rref \begin{bmatrix} 2 & 1 & | & 4 \\ -1 & 1 & | & 1 \\ 3 & 1 & | & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & | & 1 \\ 0 & 1 & | & 2 \\ 0 & 0 & | & 0 \end{bmatrix}$$

hence x = 1 and y = 2. Indeed, it is easy to check that v + 2w = b.

The geometric question which is equivalent to the previous question is as follows: "is the vector b found in the plane which contains v and w"? Here's a picture of the calculation we just performed:



The set of all linear combinations of several vectors in \mathbb{R}^n is called the *span* of those vectors. To be precise

Definition 4.2.2.

Let $S = \{v_1, v_2, \dots, v_k\} \subset \mathbb{R}^n$ be a finite set of *n*-vectors then span(S) is defined to be the set of all linear combinations formed from vectors in S:

$$span\{v_1, v_2, \dots, v_k\} = \{\sum_{i=1}^k c_i v_i \mid c_i \in \mathbb{R} \text{ for } i = 1, 2, \dots, k\}$$

If W = span(S) then we say that S is a generating set for W.

If we have one vector then it has a span which could be a line. With two vectors we might generate a plane. With three vectors we might generate a volume. With four vectors we might generate a hypervolume or 4-volume. We'll return to these geometric musings in § 4.3 and explain why I have used the word "might" rather than an affirmative "will" in these claims. For now, we return to the question of how to decide if a given vector is in the span of another set of vectors.

Example 4.2.3. Problem: Let $b_1 = (1, 1, 0), b_2 = (0, 1, 1)$ and $b_3 = (0, 1, -1)$. Is¹ $e_3 \in span\{b_1, b_2, b_3\}$?

Solution: Find the explicit linear combination of b_1, b_2, b_3 that produces e_3 . We seek to find $x, y, z \in \mathbb{R}$ such that $xb_1 + yb_2 + zb_3 = e_3$,

$$x \begin{bmatrix} 1\\1\\0 \end{bmatrix} + y \begin{bmatrix} 0\\1\\1 \end{bmatrix} + z \begin{bmatrix} 0\\1\\-1 \end{bmatrix} = \begin{bmatrix} 0\\0\\1 \end{bmatrix} \implies \begin{bmatrix} x\\x+y+z\\y-z \end{bmatrix} = \begin{bmatrix} 0\\0\\1 \end{bmatrix}$$

Following essentially the same arguments as the last example we find this question of solving the system formed by gluing the given vectors into a matrix and doing row reduction. In particular, we

¹challenge: once you understand this example for e_3 try answering it for other vectors or for an arbitrary vector $v = (v_1, v_2, v_3)$. How would you calculate $x, y, z \in \mathbb{R}$ such that $v = xb_1 + yb_2 + zb_3$?

can solve the vector equation above by solving the corresponding system below:

$$\begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 1 & 1 & 1 & | & 0 \\ 0 & 1 & -1 & | & 1 \end{bmatrix} \xrightarrow{r_2 - r_1 \to r_2} \begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & 1 & | & 0 \\ 0 & 1 & -1 & | & 1 \end{bmatrix} \xrightarrow{r_3 - r_2 \to r_3} \begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & -1 & | & 1 \end{bmatrix} \xrightarrow{r_3 - r_2 \to r_3} \begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & -2 & | & 1 \end{bmatrix} \xrightarrow{r_3 - r_2 \to r_3} \begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & | & 1/2 \\ 0 & 0 & 1 & | & -1/2 \end{bmatrix}$$

Therefore, $x = 0, y = \frac{1}{2}$ and $z = -\frac{1}{2}$. We find that $e_3 = \frac{1}{2}b_1 + \frac{1}{2}b_2 - \frac{1}{2}b_3$ thus $e_3 \in span\{b_1, b_2, b_3\}$.

The power of the matrix technique is shown in the next example.

Example 4.2.4. Problem: Let $b_1 = (1, 2, 3, 4)$, $b_2 = (0, 1, 0, 1)$ and $b_3 = (0, 0, 1, 1)$. Is $w = (1, 1, 4, 4) \in span\{b_1, b_2, b_3\}$?

Solution: Following the same method as the last example we seek to find x_1, x_2 and x_3 such that $x_1b_1 + x_2b_2 + x_3b_3 = w$ by solving the aug. coeff. matrix as is our custom:

$$\begin{bmatrix} b_1|b_2|b_3|w] = \begin{bmatrix} 1 & 0 & 0 & | & 1\\ 2 & 1 & 0 & | & 1\\ 3 & 0 & 1 & | & 4\\ 4 & 1 & 1 & | & 4 \end{bmatrix} \xrightarrow{\begin{array}{c} r_2 - 2r_1 \to r_2\\ r_3 - 3r_1 \to r_3\\ r_4 - 4r_1 \to r_4 \end{array}} \begin{bmatrix} 1 & 0 & 0 & | & 1\\ 0 & 1 & 0 & | & -1\\ 0 & 0 & 1 & | & 1\\ 0 & 1 & 1 & | & 0 \end{bmatrix} \xrightarrow{\begin{array}{c} r_4 - r_2 \to r_4\\ r_4 - r_2 \to r_4\\ \hline 0 & 1 & 1 & | & 0 \end{bmatrix}} \begin{bmatrix} 1 & 0 & 0 & | & 1\\ 0 & 1 & 0 & | & 1\\ 0 & 0 & 1 & | & 1\\ 0 & 0 & 1 & | & 1 \end{bmatrix} \xrightarrow{\begin{array}{c} r_4 - r_3 \to r_4\\ r_4 - r_3 \to r_4\\ \hline 0 & 0 & 0 & | & 0 \end{bmatrix}} \begin{bmatrix} 1 & 0 & 0 & | & 1\\ 0 & 1 & 0 & | & -1\\ 0 & 0 & 1 & | & 1\\ 0 & 0 & 0 & | & 0 \end{bmatrix} = rref[b_1|b_2|b_3|w]$$

We find $x_1 = 1, x_2 = -1, x_3 = 1$ thus $w = b_1 - b_2 + b_3$. Therefore, $w \in span\{b_1, b_2, b_3\}$.

Pragmatically, if the question is sufficiently simple you may not need to use the augmented coefficient matrix to solve the question. I use them here to illustrate the method.

Example 4.2.5. Problem: Let $b_1 = (1, 1, 0)$ and $b_2 = (0, 1, 1)$. Is $e_2 \in span\{b_1, b_2\}$?

Solution: Attempt to find the explicit linear combination of b_1, b_2 that produces e_2 . We seek to find $x, y \in \mathbb{R}$ such that $xb_1 + yb_2 = e_3$,

$$x \begin{bmatrix} 1\\1\\0 \end{bmatrix} + y \begin{bmatrix} 0\\1\\1 \end{bmatrix} = \begin{bmatrix} 0\\1\\0 \end{bmatrix} \implies \begin{bmatrix} x\\x+y\\y \end{bmatrix} = \begin{bmatrix} 0\\1\\0 \end{bmatrix}$$

We don't really need to consult the augmented matrix to solve this problem. Clearly x = 0 and y = 0 is found from the first and third components of the vector equation above. But, the second component yields x + y = 1 thus 0 + 0 = 1. It follows that this system is inconsistent and we may

conclude that $w \notin span\{b_1, b_2\}$. For the sake of curiousity let's see how the augmented solution matrix looks in this case: omitting details of the row reduction,

$$rref \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

note the last row again confirms that this is an inconsistent system.

4.2.1 solving several spanning questions simultaneously

If we are given $B = \{b_1, b_2, \ldots, b_k\} \subset \mathbb{R}^n$ and $T = \{w_1, w_2, \ldots, w_r\} \subset \mathbb{R}^n$ and we wish to determine if $T \subset span(B)$ then we can answer the question by examining if $[b_1|b_2|\cdots|b_k]x = w_j$ has a solution for each $j = 1, 2, \ldots r$. Or we could make use of Proposition 3.7.1 and solve it in one sweeping matrix calculation;

$$rref[b_1|b_2|\cdots|b_k|w_1|w_2|\cdots|w_r]$$

If there is a row with zeros in the first k-columns and a nonzero entry in the last r-columns then this means that at least one vector w_k is not in the span of B(moreover, the vector not in the span corresponds to the nonzero entrie(s)). Otherwise, each vector is in the span of B and we can read the precise linear combination from the matrix. I will illustrate this in the example that follows.

Example 4.2.6. Let $W = span\{e_1 + e_2, e_2 + e_3, e_1 - e_3\}$ and suppose $T = \{e_1, e_2, e_3 - e_1\}$. Is $T \leq W$? If not, which vectors in T are not in W? Consider,

$$\begin{bmatrix} e_1 + e_1 | e_2 + e_3 | e_1 - e_3 | | e_1 | e_2 | e_3 - e_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & | & 1 & 0 & -1 \\ 1 & 1 & 0 & | & 0 & 1 & 0 \\ 0 & 1 & -1 & | & 0 & -1 \\ 0 & 1 & -1 & | & -1 & 1 & 1 \\ 0 & 1 & -1 & | & 0 & 0 & 1 \end{bmatrix} \xrightarrow{r_2 - r_1 \to r_2} \begin{bmatrix} 1 & 0 & 1 & | & 1 & 0 & -1 \\ 0 & 1 & -1 & | & 0 & 0 & 1 \end{bmatrix} \xrightarrow{r_2 + r_3 \to r_2} \begin{bmatrix} 1 & 0 & 1 & | & 1 & 0 & -1 \\ 0 & 1 & -1 & | & 0 & 0 & | & 1 & -1 & 0 \end{bmatrix} \xrightarrow{r_2 + r_3 \to r_2} \xrightarrow{r_1 \to r_2} \begin{bmatrix} 1 & 0 & 1 & | & 1 & 0 & -1 \\ 0 & 1 & -1 & | & 0 & 0 & | & 1 & -1 & 0 \end{bmatrix} \xrightarrow{r_2 + r_3 \to r_2} \xrightarrow{r_1 \to r_2} \xrightarrow{r_2 \to r_3} \xrightarrow{r_1 \to r_2} \xrightarrow{r_2 \to r_3} \xrightarrow{r_1 \to r_2} \xrightarrow{r_2 \to r_3} \xrightarrow{r_3 \to r_4} \xrightarrow{r_4 \to r_4} \xrightarrow{r_4$$

Let me summarize the calculation:

$$rref[e_1 + e_2|e_2 + e_3||e_1 - e_3|e_1|e_2|e_3 - e_1] = \begin{bmatrix} 1 & 0 & 1 & | & 0 & 1 & -1 \\ 0 & 1 & -1 & | & 0 & 0 & 1 \\ 0 & 0 & 0 & | & 1 & -1 & 0 \end{bmatrix}$$

We deduce that e_1 and e_2 are not in W. However, $e_1 - e_3 \in W$ and we can read from the matrix $-(e_1 + e_2) + (e_2 + e_3) = e_3 - e_1$. I added the double vertical bar for book-keeping purposes, as usual the vertical bars are just to aid the reader in parsing the matrix.

4.3 linear independence

In the previous sections we have only considered questions based on a fixed spanning set². We asked if $b \in span\{v_1, v_2, \ldots, v_n\}$ and we even asked if it was possible for all b. What we haven't thought about yet is the following:

PROBLEM: Given vectors v_1, v_2, \ldots, v_k and a vector $b = c_1v_1 + c_2v_2 + \cdots + c_kv_k$ for some constants c_j is it possible that b can be written as a linear combination of some subset of $\{v_1, v_2, \ldots, v_k\}$? If so, how should we determine which vectors can be taken away from the spanning set? How should we decide which vectors to keep and which are redundant?

The span of a set of vectors is simply all possible finite linear combinations of vectors from the set. If you think about it, we don't need a particular vector in the generating set if that vector can be written as a linear combination of other vectors in the generating set. To solve the problem stated above we need to remove linear dependencies of the generating set.

Definition 4.3.1.

If a vector v_k can be written as a linear combination of vectors $\{v_1, v_2, \ldots, v_{k-1}\}$ then we say that the vectors $\{v_1, v_2, \ldots, v_{k-1}, v_k\}$ are **linearly dependent**. If the vectors $\{v_1, v_2, \ldots, v_{k-1}, v_k\}$ are not linear dependent then they are said to be **linearly independent**.

Example 4.3.2. Let $v = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^T$ and $w = \begin{bmatrix} 2 & 4 & 6 \end{bmatrix}^T$. Clearly v, w are linearly dependent since w = 2v.

I often quote the following proposition as the definition of linear independence, it is an equivalent statement and as such can be used as the definition(but not by us, I already made the definition above). If this was our definition then our definition would become a proposition. Math always has a certain amount of this sort of ambiguity.

Proposition 4.3.3.

Let $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$. The set of vectors $\{v_1, v_2, \ldots, v_k\}$ is linearly independent iff $c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0 \implies c_1 = c_2 = \cdots = c_k = 0.$

Proof: (\Rightarrow) Suppose $\{v_1, v_2, \ldots, v_k\}$ is linearly independent. Assume that there exist constants c_1, c_2, \ldots, c_k such that

$$c_1v_1 + c_2v_2 + \dots + c_kv_k = 0$$

and at least one constant, say c_j , is nonzero. Then we can divide by c_j to obtain

$$\frac{c_1}{c_j}v_1 + \frac{c_2}{c_j}v_2 + \dots + v_j + \dots + \frac{c_k}{c_j}v_k = 0$$

solve for v_j , (we mean for $\hat{v_j}$ to denote the deletion of v_j from the list)

$$v_j = -\frac{c_1}{c_j}v_1 - \frac{c_2}{c_j}v_2 - \dots - \hat{v_j} - \dots - \frac{c_k}{c_j}v_k$$

 $^{^{2}}$ sometimes I call it the spanning set, other times the generating set. It turns out that a given space may be generated in many different ways. This section begins the quest to unravel that puzzle

but this means that v_j linearly depends on the other vectors hence $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent. This is a contradiction, therefore $c_j = 0$. Note j was arbitrary so we may conclude $c_j = 0$ for all j. Therefore, $c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0 \Rightarrow c_1 = c_2 = \cdots = c_k = 0$.

Proof: (\Leftarrow) Assume that

$$c_1v_1 + c_2v_2 + \dots + c_kv_k = 0 \implies c_1 = c_2 = \dots = c_k = 0$$

If $v_j = b_1v_1 + b_2v_2 + \cdots + \widehat{b_jv_j} + \cdots + b_kv_k$ then $b_1v_1 + b_2v_2 + \cdots + b_jv_j + \cdots + b_kv_k = 0$ where $b_j = -1$, this is a contradiction. Therefore, for each j, v_j is not a linear combination of the other vectors. Consequently, $\{v_1, v_2, \ldots, v_k\}$ is linearly independent.

Example 4.3.4. Let $v = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^T$ and $w = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$. Let's prove these are linearly independent. Assume that $c_1v + c_2w = 0$, this yields

$$c_1 \begin{bmatrix} 1\\2\\3 \end{bmatrix} + c_2 \begin{bmatrix} 1\\0\\0 \end{bmatrix} = \begin{bmatrix} 0\\0\\0 \end{bmatrix}$$

thus $c_1 + c_2 = 0$ and $2c_1 = 0$ and $3c_1 = 0$. We find $c_1 = c_2 = 0$ thus v, w are linearly independent. Alternatively, you could explain why there does not exist any $k \in \mathbb{R}$ such that v = kw

Think about this, if the set of vectors $\{v_1, v_2, \ldots, v_k\} \subset \mathbb{R}^n$ is linearly independent then the equation $c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0$ has the unique solution $c_1 = 0, c_2 = 0, \ldots, c_k = 0$. Notice we can reformulate the problem as a matrix equation:

$$c_1v_1 + c_2v_2 + \dots + c_kv_k = 0 \iff [v_1|v_2|\cdots|v_k][c_1 \ c_2 \ \cdots \ c_k]^T = 0$$

The matrix $[v_1|v_2|\cdots|v_k]$ is an $n \times k$. This is great. We can use the matrix techniques we already developed to probe for linear independence of a set of vectors.

Proposition 4.3.5.

Let $\{v_1, v_2, \ldots, v_k\}$ be a set of vectors in \mathbb{R}^n .

- 1. If $rref[v_1|v_2|\cdots|v_k]$ has less than k pivot columns then the set of vectors $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent.
- 2. If $rref[v_1|v_2|\cdots|v_k]$ has k pivot columns then the set of vectors $\{v_1, v_2, \ldots, v_k\}$ is linearly independent.

Proof: Denote $V = [v_1|v_2|\cdots|v_k]$ and $c = [c_1, c_2, \ldots, c_k]^T$. If V contains a linearly independent set of vectors then we must find that Vc = 0 implies c = 0. Consider Vc = 0, this is equivalent to using Gaussian elimination on the augmented coefficient matrix [V|0]. We know this system is consistent since c = 0 is a solution. Thus Theorem 2.5.1 tells us that there is either a unique solution or infinitely many solutions.

Clearly if the solution is unique then c = 0 is the only solution and hence the implication Av = 0

implies c = 0 holds true and we find the vectors are linearly independent. We find

$$rref[v_1|v_2|\cdots|v_k] = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{vmatrix} = \begin{bmatrix} I_k \\ 0 \end{bmatrix}$$

where there are *n*-rows in the matrix above. If n = k then there would be no zero row. If there are infinitely many solutions then there will be free variables in the solution of Vc = 0. If we set the free variables to 1 we then find that Vc = 0 does not imply c = 0 since at least the free variables are nonzero. Thus the vectors are linearly dependent in this case, proving (2.). \Box

Before I get to the examples let me glean one more fairly obvious statement from the proof above:

Corollary 4.3.6.

```
If \{v_1, v_2, \ldots, v_k\} is a set of vectors in \mathbb{R}^n and k > n then the vectors are linearly dependent.
```

Proof: Proposition 4.3.5 tells us that the set is linearly independent if there are k pivot columns in $[v_1|\cdots|v_k]$. However, that is impossible since k > n this means that there will be at least one column of zeros in $rref[v_1|\cdots|v_k]$. Therefore the vectors are linearly dependent. \Box

This Proposition is obvious but useful. We may have at most 2 linearly independent vectors in \mathbb{R}^2 , 3 in \mathbb{R}^3 , 4 in \mathbb{R}^4 , and so forth...

Example 4.3.7. Determine if v_1, v_2, v_3 (given below) are linearly independent or dependent. If the vectors are linearly dependent show how they depend on each other.

$$v_1 = \begin{bmatrix} 1\\1\\1 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 2\\1\\0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 3\\2\\1 \end{bmatrix}$$

We seek to use the Proposition 4.3.5. Consider then,

$$\begin{bmatrix} v_1 | v_2 | v_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3\\ 1 & 1 & 2\\ 1 & 0 & 1 \end{bmatrix} \xrightarrow{r_2 - r_1 \to r_2} \begin{bmatrix} 1 & 2 & 3\\ 0 & -1 & -1\\ 0 & -2 & -2 \end{bmatrix} \xrightarrow{r_1 + 2r_2 \to r_2} \begin{bmatrix} 1 & 0 & 1\\ 0 & -1 & -1\\ 0 & 0 & 0 \end{bmatrix}$$

Thus we find that,

$$rref[v_1|v_2|v_3] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

hence the variable c_3 is free in the solution of Vc = 0. We find solutions of the form $c_1 = -c_3$ and $c_2 = -c_3$. This means that

 $-c_3v_1 - c_3v_2 + c_3v_3 = 0$

for any value of c_3 . I suggest $c_3 = 1$ is easy to plug in,

 $-v_1 - v_2 + v_3 = 0$ or we could write $v_3 = v_1 + v_2$

We see clearly that v_3 is a linear combination of v_1, v_2 .

Example 4.3.8. Determine if v_1, v_2, v_3, v_4 (given below) are linearly independent or dependent.

$$v_1 = \begin{bmatrix} 1\\0\\0\\0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 1\\1\\0\\0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 1\\1\\1\\0 \end{bmatrix} \qquad v_4 = \begin{bmatrix} 1\\1\\1\\1 \end{bmatrix}$$

We seek to use the Proposition 4.3.5. Omitting details we find,

$$rref[v_1|v_2|v_3|v_4] = rref \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

In this case no variables are free, the only solution is $c_1 = 0, c_2 = 0, c_3 = 0, c_4 = 0$ hence the set of vectors $\{v_1, v_2, v_3, v_4\}$ is linearly independent.

Example 4.3.9. Determine if v_1, v_2, v_3 (given below) are linearly independent or dependent. If the vectors are linearly dependent show how they depend on each other.

$$v_1 = \begin{bmatrix} 1\\0\\0\\3 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 3\\1\\2\\0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 2\\1\\2\\-3 \end{bmatrix}$$

We seek to use the Proposition 4.3.5. Consider $[v_1|v_2|v_3] =$

$$\begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \\ 3 & 0 & -3 \end{bmatrix} \xrightarrow{r_4 - 3r_1 \to r_4} \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & -9 & -9 \end{bmatrix} \xrightarrow{r_1 - 3r_2 \to r_1} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = rref[V].$$

Hence the variable c_3 is free in the solution of Vc = 0. We find solutions of the form $c_1 = c_3$ and $c_2 = -c_3$. This means that

$$c_3v_1 - c_3v_2 + c_3v_3 = 0$$

for any value of c_3 . I suggest $c_3 = 1$ is easy to plug in,

 $v_1 - v_2 + v_3 = 0$ or we could write $v_3 = v_2 - v_1$

We see clearly that v_3 is a linear combination of v_1, v_2 .

Example 4.3.10. Determine if v_1, v_2, v_3, v_4 (given below) are linearly independent or dependent. If the vectors are linearly dependent show how they depend on each other.

$$v_1 = \begin{bmatrix} 0\\0\\1\\0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 0\\1\\0\\0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 0\\1\\1\\0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 0\\1\\2\\0 \end{bmatrix}$$

We seek to use the Proposition 4.3.5. Consider $[v_1|v_2|v_3|v_4] =$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_3} \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{r_1 - r_2 \to r_1} \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = rref[v_1|v_2|v_3|v_4]$$

Hence the variables c_3 and c_4 are free in the solution of Vc = 0. We find solutions of the form $c_1 = -c_3 + c_4$ and $c_2 = -c_3 - c_4$. This means that

$$(c_4 - c_3)v_1 - (c_3 + c_4)v_2 + c_3v_3 + c_4v_4 = 0$$

for any value of c_3 or c_4 . I suggest $c_3 = 1, c_4 = 0$ is easy to plug in,

$$-v_1 - v_2 + v_3 = 0$$
 or we could write $v_3 = v_2 + v_1$

Likewise select $c_3 = 0, c_4 = 1$ to find

$$v_1 - v_2 + v_4 = 0$$
 or we could write $v_4 = v_2 - v_1$

We find that v_3 and v_4 are linear combinations of v_1 and v_2 .

Let's pause to reflect on the geometric meaning of the examples above.

Remark 4.3.11.

For two vectors the term "linearly dependent" can be taken quite literally: two vectors are linearly dependent if they point along the same line. For three vectors they are linearly dependent if they point along the same line or possibly lay in the same plane. When we get to four vectors we can say they are linearly dependent if they reside in the same volume, plane or line. I don't find the geometric method terribly successful for dimensions higher than two. However, it is neat to think about the geometric meaning of certain calculations in dimensions higher than 3. We can't even draw it but we can eulicidate all sorts of information with the mathematics of linear algebra.

4.4 The Column Correspondence Property (CCP)

Recall that we used Proposition 4.3.5 in Examples 4.3.7, 4.3.8, 4.3.9 and 4.3.10 to ascertain the linear independence of certain sets of vectors. If you pay particular attention to those examples you may have picked up on a pattern. The columns of the $rref[v_1|v_2|\cdots|v_k]$ depend on each other in the same way that the vectors $v_1, v_2, \ldots v_k$ depend on each other. These provide examples of the so-called "column correspondence property". In a nutshell, the property says you can read the linear dependencies right off the $rref[v_1|v_2|\cdots|v_k]$.

Proposition 4.4.1. Column Correspondence Property (CCP)

Let $A = [col_1(A)| \cdots |col_n(A)] \in \mathbb{R}^{m \times n}$ and $R = rref[A] = [col_1(R)| \cdots |col_n(R)]$. There exist constants $c_1, c_2, \ldots c_k$ such that $c_1col_1(A) + c_2col_2(A) + \cdots + c_kcol_k(A) = 0$ if and only if $c_1col_1(R) + c_2col_2(R) + \cdots + c_kcol_k(R) = 0$. If $col_j(rref[A])$ is a linear combination of other columns of rref[A] then $col_j(A)$ is likewise the same linear combination of columns of A.

We prepare for the proof of the Proposition by establishing a sick³ Lemma.

Lemma 4.4.2.

Let $A \in \mathbb{R}^{m \times n}$ then there exists an invertible matrix E such that $col_i(rref(A)) = Ecol_i(A)$ for all i = 1, 2, ..., n.

Proof of Lemma: Recall that there exist elementary matrices E_1, E_2, \ldots, E_r such that $A = E_1E_2\cdots E_r rref(A) = E^{-1}rref(A)$ where I have defined $E^{-1} = E_1E_2\cdots E_k$ for convenience. Recall the concatenation proposition: $X[b_1|b_2|\cdots|b_k] = [Xb_1|Xb_2|\cdots|Xb_k]$. We can unravel the Gaussian elimination in the same way,

$$EA = E[col_1(A)|col_2(A)|\cdots|col_n(A)]$$
$$= [Ecol_1(A)|Ecol_2(A)|\cdots|Ecol_n(A)]$$

Observe EA = rref(A) thus $col_i(rref(A)) = Ecol_i(A)$ for all i. \Box

Proof of Proposition: Suppose that there exist constants c_1, c_2, \ldots, c_k such that $c_1 col_1(A) + c_2 col_2(A) + \cdots + c_k col_k(A) = 0$. By the Lemma we know there exists E such that $col_j(rref(A)) = Ecol_j(A)$. Multiply linear combination by E to find:

$$c_1Ecol_1(A) + c_2Ecol_2(A) + \dots + c_kEcol_k(A) = 0$$

which yields

$$c_1 col_1(rref(A)) + c_2 col_2(rref(A)) + \dots + c_k col_k(rref(A)) = 0.$$

Likewise, if we are given a linear combination of columns of rref(A) we can multiply by E^{-1} to recover the same linear combination of columns of A. \Box

³Sorry, is this so 2009 now ?

Example 4.4.3. I will likely use the abbreviation "CCP" for column correspondence property. We could have deduced all the linear dependencies via the CCP in Examples 4.3.7,4.3.9 and 4.3.10. We found in 4.3.7 that

$$rref[v_1|v_2|v_3] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Obviously $col_3(R) = col_1(R) + col_2(R)$ hence by CCP $v_3 = v_1 + v_2$. We found in 4.3.9 that

$$rref[v_1|v_2|v_3] = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

By inspection, $col_3(R) = col_2(R) - col_1(R)$ hence by CCP $v_3 = v_2 - v_1$. We found in 4.3.10 that

$$rref[v_1|v_2|v_3|v_4] = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

By inspection, $col_3(R) = col_1(R) + col_2(R)$ hence by CCP $v_3 = v_1 + v_2$. Likewise by inspection, $col_4(R) = col_2(R) - col_1(R)$ hence by CCP $v_4 = v_2 - v_1$.

You should notice that the CCP saves us the trouble of expressing how the constants c_i are related. If we are only interested in how the vectors are related the CCP gets straight to the point quicker. We should pause and notice another pattern here while were thinking about these things.

Proposition 4.4.4.

The non-pivot columns of a matrix can be written as linear combinations of the pivot columns and the pivot columns of the matrix are linearly independent.

Proof: Let A be a matrix. Notice the Proposition is clearly true for rref(A). Hence, using Lemma 4.4.2 we find the same is true for the matrix A. \Box

Proposition 4.4.5.

The rows of a matrix A can be written as linear combinations of the transposes of pivot columns of A^T , and the rows which are transposes of the pivot columns of A^T are linearly independent.

Proof: Let A be a matrix and A^T its transpose. Apply Proposition 4.4.1 to A^T to find pivot columns which we denote by $col_{i_j}(A^T)$ for j = 1, 2, ..., k. These columns are linearly independent and they span $Col(A^T)$. Suppose,

$$c_1 row_{i_1}(A) + c_2 row_{i_2}(A) + \dots + c_k row_{i_k}(A) = 0.$$

Take the transpose of the equation above, use Proposition 3.8.3 to simplify:

$$c_1(row_{i_1}(A))^T + c_2(row_{i_2}(A))^T + \dots + c_k(row_{i_k}(A))^T = 0.$$

Recall $(row_j(A))^T = col_j(A^T)$ thus,

$$c_1 col_{i_1}(A^T) + c_2 col_{i_2}(A^T) + \dots + c_k col_{i_k}(A^T) = 0.$$

hence $c_1 = c_2 = \cdots = c_k = 0$ as the pivot columns of A^T are linearly independendent. This shows the corresponding rows of A are likewise linearly independent. The proof that these same rows span Row(A) is similar. \Box

4.5 theoretical summary

Let's pause to think about what we've learned about spans in this section. First of all the very definition of matrix multiplication defined Av to be a linear combination of the columns of A so clearly Av = b has a solution iff b is a linear combination of the columns in A.

We have seen for a particular matrix A and a given vector b it may or may not be the case that Av = b has a solution. It turns out that certain special matrices will have a solution for each choice of b. The theorem below is taken from Lay's text on page 43. The abbreviation **TFAE** means "The Following Are Equivalent".

Theorem 4.5.1.

Suppose $A = [A_{ij}] \in \mathbb{R}^{k \times n}$ then TFAE,

- 1. Av = b has a solution for each $b \in \mathbb{R}^k$
- 2. each $b \in \mathbb{R}^k$ is a linear combination of the columns of A
- 3. columns of A span \mathbb{R}^k
- 4. A has a pivot position in each row.

Proof: the equivalence of (1.) and (2.) is immediate from the definition of matrix multiplication of a matrix and a vector. Item (3.) says that the set of all linear combinations of the columns of A is equal to \mathbb{R}^k , thus (2.) \Leftrightarrow (3.). Finally, item (4.) is not just notation.

Suppose (4.) is true. Recall that rref[A] and rref[A|b] have matching columns up to the rightmost column of rref[A|b] by the Theorem 2.5.2. It follows that rref[A|b] is a consistent system since we cannot have a row where the first nonzero entry occurs in the last column. But, this result is independent of b hence we have a solution of Av = b for each possible $b \in \mathbb{R}^k$. Hence (4.) \Rightarrow (.1).

Conversely suppose (1.) is true; suppose Av = b has a solution for each $b \in \mathbb{R}^k$. If rref[A] has a row of zeros then we could choose $b \neq 0$ with a nonzero component in that row and the equation Av = b would be inconsistent. But that contradicts (1.) hence it must be the case that rref[A] has no row of zeros hence every row must be a pivot row. We have $(1.) \Rightarrow (4.)$.

In conclusion, $(1.) \Leftrightarrow (2.) \Leftrightarrow (3.)$ and $(1.) \Leftrightarrow (4.)$ hence $(4.) \Leftrightarrow (1.) \Leftrightarrow (2.) \Leftrightarrow (3.) \square$.

In truth this theorem really only scratches the surface. We can say more if the matrix A is square. But, I leave the fun for a later chapter. This much fun for now should suffice.

Chapter 5

linear transformations of column vectors

We study linear transformations of a special type in this chapter. Keeping with the general theme of Part I. we consider only transformations from \mathbb{R}^n to \mathbb{R}^m . The first section we study the definition. Proof that linear transformations take linear objects to linear objects is given. A number of explicit pictures show how the transformations move points. Then the following section details basic properties. Surjectivity and injectivity of a transformation are studied. We see how the standard matrix of the linear transformation reveals the structure of the map. Finally, matrix multiplication is shown to follow naturally from composition of functions. We conclude with an application to more general maps.

5.1 a gallery of linear transformations

A function from $U \subset \mathbb{R}^n$ to $V \subseteq \mathbb{R}^k$ is called a **mapping** or **transformation**. We could just use the term "function", but these other terms help draw attention to the vectorial nature of the domain and codomain.

Definition 5.1.1.

Let $V = \mathbb{R}^n, W = \mathbb{R}^k$. If a mapping $L: V \to W$ satisfies

1. L(x+y) = L(x) + L(y) for all $x, y \in V$; this is called **additivity**.

2. L(cx) = cL(x) for all $x \in V$ and $c \in \mathbb{R}$; this is called **homogeneity**.

then we say L is a linear transformation. If n = m then we may say that L is a linear transformation on \mathbb{R}^n .

Example 5.1.2. Let $L : \mathbb{R} \to \mathbb{R}$ be defined by L(x) = mx + b where $m, b \in \mathbb{R}$ and $b \neq 0$. This is often called a **linear function** in basic courses. However, this is unfortunate terminology as:

L(x + y) = m(x + y) + b = mx + b + my + b - b = L(x) + L(y) - b.

Thus L is not additive hence it is not a linear transformation. It is certainly true that y = L(x) gives a line with slope m and y-intercept b. An accurate term for L is that it is an **affine** function.

Example 5.1.3. Let $f(x,y) = x^2 + y^2$ define a function from \mathbb{R}^2 to \mathbb{R} . Observe,

$$f(c(x,y)) = f(cx,cy) = (cx)^{2} + (cy)^{2} = c^{2}(x^{2} + y^{2}) = c^{2}f(x,y).$$

Clearly f is not homogeneous hence f is not linear.

Example 5.1.4. Suppose $f(t, s) = (\sqrt{t}, s^2 + t)$ note that f(1, 1) = (1, 2) and f(4, 4) = (2, 20). Note that (4, 4) = 4(1, 1) thus we should see f(4, 4) = f(4(1, 1)) = 4f(1, 1) but that fails to be true so f is not a linear transformation.

Now that we have a few examples of how not to be a linear transformation, let's take a look at some positive examples.

Example 5.1.5. Let L(x, y) = (x, 2y). This is a mapping from \mathbb{R}^2 to \mathbb{R}^2 . Notice

$$L((x,y) + (z,w)) = L(x + z, y + w) = (x + z, 2(y + w)) = (x, 2y) + (z, 2w) = L(x, y) + L(z, w)$$

and

$$L(c(x,y)) = L(cx, cy) = (cx, 2(cy)) = c(x, 2y) = cL(x, y)$$

for all $(x, y), (z, w) \in \mathbb{R}^2$ and $c \in \mathbb{R}$. Therefore, L is a linear transformation on \mathbb{R}^2 . Let's examine how this function maps the unit square in the domain: suppose $(x, y) \in [0, 1] \times [0, 1]$. This means $0 \le x \le 1$ and $0 \le y \le 1$. Label the Cartesian coordinates of the range by u, v so L(x, y) =(x, 2y) = (u, v). We have x = u thus $0 \le u \le 1$. Also, v = 2y hence $y = \frac{v}{2}$ hence $0 \le y \le 1$ implies $0 \le \frac{v}{2} \le 1$ or $0 \le v \le 2$.

To summarize: $L([0,1] \times [0,1]) = [0,1] \times [0,2]$. This mapping has stretched out the horizontal direction.



The method of analysis we used in the preceding example was a little clumsy, but for general mappings that is more or less the method of attack. You pick some shapes or curves in the domain and see what happens under the mapping. For linear mappings there is an easier way. It turns out that if we map some shape with straight sides then the image will likewise be a shape with flat sides (or faces in higher dimensions). Therefore, to find the image we need only map the corners of the shape then connect the dots. However, I should qualify that it may not be the case the type of shape is preserved. We could have a rectangle in the domain get squished into a line or point in the domain. We would like to understand when such squishing will happen and also when a given mapping will actually cover the whole codomain. For linear mappings there are very satisfying answers to these questions in terms of the theory we have already discussed in previous chapters.

Proposition 5.1.6.

If $A \in \mathbb{R}^{m \times n}$ and $L : \mathbb{R}^n \to \mathbb{R}^m$ is defined by L(x) = Ax for each $x \in \mathbb{R}^n$ then L is a linear transformation.

Proof: Let $A \in \mathbb{R}^{m \times n}$ and define $L : \mathbb{R}^n \to \mathbb{R}^m$ by L(x) = Ax for each $x \in \mathbb{R}^n$. Let $x, y \in \mathbb{R}^n$ and $c \in \mathbb{R}$,

$$L(x + y) = A(x + y) = Ax + Ay = L(x) + L(y)$$

and

$$L(cx) = A(cx) = cAx = cL(x)$$

thus L is a linear transformation. \Box

Obviously this gives us a nice way to construct examples. The following proposition is really at the heart of all the geometry in this section.

Proposition 5.1.7.

Let $\mathcal{L} = \{p + tv \mid t \in [0, 1], p, v \in \mathbb{R}^n \text{ with } v \neq 0\}$ define a line segment from p to p + v in \mathbb{R}^n . If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation then $T(\mathcal{L})$ is a either a line-segment from T(p) to T(p + v) or a point.

Proof: suppose T and \mathcal{L} are as in the proposition. Let $y \in T(\mathcal{L})$ then by definition there exists $x \in \mathcal{L}$ such that T(x) = y. But this implies there exists $t \in [0, 1]$ such that x = p + tv so T(p + tv) = y. Notice that

$$y = T(p + tv) = T(p) + T(tv) = T(p) + tT(v).$$

which implies $y \in \{T(p) + sT(v) \mid s \in [0,1]\} = \mathcal{L}_2$. Therefore, $T(\mathcal{L}) \subseteq \mathcal{L}_2$. Conversely, suppose $z \in \mathcal{L}_2$ then z = T(p) + sT(v) for some $s \in [0,1]$ but this yields by linearity of T that z = T(p+sv) hence $z \in T(\mathcal{L})$. Since we have that $T(\mathcal{L}) \subseteq \mathcal{L}_2$ and $\mathcal{L}_2 \subseteq T(\mathcal{L})$ it follows that $T(\mathcal{L}) = \mathcal{L}_2$. Note that \mathcal{L}_2 is a line-segment provided that $T(v) \neq 0$, however if T(v) = 0 then $\mathcal{L}_2 = \{T(p)\}$ and the proposition follows. \Box

My choice of mapping the unit square has no particular significance in the examples below. I merely wanted to keep it simple and draw your eye to the distinction between the examples. In each example we'll map the four corners of the square to see where the transformation takes the unit-square. Those corners are simply (0,0), (1,0), (1,1), (0,1) as we traverse the square in a counter-clockwise direction.

Example 5.1.8. Let $A = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$ for some k > 0. Define L(v) = Av for all $v \in \mathbb{R}^2$. In particular this means,

$$L(x,y) = A(x,y) = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} kx \\ ky \end{bmatrix}.$$

We find L(0,0) = (0,0), L(1,0) = (k,0), L(1,1) = (k,k), L(0,1) = (0,k). This mapping is called a dilation.



Example 5.1.9. Let $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Define L(v) = Av for all $v \in \mathbb{R}^2$. In particular this means,

$$L(x,y) = A(x,y) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix}.$$

We find L(0,0) = (0,0), L(1,0) = (-1,0), L(1,1) = (-1,-1), L(0,1) = (0,-1). This mapping is called an inversion.



Example 5.1.10. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Define L(v) = Av for all $v \in \mathbb{R}^2$. In particular this means,

$$L(x,y) = A(x,y) = \begin{bmatrix} 1 & 2\\ 3 & 4 \end{bmatrix} \begin{bmatrix} x\\ y \end{bmatrix} = \begin{bmatrix} x+2y\\ 3x+4y \end{bmatrix}$$

We find L(0,0) = (0,0), L(1,0) = (1,3), L(1,1) = (3,7), L(0,1) = (2,4). This mapping shall remain nameless, it is doubtless a combination of the other named mappings.



Example 5.1.11. Let $A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. Define L(v) = Av for all $v \in \mathbb{R}^2$. In particular this means,

$$L(x,y) = A(x,y) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} x-y \\ x+y \end{bmatrix}.$$

We find L(0,0) = (0,0), $L(1,0) = \frac{1}{\sqrt{2}}(1,1)$, $L(1,1) = \frac{1}{\sqrt{2}}(0,2)$, $L(0,1) = \frac{1}{\sqrt{2}}(-1,1)$. This mapping is a rotation by $\pi/4$ radians.



Example 5.1.12. Let $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. Define L(v) = Av for all $v \in \mathbb{R}^2$. In particular this means,

$$L(x,y) = A(x,y) = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x-y \\ x+y \end{bmatrix}$$

We find L(0,0) = (0,0), L(1,0) = (1,1), L(1,1) = (0,2), L(0,1) = (-1,1). This mapping is a rotation followed by a dilation by $k = \sqrt{2}$.



We will come back to discuss rotations a few more times this semester, you'll see they give us interesting and difficult questions later this semester. Also, if you so choose there are a few bonus applied problems on computer graphics which are built from an understanding of the mathematics in the next example.

Example 5.1.13. Let $A = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$. Define L(v) = Av for all $v \in \mathbb{R}^2$. In particular this means,

$$L(x,y) = A(x,y) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos(\theta) - y\sin(\theta) \\ x\sin(\theta) + y\cos(\theta) \end{bmatrix}.$$

We find L(0,0) = (0,0), $L(1,0) = (\cos(\theta), \sin(\theta))$, $L(1,1) = (\cos(\theta) - \sin(\theta), \cos(\theta) + \sin(\theta)) L(0,1) = (\sin(\theta), \cos(\theta))$. This mapping is a rotation by θ in the counter-clockwise direction. Of course you could have derived the matrix A from the picture below.



Example 5.1.14. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Define L(v) = Av for all $v \in \mathbb{R}^2$. In particular this means,

$$L(x,y) = A(x,y) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

We find L(0,0) = (0,0), L(1,0) = (1,0), L(1,1) = (1,1), L(0,1) = (0,1). This mapping is a rotation by zero radians, or you could say it is a dilation by a factor of 1, ... usually we call this the identity mapping because the image is identical to the preimage.



Example 5.1.15. Let $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Define $P_1(v) = A_1 v$ for all $v \in \mathbb{R}^2$. In particular this means,

$$P_1(x,y) = A_1(x,y) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}.$$

We find $P_1(0,0) = (0,0)$, $P_1(1,0) = (1,0)$, $P_1(1,1) = (1,0)$, $P_1(0,1) = (0,0)$. This mapping is a **projection** onto the first coordinate.

Let
$$A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$
. Define $L(v) = A_2 v$ for all $v \in \mathbb{R}^2$. In particular this means,
$$P_2(x, y) = A_2(x, y) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}.$$

We find $P_2(0,0) = (0,0)$, $P_2(1,0) = (0,0)$, $P_2(1,1) = (0,1)$, $P_2(0,1) = (0,1)$. This mapping is **projection** onto the second coordinate.

We can picture both of these mappings at once:



Example 5.1.16. Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Define L(v) = Av for all $v \in \mathbb{R}^2$. In particular this means,

$$L(x,y) = A(x,y) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ x+y \end{bmatrix}$$

We find L(0,0) = (0,0), L(1,0) = (1,1), L(1,1) = (2,2), L(0,1) = (1,1). This mapping is not a projection, but it does collapse the square to a line-segment.



A projection has to have the property that if it is applied twice then you obtain the same image as if you applied it only once. If you apply the transformation to the image then you'll obtain a line-segment from (0,0) to (4,4). While it is true the transformation "projects" the plane to a line it is not technically a "projection".

Remark 5.1.17.

The examples here have focused on linear transformations from \mathbb{R}^2 to \mathbb{R}^2 . It turns out that higher dimensional mappings can largely be understood in terms of the geometric operations we've seen in this section. Example 5.1.18. Let $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$. Define L(v) = Av for all $v \in \mathbb{R}^2$. In particular this means, $L(x,y) = A(x,y) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ x \\ y \end{bmatrix}.$

We find L(0,0) = (0,0,0), L(1,0) = (0,1,0), L(1,1) = (0,1,1), L(0,1) = (0,0,1). This mapping moves the xy-plane to the yz-plane. In particular, the horizontal unit square gets mapped to vertical unit square; $L([0,1] \times [0,1]) = \{0\} \times [0,1] \times [0,1]$. This mapping certainly is not surjective because no point with $x \neq 0$ is covered in the range.



Example 5.1.19. Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$. Define L(v) = Av for all $v \in \mathbb{R}^3$. In particular this means,

$$L(x,y,z) = A(x,y,z) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ x+y+z \end{bmatrix}.$$

Let's study how L maps the unit cube. We have $2^3 = 8$ corners on the unit cube,

L(0,0,0) = (0,0), L(1,0,0) = (1,1), L(1,1,0) = (2,2), L(0,1,0) = (1,1)

 $L(0,0,1) = (0,1), \ L(1,0,1) = (1,2), \ L(1,1,1) = (2,3), \ L(0,1,1) = (1,2).$

This mapping squished the unit cube to a shape in the plane which contains the points (0,0), (0,1), (1,1), (1,2), (2,2), (2,3). Face by face analysis of the mapping reveals the image is a parallelogram. This mapping is certainly not injective since two different points get mapped to the same point. In particular, I have color-coded the mapping of top and base faces as they map to line segments. The vertical faces map to one of the two parallelograms that comprise the image.



I have used terms like "vertical" or "horizontal" in the standard manner we associate such terms with three dimensional geometry. Visualization and terminology for higher-dimensional examples is not as obvious. However, with a little imagination we can still draw pictures to capture important aspects of mappings.

Example 5.1.20. Let $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$. Define L(v) = Av for all $v \in \mathbb{R}^4$. In particular this means,

$$L(x, y, z, t) = A(x, y, z, t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{vmatrix} x \\ y \\ z \\ t \end{vmatrix} = \begin{bmatrix} x \\ x \end{bmatrix}.$$

Let's study how L maps the unit hypercube $[0,1]^4 \subset \mathbb{R}^4$. We have $2^4 = 16$ corners on the unit hypercube, note L(1, a, b, c) = (1, 1) whereas L(0, a, b, c) = (0, 0) for all $a, b, c \in [0, 1]$. Therefore, the unit hypercube is squished to a line-segment from (0,0) to (1,1). This mapping is neither surjective nor injective. In the picture below the vertical axis represents the y, z, t-directions.



Obviously we have not even begun to appreciate the wealth of possibilities that exist for linear mappings. Clearly different types of matrices will decribe different types of geometric transformations from \mathbb{R}^n to \mathbb{R}^m . On the other hand, square matrices describe mappings from \mathbb{R}^n to \mathbb{R}^n and these can be thought of as coordinate transformations. A square matrix may give us a way to define new coordinates on \mathbb{R}^n . We will return to the concept of linear transformations a number of times in this course. Hopefully you already appreciate that linear algebra is not just about solving equations. It always comes back to that, but there is more here to ponder.

5.2 properties of linear transformations

If you are pondering what I am pondering then you probably would like to know if all linear mappings from \mathbb{R}^n to \mathbb{R}^m can be reduced to matrix multiplication? We saw that if a map is defined as a matrix multiplication then it will be linear. A natural question to ask: is the converse true? Given a linear transformation from \mathbb{R}^n to \mathbb{R}^m can we write the transformation as multiplication by a matrix ?

Theorem 5.2.1. fundamental theorem of linear algebra.

 $L: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if and only if there exists $A \in \mathbb{R}^{m \times n}$ such that L(x) = Ax for all $x \in \mathbb{R}^n$.

Proof: (\Leftarrow) Assume there exists $A \in \mathbb{R}^{m \times n}$ such that L(x) = Ax for all $x \in \mathbb{R}^n$. As we argued before,

$$L(x + cy) = A(x + cy) = Ax + cAy = L(x) + cL(y)$$

for all $x, y \in \mathbb{R}^n$ and $c \in \mathbb{R}$ hence L is a linear transformation.

 (\Rightarrow) Assume $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Let e_i denote the standard basis in \mathbb{R}^n and let f_j denote the standard basis in \mathbb{R}^m . If $x \in \mathbb{R}^n$ then there exist constants x_i such that $x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$ and

$$L(x) = L(x_1e_1 + x_2e_2 + \dots + x_ne_n)$$

= $x_1L(e_1) + x_2L(e_2) + \dots + x_nL(e_n)$

where we made use of Proposition 7.2.1. Notice $L(e_i) \in \mathbb{R}^m$ thus there exist constants, say A_{ij} , such that

$$L(e_i) = A_{1i}f_1 + A_{2i}f_2 + \dots + A_{mi}f_m$$

for each $i = 1, 2, \ldots, n$. Let's put it all together,

$$L(x) = \sum_{i=1}^{n} x_i L(e_i)$$
$$= \sum_{i=1}^{n} x_i \sum_{j=1}^{m} A_{ji} f_j$$
$$= \sum_{i=1}^{n} \sum_{j=1}^{m} A_{ji} x_i f_j$$
$$= Ax.$$

Notice that $A_{ji} = L(e_i)_j$ for $1 \le i \le n$ and $1 \le j \le m$ hence $A \in \mathbb{R}^{m \times n}$ by its construction. \Box

The fundamental theorem of linear algebra allows us to make the following definition.

Definition 5.2.2.

Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, the matrix $A \in \mathbb{R}^{m \times n}$ such that L(x) = Axfor all $x \in \mathbb{R}^n$ is called the **standard matrix** of L. We denote this by [L] = A or more compactly, $[L_A] = A$, we say that L_A is the linear transformation induced by A. Moreover, the components of the matrix A are found from $A_{ji} = (L(e_i))_j$.

Example 5.2.3. Given that $L([x, y, z]^T) = [x + 2y, 3y + 4z, 5x + 6z]^T$ for $[x, y, z]^T \in \mathbb{R}^3$ find the the standard matrix of L. We wish to find a 3×3 matrix such that L(v) = Av for all $v = [x, y, z]^T \in \mathbb{R}^3$. Write L(v) then collect terms with each coordinate in the domain,

$$L\left(\left[\begin{array}{c}x\\y\\z\end{array}\right]\right) = \left[\begin{array}{c}x+2y\\3y+4z\\5x+6z\end{array}\right] = x\left[\begin{array}{c}1\\0\\5\end{array}\right] + y\left[\begin{array}{c}2\\3\\0\end{array}\right] + z\left[\begin{array}{c}0\\4\\6\end{array}\right]$$

It's not hard to see that,

$$L\left(\left[\begin{array}{c}x\\y\\z\end{array}\right]\right) = \left[\begin{array}{ccc}1&2&0\\0&3&4\\5&0&6\end{array}\right]\left[\begin{array}{c}x\\y\\z\end{array}\right] \quad \Rightarrow \quad A = [L] = \left[\begin{array}{cccc}1&2&0\\0&3&4\\5&0&6\end{array}\right]$$

Notice that the columns in A are just as you'd expect from the proof of theorem 5.2.1. $[L] = [L(e_1)|L(e_2)|L(e_3)]$. In future examples I will exploit this observation to save writing.

Example 5.2.4. Suppose that L((t, x, y, z)) = (t + x + y + z, z - x, 0, 3t - z), find [L].

$$\begin{split} L(e_1) &= L((1,0,0,0)) = (1,0,0,3) \\ L(e_2) &= L((0,1,0,0)) = (1,-1,0,0) \\ L(e_3) &= L((0,0,1,0)) = (1,0,0,0) \\ L(e_4) &= L((0,0,0,1)) = (1,1,0,-1) \end{split} \Rightarrow \ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & -1 \end{bmatrix}$$

I invite the reader to check my answer here and see that L(v) = [L]v for all $v \in \mathbb{R}^4$ as claimed.

Proposition 5.2.5.

Let
$$L : \mathbb{R}^n \to \mathbb{R}^m$$
 be a linear transformation,
1. $L(0) = 0$
2. $L(c_1v_1 + c_2v_2 + \cdots + c_nv_n) = c_1L(v_1) + c_2L(v_2) + \cdots + c_nL(v_n)$ for all $v_i \in \mathbb{R}^n$ and $c_i \in \mathbb{R}$.

Proof: to prove of (1.) let $x \in \mathbb{R}^n$ and notice that x - x = 0 thus

$$L(0) = L(x - x) = L(x) + L(-1x) = L(x) - L(x) = 0.$$

To prove (2.) we use induction on n. Notice the proposition is true for n=1,2 by definition of linear transformation. Assume inductively $L(c_1v_1+c_2v_2+\cdots c_nv_n)=c_1L(v_1)+c_2L(v_2)+\cdots+c_nL(v_n)$ for all $v_i \in \mathbb{R}^n$ and $c_i \in \mathbb{R}$ where $i = 1, 2, \ldots, n$. Let $v_1, v_2, \ldots, v_n, v_{n+1} \in \mathbb{R}^n$ and $c_1, c_2, \ldots, c_n, c_{n+1} \in \mathbb{R}$ and consider, $L(c_1v_1+c_2v_2+\cdots c_nv_n+c_{n+1}v_{n+1}) =$

$$= L(c_1v_1 + c_2v_2 + \dots + c_nv_n) + c_{n+1}L(v_{n+1})$$
 by linearity of L
= $c_1L(v_1) + c_2L(v_2) + \dots + c_nL(v_n) + c_{n+1}L(v_{n+1})$ by the induction hypothesis.

Hence the proposition is true for n+1 and we conclude by the principle of mathematical induction that (2.) is true for all $n \in \mathbb{N}$. \Box

Example 5.2.6. Suppose $L : \mathbb{R} \to \mathbb{R}$ is defined by L(x) = mx + b for some constants $m, b \in \mathbb{R}$. Is this a linear transformation on \mathbb{R} ? Observe:

$$L(0) = m(0) + b = b$$

thus L is not a linear transformation if $b \neq 0$. On the other hand, if b = 0 then L is a linear transformation. You might contrast this example with Example 5.1.2.

Remark 5.2.7.

A mapping on \mathbb{R}^n which has the form T(x) = x + b is called a **translation**. If we have a mapping of the form F(x) = Ax + b for some $A \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}$ then we say F is an **affine transformation** on \mathbb{R}^n . Technically, in general, the line y = mx + b is the graph of an affine function on \mathbb{R} . I invite the reader to prove that affine transformations also map line-segments to line-segments (or points).

Very well, let's return to the concepts of injective and surjectivity of linear mappings. How do our theorems of LI and spanning inform us about the behaviour of linear transformations? The following pair of theorems summarize the situation nicely.

Theorem 5.2.8. linear map is injective iff only zero maps to zero.

 $L: \mathbb{R}^n \to \mathbb{R}^m$ is an injective linear transformation iff the only solution to the equation L(x) = 0 is x = 0.

Proof: this is a biconditional statement. I'll prove the converse direction to begin. (\Leftarrow) Suppose L(x) = 0 iff x = 0 to begin. Let $a, b \in \mathbb{R}^n$ and suppose L(a) = L(b). By linearity we have L(a - b) = L(a) - L(b) = 0 hence a - b = 0 therefore a = b and we find L is injective. (\Rightarrow) Suppose L is injective. Suppose L(x) = 0. Note L(0) = 0 by linearity of L but then by 1-1 property we have L(x) = L(0) implies x = 0 hence the unique solution of L(x) = 0 is the zero solution. \Box

The theorem above is very important to abstract algebra. It turns out this is also a useful criteria to determine if a homomorphism is a 1-1 mapping. Linear algebra is a prerequisite of abstract because linear algebra provides a robust example of what is abstracted in abstract algebra. The following theorem is special to our context this semester.

Theorem 5.2.9. linear map is injective iff only zero maps to zero.

- $L: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation with standard matrix [L] then
 - 1. L is 1-1 iff the columns of [L] are linearly independent,
 - 2. L is onto \mathbb{R}^m iff the columns of [L] span \mathbb{R}^m .

Proof: To prove (1.) use Theorem 7.2.3:

$$L \text{ is } 1\text{-}1 \Leftrightarrow \left\{ L(x) = 0 \Leftrightarrow x = 0 \right\} \Leftrightarrow \left\{ [L]x = 0 \Leftrightarrow x = 0. \right\}$$

and the last equation simply states that if a linear combination of columns of L is zero then the coefficients of that linear equation are zero so (1.) follows.

To prove (2.) recall that Theorem 4.5.1 stated that if $A \in \mathbb{R}^{m \times n}$, $v \in \mathbb{R}^n$ then Av = b is consistent for all $b \in \mathbb{R}^m$ iff the columns of A span \mathbb{R}^m . To say L is onto \mathbb{R}^m means that for each $b \in \mathbb{R}^m$ there exists $v \in \mathbb{R}^n$ such that L(v) = b. But, this is equivalent to saying that [L]v = b is consistent for each $b \in \mathbb{R}^m$ so (2.) follows. \Box

- **Example 5.2.10.** 1. You can verify that the linear mappings in Examples 5.1.8, 5.1.9, 5.1.10, 5.1.11, 5.1.12, 5.1.13 and 5.1.14 wer both 1-1 and onto. You can see the columns of the transformation matrices were both LI and spanned \mathbb{R}^2 in each of these examples.
 - In constrast, Examples 5.1.15 and 5.1.16 were neither 1-1 nor onto. Moreover, the columns
 of transformation's matrix were linearly dependent in each of these cases and they did not
 span R². Instead the span of the columns covered only a particular line in the range.
 - In Example 5.1.18 the mapping is injective and the columns of A were indeed linearly indpendent. However, the columns do not span ℝ³ and as expected the mapping is not onto ℝ³.
 - In Example 5.1.19 the mapping is not 1-1 and the columns are obviously linearly dependent. On the other hand, the columns of A do span ℝ² and the mapping is onto.
 - 5. In Example 5.1.20 the mapping is neither 1-1 nor onto and the columns of the matrix are neither linearly independent nor do they span \mathbb{R}^2 .

5.3 new linear transformations from old

We can add, subtract and scalar multiply linear transformations. Let us define these:

Definition 5.3.1.

Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^n \to \mathbb{R}^m$ are linear transformations then we define T + S, T - S and cT for any $c \in \mathbb{R}$ by the rules $(T + S)(x) = T(x) + S(x). \quad (T - S)(x) = T(x) - S(x), \quad (cT)(x) = cT(x)$ for all $x \in \mathbb{R}^n$.

The following does say something new. Notice I'm talking about adding the transformations themselves not the points in the domain or range.

Proposition 5.3.2.

The sum, difference or scalar multiple of a linear transformations from \mathbb{R}^n to \mathbb{R}^m are once more a linear transformation from \mathbb{R}^n to \mathbb{R}^m .

Proof: I'll be greedy and prove all three at once: let $x, y \in \mathbb{R}^n$ and $c \in \mathbb{R}$,

$$\begin{aligned} (T+cS)(x+by) &= T(x+by) + (cS)(x+by) & \text{defn. of sum of transformations} \\ &= T(x+by) + cS(x+by) & \text{defn. of scalar mult. of transformations} \\ &= T(x) + bT(y) + c[S(x) + bS(y)] & \text{linearity of } S \text{ and } T \\ &= T(x) + cS(x) + b[T(y) + cS(y)] & \text{vector algebra props.} \\ &= (T+cS)(x) + b(T+cS)(y) & \text{again, defn. of sum and scal. mult. of trans.} \end{aligned}$$

Let c = 1 and b = 1 to see T + S is additive. Let c = 1 and x = 0 to see T + S is homogeneous. Let c = -1 and b = 1 to see T - S is additive. Let c = -1 and x = 0 to see T - S is homogeneous. Finally, let T = 0 to see cS is additive (b = 1) and homogeneous (x = 0). \Box

Proposition 5.3.3.

Suppose $T: \mathbb{R}^n \to \mathbb{R}^m$ and $S: \mathbb{R}^n \to \mathbb{R}^m$ are linear transformations then

(1.)
$$[T+S] = [T] + [S],$$
 (2.) $[T-S] = [T] - [S],$ (3.) $[cS] = c[S].$

In words, the standard matrix of the sum, difference or scalar multiple of linear transformations likewise the sum, difference or scalar multiple of the standard matrices of the respective linear transformations.

Proof: Note $(T + cS)(e_j) = T(e_j) + cS(e_j)$ hence $((T + cS)(e_j))_i = (T(e_j))_i + c(S(e_j))_i$ for all i, j hence [T + cS] = [T] + c[S]. Set c = 1 to obtain (1.). Set c = -1 to obtain (2.). Finally, set T = 0 to obtain (3.). \Box

Example 5.3.4. Suppose T(x,y) = (x+y, x-y) and S(x,y) = (2x, 3y). It's easy to see that

$$[T] = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} and [S] = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \Rightarrow [T+S] = [T] + [S] = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$$

Therefore, $(T+S)(x,y) = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x+y \\ x+2y \end{bmatrix} = (3x+y,x+2y)$. Naturally this is the same formula that we would obtain through direct addition of the formulas of T and S.

5.3.1 composition and matrix multiplication

In this subsection we see that matrix multiplication is naturally connected to the problem of composition of linear maps. The definition that follows here is just the usual definition of composite.

Definition 5.3.5.

Suppose
$$T : \mathbb{R}^n \to \mathbb{R}^m$$
 and $S : \mathbb{R}^m \to \mathbb{R}^p$ are linear transformations then we define $S \circ T : \mathbb{R}^n \to \mathbb{R}^p$ by $(S \circ T)(x) = S(T(x))$ for all $x \in \mathbb{R}^n$.

The composite of linear maps is once more a linear map.

Proposition 5.3.6.

Suppose $T : \mathbb{R}^n \to \mathbb{R}^p$ and $S : \mathbb{R}^p \to \mathbb{R}^m$ are linear transformations then we define $S \circ T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation.

Proof: Let $x, y \in \mathbb{R}^n$ and $c \in \mathbb{R}$,

$$\begin{array}{ll} (S \circ T)(x+cy) &= S(T(x+cy)) & \text{defn. of composite} \\ &= S(T(x)+cT(y)) & T \text{ is linear trans.} \\ &= S(T(x))+cS(T(y)) & S \text{ is linear trans.} \\ &= (S \circ T)(x)+c(S \circ T)(y) & \text{defn. of composite} \end{array}$$

thus $S \circ T$ is a linear transformation. \Box
Proposition 5.3.7.

 $S: \mathbb{R}^p \to \mathbb{R}^n$ and $T: \mathbb{R}^n \to \mathbb{R}^p$ are linear transformations then $S \circ T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation with standard matrix [S][T]; that is, $[S \circ T] = [S][T]$.

Proof: Let us denote $\mathbb{R}^n = span\{e_i \mid i = 1, ..., n\}$ whereas $\mathbb{R}^p = span\{f_i \mid i = 1, ..., p\}$ and $\mathbb{R}^m = span\{g_i \mid i = 1, ..., m\}$. To find the matrix of the composite we need only calculate its action on the standard basis: by definition, $[S \circ T]_{ij} = ((S \circ T)(e_j))_i$, observe

$$(S \circ T)(e_j) = S(T(e_j)) \qquad : \text{ def. of composite} \\ = S([T]e_j) \qquad : \text{ def. of } [T] \\ = S(\sum_k [T]_{kj} f_k) \qquad : \text{ standard basis } \{f_i\} \text{ spans } \mathbb{R}^p \\ = \sum_k [T]_{kj} S(f_k) \qquad : \text{ homogeneity of } S \\ = \sum_k [T]_{kj} [S]f_k \qquad : \text{ def. of } [S] \\ = \sum_k [T]_{kj} \sum_i [S]_{ik} g_i \qquad : \text{ standard basis } \{g_i\} \text{ spans } \mathbb{R}^m \\ = \sum_k \sum_i [S]_{ik} [T]_{kj} g_i \qquad : \text{ by } (2.) \text{ of Prop. } 1.3.3 \\ = \sum_i [[S]_{ik} [T]_{kj}]g_i \qquad : \text{ by } (1.) \text{ of Prop. } 1.3.3 \\ = \sum_i ([S][T])_{ij} g_i \qquad : \text{ def. of matrix multiplication} \end{cases}$$

The *i*-th component of $(S \circ T)(e_j)$ is easily seen from the above expression. In particular, we find $[S \circ T]_{ij} = \sum_k [S]_{jk} [T]_{ki}$ and the proof is complete. \Box

Think about this: **matrix multiplication was defined to make the above proposition true**. Perhaps you wondered, why don't we just multiply matrices some other way? Well, now you have an answer. If we defined matrix multiplication differently then the result we just proved would not be true. However, as the course progresses, you'll see why it is so important that this result be true. It lies at the heart of many connections between the world of linear transformations and the world of matrices. It says we can trade composition of linear transformations for multiplication of matrices.

5.4 applications

Geometry is conveniently described by parametrizations. The number of parameters needed to map out some object is the dimension of the object. For example, the rule $t \mapsto \vec{r}(t)$ describes a curve in \mathbb{R}^n . Of course we have the most experience in the cases $\vec{r} = \langle x, y \rangle$ or $\vec{r} = \langle x, y, z \rangle$, these give so-called *planar curves* or *space curves* respectively. Generally, a mapping from $\gamma : \mathbb{R} \to S$ where S is some space¹ is called a *path*. The point set $\gamma(S)$ can be identified as a sort of copy of \mathbb{R} which resides in S.

¹here S could be a set of matrices or functions or an abstract manifold... the concept of a path is very general

Next, we can consider mappings from \mathbb{R}^2 to some space S. In the case $S = \mathbb{R}^3$ we use $X(u, v) = \langle x(u, v), y(u, v), z(u, v) \rangle$ to parametrize a surface. For example,

$$X(\phi, \theta) = \langle \cos(\theta) \sin(\phi), \sin(\theta) \sin(\phi), \cos(\phi) \rangle$$

parametrizes a sphere if we insist that the angles $0 \le \theta \le 2\pi$ and $0 \le \phi \le \pi$. We call ϕ and θ coordinates on the sphere, however, these are not coordinates in the technical sense later defined in this course. These are so-called *curvelinear coordinates*. Generally a surface in some space is sort-of a copy of \mathbb{R}^2 (well, to be more precise it resembles some subset of \mathbb{R}^2).



Past the case of a surface we can talk about volumes which are parametrized by three parameters. A volume would have to be embedded into some space which had at least 3 dimensions. For the same reason we can only place a surface in a space with at least 2 dimensions. Perhaps you'd be interested to learn that in relativity theory one considers the world-volume that a particle traces out through spacetime, this is a hyper-volume, it's a 4-dimensional subset of 4-dimensional spacetime.

Let me be a little more technical, if the space we consider is to be a k-dimensional parametric subspace of S then that means there exists an **invertible** mapping $X : U \subseteq \mathbb{R}^k \to S \subseteq \mathbb{R}^n$. It turns out that for $S = \mathbb{R}^n$ where $n \geq k$ the condition that X be invertible means that the derivative $D_pX : T_pU \to T_{X(p)}S$ must be an invertible linear mapping at each point p in the parameter space U. This in turn means that the tangent-vectors to the coordinate curves must come together to form a linearly independent set. Linear independence is key.

Curvy surfaces and volumes and parametrizations that describe them analytically involve a fair amount of theory which I have only begun to sketch here. However, if we limit our discussion to **affine subspaces** of \mathbb{R}^n we can be explicit. Let me go ahead and write the general form for a line, surface, volume etc... in terms of linearly indpendent vectors $\vec{A}, \vec{B}, \vec{C}, \ldots$

$$\vec{r}(u) = \vec{r}_o + uA$$
$$X(u, v) = \vec{r}_o + u\vec{A} + v\vec{B}$$
$$X(u, v, w) = \vec{r}_o + u\vec{A} + v\vec{B} + w\vec{C}$$

I hope you you get the idea.



In each case the parameters give an invertible map only if the vectors are linearly independent. If there was some linear dependence then the dimension of the subspace would collapse. For example,

$$X(u,v) = <1, 1, 1>+u<1, 0, 1>+v<2, 0, 2>$$

appears to give a plane, but upon further inspection you'll notice

$$X(u,v) = <1+u+2v, 1, 1+u+2v > = <1, 1, 1>+(u+2v) <1, 0, 1>$$

which reveals this is just a line with direction-vector < 1, 0, 1 > and parameter u + 2v.

Part II

abstract linear algebra

Chapter 6

vector space

Up to this point the topics we have discussed loosely fit into the category of matrix theory. The concept of a matrix is milienia old. If I trust my source, and I think I do, the Chinese even had an analog of Gaussian elimination about 2000 years ago. The modern notation likely stems from the work of Cauchy in the 19-th century. Cauchy's prolific work colors much of the notation we still use. The concept of coordinate geometry as introduced by Descartes and Fermat around 1644 is what ultimately led to the concept of a vector space.¹. Grassmann, Hamilton, and many many others worked out volumous work detailing possible transformations on what we now call $\mathbb{R}^2, \mathbb{R}^3, \mathbb{R}^4$,. Argand(complex numbers) and Hamilton(quaternions) had more than what we would call a vector space. They had a linear structure plus some rule for multiplication of vectors. A vector space with a multiplication is called an *algebra* in the modern terminology.

Honestly, I think once the concept of the Cartesian plane was discovered the concept of a vector space almost certainly must follow. That said, it took a while for the definition I state in the next section to appear. Giuseppe Peano gave the modern definition for a vector space in 1888². In addition he put forth some of the ideas concerning linear transformations. Peano is also responsible for the modern notations for intersection and unions of sets³. He made great contributions to proof by induction and the construction of the natural numbers from basic set theory.

I should mention the work of Hilbert, Lebesque, Fourier, Banach and others were greatly influential in the formation of infinite dimensional vector spaces. Our focus is on the finite dimensional case.⁴

Let me summarize what a vector space is before we define it properly. In short, a vector space over a field \mathbb{F} is simply a set which allows you to add its elements and multiply by the numbers in \mathbb{F} . A field is a set with addition and multiplication defined such that every nonzero element has a multiplicative inverse. Typical examples, $\mathbb{F} = \mathbb{R}, \mathbb{C}, \mathbb{Z}_p$ where p is prime. I'll focus on \mathbb{R} , but many of the theorems and definitions we consider readily generalized to arbitrary \mathbb{F} .

Vector spaces are found throughout modern mathematics. Moreover, the theory we cover in this chapter is applicable to a myriad of problems with real world content. This is the beauty of linear algebra: it simultaneously illustrates the power of application and abstraction in mathematics.

¹ Bourbaki 1969, ch. "Algebre lineaire et algebre multilineaire", pp. 78-91.

²Peano, Giuseppe (1888), Calcolo Geometrico secondo l'Ausdehnungslehre di H. Grassmann preceduto dalle Operazioni della Logica Deduttiva, Turin

³see Pg 87 of A Transition to Advanced Mathematics: A Survey Course By William Johnston

⁴this history is flawed, one-sided and far too short. You should read a few more books if you're interested.

6.1 definition and examples

Axioms are not derived from a more basic logic. They are the starting point. Their validity is ultimately judged by their use. However, this definition is naturally motivated by the structure of vector addition and scalar multiplication in \mathbb{R}^n (see Proposition 1.5.9)

Definition 6.1.1.

A vector space V over \mathbb{R} is a set V together with a function $+: V \times V \to V$ called **vector** addition and another function $\cdot: \mathbb{R} \times V \to V$ called **scalar multiplication**. We require that the operations of vector addition and scalar multiplication satisfy the following 10 axioms: for all $x, y, z \in V$ and $a, b \in \mathbb{R}$,

- 1. (A1) x + y = y + x for all $x, y \in V$,
- 2. (A2) (x+y) + z = x + (y+z) for all $x, y, z \in V$,
- 3. (A3) there exists $0 \in V$ such that x + 0 = x for all $x \in V$,
- 4. (A4) for each $x \in V$ there exists $-x \in V$ such that x + (-x) = 0,
- 5. (A5) $1 \cdot x = x$ for all $x \in V$,
- 6. (A6) $(ab) \cdot x = a \cdot (b \cdot x)$ for all $x \in V$ and $a, b \in \mathbb{R}$,
- 7. (A7) $a \cdot (x + y) = a \cdot x + a \cdot y$ for all $x, y \in V$ and $a \in \mathbb{R}$,
- 8. (A8) $(a+b) \cdot x = a \cdot x + b \cdot x$ for all $x \in V$ and $a, b \in \mathbb{R}$,
- 9. (A9) If $x, y \in V$ then x + y is a single element in V, (we say V is closed with respect to addition)
- 10. (A10) If $x \in V$ and $c \in \mathbb{R}$ then $c \cdot x$ is a single element in V. (we say V is closed with respect to scalar multiplication)

We call 0 in axiom 3 the **zero vector** and the vector -x is called the **additive inverse** of x. We will sometimes omit the \cdot and instead denote scalar multiplication by juxtaposition; $a \cdot x = ax$.

Axioms (9.) and (10.) are admittably redundant given that those automatically follow from the statements that $+: V \times V \to V$ and $\cdot: \mathbb{R} \times V \to V$ are <u>functions</u>. I've listed them so that you are less likely to forget they must be checked.

The terminology "vector" does not necessarily indicate an explicit geometric interpretation in this general context. Sometimes I'll insert the word "abstract" to emphasize this distinction. We'll see that matrices, polynomials and functions in general can be thought of as abstract vectors.

Example 6.1.2. \mathbb{R} is a vector space if we identify addition of real numbers as the vector addition and multiplication of real numbers as the scalar multiplication.

The preceding example is very special because we can actually multiply the vectors. Usually we cannot multiply vectors.

Example 6.1.3. Proposition 1.5.9 shows \mathbb{R}^n forms a vector space with respect to the standard vector addition and scalar multiplication.

Example 6.1.4. The set of all $m \times n$ matrices is denoted $\mathbb{R}^{m \times n}$. It forms a vector space with respect to matrix addition and scalar multiplication as we defined previously. Notice that we cannot mix matrices of differing sizes since we have no natural way of adding them.

Example 6.1.5. The set of all linear transformations from \mathbb{R}^n to \mathbb{R}^m is denoted $L(\mathbb{R}^n, \mathbb{R}^m)$. Define addition and scalar multiplication of the transformations in the natural manner: if $S, T \in L(\mathbb{R}^n, \mathbb{R}^m)$ then for $c \in \mathbb{R}$ and each $v \in \mathbb{R}^n$

$$(S+T)(v) = S(v) + T(v),$$
 $(c \cdot T)(v) = cT(v).$

we can show $S + T \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $c \cdot T \in L(\mathbb{R}^n, \mathbb{R}^m)$ and the other axioms follow easily.

Example 6.1.6. Let $\mathcal{F}(\mathbb{R})$ denote the set of all functions with domain \mathbb{R} . Let $f, g \in \mathcal{F}(\mathbb{R})$ and suppose $c \in \mathbb{R}$, define addition of functions by

$$(f+g)(x) \equiv f(x) + g(x)$$

for all $x \in \mathbb{R}$. Likewise for $f \in \mathcal{F}(\mathbb{R})$ and $c \in \mathbb{R}$ define scalar multiplication of a function by a constant in the obvious way:

$$(cf)(x) = cf(x)$$

for all $x \in \mathbb{R}$. In short, we define addition and scalar multiplication by the natural "point-wise" rules. Notice we must take functions which share the same domain since otherwise we face difficulty in choosing the domain for the new function f+g, we can also consider functions sharing a common domain $I \subset \mathbb{R}$ and denote that by $\mathcal{F}(I)$. These are called function spaces.

Example 6.1.7. Let $P_2 = \{ax^2 + bx + c \mid a, b, c \in \mathbb{R}\}$, the set of all polynomials up to quadratic order. Define addition and scalar multiplication by the usual operations on polynomials. Notice that if $ax^2 + bx + c$, $dx^2 + ex + f \in P_2$ then

$$(ax^{2} + bx + c) + (dx^{2} + ex + f) = (a + d)x^{2} + (b + e)x + (c + f) \in P_{2}$$

thus $+: P_2 \times P_2 \rightarrow P_2$ (it is a binary operation on P_2). Similarly,

$$d(ax^2 + bx + c) = dax^2 + dbx + dc \in P_2$$

thus scalar multiplication maps $\mathbb{R} \times P_2 \to P_2$ as it ought. Verification of the other 8 axioms is straightfoward. We denote the set of polynomials of order n or less via $P_n = \{a_n x^n + \cdots + a_1 x + a_0 | a_i \in \mathbb{R}\}$. Naturally, P_n also forms a vector space. Finally, if we take the set of all polynomials P it forms a vector space. Notice,

$$P_2 \subset P_3 \subset P_4 \subset \cdots \subset P$$

Example 6.1.8. Let V, W be vector spaces over \mathbb{R} . The Cartesian product $V \times W$ has a natural vector space structure inherited from V and W: if $(v_1, w_1), (v_2, w_2) \in V \times W$ then we define

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$$
 & $c \cdot (v_1, w_1) = (c \cdot v_1, c \cdot w_1)$

where the vector and scalar operations on the L.H.S. of the above equalities are given from the vector space structure of V and W. All the axioms of a vector space for $V \times W$ are early verified from the corresponding axioms for V and W.

The theorem that follows is full of seemingly obvious facts. I show how each of these facts follow from the vector space axioms.

Theorem 6.1.9.

Let V be a vector space with zero vector 0 and let $c \in \mathbb{R}$, 1. $0 \cdot x = 0$ for all $x \in V$, 2. $c \cdot 0 = 0$ for all $c \in \mathbb{R}$, 3. $(-1) \cdot x = -x$ for all $x \in V$, 4. if cx = 0 then c = 0 or x = 0.

Lemma 6.1.10. Law of Cancellation:

Let a, x, y be vectors in a vector space V. If x + a = y + a then x = y.

Proof of Lemma: Suppose x + a = y + a. By A4 there exists -a such that a + (-a) = 0. Thus x + a = y + a implies (x + a) + (-a) = (y + a) + (-a). By A2 we find x + (a + (-a)) = y + (a + (-a)) which gives x + 0 = y + 0. Continuing we use A3 to obtain x + 0 = 0 and y + 0 = y and consequently x = y. ∇ .

We now seek to prove (1.). Consider:

0

$x + 0 = 0 \cdot x$	by A3
$= [0(1+1)] \cdot x$	arithmetic in $\mathbb R$
$= 0 \cdot \left((1+1) \cdot x \right)$	by A6
$= 0 \cdot (1 \cdot x + 1 \cdot x)$	by A8
$= 0 \cdot (1 \cdot x) + 0 \cdot (1 \cdot x)$	by A7
$= (0(1)) \cdot x + (0(1)) \cdot x$	by A6
$= 0 \cdot x + 0 \cdot x$	arithmetic in \mathbb{R}

Finally, apply the cancellation lemma to conclude $0 \cdot x = 0$. Note x was arbitrary thus (1.) has been shown true. ∇

We now prove (2.). Suppose $c \in \mathbb{R}$.

$$c \cdot 0 + 0 = c \cdot 0 \qquad \qquad \text{by A3}$$

$$= c \cdot (0+0) \qquad \qquad \text{by A3}$$

 $= c \cdot 0 + c \cdot 0 \qquad \qquad \text{by A7}$

Consequently, by the cancellation lemma we find $c \cdot 0 = 0$ for all $c \in \mathbb{R}$. ∇

The proof of (3.) is similar. Consider,

$$1 \cdot x + 0 = 1 \cdot x \qquad \text{by A3}$$

$$= (2 + (-1)) \cdot x \qquad \text{arithmetic in } \mathbb{R}$$

$$= 2 \cdot x + (-1) \cdot x \qquad \text{by A8}$$

$$= [1 + 1] \cdot x + (-1) \cdot x \qquad \text{arithmetic in } \mathbb{R}$$

$$= (1 \cdot x + 1 \cdot x) + (-1) \cdot x \qquad \text{by A8}$$

$$= 1 \cdot x + (1 \cdot x + (-1) \cdot x) \qquad \text{by A2}$$

Applying the cancellation lemma we deduce $0 = 1 \cdot x + (-1) \cdot x$. However, by A5 we know $1 \cdot x = x$ and by A3 0 = x + (-x) therefore $x + (-x) = x + (-1) \cdot x$ and by the cancellation lemma we conclude $(-1) \cdot x = -x$ for all $x \in V$. ∇

To prove (4.) we make use of (1.), (2.) and (3.) as appropriate. Let $c \in \mathbb{R}$ and $x \in V$ and assume $c \cdot x = 0$. To begin, suppose $c \neq 0$ thus $\frac{1}{c}c = 1$. Use A10 to multiply $c \cdot x = 0$ by $\frac{1}{c}$ hence $\frac{1}{c} \cdot (c \cdot x) = \frac{1}{c} \cdot 0$, call this \star . Consider:

$$0 = \frac{1}{c} \cdot 0 \qquad \text{by (2.)}$$

$$= \frac{1}{c} \cdot (c \cdot x) \qquad \text{by } \star$$

$$= \left(\frac{1}{c} \cdot c\right) \cdot x \qquad \text{by A6}$$

$$= 1 \cdot x \qquad \text{arithmetic in } \mathbb{R}$$

$$= x \qquad \text{by A5.}$$

Therefore, if $c \neq 0$ then x = 0. To complete the argument we suppose $x \neq 0$ and seek to show c = 0. Suppose $c \neq 0$ towards a contradiction. By $c \cdot x = 0$ and (2.),

$$\frac{1}{c} \cdot (c \cdot x) = \frac{1}{c} \cdot 0 = 0$$

But, by A6 and $\frac{1}{c}c = 1$ and A5 we find

$$\frac{1}{c} \cdot (c \cdot x) = \left(\frac{1}{c}c\right) \cdot x = 1 \cdot x = x$$

Therefore, x = 0 which clearly contradicts $x \neq 0$. Therefore, by proof by contradiction, we find c = 0. It follows that $c \cdot x = 0$ implies c = 0 or x = 0. \Box .

Perhaps we should pause to appreciate what was not in the last page or two of proofs. There were no components, no reference to the standard basis. The arguments offered depended only on the definition of the vector space itself. This means the truths we derived above are completely general; they hold for all vector spaces. In what follows past this point we sometimes use Theorem 6.1.9 without explicit reference. That said, I would like you to understand the results of the theorem do require proof and that is why we have taken some effort here to supply that proof.

6.2 subspaces

Definition 6.2.1.

Let V be a vector space. If $W \subseteq V$ such that W is a vector space with respect to the operations of V restricted to W then we say W is a **subspace** of V and write $W \leq V$.

Example 6.2.2. Let V be a vector space. Notice that $V \subseteq V$ and obviously V is a vector space with respect to its operations. Therefore $V \leq V$. Likewise, the set containing the zero vector $\{0\} \leq V$. Notice that 0 + 0 = 0 and $c \cdot 0 = 0$ so Axioms 9 and 10 are satisfied. I leave the other axioms to the reader. The subspaces $\{0\}$ is called the **trivial subspace**.

Example 6.2.3. Let $L = \{(x, y) \in \mathbb{R}^2 | ax + by = 0\}$. Define addition and scalar multiplication by the natural rules in \mathbb{R}^2 . Note if $(x, y), (z, w) \in L$ then (x, y) + (z, w) = (x + z, y + w) and a(x + z) + b(y + w) = ax + by + az + bw = 0 + 0 = 0 hence $(x, y) + (z, w) \in L$. Likewise, if $c \in \mathbb{R}$ and $(x, y) \in L$ then ax + by = 0 implies acx + bcy = 0 thus $(cx, cy) = c(x, y) \in L$. We find that L is closed under vector addition and scalar multiplication. The other 8 axioms are naturally inherited from \mathbb{R}^2 .

Example 6.2.4. If $V = \mathbb{R}^3$ then

- 1. $\{(0,0,0)\}$ is a subspace,
- 2. any line through the origin is a subspace,
- 3. any plane through the origin is a subspace.

Example 6.2.5. Let $W = \{(x, y, z) \mid x + y + z = 1\}$. Is this a subspace of \mathbb{R}^3 with the standard⁵ vector space structure? The answer is no. There are many reasons,

- 1. $(0,0,0) \notin W$ thus W has no zero vector, axiom 3 fails. Notice we cannot change the idea of "zero" for the subspace, if (0,0,0) is zero for \mathbb{R}^3 then it is the only zero for potential subspaces. Why? Because subspaces inherit their structure from the vector space which contains them.
- 2. let $(u, v, w), (a, b, c) \in W$ then u + v + w = 1 and a + b + c = 1, however $(u + a, v + b, w + c) \notin W$ since (u + a) + (v + b) + (w + c) = (u + v + w) + (a + b + c) = 1 + 1 = 2.
- 3. let $(u, v, w) \in W$ then notice that 2(u, v, w) = (2u, 2v, 2w). Observe that 2u + 2v + 2w = 2(u + v + w) = 2 hence $(2u, 2v, 2w) \notin W$. Thus axiom 10 fails, the subset W is not closed under scalar multiplication.

Of course, one reason is all it takes.

My focus on the last two axioms is not without reason. Let me explain this obsession.

Theorem 6.2.6.

Let V be a vector space and suppose $W \subset V$ with $W \neq \emptyset$ then $W \leq V$ if and only if the following two conditions hold true

1. if $x, y \in W$ then $x + y \in W$ (W is closed under addition),

2. if $x \in W$ and $c \in \mathbb{R}$ then $c \cdot x \in W$ (W is closed under scalar multiplication).

⁵yes, there is a non-standard addition which gives this space a vector space structure

Proof: (\Rightarrow) If $W \leq V$ then W is a vector space with respect to the operations of addition and scalar multiplication thus (1.) and (2.) hold true.

(\Leftarrow) Suppose W is a nonempty set which is closed under vector addition and scalar multiplication of V. We seek to prove W is a vector space with respect to the operations inherited from V. Let $x, y, z \in W$ then $x, y, z \in V$. Use A1 and A2 for V (which were given to begin with) to find

x + y = y + x and (x + y) + z = x + (y + z).

Thus A1 and A2 hold for W. By (3.) of Theorem 6.1.9 we know that $(-1) \cdot x = -x$ and $-x \in W$ since we know W is closed under scalar multiplication. Consequently, $x + (-x) = 0 \in W$ since W is closed under addition. It follows A3 is true for W. Then by the arguments just given A4 is true for W. Let $a, b \in \mathbb{R}$ and notice that by A5,A6,A7,A8 for V we find

$$1 \cdot x = x, \quad (ab) \cdot x = a \cdot (b \cdot x), \quad a \cdot (x + y) = a \cdot x + a \cdot y, \quad (a + b) \cdot x = a \cdot x + b \cdot x.$$

Thus A5,A6,A7,A8 likewise hold for W. Finally, we assumed closure of addition and scalar multiplication on W so A9 and A10 are likewise satisfied and we conclude that W is a vector space. Thus $W \leq V$. (if you're wondering where we needed W nonempty it was to argue that there exists at least one vector x and consequently the zero vector is in W.) \Box

Remark 6.2.7.

The application of Theorem 6.2.6 is a four-step process

- 1. check that $W \subset V$
- 2. check that $0 \in W$
- 3. take arbitrary $x, y \in W$ and show $x + y \in W$
- 4. take arbitrary $x \in W$ and $c \in \mathbb{R}$ and show $cx \in W$

Step (2.) is just for convenience, you could just as well find another vector in W. We need to find at least one to show that W is nonempty. Also, usually we omit comment about (1.) since it is obvious that one set is a subset of another.

Example 6.2.8. The function space $\mathcal{F}(\mathbb{R})$ has many subspaces.

- 1. continuous functions: $C(\mathbb{R})$
- 2. differentiable functions: $C^1(\mathbb{R})$
- 3. smooth functions: $C^{\infty}(\mathbb{R})$
- 4. polynomial functions
- 5. analytic functions
- 6. solution set of a linear homogeneous ODE with no singular points

The proof that each of these follows from Theorem 6.2.6. For example, f(x) = x is continuous therefore $C(\mathbb{R}) \neq \emptyset$. Moreover, the sum of continuous functions is continuous and a scalar multiple of a continuous function is continuous. Thus $C(\mathbb{R}) \leq \mathcal{F}(\mathbb{R})$. The arguments for (2.), (3.), (4.), (5.)and (6.) are identical. The solution set example is one of the most important examples for engineering and physics, linear ordinary differential equations. Also, we should note that \mathbb{R} can be replaced with some subset I of real numbers. $\mathcal{F}(I)$ likewise has subspaces $C(I), C^1(I), C^{\infty}(I)$ etc.

Example 6.2.9. The null space of a matrix $A \in \mathbb{R}^{m \times n}$ is a subspace of \mathbb{R}^m defined as follows:

$$Null(A) \equiv \{x \in \mathbb{R}^n \mid Ax = 0\}$$

Let's prove $Null(A) \leq \mathbb{R}^n$. Observe that A0 = 0 hence $0 \in Null(A)$ so the nullspace is nonempty. Suppose $x, y \in Null(A)$ and $c \in \mathbb{R}$,

$$A(x + cy) = Ax + cAy = 0 + c(0) = 0$$

thus $x + cy \in Null(A)$. Closure of addition for Null(A) follows from c = 1 and closure of scalar multiplication follows from x = 0 in the just completed calculation. \Box

Sometimes it's easier to check both scalar multiplication and addition at once. It saves some writing. If you don't understand it then don't use the trick I just used, we should understand our work.

The example that follows here introduces an important point in abstract math. Given a particular point set, there is often more than one way to define a structure on the set. Therefore, it is important to view things as more than mere sets. Instead, think about sets paired with a structure.

Example 6.2.10. Let⁶ V_p be the set of all vectors with base point $p \in \mathbb{R}^n$,

$$V_p = \{ p + v \mid v \in \mathbb{R}^n \}$$

We define a nonstandard vector addition on V_p , if $p + v, p + w \in V_p$ and $c \in \mathbb{R}$ define:

$$(p+v) +_p (p+w) = p + v + w$$
 & $c \cdot_p (p+v) = p + cv$

Clearly $+_p: V_p \times V_p \to V_p$ and $\cdot_p: \mathbb{R} \times V_p \to V_p$ are closed and verification of the other axioms is straightforward. Observe $0_p = p$ as $(p+v) +_p (p+0) = p + v + 0 = p + v$ hence $O_p = p + 0 = p$. Mainly, the vector space axioms for V_p follow from the corresponding axioms for \mathbb{R}^n . Geometrically, $+_p$ corresponds to the **tip-to-tail** rule we use in physics to add vectors. Consider S_p defined below:

$$S_p = \{ p + v \mid v \in W \le \mathbb{R}^n \}$$

Notice $0_p \in S_p$ as $0 \in W$ and $0_p = p + 0$. Furthermore, consider $p + v, p + w \in S_p$ and $c \in \mathbb{R}$

$$(p+v) +_p (p+w) = p + (v+w)$$
 & $c \cdot_p (p+v) = p + cv$

note $v + w, cv \in W$ as $W \leq \mathbb{R}^n$ is closed under addition and scalar multiplication. We find $(p+v) +_p (p+w), c \cdot_p (p+v) \in S_p$ thus $S_p \leq V_p$ by the subspace test Theorem 6.2.6.

⁶it may be better to use the notation (p, v) for p + v, this has the advantage of making the base-point p explicit whereas p can be obscured in the more geometrically direct p + v notation. Another choice is to use v_p .

In the previous example, S_p need not be a subspace with respect to the standard vector addition of column vectors. However, with the modified addition based at p it is a subspace. We often say the solution set to Ax = b with $b \neq 0$ is not a subspace. It should be understood that what is meant is that the solution set of Ax = b is not a subspace with respect to the usual vector addition. It is possible to define a different vector addition which gives the solution set of Ax = b a vector space structure. I'll let you think about the details.

Example 6.2.11. Let $W = \{A \in \mathbb{R}^{n \times n} \mid A^T = A\}$. This is the set of symmetric matrices, it is nonempty since $I^T = I$ (of course there are many other examples, we only need one to show it's nonempty). Let $A, B \in W$ and suppose $c \in \mathbb{R}$ then

$$(A+B)^T = A^T + B^T$$
 prop. of transpose
= $A + B$ since $A, B \in W$

thus $A + B \in W$ and we find W is closed under addition. Likewise let $A \in W$ and $c \in \mathbb{R}$,

$$(cA)^T = cA^T$$
 prop. of transpose
= cA since $A, B \in W$

thus $cA \in W$ and we find W is closed under scalar multiplication. Therefore, by the subspace test Theorem 6.2.6, $W \leq \mathbb{R}^{n \times n}$.

I invite the reader to modify the example above to show the set of antisymmetric matrices also forms a subspace of the vector space of square matrices.

Example 6.2.12. Let $W = \{f \in \mathcal{F}(\mathbb{R}) \mid \int_{-1}^{1} f(x) dx = 0\}$. Notice the zero function 0(x) = 0 is in W since $\int_{-1}^{1} 0 dx = 0$. Let $f, g \in W$, use linearity property of the definite integral to calculate

$$\int_{-1}^{1} (f(x) + g(x)) \, dx = \int_{-1}^{1} f(x) \, dx + \int_{-1}^{1} g(x) \, dx = 0 + 0 = 0$$

thus $f + g \in W$. Likewise, if $c \in \mathbb{R}$ and $f \in W$ then

$$\int_{-1}^{1} cf(x) \, dx = c \int_{-1}^{1} f(x) \, dx = c(0) = 0$$

thus $cf \in W$ and by subspace test Theorem 6.2.6 $W \leq \mathcal{F}(\mathbb{R})$.

Example 6.2.13. Here we continue discussion of the product space introduced in Example 6.1.8. Suppose $V = \mathbb{C}$ and $W = P_2$ then $V \times W = \{(a + ib, cx^2 + dx + e) \mid a, b, c, d, e \in \mathbb{R}\}$. Let $U = \{(a, b) \mid a, b \in \mathbb{R}\}$. We can easily show $U \leq V \times W$ by the subspace test Theorem 6.2.6 $W \leq \mathcal{F}(\mathbb{R})$. Can you think of other subspaces? Is it possible to have a subspace of $V \times W$ which is not formed from a pair of subspaces from V and W respective?

