Lecture 31: Diagonalization of Quadratic Forms

Let's begin with what we mean by "quadratic form"

Def^n/ If $Q : \mathbb{R}^n \to \mathbb{R}$ is given by $Q(v) = v^T A v$ where A is a symmetric matrix then we say Q is a quadratic form with matrix $[Q] = A$.

We insist the matrix of Q is symmetric so that we know $A = [Q]$ is orthonormally diagonalizable.

$\boxed{E1}$ $$Q(x,y) = x^2 + y^2 + 2xy = [x, y] \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}}_{[Q]} \begin{bmatrix} x \\ y \end{bmatrix}$$

$\boxed{E2}$ $$Q(x,y,z) = x^2 + 2y^2 + 3z^2 + 6xz - 4yz$$
$$= [x, y, z] \underbrace{\begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & -2 \\ 3 & -2 & 3 \end{bmatrix}}_{[Q]} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$\boxed{E3}$ $Q(x) = x_1^2 + 3x_3^2 + 5x_5^2 - x_1 x_5$ for $x \in \mathbb{R}^5$

gives $$[Q] = \begin{bmatrix} 1 & 0 & 0 & 0 & -1/2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -1/2 & 0 & 0 & 0 & 5 \end{bmatrix}$$ and $\underline{Q(x) = x^T [Q] x}$.

$\boxed{E4}$ $$Q(x,y,z) = x^2 - y^2 + 3xz$$
$$= [x, y, z] \underbrace{\begin{bmatrix} 1 & 0 & 3/2 \\ 0 & -1 & 0 \\ 3/2 & 0 & 0 \end{bmatrix}}_{[Q]} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Thm/ If $Q : \mathbb{R}^n \to \mathbb{R}$ is quadratic form with matrix $A = [Q]$ and $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n$ is an orthonormal eigenbasis for A which defines eigencoordinates $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) = \bar{x}$ by $\bar{x} = [\bar{\beta}]^T x$ where $\bar{\beta} = \{\bar{v}_1, \dots, \bar{v}_n\}$ then

$$Q(x) = \lambda_1 \bar{x}_1^2 + \lambda_2 \bar{x}_2^2 + \cdots + \lambda_n \bar{x}_n^2$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are e-values of $v_1, \dots, v_n$ respective.

Proof: Suppose $Q(x) = x^T A x$ where $A^T = A$. Let $\bar{\beta} = \{\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n\}$ be the orthonormal e-basis with e-values $\lambda_1, \lambda_2, \dots, \lambda_n$ for A as given by the real spectral Thm. Recall $[\bar{\beta}]^T A [\bar{\beta}] = \mathrm{Diag}(\lambda_1, \dots, \lambda_n) = D$. Consider,

$$\begin{aligned}
Q(x) &= x^T A x \\
&= x^T [\bar{\beta}][\bar{\beta}]^T A [\bar{\beta}][\bar{\beta}]^T x \\
&= ([\bar{\beta}]^T x) D ([\bar{\beta}]^T x) \\
&= \bar{x}^T D \bar{x} \\
&= [\bar{x}_1, \dots, \bar{x}_n] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_n \end{bmatrix} \\
&= \underline{\lambda_1 \bar{x}_1^2 + \lambda_2 \bar{x}_2^2 + \cdots + \lambda_n \bar{x}_n^2}. //
\end{aligned}$$

Level Curves of the form $ax^2+2bx+cy^2=k$

$$x^2/a^2 + y^2/b^2 = 1 \quad \text{ellipse}$$

$$x^2/a^2 - y^2/b^2 = \pm 1 \quad \text{hyperbola}$$

$$x^2/a^2 = 1 \quad \text{pair of lines (vertical)}$$

$$y^2/b^2 = 1 \quad \text{pair of horizontal lines}$$

The curves above show the different cases which can arise for the solution set of $ax^2+2bx+cy^2=1$. If we allow $k=0$ or $k=-1$ we may get no solution ($x^2+y^2=-1$) or pair of lines ($x^2-y^2=0$ gives $y=\pm x$). Observe

$$ax^2+2bx+cy^2 = [x, y]\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \underbrace{\lambda_1 \bar{x}^2 + \lambda_2 \bar{y}^2}_{}$$

we can relate $ax^2+2bx+cy^2=1$ to either an ellipse, hyperbola or pair of lines. If we picture $\bar{x}, \bar{y}$ - coordinate axes then sketching is easy on basis of past graphing of ellipses, hyperbolas and lines.

