LECTURE 21: INDEPENDENT SUBSPACES & ORTHOGONAL PROJECTION

We've discussed linear independence of vectors. Now we take it to another level and define independence of subspaces. I'll state the definition for real vector spaces, but our main focus is simply $\mathbb{R}^n$.

Defⁿ/ Let $W_1, W_2, \ldots, W_k$ be subspaces of V then these are __independent__ subspaces if

$$\left(W_1 + \cdots + W_{j-1} + W_{j+1} + \cdots + W_k \right) \cap W_j = \{0\}$$

for $j = 1, 2, \ldots, k$. If $V = W_1 + \cdots + W_k$ and the subspaces are independent then we say that V is the (internal) direct sum of $W_1, \ldots, W_k$ and we write $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$
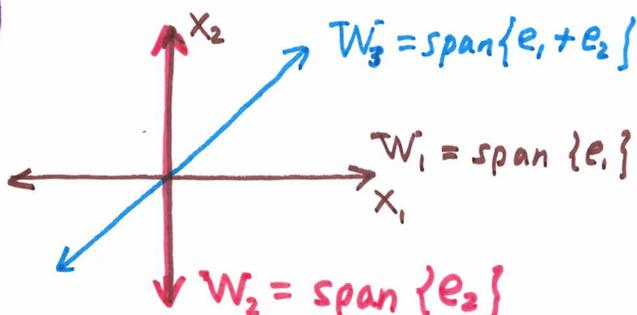
Now, in lecture during classtime I'll probably just explain this for $k = 2$ or 3.

E1 $V = W_1 \oplus W_2$ means $V = W_1 + W_2$ and $W_1 \cap W_2 = \{0\}$

E2 $V = W_1 \oplus W_2 \oplus W_3$ means $V = W_1 + W_2 + W_3$
and $(W_1 + W_2) \cap W_3 = \{0\}$, $(W_1 + W_3) \cap W_2 = \{0\}$ and $(W_2 + W_3) \cap W_1 = \{0\}$.

The following example reveals why pairwise trivial intersection is not sufficient to give direct sum

E3

$W_3 = \text{span}\{e_1 + e_2\}$
$W_1 = \text{span}\{e_1\}$
$W_2 = \text{span}\{e_2\}$

$W_1 + W_2 = \text{span}\{e_1, e_2\} = \mathbb{R}^2$
$W_1 + W_3 = \text{span}\{e_1, e_1 + e_2\} = \mathbb{R}^2$
$W_2 + W_3 = \text{span}\{e_2, e_1 + e_2\} = \mathbb{R}^2$

$W_1 \cap W_2 = W_1 \cap W_3 = W_2 \cap W_3 = \{0\}$

$\left(\dim W_1 + \dim W_2 + \dim W_3 = \dim(W_1 \oplus W_2 \oplus W_3), \text{ doesn't hold} \uparrow \right)$

The previous [E3] does not fit with the following Th⁰ which I present here w/o proof.

Th⁰/ If $W_1, W_2, .., W_k$ are subspaces of V and
$V = W_1 + W_2 + \cdots + W_k$ then the following are equivalent,

(1.) $W_1, W_2, .., W_k$ are independent subspaces

(2.) β_i basis for W_i for $i = 1, 2, .., k$ has
union $\beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ which is basis for V

(3.) for each $x \in V$ there exist unique $x_i \in W_i$
for $i = 1, 2, .., k$ and $X = X_1 + X_2 + \cdots + X_k$

I'm mainly interested in (2.) for this course. Sometime
we know $V = W_1 \oplus W_2$ and in such a case
we can find basis for W_1 and then a basis for W_2
and if we combine those bases then we're sure to
have a basis for V. I used this idea last
class in the last example, but you might not
have noticed, & I found basis for $S^\perp$ without
worrying if I messed up the basis for span (S).