Example 6.2.14. Let W be the set of real-valued functions on \mathbb{R} for which f(a) = 0 for some fixed value $a \in \mathbb{R}$. If $f, g \in W$ and $c \in \mathbb{R}$ then (f + cg)(a) = f(a) + cg(a) = 0 + c(0) = 0 thus $f + cg \in W$. Observe W is closed under addition by the case c = 1 and W is closed under scalar multiplication by the case f = 0. Furthermore, f(x) = 0 for all $x \in \mathbb{R}$ defines the zero function which is in W. Hence $W \leq \mathcal{F}(\mathbb{R})$ by subspace test Theorem 6.2.6.

6.3 spanning sets and subspaces

The expression x+cy is a "linear combination" of x and y. Subspaces must keep linear combinations of subspace vectors from escaping the subspace. We defined linear combinations in a previous chapter (see 1.5.7). Can we use linear combinations to form a subspace?

Theorem 6.3.1.

- Let V be a vector space which contains vectors v_1, v_2, \ldots, v_k then
 - 1. the set of all linear combinations of v_1, v_2, \ldots, v_k forms a subspace of V, call it W_o
 - 2. W_o is the smallest subspace of V which contains v_1, v_2, \ldots, v_k . Any other subspace which contains v_1, v_2, \ldots, v_k also contains W_o .

Proof: Define $W_o = \{c_1v_1 + c_2v_2 + \cdots + c_kv_k \mid c_i \in \mathbb{R} \text{ for } i = 1, 2, \dots, k\}$. Notice $0 \cdot v_1 = 0$ hence $0 \in W_o$. Suppose that $x, y \in W_o$ then there exist constants c_i and b_i such that

 $x = c_1 v_1 + c_2 v_2 + \cdots + c_k v_k$ $y = b_1 v_1 + b_2 v_2 + \cdots + b_k v_k$

Consider the sum of x and y,

$$\begin{aligned} x+y &= c_1v_1 + c_2v_2 + \cdots + c_kv_k + b_1v_1 + b_2v_2 + \cdots + b_kv_k \\ &= (c_1+b_1)v_1 + (c_2+b_2)v_2 + \cdots + (c_k+b_k)v_k \end{aligned}$$

thus $x + y \in W_o$ for all $x, y \in W_o$. Let $a \in \mathbb{R}$ and observe

$$ax = a(c_1v_1 + c_2v_2 + \dots + c_kv_k) = ac_1v_1 + ac_2v_2 + \dots + ac_kv_k$$

thus $cx \in W_o$ for all $x \in W_o$ and $c \in \mathbb{R}$. Thus by the subspace test theorem we find $W_o \leq V$.

To prove (2.) we suppose R is any subspace of V which contains v_1, v_2, \ldots, v_k . By definition R is closed under scalar multiplication and vector addition thus all linear combinations of v_1, v_2, \ldots, v_k must be in R hence $W_o \subseteq R$. Finally, it is clear that $v_1, v_2, \ldots, v_k \in W_o$ since $v_1 = 1v_1 + 0v_2 + \cdots + 0v_k$ and $v_2 = 0v_1 + 1v_2 + \cdots + 0v_k$ and so forth. \Box

Definition 6.3.2.

Let $S = \{v_1, v_2, \dots, v_k\}$ be a finite set of vectors in a vector space V then span(S) is defined to be the set of all linear combinations of S:

span{
$$v_1, v_2, \dots, v_k$$
} = { $\sum_{i=1}^k c_i v_i \mid c_i \in \mathbb{R} \text{ for } i = 1, 2, \dots, k$ }

If W = span(S) then we say that S is a **generating set** for W. We also say S **spans** W in this case. Furthermore, if S is an infinite set then span(S) is defined to be all possible finite linear combinations from S.

In view of Theorem 6.3.1 the definition above is equivalent to defining span(S) to be the smallest subspace which contains S.

Example 6.3.3. Proposition 1.5.8 explained how \mathbb{R}^n was spanned by the standard basis; $\mathbb{R}^n = span\{e_i\}_{i=1}^n$. Likewise, Proposition 3.3.2 showed the $m \times n$ matrix units E_{ij} spanned the set of all $m \times n$ matrices; $\mathbb{R}^{m \times n} = span\{E_{ij}\}_{i,j=1}^n$.

Example 6.3.4. Let $S = \{1, x, x^2, ..., x^n\}$ then $span(S) = P_n$. For example,

$$span\{1, x, x^2\} = \{ax^2 + bx + c \mid a, b, c \in \mathbb{R}\} = P_2$$

The set of all polynomials is spanned by $\{1, x, x^2, x^3, ...\}$. We are primarily interested in the span of finite sets however this case is worth mentioning.

Example 6.3.5. Let $W = \{[s + t, 2s + t, 3s + t]^T \mid s, t \in \mathbb{R}\}$. We can show W is a subspace of $\mathbb{R}^{3 \times 1}$. What is a generating set of W? Let $w \in W$ then by definition there exist $s, t \in \mathbb{R}$ such that

$$w = \begin{bmatrix} s+t\\2s+t\\3s+t \end{bmatrix} = s \begin{bmatrix} 1\\2\\3 \end{bmatrix} + t \begin{bmatrix} 1\\1\\1 \end{bmatrix}$$

Thus $w \in span\{[1,2,3]^T, [1,1,1]^T\}$ and it follows $W \subseteq span\{[1,2,3]^T, [1,1,1]^T\}$. Conversely, if $y \in span\{[1,2,3]^T, [1,1,1]^T\}$ then there exist $c_1, c_2 \in \mathbb{R}$ such that $y = c_1[1,2,3]^T + c_2[1,1,1]^T$. But then $y = [c_1 + c_2, 2c_1 + c_2, 3c_1 + c_2]^T$ so it is clear $y \in W$, therefore $span\{[1,2,3]^T, [1,1,1]^T\} \subseteq W$. It follows that $W = span\{[1,2,3]^T, [1,1,1]^T\}$. Finally, Theorem 6.3.1 gives us $W \leq \mathbb{R}^3$.

The lesson of the last example is that we can show a particular space is a subspace by finding its generating set. Theorem 6.3.1 tells us that any set generated by a span is a subspace. This test is only convenient for subspaces which are defined as some sort of span. In that case we can immediately conclude the subset is in fact a subspace.

Example 6.3.6. Suppose $a, b, c \in \mathbb{R}$ and $a \neq 0$. Consider the differential equation ay'' + by' + cy = 0. There is a theorem in the study of differential equations which states **every** solution can be written as a linear combination of a pair of special solutions y_1, y_2 ; we say $y = c_1y_1 + c_2y_2$ is the "general solution" in the terminology of Math 334. In other words, there exist solutions y_1, y_2 such that the solution set S of ay'' + by' + cy = 0 is

$$S = span\{y_1, y_2\}.$$

Since S is a span it is clear that $S \leq \mathcal{F}(\mathbb{R})$.

Example 6.3.7. Suppose L = P(D) where D = d/dx and P is a polynomial with real coefficients. This makes L a smooth operator on the space of smooth functions. Suppose deg(P) = n, a theorem in differential equations states that there exist solutions y_1, y_2, \ldots, y_n of L[y] = 0 such that **every** solution of L[y] = 0 can be written in the form $y = c_1y_1 + c_2y_2 + \cdots + c_ny_n$ for constants $c_1, c_2, \ldots, c_n \in \mathbb{R}$. In other words, the solution set S of L[y] = 0 is formed from a span:

$$S = span\{y_1, y_2, \dots, y_n\}.$$

Notice the last example is a subcase of this example. Simply set $L = aD^2 + bD + c$.

Perhaps the examples above were too abstract for you at this point. Let me give a couple specific examples in the same vein.

Example 6.3.8. Consider y' = y. Or, taking t as the independent variable, $\frac{dy}{dt} = y$. Separation of variables (that you are expected to know from calculus II) shows $\frac{dy}{y} = dt$ hence $\ln |y| = t + c$. It follows that $y = \pm e^c e^t$. Note y = 0 is also a solution of y' = y. In total, we find solutions of the form $y = c_1 e^t$. The solution set of this differential equation is a span; $S = span\{e^t\} \leq \mathcal{F}(\mathbb{R})$.

Example 6.3.9. Consider y'' - y = 0. I invite the reader to verify that $y_1 = \cosh(t)$ and $y_2 = \sinh(t)$ are solutions. The solution set is $S = span\{y_1, y_2\} \leq \mathcal{F}(\mathbb{R})$.

Example 6.3.10. Consider y'' + y = 0. I invite the reader to verify that $y_1 = \cos(t)$ and $y_2 = \sin(t)$ are solutions. The solution set is $S = span\{y_1, y_2\} \leq \mathcal{F}(\mathbb{R})$. Physically, this could represent Newton's equation for a spring with mass m = 1 and stiffness k = 1, the set of all possible physical motions forms a linear subspace of function space.

Example 6.3.11. Consider, y''' = 0. Integrate both sides to find $y'' = c_1$. Integrate again to find $y' = c_1t + c_2$. Integrate once more, $y = c_1\frac{1}{2}t^2 + c_2t + c_3$. The general solution of y''' = 0 is a subspace S of function space:

$$S = span\left\{\frac{1}{2}t^2, t, 1\right\} \le \mathcal{F}(\mathbb{R})$$

Physically, we often consider the situation $c_1 = -g$.

Examples 6.3.8 and 6.3.11 are fair game for test, quizzes etc... they only assume prerequisite knowledge plus linear algebra. In constrast, I don't expect you can find y_1, y_2 as in Examples 6.3.9 and 6.3.10 since the Differential Equations course is not a prerequisite.

Example 6.3.12. Let $A \in \mathbb{R}^{m \times n}$. Define column space of A as the span of the columns of A:

$$Col(A) = span\{col_j(A) \mid j = 1, 2, \dots, n\}$$

this is clearly a subspace of \mathbb{R}^m since each column has as many components as there are rows in A. We also define row space as the span of the rows:

$$Row(A) = span\{row_i(A) \mid i = 1, 2, \dots, m\}$$

this is clearly a subspace of $\mathbb{R}^{1 \times n}$ since it is formed as a span of vectors. Since the columns of A^T are the rows of A and the rows of A^T are the columns of A we can conclude that $Col(A^T) = Row(A)$ and $Row(A^T) = Col(A)$.

I would remind the reader we have the CCP and associated techniques to handle spanning questions for column vectors. In contrast, the following example requires a direct assault⁷:

Example 6.3.13. Is $E_{11} \in span\{E_{12} + 2E_{11}, E_{12} - E_{11}\}$? Assume $E_{ij} \in \mathbb{R}^{2 \times 2}$ for all i, j. We seek to find solutions of

$$E_{11} = a(E_{12} + 2E_{11}) + b(E_{12} - E_{11})$$

in explicit matrix form the equation above reads:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = a \left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \right) + b \left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \right)$$
$$= \begin{bmatrix} 2a & a \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -b & b \\ 0 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 2a - b & a + b \\ 0 & 0 \end{bmatrix}$$

⁷However, once we have the idea of *coordinates* ironed out then we can use the CCP tricks on the coordinate vectors then push back the results to the world of abstract vectors. For now we'll just confront each question by brute force. For an example such as this, the method used here is as good as our later methods.

thus 1 = 2a - b and 0 = a + b. Substitute a = -b to find 1 = 3a hence $a = \frac{1}{3}$ and $b = \frac{-1}{3}$. Indeed,

$$\frac{1}{3}(E_{12} + 2E_{11}) - \frac{1}{3}(E_{12} - E_{11}) = \frac{2}{3}E_{11} + \frac{1}{3}E_{11} = E_{11}.$$

Therefore, $E_{11} \in span\{E_{12} + 2E_{11}, E_{12} - E_{11}\}.$

Example 6.3.14. Find a generating set for the set of symmetric 2×2 matrices. That is find a set S of matrices such that $span(S) = \{A \in \mathbb{R}^{2 \times 2} \mid A^T = A\} = W$. There are many approaches, but I find it most natural to begin by studying the condition which defines W. Let $A \in W$ and note

$$A^{T} = A \& A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

this means we need b = c but we find no particular condition on a or d. Notice $A \in W$ implies

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = aE_{11} + b(E_{12} + E_{21}) + dE_{22}$$

Thus $A \in W$ implies $A \in span\{E_{11}, E_{12} + E_{21}, E_{22}\}$, hence $W \subseteq span\{E_{11}, E_{12} + E_{21}, E_{22}\}$. Conversely, if $B \in span\{E_{11}, E_{12} + E_{21}, E_{22}\}$ then there exist $c_1, c_2, c_3 \in \mathbb{R}$ such that

$$B = c_1 E_{11} + c_2 (E_{12} + E_{21}) + c_3 E_{22}$$

but this means

$$B = \left[\begin{array}{cc} c_1 & c_2 \\ c_2 & c_3 \end{array} \right]$$

so B is symmetric and it follows $span\{E_{11}, E_{12}+E_{21}, E_{22}\} \subseteq W$. Consequently $W = span\{E_{11}, E_{12}+E_{21}, E_{22}\}$ and the set $\{E_{11}, E_{12}+E_{21}, E_{22}\}$ generates W. This is not unique, there are many other sets which also generate W. For example, if we took $\bar{S} = \{E_{11}, E_{12}+E_{21}, E_{22}, E_{11}+E_{22}\}$ then the span of \bar{S} would still work out to W.

I could use the lemma below to prove the theorem that follows, however, I thought it wise to leave the proof of the theorem as it is written so you can compare the methods of argument. Index techniques save some writing, but, many students need to see the proof of the theorem before the lemma. So, you might skip past the lemma in your first read.

Lemma 6.3.15.

The linear combination of linear combinations is a linear combination.

Proof: Suppose V is a vector space. Let $s_i = \sum_{j=1}^{n_i} c_{ij} t_{ij}$ where $c_{ij} \in \mathbb{R}$ and $t_{ij} \in V$ for $n_i, i \in \mathbb{N}$ with $i = 1, 2, \ldots, k$. Let $b_1, \ldots, b_k \in \mathbb{R}$ and consider by (2.) of Proposition 1.3.3

$$\sum_{i=1}^{k} b_i s_i = \sum_{i=1}^{k} b_i \left(\sum_{j=1}^{n_i} c_{ij} t_{ij} \right) = \sum_{i=1}^{k} \sum_{j=1}^{n_i} b_i c_{ij} t_{ij}.$$

Notice, this is a linear combination as $b_i c_{ij} \in \mathbb{R}$. \Box

Theorem 6.3.16.

If $S = \{s_1, s_2, \ldots, s_k\}$ and $T = \{t_1, t_2, \ldots, t_r\}$ are subsets of a vector space V then span(S) = span(T) if and only if every vector in S is a linear combination of vectors in T and every vector in T is a linear combination of vectors in S.

Proof: (\Rightarrow) Assume span(S) = span(T). If $v \in S$ then $v \in span(S)$ hence $v \in span(T)$ and it follows that v is a linear combination of vectors in T. If $w \in T$ then $w \in span(T)$ hence $w \in span(S)$ and by definition of the span(S) we find w is a linear combination of vectors in S.

(\Leftarrow) Assume every vector in S is a linear combination of vectors in T and every vector in T is a linear combination of vectors in S. Suppose $v \in Span(S)$ then v is a linear combination of vectors in S, say

$$v = c_1 s_1 + c_2 s_2 + \dots + c_k s_k.$$

Furthermore, each vector in S is a linear combination of vectors in T by assumption so there exist constants d_{ij} such that

$$s_i = d_{i1}t_1 + d_{i2}t_2 + \dots + d_{ir}t_r$$

for each $i = 1, 2, \ldots, k$. Thus,

$$v = c_1 s_1 + c_2 s_2 + \dots + c_k s_k.$$

= $c_1 (d_{11}t_1 + d_{12}t_2 + \dots + d_{1r}t_r) + c_2 (d_{21}t_1 + d_{22}t_2 + \dots + d_{2r}t_r) + \dots + c_k (d_{k1}t_1 + d_{k2}t_2 + \dots + d_{kr}t_r)$
= $(c_1 d_{11} + c_2 d_{21} + \dots + c_k d_{k1})t_1 + (c_1 d_{12} + c_2 d_{22} + \dots + c_k d_{k2})t_2 + \dots + (c_1 d_{1r} + c_2 d_{2r} + \dots + c_k d_{kr})t_r$

thus v is a linear combination of vectors in T, in other words $v \in span(T)$ and we find $span(S) \subseteq span(T)$. Notice, we just proved that a linear combination of linear combinations is again a linear combination. Almost the same argument shows $span(T) \subseteq span(S)$ hence span(S) = span(T). \Box .

6.4 linear independence

We have seen a variety of generating sets in the preceding section. In the last example I noted that if we added an additional vector $E_{11} + E_{22}$ then the same span would be created. The vector $E_{11} + E_{22}$ is **redundant** since we already had E_{11} and E_{22} . In particular, $E_{11} + E_{22}$ is a linear combination of E_{11} and E_{22} so adding it will not change the span. How can we decide if a vector is absolutely necessary for a span? In other words, if we want to span a subspace W then how do we find a **minimal spanning set**? We want a set of vectors which does not have any linear dependences. We say such vectors are linearly independent. Let me be precise⁸:

Definition 6.4.1.

If a vector v_k can be written as a linear combination of vectors $\{v_1, v_2, \ldots, v_{k-1}\}$ then we say that the vectors $\{v_1, v_2, \ldots, v_{k-1}, v_k\}$ are **linearly dependent**. If the vectors $\{v_1, v_2, \ldots, v_{k-1}, v_k\}$ are not linear dependent then they are **linearly independent (LI)**.

⁸ if you have a sense of deja vu here, it is because I uttered many of the same words in the context of \mathbb{R}^n . Notice, in constrast, I now consider the abstract case. We cannot use the CCP directly here

Example 6.4.2. Let $v = \cos^2(t)$ and $w = 1 + \cos(2t)$. Clearly v, w are linearly dependent since w = 2v. We should remember from trigonometry $\cos^2(t) = \frac{1}{2}(1 + \cos(2t))$.

I often quote the following proposition as the definition of linear independence, it is an equivalent statement and as such can be used as the definition. If this was our definition then our definition would become a proposition. Math always has a certain amount of this sort of ambiguity.

Proposition 6.4.3.

Let
$$v_1, v_2, \ldots, v_k \in V$$
 a vector space. The set of vectors $\{v_1, v_2, \ldots, v_k\}$ is LI iff
 $c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0 \implies c_1 = c_2 = \cdots = c_k = 0.$

Proof: (\Rightarrow) Suppose $\{v_1, v_2, \ldots, v_k\}$ is linearly independent. Assume that there exist constants c_1, c_2, \ldots, c_k such that

$$c_1 v_1 + c_2 v_2 + \dots + c_k v_k = 0$$

and at least one constant, say c_j , is nonzero. Then we can divide by c_j to obtain

$$\frac{c_1}{c_j}v_1 + \frac{c_2}{c_j}v_2 + \dots + v_j + \dots + \frac{c_k}{c_j}v_k = 0$$

solve for v_j , (we mean for $\hat{v_j}$ to denote the deletion of v_j from the list)

$$v_j = -\frac{c_1}{c_j}v_1 - \frac{c_2}{c_j}v_2 - \dots - \hat{v_j} - \dots - \frac{c_k}{c_j}v_k$$

but this means that v_j linearly depends on the other vectors hence $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent. This is a contradiction, therefore $c_j = 0$. Note j was arbitrary so we may conclude $c_j = 0$ for all j. Therefore, $c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0 \implies c_1 = c_2 = \cdots = c_k = 0$.

Proof: (\Leftarrow) Assume that

$$c_1v_1 + c_2v_2 + \dots + c_kv_k = 0 \implies c_1 = c_2 = \dots = c_k = 0.$$

If $v_j = b_1v_1 + b_2v_2 + \cdots + \widehat{b_jv_j} + \cdots + b_kv_k$ then $b_1v_1 + b_2v_2 + \cdots + b_jv_j + \cdots + b_kv_k = 0$ where $b_j = -1$, this is a contradiction. Therefore, for each j, v_j is not a linear combination of the other vectors. Consequently, $\{v_1, v_2, \ldots, v_k\}$ is linearly independent. \Box

What follows next yet another equivalent definition of linear independence. In short, our ability to equate coefficients for a given set of objects is interchangeable with the LI of the set of objects.

Proposition 6.4.4.

S is a linearly independent set of vectors iff for all
$$v_1, v_2, \ldots, v_k \in S$$
,

$$a_1v_1 + a_2v_2 + \dots + a_kv_k = b_1v_1 + b_2v_2 + \dots + b_kv_k$$

implies $a_i = b_i$ for each i = 1, 2, ..., k. In other words, we can equate coefficients of linearly indpendent vectors. And, conversely if a set of vectors allows for equating coefficients then it is linearly independent.

Proof: likely homework problem. \Box

In retrospect, partial fractions is based on the LI of the basic rational functions. The technique of equating coefficients only made sense because the set of functions involved was in fact LI.

Proposition 6.4.5.

If S is a finite set of vectors which contains the zero vector then S is linearly dependent.

Proof: Let $\{\vec{0}, v_2, \dots, v_k\} = S$ and observe that

$$1\vec{0} + 0v_2 + \dots + 0v_k = 0$$

Thus $c_1 \vec{0} + c_2 v_2 + \cdots + c_k v_k = 0$ does not imply $c_1 = 0$ hence the set of vectors is not linearly independent. Thus S is linearly dependent. \Box

Proposition 6.4.6.

Let v and w be nonzero vectors.

v, w are linearly dependent $\Leftrightarrow \exists k \neq 0 \in \mathbb{R}$ such that v = kw.

Proof: Suppose v, w are linearly dependent then there exist constants c_1, c_2 , not all zero, such that $c_1v + c_2w = 0$. Suppose that $c_1 = 0$ then $c_2w = 0$ hence $c_2 = 0$ or w = 0 by (4.) of Theorem 6.1.9. But this is a contradiction since v, w are nonzero and at least one of c_1, c_2 must be nonzero. Therefore, $c_1 \neq 0$. Likewise, if $c_2 = 0$ we find a similar contradiction. Hence c_1, c_2 are both nonzero and we calculate $v = (-c_2/c_1)w$, identify that $k = -c_2/c_1$. \Box

Remark 6.4.7.

We should keep in mind that in the abstract context statements such as "v and w go in the same direction" or "u is contained in the plane spanned by v and w" are not statements about ordinary three dimensional geometry. Moreover, you **cannot** write that $u, v, w \in \mathbb{R}^n$ unless you happen to be working with that rather special vector space. These "vectors" could be matrices, polynomials or even operators. All of this said, we will find a way to correctly think of an abstract vector space V as another version of \mathbb{R}^n . We'll see how Vand \mathbb{R}^n correspond, we will not be so careless as to say they are equal.

Given a set of vectors in \mathbb{R}^n the question of LI is elegantly answered by the CCP. In examples that follow in this section we leave the comfort zone and study LI in abstract vector spaces. For now we only have brute force at our disposal. In other words, I'll argue directly from the definition without the aid of the CCP from the outset.

Example 6.4.8. Suppose $f(x) = \cos(x)$ and $g(x) = \sin(x)$ and define $S = \{f, g\}$. Is S linearly independent with respect to the standard vector space structure on $\mathcal{F}(\mathbb{R})$? Let $c_1, c_2 \in \mathbb{R}$ and assume that

$$c_1f + c_2g = 0$$

It follows that $c_1f(x) + c_2g(x) = 0$ for each $x \in \mathbb{R}$. In particular,

$$c_1\cos(x) + c_2\sin(x) = 0$$

for each $x \in \mathbb{R}$. Let x = 0 and we get $c_1 \cos(0) + c_2 \sin(0) = 0$ thus $c_1 = 0$. Likewise, let $x = \pi/2$ to obtain $c_1 \cos(\pi/2) + c_2 \sin(\pi/2) = 0 + c_2 = 0$ hence $c_2 = 0$. We have shown that $c_1 f + c_2 g = 0$ implies $c_1 = c_2 = 0$ thus $S = \{f, g\}$ is a linearly independent set.

Example 6.4.9. Let $f_n(t) = t^n$ for n = 0, 1, 2, ... Suppose $S = \{f_0, f_1, ..., f_n\}$. Show S is a linearly independent subset of function space. Assume $c_0, c_1, ..., c_n \in \mathbb{R}$ and

$$c_0 f_0 + c_1 f_1 + c_2 f_2 + \dots + c_n f_n = 0.$$
 *

I usually skip the expression above, but I'm including this extra step to emphasize the distinction between the function and its formula. The \star equation is a function equation, it implies

$$c_0 + c_1 t + c_2 t^2 + \dots + c_n t^n = 0$$
 *

for all $t \in \mathbb{R}$. Evaluate $\star\star$ at t = 0 to obtain $c_0 = 0$. Differentiate \star^2 and find

$$c_1 + 2c_2t + \dots + nc_nt^{n-1} = 0 \qquad \star^3$$

Evaluate \star^3 at t = 0 to obtain $c_1 = 0$. If we continue to differentiate and evaluate we will similarly obtain $c_2 = 0$, $c_3 = 0$ and so forth all the way up to $c_n = 0$. Therefore, \star implies $c_0 = c_1 = \cdots = c_n = 0$.

Linear dependence in function space is sometimes a source of confusion for students. The idea of evaluation doesn't help in the same way as it just has in the two examples above.

Example 6.4.10. Let f(t) = t - 1 and $g(t) = t + t^2$ is f linearly dependent on g? A common mistake is to say something like f(1) = 1 - 1 = 0 so $\{f, g\}$ is linearly independent since it contains zero. Why is this wrong? The reason is that we have confused the value of the function with the function itself. If f(t) = 0 for all $t \in \mathbb{R}$ then f is the zero function which is the zero vector in function space. Many functions will be zero at a point but that doesn't make them the zero function. To prove linear dependence we must show that there exists $k \in \mathbb{R}$ such that f = kg, but this really means that f(t) = kg(t) for all $t \in \mathbb{R}$ in the current context. I leave it to the reader to prove that $\{f, g\}$ is in fact LI. You can evaluate at t = 1 and t = 0 to obtain equations for c_1, c_2 which have a unique solution of $c_1 = c_2 = 0$.

Example 6.4.11. Let $f(t) = t^2 - 1$, $g(t) = t^2 + 1$ and $h(t) = 4t^2$. Suppose

$$c_1(t^2 - 1) + c_2(t^2 + 1) + c_3(4t^2) = 0$$

A little algebra reveals,

$$c_1 + c_2 + 4c_3)t^2 - (c_1 - c_2)1 = 0$$

Using linear independence of t^2 and 1 we find

$$c_1 + c_2 + 4c_3 = 0$$
 and $c_1 - c_2 = 0$

We find infinitely many solutions,

$$c_1 = c_2$$
 and $c_3 = -\frac{1}{4}(c_1 + c_2) = -\frac{1}{2}c_2$

Therefore, \star allows nontrivial solutions. Take $c_2 = 1$,

$$1(t^{2} - 1) + 1(t^{2} + 1) - \frac{1}{2}(4t^{2}) = 0.$$

We can write one of these functions as a linear combination of the other two,

$$f = -g + \frac{1}{2}h.$$

Once we get past the formalities of the particular vector space structure it always comes back to solving systems of linear equations.

6.5 bases and dimension

We have seen that linear combinations can generate vector spaces. We have also seen that sometimes we can remove a vector from the generating set and still generate the whole vector space. For example,

$$span\{e_1, e_2, e_1 + e_2\} = \mathbb{R}^2$$

and we can remove any one of these vector and still span \mathbb{R}^2 ,

$$span\{e_1, e_2\} = span\{e_1, e_1 + e_2\} = span\{e_2, e_1 + e_2\} = \mathbb{R}^2$$

However, if we remove another vector then we will not span \mathbb{R}^2 . A generating set which is just big enough is called a basis. We can remove vectors which are linearly dependent on the remaining vectors without changing the span. Therefore, we should expect that a minimal spanning set is linearly independent.

Definition 6.5.1.

A **basis** for a vector space V is a set of vectors S such that

1. V = span(S),

2. S is linearly independent.

Example 6.5.2. It is not hard to show that $B_1 = \{e_1, e_2\}$ and $B_2 = \{e_1, e_1 + e_2\}$ and $B_3 = \{e_2, e_1 + e_2\}$ are linearly independent sets. Furthermore, each spans \mathbb{R}^2 . Therefore, B_1, B_2, B_3 are bases for \mathbb{R}^2 . In particular, $B_1 = \{e_1, e_2\}$ is called the <u>standard basis</u>.

Example 6.5.3. I called $\{e_1, e_2, \ldots, e_n\}$ the standard basis of \mathbb{R}^n . Since $v \in \mathbb{R}^n$ can be written as

$$v = \sum_{i} v_i e_i$$

it follows $\mathbb{R}^n = span\{e_i \mid 1 \leq i \leq n\}$. Moreover, linear independence of $\{e_i \mid 1 \leq i \leq n\}$ follows from a simple calculation:

$$0 = \sum_{i} c_{i} e_{i} \Rightarrow 0 = \left\lfloor \sum_{i} c_{i} e_{i} \right\rfloor_{k} = \sum_{i} c_{i} \delta_{ik} = c_{k}$$

hence $c_k = 0$ for all k. Thus $\{e_i \mid 1 \leq i \leq n\}$ is a basis for \mathbb{R}^n , we continue to call it the standard basis of \mathbb{R}^n . The vectors e_i are also called "unit-vectors".

Example 6.5.4. Since $A \in \mathbb{R}^{m \times n}$ can be written as

$$A = \sum_{i,j} A_{ij} E_{ij}$$

it follows $\mathbb{R}^{m \times n} = span\{E_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$. Moreover, linear independence of $\{E_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ follows from a simple calculation:

$$0 = \sum_{i,j} c_{ij} E_{ij} \Rightarrow 0 = \left[\sum_{i,j} c_{ij} E_{ij} \right]_{kl} = \sum_{i,j} c_{ij} \delta_{ik} \delta_{jl} = c_{kl}$$

hence $c_{kl} = 0$ for all k, l. Thus $\{E_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ is a basis for $\mathbb{R}^{m \times n}$, we continue to call it the standard basis of $\mathbb{R}^{m \times n}$. The matrices E_{ij} are also called "unit-matrices".

Proposition 6.5.5.

Suppose $B = \{f_1, f_2, \dots, f_n\}$ is a basis for V. If $v \in V$ with $v = \sum_{i=1}^n x_i f_i$ and $v = \sum_{i=1}^n y_i f_i$ for constants $x_i, y_i \in \mathbb{R}$. Then $x_i = y_i$ for $i = 1, 2, \dots, n$.

Proof: Suppose $v = x_1f_1 + x_2f_2 + \cdots + x_nf_n$ and $v = y_1f_1 + y_2f_2 + \cdots + y_nf_n$ notice that

$$0 = v - v = (x_1 f_1 + x_2 f_2 + \dots + x_n f_n) - (y_1 f_1 + y_2 f_2 + \dots + y_n f_n)$$

= $(x_1 - y_1)f_1 + (x_2 - y_2)f_2 + \dots + (x_n - y_n)f_n$

then by LI of the basis vectors we find $x_i - y_i = 0$ for each *i*. Thus $x_i = y_i$ for all *i*. \Box

Notice that both LI and spanning were necessary for the idea of a coordinate vector (defined below) to make sense.

Definition 6.5.6.

Suppose
$$B = \{f_1, f_2, \dots, f_n\}$$
 is a basis for V . If $v \in V$ has
 $v = v_1 f_1 + v_2 f_2 + \dots + v_n f_n$
then $[v]_B = [v_1 \ v_2 \ \dots \ v_n]^T \in \mathbb{R}^n$ is called the **coordinate vector** of v with respect to B .

Technically, the each basis considered in the course is an "ordered basis". This means the set of vectors that forms the basis has an ordering to it. This is more structure than just a plain set since basic set theory does not distinguish $\{1,2\}$ from $\{2,1\}$. I should always say "we have an ordered basis" but I will not (and most people do not) say that in this course. Let it be understood that when we list the vectors in a basis they are listed in order and we cannot change that order without changing the basis. For example v = (1, 2, 3) has coordinate vector $[v]_{B_1} = (1, 2, 3)$ with respect to $B_1 = \{e_1, e_2, e_3\}$. On the other hand, if $B_2 = \{e_2, e_1, e_3\}$ then the coordinate vector of v with respect to B_2 is $[v]_{B_2} = (2, 1, 3)$.

Example 6.5.7. Let $\beta = \{E_{11}, E_{12}, E_{22}, E_{21}\}$. Observe: $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = aE_{11} + bE_{12} + dE_{22} + cE_{21}$. Therefore, $[A]_{\beta} = (a, b, d, c)$.

Example 6.5.8. Consider $\beta = \{(t+1)^2, t+1, 1\}$ and calculate the coordinate vector of $f(t) = t^2$ with respect to β . I often use an adding zero trick for such a problem:

$$f(t) = t^2 = (t+1-1)^2 = (t+1)^2 - 2(t+1) + 1.$$

From the expression above we can read that $[f(t)]_{\beta} = (1, -2, 1)$.

Example 6.5.9. Suppose Av = b has solution v = (1, 2, 3, 4). It follows that A has 4 columns. Define,

$$\beta = \{ col_4(A), col_3(A), col_2(A), col_1(A) \}$$

Given that (1, 2, 3, 4) is a solution of Av = b we know:

$$col_1(A) + 2col_2(A) + 3col_3(A) + 4col_4(A) = b$$

Given the above, we can deduce $[b]_{\beta} = (4, 3, 2, 1)$.

The three examples above were simple enough that not much calculation was needed. Understanding the definition of basis was probably the hardest part. In general, finding the coordinates of a vector with respect to a given basis is a spanning problem.

Example 6.5.10. Let $v = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ find the coordinates of v relative to B_1, B_2 and B_3 where $B_1 = \{e_1, e_2\}$ and $B_2 = \{e_1, e_1 + e_2\}$ and $B_3 = \{e_2, e_1 + e_2\}$. We'll begin with the standard basis, (I hope you could see this without writing it)

$$v = \begin{bmatrix} 1\\3 \end{bmatrix} = 1 \begin{bmatrix} 1\\0 \end{bmatrix} + 3 \begin{bmatrix} 0\\1 \end{bmatrix} = 1e_1 + 3e_2$$

thus $[v]_{B_1} = [1 \ 3]^T$. Find coordinates relative to the other two bases is not quite as obvious. Begin with B_2 . We wish to find x, y such that

$$v = xe_1 + y(e_1 + e_2)$$

we can just use brute-force,

$$v = e_1 + 3e_2 = xe_1 + y(e_1 + e_2) = (x + y)e_1 + ye_2$$

using linear independence of the standard basis we find 1 = x + y and y = 3 thus x = 1 - 3 = -2and we see $v = -2e_1 + 3(e_1 + e_2)$ so $[v]_{B_2} = [-2 \ 3]^T$. This is interesting, the same vector can have different coordinate vectors relative to distinct bases. Finally, let's find coordinates relative to B_3 . I'll try to be more clever this time: we wish to find x, y such that

$$v = xe_2 + y(e_1 + e_2) \quad \Leftrightarrow \quad \begin{bmatrix} 1\\ 3 \end{bmatrix} = \begin{bmatrix} 0 & 1\\ 1 & 1 \end{bmatrix} \begin{bmatrix} x\\ y \end{bmatrix}$$

We can solve this via the augemented coefficient matrix

$$rref \begin{bmatrix} 0 & 1 & | & 1 \\ 1 & 1 & | & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & | & 2 \\ 0 & 1 & | & 1 \end{bmatrix} \quad \Leftrightarrow \quad x = 2, \ y = 1.$$

Thus, $[v]_{B_3} = [2 \ 1]^T$. Notice this is precisely the rightmost column in the rref matrix. Perhaps my approach for B_3 is a little like squashing a fly with with a dumptruck. However, once we get to an example with 4-component vectors you may find the matric technique useful.

Example 6.5.11. Given that $B = \{b_1, b_2, b_3, b_4\} = \{e_1 + e_2, e_2 + e_3, e_3 + e_4, e_4\}$ is a basis for \mathbb{R}^4 find coordinates for $v = [1, 2, 3, 4]^T \in \mathbb{R}^4$. Given the discussion in the preceding example it is clear we can find coordinates $[x_1, x_2, x_3, x_4]^T$ such that $v = \sum_i x_i b_i$ by calculating $rref[b_1|b_2|b_3|b_4|v]$ the rightmost column will be $[v]_B$.

$$rref \begin{bmatrix} 1 & 0 & 0 & 0 & | & 1 \\ 1 & 1 & 0 & 0 & | & 2 \\ 0 & 1 & 1 & 0 & | & 3 \\ 0 & 0 & 1 & 1 & | & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & | & 1 \\ 0 & 1 & 0 & 0 & | & 1 \\ 0 & 0 & 1 & 0 & | & 2 \\ 0 & 0 & 0 & 1 & | & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} v \\ v \\ B \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \end{bmatrix}$$

This calculation should be familar. We discussed it at length in the spanning section.

Remark 6.5.12.

Curvelinear coordinate systems from calculus III are in a certain sense more general than the idea of a coordinate system in linear algebra. If we focus our attention on a single point in space then a curvelinear coordinate system will produce three linearly independent vectors which are tangent to the coordinate curves. However, if we go to a different point then the curvelinear coordinate system will produce three different vectors in general. For example, in spherical coordinates the radial unit vector is $e_{\rho} = \langle \cos \theta \sin \phi, \sin \theta \sin \phi, \cos \phi \rangle$ and you can see that different choices for the angles θ, ϕ make e_{ρ} point in different directions. In contrast, in this course we work with vector spaces. Our coordinate systems have the same basis vectors over the whole space. Vector spaces are examples of *f* lat manifolds since they allow a single global coordinate system. Vector spaces also allow for curvelinear coordinates (which are not coordinates in the sense of linear algebra). However the converse is not true; spaces with nonzero curvature do not allow for global coordinates. I digress, we may have occassion to discuss these matters more cogently in our Advanced Calculus course (Math 332)(join us)

Definition 6.5.13.

If a vector space V has a basis which consists of a finite number of vectors then we say that V is **finite-dimensional** vector space. Otherwise V is said to be **infinite-dimensional**. We define the number of elements in a finite set $S = \{s_1, s_2, \ldots, s_k\}$ to be #(S) = k.

Example 6.5.14. \mathbb{R}^n , $\mathbb{R}^{m \times n}$, P_n are examples of finite-dimensional vector spaces. On the other hand, $\mathcal{F}(\mathbb{R})$, $C^0(\mathbb{R})$, $C^1(\mathbb{R})$, $C^{\infty}(\mathbb{R})$ are infinite-dimensional.

Example 6.5.15. We can prove that S from Example 6.3.14 is linearly independent, thus symmetric 2×2 matrices have a S as a basis

$$S = \{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \}$$

thus the dimension of the vector space of 2×2 symmetric matrices is 3. (notice \overline{S} from that example is not a basis because it is linearly dependent). While we're thinking about this let's find the coordinates of $A = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$ with respect to S. Denote $[A]_S = [x, y, z]^T$. We calculate,

$$\begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} = x \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + y \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + z \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \implies [A]_S = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

6.5.1 how to calculate a basis for a span of row or column vectors

Given some subspace of \mathbb{R}^n we would like to know how to find a basis for that space. In particular, if $V = span\{v_1, v_2, \ldots, v_k\}$ then what is a basis for W? Likewise, given some set of row vectors $W = \{w_1, w_2, \ldots, w_k\} \subset \mathbb{R}^{1 \times n}$ how can we select a basis for span(W). We would like to find answers to these question since most subspaces are characterized either as spans or solution sets(see the next section on Null(A)). We already have the tools to answer these questions, we just need to apply them to the tasks at hand.

Proposition 6.5.16.

Let $W = span\{v_1, v_2, \ldots, v_k\} \subset \mathbb{R}^n$ then a basis for W can be obtained by selecting the vectors that reside in the pivot columns of $[v_1|v_2|\cdots|v_k]$.

Proof: this is immediately obvious from Proposition 4.4.1. \Box

The proposition that follows is also follows immediately from Proposition 4.4.1.

Proposition 6.5.17.

Let $A \in \mathbb{R}^{m \times n}$ the pivot columns of A form a basis for Col(A).

Example 6.5.18. Suppose A is given as below: (I omit the details of the Gaussian elimination)

	1	2	3	4]			[1]	0	5/3	0	1
A =	2	1	4	1		\Rightarrow	rref[A] =	0	1	2/3	0	.
	0	0	0	3				0	0	0	1.	

Identify that columns 1,2 and 4 are pivot columns. Moreover,

$$Col(A) = span\{col_1(A), col_2(A), col_4(A)\}$$

In particular we can also read how the second column is a linear combination of the basis vectors.

$$col_3(A) = \frac{5}{3}col_1(A) + \frac{2}{3}col_2(A)$$

= $\frac{5}{3}[1, 2, 0]^T + \frac{2}{3}[2, 1, 0]^T$
= $[5/3, 10/3, 0]^T + [4/3, 2/3, 0]^T$
= $[3, 4, 0]^T$

What if we want a basis for Row(A) which consists of rows in A itself?

Proposition 6.5.19.

Let $W = span\{w_1, w_2, \ldots, w_k\} \subset \mathbb{R}^{1 \times n}$ and construct A by concatenating the row vectors in W into a matrix A:

$$A = \begin{bmatrix} \frac{w_1}{w_2} \\ \vdots \\ w_k \end{bmatrix}$$

A basis for W is given by the transposes of the pivot columns for A^T .

Proof: this is immediately obvious from Proposition 4.4.5. \Box

The proposition that follows is also follows immediately from Proposition 4.4.5.

Proposition 6.5.20.

Let $A \in \mathbb{R}^{m \times n}$ the rows which are transposes of the pivot columns of A^T form a basis for Row(A).

Example 6.5.21.

$$A^{T} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 0 \\ 4 & 1 & 3 \end{bmatrix} \implies rref[A^{T}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Notice that each column is a pivot column in A^T thus a basis for Row(A) is simply the set of all rows of A; $Row(A) = span\{[1,2,3,4], [2,1,4,1], [0,0,1,0]\}$ and the spanning set is linearly independent.

Example 6.5.22.

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 4 & 0 \\ 5 & 6 & 2 \end{bmatrix} \qquad \Rightarrow \qquad A^T = \begin{bmatrix} 1 & 2 & 3 & 5 \\ 1 & 2 & 4 & 6 \\ 1 & 2 & 0 & 2 \end{bmatrix} \qquad \Rightarrow \qquad rref[A^T] = \begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We deduce that rows 1 and 3 or A form a basis for Row(A). Notice that $row_2(A) = 2row_1(A)$ and $row_4(A) = row_3(A) + 2row_1(A)$. We can read linear dependendcies of the rows from the corresponding linear dependencies of the columns in the rref of the transpose.

The preceding examples are nice, but what should we do if we want to find both a basis for Col(A) and Row(A) for some given matrix? Let's pause to think about how elementary row operations modify the row and column space of a matrix. In particular, let A be a matrix and let A' be the result of performing an elementary row operation on A. It is fairly obvious that

$$Row(A) = Row(A').$$

Think about it. If we swap to rows that just switches the order of the vectors in the span that makes Row(A). On the other hand if we replace one row with a nontrivial linear combination of itself and other rows then that will not change the span either. Column space is not so easy though. Notice that elementary row operations can change the column space. For example,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \implies rref[A] = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

has $Col(A) = span\{[1,1]^T\}$ whereas $Col(rref(A)) = span([1,0]^T)$. We cannot hope to use columns of ref(A) (or rref(A)) for a basis of Col(A). That's no big problem though because we already have the CCP-principle which helped us pick out a basis for Col(A). Let's collect our thoughts:

Proposition 6.5.23.

Let $A \in \mathbb{R}^{m \times n}$ then a basis for Col(A) is given by the pivot columns in A and a basis for Row(A) is given by the nonzero rows in ref(A).

This means we can find a basis for Col(A) and Row(A) by performing the forward pass on A. We need only calculate the ref(A) as the pivot columns are manifest at the end of the forward pass.

Example 6.5.24.

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \xrightarrow{r_2 - r_1 \to r_2} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 2 \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} = ref[A]$$

We deduce that $\{[1,1,1], [0,1,2]\}$ is a basis for Row(A) whereas $\{[1,1,1]^T, [1,1,2]^T\}$ is a basis for Col(A). Notice that if I wanted to reveal further linear dependencies of the non-pivot columns on the pivot columns of A it would be wise to calculate rref[A] by making the backwards pass on ref[A].

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{r_1 - r_2 \to r_1} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} = rref[A]$$

From which I can read $col_3(A) = 2col_2(A) - col_1(A)$, a fact which is easy to verify.

Example 6.5.25.

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 8 & 10 \\ 1 & 2 & 4 & 11 \end{bmatrix} \xrightarrow{r_2 - r_1 \to r_2} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 5 & 6 \\ 0 & 0 & 1 & 7 \end{bmatrix} = ref[A]$$

We find that Row(A) has basis

$$\{[1, 2, 3, 4], [0, 1, 5, 6], [0, 0, 1, 7]\}$$

and Col(A) has basis

$$\left\{ \begin{bmatrix} 1\\1\\1 \end{bmatrix}, \begin{bmatrix} 2\\3\\2 \end{bmatrix}, \begin{bmatrix} 3\\8\\4 \end{bmatrix} \right\}$$

Proposition 6.5.23 was the guide for both examples above.

6.5.2 calculating basis of a solution set

Often a subspace is described as the solution set of some equation Ax = 0. How do we find a basis for Null(A)? If we can do that we find a basis for subspaces which are described by some equation.

Proposition 6.5.26.

Let $A \in \mathbb{R}^{m \times n}$ and define W = Null(A). A basis for W is obtained from the solution set of Ax = 0 by writing the solution as a linear combination where the free variables appear as coefficients in the vector-sum.

Proof: $x \in W$ implies Ax = 0. Denote $x = [x_1, x_2, \ldots, x_n]^T$. Suppose that rref[A] has r-pivot columns (we must have $0 \le r \le n$). There will be (m - r)-rows which are zero in rref(A) and (n - r)-columns which are not pivot columns. The non-pivot columns correspond to free-variables in the solution. Define p = n - r for convenience. Suppose that $x_{i_1}, x_{i_2}, \ldots, x_{i_p}$ are free whereas $x_{j_1}, x_{j_2}, \ldots, x_{j_r}$ are functions of the free variables: in particular they are linear combinations of the free variables as prescribed by rref[A]. There exist constants b_{ij} such that

$$\begin{aligned} x_{j_1} &= b_{11} x_{i_1} + b_{12} x_{i_2} + \dots + b_{1p} x_{i_p} \\ x_{j_2} &= b_{21} x_{i_1} + b_{22} x_{i_2} + \dots + b_{2p} x_{i_p} \\ \vdots &\vdots & \dots & \vdots \\ x_{j_r} &= b_{r1} x_{i_1} + b_{r2} x_{i_2} + \dots + b_{rp} x_{i_p} \end{aligned}$$

For convenience of notation assume that the free variables are put at the end of the list. We have

$$\begin{aligned} x_1 &= b_{11}x_{r+1} + b_{12}x_{r+2} + \dots + b_{1p}x_n \\ x_2 &= b_{21}x_{r+1} + b_{22}x_{r+2} + \dots + b_{2p}x_n \\ \vdots &\vdots & \dots & \vdots \\ x_r &= b_{r1}x_{r+1} + b_{r2}x_{n-p+2} + \dots + b_{rp}x_n \end{aligned}$$

and $x_j = x_j$ for j = r + 1, r + 2, ..., r + p = n (those are free, we have no conditions on them, they can take any value). We find,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{bmatrix} = x_{r+1} \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{r1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_{r+2} \begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{r2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + x_n \begin{bmatrix} b_{1p} \\ b_{2p} \\ \vdots \\ b_{rp} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

We define the vectors in the sum above as v_1, v_2, \ldots, v_p . If any of the vectors, say v_j , was linearly dependent on the others then we would find that the variable x_{r+j} was likewise dependent on the other free variables. This would contradict the fact that the variable x_{r+j} was free. Consequently the vectors v_1, v_2, \ldots, v_p are linearly independent. Moreover, they span the null-space by virtue of their construction. \Box

Didn't follow the proof above? No problem. See the examples to follow here. These are just the proof in action for specific cases. We've done these sort of calculations in §1.3. We're just adding a little more insight here.

Example 6.5.27. Find a basis for the null space of A = [1, 2, 3, 4]. This example requires no additional calculation except this; Ax = 0 for $x = (x_1, x_2, x_3, x_4)$ yields $x_1 = -2x_2 - 3x_3 - 4x_4$ thus:

$$x = \begin{bmatrix} -2x_2 - 3x_3 - 4x_4 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -4 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Thus, $\{(-2, 1, 0, 0), (-3, 0, 1, 0), (-4, 0, 0, 1)\}$ forms a basis for Null(A).

Example 6.5.28. Find a basis for the null space of A given below,

$$A = \left[\begin{array}{rrrr} 1 & 0 & 0 & 1 & 0 \\ 2 & 2 & 0 & 0 & 1 \\ 4 & 4 & 4 & 0 & 0 \end{array} \right]$$

Gaussian elimination on the augmented coefficient matrix reveals (see Example 2.2.5 for details of the Gaussian elimination)

$$rref\left[\begin{array}{rrr}1 & 0 & 0 & 1 & 0\\2 & 2 & 0 & 0 & 1\\4 & 4 & 4 & 0 & 0\end{array}\right] = \left[\begin{array}{rrr}1 & 0 & 0 & 1 & 0\\0 & 1 & 0 & -1 & 1/2\\0 & 0 & 1 & 0 & -1/2\end{array}\right]$$

Denote $x = [x_1, x_2, x_3, x_4, x_5]^T$ in the equation Ax = 0 and identify from the calculation above that x_4 and x_5 are free thus solutions are of the form

$$x_{1} = -x_{4}$$

$$x_{2} = x_{4} - \frac{1}{2}x_{5}$$

$$x_{3} = \frac{1}{2}x_{5}$$

$$x_{4} = x_{4}$$

$$x_{5} = x_{5}$$

for all $x_4, x_5 \in \mathbb{R}$. We can write these results in vector form to reveal the basis for Null(A),

$$x = \begin{bmatrix} -x_4 \\ x_4 - \frac{1}{2}x_5 \\ \frac{1}{2}x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_4 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 0 \\ -\frac{1}{2} \\ \frac{1}{2} \\ 0 \\ 1 \end{bmatrix}$$

It follows that the basis for Null(A) is simply

$$\left\{ \begin{bmatrix} -1\\1\\0\\1\\0 \end{bmatrix}, \begin{bmatrix} 0\\-\frac{1}{2}\\\frac{1}{2}\\0\\1 \end{bmatrix} \right\}$$

Of course, you could multiply the second vector by 2 if you wish to avoid fractions. In fact there is a great deal of freedom in choosing a basis. We simply show one way to do it. **Example 6.5.29.** Find a basis for the null space of A given below,

Gaussian elimination on the augmented coefficient matrix reveals:

Denote $x = [x_1, x_2, x_3, x_4]^T$ in the equation Ax = 0 and identify from the calculation above that x_2, x_3 and x_4 are free thus solutions are of the form

$$\begin{aligned} x_1 &= -x_2 - x_3 - x_4 \\ x_2 &= x_2 \\ x_3 &= x_3 \\ x_4 &= x_4 \end{aligned}$$

for all $x_2, x_3, x_4 \in \mathbb{R}$. We can write these results in vector form to reveal the basis for Null(A),

$$x = \begin{bmatrix} -x_2 - x_3 - x_4 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

It follows that the basis for Null(A) is simply

ſ	$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$		$\begin{bmatrix} -1\\ 0 \end{bmatrix}$		$\begin{bmatrix} -1\\ 0 \end{bmatrix}$	٦
	0 0	,	$\begin{bmatrix} 1\\ 0 \end{bmatrix}$,	$\begin{array}{c} 0 \\ 1 \end{array}$	Ĵ

6.6 theory of dimensions

We prove a number of theorems in the section which show that dimension is a well-defined quantity for a finite dimensional vector space. Up to this point we have only used the phrase "finitedimensional" to mean that there exists one basis with finitely many vectors. In this section we prove that if that is the case then all other bases for the vector space must likewise have the same number of basis vectors. In addition we give several existence theorems which are of great theoretical importance. Finally, we discuss dimensions of column, row and null space of a matrix.

The proposition that follows is the baby version of Proposition 6.6.5. I include this proposition in the notes because the proof is fun.

Proposition 6.6.1.

Let V be a finite-dimensional vector space and suppose $B = \{b_1, b_2, \dots, b_n\}$ is any basis of V,

- 1. $B \cup \{v\}$ is linearly dependent
- 2. for any $1 \le k \le n$, $B \{b_k\}$ does not span V

Proof of (1.): Since B spans V it follows that v is a linear combination of vectors in B thus $B \cup \{v\}$ is linearly dependent.

Proof of (2.): We argue that $b_k \notin span(B - \{b_k\})$. Argue by contradiction. Suppose that $b_k \in span(B - \{b_k\})$ then there exist constants $c_1, c_2, \ldots, \widehat{c_k}, c_n$ such that

$$b_k = c_1 b_1 + c_2 b_2 + \dots + \widehat{c_k b_k} + \dots + c_n b_n$$

but this contradicts the linear independence of the basis as

 $c_1b_1 + c_2b_2 + \dots - b_k + \dots + c_nb_n = 0$

does not imply all the coefficients are zero. Therefore, using proof by contradiction, $span(B - \{b_k\}) \neq V$. \Box

Proposition 6.6.2.

Let V be a finite-dimensional vector space and suppose $B = \{b_1, b_2, \ldots, b_n\}$ is any basis of V then any other basis for V also has n-elements.

Proof: Suppose $B = \{b_1, b_2, \ldots, b_n\}$ and $F = \{f_1, f_2, \ldots, f_p\}$ are both bases for a vector space V. Since F is a basis it follows $b_k \in span(F)$ for all k so there exist constants c_{ik} such that

$$b_k = c_{1k}f_1 + c_{2k}f_2 + \dots + c_{pk}f_p$$

for k = 1, 2, ..., n. Likewise, since $f_j \in span(B)$ there exist constants d_{lj} such that

$$f_j = d_{1j}b_1 + d_{2j}b_2 + \dots + d_{nj}b_n$$

for $j = 1, 2, \ldots, p$. Substituting we find

$$f_j = d_{1j}b_1 + d_{2j}b_2 + \dots + d_{nj}b_n$$

= $d_{1j}(c_{11}f_1 + c_{21}f_2 + \dots + c_{p1}f_p) + d_{2j}(c_{12}f_1 + c_{22}f_2 + \dots + c_{p2}f_p) + d_{2j}(c_{12}f_1 + c_{2j}f_2 + \dots + c_{p2}f_p) + d_{2j}(c_{12}f_1 + \dots + c_{p2}f_p) + d_{2$

$$+\cdots + d_{nj}(c_{1n}f_1 + c_{2n}f_2 + \cdots + c_{pn}f_p)$$

$$= (d_{1j}c_{11} + d_{2j}c_{12} + \cdots + d_{nj}c_{1n})f_1$$

$$(d_{1j}c_{21} + d_{2j}c_{22} + \cdots + d_{nj}c_{2n})f_2 +$$

$$+ \cdots + (d_{1j}c_{p1} + d_{2j}c_{p2} + \cdots + d_{nj}c_{pn})f_p$$

Suppose j = 1. We deduce, by the linear independence of F, that

$$d_{11}c_{11} + d_{21}c_{12} + \cdots + d_{n1}c_{1n} = 1$$

from comparing coefficients of f_1 , whereas for f_2 we find,

$$d_{11}c_{21} + d_{21}c_{22} + \cdots + d_{n1}c_{2n} = 0$$

likewise, for f_q with $q \neq 1$,

$$d_{11}c_{q1} + d_{21}c_{q2} + \cdots + d_{n1}c_{qn} = 0$$

Notice we can rewrite all of these as

$$\delta_{q1} = c_{q1}d_{11} + c_{q2}d_{21} + \dots + c_{qn}d_{n1}$$

Similarly, for arbitrary j we'll find

$$\delta_{qj} = c_{q1}d_{1j} + c_{q2}d_{2j} + \cdots + c_{qn}d_{nj}$$

If we define $C = [c_{ij}] \in \mathbb{R}^{p \times n}$ and $D = [d_{ij}] \in \mathbb{R}^{n \times p}$ then we can translate the equation above into the matrix equation that follows:

 $CD = I_p$.

We can just as well argue that

 $DC = I_n$

The **trace** of a matrix is the sum of the diagonal entries in the matrix; trace $(A) = \sum_{i=1}^{n} A_{ii}$ for $A \in \mathbb{R}^{n \times n}$. It is not difficult to show that trace(AB) = trace(BA) provided the products AB and BA are both defined. Moreover, it is also easily seen $tr(I_p) = p$ and $tr(I_q) = q$. It follows that,

 $tr(CD) = tr(DC) \Rightarrow tr(I_p) = tr(I_q) \Rightarrow p = q.$

Since the bases were arbitrary this proves any pair have the same number of vectors. \Box

Given the preceding proposition the following definition is logical.

Definition 6.6.3.

If V is a finite-dimensional vector space then the **dimension** of V is the number of vectors in any basis of V and it is denoted dim(V).

Example 6.6.4. Let me state the dimensions which follow from the standard bases of \mathbb{R}^n and $\mathbb{R}^{m \times n}$ respective,

$$dim(\mathbb{R}^n) = n$$
 $dim(\mathbb{R}^{m \times n}) = mn$

these results follow from counting.

Proposition 6.6.5.

Suppose V is a vector space with dim(V) = n.

1. If S is a set with more than n vectors then S is linearly dependent.

2. If S is a set with less than n vectors then S does not generate V.

Proof of (1.): Suppose $S = \{s_1, s_2, \ldots, s_m\}$ has *m* vectors and m > n. Let $B = \{b_1, b_2, \ldots, b_n\}$ be a basis of *V*. Consider the corresponding set of coordinate vectors of the vectors in *S*, we denote

$$[S]_B = \{[s_1]_B, [s_2]_B, \dots, [s_m]_B\}$$

The set $[S]_B$ has m vectors in \mathbb{R}^n and m > n therefore by Proposition 4.3.6 we know $[S]_B$ is a linearly dependent set. Therefore at least one, say $[s_j]_B$, vector can be written as a linear combination of the other vectors in $[S]_B$ thus there exist constants c_i with (this is a vector equation)

$$[s_j]_B = c_1[s_1]_B + c_2[s_2]_B + \dots + \widehat{c_j[s_j]_B} + \dots + c_m[s_m]_B$$

Also notice that (introducing a new shorthand $B[s_j]$ which is not technically matrix multiplication since b_i are not column vectors generally, they could be chickens for all we know)

$$s_j = B[s_j] = s_{j1}b_1 + s_{j2}b_2 + \dots + s_{jn}b_n$$

We also know, using the notation $([s_j]_B)_k = s_{jk}$,

$$s_{jk} = c_1 s_{1k} + c_2 s_{2k} + \dots + \widehat{c_j s_{jk}} + \dots + c_m s_{mk}$$

for k = 1, 2, ..., n. Plug these into our s_i equation,

$$s_{j} = (c_{1}s_{11} + c_{2}s_{21} + \dots + \widehat{c_{j}s_{j1}} + \dots + c_{m}s_{m1})b_{1} + (c_{1}s_{12} + c_{2}s_{22} + \dots + \widehat{c_{j}s_{j2}} + \dots + c_{m}s_{m2})b_{2} + \dots + (c_{1}s_{1n} + c_{2}s_{2n} + \dots + \widehat{c_{j}s_{jn}} + \dots + c_{m}s_{mn})b_{n}$$
$$= c_{1}(s_{11}b_{1} + s_{12}b_{2} + \dots + s_{1n}b_{n}) + (c_{1}s_{1n} + c_{2}s_{2n} + \dots + c_{n}s_{n}) + (c_{1}s_{1n} + c_{2}s_{n}) + (c_{1}s_{1n} + c_{2$$

 $c_2(s_{21}b_1 + s_{22}b_2 + \dots + s_{2n}b_n) + \\ + \dots + c_m(s_{m1}b_1 + s_{m2}b_2 + \dots + s_{mn}b_n) : \text{ excluding } c_j(\dots)$

$$= c_1 s_1 + c_2 s_2 + \dots + \widehat{c_j s_j} + \dots + c_n s_n$$

Well this is a very nice result, the same linear combination transfers over to the abstract vectors. Clearly s_j linearly depends on the other vectors in S so S is linearly dependent. The heart of the proof was Proposition 4.3.6 and the rest was just battling notation.

Proof of (2.): Use the corresponding result for \mathbb{R}^n which was given by Proposition 4.3.5. Given m abstract vectors if we concantenate their coordinate vectors we will find a matrix [S] in $\mathbb{R}^{n \times m}$ with m < n and as such there will be some choice of the vector b for which $[S]x \neq b$. The abstract vector corresponding to b will not be covered by the span of S. \Box

Proposition 6.6.6.

Suppose V is a vector space with dim(V) = n and $W \leq V$ then there exists a basis for W and $dim(W) \leq n$.

Proof: If $W = \{0\}$ then the proposition is true. Suppose $W \neq 0$ and set S be a finite subset of W. Apply Proposition 6.6.5 to modify S to a basis β_W for W by possibly deleting or adjoining vectors from W. Again, apply Proposition 6.6.5 to see $\#(\beta_W) \leq n$ and this completes the proof. \Box .

The Proposition above is almost an immediate consquence of other theorems and propositions in these notes, I included it just for the sake of later reference. Anton calls the following proposition the "Plus/Minus" Theorem.

Proposition 6.6.7.

Let V be a vector space and suppose S is a nonempty set of vectors in V.

- 1. If S is linearly independent a nonzero vector $v \notin span(S)$ then $S \cup \{v\}$ is a linearly independent set.
- 2. If $v \in S$ is a linear combination of other vectors in S then $span(S \{v\}) = span(S)$.
Proof of (1.): Suppose $S = \{s_1, s_2, \ldots, s_k\}$ and consider,

$$c_1s_1 + c_2s_2 + \dots + c_ks_k + c_{k+1}v = 0$$

If $c_{k+1} \neq 0$ it follows that v is a linear combination of vectors in S but this is impossible so $c_{k+1} = 0$. Then since S is linear independent

 $c_1s_1 + c_2s_2 + \dots + c_ks_k = 0 \implies c_1 = c_2 = \dots = c_k = 0$

thus $S \cup \{v\}$ is linearly independent.

Proof of (2.): Suppose $v = s_i$. We are given that there exist constants d_i such that

$$s_j = d_1 s_1 + \dots + \widehat{d_j s_j} + \dots + d_k s_k$$

Let $w \in span(S)$ so there exist constants c_i such that

$$w = c_1 s_1 + c_2 s_2 + \dots + c_j s_j + \dots + c_k s_k$$

Now we can substitute the linear combination with d_i -coefficients for s_i ,

$$w = c_1 s_1 + c_2 s_2 + \dots + c_j (d_1 s_1 + \dots + \widehat{d_j s_j} + \dots + d_k s_k) + \dots + c_k s_k$$
$$= (c_1 + c_j d_1) s_1 + (c_2 + c_j d_2) s_2 + \dots + \widehat{c_j d_j s_j} + \dots + (c_k + c_j d_k) s_k$$

thus w is a linear combination of vectors in S, but not $v = s_j$, thus $w \in span(S - \{v\})$ and we find $span(S) \subseteq span(S - \{v\})$.

Next, suppose $y \in span(S - \{v\})$ then y is a linear combination of vectors in $S - \{v\}$ hence y is a linear combination of vectors in S and we find $y \in span(S)$ so $span(S - \{v\}) \subseteq span(S)$. (this inclusion is generally true even if v is linearly independent from other vectors in S). We conclude that $span(S) = span(S - \{v\})$. \Box

Proposition 6.6.8.

Let V be an n-dimensional vector space. A set S with n-vectors is a basis for V if S is either linearly independent or if span(S) = V.

Proof: Assume S has n-vectors which are linearly independent in a vector space V with dimension n. Suppose towards a contradiction that S does not span V. Then there exists $v \in V$ such that $v \notin span(S)$. Then by Proposition 6.6.7 we find $V \cup \{v\}$ is linearly independent. But, Proposition 6.6.5 the set $V \cup \{v\}$ is linearly dependent. This is a contradiction, thus S spans V and we find D is a basis.

Assume S has n-vectors which span a vector space V with dimension n. Suppose towards a contradiction that S is not linearly independent V. This means there exists $v \in S$ which is a linear combination of other vectors in S. Therefore, by 6.6.5, S does not span V. This is a contradicts the assumption span(S) = V therefore S is linearly independent and consequently S is a basis. \Box

Remark 6.6.9.

Intuitively speaking, linear independence is like injectivity for functions whereas spanning is like the onto property for functions. Suppose A is a finite set. If a function $f: A \to A$ is 1-1 then it is onto. Also if the function is onto then it is 1-1. The finiteness of A is what blurs the concepts. For a vector space, we also have a sort of finiteness in play if dim(V) = n. When a set with dim(V)-vectors spans (like onto) V then it is automatically linearly independent. When a set with dim(V)-vectors is linearly independent (like 1-1) V then it automatically spans V. However, in an infinite dimensional vector space this need not be the case. For example, d/dx is a surjective linear mapping on $\mathbb{R}[x] = span\{1, x, x^2, x^3, ...\}$ however if $f, g \in \mathbb{R}[x]$ and df/dx = dg/dx we can only conclude that f = g + c thus d/dx is not injective on vector space of polynomials in x. Many theorems we discuss do hold in the infinite dimensional context, but you have to be careful.

Theorem 6.6.10.

Let S be a subset of a finite dimensional vector space V.

- 1. If span(S) = V but S is not a basis then S can be modified to make a basis by removing redundant vectors.
- 2. If S is linearly independent and $span(S) \neq V$ then S can be modified to make a basis for V by unioning vectors outside span(S).

Proof of (1.): If span(S) = V but S is not a basis we find S is linearly dependent. (if S is linearly independent then Proposition 6.6.8 says S is a basis which is a contradiction). Since S is linearly dependent we can write some $v \in S$ as a linear combination of other vectors in S. Furthermore, by Proposition 6.6.5 $span(S) = span(S - \{v\})$. If $S - \{v\}$ is linearly independent then $S - \{v\}$ is a basis. Otherwise $S - \{v\}$ is linearly dependent and we can remove another vector. Continue until the resulting set is linearly independent (we know this happens when there are just dim(V)-vectors in the set so this is not an endless loop)

Proof of (2.): If S is linearly independent but $span(S) \neq V$ then there exists $v \in V$ but $v \notin span(S)$. Proposition 6.6.7 shows that $S \cup \{v\}$ is linearly independent. If $span(S \cup \{v\}) = V$ then $S \cup \{v\}$ is a basis. Otherwise there is still some vector outside $span(S \cup \{v\}) = V$ and we can repeat the argument for that vector and so forth until we generate a set which spans V. Again we know this is not an endless loop because V is finite dimensional and once the set is linearly independent and contains dim(V) vectors it must be a basis (see Proposition 6.6.8). \Box

Remark 6.6.11.

We already saw in the previous sections that we can implement part (1.) of the preceding proposition in \mathbb{R}^n and $\mathbb{R}^{1 \times n}$ through matrix calculations. There are really nice results about row and column spaces which show us precisely which vectors we need to remove or add to obtain a basis. I'll probably ask a homework question which tackels the question in the abstract. Once you understand the \mathbb{R}^n -case you can do the abstract case by lifting the arguments through the coordinate maps. We've already seen this "lifting" idea come into play in several proof of Proposition 6.6.5. Part (2.) involves making a choice. How do you choose a vector outside the span? I leave this question to the reader for the moment.

Proposition 6.6.12.

```
If V is a finite-dimensional vector space and W \leq V then dim(W) \leq dim(V). Moreover, if dim(V) = dim(W) then V = W.
```

Proof: Let β be a basis for W, if β is also a basis for V then dim(V) = dim(W) and $V = W = span(\beta)$. Otherwise, if $span(\beta) \neq V$, apply Theorem 6.6.10 to extend β to γ a basis for V. \Box

6.6.1 application to fundamental matrix subspaces

These were defined before, I restate them here along with their dimensions for convenience.

Definition 6.6.13.

Let $A \in \mathbb{R}^{m \times n}$. We define 1. $Col(A) = span\{col_j(A)|j = 1, 2, ..., n\}$ and r = rank(A) = dim(Col(A))2. $Row(A) = span\{row_i(A)|i = 1, 2, ..., m\}$ 3. $Null(A) = \{x \in \mathbb{R}^n | Ax = 0\}$ and $\nu = nullity(A) = dim(Null(A))$

Proposition 6.6.14.

Let $A \in \mathbb{R}^{m \times n}$ then dim(Row(A)) = dim(Col(A))

Proof: By Proposition 6.5.17 we know the number of vectors in the basis for Col(A) is the number of pivot columns in A. Likewise, Proposition 6.5.23 showed the number of vectors in the basis for Row(A) was the number of nonzero rows in ref(A). But the number of pivot columns is precisely the number of nonzero rows in ref(A) therefore, dim(Row(A)) = dim(Col(A)). \Box

Theorem 6.6.15.

Let $A \in \mathbb{R}^{m \times n}$ then n = rank(A) + nullity(A).

Proof: The proof of Proposition 6.5.26 makes is clear that if a $m \times n$ matrix A has r-pivot columns then there will be n - r vectors in the basis of Null(A). It follows that

$$rank(A) + nullity(A) = r + (n - r) = n.$$

6.7 general theory of linear systems

We've seen some rather abstract results thus far in this chapter. I thought it might be helpful to tie them back to our fundamental problem; how does dimension theory help us understand the structure of solutions to Ax = b? Let $A \in \mathbb{R}^{m \times n}$ we should notice that $Null(A) \leq \mathbb{R}^n$ is only possible since homogeneous systems of the form Ax = 0 have the nice property that linear combinations of solutions is again a solution:

Proposition 6.7.1.

Let Ax = 0 denote a homogeneous linear system of *m*-equations and *n*-unknowns. If v_1 and v_2 are solutions then any linear combination $c_1v_1 + c_2v_2$ is also a solution of Ax = 0.

Proof: Suppose $Av_1 = 0$ and $Av_2 = 0$. Let $c_1, c_2 \in \mathbb{R}$ and recall Theorem 3.2.17,

$$A(c_1v_1 + c_2v_2) = c_1Av_1 + c_2Av_2 = c_10 + c_20 = 0.$$

Therefore, $c_1v_1 + c_2v_2 \in Sol_{[A|0]}$. \Box

We proved this before, but I thought it might help to see it again.

Proposition 6.7.2.

Let $A \in \mathbb{R}^{m \times n}$. If v_1, v_2, \ldots, v_k are solutions of Av = 0 then $V = [v_1|v_2|\cdots|v_k]$ is a solution matrix of Av = 0 (V a solution matrix of Av = 0 iff AV = 0)

Proof: Let $A \in \mathbb{R}^{m \times n}$ and suppose $Av_i = 0$ for i = 1, 2, ..., k. Let $V = [v_1|v_2|\cdots|v_k]$ and use the solution concatenation Proposition 3.7.1,

$$AV = A[v_1|v_2|\cdots|v_k] = [Av_1|Av_2|\cdots|Av_k] = [0|0|\cdots|0] = 0.$$

A solution matrix of a linear system is a matrix in which each column is itself a solution.

Proposition 6.7.3.

Let $A \in \mathbb{R}^{m \times n}$. The system of equations Ax = b is consistent iff $b \in Col(A)$.

Proof: Observe,

$$Ax = b \iff \sum_{i,j} A_{ij} x_j e_i = b$$

$$\Leftrightarrow \sum_j x_j \sum_i A_{ij} e_i = b$$

$$\Leftrightarrow \sum_j x_j \operatorname{col}_j(A) = b$$

$$\Leftrightarrow b \in \operatorname{Col}(A)$$

Therefore, the existence of a solution to Ax = b is interchangeable with the statement $b \in Col(A)$. They both amount to saying that b is a linear combination of columns of A. \Box

Proposition 6.7.4.

Let $A \in \mathbb{R}^{m \times n}$ and suppose the system of equations Ax = b is consistent. We find $x \in \mathbb{R}^n$ is a solution of the system if and only if it can be written in the form

$$x = x_h + x_p = c_1 v_1 + c_2 v_2 + \dots + c_{\nu} v_{\nu} + x_p$$

where $Ax_h = 0$, $\{v_j\}_{j=1}^{\nu}$ are a basis for Null(A), and $Ax_p = b$. We call x_h the homogeneous solution and x_p is the nonhomogeneous solution.

Proof: Suppose Ax = b is consistent then $b \in Col(A)$ therefore there exists $x_p \in \mathbb{R}^n$ such that $Ax_p = b$. Let x be any solution. We have Ax = b thus observe

$$A(x - x_p) = Ax - Ax_p = Ax - b = 0 \quad \Rightarrow \quad (x - x_p) \in Null(A).$$

Define $x_h = x - x_p$ it follows that there exist constants c_i such that $x_h = c_1v_1 + c_2v_2 + \cdots + c_\nu v_\nu$ since the vectors v_i span the null space.

Conversely, suppose $x = x_p + x_h$ where $x_h = c_1v_1 + c_2v_2 + \cdots + c_\nu v_\nu \in Null(A)$ then it is clear that

$$Ax = A(x_p + x_h) = Ax_p + Ax_h = b + 0 = b$$

thus $x = x_p + x_h$ is a solution. \Box

Example 6.7.5. Consider the system of equations x + y + z = 1, x + z = 1. In matrix notation,

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \Rightarrow rref[A|b] = rref\begin{bmatrix} 1 & 1 & 1 & | & 1 \\ 1 & 0 & 1 & | & 1 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & | & 1 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

It follows that x = 1 - y - z is a solution for any choice of $y, z \in \mathbb{R}$.

$$v = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 - y - z \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

We recognize that $v_p = [1, 0, 0]^T$ while $v_h = y[-1, 1, 0]^T + z[-1, 0, 1]^T$ and $\{[-1, 1, 0]^T, [-1, 0, 1]^T\}$ is a basis for the null space of A. We call y, z parameters in the solution.

We will see that null spaces play a central part in the study of eigenvectors in Part III. In fact, about half of the eigenvector calculation is finding a basis for the null space of a certain matrix. So, don't be too disappointed if I don't have too many examples here. You'll work dozens of them later.

The following proposition simply summarizes what we just calculated:

Proposition 6.7.6.

Let $A \in \mathbb{R}^{m \times n}$. If the system of equations Ax = b is consistent then the general solution has as many parameters as the dim(Null(A)).

6.7.1 linear algebra in DEqns

A very similar story is told in differential equations. In Math 334 we spend some time unraveling the solution of L[y] = g where L = P(D) is an *n*-th order polynomial in the differentiation operator with constant coefficients. In total we learn that $y = c_1y_1 + c_2y_2 + \cdots + c_ny_n + y_p$ is the solution where y_j are the homogeneous solutions which satisfy $L[y_j] = 0$ for each $j = 1, 2, \ldots, n$ and, in contrast, y_p is the so-called "particular solution" which satisfies $L[y_p] = g$. On the one hand, the results in DEqns are very different because the solutions are functions which live in the infinitedimensional function space. However, on the other hand, L[y] = g is a finite dimensional problem thanks to the fortunate fact that $Null(L) = \{f \in \mathcal{F}(\mathbb{R}) | L(f) = 0\} = span\{y_1, y_2, \ldots, y_n\}$. For this reason there are *n*-parameters in the general solution which we typically denote by c_1, c_2, \ldots, c_n in the Math 334 course. The particular solution is not found by row reduction on a matrix in DEqns⁹. Instead, we either use the annihilator method, power series techniques, or most generally the method of variation of parameters will calculate y_p . The analogy to the linear system Av = bis striking:

1.
$$Av = b$$
 has solution $v = c_1v_1 + c_2v_2 + \cdots + c_kv_n + v_p$ where $v_i \in Null(A)$ and $Av_p = b$.

2.
$$L[y] = g$$
 has solution $v = c_1y_1 + c_2y_2 + \cdots + c_ky_n + y_p$ where $y_j \in Null(L)$ and $L[y_p] = b$.

The reason the DEqn L[y] = g possesses such an elegant solution stems from the linearity of L. If you study nonlinear DEqns the structure is not so easily described.

⁹ok, to be fair you could use coordinate vectors of the next chapter to convert $y_1, y_2, \ldots y_n$ to coordinate vectors and if you worked in a sufficiently large finite dimensional subspace of function space perhaps you could do a row reduction to find g, but this is not the typical calculation.

Chapter 7

abstract linear transformations

We already studied the structure of linear transformations from \mathbb{R}^n to \mathbb{R}^m in Chapter 5. In this chapter we study functions on abstract vector spaces for which the linear structure is preserved. Linear transformations in the abstract enjoy an interesting array of theorems. We'll spend considerable energy in detailing these theorems. Furthermore, the set of linear transformations is found to be a vector space with respect to the natural vector addition on function space. Sets of linear transformations provide interesting new examples of vector spaces which add to the wealth of examples we already saw in the previous chapter.

The theorems on dimension also find further illumination in this chapter. We study **isomor-phisms**. Roughly speaking, two vector spaces which are isomorphic are just the same set with different notation in so far as the vector space structure is concerned. Don't view this sentence as a license to trade column vectors for matrices or functions. We're not there yet. You can do that after this course, once you understand the abuse of language properly. Sort of like how certain musicians can say forbidden words since they have earned the rights through their life experience.

We also study the problem of coordinate change. Since the choice of basis is not unique the problem of comparing different pictures of vectors or transformations for abstract vector spaces requires some effort. We begin by translating our earlier work on coordinate vectors into a mapping-centered notation. Once you understand the notation properly, we can draw pictures to solve problems. This idea of **diagrammatic argument** is an important and valuable technique of modern mathematics. Modern mathematics is less concerned with equations and more concerned with functions and sets.

Finally, we study the quotient space construction. Any linear transformation induces an isomorphism from a particular quotient space formed from the domain of the linear transformation and the range of the map. This result parallels the first isomorphism theorem of group theory and is actually replicated across other *categories* of math¹. The problem of quotients and the problem of direct sum decompositions have interesting connections. We study some of the basics to better understand both subspaces and the structure of linear transformations.

Just a word on notation before we get started. Please learn my notation.

¹in group theory you'll learn that the quotient group can only be formed by a normal subgroup. Every abelian group is normal hence the additive group structure of the vector space makes the quotient here well-defined. More care is required in group theory which faces nonabelian group operations.

7.1 basic terminology

Definition 7.1.1.

Let
$$V, W$$
 be vector spaces. If a mapping $L: V \to W$ satisfies

1. L(x+y) = L(x) + L(y) for all $x, y \in V$; L is additive

2. L(cx) = cL(x) for all $x \in V$ and $c \in \mathbb{R}$; L is homogeneous

then we say L is a linear transformation. The set of all linear transformations from V to W is denoted L(V, W). Also, L(V, V) = L(V) and $L \in L(V)$ is called a linear transformation **on** V.

We already saw many examples for the column-vector case $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$. I'll focus on abstract vector space examples here.

Example 7.1.2. Define $L : \mathbb{R}^{m \times n} \to \mathbb{R}^{n \times m}$ by $L(A) = A^T$. This is clearly a linear transformation since

$$L(cA + B) = (cA + B)^T = cA^T + B^T = cL(A) + L(B)$$

for all $A, B \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}$.

Example 7.1.3. Let V, W be a vector spaces and $L : V \to W$ defined by L(x) = 0 for all $x \in V$. This is a linear transformation known as the **trivial transformation**

$$L(x+y) = 0 = 0 + 0 = L(x) + L(y)$$

and

$$L(cx) = 0 = c0 = cL(x)$$

for all $c \in \mathbb{R}$ and $x, y \in V$.

Example 7.1.4. The identity function on a vector space is also a linear transformation. Let $Id: V \to V$ satisfy L(x) = x for each $x \in V$. Observe that

$$Id(x + cy) = x + cy = x + c(y) = Id(x) + cId(y)$$

for all $x, y \in V$ and $c \in \mathbb{R}$.

Example 7.1.5. Define $L: C^0(\mathbb{R}) \to \mathbb{R}$ by $L(f) = \int_0^1 f(x) dx$. Notice that L is well-defined since all continuous functions are integrable and the value of a definite integral is a number. Furthermore,

$$L(f+cg) = \int_0^1 (f+cg)(x)dx = \int_0^1 \left[f(x) + cg(x) \right] dx = \int_0^1 f(x)dx + c \int_0^1 g(x)dx = L(f) + cL(g)$$

for all $f, g \in C^0(\mathbb{R}(and \ c \in \mathbb{R}))$. The definite integral is a linear transformation.

Example 7.1.6. Let $L: C^1(\mathbb{R}) \to C^0(\mathbb{R})$ be defined by L(f)(x) = f'(x) for each $f \in P_3$. We know from calculus that

$$L(f+g)(x) = (f+g)'(x) = f'(x) + g'(x) = L(f)(x) + L(g)(x)$$

and

$$L(cf)(x) = (cf)'(x) = cf'(x) = cL(f)(x).$$

The equations above hold for all $x \in \mathbb{R}$ thus we find function equations L(f+g) = L(f) + L(g) and L(cf) = cL(f) for all $f, g \in C^1(\mathbb{R})$ and $c \in \mathbb{R}$.

Example 7.1.7. Let $a \in \mathbb{R}$. The evaluation mapping $\phi_a : \mathcal{F}(\mathbb{R}) \to \mathbb{R}$ is defined by $\phi_a(f) = f(a)$. This is a linear transformation as (f + cg)(a) = f(a) + cg(a) by definition of function addition and scalar multiplication.

7.2 theory of linear transformations

Let us begin by pointing out two important facts which follow without much work from additivity and homgeneity. We assume V, W are vector spaces over \mathbb{R} in the remainder of this section.

Proposition 7.2.1.

Let $L: V \to W$ be a linear transformation, 1. L(0) = 02. $L(c_1v_1 + c_2v_2 + \cdots + c_nv_n) = c_1L(v_1) + c_2L(v_2) + \cdots + c_nL(v_n)$ for all $v_i \in V$ and $c_i \in \mathbb{R}$.

Proof: to prove of (1.) let $x \in V$ and notice that x - x = 0 thus

$$L(0) = L(x - x) = L(x) + L(-1x) = L(x) - L(x) = 0.$$

To prove (2.) we use induction on n. Notice the proposition is true for n=1,2 by definition of linear transformation. Assume inductively $L(c_1v_1+c_2v_2+\cdots c_nv_n) = c_1L(v_1)+c_2L(v_2)+\cdots+c_nL(v_n)$ for all $v_i \in V$ and $c_i \in \mathbb{R}$ where $i = 1, 2, \ldots, n$. Let $v_1, v_2, \ldots, v_n, v_{n+1} \in V$ and $c_1, c_2, \ldots, c_n, c_{n+1} \in \mathbb{R}$ and consider, $L(c_1v_1+c_2v_2+\cdots c_nv_n+c_{n+1}v_{n+1}) =$

$$= L(c_1v_1 + c_2v_2 + \dots + c_nv_n) + c_{n+1}L(v_{n+1})$$
 by linearity of L
= $c_1L(v_1) + c_2L(v_2) + \dots + c_nL(v_n) + c_{n+1}L(v_{n+1})$ by the induction hypothesis

Hence the proposition is true for n+1 and we conclude by the principle of mathematical induction that (2.) is true for all $n \in \mathbb{N}$. \Box

Proposition 7.2.2.

Let $L: V \to W$ be a linear transformation . If S is linearly dependent then L(S) is linearly dependent.

Proof: Suppose there exists $c_1, \ldots, c_k \in \mathbb{R}$ for which $v = \sum_{i=1}^k c_i v_i$ is a linear dependence in S. Calculate,

$$L(v) = L\left(\sum_{i=1}^{k} c_i v_i\right) = \sum_{i=1}^{k} c_i L(v_i)$$

which, noting $L(v), L(v_i) \in L(S)$ for all $i \in \mathbb{N}_k$, shows L(S) has a linear dependence. Therefore, L(S) is linearly dependent. \Box

This is very similar to Theorem 5.2.8. Actually, the proof is identical modulo the replacement of \mathbb{R}^n with V and \mathbb{R}^m with W.

Theorem 7.2.3. linear map is injective iff only zero maps to zero.

 $L: V \to W$ is an injective linear transformation iff the only solution to the equation L(x) = 0 is x = 0.

Proof: this is a biconditional statement. I'll prove the converse direction to begin. (\Leftarrow) Suppose L(x) = 0 iff x = 0 to begin. Let $a, b \in V$ and suppose L(a) = L(b). By linearity we have L(a - b) = L(a) - L(b) = 0 hence a - b = 0 therefore a = b and we find L is injective.

(⇒) Suppose L is injective. Suppose L(x) = 0. Note L(0) = 0 by linearity of L but then by 1-1 property we have L(x) = L(0) implies x = 0 hence the unique solution of L(x) = 0 is the zero solution. \Box

The image of a subspace and the inverse image of a subspace are once again subspaces. Well, to be precise, I'm assuming the function in question is a linear transformation. It is certainly not true for arbitrary functions. In general, a nonlinear function takes linear spaces and twists them into all sorts of nonlinear shapes. For example, $f(x) = (x, x^2)$ takes the line \mathbb{R} and pastes it onto the parabola $y = x^2$ in the range. We also can observe $f^{-1}\{(0,0)\} = \{0\}$ and yet the mapping is certainly not injective. The theorems we find for linear functions do not usually generalize to functions in general²

Theorem 7.2.4.

If $L: V \to W$ is a linear transformation

1. and $V_o \leq V$ then $L(V_o) \leq W$.

2. and $W_o \leq W$ then $L^{-1}(W_o) \leq V$.

Proof: to prove (1.) suppose $V_o \leq V$. It follows $0 \in V_o$ and hence L(0) = 0 implies $0 \in L(V_o)$. Suppose $x, y \in L(V_o)$ and $c \in \mathbb{R}$. By definition of image, there exist $x_o, y_o \in V_o$ such that $L(x_o) = x$ and $L(y_o) = y$. Consider then, as L is a linear transformation,

$$L(cx_o + y_o) = cL(x_o) + L(y_o)$$
$$= cx + y.$$

Note $cx + y \in V_o$ as $V_o \leq V$. Thus $cx + y \in L(V_o)$ and by the subspace theorem $L(V_o) \leq W$.

To prove (2.) suppose $W_o \leq W$ and observe $0 \in W_o$ and L(0) = 0 implies $0 \in L^{-1}(W_o)$. Hence $L^{-1}(W_o) \neq \emptyset$. Suppose $c \in \mathbb{R}$ and $x, y \in L^{-1}(W_o)$, it follows that there exist $x_o, y_o \in W_o$ such that $L(x) = x_o$ and $L(y) = y_o$. Observe, using linearity of L,

$$L(cx + y) = cL(x) + L(y)$$
$$= cx_o + y_o.$$

Moreover, $cx_o + y_o \in W_o$ as $W_o \leq W$ hence $cx + y \in L^{-1}(W_o)$. Therefore, by the subspace theorem, $L^{-1}(W_o) \leq V$. \Box

²although, perhaps it's worth noting that in advanced calculus we learn how to linearize a function at a point. Some of our results here roughly generalize locally through the linearization and what are known as the inverse and implict function theorems

The range³ and kernel of a linear transformation tell us much about the operation of T.

Definition 7.2.5.

Let V, W be vector spaces. If a mapping $T: V \to W$ is a linear transformation then

1. Ker $(T) = T^{-1}\{0\}$.

2. Range(T) = T(V).

Corollary 7.2.6.

If $T: V \to W$ is a linear transformation then $\operatorname{Range}(T) \leq W$ and $\operatorname{Ker}(T) \leq V$.

Proof: observe $V \leq V$ and $\{0\} \leq W$ hence by Theorem 7.2.4 the Corollary holds true. \Box

For future reference⁴ since the kernel and range are standard subspaces their dimensions have special names:

Definition 7.2.7.

Let V, W be vector spaces. If a mapping $T: V \to W$ is a linear transformation then

1. $\dim(\operatorname{Ker}(T)) = \operatorname{nullity}(T)$

2. $\dim(\operatorname{Range}(T)) = \operatorname{rank}(T)$.

What about LI of sets? If S is a LI subset of V and $T \in L(V, W)$ then is T(S) also LI? The answer is clearly no in general. Consider the trivial transformation of Example 7.1.3. If we require **all** LI sets be mapped to LI sets then it turns out that injectivity is the necessary condition. In fact, this is a continuation of Theorem 7.2.3

Theorem 7.2.8.

Let L: V → W be linear transformation. The following two conditions are equivalent:
1. S subset of V implies L(S) is LI subset of W for all LI subsets S of V.
2. L is injective
3. Ker(L) = {0}

Proof: Theorem 7.2.3 proves (2.) is equivalent to (3.).

Suppose (1.) is true. Let $S \subset \text{Ker}(L)$ such that S is LI. Note S is then a LI subset of V hence L(S) is also LI. But, $L(S) = \{0\}$ as $S \subset \text{Ker}(L)$. This is a contradiction, 0 is not in any LI set. Hence

³sometimes called the "image of T", in fact our definition can be read as "the range of T is the image of V under T", so the term is quite natural

 $^{{}^{4}}$ I use "dim" rather than these terms for pedagogical reasons, but eventually, we should use rank and nullity with meaning

there does not exist a LI subset of Ker(L) and it follows $\text{Ker}(L) = \{0\}$. We have shown $(1.) \Rightarrow (3.)$.

Suppose (2.) is true. Let S be a LI subset of V. Let $L(v_1), \ldots, L(v_k) \in L(S)$ where, by definition of L(S) the vectors $v_1, \ldots, v_k \in S$. Consider,

$$c_1L(v_1) + \cdots + c_kL(v_k) = 0$$

implies by Proposition 7.2.1

$$L(c_1v_1 + \cdots + c_kv_k) = L(0).$$

By injectivity of L we obtain,

$$c_1v_1 + \cdots + c_kv_k = 0$$

and by LI if S we conclude $c_1 = 0, \ldots, c_k = 0$. Thus L(S) is LI. Since S was arbitrary we have shown the implication of (1.) for all LI subset S of V. Therefore, we've shown (2.) \Rightarrow (1.) and the Theorem follows. \Box

Thus far in this section we have studied the behaviour of a particular linear transformation. In what follows, we see how to combine given linear transformations to form new linear transformations. The definition that follows is very similar to Definition 5.3.1

Definition 7.2.9.

Suppose $T: V \to W$ and $S: V \to W$ are linear transformations then we define T + S, T - Sand cT for any $c \in \mathbb{R}$ by the rules (T+S)(x) = T(x) + S(x). (T-S)(x) = T(x) - S(x), (cT)(x) = cT(x)for all $x \in V$.

I'll skip the proof of the proposition below as it is nearly identical to the proof of Proposition 5.3.3.

Proposition 7.2.10.

The sum, difference or scalar multiple of a linear transformations from V to W are once more a linear transformation from V to W.

Recall that function space of all functions from V to W is naturally a vector space according to the point-wise addition and scalar multiplication of functions. It follows from the subspace theorem and the proposition above that:

Proposition 7.2.11.

The set of all linear transformations from V to W forms a vector space with respect to the natural point-wise addition and scalar multiplication of functions; $L(V, W) \leq \mathcal{F}(V, W)$.

Proof: If $T, S \in L(V, W)$ and $c \in \mathbb{R}$ then $T + S, cT \in L(V, W)$ hence L(V, W) is closed under addition and scalar multiplication. Moreover, the trivial function T(x) = 0 for all $x \in V$ is clearly in L(V, W) hence $L(V, W) \neq \emptyset$ and we conclude by the subspace theorem that $L(V, W) \leq \mathcal{F}(V, W)$. \Box

Function composition in the context of abstract vector spaces is the same as it was in precalculus.

Definition 7.2.12.

Suppose $T: V \to U$ and $S: U \to W$ are linear transformations then we define $S \circ T: V \to W$ by $(S \circ T)(x) = S(T(x))$ for all $x \in V$.

The composite of linear maps is once more a linear map. I'll forego the proof of the proposition below as it is identical to that of Proposition 5.3.6.

Proposition 7.2.13.

Suppose $T \in L(V, U)$ and $S \in L(U, W)$ then $S \circ T \in L(V, W)$.

A vector space V together with a multiplication $m: V \times V \to V$ is called an **algebra**⁵. For example, we saw before that square matrices form an algebra with respect to addition and matrix multiplication. Notice that V = L(W, W) is likewise naturally an algebra with respect to function addition and composition. In the section which follows we'll find the needed techniques to interchange the matrix and linear transformation formulation. We already found the explicit correspondence for transformations from \mathbb{R}^n to \mathbb{R}^m in part I. In our current context a bit more fine print is required due to the rich variety of basis choice.

The theorem below says the inverse of a linear transformation is also a linear transformation.

Theorem 7.2.14.

Suppose $T \in L(V, W)$ has inverse function $S : \operatorname{Range}(T) \to V$ then $S \in L(\operatorname{Range}(T), V)$.