$\boxed{E5}$ $Q(x,y) = 4xy = 1$

Write the equation above using eigencoordinates $(\bar{x}, \bar{y})$ and graph the $\bar{x}, \bar{y}$ axes and the curve, what is it?

$$A = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} \longrightarrow \det\begin{bmatrix} -\lambda & 2 \\ 2 & -\lambda \end{bmatrix} = \lambda^2 - 4 = (\lambda+2)(\lambda-2) = 0$$

$$\therefore \underline{\lambda_1 = -2, \ \lambda_2 = 2}.$$

$\underline{\lambda_1 = -2}$ $\quad A + 2I = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \longrightarrow v_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ (e-vector with $\lambda_1 = -2$)

$\underline{\lambda_2 = 2}$ $\quad A - 2I = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \longrightarrow v_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ ($\lambda_2 = 2$ e-vector)

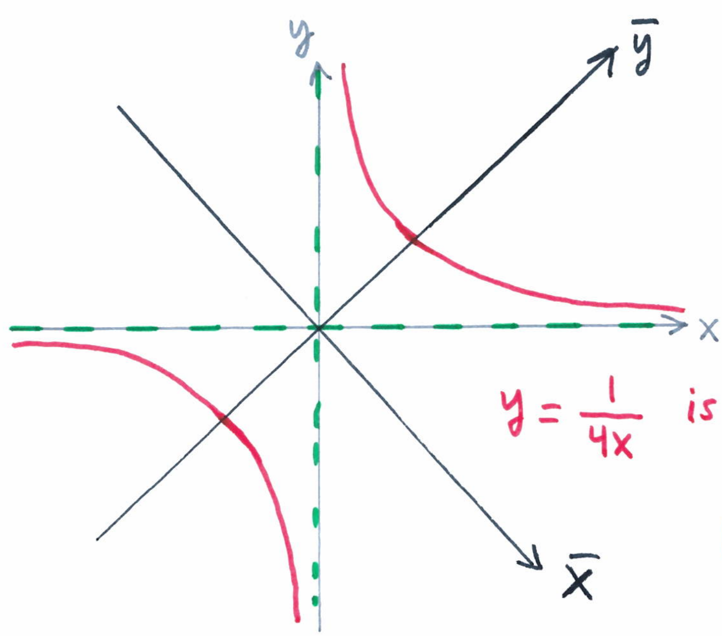
$$[\beta] = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \Rightarrow [\beta]^{-1} = [\beta]^T = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Thus $\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} x - y \\ x + y \end{bmatrix}$, $\bar{x} = \frac{x-y}{\sqrt{2}}$, $\bar{y} = \frac{x+y}{\sqrt{2}}$

and $\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} \bar{x} + \bar{y} \\ -\bar{x} + \bar{y} \end{bmatrix}$, $x = \frac{\bar{x}+\bar{y}}{\sqrt{2}}$, $y = \frac{\bar{y}-\bar{x}}{\sqrt{2}}$

Observe $4xy = 1 \longrightarrow 4\left[\frac{\bar{x}+\bar{y}}{\sqrt{2}}\right]\left[\frac{\bar{y}-\bar{x}}{\sqrt{2}}\right] = \underbrace{2(\bar{y}^2 - \bar{x}^2) = 1}$

hyperbola, with asymptotes $\bar{y} = \pm\bar{x}$ a.k.a. $y = 0, x = 0$.



$y = \frac{1}{4x}$ is a hyperbola

(**Remark**: $\bar{x} > 0$ in v_1-direction, $\bar{y} > 0$ in v_2-direction)

$\boxed{E6}$ Consider $Q(x,y) = x^2 - 2xy + 3y^2 = 1$ describe the curve, no need to explicitly change coordinates, but find the formula for $Q(x,y) = 1$ in eigencoordinates $\bar{x}, \bar{y}$

$$[Q] = \begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix}$$

$$\det \begin{bmatrix} 1-\lambda & -1 \\ -1 & 3-\lambda \end{bmatrix} = (\lambda-1)(\lambda-3) - 1 = \lambda^2 - 4\lambda + 2$$
$$= (\lambda-2)^2 - 2$$
$$= (\lambda - 2 + \sqrt{2})(\lambda - 2 - \sqrt{2})$$

We find $\lambda_1 = 2 - \sqrt{2}$

$\lambda_1 = 2 - \sqrt{2}$ and $\lambda_2 = 2 + \sqrt{2}$

Thus,

$$Q(x,y) = \underbrace{(2-\sqrt{2})\bar{x}^2 + (2+\sqrt{2})\bar{y}^2 = 1}_{\boxed{\text{ellipse}} \text{ since } \lambda_1, \lambda_2 > 0.}$$