Remark: when we have basis as above in (2.) and
$T : V \longrightarrow V$ linear such that $T(W_j) \subseteq W_j$ for
$j = 1, 2, .., k$ then the matrix of T w.r.t. β is
<u>block diagonal</u> (the $\oplus$ does double-duty here,
we have $V = W_1 \oplus W_2 \oplus + \cdots \oplus W_k$ and

$$[T]_{\beta, \beta} = M_1 \oplus M_2 \oplus \cdots \oplus M_k$$

this is important later on ...

We should begin by reminding the reader that,

Def^n/ Given $S \subseteq \mathbb{R}^n$ the __orthogonal complement__ to S is $S^\perp = \{x \in \mathbb{R}^n \mid x \cdot s = 0 \ \forall s \in S\}$

when $S = \{V_1, V_2, ..., V_k\}$ then we saw $S^\perp = \text{Null}[S]^T$. Let's examine another such example.

EY $S = \{(1, 1, 1, 1), (0, 1, 1, 0)\}$ then $[S] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$

and $[S]^T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \xrightarrow{r_1 - r_2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$ therefore,

$x \in \text{Null}[S]^T$ has $x_1 + x_4 = 0$ and $x_2 + x_3 = 0$ and

$x = (-x_4, -x_3, x_3, x_4) = x_3 \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

Therefore, $S^\perp = \text{span} \{(0, -1, 1, 0), (-1, 0, 0, 1)\}$

we probably could have guessed these w/o calculation, we need some theory to best guide such guessing...

Notice S is __not__ a subspace, yet $S^\perp$ is a subspace of $\mathbb{R}^n$.
In fact, $\text{span}(S) \oplus S^\perp$
(a subspace is closed under addition and scalar multiplication, unless we're talking about $\{0\}$, a __subspace__ has ∞ly many entries)

E5 $W = \text{span} \{(1,1,1,1), (0,1,1,0)\}$ then nothing changes, actually $W^\perp = \text{span} \{(0,-1,1,0), (-1,0,0,1)\}$
and $\underline{W \oplus W^\perp = \mathbb{R}^4}$

this makes W and $W^\perp$ __orthogonal complements__ of one another.

The $\perp$ (I read this as "perp") has nice properties,

Th^m/ Let $S \subseteq \mathbb{R}^n$ then

(1.) $S^\perp \leq \mathbb{R}^n$ ($S^\perp$ is subspace)

(2.) $\underline{\text{span}(S) \cap S^\perp = \{0\}}$; & $\underline{\text{span}(S) + S^\perp = \mathbb{R}^n}$

$$\text{span}(S) \oplus S^\perp = \mathbb{R}^n$$

span(S) and $S^\perp$ are <u>orthogonal</u> <u>complements</u>

(3.) $(S^\perp)^\perp = \text{span}(S)$ and $((\text{span}(S))^\perp)^\perp = \text{span}(S)$

Let's define orthogonality for subspaces and explain the terminology "orthogonal complement". First, an example

EG $W_1 = \text{span}\{(1,1)\}$

$W_2 = \text{span}\{(1,0)\}$

$W_3 = \text{span}\{(1,-1)\}$

$W_1 \oplus W_2 = \mathbb{R}^2$ and $W_1 \oplus W_3 = \mathbb{R}^2$ thus W_1 & W_2 as well as W_1 & W_3 are <u>complementary</u> subspaces

However, $(1,0) \in W_2$, $(1,1) \in W_1$ with $(1,0) \cdot (1,1) = 1 \neq 0$

Whereas $c_1(1,1) \cdot c_2(1,-1) = c_1 c_2 (1-1) = 0$ thus <u>every</u> pair of vectors in W_1 & W_3 are $\perp$. This makes $W_1 \oplus W_3$ an <u>orthogonal decomposition</u> of $\mathbb{R}^2$ and W_1 & W_3 are <u>orthogonal complements</u>; $W_1^\perp = W_3$

$$W_3^\perp = W_1$$

Def^n/ Subspaces $W_1, \ldots, W_k$ are <u>orthogonal</u> if for any $X_i \in W_i$, $X_j \in W_j$ with $i \neq j$ we have $X_i \cdot X_j = 0$. When the sum of the subspaces gives $\mathbb{R}^n$ and the subspaces are orthogonal then