Proof: let W = Range(T) and suppose $T \circ S = Id_W$ and $S \circ T = Id_V$. Suppose $x, y \in W$ hence there exits $a, b \in V$ for which T(a) = x and T(b) = y. Also, let $c \in \mathbb{R}$. Consider,

$$S(cx + y) = S(cT(a) + T(b)).$$

= $S(T(ca + b))$: by linearity of T
= $ca + b$: def. of identity function
= $cS(x) + S(y)$: note $a = S(T(a)) = S(x)$ and $b = S(T(b)) = S(y).$

Therefore, S is a linear transformation. \Box

Another way we can create new linear transformations from a given transformation is by restriction. Recall that the restriction of a given function is simply a new function where part of the domain has been removed. Since linear transformations are only defined on vector spaces we naturally are only permitted restrictions to subspaces of a given vector space.

Definition 7.2.15.

If $T: V \to W$ is a linear transformation and $U \subseteq V$ then we define $T|_U: U \to W$ by $T|_U(x) = T(x)$ for all $x \in U$. We say $T|_U$ is the **restriction of** T **to** U.

⁵it is somewhat ironic that all too often we often neglect to define an algebra in our modern algebra courses in the US eductional system. As students, you ought to demand more. See Dummit and Foote for a precise definition

Proposition 7.2.16.

If
$$T \in L(V, W)$$
 and $U \leq V$ then $T|_U \in L(U, W)$.

Proof: let $x, y \in U$ and $c \in \mathbb{R}$. Since $U \leq V$ it follows $cx + y \in U$ thus

$$T|_{U}(cx+y) = T(cx+y) = cT(x) + T(y) = cT|_{u}(x) + T|_{U}(y)$$

where I use linearity of T for the middle equality and the definition of $T|_U$ for the outside equalities. Therefore, $T|_U \in L(U, W)$. \Box

We can create a linear transformation on an infinity of vectors by prescribing its values on the basis alone. This is a fantastic result.

Proposition 7.2.17.

Suppose β is a basis for a vector space V and suppose W is also a vector space. Furthermore, suppose $L: \beta \to W$ is a function. There exists a unique linear extension of L to V.

Proof: to begin, let us understand the final sentence. A linear extension of L to V means a function $T: V \to W$ which is a linear transformation and $T|_{\beta} = L$. Uniqueness requires that we show if T_1, T_2 are two such extensions then $T_1 = T_2$. With that settled, let us begin the actual proof.

Suppose $\beta = \{v_1, \ldots, v_n\}$ if $x \in V$ then there exist $x_1, \ldots, x_n \in \mathbb{R}$ for which $x = \sum_{i=1}^n x_i v_i$. Therefore, define $T: V \to W$ as follows

$$T(x) = T\left(\sum_{i=1}^{n} x_i v_i\right) = \sum_{i=1}^{n} x_i L(v_i).$$

Clearly $T|_{\beta} = L$. I leave proof that $T \in L(V, W)$ to the reader. Suppose T_1, T_2 are two such extensions. Consider, $x = \sum_{i=1}^n x_i v_i$

$$T_1(x) = T_1\left(\sum_{i=1}^n x_i v_i\right) = \sum_{i=1}^n x_i L(v_i).$$

However, the same calculation holds for $T_2(x)$ hence $T_1(x) = T_2(x)$ for all $x \in V$ therefore the extension T is unique. \Box .

When we make use of the proposition above we typically use it to simplify a definition of a given linear transformation. In practice, we may define a mapping on a basis then **extend linearly**.

We conclude this section by initiating our discussion of **isomorphism**.

Definition 7.2.18.

Vector spaces V and W are **isomorphic** if there exists an invertible linear transformation $L: V \to W$. Furthermore, an invertible linear transformation is called an **isomorphism**. We write $V \approx W$ if V and W are isomorphic.

Notice that it suffices to check $L: V \to W$ is linear and invertible. Linearity of L^{-1} follows by Theorem 7.2.14. This is nice as it means we have less work to do when proving some given mapping is an isomorphism.

Theorem 7.2.19.

If $V \approx W$ then $\dim(V) = \dim(W)$

Proof: Let $L: V \to W$ be an isomorphism. Invertible linear mappings are injective we know that both L and L^{-1} must preserve LI of sets. In particular, if β is a basis for V then $L(\beta)$ must be a LI set in W. Likewise, if γ is a basis for W then $L^{-1}(\gamma)$ must be a LI set in V. Recall Theorem 6.6.10 gave that any LI subset of a finite-dimensional vector space could be extended to a basis. It follows that⁶ $\#(\beta) \leq \#(\gamma)$ and $\#(\gamma) \leq \#(\beta)$ hence $\#(\beta) = \#(\gamma)$. The theorem follows as $\#(\beta) = \dim(V)$ and $\#(\gamma) = \dim(W)$ by definition of dimension. \Box

This theorem has a converse. We need a proposition before we prove the other half.

Proposition 7.2.20.

If $T: V \to U$ and $S: U \to W$ are isomorphisms then $S \circ T$ is an isomorphism. Moreover, \approx is an equivalence relation.

Proof: let $T \in L(V, U)$ and $S \in L(U, W)$ be isomorphisms. Recall Proposition 7.2.13 gives us $S \circ T \in L(V, W)$ so, by Theorem 7.2.14, all that remains is to prove $S \circ T$ is invertible. Observe that $T^{-1} \circ S^{-1}$ serves as the inverse of $S \circ T$. In particular, calculate:

$$(S \circ T)(T^{-1} \circ S^{-1})(x)) = S(T(T^{-1}(S^{-1}(x)))) = S(S^{-1}(x)) = x.$$

Thus $(S \circ T) \circ (T^{-1} \circ S^{-1}) = Id_W$. Similarly, $(T^{-1} \circ S^{-1}) \circ (S \circ T) = id_V$. Therefore $S \circ T$ is invertible with inverse $T^{-1} \circ S^{-1}$.

The proof that \approx is an equivalence relation is not difficult. Begin by noting that $T = Id_V$ gives an isomorphism of V to V hence $V \approx V$; that is \approx is reflexive. Next, if $T: V \to W$ is an isomorphism then $T^{-1}: W \to V$ is also an isomorphism by Theorem 7.2.14 thus $V \approx W$ implies $W \approx V$; \approx is symmetric. Finally, suppose $V \approx U$ and $U \approx W$ by $T \in L(V, U)$ and $S \in L(U, W)$ isomorphisms. We proved that $S \circ T \in L(V, W)$ is an isomorphism hence $V \approx W$; that is, \approx is transitive. Therefore, \approx is an equivalence relation on the set of vector spaces of finite dimension. \Box

I included the comment about finite dimension as some of our theorems fail when the dimension is infinite. It is certainly not the case that all infinite dimensional vector spaces are isomorphic.

Theorem 7.2.21.

Let V, W be finite dimensional vector spaces.
$$V \approx W$$
 iff $\dim(V) = \dim(W)$

Proof: we already proved \Rightarrow in Theorem 7.2.19. Let us work on the converse. Suppose dim $(V) = \dim(W)$. Let β be a basis for V. In particular, denote $\beta = \{v_1, \ldots, v_n\}$. Define $\Phi_\beta : \beta \to \mathbb{R}^n$ by $\Phi_\beta(v_i) = e_i$ and extend linearly. But, if $\gamma = \{w_1, \ldots, w_n\}$ is the basis for W (we know they have the same number of elements by our supposition dim $(V) = \dim(W)$) then we may also define $\Phi_\gamma : W \to \mathbb{R}^n$ by $\Phi_\gamma(w_i) = e_i$ and extend linearly. Clearly Φ_β^{-1} and Φ_γ^{-1} exist and are easily desribed by $\Phi_\beta^{-1}(e_i) = v_i$ and $\Phi_\gamma^{-1}(e_i) = w_i$ extended linearly. Therefore, Φ_β and Φ_γ are isomorphisms. In particular, we've shown $V \approx \mathbb{R}^n$ and $W \approx \mathbb{R}^n$. By transitivity of \approx we find $V \approx W$. \Box

⁶here I use the notation that # is the function which counts the number of elements in a finite set

The proof above leads us naturally to the topic of the next section. In particular, the proof above contains a sketch of why Φ_{β} is an isomorphism. Please note I give many explicit examples of isomorphisms in the final section of this chapter. Those can be read at any point, I didn't include them here as to maintain a better flow for the theory. That said, you ought to look at them soon to get a better conceptual grasp.

7.3 matrix of linear transformation

I used the notation $[v]_{\beta}$ in the last chapter since it was sufficient. Now we need to have better notation for the coordinate maps so we can articulate the concepts clearly. Throughout this section we assume V is a vector space with basis $\beta = \{v_1, \ldots, v_n\}$.

Definition 7.3.1.

Let V be a finite dimensional vector space with basis $\beta = \{v_1, v_2, \dots, v_n\}$. The coordinate map $\Phi_{\beta} : V \to \mathbb{R}^n$ is defined by $\Phi_{\beta}(x_1v_1 + x_2v_2 + \dots + x_nv_n) = x_1e_1 + x_2e_2 + \dots + x_ne_n$ for all $v = x_1v_1 + x_2v_2 + \dots + x_nv_n \in V$.

We argued in the previous section that Φ_{β} is an invertible, linear transformation from V to \mathbb{R}^n . In other words, Φ_{β} is an **isomorphism**. It is worthwhile to note the linear extensions of

$$\Phi_{\beta}(v_i) = e_i \qquad \& \qquad \Phi_{\beta}^{-1}(e_i) = v_i$$

encapsulate the action of the coordinate map and its inverse. The coordinate map is a machine which converts an abstract basis to the standard basis.

Example 7.3.2. Let $V = \mathbb{R}^{2 \times 2}$ with basis $\beta = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ then

$$\Phi_{\beta}\left(\left[\begin{array}{cc}a&b\\c&d\end{array}\right]\right)=(a,b,c,d).$$

Example 7.3.3. Let $V = \mathbb{C}^n$ as a real vector space. Let $\beta = \{e_1, \ldots, e_n, ie_1, \ldots, ie_n\}$ be the basis of this 2n-dimensional vector space over \mathbb{R} . Observe $v \in \mathbb{C}^n$ has v = x + iy where $x, y \in \mathbb{R}^n$. In particular, if $\overline{a + ib} = a - ib$ and $\overline{v} = (\overline{v_1}, \ldots, \overline{v_n})$ then the identity below shows how to construct x, y:

$$v = \underbrace{\frac{1}{2}(v+\bar{v})}_{Re(v)=x} + \underbrace{\frac{1}{2}(v-\bar{v})}_{iIm(v)=iy}$$

and it's easy to verify $\bar{x} = x$ and $\bar{y} = y$ hence $x, y \in \mathbb{R}^n$ as claimed. The coordinate mapping is simple enough in this notation,

$$\Phi_{\beta}(x+iy) = (x,y).$$

Here we abuse notation slightly. Technically, I ought to write

$$\Phi_{\beta}(x+iy) = (x_1, \dots, x_n, y_1, \dots, y_n).$$

Example 7.3.4. Let $V = P_n$ with $\beta = \{1, (x-1), (x-1)^2, \dots, (x-1)^n\}$. To find the coordinates of an n-th order polynomial in standard form $f(x) = a_n x^n + \dots + a_1 x + a_0$ requires some calculation. We've all taken calculus II so we know Taylor's Theorem.

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(1)}{n!} (x-1)^n$$

also, clearly the series truncates for the polynomial in question hence,

$$f(x) = f(1) + f'(1)(x-1) + \dots + \frac{f^{(n)}(1)}{n!}(x-1)^n$$

Therefore,

$$\Phi_{\beta}(f(x)) = \left(f(1), f'(1), \dots, f^{(n)}(1)\right)$$

Example 7.3.5. Let $V = \{A = \sum_{i,j=1}^{2} A_{ij} E_{ij} \mid A_{11} + A_{22} = 0, A_{12} \in P_1, A_{11}, A_{22}, A_{21} \in \mathbb{C}\}$. If $A \in V$ then we can write:

$$A = \left\lfloor \frac{a+ib \quad ct+d}{x+iy \quad -a-ib} \right\rfloor$$

A natural choice for basis β is seen

$$\beta = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \begin{bmatrix} 0 & t \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ i & 0 \end{bmatrix} \right\}$$

The coordinate mapping follows easily in the notation laid out above,

$$\Phi_{\beta}(A) = (a, b, c, d, x, y).$$

Now that we have a little experience with coordinates as mappings let us turn to the central problem of this section: how can we associate a matrix with a given linear transformation $T: V \to W$?. It turns out we'll generally have to choose a basis for V and W in order to answer this question unambiguously. Therefore, let β once more serve as the basis for V and suppose γ is a basis for W. We assume $\#(\beta), \#(\gamma) < \infty$ throughout this discussion. The answer to the question is actually in the diagram below:



The matrix $[T]_{\beta,\gamma}$ induces a linear transformation from \mathbb{R}^n to \mathbb{R}^m . This means $[T]_{\beta,\gamma} \in \mathbb{R}^{m \times n}$. It is defined by the demand that the diagram above **commutes**. There are several formulas you can read into that comment. To express T explicitly as a combination of matrix multiplication and coordinate maps observe:

$$T = \Phi_{\gamma}^{-1} \circ L_{[T]_{\beta,\gamma}} \circ \Phi_{\beta}.$$

On the other hand, we could write

$$L_{[T]_{\beta,\gamma}} = \Phi_{\gamma} \circ T \circ \Phi_{\beta}^{-1}$$

if we wish to explain how to calculate $L_{[T]_{\beta,\gamma}}$ in terms of the coordinate maps and T directly. To select the *i*-th column in $[T]_{\beta,\gamma}$ we simply operate on $e_i \in \mathbb{R}^n$. This reveals,

$$col_i([T]_{\beta,\gamma}) = \Phi_{\gamma}(T(\Phi_{\beta}^{-1}(e_i)))$$

However, as we mentioned at the outset of this section, $\Phi_{\beta}^{-1}(e_i) = v_i$ hence

$$col_i([T]_{\beta,\gamma}) = \Phi_{\gamma}(T(v_i)) = [T(v_i)]_{\gamma}$$

where I have reverted to our previous notation for coordinate vectors⁷. Stringing the columns out, we find perhaps the nicest way to look at the matrix of an abstract linear transformation:

$$[T]_{\beta,\gamma} = [[T(v_1)]_{\gamma}|\cdots|[T(v_n)]_{\gamma}]$$

Each column is a W-coordinate vector which is found in \mathbb{R}^m and these are given by the *n*-basis vectors which generate V.

Alternatively, the commuting of the diagram yields:

$$\Phi_{\gamma} \circ T = L_{[T]_{\beta,\gamma}} \circ \Phi_{\beta}$$

If we feed the expression above an arbitrary vector $v \in V$ we obtain:

$$\Phi_{\gamma}(T(v)) = L_{[T]_{\beta,\gamma}}(\Phi_{\beta}(v)) \qquad \Rightarrow \qquad [T(v)]_{\gamma} = [T]_{\beta,\gamma}[v]_{\beta}$$

In practice, as I work to formulate $[T]_{\beta,\gamma}$ for explicit problems I find the boxed formulas convenient for calculational purposes. On the other hand, I have used each formula on this page for various theoretical purposes. Ideally, you'd like to understand these rather than memorize. I hope you are annoyed I have yet to define $[T]_{\beta,\gamma}$. Let us pick a definition for specificity of future proofs.

Definition 7.3.6.

Let V be a vector space with basis $\beta = \{v_1, \ldots, v_n\}$. Let W be a vector space with basis $\gamma = \{w_1, \ldots, w_m\}$. If $T: V \to W$ is a linear transformation then we define the **matrix of** T with respect to β, γ as $[T]_{\beta,\gamma}$ which is implicitly defined by

$$L_{[T]_{\beta,\gamma}} = \Phi_{\gamma} \circ T \circ \Phi_{\beta}^{-1}.$$

The discussion preceding this definition hopefully gives you some idea on what I mean by "implicitly" in the above context. In any event, we pause from our general discussion to illustrate with some explicit examples.

Example 7.3.7. Let $S: V \to W$ with $V = W = \mathbb{R}^{2 \times 2}$ are given bases $\beta = \gamma = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ and $L(A) = A + A^T$. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and calculate, $S(A) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} 2a & b + c \\ b + c & 2d \end{bmatrix}$

⁷the mapping notation supplements the $[v]_{\beta}$ notation, I use both going forward in these notes

Observe,

$$[A]_{\beta} = (a, b, c, d) \qquad \& \qquad [S(A)]_{\gamma} = (2a, b + c, b + c, 2d)$$

Moreover, we need a matrix $[S]_{\beta,\gamma}$ such that $[S(A)]_{\gamma} = [S]_{\beta,\gamma}[A]_{\beta}$. Tilt head, squint, and see

$$[S]_{\beta,\gamma} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Example 7.3.8. Let $V = P_1^{2 \times 2}$ be the set of 2×2 matrices with first order polynomials. Define T(A(x)) = A(2) where $T: V \to W$ and $W = \mathbb{R}^{2 \times 2}$. Let $\gamma = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ be the basis for W. Let β be the basis⁸ with coordinate mapping

$$\Phi_{\beta}\left(\left[\frac{a+bx \mid c+dx}{e+fx \mid g+hx}\right]\right) = (a, b, c, d, e, f, g, h)$$

We calculate for $v = \begin{bmatrix} a+bx & c+dx \\ \hline e+fx & g+hx \end{bmatrix}$ that

$$T(v) = \left[\begin{array}{c|c} a+2b & c+2d \\ \hline e+2f & g+2h \end{array} \right]$$

Therefore,

$$[T(v)]_{\gamma} = (a+2b, c+2d, e+2f, g+2h)$$

and as the coordinate vector $[v]_{\beta} = (a, b, c, d, e, f, g, h)$ the formula $[T(v)]_{\gamma} = [T]_{\beta, \gamma}[v]_{\beta}$ indicates

$$[T]_{\beta,\gamma} = \begin{bmatrix} 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

Example 7.3.9. Let $T: P_3 \to P_3$ be the derivative operator; T(f(x)) = f'(x). Give P_3 the basis $\beta = \{1, x, x^2, x^3\}$. Calculate,

$$T(a + bx + cx^{2} + dx^{3}) = b + 2cx + 3dx^{2}$$

Furthermore, note, setting $v = a + bx + cx^2 + dx^3$

$$[T(v)]_{\beta} = (b, 2c, 3d, 0) \qquad \& \qquad [v]_{\beta} = (a, b, c, d) \qquad \Rightarrow \qquad [T]_{\beta, \beta} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The results of Proposition 5.3.3 and 5.3.7 naturally generalize to our current context.

⁸you should be able to find β in view of the coordinate map formula

Proposition 7.3.10.

Suppose
$$S, T \in L(V, W)$$
 where β is a finite basis for V and γ a finite basis for W then
(1.) $[T+S]_{\beta,\gamma} = [T]_{\beta,\gamma} + [S]_{\beta,\gamma},$ (2.) $[T-S]_{\beta,\gamma} = [T]_{\beta,\gamma} - [S]_{\beta,\gamma},$ (3.) $[cS]_{\beta,\gamma} = c[S]_{\beta,\gamma}.$

Proof: the proof follows immediately from the identity below:

$$\Phi_{\gamma} \circ (T + cS) \circ \Phi_{\beta}^{-1} = \Phi_{\gamma} \circ T \circ \Phi_{\beta}^{-1} + c \Phi_{\gamma} \circ S \circ \Phi_{\beta}^{-1}.$$

This identity is true due to the linearity properties of the coordinate mappings. \Box

The generalization of Proposition 5.3.7 is a bit more interesting.

Proposition 7.3.11.

Let U, V, W be finite-dimensional vector spaces with bases β, γ, δ respectively. If $S \in L(U, W)$ and $T \in L(V, U)$ then $[S \circ T]_{\gamma, \delta} = [S]_{\beta, \delta}[T]_{\gamma, \beta}$

Proof: Notice that $L_A \circ L_B = L_{AB}$ since $L_A(L_B(v)) = L_A(Bv) = ABv = L_{AB}(v)$ for all v. Hence,

$L_{[S]_{\beta,\delta}[T]_{\gamma,\beta}} = L_{[S]_{\beta,\delta}} \circ L_{[T]_{\gamma,\beta}}$:set $A = [S]_{\beta,\delta}$ and $B = [T]_{\gamma,\beta}$,
$= (\Phi_{\delta} \circ S \circ \Phi_{\beta}^{-1}) \circ (\Phi_{\beta} \circ T \circ \Phi_{\gamma}^{-1})$:defn. of matrix of linear transformation,
$= \Phi_{\delta} \circ (S \circ T) \circ \Phi_{\gamma}^{-1}$:properties of function composition,
$=L_{[S\circ T]_{\gamma,\delta}}$:defn. of matrix of linear transformation.

The mapping $L: \mathbb{R}^{m \times n} \to L(\mathbb{R}^n, \mathbb{R}^m)$ is injective. Thus, $[S \circ T]_{\gamma, \delta} = [S]_{\beta, \delta}[T]_{\gamma, \beta}$ as we claimed. \Box

If we apply the result above to a linear transformation on a vector space V where the same basis is given to the domain and codomain some nice things occur. For example:

Example 7.3.12. Continuing Example 7.3.9. Observe that $T^2(f(x)) = T(T(f(x)) = f''(x)$. Thus if $v = a + bx + cx^2 + dx^3$ then $T^2: P_3 \to P_3$ has $T^2(v) = 2c + 6dx$ hence $[T^2(v)]_{\beta} = (2c, 6d, 0, 0)$ and we find $[T^2]_{\beta,\beta} = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. You can check that $[T^2]_{\beta,\beta} = [T]_{\beta,\beta}[T]_{\beta,\beta}$. Notice, we can

easily see that $[T^3]_{\beta,\beta} \neq 0$ whereas $[T^n]_{\beta,\beta} = 0$ for all $n \geq 4$. This makes $[T]_{\beta,\beta}$ a **nilpotent** matrix of order 4. We study the structure of nilpotent matrices in Part III of this course.

Example 7.3.13. Let V, W be vector spaces of dimension n. In addition, suppose $T : V \to W$ is a linear transformation with inverse $T^{-1} : W \to V$. Let V have basis β whereas W has basis γ . We know that $T \circ T^{-1} = Id_W$ and $T^{-1} \circ T = Id_V$. Furthermore, I invite the reader to show that $[Id_V]_{\beta,\beta} = I \in \mathbb{R}^{n \times n}$ where $n = \dim(V)$ and similarly $[Id_W]_{\gamma,\gamma} = I \in \mathbb{R}^{n \times n}$. Apply Proposition 7.3.11 to find

$$[T^{-1} \circ T]_{\beta,\beta} = [T^{-1}]_{\gamma,\beta}[T]_{\beta,\gamma}$$

but, $[T^{-1} \circ T]_{\beta,\beta} = [Id_V]_{\beta,\beta} = I$ thus $[T^{-1}]_{\gamma,\beta}[T]_{\beta,\gamma} = I$ and we conclude $([T]_{\beta,\gamma})^{-1} = [T^{-1}]_{\gamma,\beta}$. Phew, that's a relief. Wouldn't it be strange if this weren't true? Moral of story: the inverse matrix of the transformation is the matrix of the inverse transformation.

Lemma 7.3.14.

If $\Psi: V \to W$ is an isomorphism and S is a LI set then $\Psi(S)$ is a LI set with $\#(S) = \#\Psi(S)$.

Proof: If Ψ is an isomorphism then Ψ is injective. By part (3.) of Theorem 7.2.8 we have S LI implies $\Psi(S)$ is LI. Furthermore, if there is any repeated vector in S then clearly S is linearly dependent hence the vectors in S must be distinct. The lemma follows. \Box

Lemma 7.3.15.

Let $T: V \to W$ be a linear transformation where $\dim(V) = n$ and $\dim(W) = m$. Let $\Phi_{\beta}: V \to \mathbb{R}^n$ and $\Phi_{\gamma}: W \to \mathbb{R}^m$ be coordinate map isomorphisms. If β, γ are bases for V, W respective then $[T]_{\beta,\gamma}$ satisfies the following

(1.) Null($[T]_{\beta,\gamma}$) = $\Phi_{\beta}(\operatorname{Ker}(T))$, (2.) Col($[T]_{\beta,\gamma}$) = $\Phi_{\gamma}(\operatorname{Range}(T))$.

Proof of (1.): Let $v \in \text{Null}([T]_{\beta,\gamma})$ then there exists $x \in V$ for which $v = [x]_{\beta}$. By definition of nullspace, $[T]_{\beta,\gamma}[x]_{\beta} = 0$ hence, applying the identity $[T(x)]_{\gamma} = [T]_{\beta,\gamma}[x]_{\beta}$ we obtain $[T(x)]_{\gamma} = 0$ which, by injectivity of Φ_{γ} , yields T(x) = 0. Thus $x \in \text{Ker}(T)$ and it follows that $[x]_{\beta} \in \Phi_{\beta}(\text{Ker}(T))$. Therefore, $\text{Null}([T]_{\beta,\gamma}) \subseteq \Phi_{\beta}(\text{Ker}(T))$.

Conversely, if $[x]_{\beta} \in \Phi_{\beta}(\operatorname{Ker}(T))$ then there exists $v \in \operatorname{Ker}(T)$ for which $\Phi_{\beta}(v) = [x]_{\beta}$ hence, by injectivity of Φ_{β} , x = v and T(x) = 0. Observe, by linearity of Φ_{γ} , $[T(x)]_{\gamma} = 0$. Recall once more, $[T(x)]_{\gamma} = [T]_{\beta,\gamma}[x]_{\beta}$. Hence $[T]_{\beta,\gamma}[x]_{\beta} = 0$ and we conclude $[x]_{\beta} \in \operatorname{Null}([T]_{\beta,\gamma})$. Consequently, $\Phi_{\beta}(\operatorname{Ker}(T)) \subseteq \operatorname{Null}([T]_{\beta,\gamma})$.

Thus $\Phi_{\beta}(\text{Ker}(T)) = \text{Null}([T]_{\beta,\gamma})$. I leave the proof of (2.) to the reader. \Box

I should caution that the results above are basis dependendent in the following sense: If β_1, β_2 are bases with coordinate maps $\Phi_{\beta_1}, \Phi_{\beta_2}$ then it is not usually true that $\Phi_{\beta_1}(\operatorname{Ker}(T)) = \Phi_{\beta_2}(\operatorname{Ker}(T))$. It follows that $\operatorname{Null}([T]_{\beta_1,\gamma}) \neq \operatorname{Null}([T]_{\beta_2,\gamma})$ in general. That said, there is something which is common to all the nullspaces (and ranges); dimension. The dimension of the nullspace much match the dimension of the kernel. The dimension of the column space must match the dimension of the range. This result follows immediately from the two lemmas given above. See Definition 7.2.7 for rank and nullity of a linear transformation verses Definition 6.6.13 for matrices.

Proposition 7.3.16.

Let $T: V \to W$ be a linear transformation of finite dimensional vector spaces with basis β for V and γ for W then

 $\operatorname{nullity}(T) = \operatorname{nullity}([T]_{\beta,\gamma}) \qquad \& \qquad \operatorname{rank}(T) = \operatorname{rank}([T]_{\beta,\gamma}).$

You should realize the nullity and rank on the L.H.S. and R.H.S of the above proposition are quite different quantities in concept. It required some effort on our part to connect them, but, now that they are connected, perhaps you appreciated the names.

7.4 coordinate change

Vectors in abstract vector spaces do not generically come with a preferred coordinate system. There are infinitely many different choices for the basis of a given vector space. Naturally, for specific examples, we tend to have a pet-basis, but this is merely a consequence of our calculational habits. We need to find a way to compare coordinate vectors for different choices of basis. Then, the same ambiguity must be faced by the matrix of a transformation. In some sense, if you understand the diagrams then you can write all the required formulas for this section. That said, we will cut the problem for mappings of column vectors a bit more finely. There are nice matrix-theoretic formulas for \mathbb{R}^n that I'd like for you to know when you leave this course⁹.

7.4.1 coordinate change of abstract vectors

Let V be a vector space with bases β and $\overline{\beta}$. Let $\beta = \{v_1, \ldots, v_n\}$ whereas $\overline{\beta} = \{\overline{v}_1, \ldots, \overline{v}_n\}$. Let $x \in V$ then there exist **column vectors** $[x]_{\beta} = (x_1, \ldots, x_n)$ and $[x]_{\overline{\beta}} = (\overline{x}_1, \ldots, \overline{x}_n) \in \mathbb{R}^n$ such that

$$x = \sum_{i=1}^{n} x_i v_i \qquad \& \qquad x = \sum_{j=1}^{n} \bar{x}_i \bar{v}_i$$

Or, in mapping notation, $x = \Phi_{\beta}^{-1}([x]_{\beta})$ and $x = \Phi_{\overline{\beta}}^{-1}([x]_{\overline{\beta}})$. Of course x = x hence

$$\Phi_{\beta}^{-1}([x]_{\beta}) = \Phi_{\bar{\beta}}^{-1}([x]_{\bar{\beta}}).$$

Operate by $\Phi_{\bar{\beta}}$ on both sides,

$$[x]_{\bar{\beta}} = \Phi_{\bar{\beta}}(\Phi_{\beta}^{-1}([x]_{\beta})).$$

Observe that $\Phi_{\bar{\beta}} \circ \Phi_{\beta}^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation, therefore we can calculate its standard matrix. Let us collect our thoughts:

Proposition 7.4.1.

Using the notation developed in this subsection, if $P_{\beta,\bar{\beta}} = [\Phi_{\bar{\beta}} \circ \Phi_{\beta}^{-1}]$ then $[x]_{\bar{\beta}} = P_{\beta,\bar{\beta}}[x]_{\beta}$.

The diagram below contains the essential truth of the above proposition:



⁹ I mean, don't wait until then, nows a perfectly good time to learn them

Example 7.4.2. Let $V = \{A \in \mathbb{R}^{2 \times 2} | A_{11} + A_{22} = 0\}$. Observe $\beta = \{E_{12}, E_{21}, E_{11} - E_{22}\}$ gives a basis for V. On the other hand, $\bar{\beta} = \{E_{12}+E_{21}, E_{12}-E_{21}, E_{11}-E_{22}\}$ gives another basis. We denote $\beta = \{v_i\}$ and $\bar{\beta} = \{\bar{v}_i\}$. Let's work on finding the change of basis matrix. I can do this directly by our usual matrix theory. To find column i simply multiply by e_i . Or let the transformation act on e_i . The calculations below require a little thinking. I avoid algebra by thinking here.

$$\Phi_{\bar{\beta}}(\Phi_{\beta}^{-1}(e_1)) = \Phi_{\bar{\beta}}(E_{12}) = \Phi_{\bar{\beta}}\left(\frac{1}{2}\left[\bar{v}_1 + \bar{v}_2\right]\right) = (1/2, 1/2, 0).$$

$$\Phi_{\bar{\beta}}(\Phi_{\beta}^{-1}(e_2)) = \Phi_{\bar{\beta}}(E_{21}) = \Phi_{\bar{\beta}}\left(\frac{1}{2}\left[\bar{v}_1 - \bar{v}_2\right]\right) = (1/2, -1/2, 0).$$

$$\Phi_{\bar{\beta}}(\Phi_{\beta}^{-1}(e_3)) = \Phi_{\bar{\beta}}(E_{11} - E_{22}) = \Phi_{\bar{\beta}}(\bar{v}_3) = (0, 0, 1).$$

Admittably, if the bases considered were not so easily related we'd have some calculation to work through here. That said, we find:

$$P_{\beta,\bar{\beta}} = \left[\begin{array}{rrr} 1/2 & 1/2 & 0\\ 1/2 & -1/2 & 0\\ 0 & 0 & 1 \end{array} \right]$$

Let's take it for a ride. Consider $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$ clearly $[A]_{\beta} = (2,3,1)$. Calculate,

$$P_{\beta,\bar{\beta}}[A]_{\beta} = \begin{bmatrix} 1/2 & 1/2 & 0\\ 1/2 & -1/2 & 0\\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2\\ 3\\ 1 \end{bmatrix} = \begin{bmatrix} 5/2\\ -1/2\\ 1 \end{bmatrix} = [A]_{\bar{\beta}}$$

Is this correct? Check,

$$\Phi_{\bar{\beta}}^{-1}(5/2, -1/2, 1) = \frac{5}{2} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \frac{1}{2} \cdot \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} = A$$

Yep. It works.

It is often the case we face coordinate change for mappings from $\mathbb{R}^n \to \mathbb{R}^m$. Or, even more special m = n. The formulas we've detailed thus far find streamlined matrix-theoretic forms in that special context. We turn our attention there now.

7.4.2 coordinate change for column vectors

Let β be a basis for \mathbb{R}^n . In contrast to the previous subsection, we have a standard basis with which we can compare; in particular, **the** standard basis. Hazzah!¹⁰. Let $\beta = \{v_1, \ldots, v_n\}$ and note the **matrix of** β is simply defined by concatenating the basis into an $n \times n$ invertible matrix $[\beta] = [v_1|\cdots|v_n]$. If $x \in \mathbb{R}^n$ then the coordinate vector $[x]_\beta = (y_1, \ldots, y_n)$ is the column vector such that

$$x = [\beta][x]_{\beta} = y_1 v_1 + \cdots + y_n v_n$$

here I used "y" to avoid some other more annoying notation. It is not written in stone, you could use $([x]_{\beta})_i$ in place of y_i . Unfortunately, I cannot use x_i in place of y_i as the notation x_i is

¹⁰sorry, we visited Medieval Times over vacation and it hasn't entirely worn off just yet

already reserved for the Cartesian components of x. Notice, as $[\beta]$ is invertible we can solve for the coordinate vector:

$$[x]_{\beta} = [\beta]^{-1}x$$

If we had another basis $\bar{\beta}$ then

$$[x]_{\bar{\beta}} = [\bar{\beta}]^{-1}x$$

Naturally, x exists independent of these bases so we find common ground at x:

$$x = [\beta][x]_{\beta} = [\bar{\beta}][x]_{\bar{\beta}}$$

We find the coordinate vectors are related by:

$$[x]_{\bar{\beta}} = [\bar{\beta}]^{-1}[\beta][x]_{\beta}$$

Let us summarize are findings in the proposition below:

Proposition 7.4.3.

Using the notation developed in this subsection and the last, if $P_{\beta,\bar{\beta}} = [\Phi_{\bar{\beta}} \circ \Phi_{\beta}^{-1}]$ then $[x]_{\bar{\beta}} = P_{\beta,\bar{\beta}}[x]_{\beta}$ and a simple formula to calculate the change of basis matrix is given by $P_{\beta,\bar{\beta}} = [\bar{\beta}]^{-1}[\beta]$. We also note for future convenience: $[\bar{\beta}]P_{\beta,\bar{\beta}} = [\beta]$

Example 7.4.4. Let $\beta = \{(1,1), (1,-1)\}$ and $\gamma = \{(1,0), (1,1)\}$ be bases for \mathbb{R}^2 . Find $[v]_{\beta}$ and $[v]_{\gamma}$ if v = (2,4). Let me frame the problem, we wish to solve:

$$v = [\beta][v]_{\beta}$$
 and $v = [\gamma][v]_{\gamma}$

where I'm using the basis in brackets to denote the matrix formed by concatenating the basis into a single matrix,

$$[\beta] = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad and \quad [\gamma] = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

This is the 2×2 case so we can calculate the inverse from our handy-dandy formula:

$$[\beta]^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad and \quad [\gamma]^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

Then multiplication by inverse yields $[v]_{\beta} = [\beta]^{-1}v$ and $[v]_{\gamma} = [\gamma]^{-1}v$ thus:

$$[v]_{\beta} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \quad and \quad [v]_{\gamma} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Let's verify the relation of $[v]_{\gamma}$ and $[v]_{\beta}$ relative to the change of basis matrix. In particular, we expect that if $P_{\beta,\gamma} = [\gamma]^{-1}[\beta]$ then $[v]_{\gamma} = P_{\beta,\gamma}[v]_{\beta}$. Calculate,

$$P_{\beta,\gamma} = [\gamma]^{-1}[\beta] = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 1 & -1 \end{bmatrix}$$

As the last great American president said, trust, but, verify

$$P_{\beta,\gamma}[v]_{\beta} = \begin{bmatrix} 0 & 2\\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3\\ -1 \end{bmatrix} = \begin{bmatrix} -2\\ 4 \end{bmatrix} = [v]_{\gamma} \checkmark$$

It might be helpful to some to see a picture of just what we have calculated. Finding different coordinates for a given point (which corresponds to a vector from the origin) is just to prescribe different zig-zag paths from the origin along basis-directions to get to the point. In the picture below I illustrate the standard basis path and the β -basis path.



Now that we've seen an example, let's find $[v]_{\beta}$ for an arbitrary v = (x, y),

$[v]_{\beta} = \frac{1}{2} \begin{bmatrix} 1\\1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$	$\left[\begin{array}{c} x\\ y \end{array}\right] = \left $	$\left[\begin{array}{c}\frac{1}{2}(x+y)\\\frac{1}{2}(x-y)\end{array}\right]$
--	---	--	--

If we denote $[v]_{\beta} = (\bar{x}, \bar{y})$ then we can understand the calculation above as the relation between the barred and standard coordinates:

$$\bar{x} = \frac{1}{2}(x+y)$$
 $\bar{y} = \frac{1}{2}(x-y)$

Conversely, we can solve these for x, y to find the inverse transformations:

$$x = \bar{x} + \bar{y} \qquad y = \bar{x} - \bar{y}.$$

Similar calculations are possible with respect to the γ -basis.

7.4.3 coordinate change of abstract linear transformations

In Definition 7.3.6 we saw that if V is a vector space with basis $\beta = \{v_1, \ldots, v_n\}$ and W be a vector space with basis $\gamma = \{w_1, \ldots, w_m\}$. Then a linear transformation $T: V \to W$ has matrix $[T]_{\beta,\gamma}$ defined implicitly by:

$$L_{[T]_{\beta,\gamma}} = \Phi_{\gamma} \circ T \circ \Phi_{\beta}^{-1}.$$

If there was another pair of bases $\bar{\beta}$ for V and $\bar{\gamma}$ for W then we would likewise have

$$L_{[T]_{\bar{\beta},\bar{\gamma}}} = \Phi_{\bar{\gamma}} \circ T \circ \Phi_{\bar{\beta}}^{-1}$$

Solving for T relates the matrices with and without bars:

$$T = \Phi_{\gamma}^{-1} \circ L_{[T]_{\beta,\gamma}} \circ \Phi_{\beta} = \Phi_{\bar{\gamma}}^{-1} \circ L_{[T]_{\bar{\beta},\bar{\gamma}}} \circ \Phi_{\bar{\beta}}.$$

From which the proposition below follows:

Proposition 7.4.5.

Using the notation developed in this subsection

$$[T]_{\bar{\beta},\bar{\gamma}} = [\Phi_{\bar{\gamma}} \circ \Phi_{\gamma}^{-1}][T]_{\beta,\gamma}[\Phi_{\beta} \circ \Phi_{\bar{\beta}}^{-1}]$$

Moreover, recalling $P_{\beta,\bar\beta} = [\Phi_{\bar\beta} \circ \Phi_\beta^{-1}]$ we find:

$$[T]_{\bar{\beta},\bar{\gamma}} = P_{\gamma,\bar{\gamma}}[T]_{\beta,\gamma}(P_{\beta,\bar{\beta}})^{-1}$$

Note, if there exist invertible matrices P, Q such that B = PAQ then B and A are said to be **matrix congruent**. The proposition above indicates that the matrices of a given linear transformation¹¹ are congruent. In particular, $[T]_{\bar{\beta},\bar{\gamma}}$ is congruent to $[T]_{\beta,\gamma}$.

The picture below can be used to remember the formulas in the proposition above.



Example 7.4.6. Let $V = P_2$ and $W = \mathbb{C}$. Define a linear transformation $T : V \to W$ by T(f) = f(i). Thus,

$$T(ax^{2} + bx + c) = ai^{2} + bi + c = c - a + ib.$$

Use coordinate maps given below for $\beta = \{x^2, x, 1\}$ and $\gamma = \{1, i\}$:

$$\Phi_{\beta}(ax^2 + bx + c) = (a, b, c) \qquad \& \qquad \Phi_{\gamma}(a + ib) = (a, b).$$

Observe $[T(ax^2 + bx + c)]_{\gamma} = (c - a, b)$ hence $[T]_{\beta,\gamma} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

¹¹of finite dimensional vector spaces

7.4. COORDINATE CHANGE

Let us change the bases to

$$\bar{\beta} = \{(x-2)^2, (x-2), 1\}$$
 & $\bar{\gamma} = \{i, 1\}$

Calculate, if $f(x) = ax^2 + bx + c$ then f'(x) = 2ax + b and f''(x) = 2a. Observe, f(2) = 4a + 2b + c and f'(2) = 4a + b and f''(2) = 2a hence, using the Taylor expansion centered at 2,

$$f(x) = f(2) + f'(2)(x-2) + \frac{1}{2}f''(2)(x-2)^2$$

= 4a + 2b + c + (4a + b)(x - 2) + a(x - 2)^2.

Therefore,

$$\Phi_{\bar{\beta}}(ax^2 + bx + c) = (a, 4a + b, 4a + 2b + c)$$

But, $\Phi_{\beta}^{-1}(a, b, c) = ax^2 + bx + c$. Thus,

$$\Phi_{\bar{\beta}}(\Phi_{\beta}^{-1}(a,b,c)) = (a,4a+b,4a+2b+c) \qquad \Rightarrow \qquad [\Phi_{\bar{\beta}} \circ \Phi_{\beta}^{-1}] = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix}$$

Let's work out this calculation in the other direction (it's actually easier and what we need in a bit)

 $\Phi_{\beta}(a(x-2)^2 + b(x-2) + c) = \Phi_{\beta}(a(x^2 - 4x + 4) + b(x-2) + c) = (a, -4a + b, 4a - 2b + c)$ But, $\Phi_{\bar{\beta}}^{-1}(a, b, c) = a(x-2)^2 + b(x-2) + c$ therefore:

$$\Phi_{\beta}(\Phi_{\bar{\beta}}^{-1}(a,b,c)) = (4a - 2b + c, -4a + b, a) \qquad \Rightarrow \qquad [\Phi_{\beta} \circ \Phi_{\bar{\beta}}^{-1}] = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 4 & -2 & 1 \end{bmatrix}$$

On the other hand, $\Phi_{\bar{\gamma}}(a+ib) = (b,a)$. Of course, $a+ib = \Phi_{\gamma}^{-1}(a,b)$ hence $\Phi_{\bar{\gamma}}(\Phi_{\gamma}^{-1}(a,b)) = (b,a)$. It follows that $[\Phi_{\bar{\gamma}} \circ \Phi_{\gamma}^{-1}] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ We'll use the change of basis proposition to find the matrix w.r.t. $\bar{\beta}$ and $\bar{\gamma}$

$$\begin{split} [T]_{\bar{\beta},\bar{\gamma}} &= [\Phi_{\bar{\gamma}} \circ \Phi_{\gamma}^{-1}][T]_{\beta,\gamma} [\Phi_{\beta} \circ \Phi_{\bar{\beta}}^{-1}]. \\ &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 4 & -2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 4 & -2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -4 & 1 & 0 \\ 3 & -2 & 1 \end{bmatrix}. \end{split}$$

Continuing, we can check this by direct calculation of the matrix. Observe

$$T(a(x-2)^{2} + b(x-2) + c) = a(i-2)^{2} + b(i-2) + c$$

= $a[-1 - 4i + 4] + b(i-2) + c$
= $3a - 2b + c + i(-4a + b)$

Thus, $[T(a(x-2)^2+b(x-2)+c)]_{\bar{\gamma}} = (-4a+b, 3a-2b+c)$ hence $[T]_{\bar{\beta},\bar{\gamma}} = \begin{bmatrix} -4 & 1 & 0\\ 3 & -2 & 1 \end{bmatrix}$. Which agrees nicely with our previous calculation.

7.4.4 coordinate change of linear transformations of column vectors

We specialize Proposition 7.4.7 in this subsection in the case that $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$. In particular, the result of Proposition 7.4.3 makes life easy; $P_{\beta,\bar{\beta}} = [\bar{\beta}]^{-1}[\beta]$ likewise, $P_{\gamma,\bar{\gamma}} = [\bar{\gamma}]^{-1}[\gamma]$

Proposition 7.4.7.

Using the notation developed in this subsection

$$[T]_{\bar{\beta},\bar{\gamma}} = [\bar{\gamma}]^{-1}[\gamma][T]_{\beta,\gamma}[\beta]^{-1}[\bar{\beta}].$$

The standard matrix [T] is related to the non-standard matrix $[T]_{\bar{\beta},\bar{\gamma}}$ by:

$$[T]_{\bar{\beta},\bar{\gamma}} = [\bar{\gamma}]^{-1}[T][\beta].$$

Proof: Proposition 7.4.7 with $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$ together with the result of Proposition 7.4.3 give us the first equation. The second equation follows from the observation that for standard bases β and γ we have $[\beta] = I_n$ and $[\gamma] = I_m$. \Box

Example 7.4.8. Let $\bar{\beta} = \{(1,0,1), (0,1,1), (4,3,1)\}$. Furthermore, define a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ by the rule T(x, y, z) = (2x - 2y + 2z, x - z, 2x - 3y + 2z). Find the matrix of T with respect to the basis β . Note first that the standard basis is read from the rule:

$$T\left(\begin{bmatrix} x\\ y\\ z \end{bmatrix} \right) = \begin{bmatrix} 2x - 2y + 2z\\ x - z\\ 2x - 3y + 2z \end{bmatrix} = \begin{bmatrix} 2 & -2 & 2\\ 1 & 0 & -1\\ 2 & -3 & 2 \end{bmatrix} \begin{bmatrix} x\\ y\\ z \end{bmatrix}$$

Next, use the proposition with $\bar{\beta} = \bar{\gamma}$ (omitting the details of calculating $[\bar{\beta}]^{-1}$)

$$\begin{split} [\bar{\beta}]^{-1}[T][\bar{\beta}] &= \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ -1/2 & 1/2 & 1/2 \\ 1/6 & 1/6 & -1/6 \end{bmatrix} \begin{bmatrix} 2 & -2 & 2 \\ 1 & 0 & -1 \\ 2 & -3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ -1/2 & 1/2 & 1/2 \\ 1/6 & 1/6 & -1/6 \end{bmatrix} \begin{bmatrix} 4 & 0 & 4 \\ 0 & -1 & 3 \\ 4 & -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{split}$$

Therefore, in the $\bar{\beta}$ -coordinates the linear operator T takes on a particularly simple form. In particular, if $\bar{\beta} = \{f_1, f_2, f_3\}$ then¹²

$$\overline{T}(\bar{x},\bar{y},\bar{z}) = 4\bar{x}f_1 - \bar{y}f_2 + \bar{z}f_3$$

This linear transformation acts in a special way in the f_1, f_2 and f_3 directions. The basis we considered here is called an **eigenbasis** for T. We study eigenvectors and the associated problem of diagonalization in Part III.

¹²some authors just write T, myself included, but, technically $\overline{T} = T \circ \Phi_{\overline{\beta}}^{-1}$, so... as I'm being pretty careful otherwise, it would be bad form to write the prettier, but wrong, T

7.5 theory of dimensions for maps

In some sense this material is naturally paired with Section 7.2 and Section 6.6. I had to wait until this point in the presentation because I wanted to tie in some ideas about coordinate change.

This section is yet another encounter with a **classification theorem**. Previously, we learned that vector spaces are classified by their dimension; $V \approx W$ iff $\dim(V) = \dim(W)$. In this section, we'll find a nice way to lump together many linear transformations as being essentially the same function with a change of notation. However, the concept of *same* is a slippery one. In this section, **matrix congruence** is the measure of sameness. In contrast, later we study **similarity transformations** or **orthogonal transformations**. The concept that unites these discussions is classification. We seek a standard representative of an equivalence class. The type of equivalence class depends naturally on what is considered the "same". Be careful with this word "same" it might not mean the same thing to you.

The theorem below is to linear transformations what Theorem 6.6.15 is for matrices.

Theorem 7.5.1.

Let V, W be vector spaces of finite dimension over \mathbb{R} . In particular, suppose dim(V) = n and dim(W) = m. If $T: V \to W$ be a linear transformation then

 $\dim(V) = \dim(\operatorname{Ker}(T)) + \dim(\operatorname{Range}(T)).$

Proof: I'll give two proofs. The first is based on coordinates and Theorem 7.2.8 which includes the result that an injective linear transformation maps LI sets to LI sets.

Proof 1: Let β, γ be bases for V, W respectively. Define $A = [T]_{\beta,\gamma}$. Observe $A \in \mathbb{R}^{m \times n}$. Apply Theorem 6.6.15 to find

$$n = \dim(\operatorname{Null}(A)) + \dim(\operatorname{Col}(A)).$$

We found in Lemma 7.3.15 that the basis for $\operatorname{Ker}(T)$ is obtained by mapping the basis β_N for $\operatorname{Null}([T]_{\beta,\gamma})$ to V by Φ_{β}^{-1} . That is, $\Phi_{\beta}^{-1}(\beta_N) = \beta_K$ serves as a basis for $\operatorname{Ker}(T) \leq V$. On the other hand, Lemma 7.3.15 also stated the basis for the column space $\beta_C \subset \mathbb{R}^m$ is mapped to a basis for $\operatorname{Range}(T)$ in W. In particular, we define $\beta_R = \Phi_{\gamma}^{-1}(\beta_C)$ and it serves as a basis for $\operatorname{Range}(T) \leq W$. Lemma 7.3.15 also proved $\#(\beta_N) = \#(\beta_K)$ and $\#(\beta_C) = \#(\beta_R)$. Thus,

$$\dim(V) = n = \dim(\operatorname{Null}(A)) + \dim(\operatorname{Col}(A)) = \dim(\operatorname{Ker}(T)) + \dim(\operatorname{Range}(T)). \square$$

Proof 2: Note $\operatorname{Ker}(T) \leq V$ therefore we may select a basis $\beta_K = \{v_1, \ldots, v_k\}$ for $\operatorname{Ker}(T)$ by Proposition 6.6.6. By the basis extension theorem (think $W = \operatorname{Ker}(T)$ and apply Theorem 6.6.10) we can adjoin the set of vectors $\beta_{not \ K} = \{v_{k+1}, \ldots, v_n\}$ to make $\beta = \beta_K \cup \beta_{not \ K}$ a basis for V. Suppose $x = \sum_{i=1}^n x_i v_i \in V$ and calculate by linearity of T,

$$T(x) = \sum_{i=1}^{k} x_i T(v_i) + \sum_{i=k+1}^{n} x_i T(v_i) = \sum_{i=k+1}^{n} x_i T(v_i),$$

where $v_1, \ldots, v_k \in \text{Ker}(T)$ gives $T(v_1) = \cdots = T(v_k) = 0$. Observe, it follows that the set of n - k vectors $\gamma = \{T(v_{k+1}), \ldots, T(v_n)\}$ serves as a spanning set for Range(T). Moreover, we may argue

that γ is a LI set: suppose

$$c_{k+1}T(v_{k+1}) + \dots + c_nT(v_n) = 0$$

by linearity of T it follows:

 $T(c_{k+1}v_{k+1}+\cdots+c_nv_n)=0$

hence $c_{k+1}v_{k+1} + \cdots + c_nv_n \in \text{Ker}(T)$. However, by construction, $\beta_{not K} = \{v_{k+1}, \ldots, v_n\}$ are not in the kernel thus

$$c_{k+1}v_{k+1} + \dots + c_nv_n = 0.$$

Next, as $\beta_{not \ K} \subseteq \beta$ the LI of β implies the LI of $\beta_{not \ K}$ hence we conclude $c_{k+1} = 0, \ldots, c_n = 0$. Therefore, γ is a basis for Range(T). Finally, as dim(V) = n = n - k + k and dim(Ker(T)) = k and dim(Range(T)) = n - k we conclude

$$\dim(V) = \dim(\operatorname{Ker}(T)) + \dim(\operatorname{Range}(T)).$$

Proof of the theorem that follows below is essentially contained in the proof of Theorem 7.5.1. However, for the sake of completeness, I include a separate proof.

Theorem 7.5.2.

Let V, W be vector spaces of finite dimension over \mathbb{R} . If $T: V \to W$ be a linear transformation with rank $(T) = \dim(T(V)) = p$. Then, there exist bases β for V and γ for W such that:

$$[T]_{\beta,\gamma} = \begin{bmatrix} I_p & 0\\ \hline 0 & 0 \end{bmatrix}$$

where, as is our standard notation, $[T(v)]_{\gamma} = [T]_{\beta,\gamma}[v]_{\beta}$ for all $v \in V$.

Proof: Let $\dim(V) = n$ and $\dim(W) = m$ for convenience of exposition. By Theorem 7.5.1) we have $\dim(Ker(T)) = n - p$. Let $\{v_{p+1}, \ldots, v_n\}$ form a basis for $Ker(T) \leq V$. Extend the basis for Ker(T) to a basis $\beta = \{v_1, \ldots, v_p, v_{p+1}, \ldots, v_n\}$ for V. Observe, by construction, $\{T(v_1), \ldots, T(v_p)\}$ is linearly independent. Define,

$$w_1 = T(v_1), \ldots, w_p = T(v_p)$$

Clearly $\{w_1, \ldots, w_p\}$ forms a basis for the image T(V). Next, extend $\{w_1, \ldots, w_p\}$ to a basis $\gamma = \{w_1, \ldots, w_p, w_{p+1}, \ldots, w_m\}$ for W. Observe:

$$[T(v_j)]_{\gamma} = [T]_{\beta,\gamma}[v_j]_{\beta} = [T]_{\beta,\gamma}e_j = Col_j([T]_{\beta,\gamma})$$

Furthermore, for j = 1, ..., p, by construction $T(v_j) = w_j$ and hence $[T(v_j)]_{\gamma} = [w_j]_{\gamma} = \bar{e}_j \in \mathbb{R}^m$. On the other hand, for j = p + 1, ..., n we have $T(v_j) = 0$ hence $[T(v_j)]_{\gamma} = [0]_{\gamma} = 0 \in \mathbb{R}^m$. Thus,

$$[T]_{\beta,\gamma} = [e_1|\cdots|e_p|0|\cdots|0]$$

and it follows that $[T]_{\beta,\gamma} = \begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix}$. \Box .

The claim of the theorem just proved says the following: there exists a choice of coordinates which makes a given linear transformation a projection onto the range. In terms of matrix congruence, this theorem reveals the canonical form for matrices which are equivalent under matrix congruence $A \mapsto QAP^{-1}$. However, the proof above does not reveal too much about how to find such coordinates. We next investigate a calculational method to find β, γ for which the theorem is realized.

Suppose $T \in L(V, W)$ where dim(V) = n and dim(W) = m. Furthermore, suppose $\beta' = \{v'_1, \ldots, v'_n\}$ and $\gamma' = \{w'_1, \ldots, w'_m\}$ are bases for V and W respective. We define $[T]_{\beta'\gamma'}$ as usual:

$$[T]_{\beta'\gamma'} = [[T(v'_1)]_{\gamma'}] \cdots |[T(v'_n)]_{\gamma'}]$$

There exists a product of elementary $m \times m$ matrices E_1 for which

$$R_1 = \operatorname{rref}([T]_{\beta'\gamma'}) = E_1[T]_{\beta'\gamma'}$$

Let p be the number of pivot columns in R_1 . Observe that the last (m-p) rows in R_1 are zero. Therefore, the last (m-p) columns in R_1^T are zero. Gauss-Jordan elimination on R_1 is accomplished by multiplication by E_2 which is formed from a product of $n \times n$ elementary matrices.

$$R_2 = \operatorname{rref}(R_1^T) = E_2 R_1^T$$

Notice that the trivial rightmost (m-p) columns stay trivial under the Gauss-Jordan elimination. Moreover, the nonzero pivot rows in R_1 become *p*-pivot columns in R_1^T which reduce to e_1, \ldots, e_p standard basis vectors in \mathbb{R}^n for R_2 (the leading ones are moved to the top rows with row-swaps if necessary). In total, we find: (the subscripts indicate the size of the blocks)

$$E_2 R_1^T = [e_1|\cdots|e_p|0|\cdots|0] = \left[\frac{I_p | 0_{p\times(m-p)}}{0_{(n-p)\times p} | 0_{(n-p)\times(m-p)}}\right]$$

Therefore,

$$E_{2}(E_{1}[T]_{\beta'\gamma'})^{T} = \left[\begin{array}{c|c} I_{p} & 0_{p\times(m-p)} \\ \hline 0_{(n-p)\times p} & 0_{(n-p)\times(m-p)} \end{array} \right]$$

Transposition of the above equation yields the following:

$$E_{1}[T]_{\beta'\gamma'}E_{2}^{T} = \left[\begin{array}{c|c} I_{p} & 0_{p\times(n-p)} \\ \hline 0_{(m-p)\times p} & 0_{(m-p)\times(n-p)} \end{array}\right]$$

If β, γ are bases for V and W respective then we relate the matrix $[T]_{\beta,\gamma}$ to $[T]_{\beta'\gamma'}$ as follows:

$$[T]_{\beta,\gamma} = [\Phi_{\beta'} \circ \Phi_{\beta}^{-1}][T]_{\beta'\gamma'} [\Phi_{\gamma} \circ \Phi_{\gamma'}^{-1}]$$

Therefore, we ought to define β by imposing $[\Phi_{\beta'} \circ \Phi_{\beta}^{-1}] = E_1$ and γ by $[\Phi_{\gamma} \circ \Phi_{\gamma'}^{-1}] = E_2^T$. Using $L_A(v) = Av$ notation for E_1, E_2^T ,

$$L_{E_1} = \Phi_{\beta'} \circ \Phi_{\beta}^{-1} \qquad \& \qquad L_{E_2^T} = \Phi_{\gamma} \circ \Phi_{\gamma}$$

Thus,

$$\Phi_{\beta}^{-1} = \Phi_{\beta'}^{-1} \circ L_{E_1} \qquad \& \qquad \Phi_{\gamma}^{-1} = \Phi_{\gamma'} \circ L_{E_2^T}^{-1}$$

and we construct β and γ explicitly by:

$$\beta = \{ (\Phi_{\beta'}^{-1} \circ L_{E_1})(e_j) \}_{j=1}^n \qquad \gamma = \{ (\Phi_{\gamma'} \circ L_{E_2^T}^{-1})(e_j) \}_{j=1}^m.$$

Note the formulas above merely use the elementary matrices and the given pair of bases. The discussion of this page shows that β and γ so constructed will give $[T]_{\beta,\gamma} = \begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix}$.

Continuing, to implement the calculation outlined in the previous page we would like an efficient method to calculate E_1 and E_2 . We can to do this much as we did for computation of the inverse. I illustrate the idea below¹³:

Example 7.5.3. Let $A = \begin{bmatrix} 1 & 3 & 4 \\ 1 & 4 & 5 \\ 1 & 0 & 1 \\ 1 & 2 & 3 \end{bmatrix}$. If we adjoin the identity matrix to right the matrix which

is constructed in the Gauss-Joran elimination is the product of elementary matrices P for which rref(A) = PA.

$$rref[A|I_4] = rref \begin{bmatrix} 1 & 3 & 4 & | & 1 & 0 & 0 & 0 \\ 1 & 4 & 5 & | & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & | & 0 & 0 & 1 & 0 \\ 1 & 2 & 3 & | & 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & | & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & | & 0 & 0 & -1/2 & 1/2 \\ 0 & 0 & 0 & | & 1 & 0 & 1/2 & -3/2 \\ 0 & 0 & 0 & | & 0 & 1 & 1 & -2 \end{bmatrix}$$

We can read P for which rref(A) = PA from the result above, it is simply

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & -1/2 & 1/2 \\ 1 & 0 & 1/2 & -3/2 \\ 0 & 1 & 1 & -2 \end{bmatrix}.$$

Next, consider row reduction on the transpose of the reduced matrix. This corresponds to column operations on the reduced matrix.

$$rref[(rref(A))^{T}|I_{3}] = rref\begin{bmatrix} 1 & 0 & 0 & 0 & | & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & | & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & | & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & | & 0 & -1 & 1 \\ 0 & 1 & 0 & 0 & | & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & | & 1 & 1 & -1 \end{bmatrix}$$

$$Let \ Q = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix} and \ define \ R \ by:$$

$$R^{T} = Q[rref(A)]^{T} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
Finally,
$$R = (Q[rref(A)]^{T})^{T} = rref(A)Q^{T} \ hence \ R = PAQ^{T}. \ In \ total,$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & -1/2 & 1/2 \\ 1 & 0 & 1/2 & -3/2 \\ 0 & 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 3 & 4 \\ 1 & 4 & 5 \\ 1 & 0 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ -1 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

 $^{^{13}}$ see Example 2.7 on page 244 of Hefferon's Linear Algebra for a slightly different take built on explicit computation of the product of the elementary matrices needed for the reduction

There is nothing terribly special about this example. We could follow the same procedure for a general matrix to find the explicit change of basis matrices which show the matrix congruence of A where $p = \operatorname{rank}(A)$. From a coordinate change perspective, this means we can always $\begin{array}{ccc}I_p & 0\\0 & 0\end{array}$ to change coordinates on a linear transformation to make the formula for the transformation a simple projection onto the first p-coordinates; $T(y_1,\ldots,y_p,y_{p+1},\ldots,y_n) = (y_1,\ldots,y_p,0,\ldots,0) \in \mathbb{R}^m$. Of course, the richness we saw in Section 5.1 is still here, it's just hidden in the coordinate change. In Part III we'll study other problems where different types of coordinate change are allowed. When there is less freedom to modify domain and codomain coordiantes it turns out the *cannonical forms* of the object are greater in variety and structure. Just to jump ahead a bit, if we force m = n and change coordinates in domain and codomain simultaneously then the real Jordan form captures a representative of each equivalence class of matrix up to a similarity transformation. On the other hand, Sylvester's Law of Inertia reveals the cannonical form for the matrix of a quadratic form is simply a diagonal matrix with $Diag(D) = (-1, \ldots, -1, 1, \ldots, 1, 0, \ldots, 0)$. Quadratic forms are non-linear functions which happen to have an associated matrix. The coordinate change for the matrix of a quadratic form is quite different than what we've studied thus far. In any event, this is just a foreshadowing comment, we will return to this discussion once we study eigenvectors and quadratic forms in part III.

7.6 quotient space

Let us begin with a discussion of how to add sets of vectors. If $S, T \subseteq V$ a vector space over \mathbb{R} then we define S + T as follows:

$$S + T = \{s + t \mid s \in S, t \in T\}$$

In the particular case $S = \{x\}$ it is customary to write

$$x + T = \{x + t \mid t \in T\}$$

we drop the $\{\}$ around x in this special case. In the case that $T = W \leq V$ the set of all such **cosets** x + W of W has a natural vector space structure induced from V. We now work towards motivating the definition of the quotient space.

Proposition 7.6.1.

Let V be vector space over
$$\mathbb{R}$$
 and $W \leq V$. Then $x + W = y + W$ iff $x - y \in W$.

Proof: Suppose x + W = y + W. If $p \in x + W$ then it follows there exists $w_1 \in W$ for which $p = x + w_1$. However, as $x + W \subseteq y + W$ we find $x + w_1 \in y + W$ and thus there exists $w_2 \in W$ for which $x + w_1 = y + w_2$. Therefore, $y - x = w_1 - w_2 \in W$ as W is a subspace of V.

Conversely, suppose $x, y \in V$ and $x - y \in W$. Thus, there exists $w \in W$ for which x - y = w and so for future reference x = y + w or y = x - w. Let $p \in x + W$ hence there exists $w_1 \in W$ for which $p = x + w_1$. Furthermore, as W is a subspace we know $w, w_1 \in W$ implies $w + w_1 \in W$. Consider then, $p = x + w_1 = y + w + w_1 \in y + W$. Therefore, $x + W \subseteq y + W$. A similar argument shows $y + W \subseteq x + W$ hence x + W = y + W. \Box

Proposition 7.6.2.

```
Let V be vector space over \mathbb{R} and W \leq V. Then x + W = W iff x \in W.
```

Proof: if x + W = W then $x + w \in W$ for some w hence $x + w = w_1$. But, it follows $x = w_1 - w$ which makes clear that $x \in W$ as $W \leq V$.

Conversely, if $x \in W$ then consider $p = x + w_1 \in x + W$ and note $x + w_1 \in W$ hence $p \in W$ and we find $x + W \subseteq W$. Likewise, if $w \in W$ then note w = x + w - x and $w - x \in W$ thus $w \in x + W$ and we find $W \subseteq x + W$. Therefore, x + W = W. \Box

Observe that Proposition 7.6.1 can be reformulated to say x + W is the same as y + W if y = x + w for some $w \in W$. We say that x and y are coset representatives of the same coset iff x + W = y + W. Suppose $x_1 + W = x_2 + W$ and $y_1 + W = y_2 + W$; that is, suppose x_1, x_2 are representatives of the same coset and suppose y_1, y_2 are representatives of the same coset.

Proposition 7.6.3.

Let V be vector space over \mathbb{R} and $W \leq V$. If $x_1 + W = x_2 + W$ and $y_1 + W = y_2 + W$ and $c \in \mathbb{R}$ then $x_1 + y_1 + W = x_2 + y_2 + W$ and $cx_1 + W = cx_2 + W$.

Proof: Suppose $x_1 + W = x_2 + W$ and $y_1 + W = y_2 + W$ then by Proposition 7.6.1 we find $x_2 - x_1 = w_x$ and $y_2 - y_2 = w_y$ for some $w_x, w_y \in W$. Consider

$$(x_2 + y_2) - (x_1 + y_1) = x_2 - x_1 + y_2 - y_1 = w_x + w_y.$$

However, $w_x, w_y \in W$ implies $w_x + w_y \in W$ hence by Proposition 7.6.1 we find $x_1 + y_1 + W = x_2 + y_2 + W$. I leave proof that $cx_1 + W = cx_2 + W$ as an exercise to the reader. \Box

The preceding triple of propositions serves to show that the definitions given below are independent of the choice of coset representative. That is, while a particular coset representative is used to make the definition, the choice is immaterial to the outcome.

Definition 7.6.4.

We define V/W to be the **quotient space of** V by W. In particular, we define:

$$V/W = \{x + W \mid x \in V\}$$

and for all $x + W, y + W \in V/W$ and $c \in \mathbb{R}$ we define:

$$(x+W) + (y+W) = x + y + W$$
 & $c(x+W) = cx + W.$

Note, we have argued thus far that addition and scalar multiplication defined on V/W are well-defined functions. Let us complete the thought:

Theorem 7.6.5.

If $W \leq V$ a vector space over \mathbb{R} then V/W is a vector space over \mathbb{R} .

Proof: if $x + W, y + W \in V/W$ note (x + W) + (y + W) and c(x + W) are single elements of V/W thus Axioms 9 and 10 of Definition 6.1.1 are true. Axiom 1: by commutativity of addition in V we obtain commutativity in V/W:

$$(x+W) + (y+W) = x + y + W = y + x + W = (y+W) + (x+W).$$

Axiom 2: associativity of addition follows from associativity of V,

$$\begin{aligned} (x+W) + [(y+W) + (z+W)] &= x + W + [(y+z) + W] & \text{defn. of } + \text{ in } V/W \\ &= x + (y+z) + W & \text{defn. of } + \text{ in } V/W \\ &= (x+y) + z + W & \text{associativity of } + \text{ in } V \\ &= [(x+y) + W] + (z+W) & \text{defn. of } + \text{ in } V/W \\ &= [(x+W) + (y+W)] + (z+W) & \text{defn. of } + \text{ in } V/W. \end{aligned}$$

Axiom 3: note that 0 + W = W and it follows that W serves as the additive identity in the quotient:

$$(x+W) + (0+W) = x + 0 + W = x + W$$

Axiom 4: the additive inverse of x + W is simply -x + W as (x + W) + (-x + W) = W. Axiom 5: observe that

$$1(x+W) = 1 \cdot x + W = x + W.$$

I leave verification of Axioms 6,7 and 8 for V/W to the reader. I hope you can see these will easily transfer of the Axioms 6,7 and 8 for V itself. \Box

The notation x + W is at times tiresome. An alternative notation is given below:

$$[x] = x + W$$

then the vector space operations on V/W are

$$[x] + [y] = [x + y]$$
 & $c[x] = [cx].$

Naturally, the disadvantage of this notation is that it hides the particular subspace by which the quotient is formed. For a given vector space V many different subspaces are typically available and hence a wide variety of quotients may be constructed.

Example 7.6.6. Suppose $V = \mathbb{R}^3$ and $W = span\{(0,0,1)\}$. Let $[(a,b,c)] \in V/W$ note

$$[(a,b,c)] = \{(a,b,z) \mid z \in \mathbb{R}\}$$

thus a point in V/W is actually a line in V. The parameters a, b fix the choice of line so we expect V/W is a two dimensional vector space with basis $\{[(1,0,0)], [(0,1,0)]\}$.

Example 7.6.7. Suppose $V = \mathbb{R}^3$ and $W = span\{(1,0,0), (0,1,0)\}$. Let $[(a,b,c)] \in V/W$ note

$$[(a, b, c)] = \{(x, y, c) \mid x, y \in \mathbb{R}\}\$$

thus a point in V/W is actually a plane in V. In this case, each plane is labeled by a single parameter c so we expect V/W is a one-dimensional vector space with basis $\{[(0,0,1)]\}$.

Example 7.6.8. Let $V = \mathbb{R}[x]$ and let $W = \mathbb{R}$ the set of constant polynomials.

$$[a_0 + a_1 x + \dots + a_n x^n] = \{c + a_1 x + \dots + a_n x^n \mid c \in \mathbb{R}\}\$$

Perhaps, more to the point,

$$[a_0 + a_1x + \dots + a_nx^n] = [a_1x + \dots + a_nx^n]$$

In this quotient space, we identify polynomials which differ by a constant.

We could also form quotients of $\mathcal{F}(\mathbb{R})$ or P_n or $C^{\infty}(\mathbb{R})$ by \mathbb{R} and it would have the same meaning; if we quotient by constant functions then [f] = [f + c].

The quotient space construction allows us to modify a given transformation such that its reformulation is injective. For example, consider the problem of inverting the derivative operator D = d/dx.

$$D(f) = f' \qquad \& \qquad D(f+c) = f'$$

thus D is not injective. However, if we instead look at the derivative operator on¹⁴ a quotient space of differentiable functions of a connected domain where [f] = [f + c] then defining D([f]) = f'proves to be injective. Suppose D([f]) = D([g]) hence f' = g' so f - g = c and [f] = [g]. We generalize this example in the next subsection.

7.6.1 the first isomorphism theorem

The style of this section is discussion/discovery. The culmination of the section occurs at the conclusion where the totality of the given discussion justifies the so-called *first isomorphism theorem*. Many of the arguments given here generalize nicely to the context of abstract group theory. I hope this discussion seeds your intuition for such future work.

Let $T: V \to U$ be a linear transformation and $W \leq V$. We have several natural formulas which we may associate with the quotient V/W. In particular, define $\pi: V \to V/W$ by

$$\pi(x) = x + W = [x].$$

Note that π is clearly a linear transformation and $\operatorname{Ker}(\pi) = W$ as $\pi(x) = 0 + W$ implies $x \in W$. Furthermore, we define $\overline{T}: V/W \to U$ by

$$\overline{T}(x+W) = T(x)$$

for all $x + W \in V/W$. Suppose x + W = y + W. Then $y - x \in W$ thus x = y + w for some $w \in W$.

$$\overline{T}(x+W) = T(x) = T(y+w) = T(y) + T(w) = T(w) + \overline{T}(y+W).$$

Hence \overline{T} is **not** a function as it is not **single-valued**. The presence of the T(w) in the equation above suggests there may be infinitely many values for which $\overline{T}(x+W) = T(x)$. Moreover, the formula is not independent of the representative for a general subspace W. How can we repair this result? How can we create a new function on the quotient space from the given linear transformation $T: V \to U$?

¹⁴to be careful, I only modify the domain of the derivative operator here, note the output of D is not an equivalence class. Furthermore, perhaps a different symbol like \overline{D} should be used to write $\overline{D}([f]) = f'$ as $D \neq \overline{D}$
The simple answer¹⁵ is to study W = Ker(T). Note, T(w) = 0 if $w \in \text{Ker}(T)$. Hence $\overline{T} : V/\text{Ker}(T) \to U$ is a function and we can easilt verify it is linear: let $[x], [y] \in V/\text{Ker}(T)$ and $c \in \mathbb{R}$. Observe:

$$\overline{T}(c[x] + [y]) = \overline{T}([cx + y]) = T(cx + y) = cT(x) + T(y) = c\overline{T}([x]) + \overline{T}([y]).$$

Hence $\overline{T} \in \mathcal{L}(V/\operatorname{Ker}(T), U)$. In addition to linearity, \overline{T} has another exceedingly nice property. Suppose

$$\overline{T}([x]) = 0 \Rightarrow T(x) = 0 \Rightarrow x \in \operatorname{Ker}(T) \Rightarrow [x] = 0.$$

Therefore, \overline{T} is **injective**. Well, this is nice, we almost have an isomorphism. We may lack surjectivity for \overline{T} .

To obtain an surjection from $T: V \to U$ we need to remove points from the codomain which the map fails to reach. In short, just replace U with T(V) and $T': V \to T(V)$ is **surjective**. If $\iota: U \to T(V)$ is the natural projection map $\iota(x) = x$ for all $x \in T(V)$ and $\iota(x) = 0$ for $x \notin T(V)$ then this allows us to express the formula for T' explicitly at the level of maps by $T' = \iota \circ T$. Naturally, if we combine this idea with the injection creating \overline{T} construction then we'll obtain a linear map which is both injective and surjective; we obtain an isomorphism from V/Ker(T) to T(V).

Theorem 7.6.9.

If $T: V \to U$ is a linear transformation and W = Ker(T) then the mapping $\Psi: V/W \to T(V)$ defined by $\Psi(x + \text{Ker}(T)) = T(x)$ is an isomorphism. Moreover, $\Psi = T \circ \pi$ where $\pi: V \to V/W$ is the natural quotient map defined by $\pi(x) = x + \text{Ker}(T)$.

The remaining detail we have to prove is the surjectivity of Ψ . Suppose $y \in T(v)$ then by definition there exists $x \in V$ such that T(x) = y. Note that $\Psi(x + \text{Ker}(T)) = T(x) = y$ hence Ψ is surjective as claimed. This theorem is not really that difficult if we understand the quotient construction and the freedom we have to define codomains to suit our purposes.

Example 7.6.10. Consider $D: P \to P$ defined by D(f(x)) = df/dx. Here I denote $P = \mathbb{R}[x]$, the set of all polynomials with real coefficients. Notice

$$Ker(D) = \{ f(x) \in P \mid df/dx = 0 \} = \{ f(x) \in P \mid f(x) = c \}.$$

In this case D is already a surjection since we work with all polynomials hence:

$$\Psi([f(x)]) = f'(x)$$

is an isomorphism. Just to reiterate in this case:

$$\Psi([f(x)]) = \Psi([g(x)]) \quad \Rightarrow \quad f'(x) = g'(x) \quad \Rightarrow \quad f(x) = g(x) + c \quad \Rightarrow \quad [f(x)] = [g(x)] = [g(x)] + c \quad \Rightarrow \quad [f(x)] = [g(x)] = [g(x)] + c \quad \Rightarrow \quad [f(x)] = [g(x)] = [g(x)] = [g(x)] = [g(x)] = [g(x)] = [g(x$$

Essentially, Ψ is just d/dx on equivalence classes of polynomials. Notice that $\Psi^{-1}: P \to P/Ker(D)$ is a mapping you have already studied for several months! In particular,

$$\Psi^{-1}(f(x)) = \{F(x) \mid dF/dx = f(x)\}$$

¹⁵you could alternatively swap the codomain U for U/T(W) which effectively makes T(w) = 0. I'll leave U alone for our current discussion, one quotient is enough to start.

Just to be safe, let's check that my formula for the inverse is correct:

$$\Psi^{-1}(\Psi([f(x)]) = \Psi^{-1}(df/dx) = \{F(x) \mid dF/dx = df/dx\} = \{f(x) + c \mid c \in \mathbb{R}\} = [f(x)].$$

Conversely, for $f(x) \in P$,

$$\Psi(\Psi^{-1}(f(x))) = \Psi(\{F(x) \mid dF/dx = f(x)\}) = f(x).$$

Perhaps if I use a different notation to discuss the preceding example then you will see what is happening: we usually call $\Psi^{-1}(f(x)) = \int f(x) dx$ and $\Psi = d/dx$ then

$$\frac{d}{dx}\int f\,dx = f \qquad \& \qquad \int \frac{d}{dx}(f+c_1)\,dx = f+c_2$$

In fact, if your calculus instructor was careful, then he should have told you that when we calculate the indefinite integral of a function the answer is not a function. Rather, $\int f(x) dx = \{g(x) \mid g'(x) = f(x)\}$. However, nobody wants to write a set of functions every time they integrate so we instead make the custom to write g(x) + c to indicate the non-uniqueness of the answer. Antidifferentiation of f is finding a specific function F for which F'(x) = f(x). Indefinite integration of f is the process of finding the set of all functions $\int f dx$ for which $\frac{d}{dx} \int f dx = f$. In any event, I hope you see that we can claim that differentiation and integration are inverse operations, however, this is in the understanding that we work on a quotient space of functions where two functions which differ by a constant are considered the same function. In that context, $f + c_1 = f + c_2$.

Example 7.6.11. Consider $D: P_2 \rightarrow P_2$ defined by

$$D(ax^2 + bx + c) = 2ax + b$$

Observe $\Psi([ax^2 + bx + c]) = 2ax + b$ defines a natural isomorphism from P_2/\mathbb{R} to P_1 where I denote $Ker(D) = \mathbb{R}$. In other words, when I write the quotient by \mathbb{R} I am identifying the set of constant polynomials with the set of real numbers.

Example 7.6.12. Consider $S : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ defined by $S(A) = A + A^T$. Notice that the range of S(A) is simply symmetric matrices as $(S(A))^T = (A + A^T)^T = A^T + (A^T)^T = A + A^T = S(A)$. Moreover, if $A^T = A$ the clearly S(A/2) = A hence S is onto the symmetric matrices. What is the kernel of S? Suppose S(A) = 0 and note:

$$A + A^T = 0 \qquad \Rightarrow \qquad A^T = -A.$$

Thus Ker(S) is the set of antisymmetric matrices. Therefore,

$$\Psi([A]) = A + A^T$$

is an isomorphism from $\mathbb{R}^{n \times n} / Ker(S)$ to the set of symmetric $n \times n$ matrices.

Example 7.6.13. This example will be most meaningful for students of differential equations, however, there is something here for everyone to learn. An n-th order linear differential equation can be written as L[y] = g. Here y and g are functions on a connected interval $I \subseteq \mathbb{R}$. There is an existence theorem for such problems which says that **any** solution can be written as

$$y = y_h + y_p$$

where $L[y_h] = 0$ and $L[y_p] = g$. The so-called **homogeneous solution** y_h is generally formed from a linear combination of n-LI fundamental solutions y_1, y_2, \ldots, y_n as

$$y_h = c_1 y_1 + c_2 y_2 + \dots + c_n y_n$$

Here $L[y_i] = 0$ for i = 1, 2, ..., n. It follows that Null(L) is n-dimensional and the fundamental solution set forms a basis for this null-space. On the other hand the particular solution y_p can be formed through a technique known as variation of parameters. Without getting into the technical details, the point is there is an explicit, although tedious, method to calculate y_p once we know the fundamental solution set and g. Techniques for finding the fundamental solution set vary from problem to problem. For the constant coefficient case or Cauchy Euler problems it is as simple as factoring the characteristic polynomial and writing down the homogeneous solutions. Enough about that, let's think about this problem in view of quotient spaces.

The differential equation L[y] = g can be instead thought of as a function which takes g as an input and produces y as an output. Of course, given the infinity of possible homogeneous solutions this would not really be a function, it's not single-valued. If we instead associate with the differential equation a function $H: V \to V/Null(L)$ where H(g) = y + Null(L) then the formula can be compactly written as $H(g) = [y_p]$. For convenience, suppose $V = C^0(\mathbb{R})$ then dom(H) = V as variation of parameters only requires integration of the forcing function g. Thus $H^{-1}: V/Null(L) \to V$ is an isomorphism. In short, the mathematics I outline here shows us there is a one-one correspondance between forcing functions and solutions modulo homogeneous terms. Linear differential equations have this beatiful feature; the net-response of a system L to inputs g_1, \ldots, g_k is nothing more than the sum of the responses to each forcing term. This is the principal of superposition which makes linear differential equations comparitively easy to understand.

There are many things to learn about quotient space. A few more are detailed in the next section and the exercises.

7.7 structure of subspaces

I will begin this section by following an elegant construction¹⁶ I found in Morton L. Curtis' Abstract Linear Algebra pages 28-30. A bit later, I take inspiration from the section on direct sum decompositions in Jim Hefferon's Linear Algebra. The results we encounter in this section prove useful in Part III when we study eigenvectors so we best be careful to remember our work here for later.

Recall the construction in Example 6.1.8, this is known as the **external direct sum**. If V, W are vector spaces over \mathbb{R} then $V \times W$ is given the following vector space structure:

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$$
 & $c(v, w) = (cv, cw).$

In the vector space $V \times W$ the vector $(0_V, 0_W) = 0_{V \times W}$. Although, usually we just write (0, 0) = 0. Furthermore, if $\beta_V = \{v_1, \ldots, v_n\}$ and $\beta_W = \{w_1, \ldots, w_m\}$ then a basis for $V \times W$ is simply:

$$\beta = \{(v_i, 0) | i \in \mathbb{N}_n\} \cup \{(0, w_j) | j \in \mathbb{N}_m\}$$

¹⁶I don't use his notation that $A \oplus B = A \times B$, I reserve $A \oplus B$ to denote internal direct sums.

I invite the reader to check LI of β . To see how β spans, please consider the calculation below:

$$(x,y) = (x,0) + (0,y)$$

= $(x_1v_1 + \dots + x_nv_n, 0) + (0, y_1w_1 + \dots + y_mw_m)$
= $x_1(v_1,0) + \dots + x_n(v_n,0) + y_1(0,w_1) + \dots + y_m(0,w_m)$

Thus β is a basis for $V \times W$ and we can count $\#(\beta) = n + m$ hence $\dim(V \times W) = \dim(V) + \dim(W)$.

When a given vector space V is isomorphic to $A \times B$ where $A, B \leq V$ then V is said to be the **internal direct sum** of A and B. In this case, it is customary to write $V = A \oplus B$. We study the mapping η as to connect the external and internal direct sum concepts.

the study the mapping if as to connect the external and internal areas

Theorem 7.7.1.

Suppose A, B ≤ V. Let η : A × B → V be defined by η(a, b) = a + b. Then:
(i.) η is linear
(ii.) η is injective iff A ∩ B = {0}
(iii.) η is surjective iff span(A ∪ B) = V

Proof: (i.) linearity of η follows from the calculation below:

$$\eta(c(a,b) + (x,y)) = \eta((ca + x, cb + y)) = ca + x + cb + y = c(a+b) + x + y = c\eta(a,b) + \eta(x,y).$$

(ii.) If η is injective and $x \in A \cap B$ then $x \in A$ and $x \in B$. Observe $\eta(x, 0) = \eta(0, x) = x$ hence (x, 0) = (0, x) and we conclude x = 0. Thus $A \cap B \subseteq \{0\}$ and clearly $\{0\} \subseteq A \cap B$ thus $A \cap B = \{0\}$. Conversely, suppose $A \cap B = \{0\}$. Suppose $\eta(x, y) = 0$ then x + y = 0. Thus $x \in A$ and $y \in B$ with x = -y hence $x, y \in A \cap B = \{0\}$. We find $\operatorname{Ker}(\eta) = \{(0, 0)\}$ thus the linear map η is injective.

(iii.) suppose η is surjective. If $v \in V$ then there exists $(a, b) \in A \oplus B$ for which $\eta(a, b) = v$. But, this is just to say a + b = v hence $v \in \operatorname{span}(A \cup B)$ thus $V \subseteq \operatorname{span}(A \cup B)$. Clearly $\operatorname{span}(A \cup B) \subseteq V$ hence $\operatorname{span}(A \cup B) = V$. Conversely, suppose $\operatorname{span}(A \cup B) = V$. If $v \in V$ then there exist $a_i \in A$ and $b_j \in B$ for which $v = \sum_{i=1}^k c_i a_i + \sum_{j=1}^l d_j b_j$ let $a = \sum_{i=1}^k c_i a_i$ and $b = \sum_{j=1}^l d_j b_j$ note $a \in A$ and $b \in B$ as A, B are subspaces. Note, $\eta(a, b) = a + b = v$. Hence η is surjective. \Box

Let us be precise for future reference.

Definition 7.7.2.

If $A, B \leq V$ and $\eta : A \times B \to V$ defined by $\eta(a, b) = a + b$ is an isomorphism then we say V is the **internal direct sum** of A and B and write $V = A \oplus B$.

An alternative definition of internal direct sum is given as follows: if

$$V = A + B \qquad \& \qquad A \cap B = \{0\}$$

then $V = A \oplus B$. If that definition is given then we have no need of the $A \times B$ construction since all the addition takes place inside V. However, I like the definition given since it helps us understand the relation of internal and external direct sums quite explicitly.

Proposition 7.7.3.

If $V = A \oplus B$ then $A \cap B = \{0\}$ and A + B = V where $A + B = \operatorname{span}(A \cup B)$.

Proof: by definition, $\eta : A \times B \to V$ is an isomorphism. Hence η is both injective and surjective so Theorem 7.7.1 affirms the proposition. \Box

A convenient notation for spans of a single element v in V a vector space over \mathbb{R} is simply $v\mathbb{R}$. I utilize this notation in the examples below.

Example 7.7.4. The cartesian plane $\mathbb{R}^2 = e_1 \mathbb{R} \oplus e_2 \mathbb{R}$.

Example 7.7.5. The complex numbers $\mathbb{C} = \mathbb{R} \oplus i\mathbb{R}$. We could discuss how extending $i^2 = -1$ linearly gives this an algebraic structure. We have a whole course in the major to dig into this example.

Example 7.7.6. The hyperbolic numbers $\mathcal{H} = \mathbb{R} \oplus j\mathbb{R}$. We could discuss how extending $j^2 = 1$ linearly gives this an algebraic structure. This is less known, but it naturally describes problems with some hyperbolic symmetry.

Example 7.7.7. The dual numbers $\mathcal{N} = \mathbb{R} \oplus \epsilon \mathbb{R}$. We could discuss how extending $\epsilon^2 = 0$ linearly gives this an algebraic structure.

The algebraic comments above are mostly for breadth. We focus on linear algebra¹⁷ in these notes.

Proposition 7.7.8.

If $V = A \oplus B$ then $V/A \approx B$.

Proof: Since $V \approx A \times B$ under $\eta : A \times B \to V$ with $\eta(a, b) = a + b$ it follows for each $v \in V$ there exists a unique pair (a, b) such that v = a + b. Given this decomposition of each vector in V we can define a projection onto B as follows: define $\pi_B : V \to B$ by $\pi_B(a + b) = a$. It is clear π_B is linear and $\text{Ker}(\pi_B) = A$ thus the first isomorphism theorem gives $V/A \approx B$. \Box

Naturally we should consider extending the discussion to more than two subspaces.

Definition 7.7.9.

If
$$A_1, A_2, \ldots, A_k \leq V$$
 and $\eta : A_1 \times A_2 \times \cdots \times A_k \to V$ defined by
 $\eta(a_1, a_2, \ldots, a_k) = a_1 + a_2 + \cdots + a_k$
is an isomorphism then we say V is the **internal direct sum** of A_1, A_2, \ldots, A_k and write
 $V = A_1 \oplus A_2 \oplus \cdots \oplus A_k$ which may also be denoted $V = \bigoplus_{i=1}^k A_i$.

The necessary criteria for a given k-tuple of subspaces $A_1, \ldots, A_k \leq V$ to form a direct sum decomposition of V. Naturally, we do need V to be covered by the sum of the subspaces formed by the span of their union. However, the nature of the isomorphism above above forbids some overlap between the subspaces. For example, V + V = V but we would not be able to say $V \oplus V = V$. Furthermore, you may be tempted to suppose the criteria for k = 2 suffices if we extend it pair wise here. But, the example below shows pairwise intersection triviality is also insufficient.

 $^{1^{7}}$ a vector space paired with a multiplication is called an algebra. The rules $i^{2} = -1$, $j^{2} = 1$ and $\epsilon^{2} = 0$ all serve to define non-isomorphic algebraic structures on \mathbb{R}^{2} . These are isomorphic as vector spaces.

Example 7.7.10. Let $A_1 = (1,1)\mathbb{R}$ and $A_2 = (1,0)\mathbb{R}$ and $A_3 = (1,1)\mathbb{R}$. It is not hard to verify $A_1 + A_2 + A_3 = span(A_1 \cup A_2 \cup A_3) = \mathbb{R}^2$ and $A_1 \cap A_2 = A_1 \cap A_3 = A_2 \cap A_3 = \{0\}$. However, it is certainly not possible to find an isomorphism of \mathbb{R}^2 and the three dimensional vector space $A_1 \times A_2 \times A_3$.

There are three ways to describe the needed criteria. This essentially Lemma 4.8 of Hefferon's *Linear Algebra* see page 130-131.

Theorem 7.7.11.

Suppose $A_1, A_2, \ldots, A_k \leq V$ and suppose $\beta_1, \beta_2, \ldots, \beta_k$ are bases for A_1, A_2, \ldots, A_k respective. The following are equivalent:

(i.) $V = A_1 \oplus A_2 \oplus \cdots \oplus A_k$

(ii.) each $v \in V$ there exist unique $v_i \in A_i$ such that $v = v_1 + v_2 + \cdots + v_k$

(iii.) $\beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ forms a basis for V

(iv.) any finite set $\{a_1, a_2, \ldots, a_k \mid 0 \neq a_i \in A_i \text{ for all } i \in \mathbb{N}_k\}$ is LI.

Proof: see Hefferon for a nice proof of the equivalence of (ii.), (iii.) and (iv.). Let us prove the equivalence of (i.) and (ii.). Suppose (i.) is true. Let $v \in V$ then there exists a unique k-tuple $(v_1, v_2, \ldots, v_k) \in A_1 \times A_2 \times \cdots \times A_k$ for which

$$\eta(v_1, v_2, \dots, v_k) = v_1 + v_2 + \dots + v_k = v.$$

This proves (ii.). Now suppose (ii.) is true. Define η for (i.) by the unique expansion for $v = v_1 + v_2 + \cdots + v_k$ we define

$$\eta^{-1}(v_1 + v_2 + \dots + v_k) = (v_1, v_2, \dots, v_k).$$

It follows that $\eta(v_1, v_2, \ldots, v_k) = v_1 + v_2 + \cdots + v_k$ defines an isomorphism from $A_1 \times A_2 \times \cdots \times A_k$ to V. This shows (i.) is true. \Box

Example 7.7.12. Quaternions. $\mathbb{H} = \mathbb{R} \oplus i\mathbb{R} \oplus j\mathbb{R} \oplus k\mathbb{R}$ where $i^2 = j^2 = k^2 = -1$. Our notation for vectors in most calculus texts has a historical basis in Hamilton's quarternions.

There is much more to say, but I'll stop here. I hope the exercises help bring further depth to this topic. In particular, when V permits a direct sum decomposition this allows us to align the basis with the decomposition. The result is certain formulas simplify in a very nice way due to a certain block structure. On a deeper level, there are some simple but elegant things which should be said about the **universal principal** as it can be seen at work in the first isomorphism theorem.

7.8 examples of isomorphisms

In your first read of this section, you might just read the examples. I have purposely put the big-picture and extracurricular commentary outside the text of the examples.

The coordinate map is an isomorphism which allows us to trade the abstract for the concrete.

Example 7.8.1. Let V be a vector space over \mathbb{R} with basis $\beta = \{f_1, \ldots, f_n\}$ and define Φ_β by $\Phi_\beta(f_j) = e_j \in \mathbb{R}^n$ extended linearly. In particular,

$$\Phi_{\beta}(v_1f_1 + \dots + v_nf_n) = v_1e_1 + \dots + v_ne_n$$

This map is a linear bijection and it follows $V \approx \mathbb{R}^n$.

Example 7.8.2. Suppose $V = \{A \in \mathbb{C}^{2 \times 2} \mid A^T = -A\}$ find an isomorphism to $P_n \leq \mathbb{R}[x]$ for appropriate n. Note, $A_{ij} = -A_{ji}$ gives $A_{11} = A_{22} = 0$ and $A_{12} = -A_{21}$. Thus, $A \in V$ has the form:

$$A = \left[\begin{array}{cc} 0 & a+ib\\ -a-ib & 0 \end{array} \right]$$

I propose that $\Psi(a+bx) = \begin{bmatrix} 0 & a+ib \\ -a-ib & 0 \end{bmatrix}$ provides an isomorphism of P_1 to V.