$\boxed{E7}$ Describe the curve $x^2 + 4xy - 3y^2 = 1$

$$x^2 + 4xy - 3y^2 = [x, y] \underbrace{\begin{bmatrix} 1 & 2 \\ 2 & -3 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 2 \\ 2 & -3-\lambda \end{pmatrix} = (\lambda-1)(\lambda+3) - 4$$
$$= \lambda^2 + 2\lambda - 7$$
$$= (\lambda+1)^2 - 8$$
$$= (\lambda + 1 + \sqrt{8})(\lambda + 1 - \sqrt{8})$$

Thus $\lambda = -1 \pm \sqrt{8}$ are e-values for A,

$$x^2 + 4xy - 3y^2 = \underbrace{(\sqrt{8} - 1)\bar{x}^2 + (-1-\sqrt{8})\bar{y}^2 = 1}_{\boxed{\text{hyperbola}}}$$

Optimization of Multivariate Functions

If we consider $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ then the multivariate Taylor's Th$\underline{\text{m}}$ gives a multivariate power series for $f(x)$ about some point $P \in \mathbb{R}^n$.

$$f(x) = f(P) + \nabla f(P) \cdot (x-P) + \frac{1}{2} \overbrace{Q(x-P)}^{\text{quadratic form!}} + \cdots$$

Here we have just written the expansion to 2nd order,

$$\nabla f(P) = \left\langle \frac{\partial f}{\partial x_1}(P), \frac{\partial f}{\partial x_2}(P), \ldots, \frac{\partial f}{\partial x_n}(P) \right\rangle \quad \text{(Gradient)}$$

$$[Q] = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(P) \right] \quad \text{(Hessian Matrix)}$$

Extreme values for f occur locally at points P where $\nabla f(P) = 0$ (critical point). If we write $x - P = h$ at such a point then

$$\begin{aligned} f(P+h) &= f(P) + \frac{1}{2} Q(h) + \cdots \\ &= f(P) + \frac{1}{2} \underbrace{\left(\lambda_1 \bar{h}_1^2 + \lambda_2 \bar{h}_2^2 + \cdots + \lambda_n \bar{h}_n^2 \right)} + \cdots \end{aligned}$$

$f(P)$ is local max if $\lambda_1, \lambda_2, .., \lambda_n < 0$

$f(P)$ is local min if $\lambda_1, \lambda_2, .., \lambda_n > 0$

If any $\lambda_j = 0$ then we cannot use quadratic analysis to discern nature of extremal point.

The $n=2$ case is covered in Calculus III,

$$[Q] = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} = A$$

$$\det \begin{bmatrix} f_{xx}-\lambda & f_{xy} \\ f_{xy} & f_{yy}-\lambda \end{bmatrix} = (\lambda - f_{xx})(\lambda - f_{yy}) - f_{xy}^2$$
$$= \lambda^2 - (f_{xx}+f_{yy})\lambda + f_{xx}f_{yy} - f_{xy}^2$$

$$\det(A) = f_{xx}f_{yy} - f_{xy}^2 = \lambda_1\lambda_2$$
$$\text{trace}(A) = f_{xx} + f_{yy} = \lambda_1 + \lambda_2$$

LOCAL MINIMUM: $\lambda_1, \lambda_2 > 0 \Rightarrow \lambda_1\lambda_2 = D > 0$

LOCAL MAXIMUM: $\lambda_1, \lambda_2 < 0 \Rightarrow \lambda_1\lambda_2 = D > 0$

SADDLE PT.: $\left.\begin{matrix} \lambda_1 < 0,\ \lambda_2 > 0 \\ \lambda_1 > 0,\ \lambda_2 < 0 \end{matrix}\right) \lambda_1\lambda_2 = D < 0$

$\lambda_1, \lambda_2 > 0 \Rightarrow f_{xx} + f_{yy} > 0 \Rightarrow f_{xx} > 0$ (min.)

$\lambda_1, \lambda_2 < 0 \Rightarrow f_{xx} + f_{yy} < 0 \Rightarrow f_{xx} < 0$ (max.)

We recover the usual 2^{nd} derivative test from CALC III. Notice, $n \geq 3$ allows no such nice test since $\det[Q] = \lambda_1\lambda_2\lambda_3$ positive or negative does not nicely reveal signs of $\lambda_1, \lambda_2, \lambda_3$ separately. But, with linear algebra we get the test's replacement: find e-values of $[Q]$ if all positive then min, if all negative then max.