$$\mathbb{R}^n = W_1 \oplus W_2 \oplus \cdots \oplus W_k$$

is an <u>orthogonal decomposition</u> of $\mathbb{R}^n$

W_1

W_2

W_3
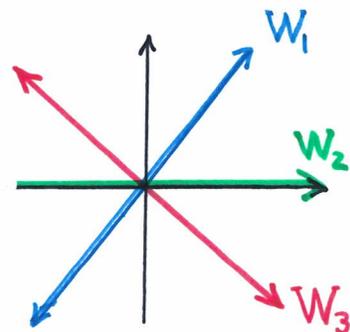
Given an orthogonal decomposition of $\mathbb{R}^n$, say
$W_1 \oplus W_2 \oplus \cdots \oplus W_k = \mathbb{R}^n$ we can parse any
subcollection of the subspaces and those will be
complementary (orthogonally) to the remaining subspaces

$$W_1^\perp = W_2 \oplus \cdots \oplus W_k$$

$$(W_1 \oplus W_2)^\perp = (W_1 + W_2)^\perp = W_3 \oplus W_4 \oplus \cdots \oplus W_k$$

$$(W_1 \oplus W_3 \oplus \cdots \oplus W_k)^\perp = W_2 \oplus W_4 \oplus \cdots \oplus W_{k-1}$$

(assume k odd)

This all ties back into our study of <u>orthonormal</u>
subsets of $\mathbb{R}^n$. To prove the assertions of the last
couple pages I would make ample use of orthonormal sets
and the G.S.A.

Th^m/ If $\beta = \{v_1, \ldots, v_k\}$ is orthogonal to $\gamma = \{w_1, \ldots, w_m\}$
then span $(\beta) \perp$ span (γ)

Proof: Suppose $\beta \perp \gamma$ where β and γ are as stated
above. Let $x \in$ span (β) and $y \in$ span (γ) then
$$x = \sum_{i=1}^{k} c_i v_i \quad \text{and} \quad y = \sum_{j=1}^{m} b_j w_j \quad \text{thus}$$

$$x \cdot y = \left(\sum_{i=1}^{k} c_i v_i \right) \cdot \left(\sum_{j=1}^{m} b_j w_j \right)$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{m} c_i b_j \underbrace{(v_i \cdot w_j)}_{0} \quad \text{since } \beta \perp \gamma.$$

$$= 0 \qquad \therefore \text{ span } (\beta) \perp \text{ span } (\gamma). \; /\!/$$

I believe the following Th^m will be calculationally helpful in that it replaces checking criteria for all points in a subspace with the much simpler task of checking a finite set's properties.

Th^m/ If $\beta_1, \beta_2, ..., \beta_k$ are spanning sets for subspaces $W_1 = \text{span}\,\beta_1$, $W_2 = \text{span}\,\beta_2$, ..., $W_k = \text{span}\,\beta_k$ and $\beta_i \perp \beta_j$ for $i \neq j$ with $1 \leq i,j \leq k$ then $W_i \perp W_j$ for $i \neq j$ with $1 \leq i,j \leq k$. More over, if additionally $|\beta_1| + |\beta_2| + \cdots + |\beta_k| = n$ then $\mathbb{R}^n = W_1 \oplus W_2 \oplus \cdots \oplus W_k$ and $\beta = \beta_1 \cup \beta_2 \cup ... \cup \beta_n$ is a basis for $\mathbb{R}^n$.

[E7] $\beta_1 = \{e_1 + e_2, e_3 + e_4\}$ $\beta_2 = \{e_1 - e_2, e_3 - e_4\}$

and $\beta_3 = \{e_5 + e_8\}$, $\beta_4 = \{e_5 - e_8, e_7, e_6\}$ as subsets of $\mathbb{R}^8$. Observe $\beta_1 \perp \beta_2$ and $\beta_3 \perp \beta_4$ etc and $|\beta_1| + |\beta_2| + |\beta_3| + |\beta_4| = 2 + 2 + 1 + 3 = 8$

Then $\underline{\text{span}(\beta_1) \oplus \text{span}(\beta_2) \oplus \text{span}(\beta_3) \oplus \text{span}(\beta_4) = \mathbb{R}^8}$.