Example 7.8.3. Let $V = (\mathbb{C} \times \mathbb{R})^{2 \times 2}$ and $W = \mathbb{C}^{2 \times 3}$. The following is an isomorphism from V to W:

$$\Psi\left[\begin{array}{ccc} (z_1, x_1) & (z_2, x_2) \\ (z_3, x_3) & (z_4, x_4) \end{array}\right] = \left[\begin{array}{cccc} z_1 & z_2 & z_3 \\ z_4 & x_1 + ix_2 & x_3 + ix_4 \end{array}\right]$$

Example 7.8.4. Consider $V \times W/(\{0\} \times W)$ and V. To show these are isomorphic we consider T(v, w) = v. It is simple to verify that $T: V \times W \to V$ is a linear surjection. Moreover, $Ker(T) = \{(0, w) \mid w \in W\} = \{0\} \times W$. The first isomorphism theorem reveals $V \times W/(\{0\} \times W) \approx V$.

Example 7.8.5. Consider $P_2(\mathbb{C}) = \{ax^2 + bx + c \mid a, b, c \in \mathbb{C}\}$. Consider the subspace of $P_2(\mathbb{C})$ defined as $V = \{f(x) \in P_2(\mathbb{C}) \mid f(i) = 0\}$. Let's find an isomorphism to \mathbb{C}^n for appropriate n. Let $f(x) = ax^2 + bx + c \in V$ and calculate

$$f(i) = a(i)^2 + bi + c = -a + bi + c = 0 \quad \Rightarrow \quad c = a - bi$$

Thus, $f(x) = ax^2 + bx + a - bi = a(x^2 + 1) + b(x - i)$. The isomorphism from V to \mathbb{C}^2 is apparent from the calculation above. If $f(x) \in V$ then we can write $f(x) = a(x^2 + 1) + b(x - i)$ and

$$\Psi(f(x)) = \Psi(a(x^2 + 1) + b(x - i)) = (a, b).$$

The inverse map is also easy to find: $\Psi^{-1}(a,b) = a(x^2+1) + b(x-i)$

Example 7.8.6. Consider $\mathcal{F}(\mathbb{R})$ the set of all functions on \mathbb{R} . Observe, any function can be written as a sum of an even and odd function:

$$f(x) = \frac{1}{2} \left(f(x) + f(-x) \right) + \frac{1}{2} \left(f(x) - f(-x) \right)$$

Furthermore, if we denote the subspaces of even and odd functions as $\mathcal{F}_{even} \leq \mathcal{F}(\mathbb{R})$ and $\mathcal{F}_{odd} \leq \mathcal{F}(\mathbb{R})$ and note $\mathcal{F}_{even} \cap \mathcal{F}_{odd} = \{0\}$ hence $\mathcal{F}(\mathbb{R}) = \mathcal{F}_{even} \oplus \mathcal{F}_{odd}$. Consider the projection $T : \mathcal{F}(\mathbb{R}) \to \mathcal{F}_{even}$ clearly $Null(T) = \mathcal{F}_{odd}$ hence by the first isomorphism theorem, $\mathcal{F}(\mathbb{R})/\mathcal{F}_{odd} \approx \mathcal{F}_{even}$.

Example 7.8.7. Let $\Psi(f(x), g(x)) = f(x) + x^{n+1}g(x)$ note this defines an isomorphism of $P_n \times P_n$ and P_{2n+1} . For example, n = 1,

$$\Psi((ax+b,cx+d)) = ax+b+x^2(cx+d) = cx^3 + dx^2 + ax + b.$$

The reason we need 2n+1 is just counting: $\dim(P_n) = n+1$ and $\dim(P_n \times P_n) = 2(n+1)$. However, $\dim(P_{2n+1}) = (2n+1)+1$.

Example 7.8.8. Let $V = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and $W = \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$. Transposition gives us a natural isomorphism as follows: for each $L \in V$ there exists $A \in \mathbb{R}^{m \times n}$ for which $L = L_A$. However, to $A^T \in \mathbb{R}^{n \times m}$ there naturally corresponds $L_{A^T} : \mathbb{R}^m \to \mathbb{R}^n$. Since V and W are spaces of functions an isomorphism is conveniently given in terms $A \mapsto L_A$ isomorphism of $\mathbb{R}^{m \times n}$ and $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$: in particular $\Psi : V \to W$ is given by:

$$\Psi(L_A) = L_{A^T}.$$

To write this isomorphism without the use of the L_A notation requires a bit more thought. Take off your shoes and socks, but them back on, then write what follows. Let $S \in V$ and $x \in \mathbb{R}^m$,

$$(\Psi(S))(x) = (x^T[S])^T = [S]^T x = L_{[S]^T}(x).$$

Since the above holds for all $x \in \mathbb{R}^m$ it can be written as $\Psi(S) = L_{[S]^T}$.

The interested reader might appreciate the example below shows Theorem 7.5.2 in action.

Example 7.8.9. Let $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 3 & 0 \end{bmatrix}$ find an isomorphism from Null(A) to Col(A). As we recall, the CCP reveals all, we can easily calculate:

$$rref(A) = \left[\begin{array}{rrr} 1 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{array} \right]$$

Null space is $x \in \mathbb{R}^4$ for which Ax = 0 hence $x_1 = -x_2 - 3x_4$ and $x_3 = 2x_4$ with x_2, x_4 free. Thus,

$$x = (-x_2 - 3x_4, x_2, 2x_4, x_4) = x_2(-1, 1, 0, 0) + x_4(-3, 0, 2, 1)$$

and we find $\beta_N = \{(-1, 1, 0, 0), (-3, 0, 2, 1)\}$ is basis for Null(A). On the other hand $\beta_C = \{(1, 2), (1, 3)\}$ forms a basis for the column space by the CCP. Let Ψ : Null(A) \rightarrow Col(A) be defined by extending

$$\Psi((-1,1,0,0)) = (1,2)$$
 & $\Psi((-3,0,2,1)) = (1,3)$

linearly. In particular, if $x \in Null(A)$ then $\Psi(x) = x_2(1,2) + x_4(1,3)$. Fun fact, with our choice of basis the matrix $[\Psi]_{\beta_N,\beta_C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

The interested reader may also note that whenever we form a linear transformation $T: V \to W$ be mapping the *j*-th β basis element of V to the *j*-th γ basis element of W this gives a blockidentity matrix in $[T]_{\beta,\gamma}$. If $\#(\beta) = \#(\gamma)$ then, as in the above example, the matrix of T is simply $[T]_{\beta,\gamma} = I$. However, if dim $(W) > \dim(V)$ then the other blocks of the matrix are zero as by construction we already mapped all non-trivial parts of V to the first *j*-dimensions of W. The remaining dim(W) - j dimensions are untouched by T as we construct it. If T is instead given and our problem is to find bases for V and W for which the matrix is all zero with a identity matrix block in the upper left block then we must choose a basis carefully as described in Section 7.5. Let us return to the considerably easier problem of constructing isomorphisms between given vector spaces. The simplest advice is just, find a basis for each space and map one to the other. I find that is a good approach for many problems. Of course, there are other tools, but first the basics. **Example 7.8.10.** Let $V = P_2$ and $W = \{f(x) \in y \mid f(1) = 0\}$. By the factor theorem of algebra we know $f(x) \in W$ implies f(x) = (x - 1)g(x) +where $g(x) \in P_2$. Define, $\Psi(f(x)) = g(x)$ where g(x)(x - 1) = f(x). We argue that Ψ is an isomorphism. Note $\Psi^{-1}(g(x)) = (x - 1)g(x)$ and it is clear that $(x - 1)g(x) \in W$ moreover, linearity of Ψ^{-1} is simply seen from the calculation below:

$$\Psi^{-1}(cg(x) + h(x)) = (x - 1)(cg(x) + h(x)) = c(x - 1)g(x) + (x - 1)g(x) = c\Psi^{-1}(g(x)) + \Psi^{-1}(h(x)) = c\Psi^{-1}(g(x)) + e\Psi^{-1}(h(x)) = c\Psi^{-1}(g(x)) + e\Psi^{-1}(g(x)) = c\Psi^{-1}(g(x)) + e\Psi^{-1}(g(x)) = c\Psi^{-1}(g(x)) = c\Psi^{-1}(g(x)) + e\Psi^{-1}(g(x)) = c\Psi^{-1}(g(x)) = c$$

Linearity of Ψ follows by Theorem 7.2.14 as $\Psi = (\Psi^{-1})^{-1}$. Thus $V \approx W$.

You might note that I found a way around using a basis in the last example. Perhaps it is helpful to see the same example done by the basis mapping technique.

Example 7.8.11. Let $V = P_2$ and $W = \{f(x) \in y \mid f(1) = 0\}$. Ignoring the fact we know the factor theorem, let us find a basis the hard way: if $f(x) = ax^3 + bx^2 + cx + d \in W$ then

$$f(1) = a + b + c + d = 0$$

Thus, d = -a - b - c and

$$f(x) = a(x^{3} - 1) + b(x^{2} - 1) + c(x - 1)$$

We find basis $\beta = \{x^3 - 1, x^2 - 1, x - 1\}$ for W. Define $\phi: W \to P_2$ by linearly extending:

$$\phi(x^3 - 1) = x^2$$
, $\phi(x^2 - 1) = x$, $\phi(x - 1) = 1$.

In this case, a moments reflection reveals:

$$\phi^{-1}(ax^2 + bx + c) = a(x^3 - 1) + b(x^2 - 1) + c(x - 1).$$

Again, these calculations serve to prove $W \approx P_2$.

It might be interesting to relate the results of Example 7.8.10 and Example 7.8.11. Examing the formula for $\Psi^{-1}(g(x)) = (x-1)g(x)$ it is evident that we should factor out (x-1) from our ϕ^{-1} formula to connect to the Ψ^{-1} formula,

$$\phi^{-1}(ax^2 + bx + c) = a(x - 1)(x^2 + x + 1) + b(x - 1)(x + 1) + c(x - 1).$$

= $(x - 1)[a(x^2 + x + 1) + b(x + 1) + c]$
= $(x - 1)[ax^2 + (a + b)x + a + b + c]$
= $\Psi^{-1}(ax^2 + (a + b)x + a + b + c).$

Evaluating the equation above by Ψ yields $(\Psi \circ \phi^{-1})(ax^2 + bx + c) = ax^2 + (a+b)x + a + b + c$. Therefore, if $\gamma = \{x^2, x, 1\}$ then we may easily deduce

$$[\Psi \circ \phi^{-1}]_{\gamma,\gamma} = \begin{bmatrix} 1 & 0 & 0\\ 1 & 1 & 0\\ 1 & 1 & 1 \end{bmatrix}$$

Example 7.8.12. Let $V = \mathbb{C}$ and $M_{\mathbb{C}}$ the set of matrices of the form: $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. Observe that the map $\Psi(a+ib) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ is a linear transformation with inverse Ψ^{-1} . $\left(\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \right) = a+ib$. Therefore, V and $M_{\mathbb{C}}$ are isomorphic as vector spaces.

Let me continue past the point of linear isomorphism. In the example above, we can show that V and $M_{\mathbb{C}}$ are isomorphic as **algebras** over \mathbb{R} . In particular, notice

$$(a+ib)(c+id) = ac - bd + i(ad + bc) \& \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} ac - bd & -(ad + bc) \\ ad + bc & ac - bd \end{bmatrix}.$$

As you can see the pattern of the multiplication is the same. To be precise,

$$\Psi(\underbrace{(a+ib)(c+id)}_{\text{complex multiplication}}) = \underbrace{\Psi(a+ib)\Psi(c+id)}_{\text{matrix multiplication}}$$

These special 2×2 matrices form a **representation** of the complex numbers. Incidentally, you can prove there is no \mathbb{R} -algebra isomorphism to the algebras described in Examples 7.7.6 and 7.7.7. In contrast, $\mathbb{R} \oplus i\mathbb{R}$, $\mathbb{R} \oplus j\mathbb{R}$ and $\mathbb{R} \oplus i\mathbb{R}$ are all linearly isomorphic. The term **isomorphism** has wide application in mathematics. In this course, the unqualified term "isomorphism" would be more descriptively termed "linear-isomorphism". An isomorphism of \mathbb{R} -algebras is a linear isomorphism which also preserves the multiplication \star of the algebra; $\Psi(v \star w) = \Psi(v)\Psi(w)$. Another related concept, a non-associative algebra on a vector space which is a generalization of the cross-product of vectors in \mathbb{R}^3 is known as¹⁸ a Lie Algebra. In short, a Lie Algebra is a vector space paired with a Lie bracket. A Lie algebra isomorphism is a linear isomorphism which also preserves the Lie bracket; $\Psi([v,w]) = [\Psi(v), \Psi(w)]$. Not all isomorphisms are linear isomorphisms. For example, in abstract algebra you will study *isomorphisms of groups* which are bijections between groups which preserves the group multiplication. My point is just this, the idea of isomorphism, our current endeavor, is one you will see repeated as you continue your study of mathematics. To quote a certain show: *it has happened before, it will happen again*.

¹⁸it is pronounced "Lee", not what Obama does

Part III applications

Chapter 8

determinants

In this chapter we motivate the determinant of a matrix as a simple criteria to judge the invertibility of a given square matrix. Once the definition is settled we prove a series of useful proposition to simplify the computations of determinants. We use the determinant to give an abstraction of length, area and volume to *n*-volume¹. In addition, the determinant serves to construct Cramer's Rule which gives us a formula to solve systems with a unique solution. Then, a formula for the inverse of a matrix itself is obtained via the transpose of matrix of cofactors rescaled by division of the determinant. Finally, we pause to again give a long list of equivalent conditions for invertibility or singularity of an $n \times n$ matrix. The determinant finds an important place on that list as there are many problems one can ask which are shockingly simple to answer with determinants and yet confound in the other approaches.

I should warn you there are some difficult calculations in this Chapter. However, the good news is these are primarily to justify the various properties of the determinant. I probably will not present these in lecture because the method used to prove them is not generally of interest in this course. Index manipulation and even the elementary matrix arguments are a means to an end in this chapter. That said, I do hope you read them so you can appreciate the nature of the tool when you use it. For example, when you solve a problem using det(AB) = det(A)det(B) you should realize that is a nontrivial algebraic step. That move carries with it the full force of the arguments we see in this chapter.

8.1 a criteria for invertibility

In this section we study the problem of invertibility and in the process we discover the formulas for the determinant of 1×1 , 2×2 and 3×3 matrices. With that settled, I give a general definition which applies to arbitrary n and we conclude the section by stating the formulas which are often used for explicit calculation. Much of this section is an attempt at motivating the definition. Fortunately, determinants have no feelings, so, if you don't understand where they come from, you can still work with them just the same.

We have studied a variety of techniques to ascertain the invertibility of a given matrix. Recall, if A is an $n \times n$ invertible matrix then Ax = b has a unique solution $x = A^{-1}b$. Alternatively, $\operatorname{rref}(A) = I$. We now seek some explicit formula in terms of the components of A. Ideally this

 $^{^{1}}$ a good slogan for the determinant is just this: the determinant gives the volume. Or more precisely, the determinant of a matrix is the volume subtended by the convex hull of its columns.

formula will **determine** if A is invertible or not.

The base case n = 1 has $A = a \in \mathbb{R}$ as we identify $\mathbb{R}^{1 \times 1}$ with \mathbb{R} . The equation ax = b has solution x = b/a provided $a \neq 0$. Thus, the simple criteria in the n = 1 case is merely that $a \neq 0$.

The n = 2 case has $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. We learned that the formula for the 2 × 2 inverse is: $A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$

The necessary and sufficient condition for invertibility here is just that $ad - bc \neq 0$. That said, it may be helpful to derive this condition from row reduction. For brevity of discussion² we assume $a, c \neq 0$.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \xrightarrow{cr_1, ar_2} \begin{bmatrix} ac & bc \\ ac & ad \end{bmatrix} \xrightarrow{r_2 - r_1} \begin{bmatrix} ac & bc \\ 0 & ad - bc \end{bmatrix}$$

Observe that $ad - bc \neq 0$ is a necessary condition to reduce the matrix A to the identity.

The
$$n = 3$$
 case has $A = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}$. I assume here for brevity that $a, b, c, d, e, f \neq 0$

$$A = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix} \xrightarrow{bcr_1, acr_2, abr_3} \begin{bmatrix} abc & dbc & gbc \\ acb & abf & abi \end{bmatrix}$$

$$\xrightarrow{r_2 - r_1, r_3 - r_1} \begin{bmatrix} abc & dbc & gbc \\ 0 & b(af - dc) & b(ai - gc) \end{bmatrix}$$

$$\xrightarrow{r_1/(bc), r_2/c, r_3/b} \begin{bmatrix} a & d & g \\ 0 & af - dc & ai - gc \end{bmatrix}$$

$$\xrightarrow{r_2/(ae - db)} \begin{bmatrix} a & d & g \\ 0 & af - dc & ai - gc \end{bmatrix}$$

$$\xrightarrow{r_3 - (af - dc)r_2} \begin{bmatrix} a & d & g \\ 0 & 1 & af - gb \\ 0 & 0 & af - gc \end{bmatrix}$$

$$\xrightarrow{(ae - db)r_3} \begin{bmatrix} a & d & g \\ 0 & 1 & af - gc \\ 0 & 0 & af - gc \end{bmatrix}$$

Apparently, we need $(ai - gc)(ae - db) - (af - dc)(ah - gb) \neq 0$. Let's see if we can simplify it, $(ai - gc)(ae - db) - (af - dc)(ah - gb) = a^{2}ie - aidb - gcae + gcdb - a^{2}fh + afgb + dcah - dcgb$ = a[aie - idb - gce - afh + fgb + dch]

 $^{^{2}}$ you could break into further cases if you want a more complete motivating discussion, our current endeavor is to explain why the determinant formula is natural

We already assumed $a \neq 0$ so it is most interesting to require:

$$aie - idb - gce - afh + fgb + dch \neq 0$$

The condition above would seem to yield invertibility of A. To be careful, the calculation above does not prove anything about matrices for which the above row operations are forbidden. Technically, you'd need to examine those cases separately to prove the boxed criteria suffices for invertibility of A. We take a different, less direct, approach in this chapter. That said, perhaps this section helps motivate why we define the following **determinants**:

$$\det[a] = a,$$
$$\det\begin{bmatrix}a & b\\c & d\end{bmatrix} = ad - bc,$$
$$\det\begin{bmatrix}a & d & g\\b & e & h\\c & f & i\end{bmatrix} = aie - idb - gce - afh + fgb + dch$$

If $x \neq 0$ then $-x \neq 0$ thus the invertibility criteria alone does not suffice to uniquely determine the determinant. We'll see in a later section that the choice of sign has geometric significance. If a set of n-1 vectors $v_1, \ldots v_{n-1}$ forms a hyperplane in \mathbb{R}^n and we consider $\det[v_1|\cdots|v_n|w]$ for some vector w then the determinant is positive if w is one one side of the hyperplane and it is negative if w is one the other side. If w is on the hyperplane then the determinant is zero. These facts serve to determine the definition given below.

Before I state the definition, I'll pause to note a few additional features of the invertibility criteria we derived thus far. You might notice the formulas we have derived are homogeneous n-th order polynomials in the components of the matrix. However, they are peculiar in that no component is repeated. Each component appears at most once in each summand of the formula. Furthermore, there is a balance between the number of positive and negative signs in the formula and the number of summands is n! for each case.

The precise definition of the determinant is intrinsically combinatorial. A permutation $\sigma : \mathbb{N}_n \to \mathbb{N}_n$ is a bijection. Every permutation can be written as a product of an even or odd composition of transpositions. The $sgn(\sigma) = 1$ if σ is formed from an even product of transpositions. The $sgn(\sigma) = -1$ if σ is formed from an odd product of transpositions. The sum below is over all possible permutations,

$$det(A) = \sum_{\sigma} sgn(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)}$$

this provides an explicit definition of the determinant. For example, in the n = 2 case we have $\sigma_o(x) = x$ or $\sigma_1(1) = 2, \sigma_1(2) = 1$. The sum over all permutations has just two terms in the n = 2 case,

$$det(A) = sgn(\sigma_o)A_{1\sigma_o(1)}A_{2\sigma_o(2)} + sgn(\sigma_1)A_{1\sigma_1(1)}A_{2\sigma_1(2)} = A_{11}A_{22} - A_{12}A_{21}$$

In the notation $A_{11} = a, A_{12} = b, A_{21} = c, A_{22} = d$ the formula above says det(A) = ad - bc.

Pure mathematicians tend to prefer the definition above to the one I am preparing below. I would argue mine has the advantage of not summing over functions. My sums are simply over integers. The calculations I make in the proofs in this Chapter may appear difficult to you, but if you gain a little more experience with index calculations I think you would find them accessible. I will not go over them all in lecture. I would recommend you at least read over them.

Definition 8.1.1.

Let $\epsilon_{i_1i_2...i_n}$ be defined to be the completely antisymmetric symbol in *n*-indices. We define $\epsilon_{12...n} = 1$ then all other values are generated by demanding the interchange of any two indices is antisymmetric. This is also known as the **Levi-Civita** symbol. In view of this notation, we define the **determinant** of $A \in \mathbb{R}^{n \times n}$ as follows:

$$det(A) = \sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, i_2, \dots, i_n} A_{1i_1} A_{2i_2} \cdots A_{ni_n}.$$

Direct implementation of the formula above is straightforward, but, tedious.

Example 8.1.2. I prefer this definition. I can actually calculate it faster, for example the n = 3 case is pretty quick:

$$det(A) = \epsilon_{123}A_{11}A_{22}A_{33} + \epsilon_{231}A_{12}A_{23}A_{31} + \epsilon_{312}A_{13}A_{21}A_{32} + \epsilon_{321}A_{13}A_{22}A_{31} + \epsilon_{213}A_{12}A_{21}A_{33} + \epsilon_{132}A_{11}A_{23}A_{32}$$

In principle there are 27 terms above but only these 6 are nontrivial because if any index is repeated the ϵ_{ijk} is zero. The only nontrivial terms are $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$ and $\epsilon_{321} = \epsilon_{213} = \epsilon_{132} = -1$. Thus,

$$det(A) = A_{11}A_{22}A_{33} + A_{12}A_{23}A_{31} + A_{13}A_{21}A_{32}$$
$$-A_{13}A_{22}A_{31} - A_{12}A_{21}A_{33} - A_{11}A_{23}A_{32}$$

There is a cute way to remember this formula by crossing diagonals in the matrix twice written.

Cute-tricks aside, we more often find it convenient to use Laplace's expansion by minor formulae to actually calculate explicit determinants. I'll postpone proof of the equivalence with the definition until Section 8.3 where you can see the considerable effort which is required to connect the formulas.³ These formulas show you how to calculate determinants of $n \times n$ matrices as an alternating sum of $(n-1) \times (n-1)$ matrix determinants. I'll begin with the 2×2 case,

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Then the 3×3 formula is:

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a \cdot \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \cdot \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \cdot \det \begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

and finally the 4×4 determinant is given by

,

$$\det \begin{pmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{pmatrix} = a \cdot \det \begin{pmatrix} f & g & h \\ j & k & l \\ n & o & p \end{pmatrix} - b \cdot \det \begin{pmatrix} e & g & h \\ i & k & l \\ m & o & p \end{pmatrix}$$
(8.1)

$$+ c \cdot \det \begin{pmatrix} e & f & h \\ i & j & l \\ m & n & p \end{pmatrix} - d \cdot \det \begin{pmatrix} e & f & g \\ i & j & k \\ m & n & o \end{pmatrix}$$
(8.2)

³those are probably the most difficult calculations contained in these notes.

8.2 determinants and geometry

What do these determinant formulas have to do with geometry? In this section I showcase a variety of examples, if you have not had Calculus III then please don't despair. This section is mostly motivational.

Example 8.2.1. Consider the vectors $\langle l, 0 \rangle$ and $\langle 0, w \rangle$. They make two sides of a rectangle with length l and width w. Notice

$$det \left[\begin{array}{cc} l & 0 \\ 0 & w \end{array} \right] = lw.$$

In contrast,

$$det \left[\begin{array}{cc} 0 & w \\ l & 0 \end{array} \right] = -lw.$$

Interestingly this works for parallellograms with sides $\langle a, b \rangle$ and $\langle c, d \rangle$ the area is given by $\pm det \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.



Maybe you can see it better in the diagram below: the point is that triangles T1 and T2 match nicely but the T3 is included in the red rectangle but is excluded from the green parallelogram. The area of the red rectangle A_1B_2 less the area of the blue square A_2B_1 is precisely the area of the green parallelogram.



Perhaps you recall from calculus III that we learned a parallelogram with sides \vec{A}, \vec{B} can be parametrized by $\vec{r}(u, v) = u\vec{A} + v\vec{B}$. We have $\vec{A} = (a, b, 0)$ and $\vec{B} = (c, d, 0)$ if you view the parallelogram from a three dimensional perspective. Moreover,

$$\vec{A} \times \vec{B} = det \begin{bmatrix} e_1 & e_2 & e_3 \\ a & b & 0 \\ c & d & 0 \end{bmatrix} = (ad - bc)e_3.$$

The sign of ad - bc indicates the **orientation** of the paralellogram. If the paralellogram lives in the xy-plane then it has an up-ward pointing normal if the determinant is positive whereas it has a downward pointing normal if the determinant is negative.

Example 8.2.2. If we look at a three dimensional box with vectors $\vec{A}, \vec{B}, \vec{C}$ pointing along three edges with from a common corner then it can be shown that the volume V is given by the determinant

$$V=\pm det \begin{bmatrix} \vec{A} \\ \vec{B} \\ \hline \vec{C} \end{bmatrix}$$

Of course it's easy to see that V = lwh if the sides have length l, width w and height h. However, this formula is more general than that, it also holds if the vectors lie along a paralell piped. Again the sign of the determinant has to do with the **orientation** of the box. If the determinant is positive then that means that the set of vectors $\{\vec{A}, \vec{B}, \vec{C}\}$ forms a righted-handed set of vectors. In terms of calculus III, \vec{C} and $\vec{A} \times \vec{B}$ both point off the same side of the plane containing \vec{A} and \vec{B} ; the ordering of the vectors is roughly consistent with the right-hand rule. If the determinant of the three vectors is negative then they will be consistent with the (inferior and evil) left-hand rule. I say "roughly" because $\vec{A} \times \vec{B}$ need not be parallel with \vec{C} .

If you study the geometry of cross and dot products it is not too hard to see that $V = |\vec{A} \cdot (\vec{B} \times \vec{C})|$. This formula is easy to reproduce,

$$det \begin{bmatrix} A_1 & A_2 & A_3 \\ \hline B_1 & B_2 & B_3 \\ \hline C_1 & C_2 & C_3 \end{bmatrix} = A_1(B_2C_3 - B_3C_2) + A_2(B_1C_3 - B_3C_1) + A_3(B_1C_2 - B_2C_1)$$
$$= \vec{A} \cdot (\vec{B} \times \vec{C}).$$

If you'd like to know more about the geometry of cross products then you should take calculus III and read more than the mainstream required calculus text. It is interesting that the determinant gives formulas for cross products and the so-called "triple product" above.

Example 8.2.3. To calculate the cross-product of \vec{A} and \vec{B} we can use the heuristic rule

$$\vec{A} \times \vec{B} = det \begin{bmatrix} e_1 & e_2 & e_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{bmatrix}$$

technically this is not a real "determinant" because there are vectors in the top row but numbers in the last two rows.

I hope the n = 2 and n = 3 cases help motivate the definition which follows.

Definition 8.2.4.

Let v_1, v_2, \ldots, v_n be vectors in \mathbb{R}^n then the *n*-volume of the *n*-piped *P* with edges v_1, v_2, \ldots, v_n is given by

 $\operatorname{Vol}(P) = \det \left[v_1 | v_2 | \cdots | v_n \right].$

Notice the terminology *n*-volume includes area as the n = 2 case and ordinary spatial volume as n = 3. Also, as a check on the definition above, if we consider the unit *n*-cube in \mathbb{R}^n it is $P = [0, 1]^n$ and we calculate⁴:

$$Vol(P) = det [e_1|e_2|\cdots|e_n] = det(I) = 1.$$

On the other hand, we also learn in a later section that if any column is repeated the determinant is zero. This matches intuition as you imagine an *n*-rectangle, if two edges from a common vertex are colinear than it's not actually an *n*-dimensional rectangle so we'd say its *n*-volume is zero. For example, a line-segment has zero area, a two-dimensional rectangle has zero 3-volume.

8.3 cofactor expansion for the determinant

The Levi-Civita definition of the determinant of an $n \times n$ matrix A is:

$$det(A) = \sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, i_2, \dots, i_n} A_{1i_1} A_{2i_2} \cdots A_{ni_n}.$$

This is our definition for the determinant. All other facts flow from that source. In some other texts, the **cofactor expansion** of the determinant is given as the definition. I already recorded the standard cofactor expansions for determinants up to order 4 in the first section of this chapter. The aim of this section is to describe the general cofactor expansions and to prove they give another equivalent characterization of the determinant.

Definition 8.3.1.

Let $A = [A_{ij}] \in \mathbb{R}^{n \times n}$. The minor of A_{ij} is denoted M_{ij} which is defined to be the determinant of the $\mathbb{R}^{(n-1)\times(n-1)}$ matrix formed by deleting the *i*-th column and the *j*-th row of A. The (i, j)-th co-factor of A is $C_{ij} = (-1)^{i+j} M_{ij}$.

Theorem 8.3.2.

The determinant of $A \in \mathbb{R}^{n \times n}$ can be calculated from a sum of cofactors either along any row or column;

- 1. $det(A) = A_{i1}C_{i1} + A_{i2}C_{i2} + \dots + A_{in}C_{in}$ (*i*-th row expansion)
- 2. $det(A) = A_{1j}C_{1j} + A_{2j}C_{2j} + \dots + A_{nj}C_{nj}$ (*j*-th column expansion)

⁴sorry, putting the cart before the horse here, we learn det(I) = 1 in future section

Proof: I'll attempt to sketch a proof of (2.) directly from the general definition. Let's try to identify A_{1i_1} with A_{1j} then A_{2i_2} with A_{2j} and so forth, keep in mind that j is a fixed but arbitrary index, it is not summed over.

$$det(A) = \sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, i_2, \dots, i_n} A_{1i_1} A_{2i_2} \cdots A_{ni_n}$$

$$= \sum_{i_2, \dots, i_n} \epsilon_{j, i_2, \dots, i_n} A_{1j} A_{2i_2} \cdots A_{ni_n} + \sum_{i_1 \neq j, i_2, \dots, i_n} \epsilon_{i_1, i_2, \dots, i_n} A_{1i_1} A_{2i_2} \cdots A_{ni_n}$$

$$= \sum_{i_2, \dots, i_n} \epsilon_{j, i_2, \dots, i_n} A_{1j} A_{2i_2} \cdots A_{ni_n} + \sum_{i_1 \neq j, i_3, \dots, i_n} \epsilon_{i_1, j, \dots, i_n} A_{1i_1} A_{2j} \cdots A_{ni_n}$$

$$+ \cdots + \sum_{i_1 \neq j, i_2 \neq j, \dots, i_{n-1} \neq j} \epsilon_{i_1, i_2, \dots, i_{n-1}, j} A_{1i_1} \cdots A_{n-1, i_{n-1}} A_{nj}$$

$$+ \sum_{i_1 \neq j, \dots, i_n \neq j} \epsilon_{i_1, \dots, i_n} A_{1i_1} A_{1i_2} \cdots A_{ni_n}$$

Consider the summand. If all the indices $i_1, i_2, \ldots, i_n \neq j$ then there must be at least one repeated index in each list of such indices. Consequently the last sum vanishes since $\epsilon_{i_1,\ldots,i_n}$ is zero if any two indices are repeated. We can pull out A_{1j} from the first sum, then A_{2j} from the second sum, and so forth until we eventually pull out A_{nj} out of the last sum.

$$det(A) = A_{1j} \left(\sum_{i_2, \dots, i_n} \epsilon_{j, i_2, \dots, i_n} A_{2i_2} \cdots A_{ni_n} \right) + A_{2j} \left(\sum_{i_1 \neq j, \dots, i_n} \epsilon_{i_1, j, \dots, i_n} A_{1i_1} \cdots A_{ni_n} \right) + \cdots + A_{nj} \left(\sum_{i_1 \neq j, i_2 \neq j, \dots, j \neq i_{n-1}} \epsilon_{i_1, i_2, \dots, j} A_{1i_1} A_{2i_2} \cdots A_{n-1, i_{n-1}} \right)$$

The terms appear different, but in fact there is a hidden symmetry. If any index in the summations above takes the value j then the Levi-Civita symbol with have two j's and hence those terms are zero. Consequently we can just as well take all the sums over all values **except** j. In other words, each sum is a completely antisymmetric sum of products of n - 1 terms taken from all columns except j. For example, the first term has an antisymmetrized sum of a product of n - 1 terms not including column j or row 1.Reordering the indices in the Levi-Civita symbol generates a sign of $(-1)^{1+j}$ thus the first term is simply $A_{1j}C_{1j}$. Likewise the next summand is $A_{2j}C_{2j}$ and so forth until we reach the last term which is $A_{nj}C_{nj}$. In other words,

$$det(A) = A_{1j}C_{1j} + A_{2j}C_{2j} + \dots + A_{nj}C_{nj}$$

The proof of (1.) is probably similar. We will soon learn that $det(A^T) = det(A)$ thus (2.) \implies (1.). since the *j*-th row of A^T is the *j*-th columns of A.

All that remains is to show why $det(A) = det(A^T)$. Recall $(A^T)_{ij} = A_{ji}$ for all i, j, thus

$$det(A^{T}) = \sum_{i_{1}, i_{2}, \dots, i_{n}} \epsilon_{i_{1}, i_{2}, \dots, i_{n}} (A^{T})_{1i_{1}} (A^{T})_{2i_{2}} \cdots (A^{T})_{ni_{n}}$$
$$= \sum_{i_{1}, i_{2}, \dots, i_{n}} \epsilon_{i_{1}, i_{2}, \dots, i_{n}} A_{i_{1}1} A_{i_{2}2} \cdots A_{i_{n}n}$$
$$= \sum_{i_{1}, i_{2}, \dots, i_{n}} \epsilon_{i_{1}, i_{2}, \dots, i_{n}} A_{1i_{1}} A_{2i_{2}} \cdots A_{ni_{n}} = det(A)$$

to make the last step one need only see that both sums contain all the same terms just written in a different order. Let me illustrate explicitly how this works in the n = 3 case,

$$det(A^{T}) = \epsilon_{123}A_{11}A_{22}A_{33} + \epsilon_{231}A_{21}A_{32}A_{13} + \epsilon_{312}A_{31}A_{12}A_{23} + \epsilon_{321}A_{31}A_{22}A_{13} + \epsilon_{213}A_{21}A_{12}A_{33} + \epsilon_{132}A_{11}A_{32}A_{23}$$

The I write the entries so the column indices go 1, 2, 3

$$det(A^{T}) = \epsilon_{123}A_{11}A_{22}A_{33} + \epsilon_{231}A_{13}A_{21}A_{32} + \epsilon_{312}A_{12}A_{23}A_{31} + \epsilon_{321}A_{13}A_{22}A_{31} + \epsilon_{213}A_{12}A_{21}A_{33} + \epsilon_{132}A_{11}A_{23}A_{32}$$

But, the indices of the Levi-Civita symbol are not in the right order yet. Fortunately, we have identities such as $\epsilon_{231} = \epsilon_{312}$ which allow us to reorder the indices without introducing any new signs,

$$det(A^{T}) = \epsilon_{123}A_{11}A_{22}A_{33} + \epsilon_{312}A_{13}A_{21}A_{32} + \epsilon_{231}A_{12}A_{23}A_{31} + \epsilon_{321}A_{13}A_{22}A_{31} + \epsilon_{213}A_{12}A_{21}A_{33} + \epsilon_{132}A_{11}A_{23}A_{32}$$

But, these are precisely the terms in det(A) just written in a different order (see Example 8.1.2). Thus $det(A^T) = det(A)$. I leave the details of how to reorder the order n sum to the reader. \Box

Remark 8.3.3.

Lay's text circumnavigates many of the difficulties I face in this chapter by using the cofactor definition as the definition of the determinant. One place you can also find a serious treatment of determinants is in *L*inear Algebra by Insel, Spence and Friedberg where you'll find the proof of the co-factor expansion is somewhat involved. However, the heart of the proof involves multilinearity. Multilinearity is practically manifest with our Levi-Civita definition. Anywho, a better definition for the determinant is as follows: **the determinant is the alternating**, *n*-**multilinear**, **real valued map such that** det(I) = 1. It can be shown this uniquely defines the determinant. All these other things like permutations and the Levi-Civita symbol are just notation. Example 8.3.4. I suppose it's about time for an example. Let

$$A = \left[\begin{array}{rrrr} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{array} \right]$$

I usually calculate by expanding across the top row out of habit,

$$det(A) = 1det \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} - 2det \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} + 3det \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix}$$
$$= 1(45 - 48) - 2(36 - 42) + 3(32 - 35)$$
$$= -3 + 12 - 9$$
$$= 0.$$

Now, we could also calculate by expanding along the middle row,

$$det(A) = -4det \begin{bmatrix} 2 & 3 \\ 8 & 9 \end{bmatrix} + 5det \begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix} - 6det \begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix}$$
$$= -4(18 - 24) + 5(9 - 21) - 6(8 - 14)$$
$$= 24 - 60 + 36$$
$$= 0.$$

Many other choices are possible, for example expan along the right column,

$$det(A) = 3det \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix} - 6det \begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix} + 9det \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}$$
$$= 3(32 - 35) - 6(8 - 14) + 9(5 - 8)$$
$$= -9 + 36 - 27$$
$$= 0.$$

which is best? Certain matrices might have a row or column of zeros, then it's easiest to expand along that row or column. Calculation completed, let's pause to appreciate the geometric significance in view of Definition 8.2.4. Our calculations show that the parallel piped spanned by (1,2,3), (4,5,6), (7,8,9) is flat, it's actually just a two-dimensional parallelogram.

If you are curious about the area of the parallelogram implicit in the example above, you could calculate the cross-product of the columns and the length of the non-zero results would give you the area of the parallelogram. See Example 8.2.3 for the formula of the cross-product.

Example 8.3.5. Let's look at an example where we can exploit the co-factor expansion to greatly reduce the difficulty of the calculation. Let

$$A = \begin{bmatrix} 1 & 2 & 3 & 0 & 4 \\ 0 & 0 & 5 & 0 & 0 \\ 6 & 7 & 8 & 0 & 0 \\ 0 & 9 & 3 & 4 & 0 \\ -1 & -2 & -3 & 0 & 1 \end{bmatrix}$$

Begin by expanding down the 4-th column,

$$det(A) = (-1)^{4+4} M_{44} = 4det \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 5 & 0 \\ 6 & 7 & 8 & 0 \\ -1 & -2 & -3 & 1 \end{bmatrix}$$

Next expand along the 2-row of the remaining determinant,

$$det(A) = (4)(5(-1)^{2+3}M_{23}) = -20det \begin{bmatrix} 1 & 2 & 4 \\ 6 & 7 & 0 \\ -1 & -2 & 1 \end{bmatrix}$$

Finish with the trick for 3×3 determinants, it helps me to write out

then calculate the products of the three down diagonals and the three upward diagonals. Subtract the up-diagonals from the down-diagonals.

$$det(A) = -20(7 + 0 - 48 - (-28) - (0) - (12)) = -20(-25) = 500.$$

It is fun to note this is the 5-volume of the 5-piped region in \mathbb{R}^5 which has the columns of A as edges from a common vertex.

I will abstain from further geometric commentary for the most part in what follows. However, one last comment, it would be interesting to understand the geometric interpretation of the cofactor expansion. Note that it relates *n*-volumes to (n-1)-volumes.

8.4 properties of determinants

In this section we learn the most important properties of the determinant. A sequence of results born of elementary matrix arguments allows us to confirm that the motivating concept for the determinant is in fact true for arbitrary order; that is, Propositon 8.4.5 proves $det(A) \neq 0$ iff A^{-1} exists. It is important that you appreciate how the results of this section are accumulated through a series of small steps, each building on the last. However, it is even more important that you learn how the results of this section can be applies to a variety of matrix problems. Your exercises will help you in that direction naturally.

The properties given in the proposition below are often useful to greatly reduce the difficulty of a determinant calculation.

Proposition 8.4.1.

Let $A \in \mathbb{R}^{n \times n}$, 1. $det(A^T) = det(A)$, 2. If there exists j such that $row_j(A) = 0$ then det(A) = 0, 3. If there exists j such that $col_j(A) = 0$ then det(A) = 0, 4. $det[A_1|A_2|\cdots|aA_k+bB_k|\cdots A_n] = adet[A_1|\cdots|A_k|\cdots|A_n] + bdet[A_1|\cdots|B_k|\cdots|A_n]$, 5. $det(kA) = k^n det(A)$ 6. if $B = \{A : r_k \leftrightarrow r_j\}$ then det(B) = -det(A), 7. if $B = \{A : r_k + ar_j \to r_k\}$ then det(B) = det(A), 8. if $row_i(A) = krow_j(A)$ for $i \neq j$ then det(A) = 0where I mean to denote $r_k \leftrightarrow r_j$ as the row interchange and $r_k + ar_j \to r_k$ as a column addition and I assume k < j.

Proof: we already proved (1.) in the proof of the cofactor expansion Theorem 8.3.2. The proof of (2.) and (3.) follows immediately from the cofactor expansion if we expand along the zero row or column. The proof of (4.) is not hard given our Levi-Civita definition, let

$$C = [A_1|A_2|\cdots|aA_k + bB_k|\cdots|A_n]$$

Calculate from the definition,

$$det(C) = \sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, i_2, \dots, i_n} C_{1i_1} \cdots C_{ki_k} \cdots C_{ni_n}$$

= $\sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, i_2, \dots, i_n} A_{1i_1} \cdots (aA_{ki_k} + bB_{ki_k}) \cdots A_{ni_n}$
= $a \left(\sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, i_2, \dots, i_n} A_{1i_1} \cdots A_{ki_k} \cdots A_{ni_n} \right)$
+ $b \left(\sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, i_2, \dots, i_n} A_{1i_1} \cdots B_{ki_k} \cdots A_{ni_n} \right)$
= $a det[A_1|A_2| \cdots |A_k| \cdots |A_n] + b det[A_1|A_2| \cdots |B_k| \cdots |A_n]$

by the way, the property above is called multilinearity. The proof of (5.) is similar,

$$det(kA) = \sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, i_2, \dots, i_n} kA_{1i_1} kA_{2i_2} \cdots kA_{ni_n}$$

= $k^n \sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, i_2, \dots, i_n} A_{1i_1} A_{2i_2} \cdots A_{ni_n}$
= $k^n det(A)$

Let B be as in (6.), this means that $col_k(B) = col_j(A)$ and vice-versa,

$$det(B) = \sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, \dots, i_k, \dots, i_j, \dots, i_n} A_{1i_1} \cdots A_{ji_k} \cdots A_{ki_j} \cdots A_{ni_n}$$
$$= \sum_{i_1, i_2, \dots, i_n} -\epsilon_{i_1, \dots, i_j, \dots, i_k, \dots, i_n} A_{1i_1} \cdots A_{ji_k} \cdots A_{ki_j} \cdots A_{ni_n}$$
$$= -det(A)$$

where the minus sign came from interchanging the indices i_j and i_k .

To prove (7.) let us define B as in the Proposition: let $row_k(B) = row_k(A) + arow_j(A)$ and $row_i(B) = row_i(A)$ for $i \neq k$. This means that $B_{kl} = A_{kl} + aA_{jl}$ and $B_{il} = A_{il}$ for each l. Consequently,

$$det(B) = \sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, \dots, i_k, \dots, i_n} A_{1i_1} \cdots (A_{ki_k} + aA_{ji_k}) \cdots A_{ni_n}$$
$$= \sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, \dots, i_n} A_{1i_1} \cdots A_{ki_k} \cdots A_{ni_n}$$
$$+ a \bigg(\sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, \dots, i_j, \dots, i_k, \dots, i_n} A_{1i_1} \cdots A_{j, i_j} \cdots A_{ji_k} \cdots A_{ni_n} \bigg)$$
$$= \sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, \dots, i_n} A_{1i_1} \cdots A_{ki_k} \cdots A_{ni_n}$$
$$= det(A).$$

The term in parenthesis vanishes because it has the sum of an antisymmetric tensor in i_j , i_k against a symmetric tensor in i_j , i_k . Here is the pattern, suppose $S_{ij} = S_{ji}$ and $T_{ij} = -T_{ji}$ for all i, j then consider

$$\sum_{i} \sum_{j} S_{ij} T_{ij} = \sum_{j} \sum_{i} S_{ji} T_{ji}$$
 switched indices
$$= \sum_{j} \sum_{i} -S_{ij} T_{ij}$$
 used sym. and antisym.
$$= -\sum_{i} \sum_{j} S_{ij} T_{ij}$$
 interchanged sums.

thus we have $\sum S_{ij}T_{ij} = -\sum S_{ij}T_{ij}$ which indicates the sum is zero. We can use the same argument on the pair of indices i_j, i_k in the expression since $A_{ji_j}A_{ji_k}$ is symmetric in i_j, i_k whereas the Levi-Civita symbol is antisymmetric in i_j, i_k .

We get (8.) as an easy consequence of (2.) and (7.), just subtract one row from the other so that we get a row of zeros. \Box

Proposition 8.4.2.

The determinant of a diagonal matrix is the product of the diagonal entries.

Proof: Use multilinearity on each row,

$$det \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} = d_1 det \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} = \cdots = d_1 d_2 \cdots d_n det \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Thus $det(D) = d_1 d_2 \cdots d_n$ as claimed. \Box

Proposition 8.4.3.

Let *L* be a lower triangular square matric and *U* be an upper triangular square matrix. 1. $det(L) = L_{11}L_{22}\cdots L_{nn}$ 2. $det(U) = U_{11}U_{22}\cdots U_{nn}$

Proof: I'll illustrate the proof of (2.) for the 3×3 case. We use the co-factor expansion across the first column of the matrix to begin,

$$det \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{bmatrix} = A_{11}det \begin{bmatrix} U_{22} & U_{23} \\ 0 & U_{33} \end{bmatrix} = U_{11}U_{22}U_{33}$$

The proof of the $n \times n$ case is essentially the same. For (1.) use the co-factor expansion across the top row of L, to get $det(L) = L_{11}C_{11}$. Not the submatrix for calculating C_{11} is again has a row of zeros across the top. We calculate $C_{11} = L_{22}C_{22}$. This continues all the way down the diagonal. We find $det(L) = L_{11}L_{22}\cdots L_{nn}$. \Box

Proposition 8.4.4.

Let $A \in \mathbb{R}^{n \times n}$ and $k \neq 0 \in \mathbb{R}$, 1. $det(E_{r_i \leftrightarrow r_j}) = -1$, 2. $det(E_{kr_i \rightarrow r_i}) = k$, 3. $det(E_{r_i + br_j \rightarrow r_i}) = 1$, 4. for any square matrix B and elementary matrix E, det(EB) = det(E)det(B)5. if E_1, E_2, \dots, E_k are elementary then $det(E_1E_2 \cdots E_k) = det(E_1)det(E_2) \cdots det(E_k)$

Proof: Proposition 8.7.2 shows us that det(I) = 1 since $I^{-1} = I$ (there are many easier ways to show that). Note that $E_{r_i \leftrightarrow r_j}$ is a row-swap of the identity matrix thus by Proposition 8.4.1 we find $det(E_{r_i \leftrightarrow r_j}) = -1$. To prove (2.) we use multilinearity from Proposition 8.4.1. For (3.) we use multilinearity again to show that:

$$det(E_{r_i+br_i\to r_i}) = det(I) + bdet(E_{ij})$$

Again det(I) = 1 and since the unit matrix E_{ij} has a row of zeros we know by Proposition 8.4.1 $det(E_{ij}) = 0$.

To prove (5.) we use Proposition 8.4.1 multiple times in the arguments below. Let $B \in \mathbb{R}^{n \times n}$ and suppose E is an elementary matrix. If E is multiplication of a row by k then det(E) = kfrom (2.). Also EB is the matrix B with some row multiplied by k. Use multilinearity to see that det(EB) = kdet(B). Thus det(EB) = det(E)det(B). If E is a row interchange then EB is B with a row swap thus det(EB) = -det(B) and det(E) = -1 thus we again find det(EB) = det(E)det(B). Finally, if E is a row addition then EB is B with a row addition and det(EB) = det(B) and det(E) = 1 hence det(EB) = det(E)det(B). Notice that (6.) follows by repeated application of (5.). \Box

Proposition 8.4.5.

A square matrix A is invertible iff $det(A) \neq 0$.

Proof: recall there exist elementary matrices E_1, E_2, \ldots, E_k such that $rref(A) = E_1 E_2 \cdots E_k A$. Thus $det(rref(A)) = det(E_1)det(E_2) \cdots det(E_k)det(A)$. Either det(rref(A)) = 0 and det(A) = 0 or they are both nonzero.

Suppose A is invertible. Then Ax = 0 has a unique solution and thus rref(A) = I hence $det(rref(A)) = 1 \neq 0$ implying $det(A) \neq 0$.

Conversely, suppose $det(A) \neq 0$, then $det(rref(A)) \neq 0$. But this means that rref(A) does not have a row of zeros. It follows rref(A) = I. Therefore $A^{-1} = E_1 E_2 \cdots E_k$. \Box

Proposition 8.4.6.

If $A, B \in \mathbb{R}^{n \times n}$ then det(AB) = det(A)det(B).

Proof: If either A or B is not invertible then the reduced row echelon form of the nonivertible matrix will have a row of zeros hence det(A)det(B) = 0. Without loss of generality, assume A is not invertible. Note $rref(A) = E_1E_2\cdots E_kA$ hence $E_3^{-1}E_2^{-1}E_1^{-1}rref(A)B = AB$. Notice that rref(A)B will have at least one row of zeros since rref(A) has a row of zeros. Thus $det(E_3^{-1}E_2^{-1}E_1^{-1}rref(A)B) = det(E_3^{-1}E_2^{-1}E_1^{-1})det(rref(A)B) = 0$.

Suppose that both A and B are invertible. Then there exist elementary matrices such that $A = E_1 \cdots E_p$ and $B = E_{p+1} \cdots E_{p+q}$ thus

$$det(AB) = det(E_1 \cdots E_p E_{p+1} \cdots E_{p+q})$$

= $det(E_1 \cdots E_p)det(E_{p+1} \cdots E_{p+q})$
= $det(A)det(B).$

We made repeated use of (6.) in Proposition 8.4.4. \Box

Proposition 8.4.7.

If $A \in \mathbb{R}^{n \times n}$ is invertible then $det(A^{-1}) = \frac{1}{det(A)}$.

Proof: If A is invertible then there exists $A^{-1} \in \mathbb{R}^{n \times n}$ such that $AA^{-1} = I$. Apply Proposition 8.4.6 to see that

$$det(AA^{-1}) = det(A)det(A^{-1}) = det(I) \implies det(A)det(A^{-1}) = 1.$$

Thus, $det(A^{-1}) = 1/det(A)$

Many of the properties we used to prove det(AB) = det(A)det(B) are easy to derive if you were simply given the assumption det(AB) = det(A)det(B). When you look at what went into the proof of Proposition 8.4.6 it's not surprising that det(AB) = det(A)det(B) is a powerful formula to know.

Proposition 8.4.8.

If A is block-diagonal with square blocks A_1, A_2, \dots, A_k then $det(A) = det(A_1)det(A_2)\cdots det(A_k).$

Proof: for a 2 × 2 matrix this is clearly true since a block diagonal matrix is simply a diagonal matrix. In the 3 × 3 nondiagonal case we have a 2 × 2 block A_1 paired with a single diagonal entry A_2 . Simply apply the cofactor expansion on the row of the diagonal entry to find that $det(A) = A_2det(A_1) = det(A_2)det(A_1)$. For a 4 × 4 we have more cases but similar arguments apply. I leave the general proof to the reader. \Box

Example 8.4.9. If $M = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ is a block matrix where A, B are square blocks then det(M) = det(A)det(B).

8.5 examples of determinants

In the preceding section we saw the derivation of determinant properties requires some effort. Thankfully, the use of the properties to solve problems typically takes much less effort.

Example 8.5.1. Notice that row 2 is twice row 1,

$$det \left[\begin{array}{rrr} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 7 & 8 & 9 \end{array} \right] = 0.$$

Example 8.5.2. To calculate this one we make a single column swap to get a diagonal matrix. The determinant of a diagonal matrix is the product of the diagonals, thus:

$$det \begin{bmatrix} 0 & 6 & 0 & 0 & 0 & 0 \\ 8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = -det \begin{bmatrix} 6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = 48.$$

Example 8.5.3. I choose the column/row for the co-factor expansion to make life easy each time:

$$det \begin{bmatrix} 0 & 1 & 0 & 2 \\ 13 & 71 & 5 & \pi \\ 0 & 3 & 0 & 4 \\ -2 & e & 0 & G \end{bmatrix} = -5det \begin{bmatrix} 0 & 1 & 2 \\ 0 & 3 & 4 \\ -2 & e & G \end{bmatrix}$$
$$= -5(-2)det \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$
$$= 10(4-6)$$
$$= -20.$$

Example 8.5.4. Find the values of λ such that the matrix $A - \lambda I$ is singular given that

$$A = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

The matrix $A - \lambda I$ is singular iff $det(A - \lambda I) = 0$,

$$det(A - \lambda I) = det \begin{bmatrix} 1 - \lambda & 0 & 2 & 3\\ 1 & -\lambda & 0 & 0\\ 0 & 0 & 2 - \lambda & 0\\ 0 & 0 & 0 & 3 - \lambda \end{bmatrix}$$
$$= (3 - \lambda)det \begin{bmatrix} 1 - \lambda & 0 & 2\\ 1 & \lambda & 0\\ 0 & 0 & 2 - \lambda \end{bmatrix}$$
$$= (3 - \lambda)(2 - \lambda)det \begin{bmatrix} 1 - \lambda & 0\\ 1 & \lambda \end{bmatrix}$$
$$= (3 - \lambda)(2 - \lambda)(1 - \lambda)(-\lambda)$$
$$= \lambda(\lambda - 1)(\lambda - 2)(\lambda - 3)$$

Thus we need $\lambda = 0, 1, 2$ or 3 in order that $A - \lambda I$ be a noninvertible matrix. These values are called the **eigenvalues** of A. We will have much more to say about that later.

Example 8.5.5. Suppose we are given the LU-factorization of a particular matrix (borrowed from the text by Spence, Insel and Friedberg see Example 2 on pg. 154-155.)

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & -1 & 7 \\ 2 & -4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} = LU$$

The LU-factorization is pretty easy to find, we do not study it directly in these notes⁵. It is an important topic if you delve into serious numerical work where you need to write your own code and so forth. Note that L, U are triangular so we can calculate the determinant with ease:

$$det(A) = det(L)det(U) = 1 \cdot 1 \cdot 1 \cdot 1 \cdot 2 \cdot 2 = 4.$$

⁵there are many additional techniques of matrix theory concerning various special ways to factor a matrix. I can recommend some reading past this course if you are interested.

From a numerical perspective, the LU-factorization is a superior method for calculating det(A) as compared to the co-factor expansion. It has much better "convergence" properties. Incidentally, you might read Insel Spence and Friedberg's Elementary Linear Algebra for more discussion of algorithmics.

Example 8.5.6. Recall that the columns in A are linearly independent iff Ax = 0 has only the x = 0 solution. We also found that the existence of A^{-1} was equivalent to that claim in the case A was square since Ax = 0 implies $A^{-1}Ax = A^{-1}0 = 0$ hence x = 0. In Proposition 8.4.5 we proved $det(A) \neq 0$ iff A^{-1} exists. Thus the following check for $A \in \mathbb{R}^{n \times n}$ is nice to know:

columns of A are linearly independent \Leftrightarrow $det(A) \neq 0$.

Observe that this criteria is only useful if we wish to examine the linear independence of preciely *n*-vectors in \mathbb{R}^n . For example, $(1,1,1), (1,0,1), (2,1,2) \in \mathbb{R}^3$ have

$$det \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} = 0.$$

Therefore, $\{(1,1,1), (1,0,1), (2,1,2)\}$ form a linearly dependent set of vectors.

A natural curiousity, what about less than *n*-vectors? Is there some formula for that? Is there some formula we can plug say k-vectors into to ascertain the LI of those k-vectors? The answer is given by the **wedge product**. In short, if $v_1 \wedge v_2 \wedge \cdots \wedge v_k \neq 0$ then $\{v_1, v_2, \ldots, v_k\}$ is LI. This ties in with determinants at order k = n by the beautiful formula: for *n*-vectors in \mathbb{R}^n ,

$$v_1 \wedge v_2 \wedge \cdots \wedge v_n = \det[v_1 | v_2 | \cdots | v_n] e_1 \wedge e_2 \wedge \cdots \wedge e_n.$$

The wedge product is an algebraic structure which can be built over any finite dimensional vector space. The external direct sum of all possible wedge products of vectors in V gives $\Omega(V)$ the $2^{\dim(V)}$ -dimensional **exterior algebra** of V. For example, $V = \mathbb{R}^2$ has $\Omega(V) = \text{span}\{1, e_1, e_2, e_1 \land e_2\}$. If you'd like to know more about this algebra and how it extends and clarifies calculus III to calculus on *n*-dimensional space then you might read my advanced calculus Lecture notes. Another nice place to read more about these things from a purely linear-algebraic perspective is the text *Abstract Linear Algebra* by Morton L. Curtis.

8.6 Cramer's Rule

The numerical methods crowd seem to think this is a loathsome brute. It is an incredibly clumsy way to calculate the solution of a system of equations Ax = b. Moreover, Cramer's rule fails in the case det(A) = 0 so it's not nearly as general as our other methods. However, it does help calculate the variation of parameters formulas in differential equations so it is still of theoretical interest at a minimum. Students sometimes like it because it gives you a *formula* to find the solution. Students sometimes incorrectly jump to the conclusion that a formula is easier than say a *method*. It is certainly wrong here, the method of Gaussian elimination beats Cramer's rule by just about every objective criteria in so far as concrete numerical examples are concerned.

Proposition 8.6.1.

If Ax = b is a linear system of equations with $x = [x_1 \ x_2 \ \cdots \ x_n]^T$ and $A \in \mathbb{R}^{n \times n}$ such that $det(A) \neq 0$ then we find solutions

$$x_1 = \frac{det(A_1)}{det(A)}, \quad x_2 = \frac{det(A_2)}{det(A)}, \quad \dots, \quad x_n = \frac{det(A_n)}{det(A)}$$

where we define A_k to be the $n \times n$ matrix obtained by replacing the k-th column of A by the inhomogeneous term b.

Proof: Since $det(A) \neq 0$ we know that Ax = b has a unique solution. Suppose $x_j = \frac{det(A_j)}{det(A)}$ where $A_j = [col_1(A)| \cdots |col_{j-1}(A)| b| col_{j+1}(A)| \cdots |col_n(A)]$. We seek to show $x = [x_j]$ is a solution to Ax = b. Notice that the *n*-vector equations

$$Ae_1 = col_1(A), \dots, Ae_{j-1} = col_{j-1}(A), Ae_{j+1} = col_{j+1}(A), \dots, Ae_n = col_n(A), Ax = b$$

can be summarized as a single matrix equation:

$$A[e_1|\dots|e_{j-1}|x|e_{j+1}|\dots|e_n] = \underbrace{[col_1(A)|\dots|col_{j-1}(A)|b|col_{j+1}(A)|\dots|col_n(A)]}_{\text{this is precisely } A_i} = A_j$$

Notice that if we expand on the j-th column it's obvious that

$$det[e_1|\ldots|e_{j-1}|x|e_{j+1}|\cdots|e_n] = x_j$$

Returning to our matrix equation, take the determinant of both sides and use that the product of the determinants is the determinant of the product to obtain:

$$det(A)x_j = det(A_j)$$

Since $det(A) \neq 0$ it follows that $x_j = \frac{det(A_j)}{det(A)}$ for all j. \Box

This is the proof that is given in Lay's text. The construction of the matrix equation is not really an obvious step in my estimation. Whoever came up with this proof originally realized that he would need to use the determinant product identity to overcome the subtlety in the proof. Once you realize that then it's natural to look for that matrix equation. This is a clever proof⁶

Example 8.6.2. Solve Ax = b given that

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 8 \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

where $x = [x_1 \ x_2]^T$. Apply Cramer's rule, note det(A) = 2,

$$x_1 = \frac{1}{2}det \begin{bmatrix} 1 & 3\\ 5 & 8 \end{bmatrix} = \frac{1}{2}(8-15) = \frac{-7}{2}.$$

and,

$$x_2 = \frac{1}{2}det \begin{bmatrix} 1 & 1\\ 2 & 5 \end{bmatrix} = \frac{1}{2}(5-2) = \frac{3}{2}.$$

The original system of equations would be $x_1 + 3x_2 = 1$ and $2x_1 + 8x_2 = 5$. As a quick check we can substitute in our answers $x_1 = -7/2$ and $x_2 = 3/2$ and see if they work.

⁶as seen from my humble vantage point naturally

Please note: the following two examples are for breadth of exposition.

Example 8.6.3. An nonhomogeneous system of linear, constant coefficient ordinary differential equations can be written as a matrix differential equation:

$$\frac{dx}{dt} = Ax + f$$

It turns out we'll be able to solve the homogeneous system dx/dt = Ax via something called the matrix exponential. Long story short, we'll find n-solutions which we can concatenate into one big matrix solution X. To solve the given nonhomogeneous problem one makes the ansatz that x = Xv is a solution for some yet unknown vector of functions. Then calculus leads to the problem of solving

$$X\frac{dv}{dt} = f$$

where X is matrix of functions, dv/dt and f are vectors of functions. X is invertible so we expect to find a unique solution dv/dt. Cramer's rule says,

$$\left(\frac{dv}{dt}\right)_i = \frac{1}{det(X)}det[\vec{x}_1|\cdots|g|\cdots|\vec{x}_n] = \frac{W_i[f]}{det(X)} \quad defining \ W_i \ in \ the \ obvious \ way$$

For each i we integrate the equation above,

$$v_i(t) = \int \frac{W_i[f]dt}{det(X)}.$$

The general solution is thus,

$$x = Xv = X\left[\int \frac{W_i[f]dt}{det(X)}\right].$$

The first component of this formula justifies n-th order variation of parameters. For example in the n = 2 case you may have learned that $y_p = y_1v_1 + y_2v_2$ solves ay'' + by' + cy = g if

$$v_1 = \int \frac{-gy_2 dt}{a(y_1 y_2' - y_2 y_1')} \qquad v_2 = \int \frac{gy_1 dt}{a(y_1 y_2' - y_2 y_1')}$$

These come from the general result above. Notice that these formulas need $y_1y'_2 - y_2y'_1 \neq 0$. This is precisely the **Wronskian** $W[y_1, y_2] = y_1y'_2 - y_2y'_1$ of the fundamental solutions y_1, y_2 . It turns out that the Wronskian is nonzero for fundamental solutions thus the formulas above are entirely general.

The example that follows is borrowed from my 2013 Advanced Calculus notes. Here I used Cramer's Rule to solve for differentials of the dependent variables.

Example 8.6.4. Suppose x+y+z+w = 3 and $x^2-2xyz+w^3 = 5$. Calculate partial derivatives of z and w with respect to the independent variables x, y. Solution: we begin by calculation of the differentials of both equations:

$$dx + dy + dz + dw = 0$$

(2x - 2yz)dx - 2xzdy - 2xydz + 3w²dw = 0

We can solve for (dz, dw). In this calculation we can treat the differentials as formal variables.

$$dz + dw = -dx - dy$$

-2xydz + 3w²dw = -(2x - 2yz)dx + 2xzdy

8.6. CRAMER'S RULE

I find matrix notation is often helpful,

$$\begin{bmatrix} 1 & 1 \\ -2xy & 3w^2 \end{bmatrix} \begin{bmatrix} dz \\ dw \end{bmatrix} = \begin{bmatrix} -dx - dy \\ -(2x - 2yz)dx + 2xzdy \end{bmatrix}$$

Use Cramer's rule, multiplication by inverse, substitution, adding/subtracting equations etc... whatever technique of solving linear equations you prefer. Our goal is to solve for dz and dw in terms of dx and dy. I'll use Cramer's rule this time:

$$dz = \frac{\det \begin{bmatrix} -dx - dy & | & 1 \\ -(2x - 2yz)dx + 2xzdy & | & 3w^2 \end{bmatrix}}{\det \begin{bmatrix} 1 & 1 \\ -2xy & 3w^2 \end{bmatrix}} = \frac{3w^2(-dx - dy) + (2x - 2yz)dx - 2xzdy}{3w^2 + 2xy}$$

Collecting terms,

$$dz = \left(\frac{-3w^2 + 2x - 2yz}{3w^2 + 2xy}\right)dx + \left(\frac{-3w^2 - 2xz}{3w^2 + 2xy}\right)dy$$

From the expression above we can read various implicit derivatives,

$$\left(\frac{\partial z}{\partial x}\right)_y = \frac{-3w^2 + 2x - 2yz}{3w^2 + 2xy} \qquad \& \qquad \left(\frac{\partial z}{\partial y}\right)_x = \frac{-3w^2 - 2xz}{3w^2 + 2xy}$$

The notation above indicates that z is understood to be a function of independent variables x, y. $\left(\frac{\partial z}{\partial x}\right)_y$ means we take the derivative of z with respect to x while holding y fixed. The appearance of the dependent variable w can be removed by using the equations G(x, y, z, w) = (3, 5). Similar ambiguities exist for implicit differentiation in calculus I. Apply Cramer's rule once more to solve for dw:

$$dw = \frac{\det \begin{bmatrix} 1 & -dx - dy \\ -2xy & -(2x - 2yz)dx + 2xzdy \end{bmatrix}}{\det \begin{bmatrix} 1 & 1 \\ -2xy & 3w^2 \end{bmatrix}} = \frac{-(2x - 2yz)dx + 2xzdy - 2xy(dx + dy)}{3w^2 + 2xy}$$

Collecting terms,

$$dw = \left(\frac{-2x + 2yz - 2xy}{3w^2 + 2xy}\right)dx + \left(\frac{2xzdy - 2xydy}{3w^2 + 2xy}\right)dy$$

We can read the following from the differential above:

$$\left(\frac{\partial w}{\partial x}\right)_y = \frac{-2x + 2yz - 2xy}{3w^2 + 2xy} \qquad \& \qquad \left(\frac{\partial w}{\partial y}\right)_x = \frac{2xzdy - 2xydy}{3w^2 + 2xy}.$$

8.7 adjoint matrix

In this section we derive a general formula for the inverse of an $n \times n$ matrix. We already saw this formula in the 2×2 case and I work it out for the 3×3 case later in this section. As with Cramer's Rule, the results of this section are not to replace our earlier row-reduction based algoriths. Instead, these simply give us another tool, another view to answer questions concerning inverses.

Definition 8.7.1.

Let $A \in \mathbb{R}^{n \times n}$ the the matrix of cofactors is called the **adjoint** of A. It is denoted adj(A) and is defined by and $adj(A)_{ij} = C_{ij}$ where C_{ij} is the (i, j)-th cofactor.

I'll keep it simple here, lets look at the 2×2 case:

$$A = \left[\begin{array}{cc} a & b \\ c & d \end{array} \right]$$

has cofactors $C_{11} = (-1)^{1+1} det(d) = d$, $C_{12} = (-1)^{1+2} det(c) = -c$, $C_{21} = (-1)^{2+1} det(b) = -b$ and $C_{22} = (-1)^{2+2} det(a) = a$. Collecting these results,

$$adj(A) = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$$

This is interesting. Recall we found a formula for the inverse of A (if it exists). The formula was

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Notice that det(A) = ad - bc thus in the 2×2 case the relation between the inverse and the adjoint is rather simple:

$$A^{-1} = \frac{1}{\det(A)} adj(A)^T$$

In fact, this is true for all n,

Proposition 8.7.2.

If $A \in \mathbb{R}^{n \times n}$ is invertible then $A^{-1} = \frac{1}{det(A)} a dj(A)^T$.

Proof I: Calculate the product of A and $adj(A)^T$,

$$Aadj(A)^{T} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

The (i, j)-th component of the product above is

$$(Aadj(A)^T)_{ij} = A_{i1}C_{j1} + A_{i2}C_{j2} + \dots + A_{in}C_{jn}.$$

Suppose that i = j then the sum above is precisely the *i*-th row co-factor expansion for det(A):

$$(Aadj(A)^T)_{ij} = A_{i1}C_{i1} + A_{i2}C_{i2} + \dots + A_{in}C_{in} = det(A)$$

8.7. ADJOINT MATRIX

If $i \neq j$ then the sum vanishes. I leave the details to the reader⁷

Proof II: To find the inverse of A we need only apply Cramer's rule to solve the equations implicit within $AA^{-1} = I$. Let $A^{-1} = [v_1|v_2|\cdots|v_n]$ we need to solve

$$Av_1 = e_1, \qquad Av_2 = e_2, \qquad \dots \qquad Av_n = e_n$$

Cramer's rule gives us $(v_1)_j = \frac{C_{1j}}{det(A)}$ where $C_{1j} = (-1)^{1+j} M_{ij}$ is the cofactor formed from deleting the first row and *j*-th column. Apply Cramer's rule to deduce the *j*-component of the *i*-th column in the inverse $(v_i)_j = \frac{C_{ij}}{det(A)}$. Therefore, $col_i(A^{-1})_j = (A^{-1})_{ji} = \frac{C_{ij}}{det(A)}$. By definition $adj(A) = [C_{ij}]$ hence $adj(A)_{ij}^T = C_{ji}$ and it follows that $A^{-1} = \frac{1}{det(A)}adj(A)^T$. \Box

Example 8.7.3. Let's calculate the general formula for the inverse of a 3×3 matrix. Assume it exists for the time being. (the criteria for the inverse existing is staring us in the face everywhere here). Let

$$A = \left[\begin{array}{rrr} a & b & c \\ d & e & f \\ g & h & i \end{array} \right]$$

Calculate the cofactors,

$$C_{11} = det \begin{bmatrix} e & f \\ h & i \end{bmatrix} = ei - fh,$$

$$C_{12} = -det \begin{bmatrix} d & f \\ g & i \end{bmatrix} = fg - di,$$

$$C_{13} = det \begin{bmatrix} d & e \\ g & h \end{bmatrix} = dh - eg,$$

$$C_{21} = -det \begin{bmatrix} b & c \\ h & i \end{bmatrix} = ch - bi,$$

$$C_{22} = det \begin{bmatrix} a & c \\ g & h \end{bmatrix} = ai - cg,$$

$$C_{23} = -det \begin{bmatrix} a & b \\ g & h \end{bmatrix} = bg - ah,$$

$$C_{31} = det \begin{bmatrix} b & c \\ e & f \end{bmatrix} = bf - ce,$$

$$C_{32} = -det \begin{bmatrix} a & c \\ d & f \end{bmatrix} = cd - af,$$

$$C_{33} = det \begin{bmatrix} a & b \\ d & e \end{bmatrix} = ae - bd.$$

Hence the transpose of the adjoint is

$$adj(A)^{T} = \begin{bmatrix} ei - fh & fg - di & dh - eg\\ \hline ch - bi & ai - cg & bg - ah\\ \hline bf - ce & cd - af & ae - bd \end{bmatrix}$$

⁷I don't have an easy proof that these terms cancel for $i \neq j$. It's simply to verify for the n = 2 or n = 3 cases but the reason appears to be a combinatorial cancellation. If you can provide a concrete and readable proof for the general case it would definitely earn you some points.

Thus, using the $A^{-1} = det(A)adj(A)^T$

$\begin{bmatrix} a \end{bmatrix}$	b	<i>c</i> -	$ ^{-1}$	1	$\int ei - fh$	ch-bi	bf - ce
d	e	f	=	$= \frac{1}{a_{0i} + b_{1}f_{2} + a_{1}db}$	fg - di	ai-cg	cd-af
$\lfloor g$	h	<i>i</i> _		aei + bfg + can - gec - hfa - iac	dh - eg	bg - ah	ae-bd

You should notice that are previous method for finding A^{-1} is far superior to this method. It required much less calculation. Let's check my formula in the case A = 3I, this means a = e = i = 3 and the others are zero.

$$I^{-1} = \frac{1}{27} \begin{bmatrix} 9 & 0 & 0 \\ \hline 0 & 9 & 0 \\ \hline 0 & 0 & 9 \end{bmatrix} = \frac{1}{3}I$$

This checks, $(3I)(\frac{1}{3}I) = \frac{3}{3}II = I$. I do not recommend that you memorize this formula to calculate inverses for 3×3 matrices.

8.8 applications

The determinant is a convenient mnemonic to create expressions which are antisymmetric. The key property is that if we switch a row or column it creates a minus sign. This means that if any two rows are repeated then the determinant is zero. Notice this is why the cross product of two vectors is naturally phrased in terms of a determinant. The antisymmetry of the determinant insures the formula for the cross-product will have the desired antisymmetry. In this section we examine a few more applications for the determinant.

Example 8.8.1. The Pauli's exclusion principle in quantum mechanics states that the wave function of a system of fermions is antisymmetric. Given N-electron wavefunctions $\chi_1, \chi_2, \ldots, \chi_N$ the following is known as the **Slater Determinant**

$$\Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = det \begin{bmatrix} \chi_1(\vec{r}_1) & \chi_2(\vec{r}_1) & \cdots & \chi_N(\vec{r}_1) \\ \chi_1(\vec{r}_2) & \chi_2(\vec{r}_2) & \cdots & \chi_N(\vec{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\vec{r}_N) & \chi_2(\vec{r}_N) & \cdots & \chi_N(\vec{r}_N) \end{bmatrix}$$

Notice that $\Psi(\vec{r}_1, \vec{r}_1, \ldots, \vec{r}_N) = 0$ and generally if any two of the position vectors $\vec{r}_i = \vec{r}_j$ then the total wavefunction $\Psi = 0$. In quantum mechanics the wavefunction's modulus squared gives the probability density of finding the system in a particular circumstance. In this example, the fact that any repeated entry gives zero means that no two electrons can share the same position. This is characteristic of particles with half-integer spin, such particles are called fermions. In contrast, bosons are particles with integer spin and they can occupy the same space. For example, light is made of photons which have spin 1 and in a laser one finds many waves of light traveling in the same space.
Example 8.8.2. This is an example of a Vandermonde determinant. Note the following curious formula:

$$det \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x & y \end{bmatrix} = 0$$

Let's reduce this by row-operations⁸

$$\begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x & y \end{bmatrix} \xrightarrow[r_3 - r_1 \to r_3]{r_3 - r_1 \to r_3} \begin{bmatrix} 1 & x_1 & y_1 \\ 0 & x_2 - x_1 & y_2 - y_1 \\ 0 & x - x_1 & y - y_1 \end{bmatrix}$$

Notice that the row operations above could be implemented by multiply on the left by $E_{r_2-r_1 \to r_2}$ and $E_{r_3-r_1 \to r_3}$. These are invertible matrices and thus $det(E_{r_2-r_1 \to r_2}) = k_1$ and $det(E_{r_3-r_1 \to r_3}) = k_2$ for some pair of nonzero constants k_1, k_2 . If X is the given matrix and Y is the reduced matrix above then $Y = E_{r_3-r_1 \to r_3}E_{r_2-r_1 \to r_2}X$ thus,

$$0 = det \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x & y \end{bmatrix} = k_1 k_2 det \begin{bmatrix} 1 & x_1 & y_1 \\ 0 & x_2 - x_1 & y_2 - y_1 \\ 0 & x - x_1 & y - y_1 \end{bmatrix}$$
$$= k_1 k_2 [(x_2 - x_1)(y - y_1) - (y_2 - y_1)(x - x_1)]$$

Divide by k_1k_2 and rearrange to find:

$$(x_2 - x_1)(y - y_1) = (y_2 - y_1)(x - x_1) \qquad \Rightarrow \qquad y = y_1 + \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x - x_1)$$

The boxed equation is the famous two-point formula for a line.

Example 8.8.3. There are many twists on the previous example. Here's one to differential equations. Suppose you want a second order linear ODE L[y] = for which a given pair of functions y_1, y_2 are solutions. A simple way to express the desired equation is L[y] = 0 where

$$L[y] = det \begin{bmatrix} y & y' & y'' \\ y_1 & y'_1 & y''_1 \\ y_2 & y'_2 & y''_2 \end{bmatrix}$$

Observe $L[y_1] = 0$ and $L[y_2] = 0$ are immediately clear as setting $y = y_1$ or $y = y_2$ gives a repeated row.

Example 8.8.4. Let us consider a linear transformation $T([x, y]^T) = [2x, x + y]^T$. Furthermore, let's see how a rectangle R with corners (0,0), (3,0), (3,1), (0,1). Since this linear transformation is invertible (I invite you to prove that) it follows that the image of a line is again a line. Therefore, if we find the image of the corners under the mapping T then we can just connect the dots in the image to see what T(R) resembles. Our goal here is to see what a linear transformation does to a rectangle.

$$T([0,0]^{T}) = [0,0]^{T}$$
$$T([3,0]^{T}) = [6,3]^{T}$$
$$T([3,1]^{T}) = [6,4]^{T}$$
$$T([0,1]^{T}) = [0,1]^{T}$$

⁸of course we could calculate it straight from the co-factor expansion, I merely wish to illustrate how we can use row operations to simplify a determinant



As you can see from the picture we have a paralellogram with base 6 and height 1 thus Area(T(R)) = 6. In constrast, Area(R) = 3. You can calculate that det(T) = 2. Curious, Area(T(R)) = det(T)Area(R). This can be derived in general, it's not too hard given our definition of n-volume and the wonderful identities we've learned for matrix multiplication and determinants.

The examples that follow illustrate how determinants arise in the study of infinitesimal areas and volumes in multivariate calculus.

Example 8.8.5. The infinitesimal area element for polar coordinate is calculated from the Jacobian:

$$dS = det \begin{bmatrix} r\sin(\theta) & -r\cos(\theta) \\ \cos(\theta) & \sin(\theta) \end{bmatrix} drd\theta = (r\sin^2(\theta) + r\cos^2(\theta))drd\theta = rdrd\theta$$

Example 8.8.6. The infinitesimal volume element for cylindrical coordinate is calculated from the Jacobian:

$$dV = det \begin{bmatrix} r\sin(\theta) & -r\cos(\theta) & 0\\ \cos(\theta) & \sin(\theta) & 0\\ 0 & 0 & 1 \end{bmatrix} dr d\theta dz = (r\sin^2(\theta) + r\cos^2(\theta)) dr d\theta dz = r dr d\theta dz$$

Jacobians are needed to change variables in multiple integrals. The Jacobian⁹ is a determinant which measures how a tiny volume is rescaled under a change of coordinates. Each row in the matrix making up the Jacobian is a tangent vector which points along the direction in which a coordinate increases when the other two coordinates are fixed.

⁹see pages 206-208 of Spence Insel and Friedberg or perhaps my advanced calculus notes where I develop differentiation from a linear algebraic viewpoint.

8.9 similarity and determinants for linear transformations

Thus far this chapter has been mainly matrix theoretic. However, the determinant is also defined and of interest for abstract linear transformations. Suppose V is an n-dimensional vector space over \mathbb{R} and consider $T: V \to V$ a linear transformation. If β, γ are finite bases for V then we can calculate $[T]_{\beta,\beta}$ and $[T]_{\gamma,\gamma}$. Note these are both $n \times n$ matrices as the domain and codomain are both n-dimensional. Furthermore, applying Proposition 7.4.7 we have:

$$[T]_{\gamma,\gamma} = [\Phi_{\gamma} \circ \Phi_{\beta}^{-1}][T]_{\beta,\beta} [\Phi_{\beta} \circ \Phi_{\gamma}^{-1}]$$

If we set $P = [\Phi_{\beta} \circ \Phi_{\gamma}^{-1}]$ then the equation above simply reduces to:

$$[T]_{\gamma,\gamma} = P^{-1}[T]_{\beta,\beta}P.$$

I've mentioned this concept in passing before, but for future reference we should give a precise definition:

Definition 8.9.1.

Let $A, B \in \mathbb{R}^{n \times n}$ then we say A and B are similar matrices if there exists $P \in \mathbb{R}^{n \times n}$ such that $B = P^{-1}AP$.

In invite the reader to verify that matrix similarity is an equivalence relation. Furthermore, you might contrast this idea of sameness with that of **matrix congruence**. To say A, B are matrix congruent it sufficed to find P, Q such that $B = P^{-1}AQ$. Here $P \in \mathbb{R}^{m \times m}$ and $Q \in \mathbb{R}^{n \times n}$ and we needed only that $A, B \in \mathbb{R}^{m \times n}$. Matrix congruence was defined for rectangular matrices whereas similarity is only for square matrices. The idea is this, two congruent matrices represent the same linear transformation $T: V \to W$. There is some choice of bases for V and W which change the formula of T from A to B or vice-versa. Moreover, Theorem 7.5.2 revealed the canonical form relative to matrix congruence classes was simply an identity matrix as big as the rank of the transformation padded with zeros. To understand the difference between congruence and similarity it is important to notice that congruence is based on adjusting both the basis in the domain and separately the basis in the codomain. In contrast, similarity is is related to changing the basis in the domain and codomain in the same exact fashion. This means it is a stronger condition for two matrices to be similar. The analog for Theorem 7.5.2 is what is known as the real Jordan form and it provides the concluding thought of this course. The criteria which will guide us to find the Jordan form is simply this: any two similar matrices should have the exact same Jordan form. With a few conventional choices made, this gives us a cannonical representative of each equivalence class of similar matrices. It is worthwhile to note the following:

Proposition 8.9.2.

- Let $A, B, C \in \mathbb{R}^{n \times n}$.
 - 1. A is similar to A.
 - 2. If A is similar to B then B is similar to A.
 - 3. If A is similar to B and B is similar to C then A similar to C.
 - 4. If A and B are similar then det(A) = det(B)
 - 5. If A and B are similar then tr(A) = tr(B)
 - 6. If A and B are similar then rank(A) = rank(B) and nullity(A) = nullity(B)

Given the proposition above we can make the following definitions without ambiguity.

Definition 8.9.3.

Let $T: V \to V$ be a linear transformation on a finite-dimensional vector space V and let β be any basis of V,

1. $det(T) = det([T]_{\beta,\beta}).$ 2. $tr(T) = tr([T]_{\beta,\beta})$ 3. $rank(T) = rank([T]_{\beta,\beta}).$

Example 8.9.4. Consider $D: P_2 \times P_2$ defined by D[f(x)] = df/dx note that $D[ax^2+bx+c] = 2ax+b$ implies that in the $\beta = \{x^2, x, 1\}$ coordinates we find:

$$[D]_{\beta,\beta} = \left[\begin{array}{rrr} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right]$$

Thus det(D) = 0.

Example 8.9.5. Consider $L : \mathbb{R}^{2 \times 2} \to \mathbb{R}^{2 \times 2}$ defined by $L(A) = A^T$. Observe:

$$L(E_{11}) = E_{11}, \ L(E_{12}) = E_{21}, \ L(E_{21}) = E_{12}, \ L(E_{122}) = E_{22}$$

Therefore, if $\beta = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ then

$$[L]_{\beta,\beta} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Swapping columns 2 and 3 brings $[L]_{\beta,\beta}$ to the identity matrix. Hence, det(L) = -1. Example 8.9.6. Consider $L : \mathbb{R}^{3\times 3} \to \mathbb{R}^{3\times 3}$ defined by $L(A) = A^T$

 $L(E_{11}) = E_{11}, \ L(E_{22}) = E_{22}, \ L(E_{33}) = E_{33}$

these explain the first three columns in $[L]_{\beta,\beta}$. Next,

$$L(E_{12}) = E_{21}, \ L(E_{13}) = E_{31}, \ L(E_{21}) = E_{12}, \ L(E_{23}) = E_{32}, \ L(E_{31}) = E_{13}, \ L(E_{32}) = E_{23}.$$

Let us order β so the diagonals come first: $\beta = \{E_{11}, E_{22}, E_{33}, E_{12}, E_{21}, E_{23}, E_{32}, E_{13}, E_{31}\}$. Thus,

Three column swaps modify the above to the identity. Thus, det(L) = -1.

8.10 conclusions

The theorem which follows collects ideas somewhat comprehensively for our course thus far.

Theorem 8.10.1.

Let A be a real $n \times n$ matrix then the following are equivalent: (a.) A is invertible, (b.) rref[A|0] = [I|0] where $0 \in \mathbb{R}^n$, (c.) Ax = 0 iff x = 0, (d.) A is the product of elementary matrices, (e.) there exists $B \in \mathbb{R}^{n \times n}$ such that AB = I, (f.) there exists $B \in \mathbb{R}^{n \times n}$ such that BA = I, (g.) rref[A] = I, (h.) rref[A|b] = [I|x] for an $x \in \mathbb{R}^n$, (i.) Ax = b is consistent for every $b \in \mathbb{R}^n$, (j.) Ax = b has exactly one solution for every $b \in \mathbb{R}^n$, (k.) A^T is invertible, (1.) $det(A) \neq 0$, (m.) Cramer's rule yields solution of Ax = b for every $b \in \mathbb{R}^n$. (n.) $Col(A) = \mathbb{R}^{n \times 1}$, (0.) $Row(A) = \mathbb{R}^{1 \times n}$, (p.) rank(A) = n, (q.) $Null(A) = \{0\},\$ (r.) $\nu = 0$ for A where $\nu = dim(Null(A))$, (s.) the columns of A are linearly independent, (t.) the rows of A are linearly independent,

This list is continued on the next page.

Let A be a real $n \times n$ matrix then the following are equivalent:

- (u.) the induced linear operator L_A is onto; $L_A(\mathbb{R}^n) = \mathbb{R}^n$.
- (v.) the induced linear operator L_A is 1-1
- (w.) the induced linear operator L_A is an isomorphism.
- (x.) the kernel of the induced linear operator is trivial; $ker(L_A) = \{0\}$.

We should pay special attention to the fact that the above comments hold only for a square matrix. If we consider a rectangular matrix then the connection between the concepts in the theorem are governed by the dimension formulas we discovered in Part II.

Next, the list of equivalent statements for a singular $n \times n$ matrix:

Theorem 8.10.2.

Let A be a real $n \times n$ matrix then the following are equivalent: (a.) A is not invertible, (b.) Ax = 0 has at least one nontrivial solution., (c.) there exists $b \in \mathbb{R}^n$ such that Ax = b is inconsistent, (d.) det(A) = 0, (e.) $Null(A) \neq \{0\}$, (f.) there are $1 \leq \nu = dim(Null(A))$ parameters in the general solution to Ax = 0, (g.) the induced linear operator L_A is **not** onto; $L_A(\mathbb{R}^n) \neq \mathbb{R}^n$. (h.) the induced linear operator L_A is **not** 1-1 (i.) the induced linear operator L_A is **not** an isomorphism.

(j.) the kernel of the induced linear operator is **non**trivial; $ker(L_A) \neq \{0\}$.

It turns out this theorem is also useful. We shall see it is fundamental to the theory of eigenvectors.

Chapter 9

euclidean geometry

The concept of a geometry is very old. Philosophers in the nineteenth century failed miserably in their analysis of geometry and the physical world. They became mired in the popular misconception that mathematics must be physical. They argued that because 3 dimensional Eulcidean geometry was the only geometry familar to everyday experience it must surely follow that a geometry which differs from Euclidean geometry must be nonsensical. However, why should physical intuition factor into the argument? We understand now that geometry is a mathematical construct, not a physical one. There are many possible geometries. On the other hand, it would seem the geometry of space and time probably takes just one form. We are tempted by this misconception every time we ask "but what is this math really". That question is usually wrong-headed. A better question is "is this math logically consistent" and if so what physical systems is it known to model.

The modern view of geometry is stated in the langauge of manifolds, fiber bundles, algebraic geometry and perhaps even more fantastic structures. There is currently great debate as to how we should model the true intrinsic geometry of the universe. Branes, strings, quivers, noncommutative geometry, twistors, ... this list is endless. However, at the base of all these things we must begin by understanding what the geometry of a flat space entails.

Vector spaces are flat manifolds. They possess a global coordinate system once a basis is chosen. Up to this point we have only cared about algebraic conditions of linear independence and spanning. There is more structure we can assume. We can ask what is the length of a vector? Or, given two vectors we might want to know what is the angle bewtween those vectors? Or when are two vectors orthogonal?

If we desire we can also insist that the basis consist of vectors which are *orthogonal* which means "perpendicular" in a generalized sense. A geometry is a vector space plus an idea of orthogonality and length. The concepts of orthogonality and length are encoded by an inner-product. Inner-products are symmetric, positive definite, bilinear forms, they're like a dot-product. Once we have a particular geometry in mind then we often restrict the choice of bases to only those bases which preserve the length of vectors.

The mathematics of orthogonality is exhibited by the dot-products and vectors in calculus III. However, it turns out the concept of an *inner-product* allows us to extend the idea or perpendicular to abstract vectors such as functions. This means we can even ask interesting questions such as "how close is one function to another" or "what is the closest function to a set of functions". Least-squares curve fitting is based on this geometry.

This chapter begins by defining dot-products and the norm (a.k.a. length) of a vector in \mathbb{R}^n . Then we discuss orthogonality, the Gram Schmidt algorithm, orthogonal complements and finally the application to the problem of least square analysis. The chapter concludes with a consideration of the similar, but abstract, concept of an inner product space. We look at how least squares generalizes to that context and we see how Fourier analysis naturally flows from our finite dimensional discussions of orthogonality. ¹

Let me digress from linear algebra for a little while. In physics it is customary to only allow coordinates which fit the physics. In classical mechanics one often works with intertial frames which are related by a rigid motion. Certain quantities are the same in all intertial frames, notably force. This means Newtons laws have the same form in all intertial frames. The geometry of special relativity is 4 dimensional. In special relativity, one considers coordinates which preserve Einstein's three axioms. Allowed coordinates are related to other coordinates by Lorentz transformations. These Lorentz transformations include rotations and velocity boosts. These transformations are designed to make the speed of a light ray invariant in all frames. For a linear algebraist the vector space is the starting point and then coordinates are something we add on later. Physics, in contrast, tends to start with coordinates and if the author is kind he might warn you which transformations are allowed.

What coordinate transformations are allowed actually tells you what kind of physics you are dealing with. This is an interesting and nearly universal feature of modern physics. The allowed transformations form what is known to physicsists as a "group" (however, strictly speaking these groups do not always have the strict structure that mathematicians insist upon for a group). In special relativity the group of interest is the Poincaire group. In quantum mechanics you use unitary groups because unitary transformations preserve probabilities. In supersymmetric physics you use the super Poincaire group because it is the group of transformations on superspace which preserves supersymmetry. In general relativity you allow general coordinate transformations which are locally lorentzian because all coordinate systems are physical provided they respect special relativity in a certain approximation. In solid state physics there is something called the renormilzation group which plays a central role in physical predictions of field-theoretic models. My point? Transformations of coordinates are important if you care about physics. We study the basic case of vector spaces in this course. If you are interested in the more sophisticated topics just ask, I can show you where to start reading.

We begin by developing all the important properties of norms and dot-products in the standard euclidean geometry of \mathbb{R}^n . Then we discuss the theory of orthogonal projections. This brings us a calculational method to find a complementary subspace for $W \leq \mathbb{R}^n$. In particular, W^{\perp} complements W meaning that $W \oplus W^{\perp} = \mathbb{R}^n$. This geometry yields a perhaps surprising result on finding approximate solutions to inconsistent systems. We devote several sections to explaining the calculational scheme of least squares data fitting. Then we abstract to inner product spaces. The definition of length and norm is modified or invented, yet the techniques we developed for \mathbb{R}^n still apply. An initiation to Fourier analysis is given. Finally, we conclude with a technical section on the so-called QR-factorization of an orthogonal matrix.

 $^{^{1}}$ we ignore analytical issues of convergence since we have only in mind a Fourier approximation, not the infinite series

9.1 Euclidean geometry of \mathbb{R}^n

The dot-product is a mapping from $\mathbb{R}^n \times \mathbb{R}^n$ to \mathbb{R} . We take in a pair of vectors and output a real number. We have used it throughout the course for the inner workings of matrix-multiplication. Now we study the geometry which the dot-product naturally induces for \mathbb{R}^n . We attempt a complete discussion here so the generalization to inner products later in this chapter is simple to envision.

Definition 9.1.1.

Let $x, y \in \mathbb{R}^n$ we define $x \cdot y \in \mathbb{R}$ by $x \cdot y = x^T y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$

Example 9.1.2. Let v = (1, 2, 3, 4, 5) and w = (6, 7, 8, 9, 10)

$$v \bullet w = 6 + 14 + 24 + 36 + 50 = 130$$

The dot-product can be used to define the length of a vector and the angle between two vectors.

Definition 9.1.3.

The **length** or **norm** of $x \in \mathbb{R}^n$ is a real number which is defined by $||x|| = \sqrt{x \cdot x}$. Furthermore, let x, y be nonzero vectors in \mathbb{R}^n we define the **angle** θ between x and y by $\cos^{-1}\left[\frac{x \cdot y}{||x|| ||y||}\right]$. \mathbb{R} together with these definitions of length and angle forms a **Euclidean Geometry**.