CONSTRAINED OPTIMIZATION ON n-sphere

The unit-circle is the 1-sphere S^1 given by $x^2+y^2=1$.
The unit-sphere is the 2-sphere S^2 given by $x^2+y^2+z^2=1$
The 3-sphere S^3 is given by $x_1^2+x_2^2+x_3^2+x_4^2=1$

Generally,

$$\boxed{\text{Def}^{n}/\quad S^n = \{ x \in \mathbb{R}^{n+1} \mid x \cdot x = 1 \}} \subseteq \mathbb{R}^{n+1}$$

Th^{m}/ any quadratic form on $\mathbb{R}^{n+1}$ is uniquely extended from its values on S^n. Moreover, the maximum and minimum values of $Q : \mathbb{R}^{n+1} \to \mathbb{R}$ restricted to S^n are specified by the largest and smallest eigenvalues of $[Q]$

<u>Proof</u>: given any $x \in \mathbb{R}^{n+1}$ with $x \neq 0$ we can write $x = \|x\| \hat{x}$ where $\hat{x} = \frac{1}{\|x\|} x$ is the unit-vector in the direction of x. Notice $\hat{x} \cdot \hat{x} = \frac{1}{\|x\|^2} x \cdot x = \frac{x \cdot x}{x \cdot x} = 1$, thus $\hat{x} \in S^n$. If $Q(x) = x^T A x$ then,

$$\begin{aligned} Q(x) &= (\|x\| \hat{x})^T A \|x\| \hat{x} \\ &= \|x\|^2 \hat{x}^T A \hat{x} \\ &= \|x\|^2 Q(\hat{x}) \end{aligned}$$

values of $Q|_{S^n}$ fix the action of Q on all of space.

Since $A^T = A$ the real spectral Th^{m} gives unit-vectors $v_1, v_2, \dots, v_{n+1} \in S^n$ and so if $u \in S^n$ and $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{n+1}$

$$Q(u) = \lambda_1 \bar{u}_1^2 + \lambda_2 \bar{u}_2^2 + \dots + \lambda_{n+1} \bar{u}_{n+1}^2$$

has $Q(\pm v_1) = \lambda_1 \leq Q(u) \leq Q(\pm v_{n+1}) = \lambda_{n+1}$.

$\boxed{E8}$ $Q(v) = v^TAv$ with $A = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 0 \end{bmatrix}$

has eigenvectors $v_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $v_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $v_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$

with $\lambda_1 = -1$, $\lambda_2 = 0$, $\lambda_3 = 3$. It follows that $\boxed{-1 \leq 2y^2 - 2xy + 2xz - 2yz \leq 3}$

for all $(x, y, z) \in S^2$. We can check,

$v = (x, y, z) = \frac{\pm 1}{\sqrt{2}}(1, 0, -1)$ gives $Q(v) = \frac{2(-1)(1)}{2} = -1$.

$v = \frac{\pm 1}{\sqrt{6}} v_3 = \frac{\pm 1}{\sqrt{6}}(1, -2, 1)$ gives $Q(v) = \frac{8+4+2+4}{6} = 3$.

$\boxed{E9}$ $Q(x, y, z) = 11x^2 + 11y^2 + 3z^2 - 12xy + 4xz + 4yz$

has matrix $[Q] = \begin{bmatrix} 11 & -6 & 2 \\ -6 & 11 & 2 \\ 2 & 2 & 3 \end{bmatrix}$ with $\lambda_1 = 1$,

$\lambda_2 = 7$ and $\lambda_3 = 17$. It follows that

$$\boxed{1 \leq 11x^2 + 11y^2 + 3z^2 - 12xy + 4xz + 4yz \leq 17}$$

for all $(x, y, z) \in S^2$.

<u>Remark</u>: Since $Q(cx) = (cx)^TAcx = c^2x^TAx = c^2\, Q(x)$ it follows $Q(\|x\|\hat{x}) = \|x\|^2 Q(\hat{x}) = Q(x)$ so for $\|x\| \leq 1$ we find $Q(x) = \|x\|^2 Q(\hat{x}) \leq Q(\hat{x})$. This means the inequalities in $\boxed{E8}$ and $\boxed{E9}$ extend to the interior of S^2, that is $\underbrace{0 \leq x^2 + y^2 + z^2 \leq 1}_{\text{unit-ball}}$.

Of course this is also true for higher dim'l examples, the e-values for $Q(x) = x^TAx$ give min/max for values of $Q(x)$ for any $x \in \mathbb{R}^{n+1}$ with $\|x\| \leq 1$.