Also, $\underline{\beta = \beta_1 \cup \beta_2 \cup \beta_3 \cup \beta_4}$ is basis for $\mathbb{R}^8$.

<u>Remark</u>: perpendicularity of β_i & β_j gives us independence of the vectors forming β_i & β_j hence forming their union gives us a larger LI set. When sets are not orthogonal things are not so easy.

[E8] $\beta_1 = \{(1,1,0), (1,1,1)\}$, $\beta_2 = \{(1,-1,0), (1,2,3)\}$

both β_1 and β_2 are LI but $\beta_1 \cup \beta_2$ not a basis for $\mathbb{R}^3$ since $|\beta_1 \cup \beta_2| = 4 \Rightarrow \beta_1 \cup \beta_2$ <u>not</u> LI.

Th^m/ If $W = \text{span}(\beta_W)$ where $\beta_W = \{W_1, ..., W_n\}$ is an orthonormal subset of $\mathbb{R}^n$ then there exists an orthonormal basis $\beta = \{W_1, ... W_n, W_{n+1}, .., W_n\}$ for $\mathbb{R}^n$ where $W^\perp = \text{span}(\beta - \beta_W) = \text{span}\{W_{n+1}, ..., W_n\}$

Proof: given β_W is LI we have $[\beta_W]$ has k-LI columns and it follows $\text{Null}[\beta_W]^T = W^\perp$ has a basis γ containing $(n-k)$-vectors. Apply the G.S.A. to $\beta_W \cup \gamma = \beta$ and the Th^m follows. //

This proof is actually how we calculate $W^\perp$ given the basis β for W, we use $W^\perp = \text{Null}[\beta]^T$. Here β_W being orthonormal just insures the G.S.A. does not alter β_W as it runs through $\beta = \beta_W \cup \gamma$.

$\boxed{E9}$ $W = \text{span}\{\underbrace{(1,1,1,1), (0,1,-1,0)}_{"\beta_W"}\}$

$$[\beta_W]^T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & -1 & 0 \end{bmatrix} = \text{rref}[\beta_W]^T$$

$\therefore$ $X \in \text{Null}[\beta]^T \Rightarrow X = (-2X_3 - X_4, X_3, X_3, X_4)$
$$= X_3(-2,1,1,0) + X_4(-1,0,0,1)$$

Here $\gamma = \{(-2,1,1,0), (-1,0,0,1)\} \perp \beta_W$.

Then to create orthonormal basis for $\mathbb{R}^4$ which is based on β_W and γ above we'd need to

① normalize $\beta_W \rightarrow \{\frac{1}{2}(1,1,1,1), \frac{1}{\sqrt{2}}(0,1,-1,0)\}$

② Run G.S.A. on γ to obtain
$$\{\frac{1}{\sqrt{6}}(-2,1,1,0), \frac{1}{\sqrt{12}}(-1,-1,-1,3)\}$$

ORTHOGONAL PROJECTIONS

Suppose we have a subspace W of dimension k which is a subset of $\mathbb{R}^n$. Then we can use the G.S.A. to orthonormalize the basis for W. We must use an orthonormal basis for W to formulate the <u>orthogonal</u> projection.