The picture below helps us understand why the definition above is a natural formula for vector length.Notice the Pythagorean theorem in two dimensions yields the same theorem in three dimensions provided our coordinate axes are set at right-angles to one another.



Technically, before we make this definition we should make sure that the formulas given above even make sense. I have not shown that $x \cdot x$ is nonnegative and how do we know that the inverse cosine is well-defined? The first proposition below shows the norm of x is well-defined and establishes several foundational properties of the dot-product.

Proposition 9.1.4.

Suppose $x, y, z \in \mathbb{R}^n$ and $c \in \mathbb{R}$ then 1. $x \cdot y = y \cdot x$ 2. $x \cdot (y + z) = x \cdot y + x \cdot z$ 3. $c(x \cdot y) = (cx) \cdot y = x \cdot (cy)$ 4. $x \cdot x \ge 0$ and $x \cdot x = 0$ iff x = 0

Proof: the proof of (1.) is easy, $x \cdot y = \sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} y_i x_i = y \cdot x$. Likewise,

$$x \cdot (y+z) = \sum_{i=1}^{n} x_i (y+z)_i = \sum_{i=1}^{n} (x_i y_i + x_i z_i) = \sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} x_i z_i = x \cdot y + x \cdot z_i$$

proves (2.) and since

$$c\sum_{i=1}^{n} x_{i}y_{i} = \sum_{i=1}^{n} cx_{i}y_{i} = \sum_{i=1}^{n} (cx)_{i}y_{i} = \sum_{i=1}^{n} x_{i}(cy)_{i}$$

we find $c(x \cdot y) = (cx) \cdot y = x \cdot (cy)$. Continuting to (4.) notice that $x \cdot x = x_1^2 + x_2^2 + \dots + x_n^2$ thus $x \cdot x$ is the sum of squares and it must be nonnegative. Suppose x = 0 then $x \cdot x = x^T x = 0^T 0 = 0$. Conversely, suppose $x \cdot x = 0$. Suppose $x \neq 0$ then we find a contradiction since it would have a nonzero component which implies $x_1^2 + x_2^2 + \dots + x_n^2 \neq 0$. This completes the proof of (4.). \Box

The formula $\cos^{-1}\left[\frac{x \cdot y}{||x|| ||y||}\right]$ is harder to justify. The inequality that we need for it to be reasonable is $\left|\frac{x \cdot y}{||x|| ||y||}\right| \leq 1$, otherwise we would not have a number in the $dom(\cos^{-1}) = range(\cos) = [-1, 1]$. An equivalent inequality is $|x \cdot y| \leq ||x|| ||y||$ which is known as the **Cauchy-Schwarz** inequality.

Proposition 9.1.5.

If $x, y \in \mathbb{R}^n$ then $|x \cdot y| \le ||x|| ||y||$

Proof: I've looked in a few linear algebra texts and I must say the proof given in Spence, Insel and Friedberg is probably the most efficient and clear. Other texts typically run up against a quadratic inequality in some part of their proof (for example the linear algebra texts by Apostle, Larson& Edwards, Anton & Rorres to name a few). That is somehow hidden in the proof that follows: let $x, y \in \mathbb{R}^n$. If either x = 0 or y = 0 then the inequality is clearly true. Suppose then that both x and y are nonzero vectors. It follows that $||x||, ||y|| \neq 0$ and we can define vectors of unit-length; $\hat{x} = \frac{x}{||x||}$ and $\hat{y} = \frac{y}{||y||}$. Notice that $\hat{x} \cdot \hat{x} = \frac{x}{||x||} \cdot \frac{x}{||x||} = \frac{1}{||x||^2} \hat{x} \cdot x = \frac{x \cdot x}{x \cdot x} = 1$ and likewise $\hat{y} \cdot \hat{y} = 1$. Consider,

$$0 \le ||\hat{x} \pm \hat{y}||^2 = (\hat{x} \pm \hat{y}) \cdot (\hat{x} \pm \hat{y})$$

$$= \hat{x} \cdot \hat{x} \pm 2(\hat{x} \cdot \hat{y}) + \hat{y} \cdot \hat{y}$$

$$= 2 \pm 2(\hat{x} \cdot \hat{y})$$

$$\Rightarrow -2 \le \pm 2(\hat{x} \cdot \hat{y})$$

$$\Rightarrow \pm \hat{x} \cdot \hat{y} \le 1$$

$$\Rightarrow |\hat{x} \cdot \hat{y}| \le 1$$

Therefore, noting that $x = ||x||\hat{x}$ and $y = ||y||\hat{y}$,

$$|x \bullet y| = ||x||\hat{x} \bullet ||y||\hat{y}| = ||x|| ||y|| |\hat{x} \bullet \hat{y}| \le ||x|| ||y||.$$

The use of unit vectors is what distinguishes this proof from the others I've found. \Box

Remark 9.1.6.

The dot-product is but one of many geometries for \mathbb{R}^n . We will explore generalizations of the dot-product in a later section. However, in this section we will work exclusively with the standard dot-product on \mathbb{R}^n . Generally, unless explicitly indicated otherwise, we assume Euclidean geometry for \mathbb{R}^n .

Example 9.1.7. Let v = (1, 2, 3, 4, 5) and w = (6, 7, 8, 9, 10) find the angle between these vectors and calculate the unit vectors in the same directions as v and w. Recall that, $v \cdot w = 6 + 14 + 24 + 36 + 50 = 130$. Furthermore,

$$||v|| = \sqrt{1^2 + 2^2 + 3^2 + 4^2 + 5^2} = \sqrt{1 + 4 + 9 + 16 + 25} = \sqrt{55}$$
$$|w|| = \sqrt{6^2 + 7^2 + 8^2 + 9^2 + 10^2} = \sqrt{36 + 49 + 64 + 81 + 100} = \sqrt{330}$$

We find unit vectors via the standard trick, you just take the given vector and multiply it by the reciprocal of its length. This is called **normalizing** the vector,

$$\hat{v} = \frac{1}{\sqrt{55}}(1, 2, 3, 4, 5)$$
 $\hat{w} = \frac{1}{\sqrt{330}}(6, 7, 8, 9, 10)$

The angle is calculated from the definition of angle,

$$\theta = \cos^{-1} \left(\frac{130}{\sqrt{55}\sqrt{330}} \right) = 15.21^{\circ}$$

It's good we have this definition, 5-dimensional protractors are very expensive.

Proposition 9.1.8.

Let $x, y \in \mathbb{R}^n$ and suppose $c \in \mathbb{R}$ then 1. ||cx|| = |c| ||x||2. $||x + y|| \le ||x|| + ||y||$

Proof: let $x \in \mathbb{R}^n$ and $c \in \mathbb{R}$ then calculate,

$$||cx||^2 = (cx) \bullet (cx) = c^2 x \bullet x = c^2 ||x||^2$$

Since $||cx|| \ge 0$ the squareroot yields $||cx|| = \sqrt{c^2} ||x||$ and $\sqrt{c^2} = |c|$ thus ||cx|| = |c|||x||. Item (2.) is called the **triangle inequality** for reasons that will be clear when we later discuss the distance function. Let $x, y \in \mathbb{R}^n$,

defn. of norm	$ x+y ^2 = (x+y) \bullet (x+y) $
prop. of dot-product	$= x \bullet (x+y) + y \bullet (x+y) $
prop. of dot-product	$= x \bullet x + x \bullet y + y \bullet x + y \bullet y $
prop. of dot-product	$= x ^2 + 2x \cdot y + y ^2 $
triangle in eq. for $\mathbb R$	$\leq x ^2 + 2 x \bullet y + y ^2$
Cauchy-Schwarz ineq.	$\leq x ^2 + 2 x y + y ^2$
algebra	$\leq (x + y)^2$

Notice that both ||x + y|| and ||x|| + ||y|| are nonnegative by (4.) of Proposition 9.1.4 hence the inequality above yields $||x + y|| \le ||x|| + ||y||$. \Box

Definition 9.1.9.

The **distance** between $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^n$ is defined to be $d(a, b) \equiv ||b - a||$.

If we draw a picture this definition is very natural. Here we are thinking of the points a, b as vectors from the origin then b - a is the vector which points from a to b (this is algebraically clear since a + (b - a) = b). Then the distance between the points is the length of the vector that points from one point to the other. If you plug in two dimensional vectors you should recognize the distance formula from middle school math:

$$d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Set $a = (x_1, y_1)$ and $b = (x_2, y_2)$ to see how $d(a, b) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$.



Actually, to be honest, the picture above is not for just n = 2. It indicates the truth which is that the distance formula d(a, b) = ||b - a|| expresses the distance between points in *n*-dimensional space. Moreover, the *n*-dimensional distance function has nice properties:

Proposition 9.1.10.

Let
$$d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$
 be the distance function then
1. $d(x, y) = d(y, x)$
2. $d(x, y) \ge 0$
3. $d(x, x) = 0$
4. $d(x, y) + d(y, z) \ge d(x, z)$

Proof: I leave the proof of (1.), (2.) and (3.) to the reader. Item (4.) is also known as the **triangle inequality**. Think of the points x, y, z as being the vertices of a triangle, this inequality says the sum of the lengths of two sides cannot be smaller than the length of the remaining side. Let $x, y, z \in \mathbb{R}^n$ and note by the triangle inequality for $|| \cdot ||$,

$$d(x,z) = ||z - x|| = ||z - y + y - x|| \le ||z - y|| + ||y - x|| = d(y,z) + d(x,y). \quad \Box$$

We study the 2 and 3 dimensional case in some depth in calculus III. Differential calculus helps to unravel the geometry of graphs and level functions and surfaces. In constrast, our objects of interest are linear so calculus is not a necessary ingredient.

orthogonality in \mathbb{R}^n 9.2

Two vectors are orthogonal if the vectors point in mutually exclusive directions. We saw in calculus III the dot-product allowed us to pick apart vectors into pieces. The same is true in *n*-dimensions: we can take a vector an disassemble it into component vectors which are orthogonal.

Definition 9.2.1.

Let $v, w \in \mathbb{R}^n$ then we say v and w are **orthogonal** iff $v \cdot w = 0$.

Example 9.2.2. Let $v = [1,2,3]^T$ describe the set of all vectors which are orthogonal to v. Let $r = [x, y, z]^T$ be an arbitrary vector and consider the orthogonality condition:

$$0 = v \bullet r = [1, 2, 3][x, y, z]^T = x + 2y + 3z = 0.$$

If you've studied 3 dimensional Cartesian geometry you should recognize this as the equation of a plane through the origin with normal vector < 1, 2, 3 >.

.....

Proposition 9.2.3. Pythagorean Theorem in n-dimensions

If
$$x, y \in \mathbb{R}^n$$
 are orthogonal vectors then $||x||^2 + ||y||^2 = ||x+y||^2$.

Proof: Calculate $||x + y||^2$ from the dot-product,

 $||x+y||^{2} = (x+y) \cdot (x+y) = x \cdot x + x \cdot y + y \cdot x + y \cdot y = ||x||^{2} + ||y||^{2}.$

Proposition 9.2.4.

The zero vector is orthogonal to all other vectors in \mathbb{R}^n .

Proof: let $x \in \mathbb{R}^n$ note 2(0) = 0 thus $0 \cdot x = 2(0) \cdot x = 2(0 \cdot x)$ which implies $0 \cdot x = 0$. \Box

Definition 9.2.5.

A set S of vectors in \mathbb{R}^n is **orthogonal** iff every pair of vectors in the set is orthogonal. If S is orthogonal and all vectors in S have length one then we say S is orthonormal.

Example 9.2.6. Let u = (1, 1, 0), v = (1, -1, 0) and w = (0, 0, 1). We calculate

 $u \bullet v = 0, \quad u \bullet w, \quad v \bullet w = 0$

thus $S = \{u, v, w\}$ is an orthogonal set. However, it is not orthonormal since $||u|| = \sqrt{2}$. It is easy to create an orthonormal set, we just normalize the vectors; $T = \{\hat{u}, \hat{v}, \hat{w}\}$ meaning,

$$T = \left\{ \frac{1}{\sqrt{2}}(1,1,0), \ \frac{1}{\sqrt{2}}(1,-1,0), \ (0,0,1) \right\}$$

Proposition 9.2.7. Extended Pythagorean Theorem in n-dimensions

If
$$x_1, x_2, \dots x_k$$
 are orthogonal then
 $||x_1||^2 + ||x_2||^2 + \dots + ||x_k||^2 = ||x_1 + x_2 + \dots + x_k||^2$

Proof: we can prove the second statement by applying the Pythagorean Theorem for two vectors repeatedly, starting with

$$||x_1 + (x_2 + \dots + x_k)||^2 = ||x_1||^2 + ||x_2 + \dots + x_k||^2$$

but then we can apply the Pythagorean Theorem to the rightmost term

$$||x_2 + (x_3 + \dots + x_k)||^2 = ||x_2||^2 + ||x_3 + \dots + x_k||^2.$$

Continuing in this fashion until we obtain the Pythagorean Theorem for k-orthogonal vectors. \Box



I have illustrated the proof above in the case of three dimensions and k-dimensions, however my k-dimensional diagram takes a little imagination. Another thing to think about: given $v = v_1e_1 + v_2e_2 + \cdots + v_ne_n$ if e_i are orthonormal then $||v||^2 = v_1^2 + v_2^2 + \cdots + v_n^2$. Therefore, if we use a basis which is orthonormal then we obtain the standard formula for length of a vector with respect to the coordinates. If we were to use a basis of vectors which were not orthogonal or normalized then the formula for the length of a vector in terms of the coordinates could look quite different.

Example 9.2.8. Use the basis $\{v_1 = [1, 1]^T, v_2 = [2, 0]^T\}$ for $\mathbb{R}^{2 \times 1}$. Notice that $\{v_1, v_2\}$ is not orthogonal or normal. Given $x, y \in \mathbb{R}$ we wish to find $a, b \in \mathbb{R}$ such that $r = [x, y]^T = av_1 + bv_2$, this amounts to the matrix calculation:

$$rref[v_1|v_2|r] = rref \begin{bmatrix} 1 & 2 & x \\ 1 & 0 & y \end{bmatrix} = \begin{bmatrix} 1 & 0 & y \\ 0 & 1 & \frac{1}{2}(x-y) \end{bmatrix}$$

Thus a = y and $b = \frac{1}{2}(x - y)$. Let's check my answer,

$$av_1 + bv_2 = y[1,1]^T + \frac{1}{2}(x-y)[2,0]^T = [y+x-y,y+0]^T = [x,y]^T.$$

Furthermore, solving for x, y in terms of a, b yields x = 2b + a and y = a. Therefore, $||[x, y]^T||^2 = x^2 + y^2$ is modified to

$$||av_1 + bv_2||^2 = (2b+a)^2 + a^2 \neq ||av_1||^2 + ||bv_2||^2.$$

If we use a basis which is not orthonormal then we should take care not to assume formulas given for the standard basis equally well apply. However, if we trade the standard basis for a new basis which is orthogonal then we have less to worry about. The Pythagorean Theorem only applies in the orthogonal case. For two normalized, but possibly non-orthogonal, vectors we can replace the Pythagorean Theorem with a generalization of the Law of Cosines in \mathbb{R}^n .

$$||av_1 + bv_2||^2 = a^2 + b^2 + 2ab\cos\theta$$

where $v_1 \cdot v_2 = \cos \theta$. (I leave the proof to the reader)

Proposition 9.2.9.

If $S = \{v_1, v_2, \dots, v_k\} \subset \mathbb{R}^n$ is an orthogonal set of nonzero vectors then S is linearly independent.

Proof: suppose $c_1, c_2, \ldots, c_k \in \mathbb{R}$ such that

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0$$

Take the dot-product of both sides with respect to $v_i \in S$,

$$c_1v_1 \bullet v_j + c_2v_2 \bullet v_j + \dots + c_kv_k \bullet v_j = 0 \bullet v_j = 0$$

Notice all terms in the sum above vanish by orthogonality except for one term and we are left with $c_j v_j \cdot v_j = 0$. However, $v_j \neq 0$ thus $v_j \cdot v_j \neq 0$ and it follows we can divide by the nonzero scalar $v_j \cdot v_j$ leaving $c_j = 0$. But j was arbitrary hence $c_1 = c_2 = \cdots = c_k = 0$ and hence S is linearly independent. \Box

The converse of the proposition above is false. Given a linearly indepdent set of vectors it is not necessarily true that set is also orthogonal. However, we can modify any linearly independent set of vectors to obtain a linearly indepedent set. The procedure for this modification is known as the *Gram-Schmidt orthogonalization*. It is based on a generalization of the idea the vector projection from calculus III. Let me remind you: we found the projection operator to be a useful construction in calculus III. The projection operation allowed us to select the vector component of one vector that pointed in the direction of another given vector. We used this to find the distance from a point to a plane.

Definition 9.2.10.

Let $\vec{A} \neq 0, \vec{B}$ be vectors then we define

 $Proj_{\vec{A}}(\vec{B}) = (\vec{B} \bullet \hat{A})\hat{A}$

where $\hat{A} = \frac{1}{||A||}A$. Moreover, the length of $Proj_{\vec{A}}(\vec{B})$ is called the component of \vec{B} in the \vec{A} -direction and is denoted $Comp_{\vec{A}}(\vec{B}) = ||Proj_{\vec{A}}(\vec{B})||$. Finally, the **orthogonal complement** is defined by $Orth_{\vec{A}}(\vec{B}) = \vec{B} - Proj_{\vec{A}}(\vec{B})$.



Example 9.2.11. Suppose $\vec{A} = \langle 2, 2, 1 \rangle$ and $\vec{B} = \langle 2, 4, 6 \rangle$ notice that we can also express the projection opertation by $\operatorname{Proj}_{\vec{A}}(\vec{B}) = (\vec{B} \cdot \hat{A})\hat{A} = \frac{1}{||\vec{A}||^2}(\vec{B} \cdot \vec{A})\vec{A}$ thus

$$Proj_{\vec{A}}(\vec{B}) = \frac{1}{9}(\langle 2,4,6\rangle \bullet \langle 2,2,1\rangle) \langle 2,2,1\rangle = \frac{4+8+6}{9}\langle 2,2,1\rangle = \langle 4,4,2\rangle$$

The length of the projection vector gives $\operatorname{Comp}_{\vec{A}}(\vec{B}) = \sqrt{16 + 16 + 4} = 6$. One application of this algebra is to calculate the distance from the plane 2x + 2y + z = 0 to the point (2, 4, 6). The "distance" from a plane to a point is defined to be the shortest distance. It's geometrically clear that the shortest path from the plane is found along the normal to the plane. If you draw a picture its not hard to see that $(2, 4, 6) - \operatorname{Proj}_{\vec{A}}(\vec{B}) = \langle 2, 4, 6 \rangle - \langle 4, 4, 2 \rangle = (-2, 0, 4)$ is the closest point to (2, 4, 6) that lies on the plane 2x + 2y + z = 0. Moreover the distance from the plane to the point is just 6.



Example 9.2.12. We studied $\vec{A} = \langle 2, 2, 1 \rangle$ and $\vec{B} = \langle 2, 4, 6 \rangle$ in the preceding example. We found that notice that $\operatorname{Proj}_{\vec{A}}(\vec{B}) = \langle 4, 4, 2 \rangle$. The projection of \vec{B} onto \vec{A} is the part of \vec{B} which points in the direction of \vec{A} . It stands to reason that if we subtract away the projection then we will be left with the part of \vec{B} which does not point in the direction of \vec{A} , it should be orthogonal.

 $Orth_{\vec{A}}(\vec{B}) = \vec{B} - Proj_{\vec{A}}(\vec{B}) = \langle 2, 4, 6 \rangle - \langle 4, 4, 2 \rangle = \langle -2, 0, 4 \rangle$

Let's verify $Orth_{\vec{A}}(\vec{B})$ is indeed orthogonal to \vec{A} ,

$$Orth_{\vec{A}}(\vec{B}) \bullet \vec{A} = \langle -2, 0, 4 \rangle \bullet \langle 2, 2, 1 \rangle = -4 + 4 = 0.$$

Notice that the projection operator has given us the following orthogonal decomposition of \vec{B} :

$$\langle 2,4,6\rangle=\vec{B}=Proj_{\vec{A}}(\vec{B})+Orth_{\vec{A}}(\vec{B})=\langle 4,4,2\rangle+\langle -2,0,4\rangle.$$

If \vec{A}, \vec{B} are any two nonzero vectors it is probably clear that we can perform the decomposition outlined in the example above. It would not be hard to show that if $S = \{\vec{A}, \vec{B}\}$ is linearly independent then $S' = \{\vec{A}, Orth_{\vec{A}}(\vec{B})\}$ is an orthogonal set, moreover they have the same span. This is a partial answer to the converse of Proposition 9.2.9. But, what if we had three vectors instead of two? How would we orthogonalize a set of three linearly independent vectors?

Remark 9.2.13.

I hope you can forgive me for reverting to calculus III notation in the last page or two. It should be clear enough to the reader that the orthogonalization and projection operations can be implemented on either rows or columns. I return to our usual custom of thinking primarily about column vectors at this point. We've already seen the definition from Calculus III, now we turn to the *n*-dimensional case in matrix notation.

Definition 9.2.14.

Suppose $a \neq 0 \in \mathbb{R}^n$, define the **projection of** b **onto** a to be the mapping $Proj_a$: $\mathbb{R}^n \to \mathbb{R}^n$ such that $Proj_a(b) = \frac{1}{a^T a}(a^T b)a$. Moreover, we define $Orth_a : \mathbb{R}^n \to \mathbb{R}^n$ by $Orth_a(b) = b - Proj_a(b) = b - \frac{1}{a^T a}(a^T b)a$ for all $b \in \mathbb{R}^n$.

Proposition 9.2.15.

If
$$a \neq 0 \in \mathbb{R}^n$$
 then $Proj_a$ and $Orth_a$ are linear transformations.

- 1. $Orth_a(b) \bullet a = 0$ for all $b \in \mathbb{R}^n$,
- 2. $Orth_a(b) \bullet Proj_a(y) = 0$ for all $b, y \in \mathbb{R}^n$,
- 3. the projection is idempotent; $Proj_a \circ Proj_a = Proj_a$.

I leave the proof of linearity as an exercise. Begin with (1.): let $a \neq 0 \in \mathbb{R}^n$ and let $b \in \mathbb{R}^n$,

$$a \bullet Orth_a(b) = a^T (b - \frac{1}{a^T a} (a^T b)a)$$

= $a^T b - a^T (\frac{1}{a^T a} (a^T b)a)$
= $a^T b - \frac{1}{a^T a} (a^T b) a^T a$
= $a^T b - a^T b = 0.$

notice I used the fact that $a^T b, a^T a$ were scalars to commute the a^T to the end of the expression. Notice that (2.) follows since $Proj_a(y) = ka$ for some constant k. Next, let $b \in \mathbb{R}^n$ and consider:

$$(Proj_a \circ Proj_a)(b) = Proj_a(Proj_a(b))$$

= $Proj_a(\frac{1}{a^Ta}(a^Tb)a)$
= $\frac{1}{a^Ta}(a^T[\frac{1}{a^Ta}(a^Tb)a])a$
= $\frac{1}{a^Ta}(\frac{a^Tb}{a^Ta}a^Ta)a$
= $\frac{1}{a^Ta}(a^Tb)a$
= $Proj_a(b)$

since the above holds for all $b \in \mathbb{R}^n$ we find $Proj_a \circ Proj_a = Proj_a$. This can also be denoted $Proj_a^2 = Proj_a$. \Box

To create an orthogonal set from a given LI set we just repeated apply the orthogonal projections:

Proposition 9.2.16.

If $S = \{a, b, c\}$ be a linearly independent set of vectors in \mathbb{R}^n then $S' = \{a', b', c'\}$ is an orthogonal set of vectors in \mathbb{R}^n if we define a', b', c' as follows:

a' = a, $b' = Orth_{a'}(b),$ $c' = Orth_{a'}(Orth_{b'}(c)).$

Proof: to prove S' orthogonal we must show that $a' \cdot b' = 0$, $a' \cdot c' = 0$ and $b' \cdot c' = 0$. We already proved $a' \cdot b' = 0$ in the Proposition 9.2.15. Likewise, $a' \cdot c' = 0$ since $Orth_{a'}(x)$ is orthogonal to a' for any x. Consider:

$$b' \bullet c' = b' \bullet Orth_{a'}(Orth_{b'}(c))$$

= b' \u03c6 [Orth_{b'}(c) - Proj_{a'}(Orth_{b'}(c))]
= b' \u03c6 Orth_{b'}(c) - Orth_a(b) \u03c6 Proj_a(Orth_{b'}(c))
= 0

Where we again used (1.) and (2.) of Proposition 9.2.15 in the critical last step. The logic of the formulas is very natural. To construct b' we simply remove the part of b which points in the direction of a'. Then to construct c' we first remove the part of c in the b' direction and then the part in the a' direction. This means no part of c' will point in the a' or b' directions. In principle, one might worry we would subtract away so much that nothing is left, but the linear independence of the vectors insures that is not possible. If it were that would imply a linear dependence of the original set of vectors. \Box

For convenience let me work out the formulas we just discovered in terms of an explicit formula with dot-products. We can also perform the same process for a set of 4 or 5 or more vectors. I'll state the process for arbitrary order, you'll forgive me if I skip the proof this time. There is a careful proof on page 379 of Spence, Insel and Friedberg. The connection between my *Orth* operator approach

and the formulas in the proposition that follows is *just algebra*:

$$\begin{aligned} v_3' &= Orth_{v_1'}(Orth_{v_2'}(v_3)) \\ &= Orth_{v_2'}(v_3) - Proj_{v_1'}(Orth_{v_2'}(v_3)) \\ &= v_3 - Proj_{v_2'}(v_3) - Proj_{v_1'}(v_3 - Proj_{v_2'}(v_3)) \\ &= v_3 - Proj_{v_2'}(v_3) - Proj_{v_1'}(v_3) - Proj_{v_1'}(Proj_{v_2'}(v_3)) \\ &= v_3 - \frac{v_3 \cdot v_2'}{v_2' \cdot v_2'}v_2' - \frac{v_3 \cdot v_1'}{v_1' \cdot v_1'}v_1' \end{aligned}$$

The last term vanished because $v'_1 \cdot v'_2 = 0$ and the projections are just scalar multiples of those vectors.

Proposition 9.2.17. The Gram-Schmidt Process

If $S = \{v_1, v_2, \dots, v_k\}$ is a linearly independent set of vectors in \mathbb{R}^n then $S' = \{v'_1, v'_2, \dots, v'_k\}$ is an orthogonal set of vectors in \mathbb{R}^n if we define v'_i as follows: $v'_1 = v_1$ $v'_2 = v_2 - \frac{v_2 \cdot v'_1}{v'_1 \cdot v'_1} v'_1$ $v'_3 = v_3 - \frac{v_3 \cdot v'_2}{v'_2 \cdot v'_2} v'_2 - \frac{v_3 \cdot v'_1}{v'_1 \cdot v'_1} v'_1$ $v'_k = v_k - \frac{v_k \cdot v'_{k-1}}{v'_{k-1} \cdot v'_{k-1}} v'_{k-1} - \frac{v_k \cdot v'_{k-2}}{v'_{k-2} \cdot v'_{k-2}} v'_{k-2} - \dots - \frac{v_k \cdot v'_1}{v'_1 \cdot v'_1} v'_1.$

Example 9.2.18. Suppose $v_1 = (1, 0, 0, 0), v_2 = (3, 1, 0, 0), v_3 = (3, 2, 0, 3)$. Let's use the Gram-Schmidt Process to orthogonalize these vectors: let $v'_1 = v_1 = (1, 0, 0, 0)$ and calculate:

$$v_2' = v_2 - \frac{v_2 \cdot v_1}{v_1 \cdot v_1} v_1 = (3, 1, 0, 0) - 3(1, 0, 0, 0) = (0, 1, 0, 0).$$

Next,

$$v_3' = v_3 - \frac{v_3 \cdot v_2'}{v_2' \cdot v_2'} v_2' - \frac{v_3 \cdot v_1'}{v_1' \cdot v_1'} v_1' = (3, 2, 0, 3) - 2(0, 1, 0, 0) - 3(1, 0, 0, 0) = (0, 0, 0, 3).$$

We find the orthogonal set of vectors $\{e_1, e_2, e_4\}$. It just so happens this is also an orthonormal set of vectors.

Proposition 9.2.19. Normalization

If $S' = \{v'_1, v'_2, \dots, v'_k\}$ is an orthogonal subset of \mathbb{R}^n then $S'' = \{v''_1, v''_2, \dots, v''_k\}$ is an orthonormal set if we define $v''_i = \widehat{v'_i} = \frac{1}{||v'_i||}v'_i$ for each $i = 1, 2, \dots, k$.

Example 9.2.20. Suppose $v_1 = (1, 1, 1), v_2 = (1, 2, 3), v_3 = (0, 0, 3)$ find an orthonormal set of vectors that spans $span\{v_1, v_2, v_3\}$. We can use Gram-Schmidt followed by a normalization, let $v'_1 = (1, 1, 1)$ then calculate

$$v_2' = (1,2,3) - \left(\frac{1+2+3}{3}\right)(1,1,1) = (1,2,3) - (2,2,2) = (-1,0,1).$$

as a quick check on my arthimetic note $v'_1 \bullet v'_2 = 0$ (good). Next,

$$\begin{split} v_3' &= (0,0,3) - \left(\frac{0(-1) + 0(0) + 3(1)}{2}\right)(-1,0,1) - \left(\frac{0(1) + 0(1) + 3(1)}{3}\right)(1,1,1) \\ \\ \Rightarrow v_3' &= (0,0,3) + (\frac{3}{2},0,-\frac{3}{2}) - (1,1,1) = (\frac{1}{2},-1,\frac{1}{2}) \end{split}$$

again it's good to check that $v'_2 \cdot v'_3 = 0$ and $v'_1 \cdot v'_3 = 0$ as we desire. Finally, note that $||v'_1|| = \sqrt{3}$, $||v'_2|| = \sqrt{2}$ and $||v'_3|| = \sqrt{3/2}$ hence

$$v_1'' = \frac{1}{\sqrt{3}}(1,1,1), \quad v_2'' = \frac{1}{\sqrt{2}}(-1,0,1), \quad v_3'' = \sqrt{\frac{2}{3}}(\frac{1}{2},-1,\frac{1}{2})$$

are orthonormal vectors.

Definition 9.2.21.

A basis for a subspace W of \mathbb{R}^n is an **orthogonal** basis for W iff it is an orthogonal set of vectors which is a basis for W. Likewise, an **orthonormal** basis for W is a basis which is **orthonormal**.

Proposition 9.2.22. Existence of Orthonormal Basis

If $W \leq \mathbb{R}^n$ then there exists an orthonormal basis of W

Proof: since W is a subspace it has a basis. Apply Gram-Schmidt to that basis then normalize the vectors to obtain an orthnormal basis. \Box

Example 9.2.23. Let $W = span\{(1, 0, 0, 0), (3, 1, 0, 0), (3, 2, 0, 3)\}$. Find an orthonormal basis for $W \leq \mathbb{R}^4$. Recall from Example 9.2.18 we applied Gram-Schmidt and found the orthonormal set of vectors $\{e_1, e_2, e_4\}$. That is an orthonormal basis for W.

Example 9.2.24. In Example 9.2.20 we found $\{v_1'', v_2'', v_3''\}$ is an orthonormal set of vectors. Since orthogonality implies linear independence it follows that this set is in fact a basis for $\mathbb{R}^{3\times 1}$. It is an **orthonormal basis**. Of course there are other bases which are orthogonal. For example, the standard basis is orthonormal.

Example 9.2.25. Let us define $S = \{v_1, v_2, v_3, v_4\} \subset \mathbb{R}^4$ as follows:

$$v_{1} = \begin{bmatrix} 1\\0\\1\\1 \end{bmatrix}, v_{2} = \begin{bmatrix} 1\\1\\1\\1 \end{bmatrix}, v_{3} = \begin{bmatrix} 0\\0\\2\\3 \end{bmatrix}, v_{4} = \begin{bmatrix} 3\\2\\0\\3 \end{bmatrix}$$

It is easy to verify that S defined below is a linearly independent set vectors basis for $span(S) \leq \mathbb{R}^{4\times 1}$. Let's see how to find an orthonormal basis for span(S). The procedure is simple: apply the

Gram-Schmidt algorithm then normalize the vectors.

$$\begin{aligned} v_1' &= v_1 = \begin{bmatrix} 1\\0\\1\\1\\1 \end{bmatrix} \\ v_2' &= v_2 - \left(\frac{v_2 \cdot v_1'}{v_1' \cdot v_1'}\right) v_1' = \begin{bmatrix} 1\\1\\1\\1\\1 \end{bmatrix} - \frac{3}{3} \begin{bmatrix} 1\\0\\1\\1\\1 \end{bmatrix} = \begin{bmatrix} 0\\1\\0\\0 \end{bmatrix} \\ v_3' &= v_3 - \left(\frac{v_3 \cdot v_2'}{v_2' \cdot v_2'}\right) v_2' - \left(\frac{v_3 \cdot v_1'}{v_1' \cdot v_1'}\right) v_1' = \begin{bmatrix} 0\\0\\2\\3 \end{bmatrix} - \frac{0}{1} \begin{bmatrix} 0\\1\\0\\0 \end{bmatrix} - \frac{5}{3} \begin{bmatrix} 1\\0\\1\\1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -5\\0\\1\\4 \end{bmatrix} \\ v_4' &= v_4 - \left(\frac{v_4 \cdot v_3'}{v_3' \cdot v_3'}\right) v_3' - \left(\frac{v_3 \cdot v_2'}{v_2' \cdot v_2'}\right) v_2' - \left(\frac{v_3 \cdot v_1'}{v_1' \cdot v_1'}\right) v_1' \\ &= \begin{bmatrix} 3\\2\\0\\3 \end{bmatrix} - \frac{1}{14} \begin{bmatrix} -5\\0\\1\\4 \end{bmatrix} - \begin{bmatrix} 0\\2\\0\\0 \end{bmatrix} - \begin{bmatrix} 2\\0\\2\\2 \end{bmatrix} = \frac{1}{14} \begin{bmatrix} 9\\0\\-27\\18 \end{bmatrix} \end{aligned}$$

Then normalize to obtain the orthonormal basis for Span(S) below:

$$\beta = \frac{1}{\sqrt{3}} \begin{bmatrix} 1\\0\\1\\1 \end{bmatrix}, \begin{bmatrix} 0\\1\\0\\0 \end{bmatrix}, \frac{1}{\sqrt{42}} \begin{bmatrix} -5\\0\\1\\4 \end{bmatrix}, \frac{1}{9\sqrt{14}} \begin{bmatrix} 9\\0\\-27\\18 \end{bmatrix} \right\}$$

Proposition 9.2.26. Coordinates with respect to an Orthonormal Basis

If $W \leq \mathbb{R}^n$ has an orthonormal basis $\{v_1, v_2, \dots, v_k\}$ and if $w = \sum_{i=1}^k w_i v_i$ then $w_i = w \cdot v_i$ for all $i = 1, 2, \dots, k$. In other words, each vector $w \in W$ may be expressed as $w = (w \cdot v_1)v_1 + (w \cdot v_2)v_2 + \dots + (w \cdots v_k)v_k$

Proof: Let $w = w_1v_1 + w_2v_2 + \cdots + w_kv_k$ and take the dot-product with v_j ,

$$w \bullet v_j = (w_1v_1 + w_2v_2 + \dots + w_kv_k) \bullet v_j = w_1(v_1 \bullet v_j) + w_2(v_2 \bullet v_j) + \dots + w_k(v_k \bullet v_j)$$

Orthonormality of the basis is compactly expressed by the Kronecker Delta; $v_i \cdot v_j = \delta_{ij}$ this is zero if $i \neq j$ and it is 1 if they are equal. The whole sum collapses except for the *j*-th term which yields: $w \cdot v_j = w_j$. But, *j* was arbitrary hence the proposition follows. \Box .

The proposition above reveals the real reason we like to work with orthonormal coordinates. It's easy to figure out the coordinates, we simply take dot-products. This technique was employed with great success in (you guessed it) Calculus III. The standard $\{\hat{i}, \hat{j}, \hat{k}\}$ is an orthonormal basis and one of the first things we discuss is that if $\vec{v} = \langle A, B, C \rangle$ then $A = \vec{v} \cdot \hat{i}$, $B = \vec{v} \cdot \hat{j}$ and $C = \vec{v} \cdot \hat{k}$.

Example 9.2.27. For the record, the standard basis of \mathbb{R}^n is an orthonormal basis and

$$v = (v \bullet e_1)e_1 + (v \bullet e_2)e_2 + \dots + (v \bullet e_n)e_n$$

for any vector v in \mathbb{R}^n .

Example 9.2.28. Let v = [1, 2, 3, 4]. Find the coordinates of v with respect to the orthonormal basis β found in Example 9.2.25.

$$\beta = \{f_1, f_2, f_3, f_4\} = \left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1\\0\\1\\1 \end{bmatrix}, \begin{bmatrix} 0\\1\\0\\0 \end{bmatrix}, \frac{1}{\sqrt{42}} \begin{bmatrix} -5\\0\\1\\4 \end{bmatrix}, \frac{1}{\frac{1}{9\sqrt{14}}} \begin{bmatrix} 9\\0\\-27\\18 \end{bmatrix} \right\}$$

Let us denote the coordinates vector $[v]_{\beta} = [w_1, w_2, w_3, w_4]$ we know we can calculate these by taking the dot-products with the vectors in the orthonormal basis β :

$$w_{1} = v \bullet f_{1} = \frac{1}{\sqrt{3}} [1, 2, 3, 4] [1, 0, 1, 1]^{T} = \frac{8}{\sqrt{3}}$$
$$w_{2} = v \bullet f_{2} = [1, 2, 3, 4] [0, 1, 0, 0]^{T} = 2$$
$$w_{3} = v \bullet f_{3} = \frac{1}{\sqrt{42}} [1, 2, 3, 4] [-5, 0, 1, 4]^{T} = \frac{14}{\sqrt{42}}$$
$$w_{4} = v \bullet f_{4} = \frac{1}{9\sqrt{14}} [1, 2, 3, 4] [9, 0, -27, 18]^{T} = \frac{0}{9\sqrt{14}} = 0$$

Therefore, $[v]_{\beta} = [\frac{8}{\sqrt{3}}, 2, \frac{14}{\sqrt{42}}, 0]$. Now, let's check our answer. What should this mean if it is correct? We should be able verify $v = w_1 f_1 + w_2 f_2 + w_3 f_3 + w_4 f_4$:

$$w_{1}f_{1} + w_{2}f_{2} + w_{3}f_{3} + w_{4}f_{4} = \frac{8}{\sqrt{3}} \frac{1}{\sqrt{3}} \begin{bmatrix} 1\\0\\1\\1\\1 \end{bmatrix} + 2\begin{bmatrix} 0\\1\\0\\0 \end{bmatrix} + \frac{14}{\sqrt{42}} \frac{1}{\sqrt{42}} \begin{bmatrix} -5\\0\\1\\4 \end{bmatrix}$$
$$= \frac{8}{3} \begin{bmatrix} 1\\0\\1\\1\\1 \end{bmatrix} + 2\begin{bmatrix} 0\\1\\0\\0 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} -5\\0\\1\\4 \end{bmatrix}$$
$$= \begin{bmatrix} 8/3 - 5/3\\2\\8/3 + 1/3\\8/3 + 4/3 \end{bmatrix}$$
$$= \begin{bmatrix} 1\\2\\3\\4 \end{bmatrix}$$

Well, that's a relief.

9.3 orthogonal complements and projections

Upto now we have discussed projections with respect to one vector at a time, however we can just as well discuss the projection onto some subspace of \mathbb{R}^n . We need a few definitions to clarify and motivate the projection.

Definition 9.3.1.

Suppose $W_1, W_2 \subseteq \mathbb{R}^n$ then we say W_1 is **orthogonal** to W_2 iff $w_1 \cdot w_2 = 0$ for all $w_1 \in W_1$ and $w_2 \in W_2$. We denote orthogonality by writing $W_1 \perp W_2$.

Example 9.3.2. Let $W_1 = span\{e_1, e_2\}$ and $W_2 = span\{e_3\}$ then $W_1, W_2 \leq \mathbb{R}^n$. Let $w_1 = ae_1 + be_2 \in W_1$ and $w_2 = ce_3 \in W_2$ calculate,

$$w_1 \bullet w_2 = (ae_1 + be_2) \bullet (ce_3) = ace_1 \bullet e_3 + bce_2 \bullet e_3 = 0$$

Hence $W_1 \perp W_2$. Geometrically, we have shown the xy-plane is orthogonal to the z-axis.

We notice that orthogonality relative to the basis will naturally extend to the span of the basis since the dot-product has nice linearity properties.

Proposition 9.3.3.

Suppose $W_1, W_2 \leq \mathbb{R}^n$ the subspace W_1 is **orthogonal** to the subspace W_2 iff $w_i \cdot v_j = 0$ for all i, j relative to a pair of bases $\{w_i\}$ for W_1 and $\{v_j\}$ for W_2 .

Proof: Suppose $\{w_i\}_{i=1}^r$ is a basis for $W_1 \leq \mathbb{R}^n$ and $\{v_j\}_{j=1}^s$ for $W_2 \leq \mathbb{R}^n$. If $W_1 \perp W_2$ then clearly $\{w_i\}_{i=1}^r$ is orthogonal to $\{v_j\}_{j=1}^s$. Conversely, suppose $\{w_i\}_{i=1}^r$ is orthogonal to $\{v_j\}_{j=1}^s$ then let $x \in W_1$ and $y \in W_2$:

$$x \bullet y = \left(\sum_{i=1}^r x_i w_i\right) \bullet \left(\sum_{i=1}^s y_j w_j\right) = \sum_{i=1}^r \sum_{j=1}^s x_i y_j (w_i \bullet v_j) = 0. \qquad \Box$$

Given a subspace W which lives in \mathbb{R}^n we might wonder what is the largest subspace which is orthogonal to W? In $\mathbb{R}^{3\times 1}$ it is clear that the xy-plane is the largest subspace which is orthogonal to the z-axis, however, if the xy-plane was viewed as a subset of $\mathbb{R}^{4\times 1}$ we could actually find a volume which was orthogonal to the z-axis (in particular $span\{e_1, e_2, e_4\} \perp span\{e_3\}$).

Definition 9.3.4.

Let $W \subseteq \mathbb{R}^n$ then W^{\perp} is defined as follows:

$$W^{\perp} = \{ v \in \mathbb{R}^n | v \bullet w = 0 \text{ for all } w \in W \}$$

It is clear that W^{\perp} is the largest subset in \mathbb{R}^n which is orthogonal to W. Better than just that, it's the largest subspace orthogonal to W.

Proposition 9.3.5.

Let $S \subset \mathbb{R}^n$ then $S^{\perp} \leq \mathbb{R}^n$.

Proof: Let $x, y \in S^{\perp}$ and let $c \in \mathbb{R}$. Furthermore, suppose $s \in S$ and note

$$(x + cy) \cdot s = x \cdot s + c(y \cdot s) = 0 + c(0) = 0$$

Thus an aribitrary linear combination of elements of S^{\perp} are again in S^{\perp} which is nonempty as $0 \in S^{\perp}$ hence by the subspace test $S^{\perp} \leq \mathbb{R}^n$. It is interesting that S need not be a subspace for this argument to hold. \Box

Example 9.3.6. Find the orthogonal complement to $W = span\{v_1 = (1, 1, 0, 0), v_2 = (0, 1, 0, 2)\}$. Let's treat this as a matrix problem. We wish to describe a typical vector in W^{\perp} . Towards that goal, let $r = (x, y, z, w) \in W^{\perp}$ then the conditions that r must satisfy are $v_1 \cdot r = v_1^T r = 0$ and $v_2 \cdot r = v_2^T r = 0$. But this is equivalent to the single matrix equation below:

$\left[\begin{array}{rrrrr} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 2 \end{array}\right] \left[$	$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$	\Rightarrow $r =$	$\begin{bmatrix} 2w \\ -2w \\ z \\ w \end{bmatrix}$	= z	$\left[\begin{array}{c}0\\0\\1\\0\end{array}\right]$	+w	$\begin{array}{c}2\\-2\\0\\1\end{array}$
---	---	---------------------	---	-----	--	----	--

Thus, $W^{\perp} = span\{(0,0,1,0), (2,-2,0,1)\}.$

If you study the preceding example it becomes clear that finding the orthogonal complement of a set of vectors is equivalent to calculating the null space of a particular matrix. We have considerable experience in such calculations so this is a welcome observation.

Proposition 9.3.7.

If
$$S = \{v_1, v_2, \dots, v_k\} \subseteq \mathbb{R}^n$$
 and $A = [v_1|v_2|\cdots|v_k]$ then $S^{\perp} = Null(A^T)$

Proof: Denote $A = [v_1|v_2|\cdots|v_k] \in \mathbb{R}^{n \times k}$ and $x = [x_1, x_2, \dots, x_k]^T$. Observe that:

$$\begin{aligned} x \in Null(A^T) \Leftrightarrow A^T x &= 0 \\ \Leftrightarrow [row_1(A^T)x, row_2(A^T)x, \cdots, row_k(A^T)x] &= 0 \\ \Leftrightarrow [(col_1(A))^T x, (col_2(A))^T x, \cdots, (col_k(A))^T x] &= 0 \\ \Leftrightarrow [v_1 \bullet x, v_2 \bullet x, \cdots, v_k \bullet x] &= 0 \\ \Leftrightarrow v_j \bullet x &= 0 \text{ for } j = 1, 2, \dots, k \\ \Leftrightarrow x \in S^{\perp} \end{aligned}$$

Therefore, $Null(A^T) = S^{\perp}$. \Box

Given the correspondence above we should be interested in statements which can be made about the row and column space of a matrix. It turns out there are two simple statements to be made in general:

Proposition 9.3.8.

Let $A \in \mathbb{R}^{m \times n}$ then 1. $Null(A^T) \perp Col(A)$. 2. $Null(A) \perp Row(A)$.

Proof: Let $S = \{col_1(A), col_2(A), \dots, col_n(A)\}$ and use Proposition 9.3.7 to deduce $S^{\perp} = Null(A^T)$. Therefore, each column of A is orthogonal to all vectors in $Null(A^T)$, in particular each column is orthogonal to the basis for $Null(A^T)$. Since the pivot columns are a basis for Col(A) we can use Proposition 9.3.3 to conclude $Null(A^T) \perp Col(A)$.

To prove of (2.) apply (1.) to $B = A^T$ to deduce $Null(B^T) \perp Col(B)$. Hence, $Null((A^T)^T) \perp Col(A^T)$ and we find $Null(A) \perp Col(A^T)$. But, $Col(A^T) = Row(A)$ thus $Null(A) \perp Row(A)$. \Box

The proof above makes ample use of previous work. I encourage the reader to try to prove this proposition from scratch. I don't think it's that hard and you might learn something. Just take an arbitrary element of each subspace and argue why the dot-product is zero.

Proposition 9.3.9.

Let $W_1, W_2 \leq \mathbb{R}^n$, if $W_1 \perp W_2$ then $W_1 \cap W_2 = \{0\}$

Proof: let $z \in W_1 \cap W_2$ then $z \in W_1$ and $z \in W_2$ and since $W_1 \perp W_2$ it follows $z \cdot z = 0$ hence z = 0 and $W_1 \cap W_2 \subseteq \{0\}$. The reverse inclusion $\{0\} \subseteq W_1 \cap W_2$ is clearly true since 0 is in every subspace. Therefore, $W_1 \cap W_2 = \{0\} \square$

We defined the direct sum of two subspaces in the Section 7.7. The fact that $W_1 + W_2 = V$ and $W_1 \cap W_2 = \{0\}$ was sufficient to prove $V \approx W_1 \times W_2$ so, by our definition, we can write $V = W_1 \oplus W_2$. The theorem below is at the heart of many geometric arguments in multivariate calculus. Intuitively I think of it like this: if we show $x \notin W$ then by process of elimination it must be in W^{\perp} . Intuition fails unless W^{\perp} is a **complementary subspace**. We say W_1 and W_2 are complementary subspaces of V iff $V = W_1 \oplus W_2$.

Theorem 9.3.10.

Let $W \leq \mathbb{R}^n$ then 1. $\mathbb{R}^n = W \oplus W^{\perp}$. 2. $dim(W) + dim(W^{\perp}) = n$, 3. $(W^{\perp})^{\perp} = W$.

Proof: Let $W \leq \mathbb{R}^n$ and choose an orthonormal basis $\beta = \{v_1, v_2, \dots, v_k\}$ for S. Let $z \in \mathbb{R}^n$ and define

$$Proj_W(z) = \sum_{i=1}^{\kappa} (z \cdot v_i) v_i$$
 and $Orth_W(z) = z - Proj_W(z).$

Observe that $z = Proj_W(z) + Orth_W(z)$ and clearly $Proj_W(z) \in S$. We now seek to argue that $Orth_W(z) \in S^{\perp}$. Let $v_j \in \beta$ then

$$v_{j} \cdot Orth_{W}(z) = v_{j} \cdot (z - Proj_{W}(z))$$

$$= v_{j} \cdot z - v_{j} \cdot \left(\sum_{i=1}^{k} (z \cdot v_{i})v_{i}\right)$$

$$= v_{j} \cdot z - \sum_{i=1}^{k} (z \cdot v_{i})(v_{j} \cdot v_{i})$$

$$= v_{j} \cdot z - \sum_{i=1}^{k} (z \cdot v_{i})\delta_{ij}$$

$$= v_{j} \cdot z - z \cdot v_{j}$$

$$= 0$$

Therefore, $\mathbb{R}^n = W \oplus W^{\perp}$. To prove (2.) notice we know by Proposition 9.3.5 that $W^{\perp} \leq \mathbb{R}^n$ and consequently there exists an orthonormal basis $\Gamma = \{w_1, w_2, \ldots, w_l\}$ for W^{\perp} . Furthermore, by Proposition 9.3.9 we find $\beta \cap \Gamma = \emptyset$ since 0 is not in either basis. We argue that $\beta \cup \Gamma$ is a basis for \mathbb{R}^n . Observe that $\beta \cup \Gamma$ clearly spans \mathbb{R}^n since $z = \operatorname{Proj}_W(z) + \operatorname{Orth}_W(z)$ for each $z \in \mathbb{R}^n$ and $\operatorname{Proj}_W(z) \in \operatorname{span}(\beta)$ while $\operatorname{Orth}_W(z) \in \operatorname{span}(\Gamma)$. Furthermore, I argue that $\beta \cup \Gamma$ is an orthonormal set. By construction β and Γ are orthonormal, so all we need prove is that the dot-product of vectors from β and Γ is zero, but that is immediate from the construction of Γ . We learned in Proposition 9.2.9 that orthogonality for set of nonzero vectors implies linearly independence. Hence, $\beta \cup \Gamma$ is a linearly independent spanning set for \mathbb{R}^n . By the dimension theorem we deduce that there must be *n*-vectors in $\beta \cup \Gamma$ since it must have the same number of vectors as any other basis for \mathbb{R}^n (the standard basis obviously has *n*-vectors). Therefore,

$$\dim(W) + \dim(W^{\perp}) = n.$$

in particular, we count $dim(W^{\perp}) = n - k$ in my current notation. Now turn to ponder the proof of (3.). Let $z \in (W^{\perp})^{\perp}$ and expand z in the basis $\beta \cup \Gamma$ to gain further insight, $z = z_1v_1 + z_2v_2 + \cdots + z_kv_k + z_{k+1}w_1 + z_{k+2}w_2 + \cdots + z_nw_{n-k}$. Since $z \in (W^{\perp})^{\perp}$ then $z \cdot w_{\perp} = 0$ for all $w_{\perp} \in W^{\perp}$, in particular $z \cdot w_j = 0$ for all $j = 1, 2, \ldots, n-k$. But, this implies $z_{k+1} = z_{k+2} = \cdots = z_n = 0$ since Proposition 9.2.26 showed the coordinates w.r.t. an orthonormal basis are given by dot-products. Therefore, $z \in span(\beta) = W$ and we have shown $(W^{\perp})^{\perp} \subseteq W$. In invite the reader to prove the reverse inclusion to complete this proof. \Box

Two items I defined for the purposes of the proof above have application far beyond the proof. Let's state them again for future reference. I give two equivalent definitions, technically we should prove that the second basis dependent statement follows from the first basis-independent statement. Primary definitions are, as a point of mathematical elegance, stated in a coordinate free langauge in as much as possible. However the second statement is how we calculate projections in many cases.

Definition 9.3.11.

Let $W \leq \mathbb{R}^n$ if $z \in \mathbb{R}^n$ and z = u + w for some $u \in W$ and $w \in W^{\perp}$ then we define $u = Proj_W(z)$ and $w = Orth_W(z)$. Equivalently, choose an orthonormal basis $\beta = \{v_1, v_2, \ldots v_k\}$ for W then if $z \in \mathbb{R}^n$ we define

$$Proj_W(z) = \sum_{i=1}^k (z \cdot v_i)v_i$$
 and $Orth_W(z) = z - Proj_W(z).$

Perhaps the following picture helps: here I show projections onto a plane with basis $\{\vec{u}_1, \vec{u}_2\}$ and its normal \vec{n} .



Example 9.3.12. Let $W = span\{e_1 + e_2, e_3\}$ and x = (1, 2, 3) calculate $Proj_W(x)$. To begin I note that the given spanning set is orthogonal and hence linear indpendent. We need only orthonormalize to obtain an orthonormal basis β for W

$$\beta = \{v_1, v_2\}$$
 with $v_1 = \frac{1}{\sqrt{2}}(1, 1, 0), v_2 = (0, 0, 1)$

Calculate, $v_1 \cdot x = \frac{3}{\sqrt{2}}$ and $v_2 \cdot x = 3$. Thus,

$$Proj_W((1,2,3)) = (v_1 \bullet x)v_1 + (v_2 \bullet x)v_2 = \frac{3}{\sqrt{2}}v_1 + 3v_2 = (\frac{3}{2}, \frac{3}{2}, 3)$$

Then it's easy to calculate the orthogonal part,

$$Orth_W((1,2,3)) = (1,2,3) - (\frac{3}{2},\frac{3}{2},3) = (-\frac{1}{2},\frac{1}{2},0)$$

As a check on the calculation note that $Proj_W(x) + Orth_W(x) = x$ and $Proj_W(x) \cdot Orth_W(x) = 0$.

Example 9.3.13. Let $W = span\{u_1, u_2, u_3\} \leq \mathbb{R}^4$ where

$$u_{1} = \begin{bmatrix} 2\\1\\2\\0 \end{bmatrix} \qquad u_{2} = \begin{bmatrix} 0\\-2\\1\\1 \end{bmatrix} \qquad u_{3} = \begin{bmatrix} -1\\2\\0\\-1 \end{bmatrix}$$

calculate $Proj_W([0, 6, 0, 6]^T)^2$. Notice that the given spanning set appears to be linearly independent but it is not orthogonal. Apply Gram-Schmidt to fix it:

$$\begin{aligned} v_1 &= u_1 = [2, 1, 2, 0]^T \\ v_2 &= u_2 - \frac{u_2 \cdot v_1}{v_1 \cdot v_1} v_1 = u_2 = [0, -2, 1, 1]^T \\ v_3 &= u_3 - \frac{u_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{u_3 \cdot v_2}{v_2 \cdot v_2} v_2 = u_3 + \frac{5}{6} v_2 = [-1, 2, 0, -1]^T + [0, -\frac{10}{6}, \frac{5}{6}, \frac{5}{6}]^T \end{aligned}$$

We calculate,

$$v_3 = \begin{bmatrix} -1, \ 2 - \frac{5}{3}, \ \frac{5}{6}, \ -1 + \frac{5}{6} \end{bmatrix}^T = \begin{bmatrix} -1, \ \frac{1}{3}, \ \frac{5}{6}, \ -\frac{1}{6} \end{bmatrix}^T = \frac{1}{6} \begin{bmatrix} -6, 2, 5, -1 \end{bmatrix}^T$$

The normalized basis follows easily,

$$v_1' = \frac{1}{3}[2, 1, 2, 0]^T$$
 $v_2' = \frac{1}{\sqrt{6}}[0, -2, 1, 1]^T$ $v_3' = \frac{1}{\sqrt{66}}[-6, 2, 5, -1]^T$

Calculate dot-products in preparation for the projection calculation,

$$v'_1 \bullet x = \frac{1}{3} [2, 1, 2, 0] [0, 6, 0, 6]^T = 2$$
$$v'_2 \bullet x = \frac{1}{\sqrt{6}} [0, -2, 1, 1] [0, 6, 0, 6]^T = \frac{1}{\sqrt{6}} (-12 + 6) = -\sqrt{6}$$
$$v'_3 \bullet x = \frac{1}{\sqrt{66}} [-6, 2, 5, -1] [0, 6, 0, 6]^T = \frac{1}{\sqrt{66}} (12 - 6) = \frac{6}{\sqrt{66}}$$

Now we calculate the projection of $x = [0, 6, 0, 6]^T$ onto W with ease:

$$\begin{aligned} Proj_W(x) &= (x \bullet v_1')v_1' + (x \bullet v_2')v_2' + (x \bullet v_3')v_3' \\ &= (2)\frac{1}{3}[2, 1, 2, 0]^T - (\sqrt{6})\frac{1}{\sqrt{6}}[0, -2, 1, 1]^T + (\frac{6}{\sqrt{66}})\frac{1}{\sqrt{66}}[-6, 2, 5, -1]^T \\ &= [\frac{4}{3}, \frac{2}{3}, \frac{4}{3}, 0]^T + [0, 2, -1, -1]^T + [\frac{-6}{11}, \frac{2}{11}, \frac{5}{11}, \frac{-1}{11}]^T \\ &= [\frac{26}{33}, \frac{94}{33}, \frac{26}{33}, \frac{-36}{33}]^T \end{aligned}$$

and,

$$Orth_W(x) = \begin{bmatrix} \frac{-26}{33}, & \frac{104}{33}, & \frac{-26}{33}, & \frac{234}{33} \end{bmatrix}^T$$

²this problem is inspired from Anton & Rorres' §6.4 homework problem 3 part d.Sorry about the notation here, I'm afraid I'll make a typo as I change it so here it stays.

9.4 orthogonal transformations and geometry

If we begin with an orthogonal subset of \mathbb{R}^n and we preform a linear transformation then will the image of the set still be orthogonal? We would like to characterize linear transformations which maintain orthogonality. These transformations should take an orthogonal basis to a new basis which is still orthogonal.

Definition 9.4.1.

If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation such that $T(x) \bullet T(y) = x \bullet y$ for all $x, y \in \mathbb{R}^n$ then we say that T is an **orthogonal transformation**

Example 9.4.2. Let $\{e_1, e_2\}$ be the standard basis for \mathbb{R}^2 and let $R(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ be a rotation of the coordinates by angle θ in the clockwise direction,

$$\begin{bmatrix} x'\\y'\end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta\\\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x\\y\end{bmatrix} = \begin{bmatrix} x\cos\theta + y\sin\theta\\-x\sin\theta + y\cos\theta\end{bmatrix}$$

As a check on my sign conventions, consider rotating (1,0) by $R(\pi/2)$, we obtain (x', y') = (0, 1). Intuitively, a rotation should not change the length of a vector, let's check the math: let $v, w \in \mathbb{R}^2$,

$$R(\theta)v \bullet R(\theta)w = [R(\theta)v]^T R(\theta)w$$
$$= v^T R(\theta)^T R(\theta)w$$

Now calculate $R(\theta)^T R(\theta)$,

$$R(\theta)^T R(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \sin^2\theta + \cos^2\theta \end{bmatrix} = I$$

Therefore, $R(\theta)v \bullet R(\theta) = v^T I w = v^T w = v \bullet w$ for all $v, w \in \mathbb{R}^2$ and we find $L_{R(\theta)}$ is an orthogonal transformation.

This shows the matrix of a rotation L_R satisfies $R^T R = I$. Is this always true or was this just a special formula for rotations? Or is this just a two-dimensional thing? What if we look at orthhogonal transformations on \mathbb{R}^n what general condition is there on the matrix of the transformation?

Definition 9.4.3.

Let $A \in \mathbb{R}^{n \times n}$ then we say that A is an **orthogonal matrix** iff $A^T A = I$. Moreover, we say A is a **reflection matrix** if A is orthogonal and det(A) = -1 whereas we say A is a **rotation matrix** if A is orthogonal with det(A) = 1. The set of all orthogonal $n \times n$ matrices is denoted O(n) and the set of all $n \times n$ rotation matrices is denoted SO(n).

Proposition 9.4.4. matrix of an orthogonal transformation is orthogonal

If A is the matrix of an orthogonal transformation on \mathbb{R}^n then $A^T A = I$ and either A is a rotation matrix or A is a reflection matrix.

Proof: Suppose L(x) = Ax and L is an orthogonal transformation on \mathbb{R}^n . Notice that

$$L(e_i) \bullet L(e_j) = [Ae_i]^T Ae_j = e_i^T [A^T A]e_j$$

and

$$e_i \bullet e_j = e_i^T e_j = e_i^T I e_j$$

hence $e_i^T [A^T A - I] e_j = 0$ for all i, j thus $A^T A - I = 0$ by Example 3.3.11 and we find $A^T A = I$. Following a homework you did earlier in the course,

$$det(A^T A) = det(I) \iff det(A)det(A) = 1 \iff det(A) = \pm 1$$

Thus $A \in SO(n)$ or A is a reflection matrix. \Box

The proposition below is immediate from the definitions of length, angle and linear transformation.

Proposition 9.4.5. orthogonal transformations preserve lengths and angles

If $v, w \in \mathbb{R}^n$ and L is an orthogonal transformation such that v' = L(v) and w' = L(w) then the angle between v' and w' is the same as the angle between v and w, in addition the length of v' is the same as v.

Remark 9.4.6.

Reflections, unlike rotations, will spoil the "handedness" of a coordinate system. If we take a right-handed coordinate system and perform a reflection we will obtain a new coordinate system which is left-handed. If you'd like to know more just ask me sometime.

If orthogonal transformations preserve the geometry of \mathbb{R}^n you might wonder if there are other non-linear transformations which also preserve distance and angle. The answer is yes, but we need to be careful to distinguish between the length of a vector and the distance between points. It turns out that the translation defined below will preserve the distance, but not the norm or length of a vector.

Definition 9.4.7.

Fix $b \in \mathbb{R}^n$ then a translation by b is the mapping $T_b(x) = x + b$ for all $x \in \mathbb{R}^n$.

This is known as an **affine transformation**, it is not linear since $T(0) = b \neq 0$ in general. (if b = 0 then the translation is both affine and linear). Anyhow, affine transformations should be familar to you: y = mx + b is an affine transformation on \mathbb{R} .

Proposition 9.4.8. translations preserve geometry

Suppose $T_b : \mathbb{R}^n \to \mathbb{R}^n$ is a translation then

- 1. If $\angle (xyz)$ denotes the angle formed by line segments \bar{xy}, \bar{yz} which have endpoints x, yand y, z respectively then $\angle (T_b(x)T_b(y)T_b(z)) = \angle (xyz)$
- 2. The distance from x to y is the equal to the distance from $T_b(x)$ to $T_b(y)$.

Proof: I'll begin with (2.) since it's easy:

$$d(T_b(x), T_b(y)) = ||T_b(y) - T_b(x)|| = ||y + b - (x + b)|| = ||y - x|| = d(x, y)$$

Next, the angle $\angle (xyz)$ is the angle between x - y and z - y. Likewise the angle $\angle T_b(x)T_b(y)T_b(z)$ is the angle between $T_b(x) - T_b(y)$ and $T_b(z) - T_b(y)$. But, these are the same vectors since $T_b(x) - T_b(y) = x + b - (y + b) = x - y$ and $T_b(z) - T_b(y) = z + b - (y + b) = z - y$. \Box

Definition 9.4.9.

Suppose T(x) = Ax + b where $A \in SO(n)$ and $b \in \mathbb{R}^n$ for all $x \in \mathbb{R}^n$ then we say T is a **rigid motion**.

In high-school geometry you studied the concept of *congruence*. To objects were congruent if they had the same size and shape. From the viewpoint of analytic geometry we can say two objects are congruent iff one is the image of the other with respect to some rigid motion. We leave further discussion of such matters to the modern geometry course where you study these concepts in depth.

Remark 9.4.10.

In Chapter 6 of my *Mathematical Models in Physics* notes I describe how Euclidean geometry is implicit and foundational in classical Newtonian Mechanics. The concept of a rigid motion is used to define what is meant by an *intertial frame*. I have these notes posted on my website, ask if your interested. Chapter 7 of the same notes describes how Special Relativity has hyperbolic geometry as its core. The dot-product is replaced with a Minkowski-product which yields all manner of curious results like time-dilation, length contraction, and the constant speed of light. If your interested in hearing a lecture or two on the geometry of Special Relativity please ask and I'll try to find a time and a place, I mean, we'll make it an event.

This concludes our short tour of Euclidean geometry. Incidentally, you might look at Barret Oneil's *Elementary Differential Geometry* if you'd like to see a more detailed study of isometries of \mathbb{R}^3 . Some notes are posted on my website from the Math 497, Spring 2014 course. We now generalize to inner product spaces which include the dot-product as a particular case. The dot-product is the most important and common inner-product, however it is not the only case of interest in this course.

9.5 least squares analysis

In this section we consider results which ultimately show how to find the best approximation to problems which have no exact solution. In other words, we consider how to almost solve inconsistent systems in the best way possible.

9.5.1 the closest vector problem

Suppose we are given a subspace and a vector not in the subspace, which vector in the subspace is closest to the external vector ? Naturally the projection answers this question. The projection of the external vector onto the subspace will be closest. Let me be a bit more precise:

Proposition 9.5.1. Closest vector inequality.

If $S \leq \mathbb{R}^n$ and $b \in \mathbb{R}^n$ such that $b \notin S$ then for all $u \in S$ with $u \neq Proj_S(b)$,

 $||b - Proj_S(b)|| < ||b - u||.$

This means $Proj_S(b)$ is the closest vector to b in S.

Proof: Noice that $b - u = b - Proj_S(b) + Proj_S(b) - u$. Furthermore note that $b - Proj_S(b) = Orth_S(b) \in S^{\perp}$ whereas $Proj_S(b) - u \in S$ hence these are orthogonal vectors and we can apply the Pythagorean Theorem,

$$||b - u||^{2} = ||b - Proj_{S}(b)||^{2} + ||Proj_{S}(b) - u||^{2}$$

Notice that $u \neq Proj_S(b)$ implies $Proj_S(b) - u \neq 0$ hence $||Proj_S(b) - u||^2 > 0$. It follows that $||b - Proj_S(b)||^2 < ||b - u||^2$. And as the $|| \cdot ||$ is nonnegative³ we can take the squareroot to obtain $||b - Proj_S(b)|| < ||b - u||$. \Box

Remark 9.5.2.

In calculus III I show at least three distinct methods to find the point off a plane which is closest to the plane. We can minimize the distance function via the 2nd derivative test for two variables, or use Lagrange Multipliers or use the geometric solution which invokes the projection operator. It's nice that we have an explicit proof that the geometric solution is valid. We had argued on the basis of geometric intuition that $Orth_S(b)$ is the shortest vector from the plane S to the point b off the plane⁴ Now we have proof. Better yet, our proof equally well applies to subspaces of \mathbb{R}^n . In fact, this discussion extends to the context of inner product spaces.

Example 9.5.3. Consider \mathbb{R}^2 let $S = span\{(1,1)\}$. Find the point on the line S closest to the point (4,0).

$$Proj_S((4,0)) = \frac{1}{2}((1,1) \bullet (4,0))(1,1) = (2,2)$$

Thus, $(2,2) \in S$ is the closest point to (4,0). Geometrically, this is something you should have been able to derive for a few years now. The points (2,2) and (4,0) are on the perpendicular bisector of y = x (the set S is nothing more than the line y = x making the usual identification of points and vectors)

Example 9.5.4. In Example 9.3.13 we found that $W = span\{u_1, u_2, u_3\} \leq \mathbb{R}^4$ where

$u_1 =$	$\begin{bmatrix} 2\\1\\2\\0 \end{bmatrix}$	$u_2 =$	$\begin{bmatrix} 0\\ -2\\ 1\\ 1 \end{bmatrix}$	$u_3 =$	$\begin{bmatrix} -1\\2\\0\\-1 \end{bmatrix}$

³notice $a^2 < b^2$ need not imply a < b in general. For example, $(5)^2 < (-7)^2$ yet $5 \not< -7$. Generally, $a^2 < b^2$ together with the added condition a, b > 0 implies a < b.

has $Proj_W((0,6,0,6)) = (\frac{26}{33}, \frac{94}{33}, \frac{26}{33}, \frac{-36}{33})$. We can calculate that

$$rref \begin{bmatrix} 2 & 0 & -1 & 0 \\ 1 & -2 & 2 & 6 \\ 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

This means that $(0, 6, 0, 6) \notin W$. However, we learned in Proposition 9.5.1 that $\operatorname{Proj}_W((0, 6, 0, 6))$ is the vector in W which is closest to (0, 6, 0, 6). Notice that we can deduce that the orthogonal basis from Example 9.3.13 unioned with $\operatorname{Orth}_W((0, 6, 0, 6))$ will form an orthogonal basis for \mathbb{R}^4 . To modify it to an orthonormal basis we could simply normalize each vector to length one.

Example 9.5.5. Example 9.3.12 shows that $W = span\{e_1 + e_2, e_3\}$ and x = (1, 2, 3) yields $Proj_W(x) = (\frac{3}{2}, \frac{3}{2}, 3)$. Again we can argue that $x \notin Col[e_1 + e_2|e_3] = W$ but $Proj_W(x)$ is in fact in W. Moreover, $Proj_W(x)$ is the closest vector to x which is in W. The geometric interpretation here is that $Orth_W(x) = (-\frac{1}{2}, \frac{1}{2}, 0)$ is precisely the normal vector to the plane W.

The examples above are somewhat special in that the subspaces considered have only one dimension less than the total vector space. This means that the orthogonal projection of any vector outside the subspace will return the same vector modulo a nonzero constant. In other words, the orthogonal complement is selecting the normal vector to our subspace. In general if we had a subspace which was two or more dimensions smaller than the total vector space then there would be more variety in the output of the orthogonal projection with respect to the subspace. For example, if we consider a plane inside \mathbb{R}^4 then there is more than just one direction which is orthogonal to the plane, the orthogonal projection would itself fill out a plane in \mathbb{R}^4 .

9.5.2 inconsistent equations

We've spent considerable time solving systems of equations which were *consistent*. What if a system of equations Ax = b is inconsistent? What if anything can we say? Let $A \in \mathbb{R}^{m \times n}$ then we found in Proposition 6.7.3 Ax = b is consistent iff $b \in Col(A)$. In other words, the system has a solution iff there is some linear combination of the columns of A such that we obtain b. Here the columns of A and b are both m-dimensional vectors. If rank(A) = dim(Col(A)) = m then the system is consistent no matter which choice for b is made. However, if rank(A) < m then there are some vectors in \mathbb{R}^m which are not in the column space of A and if $b \notin Col(A)$ then there will be no $x \in \mathbb{R}^n$ such that Ax = b. We can picture it as follows: the Col(A) is a subspace of \mathbb{R}^m and $b \in \mathbb{R}^m$ is a vector pointing out of the subspace. The shadow of b onto the subspace Col(A) is given by $Proj_{Col(A)}(b)$.



Notice that $Proj_{Col(A)}(b) \in Col(A)$ thus the system $Ax = Proj_{Col(A)}(b)$ has a solution for any $b \in \mathbb{R}^m$. In fact, we can argue that x which solves $Ax = Proj_{Col(A)}(b)$ is the solution which comes closest to solving Ax = b. Closest in the sense that $||Ax - b||^2$ is minimized. We call such x the least squares solution to Ax = b (which is kind-of funny terminology since x is not actually a solution, perhaps we should really call it the "least squares approximation").

Theorem 9.5.6. Least Squares Solution:

If Ax = b is inconsistent then the solution of $Au = Proj_{col(A)}(b)$ minimizes $||Ax - b||^2$.

Proof: We can break-up the vector b into a vector $Proj_{Col(A)}(b) \in Col(A)$ and $Orth_{col(A)}(b) \in Col(A)^{\perp}$ where

$$b = Proj_{Col(A)}(b) + Orth_{Col(A)}(b).$$

Since Ax = b is inconsistent it follows that $b \notin Col(A)$ thus $Orth_{Col(A)}(b) \neq 0$. Observe that:

$$||Ax - b||^{2} = ||Ax - Proj_{Col(A)}(b) - Orth_{Col(A)}(b)||^{2}$$
$$= ||Ax - Proj_{Col(A)}(b)||^{2} + ||Orth_{Col(A)}(b)||^{2}$$

Therefore, the solution of $Ax = Proj_{Col(A)}(b)$ minimizes $||Ax - b||^2$ since any other vector will make $||Ax - Proj_{Col(A)}(b)||^2 > 0$. \Box

Admittably, there could be more than one solution of $Ax = Proj_{Col(A)}(b)$, however it is usually the case that this system has a unique solution. Especially for experimentally determined data sets.

We already have a technique to calculate projections and of course we can solve systems but it is exceedingly tedious to use the proposition above from scratch. Fortunately there is no need:

Proposition 9.5.7.

If Ax = b is inconsistent then the solution(s) of $Au = Proj_{Col(A)}(b)$ are solutions of the so-called **normal equations** $A^T Au = A^T b$.

Proof: Observe that,

$$Au = Proj_{Col(A)}(b) \Leftrightarrow b - Au = b - Proj_{Col(A)}(b) = Orth_{Col(A)}(b)$$

$$\Leftrightarrow b - Au \in Col(A)^{\perp}$$

$$\Leftrightarrow b - Au \in Null(A^{T})$$

$$\Leftrightarrow A^{T}(b - Au) = 0$$

$$\Leftrightarrow A^{T}Au = A^{T}b,$$

where we used Proposition 9.3.8 in the third step. \Box

The proposition below follows immediately from the preceding proposition.

Proposition 9.5.8.

If $det(A^T A) \neq 0$ then there is a unique solution of $Au = Proj_{Col(A)}(b)$.

Examples are given in the next section. The proposition above is the calculational core of the least squares method.

9.5.3 the least squares problem

In experimental studies we often have some model with coefficients which appear linearly. We perform an experiment, collect data, then our goal is to find coefficients which make the model fit the collected data. Usually the data will be inconsistent with the model, however we'll be able to use the idea of the last section to find the so-called *best-fit* curve. I'll begin with a simple linear model. This linear example contains all the essential features of the least-squares analysis.

linear least squares problem

Problem: find values of c_1, c_2 such that $y = c_1x + c_2$ most closely models a given data set: $\{(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)\}$

Solution: Plug the data into the model and see what equations result:

$$y_1 = c_1 x_1 + c_2, \quad y_2 = c_1 x_2 + c_2, \quad \dots \quad y_k = c_1 x_k + c_2$$

arrange these as a matrix equation,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_k & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \Rightarrow \vec{y} = M\vec{v}$$

where $\vec{y} = (y_1, y_2, \dots, y_k)$ and $v = (c_1, c_2)$ and M is defined in the obvious way. The system $\vec{y} = M\vec{v}$ will be inconsistent due to the fact that error in the data collection will⁵ make the results bounce above and below the true solution. We can solve the normal equations $M^T \vec{y} = M^T M \vec{v}$ to find c_1, c_2 which give the best-fit curve⁶.

Example 9.5.9. Find the best fit line through the points (0,2), (1,1), (2,4), (3,3). Our model is $y = c_1 + c_2 x$. Assemble M and \vec{y} as in the discussion preceding this example:

$$\vec{y} = \begin{bmatrix} 2\\1\\4\\3 \end{bmatrix} \quad M = \begin{bmatrix} 0 & 1\\1 & 1\\2 & 1\\3 & 1 \end{bmatrix} \quad \Rightarrow \quad M^T M = \begin{bmatrix} 0 & 1 & 2 & 3\\1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1\\1 & 1\\2 & 1\\3 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 6\\6 & 4 \end{bmatrix}$$

and we calculate:
$$M^T y = \begin{bmatrix} 0 & 1 & 2 & 3\\1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2\\1\\4\\3 \end{bmatrix} = \begin{bmatrix} 18\\10 \end{bmatrix}$$

The normal equations⁷ are $M^T M \vec{v} = M^T \vec{y}$. Note that $(M^T M)^{-1} = \frac{1}{20} \begin{bmatrix} 4 & -6 \\ -6 & 14 \end{bmatrix}$ thus the solution of the normal equations is simply,

$$\vec{v} = (M^T M)^{-1} M^T \vec{y} = \frac{1}{20} \begin{bmatrix} 4 & -6 \\ -6 & 14 \end{bmatrix} \begin{bmatrix} 18 \\ 10 \end{bmatrix} = \begin{bmatrix} \frac{3}{5} \\ \frac{8}{5} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

⁵almost always

⁶notice that if x_i are not all the same then it is possible to show $det(M^T M) \neq 0$ and then the solution to the normal equations is unique

⁷notice my choice to solve this system of 2 equations and 2 unknowns is just a choice, You can solve it a dozen different ways, you do it the way which works best for you.

Thus, y = 0.6x + 1.6 is the best-fit line. This solution minimizes the vertical distances squared between the data and the model.

It's really nice that the order of the normal equations is only as large as the number of coefficients in the model. If the order depended on the size of the data set this could be much less fun for real-world examples. Let me set-up the linear least squares problem for 3-coefficients and data from \mathbb{R}^3 , the set-up for more coefficients and higher-dimensional data is similar. We already proved this in general in the last section, the proposition simply applies mathematics we already derived. I state it for your convenience.

Proposition 9.5.10.

Given data $\{\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n\} \subset \mathbb{R}^3$, with $\vec{r}_k = [x_k, y_k, z_k]^T$, the best-fit of the linear model $z = c_1 x + c_2 y + c_3$ is obtained by solving the normal equations $M^T M \vec{v} = M^T \vec{z}$ where $\vec{z} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \qquad M = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & 1 \end{bmatrix} \qquad \vec{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}.$

Example 9.5.11. Find the plane which is closest to the points (0, 0, 0), (1, 2, 3), (4, 0, 1), (0, 3, 0), (1, 1, 1). An arbitrary⁸ plane has the form $z = c_1x + c_2y + c_3$. Work on the normal equations,

$$M = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 2 & 1 \\ 4 & 0 & 1 \\ 0 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \vec{z} = \begin{bmatrix} 0 \\ 3 \\ 1 \\ 0 \\ 1 \end{bmatrix} \quad \Rightarrow \quad M^T M = \begin{bmatrix} 0 & 1 & 4 & 0 & 1 \\ 0 & 2 & 0 & 3 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 2 & 1 \\ 4 & 0 & 1 \\ 0 & 3 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 18 & 3 & 6 \\ 3 & 14 & 6 \\ 6 & 6 & 5 \end{bmatrix}$$
$$also, \quad M^T \vec{z} = \begin{bmatrix} 0 & 1 & 4 & 0 & 1 \\ 0 & 2 & 0 & 3 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \\ 5 \end{bmatrix}$$

We solve $M^T M \vec{v} = M^T \vec{z}$ by row operations, after some calculation we find:

$$rref[M^{T}M|M^{T}\vec{z}] = \begin{bmatrix} 1 & 0 & 1 & 89/279 \\ 0 & 1 & 1 & 32/93 \\ 0 & 0 & 1 & 19/93 \end{bmatrix} \xrightarrow{c_{1} = 89/279} c_{2} = 32/93 \\ c_{3} = 19/93$$

Therefore, $z = \frac{89}{293}x + \frac{32}{93}y + \frac{19}{93}$ is the plane which is "closest" to the given points. Technically, I'm not certain that is is the absolute closest. We used the vertical distance squared as a measure of distance from the point. Distance from a point to the plane is measured along the normal direction, so there is no garauntee this is really the absolute "best" fit. For the purposes of this course we will ignore this subtle and annoying point. When I say "best-fit" I mean the least squares fit of the model.

⁸technically, the general form for a plane is ax + by + cz = d, if c = 0 for the best solution then our model misses it. In such a case we could let x or y play the role that z plays in our set-up.
nonlinear least squares

Problem: find values of c_1, c_2 such that $y = c_1 f_1(x)x + c_2 f_2(x) + \cdots + c_n f_n(x)$ most closely models a given data set: $\{(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)\}$. We assume the coefficients c_1, c_2 appear linearly on (possibly nonlinear) functions $f_1, f_2, \dots f_n$.

Solution: Plug the data into the model and see what equations result:

$$y_1 = c_1 f_1(x_1) + c_2 f_2(x_1) + \dots + c_n f_n(x_1),$$

$$y_2 = c_1 f_1(x_2) + c_2 f_2(x_2) + \dots + c_n f_n(x_2),$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$y_k = c_1 f_1(x_k) + c_2 f_2(x_k) + \dots + c_n f_n(x_k)$$

arrange these as a matrix equation,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_n(x_1) \\ f_1(x_1) & f_2(x_1) & \cdots & f_n(x_1) \\ \vdots & \vdots & \vdots & \vdots \\ f_1(x_k) & f_2(x_k) & \cdots & f_n(x_k) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \Rightarrow \vec{y} = M\vec{v}$$

where $\vec{y} = [y_1, y_2, \dots, y_k]^T$, $v = [c_1, c_2, \dots, c_n]^T$ and M is defined in the obvious way. The system $\vec{y} = M\vec{v}$ will be inconsistent due to the fact that error in the data collection will⁹ make the results bounce above and below the true solution. We can solve the normal equations $M^T\vec{y} = M^T M\vec{v}$ to find c_1, c_2, \dots, c_n which give the best-fit curve¹⁰.

Remark 9.5.12.

Nonlinear least squares includes the linear case as a subcase, take $f_1(x) = x$ and $f_2(x) = 1$ and we return to the linear least squares examples. We will use data sets from \mathbb{R}^2 in this subsection. These techniques do extend to data sets with more variables as I demonstrated in the simple case of a plane.

Example 9.5.13. Find the best-fit parabola through the data (0,0), (1,3), (4,4), (3,6), (2,2). Our model has the form $y = c_1x^2 + c_2x + c_3$. Identify that $f_1(x) = x^2, f_2(x) = x$ and $f_3(x) = 1$ thus we should study the normal equations: $M^T M \vec{v} = M^T \vec{y}$ where:

$$M = \begin{bmatrix} f_1(0) & f_2(0) & f_3(0) \\ f_1(1) & f_2(1) & f_3(1) \\ f_1(4) & f_2(4) & f_3(4) \\ f_1(3) & f_2(3) & f_3(3) \\ f_1(2) & f_2(2) & f_3(2) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 16 & 4 & 1 \\ 9 & 3 & 1 \\ 4 & 2 & 1 \end{bmatrix} \quad and \quad \vec{y} = \begin{bmatrix} 0 \\ 3 \\ 4 \\ 6 \\ 2 \end{bmatrix}.$$

Hence, calculate

$$M^{T}M = \begin{bmatrix} 0 & 1 & 16 & 9 & 4 \\ 0 & 1 & 4 & 3 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 16 & 4 & 1 \\ 9 & 3 & 1 \\ 4 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 354 & 100 & 30 \\ 100 & 30 & 10 \\ 30 & 10 & 5 \end{bmatrix}$$

⁹almost always

¹⁰notice that if $f_j(x_i)$ are not all the same then it is possible to show $det(M^T M) \neq 0$ and then the solution to the normal equations is unique

and,

$$M^{T}\vec{y} = \begin{bmatrix} 0 & 1 & 16 & 9 & 4 \\ 0 & 1 & 4 & 3 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \\ 4 \\ 6 \\ 2 \end{bmatrix} = \begin{bmatrix} 129 \\ 41 \\ 15 \end{bmatrix}$$

After a few row operations we can deduce,

$$rref[M^{T}M|M^{T}\vec{y}] = \begin{bmatrix} 1 & 0 & 1 & -5/14 \\ 0 & 1 & 1 & 177/70 \\ 0 & 0 & 1 & 3/35 \end{bmatrix} \xrightarrow{c_{1} = -5/14 \approx -0.357}{c_{2} = 177/70 \approx 2.529} c_{3} = 3/35 = 0.086$$

We find the best-fit parabola is $y = -0.357x^2 + 2.529x + 0.086$



Yes..., but what's this for?

Example 9.5.14. Suppose you land on a mysterious planet. You find that if you throw a ball it's height above the ground y at time t is measured at times t = 0, 1, 2, 3, 4 seconds to be y = 0, 2, 3, 6, 4 meters respective. Assume that Newton's Law of gravity holds and determine the gravitational acceleration from the data. We already did the math in the last example. Newton's law approximated for heights near the surface of the planet simply says y'' = -g which integrates twice to yield $y(t) = -gt^2/2 + v_0t + y_0$ where v_0 is the initial velocity in the vertical direction. We find the best-fit parabola through the data set $\{(0,0), (1,3), (4,4), (3,6), (2,2)\}$ by the math in the last example,

$$y(t) = -0.357t^2 + 2.529 + 0.086$$

we deduce that $g = 2(0.357)m/s^2 = 0.714m/s^2$. Apparently the planet is smaller than Earth's moon (which has $g_{moon} \approx \frac{1}{6}9.8m/s^2 = 1.63m/s^2$.

Remark 9.5.15.

If I know for certain that the ball is at y = 0 at t = 0 would it be equally reasonable to assume y_o in our model? If we do it simplifies the math. The normal equations would only be order 2 in that case.

Example 9.5.16. Find the best-fit parabola that passes through the origin and the points (1,3), (4,4), (3,6), (2,2). To begin we should state our model: since the parabola goes through the origin we know the y-intercept is zero hence $y = c_1x^2 + c_2x$. Identify $f_1(x) = x^2$ and $f_2(x) = x$. As usual set-up the M and \vec{y} ,

$$M = \begin{bmatrix} f_1(1) & f_2(1) \\ f_1(4) & f_2(4) \\ f_1(3) & f_2(3) \\ f_1(2) & f_2(2) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 16 & 4 \\ 9 & 3 \\ 4 & 2 \end{bmatrix} \quad and \quad \vec{y} = \begin{bmatrix} 3 \\ 4 \\ 6 \\ 2 \end{bmatrix}.$$

Calculate,

$$M^{T}M = \begin{bmatrix} 1 & 16 & 9 & 4 \\ 1 & 4 & 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 16 & 4 \\ 9 & 3 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} 354 & 100 \\ 100 & 30 \end{bmatrix} \Rightarrow (M^{T}M)^{-1} = \frac{1}{620} \begin{bmatrix} 30 & -100 \\ -100 & 354 \end{bmatrix}$$

and,

$$M^{T}\vec{y} = \begin{bmatrix} 1 & 16 & 9 & 4 \\ 1 & 4 & 3 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \\ 6 \\ 2 \end{bmatrix} = \begin{bmatrix} 129 \\ 41 \end{bmatrix}$$

We solve $M^T M \vec{v} = M^T \vec{y}$ by multiplying both sides by $(M^T M)^{-1}$ which yields,

$$\vec{v} = (M^T M)^{-1} M^T \vec{y} = \frac{1}{620} \begin{bmatrix} 30 & -100 \\ -100 & 354 \end{bmatrix} \begin{bmatrix} 129 \\ 41 \end{bmatrix} = \begin{bmatrix} -23/62 \\ 807/310 \end{bmatrix} \Rightarrow \begin{array}{c} c_1 = -23/62 \approx -0.371 \\ c_2 = 807/310 \approx 2.603 \end{array}$$

Thus the best-fit parabola through the origin is $y = -0.371x^2 + 2.603x$

Sometimes an application may not allow for direct implementation of the least squares method, however a rewrite of the equations makes the unknown coefficients appear linearly in the model.

Example 9.5.17. Newton's Law of Cooling states that an object changes temperature T at a rate proportional to the difference between T and the room-temperature. Suppose room temperature is known to be 70° then dT/dt = -k(T-70) = -kT + 70k. Calculus reveals solutions have the form $T(t) = c_0 e^{-kt} + 70$. Notice this is very intuitive since $T(t) \rightarrow 70$ for t >> 0. Suppose we measure the temperature at successive times and we wish to find the best model for the temperature at time t. In particular we measure: T(0) = 100, T(1) = 90, T(2) = 85, T(3) = 83, T(4) = 82. One unknown coefficient is k and the other is c_1 . Clearly k does not appear linearly. We can remedy this by working out the model for the natural log of T - 70. Properties of logarithms will give us a model with linearly appearing unknowns:

$$\ln(T(t) - 70) = \ln(c_0 e^{-kt}) = \ln(c_0) + \ln(e^{-kt}) = \ln(c_0) - kt$$

Let $c_1 = \ln(c_0)$, $c_2 = -k$ then identify $f_1(t) = 1$ while $f_2(t) = t$ and $y = \ln(T(t) - 70)$. Our model is $y = c_1 f_1(t) + c_2 f_2(t)$ and the data can be generated from the given data for T(t):

$$t_1 = 0: y_1 = \ln(T(0) - 70) = \ln(100 - 70) = \ln(30)$$

$$t_2 = 1: y_2 = \ln(T(1) - 90) = \ln(90 - 70) = \ln(20)$$

$$t_3 = 2: y_3 = \ln(T(2) - 85) = \ln(85 - 70) = \ln(15)$$

$$t_4 = 3: y_4 = \ln(T(2) - 83) = \ln(83 - 70) = \ln(13)$$

$$t_5 = 4: y_5 = \ln(T(2) - 82) = \ln(82 - 70) = \ln(12)$$

Our data for (t, y) is $(0, \ln 30), (1, \ln 20), (2, \ln 15), (3, \ln 13), (4, \ln 12)$. We should solve normal equations $M^T M \vec{v} = M^T \vec{y}$ where

$$M = \begin{bmatrix} f_1(0) & f_2(0) \\ f_1(1) & f_2(1) \\ f_1(2) & f_2(2) \\ f_1(3) & f_2(3) \\ f_1(4) & f_2(4) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \quad and \quad \vec{y} = \begin{bmatrix} \ln 30 \\ \ln 20 \\ \ln 15 \\ \ln 13 \\ \ln 12 \end{bmatrix}$$

We can calculate $M^T M = \begin{bmatrix} 5 & 10 \\ 10 & 30 \end{bmatrix}$ and $M^T \vec{y} \approx \begin{bmatrix} 14.15 \\ 26.05 \end{bmatrix}$. Solve $M^T M \vec{v} = M^T \vec{y}$ by multiplication by inverse of $M^T M$:

$$\vec{y} = (M^T M)^{-1} M^T \vec{y} = \begin{bmatrix} 3.284 \\ -0.2263 \end{bmatrix} \Rightarrow \begin{array}{c} c_1 \approx 3.284 \\ c_2 \approx -0.2263 \end{array}$$

Therefore, $y(t) = \ln(T(t) - 70) = 3.284 - 0.2263$ we identify that k = 0.2263 and $\ln(c_0) = 3.284$ which yields $c_0 = e^{3.284} = 26.68$. We find the best-fit temperature function is

$$T(t) = 26.68e^{-0.2263t} + 70.$$

Now we could give good estimates for the temperature T(t) for other times. If Newton's Law of cooling is an accurate model and our data was collected carefully then we ought to be able to make accurate predictions with our model.

Remark 9.5.18.

The accurate analysis of data is more involved than my silly examples reveal here. Each experimental fact comes with an error which must be accounted for. A real experimentalist never gives just a number as the answer. Rather, one must give a number and an uncertainty or error. There are ways of accounting for the error of various data. Our approach here takes all data as equally valid. There are weighted best-fits which minimize a weighted least squares. Technically, this takes us into the realm of math of inner-product spaces. Finite dimensional inner-product spaces also allows for least-norm analysis. The same philosophy guides the analysis: the square of the norm measures the sum of the squares of the errors in the data. The collected data usually does not precisely fit the model, thus the equations are inconsistent. However, we project the data onto the plane representative of model solutions and this gives us the best model for our data. Generally we would like to minimize χ^2 , this is the notation for the sum of the squares of the error often used in applications. In statistics finding the best-fit line is called doing "linear regression".

9.6 inner products

The definition of an inner product is based on the idea of the dot product. Proposition 9.1.4 summarized the most important properties. These properties form the definition for an inner product. If you examine proofs in § 9.1 you'll notice most of what I argued was based on using these 4 simple facts for the dot-product¹¹.

Definition 9.6.1.

Let V be a vector space over \mathbb{R} . If there is a function $\langle , \rangle : V \times V \to \mathbb{R}$ such that for all $x, y, z \in V$ and $c \in \mathbb{R}$, 1. $\langle x, y \rangle = \langle y, x \rangle$ (symmetric), 2. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$, 3. $\langle cx, y \rangle = c \langle x, y \rangle$, 4. $\langle x, x \rangle \ge 0$ and $\langle x, x \rangle = 0$ iff x = 0, then we say \langle , \rangle is an inner product on V. In this case we say V with $\langle \rangle$ is an inner product space. Items (1.), (2.) and (3.) together allow us to call \langle , \rangle a real-valued superscript bilinger form on V.

symmetric-bilinear-form on V. We may find it useful to use the notation $g(x, y) = \langle x, y \rangle$ for some later arguments, one should keep in mind the notation \langle , \rangle is not the only choice.