Def:/ Suppose $W = \text{span}(\beta_w)$ where $\beta_w = \{W_1, \ldots, W_k\}$ is an orthonormal subset of $\mathbb{R}^n$ then for $X \in \mathbb{R}^n$,

$$\text{Proj}_W(x) = \sum_{j=1}^{k} (x \cdot W_j) W_j$$

and then,

$$\text{Orth}_W(x) = x - \text{Proj}_W(x)$$

[E10] Let's build off E9, we found orthonormal bases for W and $W^{\perp}$

$$W = \text{span}\left\{ \frac{1}{2}(1,1,1,1), \frac{1}{\sqrt{2}}(0,1,-1,0) \right\}$$

$$W^{\perp} = \text{span}\left\{ \frac{1}{\sqrt{6}}(-2,1,1,0), \frac{1}{\sqrt{12}}(-1,-1,-1,3) \right\}$$

Therefore, for $X = (a,b,c,d) \in \mathbb{R}^4$ we calculate,

$$\text{Proj}_W(x) = (x \cdot W_1)W_1 + (x \cdot W_2)W_2$$

$$= \left(\frac{a+b+c+d}{2}\right)W_1 + \left(\frac{b-c}{\sqrt{2}}\right)W_2$$

$$= \frac{a+b+c+d}{4}(1,1,1,1) + \frac{b-c}{2}(0,1,-1,0) \quad *$$

Likewise,

$$\text{Proj}_{W^{\perp}}(x) = (x \cdot W_3)W_3 + (x \cdot W_4)W_4$$

$$= \frac{-2a+b+c}{6}(-2,1,1,0) + \frac{-a-b-c+3d}{12}(-1,-1,-1,3) \quad **$$

You can check, $\text{Proj}_W(x) + \text{Proj}_{W^{\perp}}(x) = (a,b,c,d) = X.$

E10 continued

We found $\text{Proj}_W + \text{Proj}_{W^\perp} : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ is in fact the __identity operator__; $\text{Proj}_W + \text{Proj}_{W^\perp} = Id_{\mathbb{R}^4}$.

$$\text{Orth}_W(x) = (a,b,c,d) - \text{Proj}_W(a,b,c,d)$$

$$= (a,b,c,d) - \left(\frac{a+b+c+d}{4}\right)(1,1,1,1) - \left(\frac{b-c}{2}\right)(0,1,-1,0)$$

$$= \text{Proj}_{W^\perp}(x)$$

I know this had to happen because the theory of Proj and Orth is very simple,

Th^m/ Suppose $\beta = \{W_1, W_2, .., W_k, W_{k+1}, .., W_n\}$ is orthonormal basis for $\mathbb{R}^n$ and $W = \text{span}\{W_1, ..., W_k\}$ then, for $x \in \mathbb{R}^n$,

1.) $\text{Orth}_W(x) = \sum\limits_{j=k+1}^{n}(x \cdot W_j)W_j = \text{Proj}_{W^\perp}(x)$

2.) $\text{Proj}_W(x) + \text{Proj}_{W^\perp}(x) = x$ and $\text{Proj}_W(x) + \text{Orth}_W(x) = x$

$\text{Proj}_W(x) \cdot \text{Proj}_{W^\perp}(x) = 0$ and $\text{Proj}_W(x) \cdot \text{Orth}_W(x) = 0$

3.) $\|x\|^2 = \|\text{Proj}_W(x)\|^2 + \|\text{Orth}_W(x)\|^2$

4.) $\text{Proj}_W : \mathbb{R}^n \longrightarrow W \subseteq \mathbb{R}^n$ is linear transformation and $\text{Proj}_W \circ \text{Proj}_W = \text{Proj}_W$.

__Proof:__ mostly follows from β orthonormal since $x \in \mathbb{R}^n$ has $x = \sum\limits_{j=1}^{n}(x \cdot W_j)W_j = \underbrace{\sum\limits_{j=1}^{k}(x \cdot W_j)W_j}_{\text{Proj}_W(x)} + \underbrace{\sum\limits_{j=k+1}^{n}(x \cdot W_j)W_j}_{\text{Proj}_{W^\perp}(x)}$

$\left(\begin{array}{l}\text{since } \{W_1, ..., W_k\} \text{ basis for } W \\ \text{and } \{W_{k+1}, .., W_n\} \text{ basis for } W^\perp\end{array}\right)$

I'll leave the proof of $\text{Proj}_W(x) \perp \text{Orth}_W(x)$ for next lecture ☺.