Technically, items (2.) and (3.) give us "linearity in the first slot". To obtain bilinearity we need to have linearity in the second slot as well. This means $\langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle$ and $\langle x, cy \rangle = c \langle x, y \rangle$ for all $x, y, z \in V$ and $c \in \mathbb{R}$. Fortunately, the symmetry property will transfer the linearity to the second slot. I leave that as an exercise for the reader.

Example 9.6.2. Obviously \mathbb{R}^n together with the dot-product forms an inner product space. Moreover, the dot-product is an inner product.

Once we have an inner product for a vector space then we also have natural definitions for the length of a vector and the distance between two points.

Definition 9.6.3.

Let V be an inner product vector space with inner product \langle , \rangle . The **norm** or **length** of a vector is defined by $||x|| = \sqrt{\langle x, x \rangle}$ for each $x \in V$. Likewise the **distance** between $a, b \in V$ is defined by $d(a, b) = \sqrt{\langle b - a, b - a \rangle} = ||b - a||$ for all $a, b \in V$. We say these are the length and distance functions induced by \langle , \rangle . Likewise the **angle** between two nonzero vectors is defined implicitly by $\langle v, w \rangle = ||v||||w|| \cos(\theta)$.

As before the definition above is only logical if certain properties hold for the inner product, norm and distance function. Happily we find all the same general properties for the inner product and its induced norm and distance function.

¹¹ WARNING: the next couple pages is dense. It's a reiteration of the main theoretical accomplishments of this chapter in the context of inner product spaces. If you need to see examples first then skip ahead as needed.

Proposition 9.6.4.

```
If V is an inner product space with induced norm ||\cdot|| and x, y \in V then |\langle x, y \rangle| \leq ||x|| ||y||.
```

Proof: since $||x|| = \sqrt{\langle x, x \rangle}$ the proof we gave for the case of the dot-product equally well applies here. You'll notice in retrospect I only used those 4 properties which we take as the defining axioms for the inner product. \Box

In fact, all the propositions from §9.1 apply equally well to an arbitrary finite-dimensional inner product space. The proof of the proposition below is similar to those I gave in §9.1

Proposition 9.6.5. Properties for induced norm and distance function on an inner product space.

If V is an inner product space with inner product \langle , \rangle and norm $||x|| = \sqrt{x, x}$ and distance function d(x, y) = ||y - x|| then for all $x, y, z \in V$ and $c \in \mathbb{R}$ $(i.) ||x|| \ge 0 \qquad (v.) \ d(x, y) \ge 0$

 $\begin{array}{ll} (i.) \ ||x|| \geq 0 & (v.) \ d(x,y) \geq 0 \\ (ii.) \ ||x|| = 0 \Leftrightarrow x = 0 & (vi.) \ d(x,y) = 0 \Leftrightarrow x = y \\ (iii.) \ ||cx|| = |c|||x|| & (vii.) \ d(x,y) = d(y,x) \\ (iv.) \ ||x+y|| \leq ||x|| + ||y|| & (viii.) \ d(x,z) \leq d(x,y) + d(y,z) \end{array}$

An norm is simply an operation which satisfies (i.) - (iv.). If we are given a vector space with a norm then that is called a normed linear space. If in addition all Cauchy sequences converge in the space it is said to be a complete normed linear space. A **Banach Space** is defined to be a complete normed linear space. A distance function is simply an operation which satisfies (v.) - (viii.). A set with a distance function is called a **metric space**. I'll let you ponder all these things in some other course, I mention them here merely for breadth. These topics are more interesting infinite-dimensional case.

What is truly interesting is that the orthogonal complement theorems and closest vector theory transfer over to the case of an inner product space.

Definition 9.6.6.

Let V be an inner product space with inner product \langle , \rangle . Let $x, y \in V$ then we say x is **orthogonal** to y iff $\langle x, y \rangle = 0$. A set S is said to be orthogonal iff every pair of vectors in S is orthogonal. If $W \leq V$ then the **orthogonal complement** of W is defined to be $W^{\perp} = \{v \in V \mid v \cdot w = 0 \ \forall w \in W\}.$

Proposition 9.6.7. Orthogonality results for inner product space.

If V is an inner product space with inner product $\langle \ , \ \rangle$ and norm $||x|| = \sqrt{x, x}$ then for all $x, y, z \in V$ and $W \leq V$, (i.) $\langle x, y \rangle = 0 \Rightarrow ||x + y||^2 = ||x||^2 + ||y||^2$ (ii.) if $S \subset V$ is orthonormal $\Rightarrow S$ is linearly independent (iii.) $S \subset V \Rightarrow S^{\perp} \leq V$ (iv.) $W^{\perp} \cap W = \{0\}$ (v.) $V = W \oplus W^{\perp}$

Definition 9.6.8.

Let V be an inner product space with inner product \langle , \rangle . A basis of \langle , \rangle -orthogonal vectors is an **orthogonal basis**. Likewise, if every vector in an orthogonal basis has length one then we call it an **orthonormal basis**.

Every finite dimensional inner product space permits a choice of an orthonormal basis. Examine my proof in the case of the dot-product. You'll find I made all arguments on the basis of the axioms for an inner-product. The Gram-Schmidt process works equally well for inner product spaces, we just need to exchange dot-products for inner-products as appropriate.

Proposition 9.6.9. Orthonormal coordinates and projection results.

If V is an inner product space with inner product \langle , \rangle and $\beta = \{v_1, v_2, \ldots, v_k\}$ is a orthonormal basis for a subspace W then

(i.) $w = \langle w, v_1 \rangle v_1 + \langle w, v_2 \rangle v_2 + \dots + \langle w, v_k \rangle v_k$ for each $w \in W$, (ii.) $Proj_W(x) \equiv \langle x, v_1 \rangle v_1 + \langle x, v_2 \rangle v_2 + \dots + \langle x, v_k \rangle v_k \in W$ for each $x \in V$, (iii.) $Orth_W(x) \equiv x - Proj_W(x) \in W^{\perp}$ for each $x \in V$, (iv.) $x = Proj_W(x) + Orth_W(x)$ and $\langle Proj_W(x), Orth_W(x) \rangle = 0$ for each $x \in V$, (v.) $||x - Proj_W(x)|| \langle ||x - y||$ for all $y \notin W$.

Notice that we can use the Gram-Schmidt idea to implement the least squares analysis in the context of an inner-product space. However, we cannot multiply abstract vectors by matrices so the short-cut normal equations may not make sense in this context. We have to implement the closest vector idea without the help of those normal equations. I'll demonstrate this idea in the Fourier analysis section.

9.6.1 examples of inner-products

The dot-product is just one of many inner products. We examine an assortment of other innerproducts for various finite dimensional vector spaces.

Example 9.6.10. Let $V = \mathbb{R}^2$ and define $\langle v, w \rangle = v_1 w_1 + 3v_2 w_2$ for all $v = (v_1, v_2)$, $w = (w_1, w_2) \in V$. Let $u, v, w \in V$ and $c \in \mathbb{R}$,

1. symmetric property,

$$\langle v, w \rangle = v_1 w_1 + 3 v_2 w_2 = w_1 v_1 + 3 w_2 v_2 = \langle w, v \rangle$$

2. additive property:

3. homogeneous property:

4. positive definite property:

$$\langle v,v \rangle = v_1^2 + 3v_2^2 \ge 0 \text{ and } \langle v,v \rangle = 0 \iff v = 0.$$

Notice $e_1 = (1,0)$ is an orthonormalized vector with respect to \langle , \rangle but $e_2 = (0,1)$ not unit-length. Instead, $\langle e_2, e_2 \rangle = 3$ thus $||e_2|| = \sqrt{3}$ so the unit-vector in the e_2 -direction is $u = \frac{1}{\sqrt{3}}(0,1)$ and with respect to \langle , \rangle we have an orthonormal basis $\{e_1, u\}$.

The inner-product above might be used in an application where the second variable carries more weight. For example, the coordinates could represent inventory of items in some shop. The different coefficients of a non-standard inner product could reflect the prices associated with each unit-item.

Example 9.6.11. Let $V = \mathbb{R}^{m \times n}$ we define the Frobenious inner-product as follows:

$$\langle A, B \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} B_{ij}.$$

It is clear that $\langle A, A \rangle \ge 0$ since it is the sum of squares and it is also clear that $\langle A, A \rangle = 0$ iff A = 0. Symmetry follows from the calculation

$$\langle A, B \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} B_{ij} = \sum_{i=1}^{m} \sum_{j=1}^{n} B_{ij} A_{ij} = \langle B, A \rangle$$

where we can commute B_{ij} and A_{ij} for each pair i, j since the components are just real numbers. Linearity and homogeneity follow from:

$$\langle \lambda A + B, C \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} (\lambda A + B)_{ij} C_{ij} = \sum_{i=1}^{m} \sum_{j=1}^{n} (\lambda A_{ij} + B_{ij}) C_{ij}$$
$$= \lambda \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} C_{ij} + \sum_{i=1}^{m} \sum_{j=1}^{n} B_{ij} C_{ij} = \lambda \langle A, C \rangle + \langle B, C \rangle$$

Therefore. the Frobenius inner-product is in fact an inner product. The Frobenious norm of a matrix is induced as usual:

$$||A|| = \sqrt{\langle A, A \rangle}$$

as a consequence of the theory in this chapter we already know a few interesting properties form the matrix-norm, in particular $||\langle A, B \rangle|| \leq ||A||||B||$. The particular case of square matrices allows further comments. If $A, B \in \mathbb{R}^{n \times n}$ then notice

$$\langle A, B \rangle = \sum_{i,j} A_{ij} B_{ij} = \sum_{i} \sum_{j} A_{ij} (B^T)_{ji} = trace(AB^T) \quad \Rightarrow \quad ||A|| = trace(AA^T)$$

We find an interesting identity for any square matrix

$$|trace(AB^T)| \le \sqrt{trace(AA^T)trace(BB^T)}.$$

The work of Frobenious was vast. As you take various courses you'll come across his work. His linear algebra text was one of the first to treat problems in *n*-dimensions across most topics. In the theory of partial differential equations he found an existence theorem which still is of great utility in the study of submanifolds which foliate a space. The method of Frobenius in ordinary differential equations is also extremely important and lies at the genesis of many special functions. That said, I don't think Frobenius is well-known outside mathematical circles. Explaining his life's work would make an interesting topic for a math history project.

Example 9.6.12. Let C[a, b] denote the set of functions which are continuous on [a, b]. This is an infinite dimensional vector space. We can define an inner-product via the definite integral of the product of two functions: let $f, g \in C[a, b]$ define

$$\langle f,g\rangle = \int_a^b f(x)g(x)dx.$$

We can prove this is an inner-product. I'll just show additivity,

$$\begin{split} \langle f+g,h\rangle &= \int_{a}^{b} (f(x)+g(x))(x)h(x)dx\\ &= \int_{a}^{b} f(x)h(x)dx + \int_{a}^{b} g(x)h(x)dx = \langle f,h\rangle + \langle g,h\rangle \end{split}$$

I leave the proof of the other properties to the reader.

Example 9.6.13. Consider the inner-product $\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)dx$ for $f, g \in C[-1, 1]$. Let's calculate the length squared of the standard basis:

$$\langle 1,1\rangle = \int_{-1}^{1} 1 \cdot 1 dx = 2, \qquad \langle x,x\rangle = \int_{-1}^{1} x^2 dx = \frac{x^3}{3} \Big|_{-1}^{1} = \frac{2}{3} \langle x^2,x^2\rangle = \int_{-1}^{1} x^4 dx = \frac{x^5}{5} \Big|_{-1}^{1} = \frac{2}{5}$$

Notice that the standard basis of P_2 are not all \langle , \rangle -orthogonal:

$$\langle 1, x \rangle = \int_{-1}^{1} x dx = 0$$
 $\langle 1, x^2 \rangle = \langle x, x \rangle = \int_{-1}^{1} x^2 dx = \frac{2}{3}$ $\langle x, x^2 \rangle = \int_{-1}^{1} x^3 dx = 0$

We can use the Gram-Schmidt process on $\{1, x, x^2\}$ to find an orthonormal basis for P_2 on [-1, 1]. Let, $u_1(x) = 1$ and

$$u_2(x) = x - \frac{\langle x, 1 \rangle}{\langle 1, 1 \rangle} = x$$
$$u_3(x) = x^2 - \frac{\langle x^2, x \rangle}{\langle x, x \rangle} x - \frac{\langle x^2, 1 \rangle}{\langle 1, 1 \rangle} = x^2 - \frac{1}{3}$$

We have an orthogonal set of functions $\{u_1, u_2, u_3\}$ we already calculated the length of u_1 and u_2 so we can immediately normalize those by dividing by their lengths; $v_1(x) = \frac{1}{\sqrt{2}}$ and $v_2(x) = \sqrt{\frac{3}{2}x}$. We need to calculate the length of u_3 so we can normalize it as well:

$$\langle u_3, u_3 \rangle = \int_{-1}^{1} \left(x^2 - \frac{1}{3} \right)^2 dx = \int_{-1}^{1} \left(x^4 - \frac{2}{3}x^2 + \frac{1}{9} \right) dx = \frac{2}{5} - \frac{4}{9} + \frac{2}{9} = \frac{8}{45}$$

Thus $v_3(x) = \sqrt{\frac{8}{45}} \left(x^2 - \frac{1}{3}\right)$ has length one. Therefore, $\left\{\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}x, \sqrt{\frac{8}{45}} \left(x^2 - \frac{1}{3}\right)\right\}$ is an orthonormal basis for P_2 restricted to [-1, 1]. Other intervals would not have the same basis. This construction depends both on our choice of inner-product and the interval considered. Incidentally, these are the first three **Legendre Polynomials**. These arise naturally as solutions to certain differential equations. The theory of **orthogonal polynomials** is full of such calculations. Orthogonal polynomials are quite useful as approximating functions. If we offered a second course in differential equations we could see the full function of such objects.

Example 9.6.14. Clearly $f(x) = e^x \notin P_2$. What is the least-squares approximation of f? Use the projection onto P_2 : Proj $_{P_2}(f) = \langle f, v_1 \rangle v_1 + \langle f, v_2 \rangle v_2 + \langle f, v_3 \rangle v_3$. We calculate,

$$\langle f, v_1 \rangle = \int_{-1}^1 \frac{1}{\sqrt{2}} e^x dx = \frac{1}{\sqrt{2}} (e^1 - e^{-1}) \approx 1.661$$
$$\langle f, v_2 \rangle = \int_{-1}^1 \sqrt{\frac{3}{2}} x e^x dx = \sqrt{\frac{3}{2}} (x e^x - e^x) |_{-1}^1 = \sqrt{\frac{3}{2}} [-(-e^{-1} - e^{-1})] = \sqrt{6}e^{-1} \approx 0.901$$
$$\langle f, v_3 \rangle = \int_{-1}^1 \sqrt{\frac{8}{45}} (x^2 - \frac{1}{3}) e^x dx = \frac{2e}{3} - \frac{14e^{-1}}{3} \approx 0.0402$$

Thus,

$$Proj_{P_2}(f)(x) = 1.661v_1(x) + 0.901v_2(x) + 0.0402v_3(x)$$
$$= 1.03 + 1.103x + 0.017x^2$$

This is closest a quadratic can come to approximating the exponential function on the interval [-1, 1]. What's the giant theoretical leap we made in this example? We wouldn't face the same leap if we tried to approximate $f(x) = x^4$ with P_2 . What's the difference? Where does e^x live?

Example 9.6.15. Consider $C[-\pi,\pi]$ with inner product $\langle f,g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$. The set of sine and cosine functions $\{1, \cos(x), \sin(x), \cos(2x), \sin(2x), \ldots, \cos(kx), \sin(kx)\}$ is an orthogonal set of functions.

$$\langle \cos(mx), \cos(nx) \rangle = \int_{-\pi}^{\pi} \cos(mx) \cos(nx) dx = \pi \delta_{mn}$$
$$\langle \sin(mx), \sin(nx) \rangle = \int_{-\pi}^{\pi} \sin(mx) \sin(nx) dx = \pi \delta_{mn}$$
$$\langle \sin(mx), \cos(nx) \rangle = \int_{-\pi}^{\pi} \sin(mx) \cos(nx) dx = 0$$

Thus we find the following is a set of <u>orthonormal</u> functions

$$\beta_{trig} = \{ \frac{1}{\sqrt{2\pi}}, \ \frac{1}{\sqrt{\pi}} \cos(x), \ \frac{1}{\sqrt{\pi}} \sin(x), \ \frac{1}{\sqrt{\pi}} \cos(2x), \ \frac{1}{\sqrt{\pi}} \sin(2x), \ \dots, \ \frac{1}{\sqrt{\pi}} \cos(kx), \ \frac{1}{\sqrt{\pi}} \sin(kx) \}$$

9.6.2 Fourier analysis

The idea of Fourier analysis is based on the least-squares approximation and the last example of the preceding section. We wish to represent a function with a sum of sines and cosines, this is called a **Fourier sum**. Much like a power series, the more terms we use to approximate the function the closer the approximating sum of functions gets to the real function. In the limit the approximation can become exact, the Fourier sum goes to a Fourier series. I do not wish to confront the analytical issues pertaining to the convergence of Fourier series. As a practical matter, it's difficult to calculate infinitely many terms so in practice we just keep the first say 10 or 20 terms and it will come very close to the real function. The advantage of a Fourier sum over a polynomial is that sums of trigonometric functions have natural periodicities. If we approximate the function over the interval $[-\pi, \pi]$ we will also find our approximation repeats itself outside the interval. This is desireable if one wishes to model a wave-form of some sort. Enough talk. Time for an example. (there also an example in your text on pages 540-542 of Spence, Insel and Friedberg)

Example 9.6.16. Suppose $f(t) = \begin{cases} 1 & 0 < t < \pi \\ -1 & -\pi < t < 0 \end{cases}$ and $f(t + 2n\pi) = f(t)$ for all $n \in \mathbb{Z}$. This is called a square wave for the obvious reason (draw its graph). Find the first few terms in

This is called a square wave for the obvious reason (draw its graph). Find the first few terms in a Fourier sum to represent the function. We'll want to use the projection: it's convenient to bring the normalizing constants out so we can focus on the integrals without too much clutter.¹²

$$Proj_W(f)(t) = \frac{1}{2\pi} \langle f, 1 \rangle + \frac{1}{\pi} \langle f, \cos t \rangle \cos t + \frac{1}{\pi} \langle f, \sin t \rangle \sin t + \frac{1}{\pi} \langle f, \cos 2t \rangle \cos 2t + \frac{1}{\pi} \langle f, \sin 2t \rangle \sin 2t + \cdots$$

Where $W = span(\beta_{trig})$. The square wave is constant on $(0, \pi]$ and $[-\pi, 0)$ and the value at zero is not defined (you can give it a particular value but that will not change the integrals that calculate the Fourier coefficients). Calculate,

$$\langle f,1\rangle = \int_{-\pi}^{\pi} f(t)dt = 0$$

$$\langle f, \cos t \rangle = \int_{-\pi}^{\pi} \cos(t) f(t) dt = 0$$

Notice that f(t) and $\cos(t)f(t)$ are odd functions so we can conclude the integrals above are zero without further calculation. On the other hand, $\sin(-t)f(-t) = (-\sin t)(-f(t)) = \sin tf(t)$ thus $\sin(t)f(t)$ is an even function, thus:

$$\langle f, \sin t \rangle = \int_{-\pi}^{\pi} \sin(t) f(t) dt = 2 \int_{0}^{\pi} \sin(t) f(t) dt = 2 \int_{0}^{\pi} \sin(t) dt = 4$$

Notice that $f(t)\cos(kt)$ is odd for all $k \in \mathbb{N}$ thus $\langle f, \cos(kt) \rangle = 0$. Whereas, $f(t)\sin(kt)$ is even for all $k \in \mathbb{N}$ thus

$$\begin{split} \langle f, \sin kt \rangle &= \int_{-\pi}^{\pi} \sin(kt) f(t) dt = 2 \int_{0}^{\pi} \sin(kt) f(t) dt \\ &= 2 \int_{0}^{\pi} \sin(kt) dt = \frac{2}{k} \left[1 - \cos(k\pi) \right] = \begin{cases} 0, & k \text{ even} \\ \frac{4}{k}, & k \text{ odd} \end{cases} \end{split}$$

Putting it all together we find (the \sim indicates the functions are nearly the same except for a finite subset of points),

$$f(t) \sim \frac{4}{\pi} \left(\sin t + \frac{1}{3} \sin 3t + \frac{1}{5} \sin 5t + \cdots \right) = \sum_{n=1}^{\infty} \frac{4}{(2n-1)\pi} \sin((2n-1)t)$$

 $^{^{12}}$ In fact, various texts put these little normalization factors in different places so when you look up results on Fourier series beware conventional discrepancies



I have graphed the Fourier sums up the sum with 11 terms.

Remark 9.6.17.

The treatment of Fourier sums and series is by no means complete in these notes. There is much more to say and do. Our goal here is simply to connect Fourier analysis with the more general story of orthogonality. In the math 334 course we use Fourier series to construct solutions to partial differential equations. Those calculations are foundational to describe interesting physical examples such as the electric and magnetic fields in a waveguide, the vibrations of a drum, the flow of heat through some solid, even the vibrations of a string instrument.

9.7 orthogonal matrices and the QR factorization

This section could be covered earlier. Here we discover a particular factorization which is possible for an orthogonal matrix A ($A^T A = I$). Some semesters we do not require this material.

Suppose we have an orthogonal basis $\beta = \{v_1, v_2, \dots, v_n\}$ for \mathbb{R}^n . Let's investigate the properties of the matrix of this basis. Note that $||v_j|| \neq 0$ for each j since β is linearly independent set of vectors. Moreover, if we denote $||v_j|| = l_j$ then we can compactly summarize orthogonality of β with the following relation:

$$v_j \bullet v_k = l_j^2 \delta_{jk}.$$

As a matrix equation we recognize that $[v_j]^T v_k$ is also the jk - th component of the product of $[\beta]^T$ and $[\beta]$. Let me expand on this in matrix notation:

$$[\beta]^{T}[\beta] = \begin{bmatrix} v_{1}^{T} \\ v_{2}^{T} \\ \vdots \\ v_{n}^{T} \end{bmatrix} [v_{1}|v_{2}|\cdots|v_{n}] = \begin{bmatrix} v_{1}^{T}v_{1} & v_{1}^{T}v_{2} & \cdots & v_{1}^{T}v_{n} \\ v_{2}^{T}v_{1} & v_{2}^{T}v_{2} & \cdots & v_{2}^{T}v_{n} \\ \vdots & \vdots & \cdots & \vdots \\ v_{n}^{T}v_{1} & v_{n}^{T}v_{2} & \cdots & v_{n}^{T}v_{n} \end{bmatrix} = \begin{bmatrix} l_{1}^{2} & 0 & \cdots & 0 \\ 0 & l_{2}^{2} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & l_{n}^{2} \end{bmatrix}$$

This means that $[\beta]^T$ is almost the inverse of $[\beta]$. Observe if we had $l_j = 1$ for j = 1, 2, ..., n then $[\beta]^T = [\beta]^{-1}$. In other words, if we use an orthonormal basis then the inverse of the basis matrix is obtained by transposition. In fact, matrices with this property have a name:

Definition 9.7.1.

Let $A \in \mathbb{R}^{n \times n}$ then we say that A is an **orthogonal matrix** iff $A^T A = I$. The set of all orthogonal $n \times n$ matrices is denoted O(n).

The discussion preceding the definition provides a proof for the following proposition:

Proposition 9.7.2. matrix of an orthonormal basis is an orthogonal matrix

If β is an orthonormal basis then $[\beta]^T[\beta] = I$ or equivalently $[\beta]^T = [\beta]^{-1}$.

So far we have considered only bases for all of \mathbb{R}^n but we can also find similar results for a subspace $W \leq \mathbb{R}^n$. Suppose dim(W) < n. If β is an orthonormal basis for W then it is still true that $[\beta]^T[\beta] = I_{dim(W)}$ however since $[\beta]$ is not a square matrix it does not make sense to say that $[\beta]^T = [\beta]^{-1}$. The QR-factorization of a matrix is tied to this discussion.

Proposition 9.7.3. QR factorization of a full-rank matrix

If $A \in \mathbb{R}^{m \times n}$ is a matrix with linearly independent columns then there exists a matrix $Q \in \mathbb{R}^{m \times n}$ whose columns form an orthonormal basis for Col(A) and square matrix $R \in \mathbb{R}^{n \times n}$ which is upper triangular and has $R_{ii} > 0$ for i = 1, 2, ..., n.

Proof: begin by performing the Gram-Schmidt procedure on the columns of A. Next, normalize that orthogonal basis to obtain an orthonormal basis $\beta = \{u_1, u_2, \ldots, u_n\}$ for Col(A). Note that since each column in A is in Col(A) it follows that some linear combination of the vectors in β will produce that column;

$$col_j(A) = R_{1j}u_1 + R_{2j}u_2 + \dots + R_{nj}u_n = [u_1|u_2|\cdots|u_n][R_{1j}, R_{2j}, \cdots, R_{nj}]^T$$

for some constants $R_{1j}, R_{2j}, \dots, R_{nj} \in \mathbb{R}$. Let R be the matrix formed from the coefficients of the linear combinations that link columns of A and the orthonormal basis; in particular define R such that $col_j(R) = (R_{1j}, R_{2j}, \dots, R_{nj})$. It follows that if we denote $[\beta] = Q$ we have for each $j = 1, 2, \dots, n$ the relation

$$col_j(A) = Qcol_j(R)$$

Hence,

$$A = [col_1(A)|col_2(A)|\cdots|col_n(A)] = [Qcol_1(R)|Qcol_2(R)|\cdots|Qcol_n(R)]$$

and we find by the concatenation proposition

$$A = Q[col_1(R)|col_2(R)|\cdots|col_n(R)] = QR$$

where $R \in \mathbb{R}^{n \times n}$ as we wished. It remains to show that R is upper triangular with positive diagonal entries. Recall how Gram-Schmidt is accomplished (I'll do normalization along side the orthogonalization for the purposes of this argument). We began by defining $u_1 = \frac{1}{||col_1(A)||} col_1(A)$ hence $col_1(A) = ||col_1(A)||u_1$ and we identify that $col_1(R) = (||col_1(A)||, 0, \ldots, 0)$. The next step in the algorithm is to define u_2 by calculating v_2 (since we normalized $u_1 \cdot u_1 = 1$)

$$v_2 = col_2(A) - (col_2(A) \bullet u_1)u_1$$

and normalizing (I define l_2 in the last equality below)

$$u_2 = \frac{1}{||col_2(A) - (col_2(A) \bullet u_1)u_1||} v_2 = \frac{1}{l_2} v_2$$

In other words, $l_2u_2 = v_2 = col_2(A) - (col_2(A) \cdot u_1)u_1$ hence

$$col_2(A) = l_2u_2 - (col_2(A) \bullet u_1)u_1$$

From which we can read the second column of R as

$$col_2(R) = (-(col_2(A) \bullet u_1), l_2, 0, \dots, 0).$$

Continuing in this fashion, if we define l_j to be the length of the orthogonalization of $col_j(A)$ with respect to the preceding $\{u_1, u_2, \ldots, u_{j-1}\}$ orthonormal vectors then a calculation similar to the one just performed will reveal that

$$col_j(R) = (\star, \dots, \star, l_j, 0, \dots, 0)$$

and \star are possibly nonzero components in rows 1 through j-1 of the column vector and l_j is the j-th component which is necessarily positive since it is the length of some nonzero vector. Put all of this together and we find that R is upper triangular with positive diagonal entries¹³. \Box

Very well, we now know that a QR-factorization exists for a matrix with LI columns. This leaves us with two natural questions:

- 1. how do we calculate the factorization of a given matrix A?
- 2. what is the use of the QR factorization ?

We will answer (1.) with an example or two and I will merely scratch the surface for question (2.). If you took a serious numerical linear algebra course then it is likely you would delve deeper.

Example 9.7.4. If $Q^TQ = I$ then A = QR iff $R = Q^TA$. Suppose A is given below and form Q as the orthonormalization of the columns in A: In particular, we use Example 9.2.25 to form Q below.

$$A = \begin{bmatrix} 1 & 1 & 0 & 3\\ 0 & 1 & 0 & 2\\ 1 & 1 & 2 & 0\\ 1 & 1 & 3 & 3 \end{bmatrix} \qquad \& \qquad Q = \begin{bmatrix} 1/\sqrt{3} & 0 & -5/\sqrt{42} & 1/\sqrt{14}\\ 0 & 1 & 0 & 0\\ 1/\sqrt{3} & 0 & 1/\sqrt{42} & -3/\sqrt{14}\\ 1/\sqrt{3} & 0 & 4/\sqrt{42} & 2/\sqrt{14} \end{bmatrix}$$

Finally, multiply Q^T on A to find R:

$$R = Q^{T}A = \begin{bmatrix} 1/\sqrt{3} & 0 & 1/\sqrt{3} & 1/\sqrt{3} \\ 0 & 1 & 0 & 0 \\ -5/\sqrt{42} & 0 & 1/\sqrt{42} & 4/\sqrt{14} \\ 1/\sqrt{14} & 0 & -3/\sqrt{42} & 2/\sqrt{14} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 1 & 1 & 2 & 0 \\ 1 & 1 & 3 & 3 \end{bmatrix}$$
$$= \begin{bmatrix} 3/\sqrt{3} & 3/\sqrt{3} & 5/\sqrt{3} & 6/\sqrt{3} \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 14/\sqrt{42} & -3/\sqrt{42} \\ 0 & 0 & 0 & 9/\sqrt{14} \end{bmatrix}.$$

Note this is upper triangular as claimed.

Finally, returning to (2.). One nice use of the QR-factorization is to simplify calculation of the normal equations. We sought to solve $A^T A u = A^T b$. Suppose that A = QR to obtain:

$$(QR)^T (QR)u = (QR)^T b \Rightarrow R^T Q^T QR u = R^T Q^T b \Rightarrow Ru = Q^T b$$

This problem is easily solved by back-substitution since R is upper-triangular.

 $^{^{13}\}mathrm{see}$ Lay pg. 405-406 if you don't like my proof

Chapter 10

complex vectorspaces

In this brief chapter we study how our work over real vector spaces naturally extends to a vector space over \mathbb{C} . It is interesting to note the construction of the complexification of V as a particular structure on $V \times V$ is the same in essence as Gauss' construction of the complex numbers from \mathbb{R}^2 . Ideally this chapter would contain further discussion of complex linear algebra including the theory of hermitian matrices and normal operators, the spectral theorem etc... but, time is short this semester. If you wish to read further I recommend Insel Spence and Friedberg, however, there are dozens of great texts to read on this topic.

10.0.1 concerning matrices and vectors with complex entries

To begin, we denote the complex numbers by \mathbb{C} . As a two-dimensional real vector space we can decompose the complex numbers into the direct sum of the real and pure-imaginary numbers; $\mathbb{C} = \mathbb{R} \oplus i\mathbb{R}$. In other words, any complex number $z \in \mathbb{R}$ can be written as z = a + ib where $a, b \in \mathbb{R}$. It is convenient to define

If
$$\lambda = \alpha + i\beta \in \mathbb{C}$$
 for $\alpha, \beta \in \mathbb{R}$ then $Re(\lambda) = \alpha$, $Im(\lambda) = \beta$

The projections onto the real or imaginary part of a complex number are actually linear transformations from \mathbb{C} to \mathbb{R} ; $Re : \mathbb{C} \to \mathbb{R}$ and $Im : \mathbb{C} \to \mathbb{R}$. Next, complex vectors are simply *n*-tuples of complex numbers:

$$\mathbb{C}^n = \{(z_1, z_2, \dots, z_n) \mid z_j \in \mathbb{C} \}.$$

Definitions of scalar multiplication and vector addition follow the obvious rules: if $z, w \in \mathbb{C}^n$ and $c \in \mathbb{C}$ then

$$(z+w)_j = z_j + w_j \qquad (cz)_j = cz_j$$

for each j = 1, 2, ..., n. The complex *n*-space is again naturally decomposed into the direct sum of two *n*-dimensional real spaces; $\mathbb{C}^n = \mathbb{R}^n \oplus i\mathbb{R}^n$. In particular, any complex *n*-vector can be written uniquely as the sum of real vectors are known as the real and imaginary vector components:

If
$$v = a + ib \in \mathbb{C}^n$$
 for $a, b \in \mathbb{R}^n$ then $Re(v) = a$, $Im(v) = b$.

Recall $z = x + iy \in \mathbb{C}$ has complex conjugate $z^* = x - iy$. Let $v \in \mathbb{C}^n$ we define the complex conjugate of the vector v to be v^* which is the vector of complex conjugates;

$$(v^*)_j = (v_j)^*$$

for each j = 1, 2, ..., n. For example, $[1 + i, 2, 3 - i]^* = [1 - i, 2, 3 + i]$. It is easy to verify the following properties for complex conjugation of numbers and vectors:

$$(v+w)^* = v^* + w^*,$$
 $(cv)^* = c^*v^*,$ $v^* = v.$

Complex matrices $\mathbb{C}^{m \times n}$ can be added, subtracted, multiplied and scalar multiplied in precisely the same ways as real matrices in $\mathbb{R}^{m \times n}$. However, we can also identify them as $\mathbb{C}^{m \times n} = \mathbb{R}^{m \times n} \oplus$ $i\mathbb{R}^{m \times n}$ via the real and imaginary part maps $(Re(Z))_{ij} = Re(Z_{ij})$ and $(Im(Z))_{ij} = Im(Z_{ij})$ for all i, j. There is an obvious isomorphism $\mathbb{C}^{m \times n} \cong \mathbb{R}^{2m \times 2n}$ which follows from stringing out all the real and imaginary parts. Again, complex conjugation is also defined component-wise: $((X + iY)^*)_{ij} = X_{ij} - iY_{ij}$. It's easy to verify that

$$(Z+W)^* = Z^* + W^*,$$
 $(cZ)^* = c^*Z^*,$ $(ZW)^* = Z^*W^*$

for appropriately sized complex matrices Z, W and $c \in \mathbb{C}$. Conjugation gives us a natural operation to characterize the *reality* of a variable. Let $c \in \mathbb{C}$ then c is **real** iff $c^* = c$. Likewise, if $v \in \mathbb{C}^n$ then we say that v is **real** iff $v^* = v$. If $Z \in \mathbb{C}^{m \times n}$ then we say that Z is **real** iff $Z^* = Z$. In short, an object is real if all its imaginary components are zero. Finally, while there is of course much more to say we will stop here for now.

10.1 the complexification

Suppose V is a vector space over \mathbb{R} , we seek to construct a new vector space $V_{\mathbb{C}}$ which is a natural extension of V. In particular, define:

$$V_{\mathbb{C}} = \{ (x, y) \mid x, y \in V \}$$

Suppose $(x, y), (v, w) \in V_{\mathbb{C}}$ and $a + ib \in \mathbb{C}$ where $a, b \in \mathbb{R}$. Define:

$$(x,y) + (v,w) = (x+v,y+w)$$
 & $(a+ib) \cdot (x,y) = (ax - by, ay + bx).$

I invite the reader to verify that $V_{\mathbb{C}}$ given the addition and scalar multiplication above forms a vector space over \mathbb{C} . In particular we may argue (0,0) is the zero in $V_{\mathbb{C}}$ and $1 \cdot (x,y) = (x,y)$. Moreover, as $x, y \in V$ and $a, b \in \mathbb{R}$ the fact that V is a real vector space yields $ax - by, ay + bx \in V$. The other axioms all follow from transferring the axioms over \mathbb{R} for V to $V_{\mathbb{C}}$. Our current notation for $V_{\mathbb{C}}$ is a bit tiresome. Note:

$$(1+0i) \cdot (x,y) = (x,y)$$
 & $(0+i) \cdot (x,y) = (-y,x).$

Since $\mathbb{R} \subset \mathbb{C}$ the fact that $V_{\mathbb{C}}$ is a complex vector space automatically makes $V_{\mathbb{C}}$ a real vector space. Moreover, with respect to the real vector space structure of $V_{\mathbb{C}}$, there are two natural subspaces of $V_{\mathbb{C}}$ which are isomorphic to V.

$$W_1 = (1+i0) \cdot V = V \times \{0\}$$
 & $W_2 = (0+i) \cdot V = \{0\} \times V$

Note $W_1 + W_2 = V_{\mathbb{C}}$ and $W_1 \cap W_2 = \{(0,0)\}$ hence $V_{\mathbb{C}} = W_1 \oplus W_2$. Here \oplus could be denoted $\oplus_{\mathbb{R}}$ to emphasize it is a direct sum with respect to the real vector space structure of $V_{\mathbb{C}}$. Moreover, it is convenient to simply write $V_{\mathbb{C}} = V \oplus iV$. Another notation for this is $V_{\mathbb{C}} = \mathbb{C} \otimes V$ where \otimes is the **tensor** product. This is perhaps the simplest way to think of the complexification:

To find the complexification of $V(\mathbb{R})$ we simply consider $V(\mathbb{C})$. In other words, replace the real scalars with complex scalars. This slogan is just a short-hand for the explicit construction outlined thus far in this section.

Example 10.1.1. If $V = \mathbb{R}$ then $V_{\mathbb{C}} = \mathbb{R} \oplus i\mathbb{R} = \mathbb{C}$.

Example 10.1.2. If $V = \mathbb{R}^n$ then $V_{\mathbb{C}} = \mathbb{R}^n \oplus i\mathbb{R}^n = \mathbb{C}^n$.

Example 10.1.3. If $V = \mathbb{R}^{m \times n}$ then $V_{\mathbb{C}} = \mathbb{R}^{m \times n} \oplus i\mathbb{R}^{m \times n} = \mathbb{C}^{m \times n}$.

We might notice a simple result about the basis of $V_{\mathbb{C}}$ which is easy to verify in the examples given thus far: if $\operatorname{span}_{\mathbb{R}}(\beta) = V$ then $\operatorname{span}_{\mathbb{C}}(\beta) = V_{\mathbb{C}}$. Furthermore, viewing $V_{\mathbb{C}}$ as real vector space, if β is a basis for V then $\beta \cup i\beta$ is a natural basis for $V_{\mathbb{C}}$. Although, it is often useful to order the real basis for $V_{\mathbb{C}}$ as follows: given $\beta = \{v_1, v_2, \ldots, v_n\}$ construct $\beta_{\mathbb{C}}$ as

$$\beta_{\mathbb{C}} = \{v_1, iv_1, v_2, iv_2, \dots, v_n, iv_n\}$$

Example 10.1.4. If $V = \mathbb{R}[t]$ then $V_{\mathbb{C}} = \mathbb{R}[t] \oplus i\mathbb{R}[t] = \mathbb{C}[t]$. Likewise for polynomials of limited degree. For example $W = P_2$ is given by $span_{\mathbb{R}}\{1, t, t^2\}$ whereas $W_{\mathbb{C}} = span_{\mathbb{R}}\{1, i, t, it, t^2, it^2\}$

From a purely complex perspective viewing an *n*-complex-dimensional space as a 2*n*-dimensional real vector space is ackward. However, in the application we are most interested, the complex vector space viewed as a real vector space yields data of interest to our study. We are primarily interested in solving real problems, but a complexification of the problem at times yields a simpler problem which is easily solved. Once the complexification has served its purpose of solvability then we have to drop back to the context of real vector spaces. This is the game plan, and the reason we are spending some effort to discuss complex vector spaces here.

Example 10.1.5. If $V = \mathcal{L}(U, W)$ then $V_{\mathbb{C}} = \mathcal{L}(U, W) \oplus i\mathcal{L}(U, W)$. If $T \in V_{\mathbb{C}}$ then $T = L_1 + iL_2$ for some $L_1, L_2 \in V$. However, if β is a basis for U then β is a complex basis for $U_{\mathbb{C}}$ thus T extends uniquely to a complex linear map $T : U_{\mathbb{C}} \to W_{\mathbb{C}}$. Therefore, we find $V_{\mathbb{C}} = \mathcal{L}_{\mathbb{C}}(U_{\mathbb{C}}, W_{\mathbb{C}})$

Example 10.1.6. As a particular application of the discussion in the last example: if $V = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ then $V_{\mathbb{C}} = \mathcal{L}_{\mathbb{C}}(\mathbb{C}^n, \mathbb{C}^m)$. Note that isomorphism and complexification intertwine nicely: $V \approx \mathbb{R}^{m \times n}$ and $\mathbb{C} \otimes V \approx \mathbb{C} \otimes \mathbb{R}^{m \times n}$ as $V_{\mathbb{C}} \approx \mathbb{C}^{m \times n}$.

The last example brings us to the main-point of this discussion. If we consider $T : \mathbb{R}^n \to \mathbb{R}^m$ s and we extend to $T : \mathbb{C}^n \to \mathbb{C}^m$ then this simply amounts to allowing the matrix of T be complex. Also, conversely, if we allow the matrix to be complex then it implies we have extended to a complex domain. The formula which defines the complexified version of a real linear transformation is simply:

$$T(x+iy) = T(x) + iT(y)$$

for all $x, y \in V$. This idea is at times tacitly used without any explicit mention of the complexification. In view of our discussion in this chapter that omission is not too dangerous. Indeed, that is why in other courses I at times just *allow* the variable to be complex. This amounts to the complexification procedure defined in this chapter.

Chapter 11

eigenvalues and eigenvectors

The terms eigenvalue and vector originate from the German school of mathematics which was very influential in the early 20-th century. Heisenberg's formulation of quantum mechanics gave new importance to linear algebra and in particular the algebraic structure of matrices. In finite dimensional quantum systems the symmetries of the system were realized by linear operators. These operators acted on states of the system which formed a complex vector space called Hilbert Space.¹

Operators representing momentum, energy, spin or angular momentum operate on a physical system represented by a sum of eigenfunctions. The eigenvalues then account for possible value which could be measured in an experiment. Generally, quantum mechanics involves complex vector spaces and infinite dimensional vector spaces however many of the mathematical difficulties are already present in our study of linear algebra. For example, one important question is how does one pick a set of states which diagonalize an operator? Moreover, if one operator is diagonalized by a particular basis then can a second operator be diagonalized simultaneously? Linear algebra, and in particular eigenvectors help give answers for these questions. ²

Beyond, or perhaps I should say before, quantum mechanics eigenvectors have great application in classical mechanics, optics, population growth, systems of differential equations, chaos theory, difference equations and much much more. They are a fundmental tool which allow us to pick apart a matrix into its very core. Diagonalization of matrices almost always allow us to see the nature of a system more clearly.

However, not all matrices are diagonalizable. It turns out that any matrix is similar to a real Jordan Block matrix. Moreover, the similarity transformation is accomplished via a matrix formed from concatenating generalized eigenvectors and certain parts of complex eigenvectors. When there are enough ordinary eigenvectors then the Jordan Form of the matrix is actually a diagonal matrix. However, when there is no eigen-basis then we must turn to generalized e-vectors and/or complex e-vectors and generalized complex e-vectors to perform a similarity transformation to the real Jordan form. The existence of the real Jordan form for any matrix serves to be a useful tool to solve a variety of problems from any field which uses linear algebra to solve a coupled system. In some sense, the coordinates paired with the Jordan form correspond to a coordinate system which

¹Hilbert Spaces and infinite dimensional linear algebra are typically discussed in graduate linear algebra and/or the graduate course in functional analysis, we focus on the basics in this course.

²in addition to linear algebra one should also study group theory. In particular, matrix Lie groups and their representation theory explains most of what we call "chemistry". Magic numbers, electronic numbers, etc... all of these are eigenvalues which label the states of the so-called Casimir operators

presents the system with minimal coupling. Complex eigenvalues correspond to rotation/dilation blocks. Also, in the study of systems of ODEs with constant coefficients the Jordan form again allows for an elegant and general solution of any system in normal form. I include a section at the end of this chapter to show you the magic of the matrix exponential paired with the Jordan basis.

11.1 why eigenvectors?

In this section I attempt to motivate why eigenvectors are natural to study for both mathematical and physical reasons. In fact, you probably could write a book just on this question.

11.1.1 quantum mechanics

Physically measureable quantities are described by operators and states in quantum mechanics³. The operators are linear operators and the states are usually taken to be the eigenvectors with respect to a physical quantity of interest. For example:

$$\hat{p}|p>=p|p>$$
 $\hat{J}^{2}|j>=j(j+1)|j>$ $\hat{H}|E>=E|E>$

In the above the eigenvalues were p, j(j+1) and E. Physically, p is the momentum, j(j+1) is the value of the square of the magnitude of the total angular momentum and E is the energy. The exact mathematical formulation of the eigenstates of momentum, energy and angular momentum is in general a difficult problem and well-beyond the scope of the mathematics we cover this semester. You have to study Hilbert space which is an infinite-dimensional vector space with rather special properties. In any event, if the physical system has nice boundary conditions then the quantum mechanics gives mathematics which is within the reach of undergraduate linear algebra. For example, one of the very interesting aspects of quantum mechanics is that we can only measure a certain pairs of operators simultaneously. Such operators have eigenstates which are simultaneously eigenstates of both operators at once. The careful study of how to label states with respect to the energy operator (called the Hamiltonian) and some other commuting symmetry operator (like isospin or angular momentum etc...) gives rise to what we call Chemistry. In other words, Chemistry is largely the tabulation of the specific interworkings of eigenstates as the correlate to the energy, momentum and spin operators (this is a small part of what is known as representation theory in modern mathematics). I may ask a question about simultaneous diagonalization. This is a hard topic compared to most we study.

11.1.2 stochastic matrices

Definition 11.1.1.

Let $P \in \mathbb{R}^{n \times n}$ with $P_{ij} \ge 0$ for all i, j. If the sum of the entries in any column of P is one then we say P is a stochastic matrix.

Example 11.1.2. Stochastic Matrix: A medical researcher⁴ is studying the spread of a virus in 1000 lab. mice. During any given week it's estimated that there is an 80% probability that a mouse will overcome the virus, and during the same week there is an 10% likelyhood a healthy mouse will

³you can skip this if you're not a physics major, but maybe you're interested despite the lack of direct relevance to your major. Maybe your interested in an education not a degree. I believe this is possible so I write these words

⁴this example and most of the other applied examples in these notes are borrowed from my undergraduate linear algebra course taught from Larson's text by Dr. Terry Anderson of Appalachian State University

become infected. Suppose 100 mice are infected to start, (a.) how many sick next week? (b.) how many sick in 2 weeks? (c.) after many many weeks what is the steady state solution?

$$I_{k} = infected mice at beginning of week k$$
$$P = \begin{bmatrix} 0.2 & 0.1 \\ 0.8 & 0.9 \end{bmatrix}$$

We can study the evolution of the system through successive weeks by multiply the state-vector $X_k = [I_k, N_k]$ by the probability transition matrix P given above. Notice we are given that $X_1 = [100, 900]^T$. Calculate then,

$$X_2 = \left[\begin{array}{cc} 0.2 & 0.1\\ 0.8 & 0.9 \end{array}\right] \left[\begin{array}{c} 100\\ 900 \end{array}\right] = \left[\begin{array}{c} 110\\ 890 \end{array}\right]$$

After one week there are 110 infected mice Continuing to the next week,

$$X_3 = \left[\begin{array}{cc} 0.2 & 0.1 \\ 0.8 & 0.9 \end{array} \right] \left[\begin{array}{c} 110 \\ 890 \end{array} \right] = \left[\begin{array}{c} 111 \\ 889 \end{array} \right]$$

After two weeks we have 111 mice infected. What happens as $k \to \infty$? Generally we have $X_k = PX_{k-1}$. Note that as k gets large there is little difference between k and k-1, in the limit they both tend to infinity. We define the steady-state solution to be $X^* = \lim_{k\to\infty} X_k$. Taking the limit of $X_k = PX_{k-1}$ as $k \to \infty$ we obtain the requirement $X^* = PX^*$. In other words, the steady state solution is found from solving $(P-I)X^* = 0$. For the example considered here we find,

$$(P-I)X^* = \begin{bmatrix} -0.8 & 0.1\\ 0.8 & -0.1 \end{bmatrix} \begin{bmatrix} u\\ v \end{bmatrix} = 0 \qquad v = 8u \qquad X^* = \begin{bmatrix} u\\ 8u \end{bmatrix}$$

However, by conservation of mice, u + v = 1000 hence 9u = 1000 and $u = 111.\overline{11}$ thus the steady state can be shown to be $X^* = [111.\overline{11}, 888.\overline{88}]$

Example 11.1.3. Diagonal matrices are nice: Suppose that demand for doorknobs halves every week while the demand for yo-yos it cut to 1/3 of the previous week's demand every week due to an amazingly bad advertising campaign⁵. At the beginning there is demand for 2 doorknobs and 5 yo-yos.

$$\begin{aligned} D_k &= demand \ for \ doorknobs \ at \ beginning \ of \ week \ k \\ Y_k &= demand \ for \ yo-yos \ at \ beginning \ of \ week \ k \end{aligned} \qquad P = \left[\begin{array}{cc} 1/2 & 0 \\ 0 & 1/3 \end{array} \right] \end{aligned}$$

We can study the evolution of the system through successive weeks by multiply the state-vector $X_k = [D_k, Y_k]$ by the transition matrix P given above. Notice we are given that $X_1 = [2, 5]^T$. Calculate then,

$$X_2 = \begin{bmatrix} 1/2 & 0\\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} 2\\ 5 \end{bmatrix} = \begin{bmatrix} 1\\ 5/3 \end{bmatrix}$$

Notice that we can actually calculate the k-th state vector as follows:

$$X_{k} = P^{k}X_{1} = \begin{bmatrix} 1/2 & 0\\ 0 & 1/3 \end{bmatrix}^{k} \begin{bmatrix} 2\\ 5 \end{bmatrix} = \begin{bmatrix} 2^{-k} & 0\\ 0 & 3^{-k} \end{bmatrix}^{k} \begin{bmatrix} 2\\ 5 \end{bmatrix} = \begin{bmatrix} 2^{-k+1}\\ 5(3^{-k}) \end{bmatrix}$$

⁵insert your own more interesting set of quantities that doubles/halves or triples during some regular interval of time

Therefore, assuming this silly model holds for 100 weeks, we can calculate the 100-the step in the process easily,

$$X_{100} = P^{100} X_1 = \begin{bmatrix} 2^{-101} \\ 5(3^{-100}) \end{bmatrix}$$

Notice that for this example the analogue of X^* is the zero vector since as $k \to \infty$ we find X_k has components which both go to zero.

For some systems we'll find a special state we called the "steady-state" for the system. If the system was attracted to some particular final state as $t \to \infty$ then that state satisfied $PX^* = X^*$. We will learn in this chapter to identify this makes X^* is an eigenvector of P with eigenvalue 1.

11.1.3 motion of points under linear transformations

Remark 11.1.4.

What follows here is just intended to show you how you might stumble into the concept of an eigenvector even if you didn't set out to find it. The calculations we study here are not what we aim to ultimately disect in this chapter. This is purely a mathematical experiment to show how eigenvectors arise naturally through repeated matrix multiplication on a given point. Physically speaking the last two subsections were way more interesting.

I'll focus on two dimensions to begin for the sake of illustration. Let's take a matrix A and a point x_o and study what happens as we multiply by the matrix. We'll denote $x_1 = Ax_o$ and generally $x_{k+1} = Ax_k$. It is customary to call x_k the "k-th state of the system". As we multiply the k-th state by A we generate the k + 1-th state.⁶

Example 11.1.5. Let $A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$ and let $x_o = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Calculate,

x_1	=	$\begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$
x_2	=	$\begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \end{bmatrix} = \begin{bmatrix} 9 \\ 18 \end{bmatrix}$
x_3	=	$\begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 9 \\ 18 \end{bmatrix} = \begin{bmatrix} 27 \\ 54 \end{bmatrix}$
x_4	=	$\left[\begin{smallmatrix}3&0\\8&-1\end{smallmatrix}\right]\left[\begin{smallmatrix}27\\54\end{smallmatrix}\right]=\left[\begin{smallmatrix}81\\162\end{smallmatrix}\right]$

Each time we multiply by A we scale the vector by a factor of three. If you want to look at x_o as a point in the plane the matrix A pushes the point x_k to the point $x_{k+1} = 3x_k$. Its not hard to see that $x_k = 3^k x_o$. What if we took some other point, say $y_o = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ then what will A do?

$y_1 =$	$\begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$	$\begin{bmatrix} 1\\0\end{bmatrix} = \begin{bmatrix} 3\\8\end{bmatrix}$
$y_2 =$	$\left[\begin{smallmatrix}3&0\\8&-1\end{smallmatrix}\right]$	$\left[\begin{smallmatrix}3\\8\end{smallmatrix}\right] = \left[\begin{smallmatrix}9\\16\end{smallmatrix}\right]$
$y_3 =$	$\left[\begin{smallmatrix}3&0\\8&-1\end{smallmatrix}\right]$	$\left[\begin{smallmatrix}9\\16\end{smallmatrix}\right]=\left[\begin{smallmatrix}27\\56\end{smallmatrix}\right]$
$y_4 =$	$\left[\begin{smallmatrix}3&0\\8&-1\end{smallmatrix}\right]$	$\left[\begin{smallmatrix}27\\48\end{smallmatrix}\right] = \left[\begin{smallmatrix}81\\160\end{smallmatrix}\right]$

Now, what happens for arbitrary k? Can you find a formula for y_k ? This point is not as simple as

⁶ask Dr. Mavinga and he will show you how a recursively defined linear difference equation can be converted into a matrix equation of the form $x_{k+1} = Ax_k$, this is much the same idea as saying that an n - th order ODE can be converted into a system of n- first order ODEs.

 x_o . The vector x_o is apparently a special vector for this matrix. Next study, $z_o = \begin{bmatrix} 0\\2 \end{bmatrix}$,

$$z_{1} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$$
$$z_{2} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$$
$$z_{3} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ -8 \end{bmatrix}$$
$$z_{4} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -8 \end{bmatrix} = \begin{bmatrix} 0 \\ 16 \end{bmatrix}$$

Let me illustrate what is happening with a picture. I have used color to track the motion of a particular point. You can see that all points tend to get drawn into the line with direction vector x_o with the sole exception of the points along the y-axis which I have denoted via diamonds in the picture below:



The directions [1,2] and [0,1] are special, the following picture illustrates the motion of those points under A:



The line with direction vector [1, 2] seems to attract almost all states to itself. On the other hand, if you could imagine yourself a solution walking along the y-axis then if you took the slightest mis-step to the right or left then before another dozen or so steps you'd find yourself stuck along the line in the [1, 2]-direction. There is a connection of the system $x_{k+1} = Ax_k$ and the system of differential equations dx/dt = Bx if we have B = I + A. Perhaps we'll have time to explore the questions

posed in this example from the viewpoint of the corresponding system of differential equations. In this case the motion is very discontinuous. I think you can connect the dots here to get a rough picture of what the corresponding system's solutions look like. In the differential equations Chapter we develop these ideas a bit further. For now we are simply trying to get a feeling for how one might discover that there are certain special vector(s) associated with a given matrix. We call these vectors the "eigenvectors" of A.

The next matrix will generate rather different motions on points in the plane.

Example 11.1.6. Let
$$A = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$
. Consider the trajectory of $x_o = [1, 0]^T$,
 $x_1 = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix}$
 $x_2 = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix}$
 $x_3 = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$
 $x_4 = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{bmatrix}$
 $x_5 = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
 $x_6 = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

Past this point we just cycle back to the same points, clearly $x_k = x_{k+6}$ for all $k \ge 0$. If we started with a different initial point we would find this pattern again. The reason for this is that A is the matrix which rotates vectors by $\pi/3$ radians. The trajectories generated by this matrix are quite different then the preceding example, there is no special direction in this case.



Although, generally this type of matrix generates elliptical orbits and then there are two special directions. Namely the major and minor axis of the elliptical orbits. Finally, this sort of matrix could

have a scaling factor built in so that the trajectories spiral in or out of the origin. I provide a picture illustrating the various possibilities. The red dots in the picture below are generated from A as was given in the preceding example, the blue dots are generated from the matrix $[\frac{1}{2}col_1(A)|col_2(A)]$ whereas the green dots are obtained from the matrix $[2col_1(A)|col_2(A)]$. In each case I started with the point (1,0) and studied the motion of the point under repeated multiplications of matrix:



Let's summarize our findings so far: if we study the motion of a given point under successive multiplications of a matrix it may be pushed towards one of several directions or it may go in a circular/spiral-type motion.

Observation: a rotation does not move vectors along its axis; if R is a rotation in \mathbb{R}^3 and if x_o points along the axis of the rotation then it is geometrically obvious that $Rx_o = x_o$. For a two-dimensional rotation the axis of rotation is not contained in the space so there is no vector like x_o . By the end of our study in this chapter we can replace geometric intuition here with an explicit algebraic analysis. The basis for our conclusions will be the simple observation that a vector along the axis of a rotation is an eigenvector with eigenvalue 1. Moreover, through fairly simple algebra, we can show there is always an axis of rotation inside \mathbb{R}^n if n is odd. However, for even n, the the axis of the rotation may not reside inside the space. These abstract geometric claims fall out of polynomial algebra. I hope by now you see eigenvectors appear naturally in a variety of applications and we should like to understand their general properties.

11.2 basic theory of eigenvectors

Eigenvectors are of special significance to a linear transformation. Let us define them precisely to begin our careful study⁷.

Definition 11.2.1. eigenvector of a linear transformation on V

Let $T: V \to V$ be a linear transformation on a vector space V over \mathbb{R} . If there exists $v \in V$ such that $v \neq 0$ such that $T(v) = \lambda v$ for some constant $\lambda \in \mathbb{R}$ then we say v is an **eigenvector** of T with **eigenvalue** λ .

We often abbreviate *eigenvector* with **e-vector** and *eigenvalue* with **e-value**.

Example 11.2.2. Let T(f) = Df where D is the derivative operator. This defines a linear transformation on function space \mathcal{F} . An eigenvector for T would be a function which is proportional to its own derivative function... in other words solve $\frac{dy}{dt} = \lambda y$. Separation of variables yields $y = ce^{\lambda t}$. The eigenfunctions for T are simply exponential functions.

Example 11.2.3. Let $T(A) = A^T$ for $A \in \mathbb{R}^{n \times n}$. If $A^T = A$ the $n T(A) = A^T = A$ so a symmetric matrix is an e-vector with e-value $\lambda_1 = 1$. On the other hand, if $A^T = -A$ then $T(A) = A^T = -A$ hence an antisymmetric matrix is an e-vector of T with e-value $\lambda_2 = -1$.

Notice that there are infinitely many eigenvectors for a given eigenvalue in both of the examples above. The number of eigenvalues for the function space example is infinite since any $\lambda \in \mathbb{R}$ will do. On the other hand, the matrix example only had two eigenvalues. The distinction between these examples is that function space is infinite dimensional whereas the matrix example is finitedimensional. The following proposition gives us a criteria to find e-values for $T: V \to V$ in the case dim $(V) < \infty$. Recall for the proposition the follows: $Id_V: V \to V$ is the identity mapping defined by $Id_V(x) = x$ for all $x \in V$.

Proposition 11.2.4.

Suppose V is a finite-dimensional vector space over \mathbb{R} . Let $T: V \to V$ be a linear transformation. Then λ is an eigenvalue of T iff $det(T - \lambda Id_V) = 0$. We say $P(\lambda) = det(T - \lambda Id_V)$ the **characteristic polynomial** and $det(T - \lambda Id_V) = 0$ is the **characteristic equation**.

Proof: Observe $\lambda \in \mathbb{R}$ is an e-value of T iff there exists nonzero $v \in V$ for which $T(v) = \lambda v$ which is equivalent to the existence of a nontrivial solution of $(T - \lambda I d_V)(v) = 0$. However, nullity $(T - \lambda I d_V) \geq 1$ iff $T - \lambda I d_V$ is not invertible which is true iff $\det(T - \lambda I d_V) = 0$. The proposition follows. \Box

Often the calculation of some quantity for a linear transformation is made clear by choice of a basis. This is certainly the case here:

Proposition 11.2.5.

If V has basis $\beta = \{f_1, \ldots, f_n\}$ and $T: V \to V$ is a linear transformation. Then λ is an eigenvalue of T iff $det([T]_{\beta,\beta} - \lambda I) = 0$.

⁷in past courses I allowed $\lambda \in \mathbb{C}$ here, however, this time I will treat that case separately as the proper discussion for $\lambda \in \mathbb{C}$ requires we discuss the complexification. A complex eigenvector is technically not an eigenvector in our langauge. An eigenvector in this chapter is by default a **real** eigenvector.

Proof: Suppose $\lambda \in \mathbb{R}$ and let $Id : V \to V$. Consider $T - \lambda Id : V \to V$ has matrix representative with respect to the β basis as follows:

$$[T - \lambda Id]_{\beta,\beta} = [\Phi_{\beta}^{-1} \circ (T - \lambda Id) \circ \Phi_{\beta}]$$

= $[\Phi_{\beta}^{-1} \circ T \circ \Phi_{\beta} - \lambda \Phi_{\beta}^{-1} \circ Id \circ \Phi_{\beta}]$
= $[\Phi_{\beta}^{-1} \circ T \circ \Phi_{\beta}] - \lambda [\Phi_{\beta}^{-1} \circ Id \circ \Phi_{\beta}]$
= $[T]_{\beta,\beta} - \lambda I.$

In the calculation above we have used the fact that the coordinate map is an isomorphism and we may easily calculate $[T - \lambda Id]_{\beta,\beta} = \lambda I$ where I is the $n \times n$ identity matrix in $\mathbb{R}^{n \times n}$. Furthermore, recall the definition of determinant for a linear transformation was that the determinant of a linear transformation is the determinant of the matrix representative:

$$\det(T - \lambda Id) = \det([T - \lambda Id]_{\beta,\beta}) = \det([T]_{\beta,\beta} - \lambda I).$$

The above is an identity for any real value λ . The proposition follows. \Box

The proposition above shows that we can narrow the focus of our study to $\mathbb{R}^{n \times n}$. If $T : V \to V$ has e-values then its matrix will share the same e-values since the matrix and the transformation share the same characteristic polynomials. For clarity of exposition we define e-vector and e-value once again for a matrix.

Definition 11.2.6.

Let $A \in \mathbb{R}^{n \times n}$. If $v \in \mathbb{R}^n$ is **nonzero** and $Av = \lambda v$ for some $\lambda \in \mathbb{R}$ then we say v is an **eigenvector** with **eigenvalue** λ of the matrix A.

The definition above simply says the e-values and e-vectors of A are the e-values and e-vectors of $L_A : \mathbb{R}^n \to \mathbb{R}^n$. We now turn our focus to matrices in what follows. The Proposition that follows is just a Corollary of Proposition 11.2.4. However, since these comments are so important I reiterate them once again in the matrix context:

Proposition 11.2.7.

Let $A \in \mathbb{R}^{n \times n}$ then λ is an eigenvalue of A iff $det(A - \lambda I) = 0$. We say $P(\lambda) = det(A - \lambda I)$ the **characteristic polynomial** and $det(A - \lambda I) = 0$ is the **characteristic equation**.

Proof: Suppose λ is an eigenvalue of A then there exists a nonzero vector v such that $Av = \lambda v$ which is equivalent to $Av - \lambda v = 0$ which is precisely $(A - \lambda I)v = 0$. Notice that $(A - \lambda I)0 = 0$ thus the matrix $(A - \lambda I)$ is singular as the equation $(A - \lambda I)x = 0$ has more than one solution. Consequently $det(A - \lambda I) = 0$.

Conversely, suppose $det(A - \lambda I) = 0$. It follows that $(A - \lambda I)$ is singular. Clearly the system $(A - \lambda I)x = 0$ is consistent as x = 0 is a solution hence we know there are infinitely many solutions. In particular there exists at least one vector $v \neq 0$ such that $(A - \lambda I)v = 0$ which means the vector v satisfies $Av = \lambda v$. Thus v is an eigenvector with eigenvalue λ for $A \square$

Let's collect the observations of the above proof for future reference.

Proposition 11.2.8.

The following are equivalent for $A \in \mathbb{R}^{n \times n}$ and $\lambda \in \mathbb{R}$,

- 1. λ is an eigenvalue of A
- 2. there exists $v \neq 0$ such that $Av = \lambda v$
- 3. there exists $v \neq 0$ such that $(A \lambda I)v = 0$
- 4. λ is a solution to $det(A \lambda I) = 0$
- 5. $(A \lambda I)v = 0$ has infinitely many solutions.

Many examples are given in Section 11.4. For clarity of logical structure we continue discussion of the theory of eigenvectors here.

Proposition 11.2.9.

There exist $n \in \mathbb{N}$ and $A \in \mathbb{R}^{n \times n}$ for which no eigenvectors exist.

Proof: choose n = 2 and study $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. We calculate $\det(A - \lambda I) = \det \begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 + 1.$

Therefore, no real solutions to the characteristic equation exist. Notice, if there existed an e-vector v with e-value $\lambda_r \in \mathbb{R}$ then that would imply λ_r solves $\lambda^2 + 1 = 0$ by Proposition 11.2.8. However, this is impossible as $\lambda^2 + 1 = 0$ has no real solutions. \Box

The theory of polynomial factoring plays a large role in the theory of eigenvectors. We just saw how an irreducible quadratic was tied to the non-existence of e-vectors, we next see how the existence of real roots for odd-order polynomials forces all odd-sized matrices to have at least one e-vector.

Proposition 11.2.10.

If $k \in \mathbb{N}$ and n = 2k - 1 then $A \in \mathbb{R}^{n \times n}$ has at least one eigenvector.

Proof: Observe $P(x) = \det(A - xI)$ has $\deg(P) = n$ with *n*-odd. Therefore, there exists $r \in \mathbb{R}$ for which P(r) = 0 and by Proposition 11.2.8 there exists $v \neq 0$ for which Av = rv. \Box

If the reader forgot, you can argue for sufficiently large magnitude a an odd-order polynomial must change sign; P(a)P(-a) < 0. The intermediate value theorem applies as polynomials are continuous and thus we find a real root must exist. Of course, there is also a purely algebraic argument which is derived from the fundamental theorem of algebra. We'll discuss the algebraic structure of polynomials in greater depth as we elevate the discussion to include complex eigenvectors in Section 11.3.

Proposition 11.2.11.

If P is the characteristic polynomial of $A \in \mathbb{R}^{n \times n}$ then $P(\lambda) = c_o + c_1 \lambda + \dots + (-1)^n \lambda^n$ and $c_0 = \det(A)$.

Proof: by definition $P(\lambda) = det(A - \lambda I)$ and clearly $P(\lambda)$ is an *n*-th order polynomial with leading coefficient $(-1)^n$ hence there exist constants $c_0, c_1, \dots, c_{n-1} \in \mathbb{R}$ for which $P(\lambda) = c_o + c_1 \lambda + \dots + (-1)^n \lambda^n$. Set $\lambda = 0$ to obtain $c_o = det(A)$. \Box

The proposition above is an interesting check on a set of proposed e-values. Incidentally, it continues to hold for complex e-values. For example, the proof of Proposition 11.2.9 had characteristic polynomial $\lambda^2 + 1 = 0$ which corresponds to $\lambda_1 = i$ and $\lambda_2 = -i$ of course $\lambda_1 \lambda_2 = -i^2 = 1 = \det(A)$. Naturally, you should complain that I have not yet defined complex eigenvalues. We shall soon.

Proposition 11.2.12.

Zero is an eigenvalue of A iff A is a singular matrix.

Proof: Let $P(\lambda)$ be the characteristic polynomial of A. If zero is an eigenvalue then λ must factor the characteristic polynomial. Moreover, the factor theorem tells us that P(0) = 0 since $(\lambda - 0)$ factors $P(\lambda)$. Thus $c_0 = 0$ and we deduce using the previous proposition that $det(A) = c_0 = 0$. Which shows that A is singular. The converse follows by the same argument reversed. \Box

Proposition 11.2.13.

If $A \in \mathbb{R}^{n \times n}$ then A has n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ then $det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$.

Proof: If $A \in \mathbb{R}^{n \times n}$ then A has n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ then the characteristic polynomial factors over \mathbb{R} :

$$det(A - \lambda I) = k(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

Moreover, the leading term in $P(\lambda)$ obtains a coefficient of $(-1)^n$ hence $k = (-1)^n$. If c_0 is the constant term in the characteristic polynomial then algebra reveals that $c_0 = (-1)^n (-\lambda_1)(-\lambda_2) \cdots (-\lambda_n) = \lambda_1 \lambda_2 \dots \lambda_n$. Therefore, using Proposition 11.2.11, $det(A) = \lambda_1 \lambda_2 \dots \lambda_n$. \Box .

Proposition 11.2.14.

If $A \in \mathbb{R}^{n \times n}$ has e-vector v with eigenvalue λ then v is a e-vector of A^k with e-value λ^k .

Proof: let $A \in \mathbb{R}^{n \times n}$ have e-vector v with eigenvalue λ . Consider,

$$A^{k}v = A^{k-1}Av = A^{k-1}\lambda v = \lambda A^{k-2}Av = \lambda^{2}A^{k-2}v = \dots = \lambda^{k}v$$

The \cdots is properly replaced by a formal induction argument. \Box .

Proposition 11.2.15.

Let A be a upper or lower triangular matrix then the eigenvalues of A are the diagonal entries of the matrix.

Proof: follows immediately from Proposition 8.4.3 since the diagonal entries of $A - \lambda I$ are of the form $A_{ii} - \lambda$ hence the characteristic equation has the form $det(A - \lambda I) = (A_{11} - \lambda)(A_{22} - \lambda) \cdots (A_{nn} - \lambda)$ which has solutions $\lambda = A_{ii}$ for i = 1, 2, ..., n. \Box

Proposition 11.2.16.

Let $A \in \mathbb{R}^{2 \times 2}$. The eigenvalues are determine the det(A) and trace(A): det $(A) = \lambda_1 \lambda_2$ & trace $(A) = \lambda_1 + \lambda_2$.

Proof: we know Proposition 11.2.13 yields $det(A) = \lambda_1 \lambda_2$. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then $P(x) = det \begin{bmatrix} a-x & b \\ c & d-x \end{bmatrix} = (x-a)(x-d) - bc$. Algebra reveals $P(x) = x^2 - (a+d)x + ad - bc$ and completing the square yields:

$$\lambda_{\pm} = \frac{a+d \pm \sqrt{(a+d)^2 + 4bc}}{2}$$

Let $\lambda_1 = \lambda_+$ and $\lambda_2 = \lambda_-$. Observe $\lambda_1 + \lambda_2 = a + d = \text{trace}(A)$. \Box

Perhaps you will not find it suprising that the algebra above equally well applies if $\lambda_{\pm} \in \mathbb{C}$. In fact, the proposition above also applies to $A \in \mathbb{R}^{n \times n}$. We can show, $\operatorname{trace}(A) = \sum_{j=1}^{n} \lambda_j$ where λ_j are eigenvalues of A. That general result also applies to the case of complex eigenvalues. I think proving that in the same way as we did for n = 2 would be nearly impossible. Instead, we turn to the question of linear independence. We saw orthonormality implied LI with little effort. We now learn that distinct e-values also provide LI.

Proposition 11.2.17.

If $A \in \mathbb{R}^{n \times n}$ has e-vector v_1 with e-value λ_1 and e-vector v_2 with e-value λ_2 such that $\lambda_1 \neq \lambda_2$ then $\{v_1, v_2\}$ is linearly independent.

Proof: Let v_1, v_2 have e-values λ_1, λ_2 respective and assume towards a contradiction that $v_2 = kv_2$ for some nonzero constant k. Multiply by the matrix A,

$$Av_1 = A(kv_2) \quad \Rightarrow \quad \lambda_1 v_1 = k\lambda_2 v_2$$

But we can replace v_1 on the l.h.s. with kv_2 hence,

$$\lambda_1 k v_2 = k \lambda_2 v_2 \quad \Rightarrow \quad k(\lambda_1 - \lambda_2) v_2 = 0$$

Note, $k \neq 0$ and $v_2 \neq 0$ by assumption thus the equation above indicates $\lambda_1 - \lambda_2 = 0$ therefore $\lambda_1 = \lambda_2$ which is a contradiction. Therefore there does not exist such a k and the vectors are linearly independent. \Box

A direct argument is also possible. Suppose $\{v_1, v_2\}$ is a set of nonzero vectors with $Av_1 = \lambda_1 v_1$ and $Av_2 = \lambda_2 v_2$ suppose $c_1 v_1 + c_2 v_2 = 0$. Multiply by $A - \lambda_1 I$,

$$c_1(A - \lambda_1 I)v_1 + c_2(A - \lambda_1 I)v_2 = 0 \quad \Rightarrow \quad c_2(\lambda_2 - \lambda_1)v_2 = 0$$

as $\lambda_2 - \lambda_1 \neq 0$ and $v_2 \neq 0$ hence $c_2 = 0$. Multiplication by $A - \lambda_2 I$ likewise reveals $c_1 = 0$. Therefore, $\{v_1, v_2\}$ is LI. You can choose which proof you think is best.

Proposition 11.2.18.

If $A \in \mathbb{R}^{n \times n}$ has eigenvectors v_1, v_2, \ldots, v_k with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ such that $\lambda_i \neq \lambda_j$ for all $i \neq j$ then $\{v_1, v_2, \ldots, v_k\}$ is linearly independent.

Proof: I begin with a direct proof. Suppose v_1, v_2, \ldots, v_k are e-vectors with e-values $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ such that $\lambda_i \neq \lambda_j$ for all $i \neq j$. Suppose $c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0$. Multiply by $\prod_{i=1}^{k-1} (A - \lambda_i I)$,

$$c_1 \prod_{i=1}^{k-1} (A - \lambda_i I) v_1 + \dots + c_{k-1} \prod_{i=1}^{k-1} (A - \lambda_i I) v_{k-1} + c_k \prod_{i=1}^{k-1} (A - \lambda_i I) v_k = 0 \star$$

Consider that the terms in the product commute as:

$$(A - \lambda_i I)(A - \lambda_j I) = A^2 - (\lambda_i - \lambda_j)A + \lambda_i \lambda_j I = (A - \lambda_j I)(A - \lambda_i I).$$

It follows that we can bring $(A - \lambda_j I)$ to the right of the product multiplying the *j*-th summand:

$$c_1 \prod_{i \neq 1}^{k-1} (A - \lambda_i I) (A - \lambda_1 I) v_1 + \dots + c_{k-1} \prod_{i \neq k-1}^{k-1} (A - \lambda_i I) (A - \lambda_{k-1} I) v_{k-1} + c_k \prod_{i=1}^{k-1} (A - \lambda_i I) v_k = 0 \star^2$$

Notice, for $i \neq j$, $(A - \lambda_j I)v_i = \lambda_i v_i - \lambda_j v_i = (\lambda_i - \lambda_j)v_i \neq 0$ as $\lambda_i \neq \lambda_j$ and $v_i \neq 0$. On the other hand, if i = j then $(A - \lambda_i I)v_i = \lambda_i v_i - \lambda_i v_i = 0$. Therefore, in \star we find that terms with coefficients $c_1, c_2, \ldots, c_{k-1}$ all vanish. All that remains is:

$$c_k \prod_{i=1}^{k-1} (A - \lambda_i I) v_k = 0 \quad \star^3$$

We calculate,

$$\begin{split} \prod_{i=1}^{k-1} (A - \lambda_i I) v_k &= \prod_{i=1}^{k-2} (A - \lambda_i I) (A - \lambda_{k-1} I) v_k = (\lambda_k - \lambda_{k-1}) \prod_{i=1}^{k-2} (A - \lambda_i I) v_k \\ &= (\lambda_k - \lambda_{k-1}) (\lambda_k - \lambda_{k-2}) \prod_{i=1}^{k-3} (A - \lambda_i I) v_k \\ &= (\lambda_k - \lambda_{k-1}) (\lambda_k - \lambda_{k-2}) \cdots (\lambda_k - \lambda_1) v_k. \end{split}$$

However, as $v_k \neq 0$ and $\lambda_k \neq \lambda_i$ for $i = 1, \ldots k - 1$ it follows that \star^3 implies $c_k = 0$. Next, we repeat the argument, except only multiply \star by $\prod_{i=1}^{k-1} (A - \lambda_i)$ which yields $c_{k-1} = 0$. We continue in this fashion until we have shown $c_1 = c_2 = \cdots = c_k = 0$. Hence $\{v_1, \ldots, v_k\}$ is LI as claimed. \Box

I am fond of the argument which was just offered. Technically, it could be improved by including explicit induction arguments in place of \cdots . The next argument is similar to our initial argument for two vectors.

Proof: Let e-vectors v_1, v_2, \ldots, v_k have e-values $\lambda_1, \lambda_2, \ldots, \lambda_k$. Let us prove the claim by induction on k. Note k = 1 and k = 2 we have already shown in previous work. Suppose inductively the claim is true for k - 1. Consider, towards a contradiction, that there is some vector v_j which is a nontrivial linear combination of the other vectors:

$$v_j = c_1 v_1 + c_2 v_2 + \dots + \widehat{c_j v_j} + \dots + c_k v_k$$

Multiply by A,

$$Av_j = c_1 Av_1 + c_2 Av_2 + \dots + \widehat{c_j Av_j} + \dots + c_k Av_k$$

Which yields,

$$\lambda_j v_j = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \dots + \widetilde{c_j} \lambda_j \widetilde{v_j} + \dots + c_k \lambda_k v_k$$

But, we can replace v_i on the l.h.s with the linear combination of the other vectors. Hence

$$\lambda_j \left[c_1 v_1 + c_2 v_2 + \dots + \widehat{c_j v_j} + \dots + c_k v_k \right] = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \dots + \widehat{c_j \lambda_j v_j} + \dots + c_k \lambda_k v_k$$

Consequently,

$$c_1(\lambda_j - \lambda_1)v_1 + c_2(\lambda_j - \lambda_2)v_2 + \dots + c_j(\widehat{\lambda_j - \lambda_j})v_j + \dots + c_k(\lambda_j - \lambda_k)v_k = 0$$

However, this is a set of k-1 e-vectors with distinct e-values linearly combined to give zero. It follows from the induction claim that each coefficient is trivial. As $\lambda_j \neq \lambda_i$ for $i \neq j$ it is thus necessary that $c_1 = c_2 = \cdots = c_k = 0$. But, this implies $v_j = 0$ which contradicts $v_j \neq 0$ as is known since v_j was assumed an e-vector. Hence $\{v_1, \ldots, v_k\}$ is LI as claimed and by induction on $k \in \mathbb{N}$ we find the proposition is true. \Box

Doubtless there are improvements and refinements of both versions of the proofs I offer here. Moreover, you may be annoyed to have me point out yet again these LI results also transfer to the context of distinct complex eigenvalues. That said, I suppose I should finally get to the task of defining the complex eigenvalue.

11.3 complex eigenvalues and vectors

By now it should be clear that as we consider problems of real vector spaces the general results, especially those algebraic in nature, invariably involve some complex case. However, technically it usually happens that the construction from which the complex algebra arose is no longer valid if the algebra requires complex solutions. The technique to capture data in the complex cases of the real problems is to **complexify** the problem. What this means is we replace the given vector spaces with their complexifications and we extend the linear transformations of interest in the same fashion. It turns out that solutions to the complexification of the problem reveal both the real solutions of the original problem as well as complex solutions which, while not real solutions, still yield useful data for unwrapping the general real problem. If this all seems a little vague, don't worry, we will get into all the messy details for the eigenvector problem.

Definition 11.3.1.

If $T: V \to V$ is a linear transformation over \mathbb{R} then the **complexification of** T is the natural extension of T to $T_{\mathbb{C}}: V_{\mathbb{C}} \to V_{\mathbb{C}}$ where $V_{\mathbb{C}} = V \oplus iV$ given by:

$$T_{\mathbb{C}}(x+iy) = T(x) + iT(y)$$

for all $x + iy \in V_{\mathbb{C}}$. If $v \in V_{\mathbb{C}}$ is a nonzero vector and $\lambda \in \mathbb{C}$ for which $T_{\mathbb{C}}(v) = \lambda v$ then we say v is a **complex eigenvector with eigenvalue** λ for T.

Example 11.3.2. Consider T = D where D = d/dx. If $\lambda = \alpha + i\beta$ then $e^{\lambda x} = e^{\alpha x}(\cos(\beta x) + i\sin(\beta x))$ by definition of the complex exponential. It is first semester calculus to show $D_{\mathbb{C}}(e^{\lambda x}) = \lambda e^{\lambda x}$. Thus $e^{\lambda x}$ is a complex e-vector of $T_{\mathbb{C}}$ with complex e-value λ . In other words, $e^{\lambda x}$ for complex λ are complex eigenfunctions of the differentiation operator.

Suppose $\beta = \{f_1, \ldots, f_n\}$ is a basis for V; $\operatorname{span}_{\mathbb{R}}(\beta) = V$. On the other hand, β also serves as a complex basis for $V_{\mathbb{C}}$, $\operatorname{span}_{\mathbb{C}}(\beta) = V_{\mathbb{C}}$. It follows that the matrix of $T_{\mathbb{C}}$ with respect to β over \mathbb{C} is the same as the matrix of T with respect to β over \mathbb{R} . In particular:

$$[T_{\mathbb{C}}(f_i)]_{\beta} = [T(f_i)]_{\beta}.$$

Suppose v is a complex e-vector with e-value λ then note $T_{\mathbb{C}}(v) = \lambda v$ implies $[T_{\mathbb{C}}]_{\beta,\beta}[v]_{\beta} = \lambda[v]_{\beta}$ where $[v]_{\beta} \in \mathbb{C}^n$. However, $[T_{\mathbb{C}}]_{\beta,\beta} = [T]_{\beta,\beta}$. Conversely, if $[T]_{\beta,\beta}$ viewed as a matrix in $\mathbb{C}^{n \times n}$ has complex e-vector w with e-value λ then $v = \Phi_{\beta}^{-1}(w)$ is a complex e-vector for $T_{\mathbb{C}}$ with e-value λ . My point is simply this: we can exchange the problem of complex e-vectors of T for the associated problem of finding complex e-vectors of $[T]_{\beta,\beta}$. Just as we found in the case of real e-vectors it suffices to study the matrix problem.

Definition 11.3.3.

Let $A \in \mathbb{C}^{n \times n}$. If $v \in \mathbb{C}^n$ is nonzero and $Av = \lambda v$ for some $\lambda \in \mathbb{C}$ then we say v is a **complex eigenvector** with **complex eigenvalue** λ of the matrix A.

The proposition below is simply the complex analog of Proposition 11.2.7.

Proposition 11.3.4.

Let $A \in \mathbb{C}^{n \times n}$ then $\lambda \in \mathbb{C}$ is an eigenvalue of A iff $det(A - \lambda I) = 0$. We say $P(\lambda) = det(A - \lambda I)$ the characteristic polynomial and $det(A - \lambda I) = 0$ is the characteristic equation.

Proof: the argument given for the real case works here also. \Box

The complex case is different than the real case for one main reason: the complex numbers are an algebraically closed field. In particular we have the Fundamental Theorem of Algebra⁸

 $^{^{8}}$ sometimes this is stated as "there exists at least one complex solution to an *n*-th order complex polynomial equation" then the factor theorem repeated applied leads to the theorem I quote here.

Theorem 11.3.5.

Fundamental Theorem of Algebra: if P(x) is an *n*-th order polynomial complex coefficients then the equation P(x) = 0 has *n*-solutions where some of the solutions may be repeated. Moreover, if P(x) is an *n*-th order polynomial with real coefficients then complex solutions to P(x) = 0 come in conjugate pairs. It follows that any polynomial with real coefficients can be factored into a unique product of repeated real and irreducible quadratic factors.

A proof of this theorem would take us far of topic here⁹. I state it here to remind us of the possibilities for solutions of the characteristic equation $P(\lambda) = \det(A - \lambda I) = 0$ which is simply an *n*-th order polynomial equation in λ .

Proposition 11.3.6.

If $A \in \mathbb{C}^{n \times n}$ then A has n eigenvalues, however, some may be repeated and/or complex. If $A \in \mathbb{R}^{n \times n}$ then complex eigenvalues arise in conjugate pairs.

Proof: observe $P(\lambda) = \det(A - \lambda I) = 0$ is an *n*-th order polynomial equation in λ . In the case $A \in \mathbb{R}^{n \times n}$ we also have $P(\lambda)$ is a polynomial with real coefficients. The proposition then follows from Theorem 11.3.5 \Box

It is interesting to contrast the proposition above with Proposition 11.2.9. On the other hand, Propositions 11.2.11, 11.2.12, 11.2.13, 11.2.14, 11.2.15, 11.2.16, 11.2.17, and 11.2.18 all naturally extend to the case of complex eigenvectors. A set of complex eigenvectors with distinct complex eigenvalues is LI as a set of complex vectors. In the case $A \in \mathbb{R}^{n \times n}$ the complex e-vectors have special structure.

Proposition 11.3.7.

If $A \in \mathbb{R}^{n \times n}$ has complex eigenvalue λ and complex eigenvector v then λ^* is likewise a complex eigenvalue with complex eigenvector v^* for A.

Proof: We assume $Av = \lambda v$ for some $\lambda \in \mathbb{C}$ and $v \in \mathbb{C}^{n \times 1}$ with $v \neq 0$. Take the complex conjugate of $Av = \lambda v$ to find $A^*v^* = \lambda^*v^*$. But, $A \in \mathbb{R}^{n \times n}$ thus $A^* = A$ and we find $Av^* = \lambda^*v^*$. Moreover, if $v \neq 0$ then $v^* \neq 0$. Therefore, v^* is an e-vector with e-value λ^* . \Box

This is a useful proposition. It means that once we calculate one complex e-vectors we almost automatically get a second e-vector merely by taking the complex conjugate.

Proposition 11.3.8.

If $A \in \mathbb{R}^{m \times n}$ has complex e-value $\lambda = \alpha + i\beta$ such that $\beta \neq 0$ and e-vector $v = a + ib \in \mathbb{C}^{n \times 1}$ such that $a, b \in \mathbb{R}^n$ then $\lambda^* = \alpha - i\beta$ is a complex e-value with e-vector $v^* = a - ib$ and $\{v, v^*\}$ is a linearly independent set of vectors over \mathbb{C} .

Proof: Proposition 11.3.7 showed that v^* is an e-vector with e-value $\lambda^* = \alpha - i\beta$. Notice that $\lambda \neq \lambda^*$ since $\beta \neq 0$. Therefore, v and v^* are e-vectors with distinct e-values. Note that Proposition 11.2.18 is equally valid for complex e-values and e-vectors. Hence, $\{v, v^*\}$ is linearly independent since these are complex e-vectors with distinct complex e-values. \Box

⁹there is a nice proof which can be given in our complex variables course

Proposition 11.3.9.

If $A \in \mathbb{R}^{m \times n}$ has complex e-value $\lambda = \alpha + i\beta$ such that $\beta \neq 0$ and e-vector $v = a + ib \in \mathbb{C}^{n \times 1}$ such that $a, b \in \mathbb{R}^n$ then $a \neq 0$ and $b \neq 0$.

Proof: Expand $Av = \lambda v$ into the real components,

$$\lambda v = (\alpha + i\beta)(a + ib) = \alpha a - \beta b + i(\beta a + \alpha b)$$

and

$$Av = A(a+ib) = Aa+iAb$$

Equating real and imaginary components yields two real matrix equations,

$$Aa = \alpha a - \beta b$$
 and $Ab = \beta a + \alpha b$

Suppose a = 0 towards a contradiction, note that $0 = -\beta b$ but then b = 0 since $\beta \neq 0$ thus v = 0 + i0 = 0 but this contradicts v being an e-vector. Likewise if b = 0 we find $\beta a = 0$ which implies a = 0 and again v = 0 which contradicts v being an e-vector. Therefore, $a, b \neq 0$. \Box

Let T be a linear transformation on a \mathbb{R}^2 such that v = a + ib is a complex eigenvector with $\lambda = \alpha + i\beta$. The calculations above make it clear that if we set $\gamma = \{a, b\}$ then

$$[T]_{\gamma,\gamma} = \left[\ [T(a)]_{\gamma} \mid [T(b)]_{\gamma} \ \right] = \left[\begin{array}{cc} \alpha & \beta \\ -\beta & \alpha \end{array} \right]$$

Of course, to be careful, we should prove $\{a, b\}$ is a LI before are certain γ is a basis.

Proposition 11.3.10.

If $A \in \mathbb{R}^{n \times n}$ and $\lambda = \alpha + i\beta \in \mathbb{C}$ with $\alpha, \beta \in \mathbb{R}$ and $\beta \neq 0$ is an e-value with e-vector $v = a + ib \in \mathbb{C}^{n \times 1}$ and $a, b \in \mathbb{R}^n$ then $\{a, b\}$ is a linearly independent set of real vectors.

Proof: Add and subtract the equations v = a + ib and $v^* = a - ib$ to deduce

$$a = \frac{1}{2}(v + v^*)$$
 and $b = \frac{1}{2i}(v - v^*)$

Let $c_1, c_2 \in \mathbb{R}$ then consider,

$$c_1 a + c_2 b = 0 \quad \Rightarrow \quad c_1[\frac{1}{2}(v + v^*)] + c_2[\frac{1}{2i}(v - v^*)] = 0$$
$$\Rightarrow \quad [c_1 - ic_2]v + [c_1 + ic_2]v^* = 0$$

But, $\{v, v^*\}$ is linearly independent hence $c_1 - ic_2 = 0$ and $c_1 + ic_2 = 0$. Adding these equations gives $2c_1 = 0$. Subtracting yields $2ic_2 = 0$. Thus $c_1 = c_2 = 0$ and we conclude $\{a, b\}$ is linearly independent set of real vectors. \Box

11.4 examples of real and complex eigenvectors

And now, the examples! Note, we should see all the propositions exhibited in these examples.

11.4.1 characteristic equations

Example 11.4.1. Let $A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$. Find the eigenvalues of A from the characteristic equation:

$$det(A - \lambda I) = det \begin{bmatrix} 3 - \lambda & 0\\ 8 & -1 - \lambda \end{bmatrix} = (3 - \lambda)(-1 - \lambda) = (\lambda + 1)(\lambda - 3) = 0$$

Hence the eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = 3$. Notice this is precisely the factor of 3 we saw scaling the vector in the first example of this chapter.

Example 11.4.2. Let $A = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$. Find the eigenvalues of A from the characteristic equation:

$$det(A - \lambda I) = det \begin{bmatrix} \frac{1}{2} - \lambda & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} - \lambda \end{bmatrix} = (\frac{1}{2} - \lambda)^2 + \frac{3}{4} = (\lambda - \frac{1}{2})^2 + \frac{3}{4} = 0$$

Well how convenient is that? The determinant completed the square for us. We find: $\lambda = \frac{1}{2} \pm i \frac{\sqrt{3}}{2}$. It would seem that elliptical orbits somehow arise from complex eigenvalues

Proposition 8.4.3 proved that taking the determinant of a triagular matrix was easy. We just multiply the diagonal entries together. This has interesting application in our discussion of eigenvalues.

Example 11.4.3. Given A below, find the eigenvalues. Use Proposition 8.4.3 to calculate the determinant,

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 0 & 5 & 6 \\ 0 & 0 & 7 \end{bmatrix} \implies det(A - \lambda I) = \begin{bmatrix} 2 - \lambda & 3 & 4 \\ 0 & 5 - \lambda & 6 \\ 0 & 0 & 7 - \lambda \end{bmatrix} = (2 - \lambda)(5 - \lambda)(7 - \lambda)$$

Therefore, $\lambda_1 = 2, \lambda_2 = 5$ and $\lambda_3 = 7$.

Remark 11.4.4. eigenwarning

Calculation of eigenvalues does not need to be difficult. However, I urge you to be careful in solving the characteristic equation. More often than not I see students make a mistake in calculating the eigenvalues. If you do that wrong then the eigenvector calculations will often turn into inconsistent equations. This should be a clue that the eigenvalues were wrong, but often I see what I like to call the "principle of minimal calculation" take over and students just adhoc set things to zero, hoping against all logic that I will somehow not notice this. Don't be this student. If the eigenvalues are correct, the eigenvector equations are consistent and you will be able to find nonzero eigenvectors. And don't forget, the eigenvectors must be nonzero.
11.4.2 real eigenvector examples

Example 11.4.5. Let
$$A = \begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix}$$
 find the e-values and e-vectors of A .

$$det(A - \lambda I) = det \begin{bmatrix} 3 - \lambda & 1 \\ 3 & 1 - \lambda \end{bmatrix} = (3 - \lambda)(1 - \lambda) - 3 = \lambda^2 - 4\lambda = \lambda(\lambda - 4) = 0$$

We find $\lambda_1 = 0$ and $\lambda_2 = 4$. Now find the e-vector with e-value $\lambda_1 = 0$, let $u_1 = [u, v]^T$ denote the e-vector we wish to find. Calculate,

$$(A - 0I)u_1 = \begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 3u + v \\ 3u + v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Obviously the equations above are redundant and we have infinitely many solutions of the form 3u + v = 0 which means v = -3u so we can write, $u_1 = \begin{bmatrix} u \\ -3u \end{bmatrix} = u \begin{bmatrix} 1 \\ -3 \end{bmatrix}$. In applications we often make a choice to select a particular e-vector. Most modern graphing calculators can calculate e-vectors. It is customary for the e-vectors to be chosen to have length one. That is a useful choice for certain applications as we will later discuss. If you use a calculator it would likely give $u_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ although the $\sqrt{10}$ would likely be approximated unless your calculator is smart.

Continuing we wish to find eigenvectors $u_2 = [u, v]^T$ such that $(A - 4I)u_2 = 0$. Notice that u, v are disposable variables in this context, I do not mean to connect the formulas from the $\lambda = 0$ case with the case considered now.

$$(A-4I)u_1 = \begin{bmatrix} -1 & 1 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -u+v \\ 3u-3v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Again the equations are redundant and we have infinitely many solutions of the form v = u. Hence, $u_2 = \begin{bmatrix} u \\ u \end{bmatrix} = u \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector for any $u \in \mathbb{R}$ such that $u \neq 0$.

Remark 11.4.6.

It was obvious the equations were redundant in the example above. However, we need not rely on pure intuition. The problem of calculating all the e-vectors is precisely the same as finding all the vectors in the null space of a matrix. We already have a method to do that without ambiguity. We find the rref of the matrix and the general solution falls naturally from that matrix. I don't bother with the full-blown theory for simple examples because there is no need. However, with 3×3 examples it may be advantageous to keep our earlier null space calculational scheme in mind.

Example 11.4.7. Let $A = \begin{bmatrix} 0 & 0 & -4 \\ 2 & 4 & 2 \\ 2 & 0 & 6 \end{bmatrix}$ find the e-values and e-vectors of A. $0 = det(A - \lambda I) = det \begin{bmatrix} -\lambda & 0 & -4 \\ 2 & 4 - \lambda & 2 \\ 2 & 0 & 6 - \lambda \end{bmatrix}$ $= (4 - \lambda) [-\lambda(6 - \lambda) + 8]$ $= (4 - \lambda) [\lambda^2 - 6\lambda + 8]$ $= -(\lambda - 4)(\lambda - 4)(\lambda - 2)$ Thus we have a repeated e-value of $\lambda_1 = \lambda_2 = 4$ and $\lambda_3 = 2$. Let's find the eigenvector $u_3 = [u, v, w]^T$ such that $(A - 2I)u_3 = 0$, we find the general solution by row reduction

$$rref\begin{bmatrix} -2 & 0 & -4 & 0\\ 2 & 2 & 2 & 0\\ 2 & 0 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 & 0\\ 0 & 1 & -1 & 0\\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{array}{c} u + 2w = 0\\ v - w = 0 \end{array} \Rightarrow \begin{array}{c} u_3 = w \begin{bmatrix} -2\\ 1\\ 1 \end{bmatrix}$$

Next find the e-vectors with e-value 4. Let $u_1 = [u, v, w]^T$ satisfy $(A - 4I)u_1 = 0$. Calculate,

$$rref \begin{bmatrix} -4 & 0 & -4 & | & 0 \\ 2 & 0 & 2 & | & 0 \\ 2 & 0 & 2 & | & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & | & 0 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \implies u + w = 0$$

Notice this case has two free variables, we can use v, w as parameters in the solution,

$$u_{1} = \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} -w \\ v \\ w \end{bmatrix} = v \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + w \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \implies u_{1} = v \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} and u_{2} = w \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

I have boxed two linearly independent eigenvectors u_1, u_2 . These vectors will be linearly independent for any pair of nonzero constants v, w.

You might wonder if it is always the case that repeated e-values get multiple e-vectors. In the preceding example the e-value 4 had *algebraic multiplicity* two and there were likewise two linearly independent e-vectors. The next example shows that is not the case.

Example 11.4.8. Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ find the e-values and e-vectors of A. $det(A - \lambda I) = det \begin{bmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{bmatrix} = (1 - \lambda)(1 - \lambda) = 0$

Hence we have a repeated e-value of $\lambda_1 = 1$. Find all e-vectors for $\lambda_1 = 1$, let $u_1 = [u, v]^T$,

$$(A-I)u_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies v = 0 \implies u_1 = u \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

We have only one e-vector for this system.

Incidentally, you might worry that we could have an e-value (in the sense of having a zero of the characteristic equation) and yet have no e-vector. Don't worry about that, we always get at least one e-vector for each distinct e-value. See Proposition 11.2.8

Example 11.4.9. Let
$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$
 find the e-values and e-vectors of A .

$$0 = det(A - \lambda I) = det \begin{bmatrix} 1 - \lambda & 2 & 3 \\ 4 & 5 - \lambda & 6 \\ 7 & 8 & 9 - \lambda \end{bmatrix}$$

$$= (1 - \lambda) [(5 - \lambda)(9 - \lambda) - 48] - 2[4(9 - \lambda) - 42] + 3[32 - 7(5 - \lambda)]$$

$$= -\lambda^3 + 15\lambda^2 + 18\lambda$$

$$= -\lambda(\lambda^2 - 15\lambda - 18)$$

Therefore, using the quadratic equation to factor the ugly part,

$$\lambda_1 = 0, \quad \lambda_2 = \frac{15 + 3\sqrt{33}}{2}, \quad \lambda_3 = \frac{15 - 3\sqrt{33}}{2}$$

The e-vector for e-value zero is not too hard to calculate. Find $u_1 = [u, v]^T$ such that $(A-0I)u_1 = 0$. This amounts to row reducing A itself:

$$rref \begin{bmatrix} 1 & 2 & 3 & | & 0 \\ 4 & 5 & 6 & | & 0 \\ 7 & 8 & 9 & | & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & | & 0 \\ 0 & 1 & 2 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \implies \begin{array}{c} u - w = 0 \\ v + 2w = 0 \end{array} \implies \begin{array}{c} u_1 = w \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

The e-vectors corresponding e-values λ_2 and λ_3 are hard to calculate without numerical help. Let's discuss Texas Instrument calculator output. To my knowledge, TI-85 and higher will calculate both e-vectors and e-values. For example, my ancient TI-89, displays the following if I define our matrix A = mat2,

$$eigVl(mat2) = \{16.11684397, -1.11684397, 1.385788954e - 13\}$$

Calculators often need a little interpretation, the third entry is really zero in disguise. The e-vectors will be displayed in the same order, they are given from the "eigVc" command in my TI-89,

$$eigVc(mat2) = \begin{bmatrix} .2319706872 & .7858302387 & .4082482905 \\ .5253220933 & .0867513393 & -.8164965809 \\ .8186734994 & -.6123275602 & .4082482905 \end{bmatrix}$$

From this we deduce that eigenvectors for λ_1, λ_2 and λ_3 are

$$u_1 = \begin{bmatrix} .2319706872 \\ .5253220933 \\ .8186734994 \end{bmatrix} \qquad u_2 = \begin{bmatrix} .7858302387 \\ .0867513393 \\ -.6123275602 \end{bmatrix} \qquad u_3 = \begin{bmatrix} .4082482905 \\ -.8164965809 \\ .4082482905 \end{bmatrix}$$

Notice that $1/\sqrt{6} \approx 0.408248905$ so you can see that u_3 above is simply the $u = 1/\sqrt{6}$ case for the family of e-vectors we calculated by hand already. The calculator chooses e-vectors so that the vectors have length one.

While we're on the topic of calculators, perhaps it is worth revisiting the example where there was only one e-vector. How will the calculator respond in that case? Can we trust the calculator?

Example 11.4.10. Recall Example 11.4.8. We let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and found a repeated e-value of $\lambda_1 = 1$ and single e-vector $u_1 = u \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Hey now, it's time for technology, let A = a, $eiaVl(a) = \{1, 1\}$ and $eiaVc(a) = \begin{bmatrix} 1 & -1 \\ 0 \end{bmatrix}$

$$eigVl(a) = \{1,1\}$$
 and $eigVc(a) = \begin{bmatrix} 1. & -1. \\ 0. & 1.e - 15 \end{bmatrix}$

Behold, the calculator has given us two alleged e-vectors. The first column is the genuine e-vector we found previously. The second column is the result of machine error. The calculator was tricked by round-off error into claiming that [-1, 0.0000000000000001] is a distinct e-vector for A. It is not. Moral of story? When using calculator you must first master the theory or else you'll stay mired in ignorance as presribed by your robot masters.

Finally, I should mention that TI-calculators may or may not deal with complex e-vectors adequately. There are doubtless many web resources for calculating e-vectors/values. I would wager if you Googled it you'd find an online calculator that beats any calculator. Many of you have a laptop with wireless so there is almost certainly a way to check your answers if you just take a minute or two. I don't mind you checking your answers. If I assign it in homework then I do want you to work it out <u>without</u> technology. Otherwise, you could get a false confidence before the test. Technology is to supplement not replace calculation.

Remark 11.4.11.

I would also remind you that there are oodles of examples beyond these lecture notes in the homework solutions from previous year(s). If these notes do not have enough examples on some topic then you should seek additional examples elsewhere, ask me, etc... Do not suffer in silence, ask for help. Thanks.

11.4.3 complex eigenvector examples

Example 11.4.12. Let $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and find the e-values and e-vectors of the matrix. Observe that $det(A - \lambda I) = \lambda^2 + 1$ hence the eigevalues are $\lambda = \pm i$. Find $u_1 = [u, v]^T$ such that $(A - iI)u_1 = 0$

$$0 = \begin{bmatrix} -i & 1 \\ -1 & -i \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -iu + v \\ -u - iv \end{bmatrix} \Rightarrow \begin{array}{c} -iu + v = 0 \\ -u - iv = 0 \end{array} \Rightarrow \begin{array}{c} v = iu \Rightarrow \\ u_1 = u \begin{bmatrix} 1 \\ i \end{bmatrix} \end{bmatrix}$$

We find infinitely many complex eigenvectors, one for each nonzero complex constant u. In applications, in may be convenient to set u = 1 so we can write, $u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + i \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Let's generalize the last example.

Example 11.4.13. Let $\theta \in \mathbb{R}$ and define $A = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$ and find the e-values and e-vectors of the matrix. Observe

$$0 = det(A - \lambda I) = det \begin{bmatrix} \cos \theta - \lambda & \sin \theta \\ -\sin \theta & \cos \theta - \lambda \end{bmatrix}$$
$$= (\cos \theta - \lambda)^2 + \sin^2 \theta$$
$$= \cos^2 \theta - 2\lambda \cos \theta + \lambda^2 + \sin^2 \theta$$
$$= \lambda^2 - 2\lambda \cos \theta + 1$$
$$= (\lambda - \cos \theta)^2 - \cos^2 \theta + 1$$
$$= (\lambda - \cos \theta)^2 + \sin^2 \theta$$

Thus $\lambda = \cos \theta \pm i \sin \theta = e^{\pm i\theta}$. Find $u_1 = [u, v]^T$ such that $(A - e^{i\theta}I)u_1 = 0$ $0 = \begin{bmatrix} -i \sin \theta & \sin \theta \\ -\sin \theta & -i \sin \theta \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies -iu \sin \theta + v \sin \theta = 0$

If $\sin \theta \neq 0$ then we divide by $\sin \theta$ to obtain v = iu hence $u_1 = [u, iu]^T = u[1, i]^T$ which is precisely what we found in the preceding example. However, if $\sin \theta = 0$ we obtain no condition what-so-ever on the matrix. That special case is not complex. Moreover, if $\sin \theta = 0$ it follows $\cos \theta = 1$ and in fact A = I in this case. The identity matrix has the repeated eigenvalue of $\lambda = 1$ and every vector in $\mathbb{R}^{2 \times 1}$ is an e-vector.

Example 11.4.14. Let
$$A = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$
 find the e-values and e-vectors of A .
$$0 = det(A - \lambda I) = \begin{bmatrix} 1 - \lambda & 1 & 0 \\ -1 & 1 - \lambda & 0 \\ 0 & 0 & 3 - \lambda \end{bmatrix}$$
$$= (3 - \lambda) [(1 - \lambda)^2 + 1]$$

Hence $\lambda_1 = 3$ and $\lambda_2 = 1 \pm i$. We have a pair of complex e-values and one real e-value. Notice that for any $n \times n$ matrix we must have at least one real e-value since all odd polynomials possess at least one zero. Let's begin with the real e-value. Find $u_1 = [u, v, w]^T$ such that $(A - 3I)u_1 = 0$:

$$rref \begin{bmatrix} -2 & 1 & 0 & | & 0 \\ -1 & -2 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \implies u_1 = w \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Next find e-vector with $\lambda_2 = 1 + i$. We wish to find $u_2 = [u, v, w]^T$ such that $(A - (1 + i)I)u_2 = 0$:

$$\begin{bmatrix} -i & 1 & 0 & | & 0 \\ -1 & -i & 0 & | & 0 \\ 0 & 0 & -1 - i & | & 0 \end{bmatrix} \xrightarrow[]{\frac{1}{-1-i}r_3 \to r_3} \begin{bmatrix} -i & 1 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \end{bmatrix}$$

One more row-swap and a rescaling of row 1 and it's clear that

$$rref\begin{bmatrix} -i & 1 & 0 & | & 0 \\ -1 & -i & 0 & | & 0 \\ 0 & 0 & -1 - i & | & 0 \end{bmatrix} = \begin{bmatrix} 1 & i & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \implies \begin{array}{c} u + iv = 0 \\ w = 0 \end{array} \implies \begin{array}{c} u_2 = v \begin{bmatrix} i \\ 1 \\ 0 \end{bmatrix}$$

I chose the free parameter to be v. Any choice of a nonzero complex constant v will yield an e-vector with e-value $\lambda_2 = 1 + i$. For future reference, it's worth noting that if we choose v = 1 then we find

$$u_2 = \begin{bmatrix} 0\\1\\0 \end{bmatrix} + i \begin{bmatrix} 1\\0\\0 \end{bmatrix}$$

We identify that $Re(u_2) = e_2$ and $Im(u_2) = e_1$

Example 11.4.15. Let $B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and let $C = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$. Define A to be the block matrix

$$A = \begin{bmatrix} B & 0 \\ \hline 0 & C \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ \hline 0 & 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

find the e-values and e-vectors of the matrix. Block matrices have nice properties: the blocks behave like numbers. Of course there is something to prove here, and I have yet to discuss block multiplication in these notes.

$$det(A - \lambda I) = det \begin{bmatrix} B - \lambda I & 0\\ 0 & C - \lambda I \end{bmatrix} = det(B - \lambda I)det(C - \lambda I)$$

Notice that both B and C are rotation matrices. B is the rotation matrix with $\theta = \pi/2$ whereas C is the rotation by $\theta = \pi/3$. We already know the e-values and e-vectors for each of the blocks if we ignore the other block. It would be nice if a block matrix allowed for analysis of each block one at a time. This turns out to be true, I can tell you without further calculation that we have e-values $\lambda_1 = \pm i$ and $\lambda_2 = \frac{1}{2} \pm i \frac{\sqrt{3}}{2}$ which have complex e-vectors

$$u_1 = \begin{bmatrix} 1\\i\\0\\0 \end{bmatrix} = e_1 + ie_2 \qquad u_2 = \begin{bmatrix} 0\\0\\1\\i \end{bmatrix} = e_3 + ie_4$$

I invite the reader to check my results through explicit calculation. Technically, this is bad form as I have yet to prove anything about block matrices. Perhaps this example gives you a sense of why we should talk about the blocks at some point.

Finally, you might wonder are there matrices which have a repeated complex e-value. And if so are there always as many complex e-vectors as there are complex e-values? The answer: sometimes. Take for instance $A = \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix}$ (where B is the same B as in the preceding example) this matrix will have a repeated e-value of $\lambda = \pm i$ and you'll be able to calculate $u_1 = e_1 \pm ie_2$ and $u_2 = e_3 \pm ie_4$ are linearly independent e-vectors for A. However, there are other matrices for which only one complex e-vector is available despite a repeat of the e-value.

Example 11.4.16. Let
$$A = \begin{bmatrix} 2 & 3 & 1 & 0 \\ -3 & 2 & 0 & 1 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & -3 & 2 \end{bmatrix}$$
 you can calculate $\lambda = 2 \pm 3i$ is repeated and yet

there are only two LI complex eigenvectors for A. In particular, v = a + ib for $\lambda = 2 + 3i$ and v^* for $\lambda^* = 2 - 3i$. From this pair, or just one of the complex eigenvectors, we find just two LI real vectors: $\{a, b\}$. Naturally, if we wish to associate some basis of \mathbb{R}^4 with A then we are missing two vectors. We return to this mystery in the next section. Note:

$$A = \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The \otimes is the tensor product. Can you see how it is defined?

11.5 eigenbases and eigenspaces

If we have a basis of eigenvectors then it is called an *eigenbasis*. For a linear transformation:

Definition 11.5.1. eigenbasis for linear transformation

Let $T: V \to V$ be a linear transformation on a vector space V over \mathbb{R} . If there exists a basis $\beta = \{v_1, v_2, \ldots, v_n\}$ of V such that such that $T(v_j) = \lambda_j v_j$ for some constant $\lambda_j \in \mathbb{R}$ then we say β is an **eigenbasis** of T.

Recall, a **diagonal** matrix D is one for which $D_{ij} = 0$ for $i \neq j$. The matrix of a linear transformation with respect to an eigenbasis will be diagonal with e-values as the diagonal entries:

Proposition 11.5.2.

If $T: V \to V$ is a linear transformation and T has an eigenbasis $\beta = \{f_1, \ldots, f_n\}$ where f_j is an eigenvector with eigenvalue λ_j for $j = 1, \ldots, n$ then

$[T]_{\beta,\beta} =$	$\begin{bmatrix} \lambda_1 \\ 0 \end{bmatrix}$	$\begin{array}{c} 0 \ \lambda_2 \end{array}$	· · · ·	$\begin{array}{c} 0 \\ 0 \end{array}$
	: 0	: 0	· · · ·	$\vdots \\ \lambda_n$

Proof: In general, $[T]_{\beta,\beta} = [[T(f_1)]_{\beta} | [T(f_2)]_{\beta} | \cdots | [T(f_n)]_{\beta}]$. However, as v_j is an eigenvector we have $T(v_j) = \lambda_j v_j$. Moreover, by definition of β coordinates, $[f_j]_{\beta} = e_j \in \mathbb{R}^n$ hence:

$$[T]_{\beta,\beta} = [[\lambda_1 f_1)]_{\beta} \mid [\lambda_2 f_2]_{\beta} \mid \cdots \mid [\lambda_n f_n]_{\beta}]$$
$$= [\lambda_1 e_1 \mid \lambda_2 e_2 \mid \cdots \mid \lambda_n e_n \mid.$$

Thus, $[T]_{\beta,\beta}$ is diagonal with $\lambda_1, \lambda_2, \ldots, \lambda_n$ on the diagonal as claimed. \Box

Now would be a good time to read Example 7.4.8 again. There we found the matrix of a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ is diagonal with respect to an eigenbasis. It turns out that there exist linear transformations which can not be diagonalized. However, even for those transformations, we may still be able to find a basis which partially diagonalizes the matrix. In particular, this brings us to the definition of the λ_j -eigenspace. We will soon see that the restriction of the linear transformation to this space will be diagonal.

Definition 11.5.3. eigenspace and geometric vs. algebraic multiplicity

Let $T: V \to V$ be a linear transformation. We define the set of all eigenvectors of T with eigenvalue λ_j adjoined the zero-vector is the λ_j -eigenspace denoted by W_{λ_j} . The dimension of W_{λ_j} is known as the **geometric multiplicity** of λ_j . The **algebraic multiplicity** of λ_j is the largest $m \in \mathbb{N}$ for which number of times $(\lambda - \lambda_j)^m$ appears as a factor of the characteristic polynomial.

I will provide examples once we focus on the matrix analog of the definition above. For the moment, we just have a few more theoretical items to clarify:

Proposition 11.5.4.

If $T: V \to V$ is a linear transformation and W_{λ} is an eigenspace of T then $W_{\lambda} \leq V$.

Proof: exercise for reader. \Box

Proposition 11.5.5.

If $T: V \to V$ is a linear transformation and W_{λ} is an eigenspace of T then $T|_{W_{\lambda}} = \lambda Id|_{W_{\lambda}}$. Moreover, if β is a basis for W_{λ} then $[T|_{W_{\lambda}}]_{\beta,\beta} = \lambda I$.

Proof: if $w \in W_{\lambda}$ then $T(w) = \lambda w = \lambda I d_{W_{\lambda}}(w)$ hence $T|_{W_{\lambda}} = \lambda I d_{W_{\lambda}}$. The fact that $[T|_{W_{\lambda}}]_{\beta,\beta} = \lambda I$ follows from the same argument as was given in Proposition 11.5.2. \Box

Theorem 11.5.6.

If $T: V \to V$ is a linear transformation with **distinct** real e-values $\lambda_1, \lambda_2, \ldots, \lambda_k$ with geometric multiplicities g_1, g_2, \ldots, g_k and algebraic multiplicities a_1, a_2, \ldots, a_k respective such that $a_j = g_j$ for all $j \in \mathbb{N}_k$. Then $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$ where $W_j = \{x \in V | T(x) = \lambda_j x\}$. Moreover, the matrix of T with respect to a basis $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ where β_j is basis for W_j from $j = 1, 2, \ldots, k$ is diagonal with: Diag $([T]a_j a_j) = (\lambda_1, \ldots, \lambda_1, \lambda_2, \ldots, \lambda_k = \lambda_k)$.

$$\operatorname{Diag}([T]_{\beta,\beta}) = (\underbrace{\lambda_1, \dots, \lambda_1}_{g_1}, \underbrace{\lambda_2, \dots, \lambda_2}_{g_2}, \dots, \underbrace{\lambda_k, \dots, \lambda_k}_{g_k}).$$

Proof: Suppose the presuppositions of the theorem are true. We intend to use criteria (iv.) of Theorem 7.7.11 to show V is a direct sum of the eigenspaces. Take nonzero $v_j \in W_j$ for j = 1, 2, ..., k. Consider,

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0.$$

By Proposition 11.2.18 adapated to linear transformations (exercise for reader) we find $\{v_1, v_2, \ldots, v_k\}$ LI as these are e-vectors of T with distinct e-values. Hence $c_1 = c_2 = \cdots = c_k = 0$. Therefore, criteria (iv.) of Theorem 7.7.11 is met and we find $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$. The remaining claim of the theorem is immediate upon application of Proposition 11.5.2. \Box

It is not generally true that algebraic and geometric multiplicities are equal. Moreover, we also know that some eigenvalues may be complex. Both of these facts make diagonalization of a given linear transformation an uncertain task. The proposition below will help us judge if diagonalization is an impossibility for T.

Proposition 11.5.7.

If $T: V \to V$ is a linear transformation with eigenvalue λ with algebraic multiplicity a and geometric multiplicity g then $g \leq a$.

Proof: Suppose g is the geometric multiplicity of λ . Then there exists a basis $\{v_1, v_2, \ldots, v_g\}$ for $W_{\lambda} \leq V$. Extend this to a basis $\beta = \{v_1, \ldots, v_g, v_{g+1}, \ldots, v_n\}$ for V. Observe,

$$T\left(\sum_{i=1}^{n} x_{i}v_{i}\right) = \sum_{i=1}^{g} x_{i}T(v_{i}) + \sum_{i=g+1}^{n} x_{i}T(v_{i})$$

$$= \sum_{i=1}^{g} \lambda x_{i}v_{i} + \sum_{i=g+1}^{n} x_{i}T(v_{i}).$$
(11.1)

Recall, $[T]_{\beta,\beta} = [[T(v_1)]_{\beta} | \cdots | [T(v_n)]_{\beta}]$. Our calculation above implies that first g columns are given as follows:

$$[T]_{\beta,\beta} = [\lambda e_1|\cdots|\lambda e_g|[T(v_{g+1})]_\beta|\cdots|[T(v_n)]_\beta]$$

Thus, the matrix of T with respect to basis β has the following block-structure:

$$[T]_{\beta,\beta} = \begin{bmatrix} \lambda I_g & B \\ \hline 0 & C \end{bmatrix}$$

We calculate the characteristic polynomial in x by an identity of the determinant: the determinant of an upper-block-triangular matrix is the product of the determinants of the blocks on the diagonal

$$\det([T]_{\beta,\beta} - xI) = \det(\lambda I_g - xI_g)\det(C - xI_{n-g}) = (\lambda - x)^g \det(C - xI_{n-g}).$$

Thus there are at least g factors of $(x - \lambda)$ in P(x) hence $a \ge g$. \Box

Proposition 11.2.4 implies that for each eigenvalue λ there exists at least one eigenvector $v \in \text{Null}(T - \lambda Id)$. This fact together with the proposition above shows that for each real eigenvalue λ_j of a linear transformation T we have $1 \leq g_j \leq a_j$. Moreover, it is only possible to diagonalize T when all the eigenvalues are real and the algebraic and geometric multiplicities all match¹⁰.

Everything we have discussed for linear transformations transfers to matrices. In particular, $A \in \mathbb{R}^{n \times n}$ has a given property if $L_A : \mathbb{R}^n \to \mathbb{R}^n$ has that given property. That said, let us be explicit:

Definition 11.5.8.

Let $A \in \mathbb{R}^{n \times n}$ then a basis $\{v_1, v_2, \ldots, v_n\}$ for \mathbb{R}^n is called an **eigenbasis** of A if each vector in the basis is an e-vector for A.

Example 11.5.9. We calculated in Example 11.4.7 the e-values and e-vectors of $A = \begin{bmatrix} 0 & 0 & -4 \\ 2 & 4 & 2 \\ 2 & 0 & 6 \end{bmatrix}$

were $\lambda_1 = \lambda_2 = 4$ and $\lambda_3 = 2$ with e-vectors

$$u_1 = \begin{bmatrix} 0\\1\\0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} -1\\0\\1 \end{bmatrix} \qquad u_3 = \begin{bmatrix} -2\\1\\1 \end{bmatrix}$$

Linear indpendence of u_3 from u_1, u_2 is given from the fact the e-values of u_3 and u_1, u_2 are distinct. Then is is clear that u_1 is not a multiple of u_2 thus they are linearly independent. It follows that $\{u_1, u_2, u_3\}$ form a linearly independent set of vectors in \mathbb{R}^3 , therefore $\{u_1, u_2, u_3\}$ is an eigenbasis.

 $^{^{10}}$ this is an iff claim, the careful reader will not I have not supplied the converse in these notes, I have not yet shown that if a linear transformation is diagonalizable then it has an eigenbasis. That is easier than what we have worked through and I leave it as an exercise for the reader.

Definition 11.5.10.

Let $A \in \mathbb{R}^{n \times n}$ then we call the set of all real e-vectors with real e-value λ unioned with the zero-vector the λ -eigenspace and we denote this set by W_{λ} .

Example 11.5.11. Again using Example 11.4.7 we have two eigenspaces,

$$W_4 = span\left\{ \begin{bmatrix} 0\\1\\0 \end{bmatrix}, \begin{bmatrix} -1\\0\\1 \end{bmatrix} \right\} \qquad W_2 = span\left\{ \begin{bmatrix} -2\\1\\1 \end{bmatrix} \right\}$$

In the example below we study how the eigenspaces of similar matrices compare. We already anticipate the result as we know that similar matrices are just different pictures of a given linear transformation. Furthermore, the eigenvalues as well as algebraic and geometric multiplicities are invariants of the underlying linear transformation. All pictures of the linear transformation must share these same traits¹¹.

Example 11.5.12. Consider the matrix

$$B = \left[\begin{array}{rrr} 4 & 2 & 2 \\ 0 & 0 & -4 \\ 0 & 2 & 6 \end{array} \right]$$

You can calculate the characteristic polynomial for B is $P_B(\lambda) = det(B - \lambda I) = (\lambda - 4)^2(\lambda - 2)$ thus we find e-values of $\lambda_1 = 4$ and $\lambda_2 = 2$. Its also easy to calculate two LI e-vectors for $\lambda_1 = 4$ namely (1,0,0) and (0,1,-1) and one e-vector (1,-2,1) with e-value $\lambda_2 = 2$. The e-spaces have the form

$$W_{\lambda_1}^B = span\{(1,0,0), (0,1,-1)\} \qquad W_{\lambda_2}^B = span\{(1,-2,1)\}$$

Clearly $dim W^B_{\lambda_1} = 2$ and $dim W^B_{\lambda_2} = 1$.

Perhaps these seem a bit familar. Recall from Example 11.4.7 that the matrix

$$A = \begin{bmatrix} 0 & 0 & -4 \\ 2 & 4 & 2 \\ 2 & 0 & 6 \end{bmatrix}$$

also had e-values $\lambda_1 = 4$ and $\lambda_2 = 2$. However, the e-spaces have the form

$$W^A_{\lambda_1} = span\{(0,1,0), (-1,0,1)\} \qquad W^A_{\lambda_2} = span\{(-2,1,1)\}$$

I constructed B by performing a similarity transformation by $P = E_{1\leftrightarrow 2}$ so it is in fact true that $B \sim A$. Therefore, we can take the following view of this example: the matrix A defines a linear operator $T : \mathbb{R}^n \to \mathbb{R}^n$ by T(v) = Av. The e-values of T are $\lambda_1 = 4$ and $\lambda_2 = 2$ and the dimensions of the corresponding e-spaces are 2 and 1 respective. If we calculate the e-spaces $W^B_{\lambda_1}, W^B_{\lambda_2}$ for $[T]_{\beta,\beta} = B$ with respect to a nonstandard basis β then the e-spaces will not be the same subspaces of \mathbb{R}^3 as $W^A_{\lambda_1}, W^A_{\lambda_2}$. However, $\dim W^B_{\lambda_1} = \dim W^A_{\lambda_1}$ and $\dim W^B_{\lambda_2} = \dim W^A_{\lambda_2}$.

¹¹the proof of the eigenvalues and algebraic multiplicities being invariant in different representations of a given T is implicit within the proof of Proposition 11.2.5. The claim that T and $[T]_{\beta,\beta}$ share the same geometric multiplicity is left as an exercise for the reader.

Definition 11.5.13.

Let A be a real square matrix with real e-value λ . The dimension of W_{λ} is called the **geometric multiplicity** of λ . The number of times the λ solution is repeated in the characteristic equation's solution is called the **algebraic multiplicity** of λ .

We already know from the examples we've considered thus far that there will not always be an eigenbasis for a given matrix A. In general, here are the problems we'll face:

- 1. we could have complex e-vectors (see Example 11.4.12)
- 2. we could have less e-vectors than needed for a basis (see Example 11.4.8)

We can say case 2 is caused from the geometric multiplicity being less than the algebraic multiplicity. What can we do about this? If we want to adjoin vectors to make-up for the lack of e-vectors then how should we find them in case 2? This question is answered in the next section.

If a matrix has *n*-linearly independent e-vectors then we'll find that we can perform a similarity transformation to transform the matrix into a diagonal form. Let me briefly summarize what is required for us to have *n*-LI e-vectors. This is the natural extension of Theorem 11.5.6. A simple proof of what follows is to apply Theorem 11.5.6 to $T = L_A$ where $A \in \mathbb{R}^{n \times n}$.

Proposition 11.5.14. criteria for real diagonalizability

Suppose that $A \in \mathbb{R}^{n \times n}$ has distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ such that the characteristic polynomial factors as follows:

$$P_A(\lambda) = \pm (\lambda - \lambda_1)^{a_1} (\lambda - \lambda_2)^{a_2} \cdots (\lambda - \lambda_k)^{a_k}.$$

We identify a_1, a_2, \ldots, a_k are the **algebraic mulitplicities** of $\lambda_1, \lambda_2, \ldots, \lambda_k$ respective and $a_1 + a_2 + \cdots + a_k = n$. Furthermore, suppose we say that the *j*-th eigenspace $W_{\lambda_j} = \{x \in \mathbb{R} \mid Ax = \lambda_j x\}$ has $dim(W_{\lambda_j}) = g_j$ for $j = 1, 2, \ldots, k$. The values g_1, g_2, \ldots, g_k are called the **geometric mulitplicities** of $\lambda_1, \lambda_2, \ldots, \lambda_k$ respective. With all of the language above in mind we can state that if $a_j = g_j$ for all $j = 1, 2, \ldots, k$ then A is diagonalizable.

Another way to understand the proposition above is that it really says is that if there exists an eigenbasis for A then it is diagonalizable. Simply take the union of the basis for each eigenspace and note the LI of this union follows immediately from Proposition 11.2.18 and the fact they are bases¹². Once we have an eigenbasis we still need to prove diagonalizability follows. Since that is what is most interesting I'll restate it once more. Note in the proposition below the e-values may be repeated. Technically, I don't really need to give the proof as we could easily derive this from Theorem 11.5.6. However, I leave the proof as it illustrates an important calculational technique.

Proposition 11.5.15.

Suppose that $A \in \mathbb{R}^{n \times n}$ has e-values $\lambda_1, \lambda_2, \ldots, \lambda_n$ with linearly independent e-vectors v_1, v_2, \ldots, v_n . If we define $V = [v_1|v_2|\cdots|v_n]$ then $D = V^{-1}AV$ where D is a diagonal matrix with the eigenvalues down the diagonal: $D = [\lambda_1 e_1 | \lambda_2 e_2 | \cdots | \lambda_n e_n]$.

¹²actually there is something to show here but I leave it to the reader for now

Proof: Notice that V is invertible since we assume the e-vectors are linearly independent. Moreover, $V^{-1}V = I$ in terms of columns translates to $V^{-1}[v_1|v_2|\cdots|v_n] = [e_1|e_2|\cdots|e_n]$. From which we deduce that $V^{-1}v_j = e_j$ for all j. Also, since v_j has e-value λ_j we have $Av_j = \lambda_j v_j$. Observe,

$$V^{-1}AV = V^{-1}A[v_{1}|v_{2}|\cdots|v_{n}]$$

= $V^{-1}[Av_{1}|Av_{2}|\cdots|Av_{n}]$
= $V^{-1}[\lambda_{1}v_{1}|\lambda_{2}v_{2}|\cdots|\lambda_{n}v_{n}]$
= $V^{-1}[\lambda_{1}v_{1}|\lambda_{2}v_{2}|\cdots|\lambda_{n}v_{n}]$
= $[\lambda_{1}V^{-1}v_{1}|\lambda_{2}V^{-1}v_{2}|\cdots|\lambda_{n}V^{-1}v_{n}]$
= $[\lambda_{1}e_{1}|\lambda_{2}e_{2}|\cdots|\lambda_{n}e_{n}]$

Example 11.5.16. Revisit Example 11.4.5 where we learned $A = \begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix}$ had e-values $\lambda_1 = 0$ and $\lambda_2 = 4$ with e-vectors: $u_1 = [1, -3]^T$ and $u_2 = [1, 1]^T$. Let's follow the advice of the proposition above and diagonalize the matrix. We need to construct $U = [u_1|u_2]$ and calculate U^{-1} , which is easy since this is a 2×2 case:

$$U = \begin{bmatrix} 1 & 1 \\ -3 & 1 \end{bmatrix} \qquad \Rightarrow \qquad U^{-1} = \frac{1}{4} \begin{bmatrix} 1 & -1 \\ 3 & 1 \end{bmatrix}$$

Now multiply,

$$U^{-1}AU = \frac{1}{4} \begin{bmatrix} 1 & -1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -3 & 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & -1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 0 & 4 \\ 0 & 4 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 0 & 0 \\ 0 & 16 \end{bmatrix}$$

Therefore, we find confirmation of the proposition, $U^{-1}AU = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$.

Notice there is one very unsettling aspect of diagonalization; we need to find the inverse of a matrix. Generally this is not pleasant. Orthogonality will offer an insight to help us here in Section 11.8.

Calculational inconvieniences aside, we have all the tools we need to diagonalize a matrix. What then is the point? Why would we care if a matrix is diagonalized? One reason is that we can calculate arbitrary powers of the matrix with a simple calculation. Note that: if $A \sim D$ then $A^k \sim D^k$. In particular: if $D = P^{-1}AP$ then $A = PDP^{-1}$ thus:

$$A^{k} = \underbrace{AA \cdots A}_{k-factors} = (PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1}) = PD^{k}P^{-1}.$$

Note, D^k is easy to calculate. Try this formula out on the last example. Try calculating A^{25} directly and then indirectly via this similarity transformation idea.

Beyond this there are applications of diagonalization too numerous to list. One reason I particularly like the text by Lay is he adds much detail on possible applications that I do not go into here. See sections 4.8, 4.9, 5.6, 5.7 for more on the applications of eigenvectors and diagonalization. Section 11.9 shows how e-vectors allow an elegant analysis of systems of differential equations and the geometry of quadratic forms. Chapters 12 give greater insight into e-vector-based analysis of quadratic forms and Chapter 13 details how generalized, possibly complex, eigenvectors derive the general solution of $\frac{dx}{dt} = Ax$ for $A \in \mathbb{R}^{n \times n}$.

11.6 generalized eigenvectors

We begin again with the definition as it applies to a linear transformation.

Definition 11.6.1.

A generalized eigenvector of order k with eigenvalue λ with respect to a linear transformation $T: V \to V$ is a nonzero vector v such that

$$(T - \lambda Id)^k v = 0 \qquad \& \qquad (T - \lambda Id)^{k-1} v \neq 0.$$

The existence of a generalized eigenvector of order k with eigenvalue λ implies the null space $\operatorname{Null}[(T - \lambda Id)^{k-1}] \neq 0$. However, if $k \geq 2$, this also implies $\operatorname{Null}[(T - \lambda Id)^{k-2}] \neq 0$. Indeed, if there exists a single generalized eigenvector of order k it follows that:

$$(T - \lambda Id)^{k-1}, \ (T - \lambda Id)^{k-2}, \ \dots, \ T - \lambda Id$$

all have nontrivial null spaces. This claim is left to the reader as an exercise. If you would like more complete exposition of this topic you can read Insel Spence and Friedberg. I am trying to get to the point without too much detail here.

Definition 11.6.2.

A k-chain with eigenvalue λ of a linear transformation $T: V \to V$ is set of k nonzero vectors v_1, v_2, \ldots, v_k such that $(T - \lambda Id)(v_j) = v_{j-1}$ for $j = 1, 2, \ldots, k$ and v_1 is an eigenvector with eigenvalue $\lambda; T - \lambda Id)(v_1) = 0$.

Of course, the reason we care about the chain is what follows:

Theorem 11.6.3.

A k-chain with e-value λ for $T: V \to V$ is a set of LI generalized e-vectors order $1, \ldots, k$.

Proof: Let $\{v_1, \ldots, v_k\}$ be a k-chain with e-value λ for T. By definition $(T - \lambda Id)(v_1) = 0$. Consider:

$$(T - \lambda Id)(v_2) = v_1 \Rightarrow (T - \lambda Id)^2(v_2) = (T - \lambda Id)(v_1) = 0$$

Thus v_2 is a generalized e-vector of order 2. Next, observe

$$(T - \lambda Id)(v_3) = v_2 \quad \Rightarrow \quad (T - \lambda Id)^3(v_3) = (T - \lambda Id)^2(v_2) = 0.$$

Thus v_3 is a generalized e-vector of order 3. We continue in this fashion until we reach the k-th vector in the chain:

$$(T - \lambda Id)(v_k) = v_{k-1} \quad \Rightarrow \quad (T - \lambda Id)^k(v_k) = (T - \lambda Id)^{k-1}(v_{k-1}) = 0.$$

Thus v_k is a generalized e-vector of order k. To prove LI of the chain suppose that:

$$c_1 v_1 + c_2 v_2 + \dots + c_k v_k = 0.$$

Operate successively by $(T - \lambda Id)^j$ for j = k - 1, k - 2, ..., 2, 1 to derive first $c_k = 0$ then $c_{k-1} = 0$ then continuing until we reach $c_2 = 0$ and finally $c_1 = 0$. \Box .

It turns out that we can always choose generalized eigenvectors such thay they line-up into chains. The details of the proof of the theorem that follow can be found in Insel Spence and Friedberg's *Linear Algebra* and most graduate linear algebra texts. They introduce an organizational tool known as **dot-diagrams** to see how to arrange the chains.

Theorem 11.6.4. Jordan basis theorem

If $T: V \to V$ is a linear transformation with real eigenvalues then there exists a basis for V formed from chains of generalized e-vectors. Such a basis is a **Jordan basis**. Moreover, up to ordering of the chains, the matrix of T is unique and is called the **Jordan form** of T

Proof: see Chapter 7 of Insel Spence and Friedberg's third¹³ ed. of *Linear Algebra*. \Box

The matrix of T with respect to a Jordan basis will be block-diagonal and each block will be a **Jordan block**. For brevity of exposition¹⁴ consider $T: V \to V$ which has a single k-chain as it a basis for $V, \beta = \{v_1, v_2, \ldots, v_k\}$ is a k-chain with e-value λ for T:

$$T(v_1) = \lambda v_1, \ T(v_2) = \lambda v_2 + v_1, \ \dots, \ T(v_k) = \lambda v_k + v_{k-1}$$

Thus the matrix of T has the form:

$$[T]_{\beta,\beta} = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

To be clear, all the diagonal entried as λ and there is a string of 1's along the superdiagonal. All other entries are zero. In some other texts, for example Hefferon, it should be noted the Jordan block has 1's right below the diagonal. This stems from a different formulation of the chains.

Perhaps you wonder why even look at chains? Of course, the Jordan basis theorem is reason enough, but another reason is that they appear somewhat naturally in differential equations. Let's examine how in a simple example.

Example 11.6.5. Consider T = D on $P_2 = span\{1, x, x^2\}$. Clearly T(1) = 0 hence $v_1 = 1$ is an eigenvector with eigenvalue $\lambda = 0$ for T. Furthermore, as T(x) = 1 and $T(x^2) = 2x$ it follows

 $[T]_{\beta,\beta} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$ Thus T has only zero as an e-value and its algebraic multiplicity is three.

If we consider $\gamma = \{\overline{1}, x, x^2/2\}$ then this is a 3-chain with e-value $\lambda = 0$. Note:

$$T(1) = 0, T(x) = 1, T(x^2/2) = x \quad \Rightarrow \quad [T]_{\gamma,\gamma} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

There are more exciting reasons attached to the study of the *matrix exponential*, see Chapter 13.

¹³maybe the fourth edition is better, some student has mine currently

¹⁴you doubt this?

It's deja vu all over again.

Definition 11.6.6.

A generalized eigenvector of order k with eigenvalue λ with respect to a $A \in \mathbb{R}^{n \times n}$ is a nonzero vector v such that

$$(A - \lambda I)^k v = 0 \qquad \& \qquad (A - \lambda I)^{k-1} v \neq 0.$$

Naturally, the chains are also of interest in the matrix case:

Definition 11.6.7.

A k-chain with eigenvalue λ of $A \in \mathbb{R}^{n \times n}$ is a set of k nonzero vectors v_1, v_2, \ldots, v_k such that $(A - \lambda I)v_j = v_{j-1}$ for $j = 1, 2, \ldots, k$ and v_1 is an eigenvector with eigenvalue λ ; $(A - \lambda I)v_1 = 0$.

The analog of Theorem 11.6.3 is true for the matrix case. However, perhaps this special case with the contradiction-based proof will add some insight for the reader.

Proposition 11.6.8.

Suppose $A \in \mathbb{R}^{n \times n}$ has e-value λ and e-vector v_1 then if $(A - \lambda I)v_2 = v_1$ has a solution then v_2 is a generalized e-vector of order 2 which is linearly independent from v_1 .

Proof: Suppose $(A - \lambda I)v_2 = v_1$ is consistent then multiply by $(A - \lambda I)$ to find $(A - \lambda I)^2 v_2 = (A - \lambda I)v_1$. However, we assumed v_1 was an e-vector hence $(A - \lambda I)v_1 = 0$ and we find v_2 is a generalized e-vector of order 2. Suppose $v_2 = kv_1$ for some nonzero k then $Av_2 = Akv_1 = k\lambda v_1 = \lambda v_2$ hence $(A - \lambda I)v_2 = 0$ but this contradicts the construction of v_2 as the solution to $(A - \lambda I)v_2 = v_1$. Consequently, v_2 is linearly independent from v_1 by virtue of its construction. \Box .

Example 11.6.9. Let's return to Example 11.4.8 and look for a generalized e-vector of order 2. Recall $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and we found a repeated e-value of $\lambda_1 = 1$ and single e-vector $u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (fix u = 1 for convenience). Let's complete the chain: find $v_2 = [u, v]^T$ such that $(A - I)u_2 = u_1$,

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \implies v = 1, u \text{ is free}$$

Any choice of u will do, in this case we can even set u = 0 to find

$$u_2 = \left[\begin{array}{c} 0\\1 \end{array} \right]$$

Clearly, $\{u_1, u_2\}$ forms a basis of $\mathbb{R}^{2 \times 1}$. It is not an eigenbasis with respect to A, however it is what is known as a Jordan basis for A.

Theorem 11.6.10.

Any matrix with real eigenvalues can be transformed to Jordan form J by a similarity transformation based on conjugation by the matrix $[\beta]$ of a Jordan basis β . That is, there exists Jordan basis β for \mathbb{R}^n such that $[\beta]^{-1}A[\beta] = J$

Proof: apply Theorem 11.6.4 to the linear transformation $T = L_A : \mathbb{R}^n \to \mathbb{R}^n$. \Box

The nicest examples are those which are already in Jordan form at the beginning:

Example 11.6.11. Suppose
$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
 it is not hard to show that $\det(A - \lambda I) = (\lambda - 1)^4 = 0$. We have a quadruple e-value $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 1$.

$$0 = (A - I)\vec{u} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \implies \vec{u} = \begin{bmatrix} s_1 \\ 0 \\ s_3 \\ 0 \end{bmatrix}$$

Any nonzero choice of s_1 or s_3 gives us an e-vector. Let's define two e-vectors which are clearly linearly independent, $\vec{u}_1 = [1, 0, 0, 0]^T$ and $\vec{u}_2 = [0, 0, 1, 0]^T$. We'll find a generalized e-vector to go with each of these. There are two length two chains to find here. In particular,

$$(A-I)\vec{u}_3 = \vec{u}_1 \quad \Rightarrow \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad s_2 = 1, s_4 = 0, \ s_1, s_3 \ free$$

I choose $s_1 = 0$ and $s_3 = 1$ since I want a new vector, define $\vec{u}_3 = [0, 0, 1, 0]^T$. Finally solving $(A - I)\vec{u}_4 = \vec{u}_2$ for $\vec{u}_4 = [s_1, s_2, s_3, s_4]^T$ yields conditions $s_4 = 1, s_2 = 0$ and s_1, s_3 free. I choose $\vec{u}_4 = [0, 0, 0, 1]^T$. To summarize we have four linearly independent vectors which form two chains:

$$(A-I)\vec{(u)}_3 = \vec{u}_1, \ (A-I)\vec{u}_1 = 0 \qquad (A-I)\vec{u}_4 = \vec{u}_2, \ (A-I)\vec{u}_2 = 0$$

The matrix above was in an example of a matrix in Jordan form. When the matrix is in Jordan form then the each elemement of then standard basis is an e-vector or generalized e-vector.

Example 11.6.12.

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}$$

Here we have the chain $\{e_1, e_2, e_3\}$ with e-value $\lambda_1 = 2$, the chain $\{e_4.e_5, e_6, e_7\}$ with e-value $\lambda_2 = 3$ and finally a lone e-vector e_8 with e-value $\lambda_3 = 4$

Usually we can find a chain of generalized e-vectors for each e-value and that will produce a Jordan basis. However, there is a trap that you will not likely get caught in for a while. It is not always possible to use a single chain for each e-value. Sometimes it takes a couple chains for a single e-value. So, to be safe, you should start with finding the highest vector in the chain then work your way down to eigenvectors. That said, we typically calculate by finding e-vectors first and working up the chain to the generalized e-vectors. I make this comment to warn you of the danger.

11.7 real Jordan form

Consider $A \in \mathbb{R}^{n \times n}$. It may not have a Jordan form. Why? Because we assumed that the matrix has only real eigenvalues in the previous section. Therefore, if we remove that restriction then we must account for the possibility of complex eigenvalues. We continue the work we began in Section 11.3 here. The theorem that follows collects the main thought for the complex case: basically, what this theorem says is that everything we did over \mathbb{R} also holds for complex vector spaces and in particular, this implies the complexification of a real linear transformation always permits a complex Jordan form.

Theorem 11.7.1.

If V is an n-dimensional real vector space and $T: V \to V$ is a linear transformation then T has n-complex e-values. Furthermore, if the geometric multiplicity of the complexification of T matches the algebraic multiplicity for each complex e-value then the complexification is diagonalizable; in particular, $T: V_{\mathbb{C}} \to V_{\mathbb{C}}$ permits a complex eigenbasis β for $V_{\mathbb{C}} = V \oplus iV$ such that $[T]_{\beta,\beta} \in \mathbb{C}^{n \times n}$ is diagonal with the complex e-values on the diagonal. If the geometric multiplicity of the complexification does not match the algebraic multiplicity for some complex eigenvalue(s) then it is possible to find a basis of generalized complex e-vectors for $V_{\mathbb{C}}$ for which the matrix of the complexified T has complex Jordan form. Furthermore, up to the ordering of the chains of complex generalized e-vectors the Jordan form of the complexification of T is unique.

Proof: the characteristic equation for the matrix of the complexification with respect to any basis is a *n*-th order complex polynomial equation hence it has *n*-complex solutions. Those are, by definition, complex e-values for T. Furthermore, the theorems about diagonalization over \mathbb{R} equally well apply to linear transformations on complex vector space and the diagonalization result follows upon transfer of the arguments for Theorem 11.5.6. Similar comments apply to the claims concerning the complex Jordan form. \Box

Diagonalization of $T: V_{\mathbb{C}} \to V_{\mathbb{C}}$ is interesting, but, we are mostly interested in what the diagonalization reveals about $T: V \to V$. The simplest case is two-dimensional.

Theorem 11.7.2.

If V is an 2-dimensional real vector space and $T: V \to V$ is a linear transformation with complex eigenvalue $\lambda = \alpha + i\beta$ where $\beta \neq 0$ with complex eigenvector $v = a + ib \in V_{\mathbb{C}}$ then the matrix of T with respect to $\gamma = \{a, b\}$ is $[T]_{\gamma, \gamma} = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$

Proof: If T has complex eigenvalue $\lambda = \alpha + i\beta$ where $\beta \neq 0$ corresponding to complex eigenvector v = a + ib for $a, b \in V$. We assume $T(v) = \lambda v$ hence:

$$T(a+ib) = (\alpha + i\beta)(a+ib)$$

thus, by definition of the complexification,

$$T(a) + iT(b) = \alpha a - \beta b + i(\beta a + \alpha b) \star$$

Then, by a modification of the arguments for Proposition 11.3.10 to the abstract context, we have that $\{a, b\}$ forms a LI set of vectors for V. Since dim(V) = 2 it follows $\gamma = \{a, b\}$ forms a basis. Moreover, from \star we obtain:

$$T(a) = \alpha a - \beta b$$
 & $T(b) = \beta a + \alpha b.$

Recall, the matrix $[T]_{\gamma,\gamma} = [[T(a)]_{\gamma} | [T(b)]_{\gamma}]$. Therefore, the theorem follows as $[T(a)]_{\gamma} = (\alpha, -\beta)$ and $[T(b)]_{\gamma} = (\beta, \alpha)$ are clear from the equations above. \Box

It might be instructive to note the complexification has a different complex matrix than the real matrix we just exhibited. The key equations are $T(v) = \lambda v$ and $T(v^*) = \lambda^* v$ thus if $\delta = \{v, v^*\}$ is a basis for $V_{\mathbb{C}} = V \oplus iV$ then the complexification $T : V_{\mathbb{C}} \to V_{\mathbb{C}}$ has matrix:

$$[T]_{\delta,\delta} = \left[\begin{array}{cc} \alpha + i\beta & 0 \\ 0 & \alpha - i\beta \end{array} \right].$$

The matrix above is complex, but it clearly contains information about the linear transformation T of the real vector space V. Next, we study a repeated complex eigenvalue where the complexification is not complex diagonalizable.

Theorem 11.7.3.

If V is an 4-dimensional real vector space and $T: V \to V$ is a linear transformation with repeated complex eigenvalue $\lambda = \alpha + i\beta$ where $\beta \neq 0$ with complex eigenvector $v_1 = a_1 + ib_1 \in V_{\mathbb{C}}$ and generalized complex eigenvector $v_2 = a_2 + ib_2$ where $(T - \lambda Id)(v_2) = v_1$ then the matrix of T with respect to $\gamma = \{a_1, b_1, a_2, b_2\}$ is $[T]_{\gamma,\gamma} = \begin{bmatrix} \alpha & \beta & 1 & 0 \\ -\beta & \alpha & 0 & 1 \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & -\beta & \alpha \end{bmatrix}$

Proof: we are given $T(v_1) = \lambda v_1$ and $T(v_2) = \lambda v_2 + v_1$. We simply need to extract real equations from this data: note $v_1 = a_1 + ib_1$ and $v_2 = a_2 + ib_2$ where $a_1, a_2, b_1, b_2 \in V$ and $\lambda = \alpha + i\beta$. Set $\gamma = \{a_1, b_1, a_2, b_2\}$. The first two columns follow from the same calculation as in the proof of Theorem 11.7.2. Calculate,

$$T(a_2 + ib_2) = (\alpha + i\beta)(a_2 + ib_2) + (a_1 + ib_1) = \alpha a_2 - \beta b_2 + a_1 i(\beta a_2 + \alpha b_2 + b_1).$$

Note $T(a_2 + ib_2) = T(a_2) + iT(b_2)$. Thus $T(a_2) = a_1 + \alpha a_2 - \beta b_2$ hence $[T(a_2)]_{\gamma} = (1, 0, \alpha, -\beta)$. Also, $T(b_2) = b_1 + \beta a_2 + \alpha b_2$ from which it follows $[T(b_2)]_{\gamma} = (0, 1, \beta, \alpha)$. The theorem follows. \Box Once more, I write the matrix of the complexification of T for the linear transformation considered above. Let $\delta = \{v_1, v_2, v_1^*, v_2^*\}$ then

$$[T]_{\delta,\delta} = \begin{bmatrix} \alpha + i\beta & 1 & 0 & 0\\ 0 & \alpha + i\beta & 0 & 0\\ \hline 0 & 0 & \alpha - i\beta & 1\\ 0 & 0 & 0 & \alpha - i\beta \end{bmatrix}$$

The next case would be a complex eigenvalue repeated three times. If $\delta = \{v_1, v_2, v_3, v_1^*, v_2^*, v_3^*\}$ where $(T - \lambda)(v_3) = v_2$, $(T - \lambda)(v_2) = v_1$ and $(T - \lambda)(v_1) = 0$. The complex Jordan matrix would have the form:

$$[T]_{\delta,\delta} = \begin{bmatrix} \lambda & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda^* & 1 & 0 \\ 0 & 0 & 0 & 0 & \lambda^* & 1 \\ 0 & 0 & 0 & 0 & 0 & \lambda^* \end{bmatrix}$$

In this case, if we use the real and imaginary components of v_1, v_2, v_3 as the basis $\gamma = \{a_1, b_1, a_2, b_2, a_3, b_3\}$ then the matrix of $T: V \to V$ will be formed as follows:

$$[T]_{\gamma,\gamma} = \begin{bmatrix} \alpha & \beta & 1 & 0 & 0 & 0 \\ -\beta & \alpha & 0 & 1 & 0 & 0 \\ 0 & 0 & \alpha & \beta & 1 & 0 \\ 0 & 0 & -\beta & \alpha & 0 & 1 \\ 0 & 0 & 0 & 0 & \alpha & \beta \\ 0 & 0 & 0 & 0 & -\beta & \alpha \end{bmatrix}$$

The proof is essentially the same as we already offered for the repeated complex eigenvalue case. In Example 11.4.16 we encountered a matrix with a repeated complex eigenvalue with geometric multiplicity of one. I observed a particular formula in terms of the tensor product. I think it warrants further comment here. In particular, we can write an analogus formula here for the 6×6 matrix above:

$$[T]_{\gamma,\gamma} = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

If T has a 4-chain of generalized complex e-vectors then we expect the pattern continues to:

$$[T]_{\gamma,\gamma} = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The term built from tensoring with the superdiagonal matrix will be **nilpotent**. Perhaps we will explore this in the exercises. Hefferon or Damiano and Little etc. has a section if you wish a second opinion on all this.

Remark 11.7.4.

I'll abstain from writing the general Jordan form of a matrix. Sufficient to say, it is block diagonal where each block is either formed as discussed thus far in this section or it is a real Jordan block. Any matrix A is similar to a unique matrix in real Jordan form up to the ordering of the blocks.

Example 11.7.5. To begin let's try an experiment using the e-vector and complex e-vectors for found in Example 11.4.14. We'll perform a similarity transformation based on this complex basis: $\beta = \{(i, 1, 0), (-i, 1, 0), (0, 0, 1)\}$. Notice that

$$[\beta] = \begin{bmatrix} i & -i & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow [\beta]^{-1} = \frac{1}{2} \begin{bmatrix} -i & 1 & 0 \\ i & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Then, we can calculate that

$$[\beta]^{-1}A[\beta] = \frac{1}{2} \begin{bmatrix} -i & 1 & 0 \\ i & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} i & -i & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1+i & 0 & 0 \\ 0 & 1-i & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Note that A is complex-diagonalizable in this case. Furthermore, A is already in real Jordan form. We should take a moment to appreciate the significance of the 2×2 complex blocks in the real

Jordan form of a matrix. It turns out there is a simple interpretation: $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Example 11.7.6. Suppose $b \neq 0$ and $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. We can calculate that $det(A - \lambda I) = (a - \lambda)^2 + b^2 = 0$ hence we have complex eigenvalues $\lambda = a \pm ib$. Denoting $r = \sqrt{a^2 + b^2}$ (the modulus of a + ib). We can work out that

$$C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = r \begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{bmatrix}$$

Therefore, a 2×2 matrix with complex eigenvalue will factor into a dilation by the modulus of the e-value $|\lambda|$ times a rotation by the argument of the eigenvalue. If we write $\lambda = rexp(i\beta)$ then we can identify that r > 0 is the modulus and β is an arugment (there is degeneracy here because angle is multiply defined).

Transforming a given matrix by a similarity transformation into real Jordan form is a generally difficult calculation. On the other hand, reading the eigenvalues as well as geometric and algebraic multiplicities is a simple matter given an explicit matrix in real Jordan form.

Example 11.7.7. Suppose $A = \begin{bmatrix} 2 & 3 & 0 & 0 \\ -3 & 2 & 0 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{bmatrix}$. I can read $\lambda_1 = 2 + 3i$ with geometric and

algebraic multiplicity one and $\lambda_2 = 5$ with geometric multiplicity one and algebraic multiplicity two. Of course, $\lambda = 2 - 3i$ is also an e-value as complex e-values come in conjugate pairs.

Example 11.7.8. Suppose
$$A = \begin{bmatrix} 0 & 3 & 1 & 0 & 0 & 0 \\ -3 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$
. I read $\lambda_1 = 3i$ with geometric multi-

plicity one and algebraic multiplicity two. Also $\lambda_2 = 5$ with geometric multiplicity and algebraic multiplicity two.

11.8 eigenvectors and orthogonality

We can apply the Gram-Schmidt process to orthogonalize the set of e-vectors. If the resulting set of orthogonal vectors is still an eigenbasis then we can prove the matrix formed from e-vectors is an orthogonal matrix.

Proposition 11.8.1.

If $A \in \mathbb{R}^{n \times n}$ has e-values $\lambda_1, \lambda_2, \ldots, \lambda_n$ with orthonormal e-vectors v_1, v_2, \ldots, v_n and if we define $V = [v_1|v_2|\cdots|v_n]$ then $V^{-1} = V^T$ and $D = V^T A V$ where D is a diagonal matrix with the eigenvalues down the diagonal: $D = [\lambda_1 e_1 | \lambda_2 e_2 | \cdots | \lambda_n e_n]$.

Proof: Orthonormality implies $v_i^T v_j = \delta_{ij}$. Observe that

$$V^{T}V = \begin{bmatrix} v_{1}^{T} \\ v_{2}^{T} \\ \vdots \\ v_{n}^{T} \end{bmatrix} \begin{bmatrix} v_{1}|v_{2}|\cdots|v_{n} \end{bmatrix} = \begin{bmatrix} v_{1}^{T}v_{1} & v_{1}^{T}v_{2} & \cdots & v_{1}^{T}v_{n} \\ v_{1}^{T}v_{1} & v_{1}^{T}v_{2} & \cdots & v_{1}^{T}v_{n} \\ \vdots & \vdots & \cdots & \vdots \\ v_{n}^{T}v_{1} & v_{n}^{T}v_{2} & \cdots & v_{n}^{T}v_{n} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Thus $V^{-1} = V^T$. The proposition follows from Proposition 11.5.15. \Box

This is great news. We now have hope of finding the diagonalization of a matrix without going to the trouble of inverting the e-vector matrix. Notice that there is no gaurantee that we can find *n*-orthonormal e-vectors. Even in the case we have *n*-linearly independent e-vectors it could happen that when we do the Gram-Schmidt process the resulting vectors are not e-vectors. That said, there is one important, and common, type of example where we are in fact gauranteed the existence of an orthonormal eigenbases for A.

Theorem 11.8.2.

A matrix $A \in \mathbb{R}^{n \times n}$ is symmetric iff there exists an orthonormal eigenbasis for A.

Proof: I'll prove the reverse implication in these notes. The forward implication is difficult and is probably best seen as a natural result in the theory of adjoints. See Chapter 6 of Insel Spence and Friedberg's third ed. or look up the section where the spectral theorem is proved in any advanced linear algebra text. Assume there exists and orthonormal eigenbasis $\{v_1, v_2, \ldots, v_n\}$ for A. Let $V = [v_1|v_2|\cdots|v_n]$ and use Proposition 11.8.1, $V^TAV = D$ where D is a diagonal matrix with the e-values down the diagonal. Clearly $D^T = D$. Transposing the equation yields $(V^TAV)^T = D$. Use the socks-shoes property for transpose to see $(V^TAV)^T = V^TA^T(V^T)^T = V^TA^TV$. We find that $V^TA^TV = V^TAV$. Multiply on the left by V and on the right by V^T and we find $A^T = A$ thus A is symmetric. \Box .

This theorem is a useful bit of trivia to know. But, be careful not to overstate the result. This theorem does not state that all diagonalizable matrices are symmetric.

Example 11.8.3. In Example 11.4.7 we found the e-values and e-vectors of $A = \begin{bmatrix} 0 & 0 & -4 \\ 2 & 4 & 2 \\ 2 & 0 & 6 \end{bmatrix}$

were $\lambda_1 = \lambda_2 = 4$ and $\lambda_3 = 2$ with e-vectors

$$u_1 = \begin{bmatrix} 0\\1\\0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} -1\\0\\1 \end{bmatrix} \qquad u_3 = \begin{bmatrix} -2\\1\\1 \end{bmatrix}$$

We argued in Example 11.5.9 that $\{u_1, u_2, u_3\}$ is an eigenbasis. In view of the Theorem above we know there is no way to perform the Gram-Schmidt process and get and orthonormal set of e-vectors for A. We could orthonormalize the basis, but it would not result in a set of e-vectors. We can be certain of this since A is not symmetric. I invite you to try Gram-Schmidt and see how the process spoils the e-values. The principle calculational observation is simply that when you add e-vectors with different e-values there is no reason to expect the sum is again an e-vector. There is an exception to my last observation, what is it?

Example 11.8.4. Let $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$. Observe that $det(A - \lambda I) = -\lambda(\lambda + 1)(\lambda - 3)$ thus $\lambda_1 = 0$.

 $0, \lambda_2 = -1, \lambda_3 = 3$. We can calculate orthonormal e-vectors of $v_1 = [1, 0, 0]^T$, $v_2 = \frac{1}{\sqrt{2}}[0, 1, -1]^T$ and $v_3 = \frac{1}{\sqrt{2}}[0, 1, 1]^T$. I invite the reader to check the validity of the following equation:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Its really neat that to find the inverse of a matrix of orthonormal e-vectors we need only take the $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$

transpose; note	$\begin{bmatrix} 1\\0\\0 \end{bmatrix}$	$\begin{array}{c} 0\\ \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} \end{array}$	$\begin{bmatrix} 0 \\ \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$	$\begin{bmatrix} 1\\0\\0 \end{bmatrix}$	$\begin{array}{c} 0\\ \frac{1}{\sqrt{2}}\\ \frac{-1}{\sqrt{2}} \end{array}$	$\begin{array}{c} 0\\ \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} \end{array}$	=	$\begin{bmatrix} 1\\0\\0 \end{bmatrix}$	$\begin{array}{c} 0 \\ 1 \\ 0 \end{array}$	$\begin{array}{c} 0\\ 0\\ 1\end{array}$.
-----------------	---	--	--	---	---	--	---	---	--	---	---

11.9 select applications

The remaining chapters in these notes give much further comment on quadratic forms and systems of differential equations. These examples are selected to show the reader what eigenvalues allow for explicit problems. To understand the general method to solve other such problems, it would be wise to read the next two chapters.

Example 11.9.1. Consider the quadric form Q(x, y) = 4xy. It's not immediately obvious (to me) what the level curves Q(x, y) = k look like. We'll make use of the preceding proposition to understand those graphs. Notice $Q(x, y) = [x, y] \begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$. Denote the matrix of the form by A and calculate the e-values/vectors:

$$det(A - \lambda I) = det \begin{bmatrix} -\lambda & 2\\ 2 & -\lambda \end{bmatrix} = \lambda^2 - 4 = (\lambda + 2)(\lambda - 2) = 0$$

Therefore, the e-values are $\lambda_1 = -2$ and $\lambda_2 = 2$.

$$(A+2I)\vec{u}_1 = \begin{bmatrix} 2 & 2\\ 2 & 2 \end{bmatrix} \begin{bmatrix} u\\ v \end{bmatrix} = \begin{bmatrix} 0\\ 0 \end{bmatrix} \quad \Rightarrow \quad \vec{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\ -1 \end{bmatrix}$$

I just solved u + v = 0 to give v = -u choose u = 1 then normalize to get the vector above. Next,

$$(A-2I)\vec{u}_2 = \begin{bmatrix} -2 & 2\\ 2 & -2 \end{bmatrix} \begin{bmatrix} u\\ v \end{bmatrix} = \begin{bmatrix} 0\\ 0 \end{bmatrix} \quad \Rightarrow \quad \vec{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\ 1 \end{bmatrix}$$

I just solved u - v = 0 to give v = u choose u = 1 then normalize to get the vector above. Let $P = [\vec{u}_1 | \vec{u}_2]$ and introduce new coordinates $\vec{y} = [\vec{x}, \vec{y}]^T$ defined by $\vec{y} = P^T \vec{x}$. Note these can be inverted by multiplication by P to give $\vec{x} = P\vec{y}$. Observe that

$$P = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \Rightarrow \begin{array}{c} x &= \frac{1}{2}(\bar{x} + \bar{y}) \\ y &= \frac{1}{2}(-\bar{x} + \bar{y}) \end{array} or \begin{array}{c} \bar{x} &= \frac{1}{2}(x - y) \\ \bar{y} &= \frac{1}{2}(x + y) \end{array}$$

The proposition preceding this example shows that substitution of the formulas above into Q yield:

$$\tilde{Q}(\bar{x},\bar{y}) = -2\bar{x}^2 + 2\bar{y}^2$$

It is clear that in the barred coordinate system the level curve Q(x, y) = k is a hyperbola. If we draw the barred coordinate system superposed over the xy-coordinate system then you'll see that the graph of Q(x, y) = 4xy = k is a hyperbola rotated by 45 degrees. The graph z = 4xy is thus a hyperbolic paraboloid:



The fascinating thing about the mathematics here is that if you don't want to graph z = Q(x, y), but you do want to know the general shape then you can determine which type of quadraic surface you're dealing with by simply calculating the eigenvalues of the form.

11.9.1 linear differential equations and e-vectors: diagonalizable case

Any system of linear differential equations with constant coefficients¹⁵ can be reformulated into a single system of linear differential equations in **normal form** $\frac{d\vec{x}}{dt} = A\vec{x} + \vec{f}$ where $A \in \mathbb{R}^{n \times n}$ and $\vec{f} : \mathbb{R} \to \mathbb{R}^n$ is a vector-valued function of a real variable which is usually called the **inhomogeneous** term. To begin suppose $\vec{f} = 0$ so the problem becomes the homogeneous system $\frac{d\vec{x}}{dt} = A\vec{x}$. We wish to find a vector-valued function $\vec{x}(t) = (x_1(t), x_2(t), \dots, x_n(t))$ such that when we differentiate it we obtain the same result as if we multiplied it by A. This is what it means to solve the differential equation $\frac{d\vec{x}}{dt} = A\vec{x}$. Essentially, solving this DEqn is like performing *n*-integrations at once. For each integration we get a constant, these constants are fixed by initial conditions if we have *n* of them. In any event, the general solution has the form:

$$\vec{x}(t) = c_1 \vec{x}_1(t) + c_2 \vec{x}_2(t) + \dots + c_n \vec{x}_n(t)$$

where $\{\vec{x}_1(t), \vec{x}_2(t), \dots, \vec{x}_n(t)\}$ is a LI set of solutions to $\frac{d\vec{x}}{dt} = A\vec{x}$ meaning $\frac{d\vec{x}_j}{dt} = A\vec{x}_j$ for each $j = 1, 2, \dots, n$. Therefore, if we can find these *n*-LI solutions then we've solved the problem. It turns out that the solutions are particularly simple if the matrix is diagonalizable: suppose $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ is an eigenbasis with e-values $\lambda_1, \lambda_2, \dots, \lambda_n$. Let $\vec{x}_j = e^{\lambda_j t}\vec{u}_j$ and observe that

$$\frac{d\vec{x}_j}{dt} = \frac{d}{dt} \left[e^{\lambda_j t} \vec{u}_j \right] = \frac{d}{dt} \left[e^{\lambda_j t} \right] \vec{u}_j = e^{\lambda_j t} \lambda_j \vec{u}_j = e^{\lambda_j t} A \vec{u}_j = A e^{\lambda_j t} \vec{u}_j = A \vec{x}_j.$$

We find that each eigenvector \vec{u}_j yields a solution $\vec{x}_j = e^{\lambda_j t} \vec{u}_j$. If there are *n*-LI e-vectors then we obtain *n*-LI solutions.

Example 11.9.2. Consider for example, the system

$$x' = x + y, \qquad y' = 3x - y$$

We can write this as the matrix problem

$$\underbrace{\begin{bmatrix} x'\\y'\end{bmatrix}}_{d\vec{x}/dt} = \underbrace{\begin{bmatrix} 1 & 1\\ 3 & -1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x\\y\end{bmatrix}}_{\vec{x}}$$

It is easily calculated that A has eigenvalue $\lambda_1 = -2$ with e-vector $\vec{u}_1 = (-1,3)$ and $\lambda_2 = 2$ with e-vectors $\vec{u}_2 = (1,1)$. The general solution of $d\vec{x}/dt = A\vec{x}$ is thus

$$\vec{x}(t) = c_1 e^{-2t} \begin{bmatrix} -1\\ 3 \end{bmatrix} + c_2 e^t \begin{bmatrix} 1\\ 1 \end{bmatrix} = \begin{bmatrix} -c_1 e^{-2t} + c_2 e^{2t}\\ 3c_1 e^{-2t} + c_2 e^{2t} \end{bmatrix}$$

So, the scalar solutions are simply $x(t) = -c_1e^{-2t} + c_2e^{2t}$ and $y(t) = 3c_1e^{-2t} + c_2e^{2t}$.

Thus far I have simply told you how to solve the system $d\vec{x}/dt = A\vec{x}$ with e-vectors, it is interesting to see what this means geometrically. For the sake of simplicity we'll continue to think about the preceding example. In it's given form the DEqn is **coupled** which means the equations for the derivatives of the dependent variables x, y cannot be solved one at a time. We have to solve both at once. In the next example I solve the same problem we just solved but this time using a change of variables approach.

¹⁵there are many other linear differential equations which are far more subtle than the ones we consider here, however, this case is of central importance to a myriad of applications

Example 11.9.3. Suppose we change variables using the diagonalization idea: introduce new variables \bar{x}, \bar{y} by $P(\bar{x}, \bar{y}) = (x, y)$ where $P = [\vec{u}_1 | \vec{u}_2]$. Note $(\bar{x}, \bar{y}) = P^{-1}(x, y)$. We can diagonalize A by the similarity transformation by P; $D = P^{-1}AP$ where Diag(D) = (-2, 2). Note that $A = PDP^{-1}$ hence $d\vec{x}/dt = A\vec{x} = PDP^{-1}\vec{x}$. Multiply both sides by P^{-1} :

$$P^{-1}\frac{d\vec{x}}{dt} = P^{-1}PDP^{-1}\vec{x} \quad \Rightarrow \quad \frac{d(P^{-1}\vec{x})}{dt} = D(P^{-1}\vec{x}).$$

You might not recognize it but the equation above is decoupled. In particular, using the notation $(\bar{x}, \bar{y}) = P^{-1}(x, y)$ we read from the matrix equation above that

$$\frac{d\bar{x}}{dt} = -2\bar{x}, \qquad \frac{d\bar{y}}{dt} = 2\bar{y}.$$

Separation of variables and a little algebra yields that $\bar{x}(t) = c_1 e^{-2t}$ and $\bar{y}(t) = c_2 e^{2t}$. Finally, to find the solution back in the original coordinate system we multiply $P^{-1}\vec{x} = (c_1 e^{-2t}, c_2 e^{2t})$ by P to isolate \vec{x} ,

$$\vec{x}(t) = \begin{bmatrix} -1 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} c_1 e^{-2t} \\ c_2 e^{2t} \end{bmatrix} = \begin{bmatrix} -c_1 e^{-2t} + c_2 e^{2t} \\ 3c_1 e^{-2t} + c_2 e^{2t} \end{bmatrix}.$$

This is the same solution we found in the last example. Usually linear algebra texts present this solution because it shows more interesting linear algebra, however, from a pragmatic viewpoint the first method is clearly faster.

Finally, we can better appreciate the solutions we found if we plot the direction field (x', y') = (x+y, 3x-y) via the "pplane" tool in Matlab. I have clicked on the plot to show a few representative trajectories (solutions):



11.9.2 linear differential equations and e-vectors: non-diagonalizable case

Generally, there does not exist an eigenbasis for the matrix in $d\vec{x}/dt = A\vec{x}$. If the e-values are all real then the remaining solutions are obtained from the matrix exponential. It turns out that X = exp(tA) is a solution matrix for $d\vec{x}/dt = A\vec{x}$ thus each column in the matrix exponential is a solution. However, direct computation of the matrix exponential is not usually tractable. Instead, an indirect approach is used. One calculates generalized e-vectors which when multiplied on exp(tA) yield very simple solutions. For example, if $(A - 3I)\vec{u}_1 = 0$ and $A - 3I)\vec{u}_2 = \vec{u}_1$ and $(A - 3I)\vec{u}_3 = \vec{u}_2$ is a chain of generalized e-vectors with e-value $\lambda = 3$ we obtain solutions to $d\vec{x}/dt = A\vec{x}$ of the form:

$$\vec{x}_1(t) = e^{3t}\vec{u}_1, \qquad \vec{x}_2(t) = e^{3t}(\vec{u}_2 + t\vec{u}_1), \qquad \vec{x}_3(t) = e^{3t}(\vec{u}_3 + t\vec{u}_2 + \frac{1}{2}t^2\vec{u}_1).$$

All these formulas stem from a simplification of $\vec{x}_j = e^{tA}\vec{u}_j$ which I call the *the magic formula*. That said, if you'd like to understand what in the world this subsection really means then you probably should read the DEqns chapter. There is one case left, if we have complex e-valued then A is not real-diagonalizable and the solutions actually have the form $\vec{x}(t) = Re(e^{tA}\vec{u})$ or $\vec{x}(t) = Im(e^{tA}\vec{u})$ where \vec{u} is either a complex e-vector or a generalized complex e-vector. Again, I leave the details for the later chapter. My point here is mostly to alert you to the fact that there are deep and interesting connections between diagonalization and the Jordan form and the solutions to corresponding differential equations.

Chapter 12

quadratic forms

Quadradic forms arise in a variety of interesting applications. From geometry to physics these particular formulas arise. When there are no cross-terms it is fairly easy to analyze the behaviour of a given form. However, the appearance of cross-terms masks the true nature of a given form. Fortunately quadratic forms permit a matrix formulation and even more fantastically the matrix is necessarily symmetric and real. It follows the matrix is orthonormally diagonalizable and the spectrum (set of eigenvalues) completely describes the given form. We study this application of eigenvectors and hopefully learn a few new things about geometry and physics in the process.

12.1 conic sections and quadric surfaces

Some of you have taken calculus III others have not, but most of you still have much to learn about level curves and surfaces. Let me give two examples to get us started:

 $x^2 + y^2 = 4$ level curve; generally has form f(x, y) = k $x^2 + 4y^2 + z^2 = 1$ level surface; generally has form F(x, y, z) = k

Alternatively, some special surfaces can be written as a graph. The top half of the ellipsoid $F(x, y, z) = x^2 + 4y^2 + z^2 = 1$ is the graph(f) where $f(x, y) = \sqrt{1 - x^2 - 4y^2}$ and $graph(f) = \{x, y, f(x, y) \mid (x, y) \in dom(f)\}$. Of course there is a great variety of examples to offer here and I only intend to touch on a few standard examples in this section. Our goal is to see what linear algebra has to say about conic sections and quadric surfaces.

12.2 quadratic forms and their matrix

Definition 12.2.1.

Generally, a **quadratic form** Q is a function $Q : \mathbb{R}^n \to \mathbb{R}$ whose formula can be written $Q(\vec{x}) = \vec{x}^T A \vec{x}$ for all $\vec{x} \in \mathbb{R}^n$ where $A \in \mathbb{R}^{n \times n}$ such that $A^T = A$.

In particular, if $\vec{x} = (x, y)$ and $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ then

$$Q(\vec{x}) = \vec{x}^T A \vec{x} = ax^2 + bxy + byx + cy^2 = ax^2 + 2bxy + y^2.$$

The n = 3 case is similar, denote $A = [A_{ij}]$ and $\vec{x} = (x, y, z)$ so that

$$Q(\vec{x}) = \vec{x}^T A \vec{x} = A_{11} x^2 + 2A_{12} xy + 2A_{13} xz + A_{22} y^2 + 2A_{23} yz + A_{33} z^2.$$

Generally, if $[A_{ij}] \in \mathbb{R}^{n \times n}$ and $\vec{x} = [x_i]^T$ then the associated quadratic form is

$$Q(\vec{x}) = \vec{x}^T A \vec{x} = \sum_{i,j} A_{ij} x_i x_j = \sum_{i=1}^n A_{ii} x_i^2 + \sum_{i < j} 2A_{ij} x_i x_j.$$

In case you wondering, yes you could write a given quadratic form with a different matrix which is not symmetric, but we will find it convenient to insist that our matrix is symmetric since that choice is always possible for a given quadratic form.

Also, you may recall (from the future) I said a **bilinear form** was a mapping from $V \times V \to \mathbb{R}$ which is linear in each slot. For example, an inner-product as defined in Definition 9.6.1 is a symmetric, positive definite bilinear form. When we discussed $\langle x, y \rangle$ we allowed $x \neq y$, in contrast a quadratic form is more like $\langle x, x \rangle$. Of course the dot-product is also an inner product and we can write a given quadratic form in terms of a dot-product:

$$\vec{x}^T A \vec{x} = \vec{x} \cdot (A \vec{x}) = (A \vec{x}) \cdot \vec{x} = \vec{x}^T A^T \vec{x}$$

Some texts actually use the middle equality above to define a symmetric matrix.

Example 12.2.2.

$$2x^{2} + 2xy + 2y^{2} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Example 12.2.3.

$$2x^{2} + 2xy + 3xz - 2y^{2} - z^{2} = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 2 & 1 & 3/2 \\ 1 & -2 & 0 \\ 3/2 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Proposition 12.2.4.

The values of a quadratic form on $\mathbb{R}^n - \{0\}$ is completely determined by it's values on the (n-1)-sphere $S_{n-1} = \{\vec{x} \in \mathbb{R}^n \mid ||\vec{x}|| = 1\}$. In particular, $Q(\vec{x}) = ||\vec{x}||^2 Q(\hat{x})$ where $\hat{x} = \frac{1}{||\vec{x}||} \vec{x}$.

Proof: Let $Q(\vec{x}) = \vec{x}^T A \vec{x}$. Notice that we can write any nonzero vector as the product of its magnitude ||x|| and its direction $\hat{x} = \frac{1}{||\vec{x}||} \vec{x}$,

$$Q(\vec{x}) = Q(||\vec{x}||\hat{x}) = (||\vec{x}||\hat{x})^T A||\vec{x}||\hat{x} = ||\vec{x}||^2 \hat{x}^T A \hat{x} = ||x||^2 Q(\hat{x}).$$

Therefore $Q(\vec{x})$ is simply proportional to $Q(\hat{x})$ with proportionality constant $||\vec{x}||^2$. \Box

The proposition above is very interesting. It says that if we know how Q works on unit-vectors then we can extrapolate its action on the remainder of \mathbb{R}^n . If $f: S \to \mathbb{R}$ then we could say f(S) > 0iff f(s) > 0 for all $s \in S$. Likewise, f(S) < 0 iff f(s) < 0 for all $s \in S$. The proposition below follows from the proposition above since $||\vec{x}||^2$ ranges over all nonzero positive real numbers in the equations above.

Proposition 12.2.5.

If Q is a quadratic form on \mathbb{R}^n and we denote $\mathbb{R}^n_* = \mathbb{R}^n - \{0\}$ 1.(negative definite) $Q(\mathbb{R}^n_*) < 0$ iff $Q(S_{n-1}) < 0$ 2.(positive definite) $Q(\mathbb{R}^n_*) > 0$ iff $Q(S_{n-1}) > 0$ 3.(non-definite) $Q(\mathbb{R}^n_*) = \mathbb{R} - \{0\}$ iff $Q(S_{n-1})$ has both positive and negative values.

Before I get too carried away with the theory let's look at a couple examples.

Example 12.2.6. Consider the quadric form $Q(x, y) = x^2 + y^2$. You can check for yourself that z = Q(x, y) is a cone and Q has positive outputs for all inputs except (0, 0). Notice that $Q(v) = ||v||^2$ so it is clear that $Q(S_1) = 1$. We find agreement with the preceding proposition. Next, think about the application of Q(x, y) to level curves; $x^2 + y^2 = k$ is simply a circle of radius \sqrt{k} or just the origin. Here's a graph of z = Q(x, y):



Notice that Q(0,0) = 0 is the absolute minimum for Q. Finally, let's take a moment to write $Q(x,y) = [x,y] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ in this case the matrix is diagonal and we note that the e-values are $\lambda_1 = \lambda_2 = 1$.

Example 12.2.7. Consider the quadric form $Q(x, y) = x^2 - 2y^2$. You can check for yourself that z = Q(x, y) is a hyperboloid and Q has non-definite outputs since sometimes the x^2 term dominates whereas other points have $-2y^2$ as the dominent term. Notice that Q(1,0) = 1 whereas Q(0,1) = -2 hence we find $Q(S_1)$ contains both positive and negative values and consequently we find agreement with the preceding proposition. Next, think about the application of Q(x, y) to level curves; $x^2 - 2y^2 = k$ yields either hyperbolas which open vertically (k > 0) or horizontally (k < 0) or a pair of lines $y = \pm \frac{x}{2}$ in the k = 0 case. Here's a graph of z = Q(x, y):



The origin is a saddle point. Finally, let's take a moment to write $Q(x, y) = [x, y] \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ in this case the matrix is diagonal and we note that the e-values are $\lambda_1 = 1$ and $\lambda_2 = -2$.

Example 12.2.8. Consider the quadric form $Q(x, y) = 3x^2$. You can check for yourself that z = Q(x, y) is parabola-shaped trough along the y-axis. In this case Q has positive outputs for all inputs except (0, y), we would call this form **positive semi-definite**. A short calculation reveals that $Q(S_1) = [0,3]$ thus we again find agreement with the preceding proposition (case 3). Next, think about the application of Q(x, y) to level curves; $3x^2 = k$ is a pair of vertical lines: $x = \pm \sqrt{k/3}$ or just the y-axis. Here's a graph of z = Q(x, y):



Finally, let's take a moment to write $Q(x,y) = [x,y] \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ in this case the matrix is diagonal and we note that the e-values are $\lambda_1 = 3$ and $\lambda_2 = 0$.

Example 12.2.9. Consider the quadric form $Q(x, y, z) = x^2 + 2y^2 + 3z^2$. Think about the application of Q(x, y, z) to level surfaces; $x^2 + 2y^2 + 3z^2 = k$ is an ellipsoid. I can't graph a function of three variables, however, we can look at level surfaces of the function. I use Mathematica to plot several below:



Finally, let's take a moment to write $Q(x, y, z) = [x, y, z] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ in this case the matrix is diagonal and we note that the e-values are $\lambda_1 = 1$ and $\lambda_2 = 2$ and $\lambda_3 = 3$.

The examples given thus far are the simplest cases. We don't really need linear algebra to understand them. In contrast, e-vectors and e-values will prove a useful tool to unravel the later examples.

Proposition 12.2.10.

If Q is a quadratic form on \mathbb{R}^n with matrix A and e-values $\lambda_1, \lambda_2, \ldots, \lambda_n$ with orthonormal e-vectors v_1, v_2, \ldots, v_n then $Q(v_i) = \lambda_i^2$ for $i = 1, 2, \ldots, n$. Moreover, if $P = [v_1|v_2|\cdots|v_n]$ then $Q(\vec{x}) = (P^T \vec{x})^T P^T A P P^T \vec{x} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$ where we defined $\vec{y} = P^T \vec{x}$.

Let me restate the proposition above in simple terms: we can transform a given quadratic form to a diagonal form by finding orthonormalized e-vectors and performing the appropriate coordinate transformation. Since P is formed from orthonormal e-vectors we know that P will be either a rotation or reflection. This proposition says we can remove "cross-terms" by transforming the quadratic forms with an appropriate rotation.

Example 12.2.11. Consider the quadric form $Q(x, y) = 2x^2 + 2xy + 2y^2$. It's not immediately obvious (to me) what the level curves Q(x, y) = k look like. We'll make use of the preceding proposition to understand those graphs. Notice $Q(x, y) = [x, y] \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$. Denote the matrix of the form by A and calculate the e-values/vectors:

$$det(A - \lambda I) = det \begin{bmatrix} 2-\lambda & 1\\ 1 & 2-\lambda \end{bmatrix} = (\lambda - 2)^2 - 1 = \lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3) = 0$$

Therefore, the e-values are $\lambda_1 = 1$ and $\lambda_2 = 3$.

$$(A-I)\vec{u}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \vec{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

I just solved u + v = 0 to give v = -u choose u = 1 then normalize to get the vector above. Next,

$$(A-3I)\vec{u}_2 = \begin{bmatrix} -1 & 1\\ 1 & -1 \end{bmatrix} \begin{bmatrix} u\\ v \end{bmatrix} = \begin{bmatrix} 0\\ 0 \end{bmatrix} \quad \Rightarrow \quad \vec{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\ 1 \end{bmatrix}$$

I just solved u - v = 0 to give v = u choose u = 1 then normalize to get the vector above. Let $P = [\vec{u}_1 | \vec{u}_2]$ and introduce new coordinates $\vec{y} = [\vec{x}, \vec{y}]^T$ defined by $\vec{y} = P^T \vec{x}$. Note these can be inverted by multiplication by P to give $\vec{x} = P\vec{y}$. Observe that

$$P = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \Rightarrow \begin{array}{c} x &= \frac{1}{2}(\bar{x} + \bar{y}) \\ y &= \frac{1}{2}(-\bar{x} + \bar{y}) \end{array} or \begin{array}{c} \bar{x} &= \frac{1}{2}(x - y) \\ \bar{y} &= \frac{1}{2}(x + y) \end{array}$$

The proposition preceding this example shows that substitution of the formulas above into Q yield¹:

$$\tilde{Q}(\bar{x},\bar{y}) = \bar{x}^2 + 3\bar{y}^2$$

It is clear that in the barred coordinate system the level curve Q(x, y) = k is an ellipse. If we draw the barred coordinate system superposed over the xy-coordinate system then you'll see that the graph of $Q(x, y) = 2x^2 + 2xy + 2y^2 = k$ is an ellipse rotated by 45 degrees. Or, if you like, we can plot z = Q(x, y):



¹technically $\tilde{Q}(\bar{x}, \bar{y})$ is $Q(x(\bar{x}, \bar{y}), y(\bar{x}, \bar{y}))$

Example 12.2.12. Consider the quadric form $Q(x, y) = x^2 + 2xy + y^2$. It's not immediately obvious (to me) what the level curves Q(x, y) = k look like. We'll make use of the preceding proposition to understand those graphs. Notice $Q(x, y) = [x, y] \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$. Denote the matrix of the form by A and calculate the e-values/vectors:

$$det(A - \lambda I) = det \begin{bmatrix} 1 - \lambda & 1\\ 1 & 1 - \lambda \end{bmatrix} = (\lambda - 1)^2 - 1 = \lambda^2 - 2\lambda = \lambda(\lambda - 2) = 0$$

Therefore, the e-values are $\lambda_1 = 0$ and $\lambda_2 = 2$.

$$(A-0)\vec{u}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \vec{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

I just solved u + v = 0 to give v = -u choose u = 1 then normalize to get the vector above. Next,

$$(A-2I)\vec{u}_2 = \begin{bmatrix} -1 & 1\\ 1 & -1 \end{bmatrix} \begin{bmatrix} u\\ v \end{bmatrix} = \begin{bmatrix} 0\\ 0 \end{bmatrix} \quad \Rightarrow \quad \vec{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\ 1 \end{bmatrix}$$

I just solved u - v = 0 to give v = u choose u = 1 then normalize to get the vector above. Let $P = [\vec{u}_1 | \vec{u}_2]$ and introduce new coordinates $\vec{y} = [\vec{x}, \vec{y}]^T$ defined by $\vec{y} = P^T \vec{x}$. Note these can be inverted by multiplication by P to give $\vec{x} = P\vec{y}$. Observe that

$$P = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \Rightarrow \begin{array}{c} x &= \frac{1}{2}(\bar{x} + \bar{y}) \\ y &= \frac{1}{2}(-\bar{x} + \bar{y}) \end{array} or \begin{array}{c} \bar{x} &= \frac{1}{2}(x - y) \\ \bar{y} &= \frac{1}{2}(x + y) \end{array}$$

The proposition preceding this example shows that substitution of the formulas above into Q yield:

$$\tilde{Q}(\bar{x},\bar{y}) = 2\bar{y}^2$$

It is clear that in the barred coordinate system the level curve Q(x,y) = k is a pair of paralell lines. If we draw the barred coordinate system superposed over the xy-coordinate system then you'll see that the graph of $Q(x,y) = x^2 + 2xy + y^2 = k$ is a line with slope -1. Indeed, with a little algebraic insight we could have anticipated this result since $Q(x,y) = (x+y)^2$ so Q(x,y) = k implies $x + y = \sqrt{k}$ thus $y = \sqrt{k} - x$. Here's a plot which again verifies what we've already found:



Example 12.2.13. Consider the quadric form Q(x, y) = 4xy. It's not immediately obvious (to me) what the level curves Q(x, y) = k look like. We'll make use of the preceding proposition to understand those graphs. Notice $Q(x, y) = [x, y] \begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$. Denote the matrix of the form by A and calculate the e-values/vectors:

$$det(A - \lambda I) = det \begin{bmatrix} -\lambda & 2\\ 2 & -\lambda \end{bmatrix} = \lambda^2 - 4 = (\lambda + 2)(\lambda - 2) = 0$$

Therefore, the e-values are $\lambda_1 = -2$ and $\lambda_2 = 2$.

$$(A+2I)\vec{u}_1 = \begin{bmatrix} 2 & 2\\ 2 & 2 \end{bmatrix} \begin{bmatrix} u\\ v \end{bmatrix} = \begin{bmatrix} 0\\ 0 \end{bmatrix} \implies \vec{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\ -1 \end{bmatrix}$$

I just solved u + v = 0 to give v = -u choose u = 1 then normalize to get the vector above. Next,

$$(A-2I)\vec{u}_2 = \begin{bmatrix} -2 & 2\\ 2 & -2 \end{bmatrix} \begin{bmatrix} u\\ v \end{bmatrix} = \begin{bmatrix} 0\\ 0 \end{bmatrix} \implies \vec{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\ 1 \end{bmatrix}$$

I just solved u - v = 0 to give v = u choose u = 1 then normalize to get the vector above. Let $P = [\vec{u}_1 | \vec{u}_2]$ and introduce new coordinates $\vec{y} = [\vec{x}, \vec{y}]^T$ defined by $\vec{y} = P^T \vec{x}$. Note these can be inverted by multiplication by P to give $\vec{x} = P\vec{y}$. Observe that

$$P = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \Rightarrow \begin{array}{c} x &= \frac{1}{2}(\bar{x} + \bar{y}) \\ y &= \frac{1}{2}(-\bar{x} + \bar{y}) \end{array} or \begin{array}{c} \bar{x} &= \frac{1}{2}(x - y) \\ \bar{y} &= \frac{1}{2}(x + y) \end{array}$$

The proposition preceding this example shows that substitution of the formulas above into Q yield:

$$\tilde{Q}(\bar{x},\bar{y}) = -2\bar{x}^2 + 2\bar{y}^2$$

It is clear that in the barred coordinate system the level curve Q(x, y) = k is a hyperbola. If we draw the barred coordinate system superposed over the xy-coordinate system then you'll see that the graph of Q(x, y) = 4xy = k is a hyperbola rotated by 45 degrees. The graph z = 4xy is thus a hyperbolic paraboloid:



The fascinating thing about the mathematics here is that if you don't want to graph z = Q(x, y), but you do want to know the general shape then you can determine which type of quadraic surface you're dealing with by simply calculating the eigenvalues of the form.

Remark 12.2.14.

I made the preceding triple of examples all involved the same rotation. This is purely for my lecturing convenience. In practice the rotation could be by all sorts of angles. In addition, you might notice that a different ordering of the e-values would result in a redefinition of the barred coordinates. 2

We ought to do at least one 3-dimensional example.

Example 12.2.15. Consider the quadric form Q defined below:

$$Q(x, y, z) = [x, y, z] \begin{bmatrix} 6 & -2 & 0 \\ -2 & 6 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Denote the matrix of the form by A and calculate the e-values/vectors:

$$det(A - \lambda I) = det \begin{bmatrix} 6 - \lambda & -2 & 0 \\ -2 & 6 - \lambda & 0 \\ 0 & 0 & 5 - \lambda \end{bmatrix}$$
$$= [(\lambda - 6)^2 - 4](5 - \lambda)$$
$$= (5 - \lambda)[\lambda^2 - 12\lambda + 32](5 - \lambda)$$
$$= (\lambda - 4)(\lambda - 8)(5 - \lambda)$$

Therefore, the e-values are $\lambda_1 = 4$, $\lambda_2 = 8$ and $\lambda_3 = 5$. After some calculation we find the following orthonormal e-vectors for A:

$$\vec{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\1\\0 \end{bmatrix} \qquad \vec{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\-1\\0 \end{bmatrix} \qquad \vec{u}_3 = \begin{bmatrix} 0\\0\\1 \end{bmatrix}$$

Let $P = [\vec{u}_1 | \vec{u}_2 | \vec{u}_3]$ and introduce new coordinates $\vec{y} = [\bar{x}, \bar{y}, \bar{z}]^T$ defined by $\vec{y} = P^T \vec{x}$. Note these can be inverted by multiplication by P to give $\vec{x} = P\vec{y}$. Observe that

$$P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0\\ -1 & 1 & 0\\ 0 & 0 & \sqrt{2} \end{bmatrix} \Rightarrow \begin{array}{ccc} x & = \frac{1}{2}(\bar{x} + \bar{y}) & \bar{x} & = \frac{1}{2}(x - y)\\ y & = \frac{1}{2}(-\bar{x} + \bar{y}) & \text{or} & \bar{y} & = \frac{1}{2}(x + y)\\ z & = \bar{z} & \bar{z} & \bar{z} & = z \end{array}$$

The proposition preceding this example shows that substitution of the formulas above into Q yield:

$$\tilde{Q}(\bar{x}, \bar{y}, \bar{z}) = 4\bar{x}^2 + 8\bar{y}^2 + 5\bar{z}^2$$

It is clear that in the barred coordinate system the level surface Q(x, y, z) = k is an ellipsoid. If we draw the barred coordinate system superposed over the xyz-coordinate system then you'll see that the graph of Q(x, y, z) = k is an ellipsoid rotated by 45 degrees around the z - axis. Plotted below are a few representative ellipsoids:



Remark 12.2.16.

If you would like to read more about conic sections or quadric surfaces and their connection to e-values/vectors I reccommend sections 9.6 and 9.7 of Anton's text. I have yet to add examples on how to include translations in the analysis. It's not much more trouble but I decided it would just be an uncessary complication this semester. Also, section 7.1,7.2 and 7.3 in Lay's text show a bit more about how to use this math to solve concrete applied problems. You might also take a look in Strang's text, his discussion of tests for positive-definite matrices is much more complete than I will give here.

12.2.1 summary of quadratic form analysis

There is a connection between the shape of level curves $Q(x_1, x_2, ..., x_n) = k$ and the graph $x_{n+1} = f(x_1, x_2, ..., x_n)$ of f. I'll discuss n = 2 but these comments equally well apply to w = f(x, y, z) or higher dimensional examples. Consider a critical point (a, b) for f(x, y) then the Taylor expansion about (a, b) has the form

$$f(a+h,b+k) = f(a,b) + Q(h,k)$$

where $Q(h,k) = \frac{1}{2}h^2 f_{xx}(a,b) + hk f_{xy}(a,b) + \frac{1}{2}h^2 f_{yy}(a,b) = [h,k][Q](h,k)$. Since $[Q]^T = [Q]$ we can find orthonormal e-vectors \vec{u}_1, \vec{u}_2 for [Q] with e-values λ_1 and λ_2 respective. Using $U = [\vec{u}_1 | \vec{u}_2]$ we can introduce rotated coordinates $(\bar{h}, \bar{k}) = U(h, k)$. These will give

$$Q(\bar{h},\bar{k}) = \lambda_1 \bar{h}^2 + \lambda_2 \bar{k}^2$$

Clearly if $\lambda_1 > 0$ and $\lambda_2 > 0$ then f(a, b) yields the local minimum whereas if $\lambda_1 < 0$ and $\lambda_2 < 0$ then f(a, b) yields the local maximum. Edwards discusses these matters on pgs. 148-153. In short, supposing $f \approx f(p) + Q$, if all the e-values of Q are positive then f has a local minimum of f(p) at p whereas if all the e-values of Q are negative then f reaches a local maximum of f(p) at p. Otherwise Q has both positive and negative e-values and we say Q is non-definite and the function has a saddle point. If all the e-values of Q are positive then Q is said to be **positive-definite** whereas if all the e-values of Q are negative then Q is said to be **negative-definite**. Edwards gives a few nice tests for ascertaining if a matrix is positive definite without explicit computation of e-values. Finally, if one of the e-values is zero then the graph will be like a trough.

Remark 12.2.17. summary of the summary.

In short, the behaviour of a quadratic form $Q(x) = x^T A x$ is governed by it's spectrum $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$. Moreover, the form can be written as $Q(y) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_k y_k^2$ by choosing the coordinate system which is built from the orthonormal eigenbasis of col(A). In this coordinate system questions of optimization become trivial (see section 7.3 of Lay for applied problems)
12.3 Taylor series for functions of two or more variables

It turns out that linear algebra and e-vectors can give us great insight into locating local extrema for a function of several variables. To summarize, we can calculate the multivariate Taylor series and we'll find that the quadratic terms correspond to a *quadratic form*. In fact, each quadratic form has a symmetric matrix representative. We know that symmetric matrices are diagonalizable hence the e-values of a symmetric matrix will be real. Moreover, the eigenvalues tell you what the min/max value of the function is at a critical point (usually). This is the n-dimensional generalization of the 2nd-derivative test from calculus. If you'd like to see further detail on these please consider taking Advanced Calculus (Math 332).

Our goal here is to find an analog for Taylor's Theorem for function from \mathbb{R}^n to \mathbb{R} . Recall that if $g: U \subseteq \mathbb{R} \to \mathbb{R}$ is *smooth* at $a \in \mathbb{R}$ then we can compute as many derivatives as we wish, moreover we can generate the Taylor's series for g centered at a:

$$g(a+h) = g(a) + g'(a)h + \frac{1}{2}g''(a)h^2 + \frac{1}{3!}g''(a)h^3 + \dots = \sum_{n=0}^{\infty} \frac{g^{(n)}(a)}{n!}h^n$$

The equation above assumes that g is analytic at a. In other words, the function actually matches it's Taylor series near a. This concept can be made rigorous by discussing the remainder. If one can show the remainder goes to zero then that proves the function is analytic. You might read pages 117-127 of Edwards *Advanced Calculus* for more on these concepts, I sometimes cover parts of that material in Advanced Calculus, Theorem 6.3 is particularly interesting.

12.3.1 deriving the two-dimensional Taylor formula

The idea is fairly simple: create a function on \mathbb{R} with which we can apply the ordinary Taylor series result. Much like our discussion of directional derivatives we compose a function of two variables with linear path in the domain. Let $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ be smooth with smooth partial derivatives of all orders. Furthermore, let $(a, b) \in U$ and construct a line through (a, b) with direction vector (h_1, h_2) as usual:

$$\phi(t) = (a, b) + t(h_1, h_2) = (a + th_1, b + th_2)$$

for $t \in \mathbb{R}$. Note $\phi(0) = (a, b)$ and $\phi'(t) = (h_1, h_2) = \phi'(0)$. Construct $g = f \circ \phi : \mathbb{R} \to \mathbb{R}$ and differentiate, note we use the chain rule for functions of several variables in what follows:

$$g'(t) = (f \circ \phi)'(t) = f'(\phi(t))\phi'(t)$$

= $\nabla f(\phi(t)) \cdot (h_1, h_2)$
= $h_1 f_x(a + th_1, b + th_2) + h_2 f_y(a + th_1, b + th_2)$

Note $g'(0) = h_1 f_x(a, b) + h_2 f_y(a, b)$. Differentiate again (I omit $(\phi(t))$ dependence in the last steps),

$$g''(t) = h_1 f'_x(a + th_1, b + th_2) + h_2 f'_y(a + th_1, b + th_2)$$

= $h_1 \nabla f_x(\phi(t)) \cdot (h_1, h_2) + h_2 \nabla f_y(\phi(t)) \cdot (h_1, h_2)$
= $h_1^2 f_{xx} + h_1 h_2 f_{yx} + h_2 h_1 f_{xy} + h_2^2 f_{yy}$
= $h_1^2 f_{xx} + 2h_1 h_2 f_{xy} + h_2^2 f_{yy}$

Thus, making explicit the point dependence, $g''(0) = h_1^2 f_{xx}(a,b) + 2h_1 h_2 f_{xy}(a,b) + h_2^2 f_{yy}(a,b)$. We may construct the Taylor series for g up to quadratic terms:

$$g(0+t) = g(0) + tg'(0) + \frac{1}{2}g''(0) + \cdots$$

= $f(a,b) + t[h_1f_x(a,b) + h_2f_y(a,b)] + \frac{t^2}{2}[h_1^2f_{xx}(a,b) + 2h_1h_2f_{xy}(a,b) + h_2^2f_{yy}(a,b)] + \cdots$

Note that $g(t) = f(a + th_1, b + th_2)$ hence $g(1) = f(a + h_1, b + h_2)$ and consequently,

$$f(a+h_1,b+h_2) = f(a,b) + h_1 f_x(a,b) + h_2 f_y(a,b) + \frac{1}{2} \left[h_1^2 f_{xx}(a,b) + 2h_1 h_2 f_{xy}(a,b) + h_2^2 f_{yy}(a,b) \right] + \cdots$$

Omitting point dependence on the 2^{nd} derivatives,

$$f(a+h_1,b+h_2) = f(a,b) + h_1 f_x(a,b) + h_2 f_y(a,b) + \frac{1}{2} \left[h_1^2 f_{xx} + 2h_1 h_2 f_{xy} + h_2^2 f_{yy} \right] + \cdots$$

Sometimes we'd rather have an expansion about (x, y). To obtain that formula simply substitute $x - a = h_1$ and $y - b = h_2$. Note that the point (a, b) is fixed in this discussion so the derivatives are not modified in this substitution,

$$f(x,y) = f(a,b) + (x-a)f_x(a,b) + (y-b)f_y(a,b) + + \frac{1}{2} \left[(x-a)^2 f_{xx}(a,b) + 2(x-a)(y-b)f_{xy}(a,b) + (y-b)^2 f_{yy}(a,b) \right] + \cdots$$

At this point we ought to recognize the first three terms give the tangent plane to z = f(z, y) at (a, b, f(a, b)). The higher order terms are nonlinear corrections to the linearization, these quadratic terms form a *quadratic form*. If we computed third, fourth or higher order terms we'd find that, using $a = a_1$ and $b = a_2$ as well as $x = x_1$ and $y = x_2$,

$$f(x,y) = \sum_{n=0}^{\infty} \sum_{i_1=0}^{n} \sum_{i_2=0}^{n} \cdots \sum_{i_n=0}^{n} \frac{1}{n!} \frac{\partial^{(n)} f(a_1, a_2)}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_n}} (x_{i_1} - a_{i_1}) (x_{i_2} - a_{i_2}) \cdots (x_{i_n} - a_{i_n})$$

The multivariate Taylor formula for a function of j-variables for j > 2 is very similar. Rather than even state the formula I will show a few examples in the subsection that follows.

12.3.2 examples

Example 12.3.1. Suppose $f(x, y) = exp(-x^2 - y^2 + 2y - 1)$ expand f about the point (0, 1):

$$f(x,y) = exp(-x^2)exp(-y^2 + 2y - 1) = exp(-x^2)exp(-(y-1)^2)$$

expanding,

$$f(x,y) = (1 - x^{2} + \dots)(1 - (y - 1)^{2} + \dots) = 1 - x^{2} - (y - 1)^{2} + \dots$$

Recenter about the point (0,1) by setting x = h and y = 1 + k so

$$f(h, 1+k) = 1 - h^2 - k^2 + \cdots$$

If (h, k) is near (0, 0) then the dominant terms are simply those we've written above hence the graph is like that of a quadraic surface with a pair of negative e-values. It follows that f(0, 1) is a local maximum. In fact, it happens to be a global maximum for this function. **Example 12.3.2.** Suppose $f(x, y) = 4 - (x - 1)^2 + (y - 2)^2 + Aexp(-(x - 1)^2 - (y - 2)^2) + 2B(x - 1)(y - 2)$ for some constants A, B. Analyze what values for A, B will make (1, 2) a local maximum, minimum or neither. Expanding about (1, 2) we set x = 1 + h and y = 2 + k in order to see clearly the local behaviour of f at (1, 2),

$$f(1+h, 2+k) = 4 - h^2 - k^2 + Aexp(-h^2 - k^2) + 2Bhk$$

= 4 - h^2 - k^2 + A(1 - h^2 - k^2) + 2Bhk \cdots
= 4 + A - (A + 1)h² + 2Bhk - (A + 1)k² + ...

There is no nonzero linear term in the expansion at (1,2) which indicates that f(1,2) = 4 + Amay be a local extremum. In this case the quadratic terms are nontrivial which means the graph of this function is well-approximated by a quadratic surface near (1,2). The quadratic form Q(h,k) = $-(A+1)h^2 + 2Bhk - (A+1)k^2$ has matrix

$$[Q] = \begin{bmatrix} -(A+1) & B \\ B & -(A+1)^2 \end{bmatrix}.$$

The characteristic equation for Q is

$$det([Q] - \lambda I) = det \begin{bmatrix} -(A+1) - \lambda & B \\ B & -(A+1)^2 - \lambda \end{bmatrix} = (\lambda + A + 1)^2 - B^2 = 0$$

We find solutions $\lambda_1 = -A - 1 + B$ and $\lambda_2 = -A - 1 - B$. The possibilities break down as follows:

- 1. if $\lambda_1, \lambda_2 > 0$ then f(1,2) is local minimum.
- 2. if $\lambda_1, \lambda_2 < 0$ then f(1, 2) is local maximum.
- 3. if just one of λ_1, λ_2 is zero then f is constant along one direction and min/max along another so technically it is a local extremum.
- 4. if $\lambda_1 \lambda_2 < 0$ then f(1,2) is not a local etremum, however it is a saddle point.

In particular, the following choices for A, B will match the choices above

- 1. Let A = -3 and B = 1 so $\lambda_1 = 3$ and $\lambda_2 = 1$;
- 2. Let A = 3 and B = 1 so $\lambda_1 = -3$ and $\lambda_2 = -5$
- 3. Let A = -3 and B = -2 so $\lambda_1 = 0$ and $\lambda_2 = 4$
- 4. Let A = 1 and B = 3 so $\lambda_1 = 1$ and $\lambda_2 = -5$

Here are the graphs of the cases above, note the analysis for case 3 is more subtle for Taylor approximations as opposed to simple quadraic surfaces. In this example, case 3 was also a local minimum. In contrast, in Example 12.2.12 the graph was like a trough. The behaviour of f away from the critical point includes higher order terms whose influence turns the trough into a local minimum.



Example 12.3.3. Suppose $f(x, y) = \sin(x)\cos(y)$ to find the Taylor series centered at (0, 0) we can simply multiply the one-dimensional result $\sin(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \cdots$ and $\cos(y) = 1 - \frac{1}{2!}y^2 + \frac{1}{4!}y^4 + \cdots$ as follows:

$$f(x,y) = (x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \cdots)(1 - \frac{1}{2!}y^2 + \frac{1}{4!}y^4 + \cdots)$$

= $x - \frac{1}{2}xy^2 + \frac{1}{24}xy^4 - \frac{1}{6}x^3 - \frac{1}{12}x^3y^2 + \cdots$
= $x + \cdots$

The origin (0,0) is a critical point since $f_x(0,0) = 0$ and $f_y(0,0) = 0$, however, this particular critical point escapes the analysis via the quadratic form term since Q = 0 in the Taylor series for this function at (0,0). This is analogous to the inconclusive case of the 2nd derivative test in calculus III.

Example 12.3.4. Suppose f(x, y, z) = xyz. Calculate the multivariate Taylor expansion about the point (1, 2, 3). I'll actually calculate this one via differentiation, I have used tricks and/or calculus II results to shortcut any differentiation in the previous examples. Calculate first derivatives

$$f_x = yz \qquad f_y = xz \qquad f_z = xy,$$

and second derivatives,

$$f_{xx} = 0 \qquad f_{xy} = z \qquad f_{xz} = y$$
$$f_{yx} = z \qquad f_{yy} = 0 \qquad f_{yz} = x$$
$$f_{zx} = y \qquad f_{zy} = x \qquad f_{zz} = 0,$$

and the nonzero third derivatives,

$$f_{xyz} = f_{yzx} = f_{zxy} = f_{zyx} = f_{yxz} = f_{xzy} = 1$$

It follows,

$$\begin{aligned} f(a+h,b+k,c+l) &= \\ &= f(a,b,c) + f_x(a,b,c)h + f_y(a,b,c)k + f_z(a,b,c)l + \\ &= \frac{1}{2}(f_{xx}hh + f_{xy}hk + f_{xz}hl + f_{yx}kh + f_{yy}kk + f_{yz}kl + f_{zx}lh + f_{zy}lk + f_{zz}ll) + \cdots \end{aligned}$$

Of course certain terms can be combined since $f_{xy} = f_{yx}$ etc... for smooth functions (we assume smooth in this section, moreover the given function here is clearly smooth). In total,

$$f(1+h,2+k,3+l) = 6 + 6h + 3k + 2l + \frac{1}{2}(3hk + 2hl + 3kh + kl + 2lh + lk) + \frac{1}{3!}(6)hkl$$

Of course, we could also obtain this from simple algebra:

$$f(1+h, 2+k, 3+l) = (1+h)(2+k)(3+l) = 6 + 6h + 3k + l + 3hk + 2hl + kl + hkl.$$

Remark 12.3.5.

One very interesting application of the orthogonal complement theorem is to the method of Lagrange multipliers. The problem is to maximize an objective function $f(x_1, x_2, \ldots, x_n)$ with respect to a set of constraint functions $g_1(x_1, x_2, \ldots, x_n) = 0$, $g_2(x_1, x_2, \ldots, x_n) = 0$ and $g_k(x_1, x_2, \ldots, x_n) = 0$. One can argue that extreme values for f must satisfy

$$\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 + \dots + \lambda_k \nabla g_k$$

for a particular set of Lagrange multipliers $\lambda_1, \lambda_2, \ldots, \lambda_k$. The crucial step in the analysis relies on the orthogonal decomposition theorem. It is the fact that forces the gradient of the objective function to reside in the span of the gradients of the constraints. See my Advanced Calculus notes, or consult many advanced calculus texts.

12.4 intertia tensor, an application of quadratic forms

We can use quadratic forms to elegantly state a number of interesting quantities in classical mechanics. For example, the translational kinetic energy of a mass m with velocity v is

$$T_{trans}(v) = \frac{m}{2}v^{T}v = [v_{1}, v_{2}, v_{3}] \begin{bmatrix} m/2 & 0 & 0\\ 0 & m/2 & 0\\ 0 & 0 & m/2 \end{bmatrix} \begin{bmatrix} v_{1}\\ v_{2}\\ v_{3} \end{bmatrix}.$$

On the other hand, the rotational kinetic energy of an object with moment of intertia I and angular velocity ω with respect to a particular axis of rotation is

$$T_{rot}(v) = \frac{I}{2}\omega^T \omega.$$

In addition you might recall that the force F applied at radial arm r gave rise to a torque of $\tau = r \times F$ which made the angular momentum $L = I\omega$ have the time-rate of change $\tau = \frac{dL}{dt}$. In the first semester of physics this is primarily all we discuss. We are usually careful to limit the discussion to rotations which happen to occur with respect to a particular axis. But, what about other rotations? What about rotations with respect to less natural axes of rotation? How should we describe the rotational physics of a rigid body which spins around some axis which doesn't happen to line up with one of the nice examples you find in an introductory physics text?

The answer is found in extending the idea of the moment of intertia to what is called the inertia tensor I_{ij} (in this section I is not the identity). To begin I'll provide a calculation which motivates the definition for the inertia tensor.

Consider a rigid mass with density $\rho = dm/dV$ which is a function of position $r = (x_1, x_2, x_3)$. Suppose the body rotates with angular velocity ω about some axis through the origin, however it is otherwise not in motion. This means all of the energy is rotational. Suppose that dm is at r then we define $v = (\dot{x}_1, \dot{x}_2, \dot{x}_3) = dr/dt$. In this context, the velocity v of dm is also given by the cross-product with the angular velocity; $v = \omega \times r$. Using the einstein repeated summation notation the k-th component of the cross-product is nicely expressed via the Levi-Civita symbol; $(\omega \times r)_k = \epsilon_{klm}\omega_l x_m$. Therefore, $v_k = \epsilon_{klm}\omega_l x_m$. The infinitesimal kinetic energy due to this little bit of rotating mass dm is hence

$$dT = \frac{dm}{2} v_k v_k$$

$$= \frac{dm}{2} (\epsilon_{klm} \omega_l x_m) (\epsilon_{kij} \omega_i x_j)$$

$$= \frac{dm}{2} \epsilon_{klm} \epsilon_{kij} \omega_l \omega_i x_m x_j$$

$$= \frac{dm}{2} (\delta_{li} \delta_{mj} - \delta_{lj} \delta_{mi}) \omega_l \omega_i x_m x_j$$

$$= \frac{dm}{2} (\delta_{li} \delta_{mj} \omega_l \omega_i x_m x_j - \delta_{lj} \delta_{mi} \omega_l \omega_i x_m x_j)$$

$$= \omega_l \frac{dm}{2} (\delta_{li} \delta_{mj} x_m x_j - \delta_{lj} \delta_{mi} x_m x_j) \omega_i$$

$$= \omega_l \left[\frac{dm}{2} (\delta_{li} ||r||^2 - x_l x_i) \right] \omega_i.$$

Integrating over the mass, if we add up all the little bits of kinetic energy we obtain the total kinetic energy for this rotating body: we replace dm with $\rho(r)dV$ and the integration is over the volume of the body,

$$T = \int \omega_l \left[\frac{1}{2} (\delta_{li} ||r||^2 - x_l x_i) \right] \omega_i \rho(r) dV$$

However, the body is rigid so the angular velocity is the same for each dm and we can pull the components of the angular velocity out of the integration³ to give:

$$T = \frac{1}{2}\omega_j \underbrace{\left[\int (\delta_{jk}||r||^2 - x_j x_k)\rho(r)dV\right]}_{I_{jk}}\omega_k$$

This integral defines the intertia tensor I_{jk} for the rotating body. Given the inertia tensor I_{lk} the kinetic energy is simply the value of the quadratic form below:

$$T(\omega) = \frac{1}{2}\omega^{T}\omega = [\omega_{1}, \omega_{2}, \omega_{3}] \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix} \begin{bmatrix} \omega_{1} \\ \omega_{2} \\ \omega_{3} \end{bmatrix}$$

The matrix above is not generally diagonal, however you can prove it is symmetric (easy). Therefore, we can find an orthonormal eigenbasis $\beta = \{u_1, u_2, u_3\}$ and if $P = [\beta]$ then it follows by orthonormality of the basis that $[I]_{\beta,\beta} = P^T[I]P$ is diagonal. The eigenvalues of the inertia tensor (the matrix $[I_{jk}]$) are called the **principle moments of inertia** and the eigenbasis $\beta = \{u_1, u_2, u_3\}$ define the **principle axes** of the body.

The study of the rotational dynamics flows from analyzing the equations:

$$L_i = I_{ij}\omega_j$$
 and $\tau_i = \frac{dL_i}{dt}$

If the initial angular velocity is in the direction of a principle axis u_1 then the motion is basically described in the same way as in the introductory physics course provided that the torque is also in the direction of u_1 . The moment of intertia is simply the first principle moment of inertia and $L = \lambda_1 \omega$. However, if the torque is not in the direction of a principle axis or the initial angular velocity is not along a principle axis then the motion is more complicated since the rotational motion is connected to more than one axis of rotation. Think about a spinning top which is spinning in place. There is wobbling and other more complicated motions that are covered by the mathematics described here.

Example 12.4.1. The intertia tensor for a cube with one corner at the origin is found to be

$$I = \frac{2}{3}Ms^2 \begin{bmatrix} 1 & -3/8 & -3/8 \\ -3/8 & 1 & -3/8 \\ -3/8 & -3/8 & 1 \end{bmatrix}$$

Introduce m = M/8 to remove the fractions,

$$I = \frac{2}{3}Ms^2 \begin{bmatrix} 8 & -3 & -3\\ -3 & 8 & -3\\ -3 & -3 & 8 \end{bmatrix}$$

³I also relabled the indices to have nicer final formula, nothing profound here

You can calculate that the e-values are $\lambda_1 = 2$ and $\lambda_2 = 11 = \lambda_3$ with principle axis in the directions

$$u_1 = \frac{1}{\sqrt{3}}(1,1,1), \ u_2 = \frac{1}{\sqrt{2}}(-1,1,0), \ u_3 = \frac{1}{\sqrt{2}}(-1,0,1).$$

The choice of u_2, u_3 is not unique. We could just as well choose any other orthonormal basis for $span\{u_2, u_3\} = W_{11}$.

Finally, a word of warning, for a particular body there may be so much symmetry that no particular eigenbasis is specified. There may be many choices of an orthonormal eigenbasis for the system. Consider a sphere. Any orthonormal basis will give a set of principle axes. Or, for a right circular cylinder the axis of the cylinder is clearly a principle axis however the other two directions are arbitrarily chosen from the plane which is the orthogonal complement of the axis. I think it's fair to say that if a body has a unique (up to ordering) set of principle axes then the shape has to be somewhat ugly. Symmetry is beauty but it implies ambiguity for the choice of certain principle axes.

Chapter 13

systems of differential equations

Systems of differential equations are found at the base of many nontrivial questions in physics, math, biology, chemistry, nuclear engineering, economics, etc... Consider this, anytime a problem is described by several quantities which depend on time and each other it is likely that a simple conservation of mass, charge, population, particle number,... force linear relations between the time-rates of change of the quantities involved. This means, we get a system of differential equations. To be specific, Newton's Second Law is a system of differential equations. Maxwell's Equations are a system of differential equations. Now, generally, the methods we discover in this chapter will not allow solutions to problems I allude to above. Many of those problems are **nonlinear**. There are researchers who spend a good part of their career just unraveling the structure of a particular partial differential equation. That said, once simplifying assumptions are made and the problem is linearlized one often faces the problem we solve in this chapter. We show how to solve **any** system of first order differential equations with constant coefficients. This is accomplished by the application of Jordan basis for the matrix of the system to the **matrix exponential**. I'm not sure the exact history of the method I show in this chapter. In my opinion, the manner in which the chains of generalized eigenvectors tame the matrix exponential are reason enough to study them.

13.1 calculus of matrices

A more apt title would be "calculus of matrix-valued functions of a real variable".

Definition 13.1.1.

A matrix-valued function of a real variable is a function from $I \subseteq \mathbb{R}$ to $\mathbb{R}^{m \times n}$. Suppose $A : I \subseteq \mathbb{R} \to \mathbb{R}^{m \times n}$ is such that $A_{ij} : I \subseteq \mathbb{R} \to \mathbb{R}$ is differentiable for each i, j then we define

$$\frac{dA}{dt} = \left[\frac{dA_{ij}}{dt}\right]$$

which can also be denoted $(A')_{ij} = A'_{ij}$. We likewise define $\int Adt = [\int A_{ij}dt]$ for A with integrable components. Definite integrals and higher derivatives are also defined component-wise.

Example 13.1.2. Suppose $A(t) = \begin{bmatrix} 2t & 3t^2 \\ 4t^3 & 5t^4 \end{bmatrix}$. I'll calculate a few items just to illustrate the definition above. calculate; to differentiate a matrix we differentiate each component one at a time:

$$A'(t) = \begin{bmatrix} 2 & 6t \\ 12t^2 & 20t^3 \end{bmatrix} \qquad A''(t) = \begin{bmatrix} 0 & 6 \\ 24t & 60t^2 \end{bmatrix} \qquad A'(0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

Integrate by integrating each component:

$$\int A(t)dt = \begin{bmatrix} t^2 + c_1 & t^3 + c_2 \\ t^4 + c_3 & t^5 + c_4 \end{bmatrix} \qquad \int_0^2 A(t)dt = \begin{bmatrix} t^2 \Big|_0^2 & t^3 \Big|_0^2 \\ \\ t^4 \Big|_0^2 & t^5 \Big|_0^2 \end{bmatrix} = \begin{bmatrix} 4 & 8 \\ 16 & 32 \end{bmatrix}$$

Proposition 13.1.3.

Suppose A, B are matrix-valued functions of a real variable, f is a function of a real variable, c is a constant, and C is a constant matrix then

1. (AB)' = A'B + AB' (product rule for matrices)

2.
$$(AC)' = A'C$$

$$3. \ (CA)' = CA'$$

$$4. \ (fA)' = f'A + fA$$

5.
$$(cA)' = cA'$$

6.
$$(A+B)' = A' + B'$$

where each of the functions is evaluated at the same time t and I assume that the functions and matrices are differentiable at that value of t and of course the matrices A, B, C are such that the multiplications are well-defined. **Proof:** Suppose $A(t) \in \mathbb{R}^{m \times n}$ and $B(t) \in \mathbb{R}^{n \times p}$ consider,

$$\begin{aligned} (AB)'_{ij} &= \frac{d}{dt} ((AB)_{ij}) & \text{defn. derivative of matrix} \\ &= \frac{d}{dt} (\sum_k A_{ik} B_{kj}) & \text{defn. of matrix multiplication} \\ &= \sum_k \frac{d}{dt} (A_{ik} B_{kj}) & \text{linearity of derivative} \\ &= \sum_k \left[\frac{dA_{ik}}{dt} B_{kj} + A_{ik} \frac{dB_{kj}}{dt} \right] & \text{ordinary product rules} \\ &= \sum_k \frac{dA_{ik}}{dt} B_{kj} + \sum_k A_{ik} \frac{dB_{kj}}{dt} & \text{algebra} \\ &= (A'B)_{ij} + (AB')_{ij} & \text{defn. of matrix multiplication} \\ &= (A'B + AB')_{ij} & \text{defn. matrix addition} \end{aligned}$$

this proves (1.) as i, j were arbitrary in the calculation above. The proof of (2.) and (3.) follow quickly from (1.) since C constant means C' = 0. Proof of (4.) is similar to (1.):

 $(fA)'_{ij} = \frac{d}{dt}((fA)_{ij})$ defn. derivative of matrix $= \frac{d}{dt}(fA_{ij})$ defn. of scalar multiplication $= \frac{df}{dt}A_{ij} + f\frac{dA_{ij}}{dt}$ ordinary product rule $= (\frac{df}{dt}A + f\frac{dA}{dt})_{ij}$ defn. matrix addition $= (\frac{df}{dt}A + f\frac{dA}{dt})_{ij}$ defn. scalar multiplication.

The proof of (5.) follows from taking f(t) = c which has f' = 0. I leave the proof of (6.) as an exercise for the reader. \Box .

To summarize: the calculus of matrices is the same as the calculus of functions with the small qualifier that we must respect the rules of matrix algebra. The noncommutativity of matrix multiplication is the main distinguishing feature.

Since we're discussing this type of differentiation perhaps it would be worthwhile for me to insert a comment about complex functions here. Differentiation of functions from \mathbb{R} to \mathbb{C} is defined by splitting a given function into its real and imaginary parts then we just differentiate with respect to the real variable one component at a time. For example:

$$\frac{d}{dt}(e^{2t}\cos(t) + ie^{2t}\sin(t)) = \frac{d}{dt}(e^{2t}\cos(t)) + i\frac{d}{dt}(e^{2t}\sin(t))
= (2e^{2t}\cos(t) - e^{2t}\sin(t)) + i(2e^{2t}\sin(t) + e^{2t}\cos(t))
= e^{2t}(2+i)(\cos(t) + i\sin(t))
= (2+i)e^{(2+i)t}$$
(13.1)

where I have made use of the identity $e^{x+iy} = e^x(\cos(y) + i\sin(y))$. We just saw that

$$\frac{d}{dt}e^{\lambda t} = \lambda e^{\lambda t}$$

which seems obvious enough until you appreciate that we just proved it for $\lambda = 2 + i$. We make use of this calculation in the next section in the case we have complex e-values.

 $^{^{1}\}mathrm{or}$ definition, depending on how you choose to set-up the complex exponential, I take this as the definition in calculus II

13.2 introduction to systems of linear differential equations

A differential equation (DEqn) is simply an equation that is stated in terms of derivatives. The highest order derivative that appears in the DEqn is called the *order* of the DEqn. In calculus we learned to integrate. Recall that $\int f(x)dx = y$ iff $\frac{dy}{dx} = f(x)$. Everytime you do an integral you are solving a first order DEqn. In fact, it's an *ordinary* DEnq (ODE) since there is only one indpendent variable (it was x). A system of ODEs is a set of differential equations with a common independent variable. It turns out that any linear differential equation can be written as a system of ODEs in *normal form*. I'll define *normal form* then illustrate with a few examples.

Definition 13.2.1.

Let t be a real variable and suppose x_1, x_2, \ldots, x_n are functions of t. If A_{ij} , f_i are functions of t for all $1 \le i \le m$ and $1 \le j \le n$ then the following set of differential equations is defined to be a system of linear differential equations in **normal form**:

$$\frac{dx_1}{dt} = A_{11}x_1 + A_{12}x_2 + \dots + A_{1n}x_n + f_1$$

$$\frac{dx_2}{dt} = A_{21}x_1 + A_{22}x_2 + \dots + A_{2n}x_n + f_2$$

$$\vdots = \vdots \quad \vdots \quad \dots \quad \vdots$$

$$\frac{dx_m}{dt} = A_{m1}x_1 + A_{m2}x_2 + \dots + A_{mn}x_n + f_m$$

In matrix notation, $\frac{dx}{dt} = Ax + f$. The system is called homogeneous if f = 0 whereas the system is called nonhomogeneous if $f \neq 0$. The system is called **constant coefficient** if $\frac{d}{dt}(A_{ij}) = 0$ for all i, j. If m = n and a set of initial conditions $x_1(t_0) = y_1, x_2(t_0) =$ $y_2, \ldots, x_n(t_0) = y_n$ are given then this is called an **initial value problem** (IVP).

Example 13.2.2. If x is the number of tigers and y is the number of rabbits then

$$\frac{dx}{dt} = x + y \qquad \qquad \frac{dy}{dt} = -100x + 20y$$

is a model for the population growth of tigers and bunnies in some closed environment. My logic for my made-up example is as follows: the coefficient 1 is the growth rate for tigers which don't breed to quickly. Whereas the growth rate for bunnies is 20 since bunnies reproduce like, well bunnies. Then the y in the $\frac{dx}{dt}$ equation goes to account for the fact that more bunnies means more tiger food and hence the tiger reproduction should speed up (this is probably a bogus term, but this is my made up example so deal). Then the -100x term accounts for the fact that more tigers means more tigers eating bunnies so naturally this should be negative. In matrix form

$$\begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -100 & 20 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

How do we solve such a system? This is the question we seek to answer.

The preceding example is a *predator-prey* model. There are many other terms that can be added to make the model more realistic. Ultimately all population growth models are only useful if they can account for all significant effects. History has shown population growth models are of only limited use for humans.

Example 13.2.3. Reduction of Order in calculus II you may have studied how to solve y'' + by' + cy = 0 for any choice of constants b, c. This is a second order ODE. We can reduce it to a system of first order ODEs by introducing new variables: $x_1 = y$ and $x_2 = y'$ then we have

$$x_1' = y' = x_2$$

and,

$$x_2' = y'' = -by' - cy = -bx_2 - cx_1$$

As a matrix DEqn,

$$\left[\begin{array}{c} x_1\\ x_2 \end{array}\right]' = \left[\begin{array}{c} 0 & 1\\ -c & -b \end{array}\right] \left[\begin{array}{c} x_1\\ x_2 \end{array}\right]$$

Similarly if y''' + 2y'' + 3y'' + 4y' + 5y = 0 we can introduce variables to reduce the order: $x_1 = y, x_2 = y', x_3 = y'', x_4 = y'''$ then you can show:

$\begin{bmatrix} x_1 \end{bmatrix}$	/ =	0	1	0	0	$\left[\begin{array}{c} x_1 \end{array} \right]$
x_2		0	0	1	0	x_2
x_3		0	0	0	1	$ x_3$
$\begin{bmatrix} x_4 \end{bmatrix}$		-5	-4	-3	-2	$\begin{bmatrix} x_4 \end{bmatrix}$

is equivalent to y''' + 2y'' + 3y'' + 4y' + 5y = 0. We call the matrix above the **companion matrix** of the n-th order constant coefficient ODE. There is a beautiful interplay between solutions to n-th order ODEs and the linear algebra of the compansion matrix.

Example 13.2.4. Suppose y'' + 4y' + 5y = 0 and x'' + x = 0. The is a system of linear second order ODEs. It can be recast as a system of 4 first order ODEs by introducing new variables: $x_1 = y, x_2 = y'$ and $x_3 = x, x_4 = x'$. In matrix form the given system in normal form is:

$\begin{bmatrix} x_1 \end{bmatrix}$	΄ Γ	0	1	0	0]	$\begin{bmatrix} x_1 \end{bmatrix}$
x_2	_	-5	-4	0	0	x_2
x_3	=	0	0	0	1	x_3
$\begin{bmatrix} x_4 \end{bmatrix}$		0	0	-1	0	$\begin{bmatrix} x_4 \end{bmatrix}$

The companion matrix above will be found to have eigenvalues $\lambda = -2 \pm i$ and $\lambda = \pm i$. I know this without further calculation purely on the basis of what I know from DEqns and the interplay I alluded to in the last example.

Example 13.2.5. If y''' + 2y'' + y = 0 we can introduce variables to reduce the order: $x_1 = y, x_2 = y', x_3 = y'', x_4 = y'''$ then you can show:

$\begin{bmatrix} x_1 \end{bmatrix}$	=	0	1	0	0	ΙΓ	x_1
x_2		0	0	1	0		x_2
x_3		0	0	0	1		x_3
$\begin{bmatrix} x_4 \end{bmatrix}$		1	0	-2	0		x_4

is equivalent to y''' + 2y'' + y = 0. If we solve the matrix system then we solve the equation in y and vice-versa. I happen to know the solution to the y equation is $y = c_1 \cos t + c_2 \sin t + c_3 t \cos t + c_4 t \sin t$. From this I can deduce that the companion matrix has a repeated e-value of $\lambda = \pm i$ and just one complex e-vector and its conjugate. This matrix would answer the bonus point question I posed a few sections back. I invite the reader to verify my claims.

Remark 13.2.6.

For those of you who will or have taken math 334 my guesswork above is predicated on two observations:

- 1. the "auxillarly" or "characteristic" equation in the study of the constant coefficient ODEs is identical to the characteristic equation of the companion matrix.
- 2. ultimately eigenvectors will give us exponentials and sines and cosines in the solution to the matrix ODE whereas solutions which have multiplications by t stem from generalized e-vectors. Conversely, if the DEqn has a t or t^2 multiplying cosine, sine or exponential functions then the companion matrix must in turn have generalized e-vectors to account for the t or t^2 etc...

I will not explain (1.) in this course, however we will hopefully make sense of (2.) by the end of this section.

13.3 the matrix exponential

Perhaps the most important first order ODE is $\frac{dy}{dt} = ay$. This DEqn says that the rate of change in y is simply proportional to the amount of y at time t. Geometrically, this DEqn states the solutions value is proportional to its slope at every point in its domain. The solution² is the exponential function $y(t) = e^{at}$.

We face a new differential equation; $\frac{dx}{dt} = Ax$ where x is a vector-valued function of t and $A \in \mathbb{R}^{n \times n}$. Given our success with the exponential function for the scalar case is it not natural to suppose that $x = e^{tA}$ is the solution to the matrix DEqn? The answer is yes. However, we need to define a few items before we can understand the true structure of the claim.

Definition 13.3.1.

Let
$$A\mathbb{R}^{n \times n}$$
 define $e^A \in \mathbb{R}^{n \times n}$ by the following formula

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n = I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \cdots$$
We also denote $e^A = exp(A)$ when convenient.

This definition is the natural extension of the Taylor series formula for the exponential function we derived in calculus II. Of course, you should be skeptical of this definition. How do I even know the series converges for an arbitrary matrix A? And, what do I even mean by "converge" for a series of matrices? (skip the next subsection if you don't care)

²ok, technically separation of variables yields the general solution $y = ce^{at}$ but I'm trying to focus on the exponential function for the moment.

13.3.1 analysis for matrices

Remark 13.3.2.

The purpose of this section is to alert the reader to the gap in the development here. We will use the matrix exponential despite our inability to fully grasp the underlying analysis. Much in the same way we calculate series in calculus without proving every last theorem. I will attempt to at least sketch the analytical underpinnings of the matrix exponential. The reader will be happy to learn this is not part of the required material.

We use the Frobenius norm for $A \in \mathbb{R}^{n \times n}$, $||A|| = \sqrt{\sum_{i,j} (A_{ij})^2}$. We already proved this was a norm in a previous chapter. A sequence of square matrices is a function from \mathbb{N} to $\mathbb{R}^{n \times n}$. We say the sequence $\{A_n\}_{n=1}^{\infty}$ converges to $L \in \mathbb{R}^{n \times n}$ iff for each $\epsilon > 0$ there exists $M \in \mathbb{N}$ such that $||A_n - L|| < \epsilon$ for all n > M. This is the same definition we used in calculus, just now the norm is the Frobenius norm and the functions are replaced by matrices. The definition of a series is also analogue to the definition you learned in calculus II.

Definition 13.3.3.

Let $A_k \in \mathbb{R}^{m \times m}$ for all k, the sequence of partial sums of $\sum_{k=0}^{\infty} A_k$ is given by $S_n = \sum_{k=1}^{n} A_k$. We say the series $\sum_{k=0}^{\infty} A_k$ converges to $L \in \mathbb{R}^{m \times m}$ iff the sequence of partial sums converges to L. In other words,

$$\sum_{k=1}^{\infty} A_k = \lim_{n \to \infty} \sum_{k=1}^{n} A_k.$$

Many of the same theorems hold for matrices:

Proposition 13.3.4.

Let $t \to S_A(t) = \sum A_k(t)$ and $t \to S_B(t) = \sum_k B_k(t)$ be matrix valued functions of a real variable t where the series are uniformly convergent and $c \in \mathbb{R}$ then

- 1. $\sum_k cA_k = c \sum_k A_k$
- 2. $\sum_{k} (A_k + B_k) = \sum_{k} A_k + \sum_{k} B_k$
- 3. $\frac{d}{dt} \left[\sum_{k} A_{k} \right] = \sum_{k} \frac{d}{dt} \left[A_{k} \right]$
- 4. $\int \left[\sum_{k} A_{k} \right] dt = C + \sum_{k} \int A_{k} dt$ where C is a constant matrix.

The summations can go to infinity and the starting index can be any integer.

Uniform convergence means the series converge without regard to the value of t. Let me just refer you to the analysis course, we should discuss uniform convergence in that course, the concept equally well applies here. It is the crucial fact which one needs to interchange the limits which are implicit within \sum_k and $\frac{d}{dt}$. There are counterexamples in the case the series is not uniformly convergent. Fortunately,

Proposition 13.3.5.

Let A be a square matrix then $exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$ is a uniformly convergent series of matrices.

Basically, the argument is as follows: The set of square matrices with the Frobenius norm is isometric to \mathbb{R}^{n^2} which is a complete space. A complete space is one in which every Cauchy sequence converges. We can show that the sequence of partial sums for exp(A) is a Cauchy sequence in $\mathbb{R}^{n \times n}$ hence it converges. Obviously I'm leaving some details out here. You can look at the excellent *Calculus* text by Apostle to see more gory details. Also, if you don't like my approach to the matrix exponential then he has several other ways to look it.

13.3.2 formulas for the matrix exponential

Now for the fun part.

Proposition 13.3.6.

Let A be a square matrix then $\frac{d}{dt} \left[exp(tA) \right] = Aexp(tA)$

Proof: I'll give the proof in two notations. First,

$$\frac{d}{dt} \left[exp(tA) \right] = \frac{d}{dt} \left[\sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k \right] \qquad \text{defn. of matrix exponential} \\ = \sum_{k=0}^{\infty} \frac{d}{dt} \left[\frac{1}{k!} t^k A^k \right] \qquad \text{since matrix exp. uniformly conv.} \\ = \sum_{k=0}^{\infty} \frac{k}{k!} t^{k-1} A^k \qquad A^k \text{ constant and } \frac{d}{dt} (t^k) = k t^{k-1} \\ = A \sum_{k=1}^{\infty} \frac{1}{(k-1)!} t^{k-1} A^{k-1} \qquad \text{since } k! = k(k-1)! \text{ and } A^k = A A^{k-1}. \\ = A exp(tA) \qquad \text{defn. of matrix exponential.} \end{cases}$$

I suspect the following argument is easier to follow:

$$\frac{d}{dt}(exp(tA)) = \frac{d}{dt}(I + tA + \frac{1}{2}t^{2}A^{2} + \frac{1}{3!}t^{3}A^{3} + \cdots)
= \frac{d}{dt}(I) + \frac{d}{dt}(tA) + \frac{1}{2}\frac{d}{dt}(t^{2}A^{2}) + \frac{1}{3\cdot 2}\frac{d}{dt}(t^{3}A^{3}) + \cdots
= A + tA^{2} + \frac{1}{2}t^{2}A^{3} + \cdots
= A(I + tA + \frac{1}{2}t^{2}A^{2} + \cdots)
= Aexp(tA).$$

Notice that we have all we need to see that exp(tA) is a matrix of solutions to the differential equation x' = Ax. The following prop. follows from the preceding prop. and Prop. 3.6.2.

Proposition 13.3.7.

If X = exp(tA) then X' = Aexp(tA) = AX. This means that each column in X is a solution to x' = Ax.

Let us illustrate this proposition with a particularly simple example.

Example 13.3.8. Suppose x' = x, y' = 2y, z' = 3z then in matrix form we have:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

The coefficient matrix is diagonal which makes the k-th power particularly easy to calculate,

$$\begin{aligned} A^{k} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}^{k} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2^{k} & 0 \\ 0 & 0 & 3^{k} \end{bmatrix} \\ \Rightarrow & exp(tA) = \sum_{k=0}^{\infty} \frac{t^{k}}{k!} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2^{k} & 0 \\ 0 & 0 & 3^{k} \end{bmatrix} = \begin{bmatrix} \sum_{k=0}^{\infty} \frac{t^{k}}{k!} 1^{k} & 0 & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{t^{k}}{k!} 2^{k} & 0 \\ 0 & 0 & \sum_{k=0}^{\infty} \frac{t^{k}}{k!} 3^{k} \end{bmatrix} \\ \Rightarrow & exp(tA) = \begin{bmatrix} e^{t} & 0 & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^{3t} \end{bmatrix} \end{aligned}$$

Thus we find three solutions to x' = Ax,

$$x_1(t) = \begin{bmatrix} e^t \\ 0 \\ 0 \end{bmatrix} \qquad x_2(t) = \begin{bmatrix} 0 \\ e^{2t} \\ 0 \end{bmatrix} \qquad x_3(t) = \begin{bmatrix} 0 \\ 0 \\ e^{3t} \end{bmatrix}$$

In turn these vector solutions amount to the solutions $x = e^t$, y = 0, z = 0 or x = 0, $y = e^{2t}$, z = 0 or x = 0, y = 0, $z = e^{3t}$. It is easy to check these solutions.

Usually we cannot calculate the matrix exponential explicitly by such a straightforward calculation. We need e-vectors and sometimes generalized e-vectors to reliably calculate the solutions of interest.

Proposition 13.3.9.

If A, B are square matrices such that
$$AB = BA$$
 then $e^{A+B} = e^A e^B$

Proof: I'll show how this works for terms up to quadratic order,

$$e^{A}e^{B} = (1 + A + \frac{1}{2}A^{2} + \dots)(1 + B + \frac{1}{2}B^{2} + \dots) = 1 + (A + B) + \frac{1}{2}A^{2} + AB + \frac{1}{2}B^{2} + \dots$$

However, since AB = BA and

$$(A+B)^{2} = (A+B)(A+B) = A^{2} + AB + BA + B^{2} = A^{2} + 2AB + B^{2}.$$

Thus,

$$e^{A}e^{B} = 1 + (A+B) + \frac{1}{2}(A+B)^{2} + \dots = e^{A+B}$$

You might wonder what happens if $AB \neq BA$. In this case we can account for the departure from commutativity by the **commutator** of A and B.

Definition 13.3.10.

Let $A, B \in \mathbb{R}^{n \times n}$ then the commutator of A and B is [A, B] = AB - BA.

Proposition 13.3.11.

Let $A, B, C \in \mathbb{R}^{n \times n}$ then 1. [A, B] = -[B, A]2. [A + B, C] = [A, C] + [B, C]3. [AB, C] = A[B, C] + [A, C]B4. [A, BC] = B[A, C] + [A, B]C5. [[A, B], C] + [[B, C], A] + [[C, A], B] = 0

The proofs of the properties above are not difficult. In contrast, the following formula known as the Baker-Campbell-Hausdorff (BCH) relation takes considerably more calculation:

$$e^{A}e^{B} = e^{A+B+\frac{1}{2}[A,B]+\frac{1}{12}[[A,B],B]+\frac{1}{12}[[B,A],A]+\cdots}$$
 BCH-formula

The higher order terms can also be written in terms of nested commutators. What this means is that if we know the values of the commutators of two matrices then we can calculate the product of their exponentials with a little patience. This connection between multiplication of exponentials and commutators of matrices is at the heart of Lie theory. Actually, mathematicians have greatly abstracted the idea of Lie algebras and Lie groups way past matrices but the concrete example of matrix Lie groups and algebras is perhaps the most satisfying. If you'd like to know more just ask. It would make an excellent topic for an independent study that extended this course.

Remark 13.3.12.

In fact the *BCH* holds in the abstract as well. For example, it holds for the Lie algebra of derivations on smooth functions. A *derivation* is a linear differential operator which satisfies the product rule. The derivative operator is a derivation since D[fg] = D[f]g + fD[g]. The commutator of derivations is defined by [X, Y][f] = X(Y(f)) - Y(X(f)). It can be shown that [D, D] = 0 thus the BCH formula yields

$$e^{aD}e^{bD} = e^{(a+b)D}.$$

If the coefficient of D is thought of as position then multiplication by e^{bD} generates a translation in the position. By the way, we can state Taylor's Theorem rather compactly in this operator notation: $f(x+h) = exp(hD)f(x) = f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{3!}f'''(x) + \cdots$

Proposition 13.3.13.

Let $A, P \in \mathbb{R}^{n \times n}$ and assume P is invertible then

$$exp(P^{-1}AP) = P^{-1}exp(A)P$$

Proof: this identity follows from the following observation:

$$(P^{-1}AP)^{k} = P^{-1}APP^{-1}APP^{-1}AP \cdots P^{-1}AP = P^{-1}A^{k}P.$$

Thus $exp(P^{-1}AP) = \sum_{k=0}^{\infty} \frac{1}{k!} (P^{-1}AP)^k = P^{-1} (\sum_{k=0}^{\infty} \frac{1}{k!} A^k) P = P^{-1} exp(A) P.$

Proposition 13.3.14.

Let A be a square matrix, det(exp(A)) = exp(trace(A)).

Proof: If the matrix A is diagonalizable then the proof is simple. Diagonalizability means there exists invertible $P = [v_1|v_2|\cdots|v_n]$ such that $P^{-1}AP = D = [\lambda_1v_1|\lambda_2v_2|\cdots|\lambda_nv_n]$ where v_i is an e-vector with e-value λ_i for all *i*. Use the preceding proposition to calculate

$$\det(exp(D)) = \det(exp(P^{-1}AP) = \det(P^{-1}exp(A)P) = \det(P^{-1}P)\det(exp(A)) = \det(exp(A))$$

On the other hand, the trace is cyclic trace(ABC) = trace(BCA)

$$trace(D) = trace(P^{-1}AP) = trace(PP^{-1}A) = trace(A)$$

But, we also know D is diagonal with eigenvalues on the diagonal hence exp(D) is diagonal with e^{λ_i} on the corresponding diagonals

$$det(exp(D)) = e^{\lambda_1} e^{\lambda_2} \cdots e^{\lambda_n}$$
 and $trace(D) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$

Finally, use the laws of exponents to complete the proof,

$$e^{trace(A)} = e^{trace(D)} = e^{\lambda_1 + \lambda_2 + \dots + \lambda_n} = e^{\lambda_1} e^{\lambda_2} \cdots e^{\lambda_n} = det(exp(D)) = det(exp(A)).$$

I've seen this proof in texts presented as if it were the general proof. But, not all matrices are diagonalizable so this is a curious proof. I stated the proposition for an arbitrary matrix and I meant it. The proof, the real proof, is less obvious. Let me sketch it for you:

better proof: The preceding proof shows it may be hopeful to suppose that det(exp(tA)) = exp(t trace(A)) for $t \in \mathbb{R}$. Notice that $y = e^{kt}$ satisfies the differential equation $\frac{dy}{dt} = ky$. Conversely, if $\frac{dy}{dt} = ky$ for some constant k then the general solution is given by $y = c_o e^{kt}$ for some $c_o \in \mathbb{R}$. Let f(t) = det(exp(tA)). If we can show that f'(t) = trace(A)f(t) then we can conclude $f(t) = c_0 e^{ttrace(A)}$. Consider:

$$\begin{aligned} f'(t) &= \left. \frac{d}{dh} \left(f(t+h) \right|_{h=0} \\ &= \left. \frac{d}{dh} \left(det(exp[(t+h)A]) \right|_{h=0} \\ &= \left. \frac{d}{dh} \left(det(exp[tA+hA]) \right|_{h=0} \\ &= \left. \frac{d}{dh} \left(det(exp[tA]exp[hA]) \right|_{h=0} \\ &= det(exp[tA]) \frac{d}{dh} \left(det(exp[hA]) \right|_{h=0} \\ &= f(t) \frac{d}{dh} \left(det(I+hA+\frac{1}{2}h^2A^2+\frac{1}{3!}h^3A^3+\cdots) \right|_{h=0} \\ &= f(t) \frac{d}{dh} \left(det(I+hA) \right) \right|_{h=0} \end{aligned}$$

Let us discuss the $\frac{d}{dh}(det(I+hA))$ term separately for a moment:³

$$\begin{aligned} \frac{d}{dh}(det(I+hA)) &= \frac{d}{dh} [\sum_{i_1,\dots,i_n} \epsilon_{i_1 i_2\dots i_n} (I+hA)_{i_1 1} (I+hA)_{i_2 2} \cdots (I+hA)_{i_n n}]_{h=0} \\ &= \sum_{i_1,\dots,i_n} \epsilon_{i_1 i_2\dots i_n} \frac{d}{dh} [(I+hA)_{1 i_1} (I+hA)_{1 i_2} \cdots (I+hA)_{n i_n}]_{h=0} \\ &= \sum_{i_1,\dots,i_n} \epsilon_{i_1 i_2\dots i_n} (A_{1 i_1} I_{1 i_2} \cdots I_{n i_n} + I_{1 i_1} A_{2 i_2} \cdots I_{n i_n} + \dots + I_{1 i_1} I_{2 i_2} \cdots A_{n i_n}) \\ &= \sum_{i_1} \epsilon_{i_1 2\dots n} A_{1 i_1} + \sum_{i_2} \epsilon_{1 i_2 \dots n} A_{2 i_2} + \dots + \sum_{i_n} \epsilon_{1 2\dots I_n} A_{n i_n} \\ &= A_{11} + A_{22} + \dots + A_{nn} \\ &= trace(A) \end{aligned}$$

Therefore, f'(t) = trace(A)f(t) consequently, $f(t) = c_o e^{t trace(A)} = det(exp(tA))$. However, we can resolve c_o by calculating $f(0) = det(exp(0)) = det(I) = 1 = c_o$ hence

$$e^{t \, trace(A)} = det(exp(tA))$$

Take t = 1 to obtain the desired result. \Box

Remark 13.3.15.

The formula det(exp(A)) = exp(trace(A)) is very important to the theory of matrix Lie groups and Lie algebras. Generically, if G is the Lie group and \mathfrak{g} is the Lie algebra then they are connected via the matrix exponential: $exp : \mathfrak{g} \to G_o$ where I mean G_o to denoted the connected component of the identity. For example, the set of all nonsingular matrices GL(n) forms a Lie group which is disconnected. Half of GL(n) has positive determinant whereas the other half has negative determinant. The set of all $n \times n$ matrices is denoted gl(n) and it can be shown that exp(gl(n)) maps onto the part of GL(n) which has positive determinant. One can even define a matrix logarithm map which serves as a local inverse for the matrix exponential near the identity. Generally the matrix exponential is not injective thus some technical considerations must be discussed before we could put the matrix log on a solid footing. This would take us outside the scope of this course. However, this would be a nice topic to do a follow-up independent study. The theory of matrix Lie groups and their representations is ubiqitious in modern quantum mechanical physics.

Finally, we come to the formula that is most important to our study of systems of DEqns. Let's call this the magic formula.

Proposition 13.3.16.

Let
$$\lambda \in \mathbb{C}$$
 and suppose $A \in \mathbb{R}^{n \times n}$ then

$$exp(tA) = e^{\lambda t} (I + t(A - \lambda I) + \frac{t^2}{2} (A - \lambda I)^2 + \frac{t^3}{3!} (A - \lambda I)^3 + \cdots).$$

³I use the definition of the identity matrix $I_{ij} = \delta_{ij}$ in eliminating all but the last summation in the fourth line. Then the levi-civita symbols serve the same purpose in going to the fifth line as $\epsilon_{i_12...n} = \delta_{1i_1}, \epsilon_{1i_2...n} = \delta_{2i_2}$ etc...

Proof: Notice that $tA = t(A - \lambda I) + t\lambda I$ and $t\lambda I$ commutes with all matrices thus,

$$exp(tA) = exp(t(A - \lambda I) + t\lambda I)$$

= $exp(t(A - \lambda I))exp(t\lambda I)$
= $e^{\lambda t}exp(t(A - \lambda I))$
= $e^{\lambda t}(I + t(A - \lambda I) + \frac{t^2}{2}(A - \lambda I)^2 + \frac{t^3}{3!}(A - \lambda I)^3 + \cdots)$

In the third line I used the identity proved below,

$$exp(t\lambda I) = I + t\lambda I + \frac{1}{2}(t\lambda)^2 I^2 + \dots = I(1 + t\lambda + \frac{(t\lambda)^2}{2} + \dots) = Ie^{t\lambda}.$$

While the proofs leading up to the magic formula only dealt with real matrices it is not hard to see the proofs are easily modified to allow for complex matrices.

13.4 solutions for systems of DEqns with real eigenvalues

Let us return to the problem of solving $\vec{x}' = A\vec{x}$ for a constant square matrix A where $\vec{x} = [x_1, x_2, \ldots, x_n]$ is a vector of functions of t. I'm adding the vector notation to help distinguish the scalar function x_1 from the vector function \vec{x}_1 in this section. Let me state one theorem from the theory of differential equations. The existence of solutions theorem which is the heart of of this theorem is fairly involved to prove. It requires a solid understanding of real analysis.

Theorem 13.4.1.

If $\vec{x}' = A\vec{x}$ and A is a constant matrix then any solution to the system has the form $\vec{x}(t) = c_1\vec{x}_1(t) + c_2\vec{x}_2(t) + \dots + c_n\vec{x}_n(t)$ where $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$ is a linearly independent set of solutions defined on \mathbb{R} (this is called the **fundamental solution set**). Moreover, these fundamental solutions can be concatenated into a single invertible solution matrix called the **fundamental matrix** $X = [\vec{x}_1 | \vec{x}_2 | \dots | \vec{x}_n]$ and the general solution can be expressed as $\vec{x}(t) = X(t)\vec{c}$ where \vec{c}

solution to the IVP is $\vec{x}(t) = X^{-1}(t_o)X(t)\vec{x}_o$. We proved in the previous section that the matrix exponential exp(tA) is a solution matrix and the inverse is easy enought to guess: $exp(tA)^{-1} = exp(-tA)$. This proves the columns of exp(tA) are solutions to $\vec{x}' = A\vec{x}$ which are linearly independent and as such form a fundamental solution set.

is an arbitrary vector of real constants. If an initial condition $\vec{x}(t_o) = \vec{x}_o$ is given then the

<u>Problem</u>: we cannot directly calculate exp(tA) for most matrices A. We have a solution we can't calculate. What good is that?

When can we explicitly calculate exp(tA) without much thought? Two cases come to mind: (1.) if A is diagonal then it's easy, saw this in Example 13.3.8, (2.) if A is a **nilpotent** matrix then there is some finite power of the matrix which is zero; $A^k = 0$. In the nilpotent case the infinite series defining the matrix exponential truncates at order k:

$$exp(tA) = I + tA + \frac{t^2}{2}A^2 + \dots + \frac{t^{k-1}}{(k-1)!}A^{k-1}$$

Example 13.4.2. Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ we calculate $A^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ thus $exp(tA) = I + tA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + t \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$

Incidentally, the solution to $\vec{x}' = A\vec{x}$ is generally $\vec{x}(t) = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} t \\ 1 \end{bmatrix}$. In other words, $x_1(t) = c_2 + c_2 t$ whereas $x_2(t) = c_2$. These solutions are easily seen to solve the system $x'_1 = x_2$ and $x'_2 = 0$.

Unfortunately, the calculation we just did in the last example almost never works. For example, try to calculate an arbitrary power of $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, let me know how it works out. We would like for all examples to truncate. The **magic formula** gives us a way around this dilemma:

Proposition 13.4.3.

Let $A \in \mathbb{R}^{n \times n}$. Suppose v is an e-vector with e-value λ then $exp(tA)v = e^{\lambda t}v$.

Proof: we are given that $(A - \lambda I)v = 0$ and it follows that $(A - \lambda I)^k v = 0$ for all $k \ge 1$. Use the magic formula,

$$exp(tA)v = e^{\lambda t}(I + t(A - \lambda I) + \cdots)v = e^{\lambda t}(Iv + t(A - \lambda I)v + \cdots = e^{\lambda t}v$$

noting all the higher order terms vanish since $(A - \lambda I)^k v = 0$. \Box

We can't hope for the matrix exponential itself to truncate, but when we multiply exp(tA) on an e-vector something special happens. Since $e^{\lambda t} \neq 0$ the set of vector functions

 $\{e^{\lambda_1 t}v_1, e^{\lambda_2 t}v_2, \ldots, e^{\lambda_k t}v_k\}$ will be linearly independent if the e-vectors v_i are linearly independent. If the matrix A is diagonalizable then we'll be able to find enough e-vectors to construct a fundamental solution set using e-vectors alone. However, if A is not diagonalizable, and has only real e-values, then we can still find a Jordan basis $\{v_1, v_2, \ldots, v_n\}$ which consists of generalized e-vectors and it follows that $\{e^{tA}v_1, e^{tA}v_2, \ldots, e^{tA}v_n\}$ forms a fundamental solution set. Moreover, this is not just of theoretical use. We can actually calculate this solution set.

Proposition 13.4.4.

Let $A \in \mathbb{R}^{n \times n}$. Suppose A has a chain $\{v_1, v_2, \ldots, v_k\}$ is of generalized e-vectors with e-value λ , meaning $(A - \lambda)v_1 = 0$ and $(A - \lambda)v_{k-1} = v_k$ for $k \ge 2$, then

1.
$$e^{tA}v_1 = e^{\lambda t}v_1$$
,
2. $e^{tA}v_2 = e^{\lambda t}(v_2 + tv_1)$,
3. $e^{tA}v_3 = e^{\lambda t}(v_3 + tv_2 + \frac{t^2}{2}v_1)$,
4. $e^{tA}v_k = e^{\lambda t}(v_k + tv_{k-1} + \dots + \frac{t^{k-1}}{(k-1)!}v_1)$

Proof: Study the chain condition,

$$(A - \lambda I)v_2 = v_1 \quad \Rightarrow \quad (A - \lambda)^2 v_2 = (A - \lambda I)v_1 = 0$$

$$(A - \lambda I)v_3 = v_2 \quad \Rightarrow \quad (A - \lambda I)^2 v_3 = (A - \lambda I)v_2 = v_1$$

Continuing with such calculations⁴ we find $(A - \lambda I)^{j}v_{i} = v_{i-j}$ for all i > j and $(A - \lambda I)^{i}v_{i} = 0$. The magic formula completes the proof:

$$e^{tA}v_2 = e^{\lambda t} \left(v_2 + t(A - \lambda I)v_2 + \frac{t^2}{2}(A - \lambda I)^2 v_2 \cdots \right) = e^{\lambda t} \left(v_2 + tv_1 \right)$$

likewise,

$$e^{tA}v_3 = e^{\lambda t} \left(v_3 + t(A - \lambda I)v_3 + \frac{t^2}{2}(A - \lambda I)^2 v_3 + \frac{t^3}{3!}(A - \lambda I)^3 v_3 + \cdots \right)$$

= $e^{\lambda t} \left(v_3 + tv_2 + \frac{t^2}{2}(A - \lambda I)v_2 \right)$
= $e^{\lambda t} \left(v_3 + tv_2 + \frac{t^2}{2}v_1 \right).$

We already proved the e-vector case in the preceding proposition and the general case follows from essentially the same calculation. \Box

We have all the theory we need to solve systems of homogeneous constant coefficient ODEs.

Example 13.4.5. Recall Example 11.4.5 we found $A = \begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix}$ had e-values $\lambda_1 = 0$ and $\lambda_2 = 4$ and corresponding e-vectors

$$\vec{u}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$
 and $\vec{u}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

thus we find the general solution to $\vec{x}' = A\vec{x}$ is simply,

$$\vec{x}(t) = c_1 \begin{bmatrix} 1\\ -3 \end{bmatrix} + c_2 e^{4t} \begin{bmatrix} 1\\ 1 \end{bmatrix}$$

just to illustrate the terms: we have fundmamental solution set and matrix:

$$\left\{ \left[\begin{array}{c} 1\\ -3 \end{array} \right], \left[\begin{array}{c} e^{4t}\\ e^{4t} \end{array} \right] \right\} \qquad X = \left[\begin{array}{c} 1 & e^{4t}\\ -3 & e^{4t} \end{array} \right]$$

Notice that a different choice of e-vector scaling would just end up adjusting the values of c_1, c_2 in the event an initial condition was given. This is why different choices of e-vectors still gives us the same general solution. It is the flexibility to change c_1, c_2 that allows us to fit any initial condition.

Example 13.4.6. We can modify Example 13.2.2 and propose a different model for a tiger/bunny system. Suppose x is the number of tigers and y is the number of rabbits then

$$\frac{dx}{dt} = x - 4y \qquad \qquad \frac{dy}{dt} = -10x + 19y$$

is a model for the population growth of tigers and bunnies in some closed environment. Suppose that there is initially 2 tigers and 100 bunnies. Find the populations of tigers and bunnies at time t > 0:

 $^{^{4}}$ keep in mind these conditions hold because of our current labling scheme, if we used a different indexing system then you'd have to <u>think</u> about how the chain conditions work out, to test your skill perhaps try to find the general solution for the system with the matrix from Example 11.6.11

<u>Solution</u>: notice that we must solve $\vec{x}' = A\vec{x}$ where $A = \begin{bmatrix} 1 & -4 \\ -10 & 19 \end{bmatrix}$ and $\vec{x}(0) = \begin{bmatrix} 2, 100 \end{bmatrix}^T$. We can calculate the eigenvalues and corresponding eigenvectors:

$$det(A - \lambda I) = 0 \implies \lambda_1 = -1, \ \lambda_2 = 21 \implies u_1 = \begin{bmatrix} 2\\1 \end{bmatrix}, \ u_2 = \begin{bmatrix} -1\\5 \end{bmatrix}$$

Therefore, using Proposition 13.4.4, the general solution has the form:

$$\vec{x}(t) = c_1 e^{-t} \begin{bmatrix} 2\\1 \end{bmatrix} + c_2 e^{21t} \begin{bmatrix} -1\\5 \end{bmatrix}.$$

However, we also know that $\vec{x}(0) = [2, 100]^T$ hence

$$\begin{bmatrix} 2\\100 \end{bmatrix} = c_1 \begin{bmatrix} 2\\1 \end{bmatrix} + c_2 \begin{bmatrix} -1\\5 \end{bmatrix} \Rightarrow \begin{bmatrix} 2\\100 \end{bmatrix} = \begin{bmatrix} 2&-1\\1&5 \end{bmatrix} \begin{bmatrix} c_1\\c_2 \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} c_1\\c_2 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 5&1\\-1&2 \end{bmatrix} \begin{bmatrix} 2\\100 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 110\\198 \end{bmatrix}$$

Finally, we find the vector-form of the solution to the given initial value problem:

$$\vec{x}(t) = 10e^{-t} \begin{bmatrix} 2\\1 \end{bmatrix} + \frac{198}{11}e^{21t} \begin{bmatrix} -1\\5 \end{bmatrix}$$

Which means that $x(t) = 20e^{-t} - \frac{198}{11}e^{21t}$ and $y(t) = 1020e^{-t} + 90e^{21t}$ are the number of tigers and bunnies respective at time t.

Notice that a different choice of e-vectors would have just made for a different choice of c_1, c_2 in the preceding example. Also, notice that when an initial condition is given there ought not be any undetermined coefficients in the final answer⁵.

Example 13.4.7. We found that in Example 11.4.7 the matrix $A = \begin{bmatrix} 0 & 0 & -4 \\ 2 & 4 & 2 \\ 2 & 0 & 6 \end{bmatrix}$ has e-values

 $\lambda_1 = \lambda_2 = 4$ and $\lambda_3 = 2$ with corresponding e-vectors

$$\vec{u}_1 = \begin{bmatrix} 0\\1\\0 \end{bmatrix} \vec{u}_2 = \begin{bmatrix} -1\\0\\1 \end{bmatrix} \vec{u}_3 = \begin{bmatrix} -2\\1\\1 \end{bmatrix}$$

Hence, using Proposition 13.4.4 and Theorem 13.4.1 the general solution of $\frac{d\vec{x}}{dt} = A\vec{x}$ is simply:

$$\vec{x}(t) = c_1 e^{4t} \vec{u}_1 + c_2 e^{4t} \vec{u}_2 + c_3 e^{2t} \vec{u}_3 = c_1 e^{4t} \begin{bmatrix} 0\\1\\0 \end{bmatrix} + c_2 e^{4t} \begin{bmatrix} -1\\0\\1 \end{bmatrix} + c_3 e^{2t} \begin{bmatrix} -2\\1\\1 \end{bmatrix}$$

⁵Assuming of course that there are enough initial conditions given to pick a unique solution from the family of solutions which we call the "general solution".

Example 13.4.8. Find the general solution of $\frac{d\vec{x}}{dt} = A\vec{x}$ given that:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

We analyzed this matrix in Example 11.6.11. We found a pair of chains of generalized e-vectors all with eigenvalue $\lambda = 1$ which satisfied the following conditions:

$$(A-I)\vec{u}_3 = \vec{u}_1, \ (A-I)\vec{u}_1 = 0 \qquad (A-I)\vec{u}_4 = \vec{u}_2, \ (A-I)\vec{u}_2 = 0$$

In particular, $\vec{u}_j = e_j$ for j = 1, 2, 3, 4. We can use the magic formula to extract 4 solutions from the matrix exponential, by Proposition 13.4.4 we find:

$$\vec{x}_{1} = e^{At}\vec{u}_{1} = e^{t}\vec{u}_{1} = e^{t}e_{1}$$

$$\vec{x}_{2} = e^{At}\vec{u}_{2} = e^{t}(e_{2} + te_{1})$$

$$\vec{x}_{3} = e^{At}\vec{u}_{3} = e^{t}e_{3}$$

$$\vec{x}_{4} = e^{At}\vec{u}_{4} = e^{t}(e_{4} + te_{3})$$
(13.2)

Let's write the general solution in vector and scalar form, by Theorem 13.4.1,

$$\vec{x}(t) = c_1 \vec{x}_1 + c_2 \vec{x}_2 + c_3 \vec{x}_3 + c_4 \vec{x}_4 = c_1 e^t e_1 + c_2 e^t (e_2 + te_1) + c_3 e^t e_3 + c_4 e^t (e_4 + te_3) = \begin{bmatrix} c_1 e^t + tc_2 e^t \\ c_2 e^t \\ c_3 e^t + tc_4 e^t \\ c_4 e^t \end{bmatrix}$$

In other words, $x_1(t) = c_1e^t + tc_2e^t$, $x_2(t) = c_2e^t$, $x_3(t) = c_3e^t + tc_4e^t$ and $x_4(t) = c_4e^t$ form the general solution to the given system of differential equations.

Example 13.4.9. Find the general solution of $\frac{d\vec{x}}{dt} = A\vec{x}$ given (generalized)eigenvectors \vec{u}_i , i = 1, 2, 3, 4, 5, 6, 7, 8, 9 such that:

$$(A-I)\vec{u}_1 = 0, \quad A\vec{u}_2 = \vec{u}_2, \quad A\vec{u}_3 = 7\vec{u}_3, \quad (A-I)\vec{u}_4 = \vec{u}_1$$
$$(A+5I)\vec{u}_5 = 0, \quad (A-3I)\vec{u}_6 = \vec{u}_7 \quad A\vec{u}_7 = 3\vec{u}_7, \quad A\vec{u}_8 = 0, \quad (A-3I)\vec{u}_9 = \vec{u}_6$$

We can use the magic formula to extract 9 solutions from the matrix exponential, by Proposition 13.4.4 we find:

$$\vec{x}_{1} = e^{At}\vec{u}_{1} = e^{t}\vec{u}_{1} = e^{t}\vec{u}_{1}$$
(13.3)

$$\vec{x}_{2} = e^{At}\vec{u}_{2} = e^{t}\vec{u}_{2}$$
(13.3)

$$\vec{x}_{3} = e^{At}\vec{u}_{3} = e^{7t}\vec{u}_{3}$$
(13.3)

$$\vec{x}_{4} = e^{At}\vec{u}_{3} = e^{7t}\vec{u}_{3}$$
(13.3)

$$\vec{x}_{4} = e^{At}\vec{u}_{4} = e^{t}(\vec{u}_{4} + t\vec{u}_{1})$$
(13.3)

$$\vec{x}_{5} = e^{At}\vec{u}_{5} = e^{-5t}\vec{u}_{5}$$
(13.3)

$$\vec{x}_{6} = e^{At}\vec{u}_{6} = e^{3t}(\vec{u}_{6} + t\vec{u}_{7})$$
(13.3)

$$\vec{x}_{6} = e^{At}\vec{u}_{6} = e^{3t}(\vec{u}_{6} + t\vec{u}_{7})$$
(13.3)

$$\vec{x}_{6} = e^{At}\vec{u}_{7} = e^{3t}\vec{u}_{7}$$
(13.3)

$$\vec{x}_{7} = e^{At}\vec{u}_{7} = e^{3t}\vec{u}_{7}$$
(13.3)

$$\vec{x}_{7} = e^{At}\vec{u}_{8} = \vec{u}_{8}$$
(13.3)

$$\vec{x}_{9} = e^{At}\vec{u}_{9} = e^{3t}(\vec{u}_{9} + t\vec{u}_{6} + \frac{1}{2}t^{2}\vec{u}_{7})$$
(13.3)

$$\vec{x}_{9} = e^{At}\vec{u}_{9} = e^{3t}(\vec{u}_{9} + t\vec{u}_{6} + \frac{1}{2}t^{2}\vec{u}_{7})$$
(13.3)

$$\vec{x}_{9} = e^{At}\vec{u}_{9} = e^{3t}(\vec{u}_{9} + t\vec{u}_{6} + \frac{1}{2}t^{2}\vec{u}_{7})$$
(13.3)

$$\vec{x}_{9} = e^{At}\vec{u}_{9} = e^{3t}(\vec{u}_{9} + t\vec{u}_{9} + \frac{1}{2}t^{2}\vec{u}_{7})$$
(13.3)

$$\vec{x}_{9} = e^{At}\vec{u}_{9} = e^{3t}(\vec{u}_{9} + t\vec{u}_{9} + \frac{1}{2}t^{2}\vec{u}_{7})$$
(13.3)

$$\vec{x}_{9} = e^{At}\vec{u}_{9} = e^{3t}(\vec{u}_{9} + t\vec{u}_{9} + \frac{1}{2}t^{2}\vec{u}_{7})$$
(13.3)

$$\vec{x}_{9} = e^{At}\vec{u}_{9} = e^{At}\vec{u$$

Let's write the general solution in vector and scalar form, by Theorem 13.4.1,

$$\vec{x}(t) = \sum_{i=1}^{9} c_i \vec{x}_i$$

where the formulas for each solution \vec{x}_i was given above. If I was to give an explicit matrix A with the eigenvectors given above it would be a 9×9 matrix.

Challenge: find the matrix exponential e^{At} in terms of the given (generalized) eigenvectors.

Hopefully the examples have helped the theory settle in by now. We have one last question to settle for systems of DEqns.

Theorem 13.4.10.

The nonhomogeneous case $\vec{x}' = A\vec{x} + \vec{f}$ the general solution is $\vec{x}(t) = X(t)c + \vec{x}_p(t)$ where X is a fundamental matrix for the corresponding homogeneous system and \vec{x}_p is a particular solution to the nonhomogeneous system. We can calculate $\vec{x}_p(t) = X(t) \int X^{-1} \vec{f} dt$.

Proof: suppose that $\vec{x}_p = X\vec{v}$ for X a fundamental matrix of $\vec{x}' = A\vec{x}$ and some vector of unknown functions \vec{v} . We seek conditions on \vec{v} which make \vec{x}_p satisfy $\vec{x}_p' = A\vec{x}_p + \vec{f}$. Consider,

$$(\vec{x}_p)' = (X\vec{v})' = X'\vec{v} + X\vec{v}' = AX\vec{v} + X\vec{v}'$$

But, $\vec{x}_p' = A\vec{X}_p + \vec{f} = AX\vec{v} + \vec{f}$ hence

$$X\frac{d\vec{v}}{dt} = \vec{f} \quad \Rightarrow \quad \frac{d\vec{v}}{dt} = X^{-1}\vec{f}$$

Integrate to find $\vec{v} = \int X^{-1} \vec{f} dt$ therefore $x_p(t) = X(t) \int X^{-1} \vec{f} dt$. \Box

If you ever work through variation of parameters for higher order ODEqns then you should appreciate the calculation above. In fact, we can derive *n*-th order variation of parameters from converting the *n*-th order ODE by reduction of order to a system of *n* first order linear ODEs. You can show that the so-called Wronskian of the fundamental solution set is precisely the determinant of the fundamental matrix for the system $\vec{x}' = A\vec{x}$ where *A* is the companion matrix. I have this worked out in an old test from a DEqns course I taught at NCSU⁶

⁶see solution of Problem 6 in *www.supermath.info/ma341f07test2_sol.pdf* for the n = 2 case of this comment, also §6.4 of Nagel Saff and Snider covers *n*-th order variation of parameters if you want to see details

Example 13.4.11. Suppose that $A = \begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix}$ and $\vec{f} = \begin{bmatrix} e^t \\ e^{-t} \end{bmatrix}$, find the general solution of the nonhomogenous DEqn $\vec{x}' = A\vec{x} + \vec{f}$. Recall that in Example 13.4.5 we found $\vec{x}' = A\vec{x}$ has fundamental matrix $X = \begin{bmatrix} 1 & e^{4t} \\ -3 & e^{4t} \end{bmatrix}$. Use variation of parameters for systems of ODEs to constuct \vec{x}_p . First calculate the inverse of the fundamental matrix, for a 2 × 2 we know a formula:

$$X^{-1}(t) = \frac{1}{e^{4t} - (-3)e^{4t}} \begin{bmatrix} e^{4t} & -e^{4t} \\ 3 & 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & -1 \\ 3e^{-4t} & e^{-4t} \end{bmatrix}$$

Thus,

$$\begin{aligned} x_p(t) &= X(t) \int \frac{1}{4} \begin{bmatrix} 1 & -1 \\ 3e^{-4t} & e^{-4t} \end{bmatrix} \begin{bmatrix} e^t \\ e^{-t} \end{bmatrix} dt = \frac{1}{4} X(t) \int \begin{bmatrix} e^t - e^{-t} \\ 3e^{-3t} + e^{-5t} \end{bmatrix} dt \\ &= \frac{1}{4} \begin{bmatrix} 1 & e^{4t} \\ -3 & e^{4t} \end{bmatrix} \begin{bmatrix} e^t + e^{-t} \\ -e^{-3t} - \frac{1}{5}e^{-5t} \end{bmatrix} \\ &= \frac{1}{4} \begin{bmatrix} 1(e^t + e^{-t}) + e^{4t}(-e^{-3t} - \frac{1}{5}e^{-5t}) \\ -3(e^t + e^{-t}) + e^{4t}(-e^{-3t} - \frac{1}{5}e^{-5t}) \end{bmatrix} \\ &= \frac{1}{4} \begin{bmatrix} e^t + e^{-t} - e^t - \frac{1}{5}e^{-t} \\ -3e^t - 3e^{-t} - e^t - \frac{1}{5}e^{-t} \end{bmatrix} \\ &= \frac{1}{4} \begin{bmatrix} \frac{4}{5}e^{-t} \\ -4e^t - \frac{16}{5}e^{-t} \end{bmatrix} \end{aligned}$$

Therefore, the general solution is

$$\vec{x}(t) = c_1 \begin{bmatrix} 1\\ -3 \end{bmatrix} + c_2 e^{4t} \begin{bmatrix} 1\\ 1 \end{bmatrix} + \frac{1}{5} \begin{bmatrix} e^{-t}\\ -e^t - 4e^{-t} \end{bmatrix}.$$

The general scalar solutions implicit within the general vector solution $\vec{x}(t) = [x(t), y(t)]^T$ are

$$x(t) = c_1 + c_2 e^{4t} + \frac{1}{5} e^{-t} \qquad y(t) = -3c_1 + c_2 e^{4t} - \frac{1}{5} e^t - \frac{4}{5} e^{-t}.$$

I'll probably ask you to solve a 3×3 system in the homework. The calculation is nearly the same as the preceding example with the small inconvenience that finding the inverse of a 3×3 requires some calculation.

Remark 13.4.12.

You might wonder how would you solve a system of ODEs x' = Ax such that the coefficients A_{ij} are not constant. We will not cover such problems in this course. We do cover how to solve an n - th order ODE with nonconstant coefficients via series techniques in Math 334. It's probably possible to extend some of those techniques to systems. Laplace Transforms also extend to systems of ODEs. It's just a matter of algebra.

13.5 solutions for systems of DEqns with complex eigenvalues

The calculations in the preceding section still make sense for a complex e-value and complex e-vector. However, we usually need to find real solutions. How to fix this? The same way as always. We extract real solutions from the complex solutions. Fortunately, our previous work on linear independence of complex e-vectors insures that the resulting solution set will be linearly independent.

Proposition 13.5.1.

Let $A \in \mathbb{R}^{n \times n}$. Suppose A has a chain $\{v_1, v_2, \dots, v_k\}$ is of generalized complex e-vectors with e-value $\lambda = \alpha + i\beta$, meaning $(A - \lambda)v_1 = 0$ and $(A - \lambda)v_{k-1} = v_k$ for $k \ge 2$ and $v_j = a_j + ib_j$ for $a_j, b_j \in \mathbb{R}^n$ for each j, then 1. $e^{tA}v_1 = e^{\lambda t}v_1$, 2. $e^{tA}v_2 = e^{\lambda t}(v_2 + tv_1)$, 3. $e^{tA}v_3 = e^{\lambda t}(v_3 + tv_2 + \frac{t^2}{2}v_1)$, 4. $e^{tA}v_k = e^{\lambda t}(v_k + tv_{k-1} + \dots + \frac{t^{k-1}}{(k-1)!}v_1)$. Furthermore, the following are the 2k linearly independent real solutions that are implicit within the complex solutions above, 1. $x_1 = Re(e^{tA}v_1) = e^{\alpha t} [(\cos\beta t)a_1 - (\sin\beta t)b_1]$, 2. $x_2 = Im(e^{tA}v_1) = e^{\alpha t} [(\cos\beta t)(a_2 + ta_1) - (\sin\beta t)(b_2 + tb_1)]$, 3. $x_3 = Re(e^{tA}v_2) = e^{\alpha t} [(\cos\beta t)(a_2 + ta_1) + (\cos\beta t)(b_2 + tb_1)]$, 4. $x_4 = Im(e^{tA}v_2) = e^{\alpha t} [(\cos\beta t)(a_3 + ta_2 + \frac{t^2}{2}a_1) - (\sin\beta t)(b_3 + tb_2 + \frac{t^2}{2}b_1)]$,

6.
$$x_6 = Im(e^{tA}v_3) = e^{\alpha t} \left[(\cos\beta t)(a_3 + ta_2 + \frac{t^2}{2}a_1) - (\sin\beta t)(b_3 + tb_2 + \frac{t^2}{2}b_1) \right].$$

Proof: the magic formula calculations of the last section just as well apply to the complex case. Furthermore, we proved that

$$Re\left[e^{\alpha t + i\beta t}(v + iw)\right] = e^{\alpha t}\left[(\cos\beta t)v - (\sin\beta t)w\right]$$

and

$$Im[e^{\alpha t+i\beta t}(v+iw)] = e^{\alpha t}[(\sin\beta tv + (\cos\beta t)w],$$

the proposition follows. \Box

Example 13.5.2. This example uses the results derived in Example 11.4.12. Let $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and find the e-values and e-vectors of the matrix. Observe that $det(A - \lambda I) = \lambda^2 + 1$ hence the eigevalues are $\lambda = \pm i$. We find $u_1 = [1, i]^T$. Notice that

$$u_1 = \begin{bmatrix} 1\\i \end{bmatrix} = \begin{bmatrix} 1\\0 \end{bmatrix} + i \begin{bmatrix} 0\\1 \end{bmatrix}.$$

This means that $\vec{x}' = A\vec{x}$ has general solution:

$$\vec{x}(t) = c_1 \left(\cos(t) \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \sin(t) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) + c_2 \left(\sin(t) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \cos(t) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)$$

The solution above is the "vector-form of the solution". We can add the terms together to find the scalar solutions: denoting $\vec{x}(t) = [x(t), y(t)]^T$,

$$x(t) = c_1 \cos(t) + c_2 \sin(t)$$
 $y(t) = -c_1 \sin(t) + c_2 \cos(t)$

These are the parametric equations of a circle with radius $R = \sqrt{c_1^2 + c_2^2}$.

Example 13.5.3. We solved the e-vector problem for $A = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ in Example 11.4.14.

We found one real e-value $\lambda_1 = 3$ and a pair of complex e-values $\lambda_2 = 1 \pm i$. The corresponding e-vectors were:

$$\vec{u}_1 = \begin{bmatrix} 0\\0\\1 \end{bmatrix} \qquad \vec{u}_2 = \begin{bmatrix} 0\\1\\0 \end{bmatrix} + i \begin{bmatrix} 1\\0\\0 \end{bmatrix}$$

We identify that $Re(\vec{u}_2) = e_2$ and $Im(\vec{u}_2) = e_1$. The general solution of $\vec{x}' = A\vec{x}$ should have the form:

$$\vec{x}(t) = c_1 e^{At} \vec{u}_1 + c_2 Re(e^{At} \vec{u}_2) + c_3 Im(e^{At} \vec{u}_2)$$

The vectors above are e-vectors so these solution simplify nicely:

$$\vec{x}(t) = c_1 e^{3t} e_3 + c_2 e^t (\cos(t)e_2 - \sin(t)e_1) + c_3 e^t (\sin(t)e_2 + \cos(t)e_1)$$

For fun let's look at the scalar form of the solution. Denoting $\vec{x}(t) = [x(t), y(t), z(t)]^T$,

$$x(t) = -c_2 e^t \sin(t) + c_3 e^t \cos(t), \qquad y(t) = c_2 e^t \cos(t) + c_3 e^t \sin(t), \qquad z(t) = c_1 e^{3t}$$

Believe it or not this is a spiral helix which has an exponentially growing height and radius.

Example 13.5.4. Let's suppose we have a chain of 2 complex eigenvectors \vec{u}_1, \vec{u}_2 with eigenvalue $\lambda = 2 + i3$. I'm assuming that

$$(A - (2+i)I)\vec{u}_2 = \vec{u}_1, \qquad (A - (2+i)I)\vec{u}_1 = 0$$

We get a pair of complex-vector solutions (using the magic formula which truncates since these are *e*-vectors):

$$\vec{z}_1(t) = e^{At}\vec{u}_1 = e^{(2+i)t}\vec{u}_1, \qquad \vec{z}_2(t) = e^{At}\vec{u}_2 = e^{(2+i)t}(\vec{u}_2 + t\vec{u}_1),$$

The real and imaginary parts of these solutions give us 4 real solutions which form the general solution:

$$\begin{aligned} \vec{x}(t) &= c_1 e^{2t} \left[\cos(3t) Re(\vec{u}_1) - \sin(3t) Im(\vec{u}_1) \right] \\ &+ c_2 e^{2t} \left[\sin(3t) Re(\vec{u}_1) + \cos(3t) Im(\vec{u}_1) \right] \\ &+ c_3 e^{2t} \left[\cos(3t) \left[Re(\vec{u}_2) + tRe(\vec{u}_1) \right] - \sin(3t) \left[Im(\vec{u}_2) + tIm(\vec{u}_1) \right] \right] \\ &+ c_4 e^{2t} \left[\sin(3t) \left[Re(\vec{u}_2) + tRe(\vec{u}_1) \right] + \cos(3t) \left[Im(\vec{u}_2) + tIm(\vec{u}_1) \right] \right] \end{aligned}$$

13.6 geometry and difference equations revisited

In Example 11.1.5 we studied $A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$ and how it pushed the point $x_o = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ around the plane. We found x_i for i = 1, 2, 3, 4 by multiplication by A directly. That method is fine for small i but what is we wished to know the formula for the 1000-th state? We should hope there is some way to find that state without direct multiplication repeated 1000 times. One method is to make use of the diagonalization of the matrix. We know that e-vectors (if they exist) can be glued together to make the diagonalizing similarity transforming matrix; there exists $P \in \mathbb{R}^{n \times n}$ such that $P^{-1}AP = D$ where D is a diagonal matrix. Notice that D^k is easy to calculate. We can solve for $A = PDP^{-1}$ and find that $A^2 = PDP^{-1}PDP^{-1} = PD^2P^{-1}$. The you can prove inductively that $A^k = PD^kP^{-1}$. It is much easier to calculate PD^kP^{-1} when k >> 1.

13.6.1 difference equations vs. differential equations

I mentioned that the equation $x_{k+1} = Ax_k$ is a *difference equation*. We can think of this as a differential equation where the time-step is always one-unit. To see this I should remind you how $\vec{x}' = B\vec{x}$ is defined in terms of a limiting process:

$$\vec{x}'(t) = \lim_{h \to 0} \frac{\vec{x}(t+h) - \vec{x}(t)}{h} = B\vec{x}(t)$$

A gross approximation to the continuous limiting process would be to just take h = 1 and drop the limit. That approximation yields:

$$B\vec{x}(t) = \vec{x}(t+1) - \vec{x}(t).$$

We then suppose $t \in \mathbb{N}$ and denote $\vec{x}(t) = \vec{x}_t$ to obtain:

$$\vec{x}_{t+1} = (B+I)\vec{x}_t.$$

We see that the differential equation $\vec{x}' = B\vec{x}$ is crudely approximated by the difference equation $\vec{x}_{t+1} = A\vec{x}_t$. where A = B + I. Since we now have tools to solve differential equations directly it should be interesting to contrast the motion generated by the difference equation to the exact parametric equations which follow from the e-vector solution of the corresponding differential equation.