ECEn 671: Mathematics of Signals and Systems

Randal W. Beard

Brigham Young University

September 1, 2023

Section 1

Quadratic Forms

Quadratic Forms

Definition

A real square matrix is <u>symmetric</u> if $A^\top = A$

Definition

A real square matrix is <u>skew-symmetric</u> if $A^\top = -A$

Quadratic Forms

Lemma

Any real square matrix $B \in \mathbb{R}^{n \times n}$ *can be written as*

$$B = B_s + B_{ss}$$

where B_s *is symmetric and* B_{ss} *is skew-symmetric.*

Proof.

$$B = \frac{B + B^\top}{2} + \frac{B - B^\top}{2} \triangleq B_s + B_{ss}$$

where

$$B_s^\top = \left(\frac{B + B^\top}{2}\right)^\top = \frac{B^\top - B}{2} = \frac{B + B^\top}{2} = B_s$$

$$B_{ss}^\top = \left(\frac{B - B^\top}{2}\right)^\top = \frac{B^\top - B}{2} = -\left(\frac{B - B^\top}{2}\right) = -B_{ss}$$

Quadratic Forms

Lemma

For any real square matrix A and for all y

$$y^\top A y = y^\top A_s y$$

where A_s is the symmetric part of A.

Proof.

$$y^\top A y = y^\top A_s y + y^\top A_{ss} y$$

but

$$y^\top A_{ss} y = y^\top \left(\frac{A - A^\top}{2} \right) y = \frac{1}{2} y^\top A y - \frac{1}{2} y^\top A^\top y.$$

But since

$$y^\top A^\top y = (y^\top A^\top y)^\top = y^\top A y \implies y^\top A_{ss} y = 0.$$

Quadratic Forms

Definition

A quadratic form of a real square matrix A is $Q_A(y) = \mathbf{y}^\top A\mathbf{y}$.

w.l.o.g. A can be assumed to be symmetric. If not, we can always limit our attention to the symmetric part of A since

$$\mathbf{y}^\top A\mathbf{y} = \mathbf{y}^\top A_s\mathbf{y}.$$

Quadratic forms show up in numerous places. For example, the pdf for a Gaussian random variable is

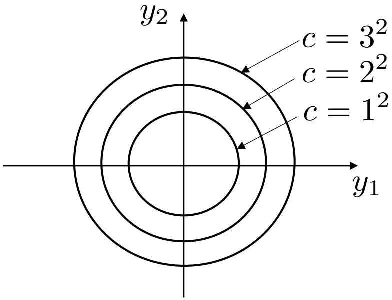
$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right).$$

Quadratic Forms

Example

Let

$$Q_A(\mathbf{y}) = \mathbf{y}^\top \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{y} = y_1^2 + y_2^2 = c$$



The level curves of $Q_A(\mathbf{y})$ are circles of radius $\sqrt{c}$.

Quadratic Forms

Example

Consider the quadratic equation

$$f(x) = 2y_1^2 + 3y_1y_2 + 4y_2^2,$$

and note that

$$\begin{aligned} f(x) &= 2y_1^2 + 3y_1y_2 + 4y_2^2 \\ &= \begin{pmatrix} y_1 & y_2 \end{pmatrix} \begin{pmatrix} 2y_1 + \frac{3}{2}y_2 \\ 4y_2 + \frac{3}{2}y_1 \end{pmatrix} \\ &= \begin{pmatrix} y_1 & y_2 \end{pmatrix} \begin{pmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & 4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ &= \mathbf{y}^\top A\mathbf{y} \end{aligned}$$

Any quadratic equation in n variables can be written in the form $\mathbf{y}^\top A\mathbf{y}$.

Quadratic Forms

By the spectral theorem, A is diagonalizable. In other words, there exists an invertible U so that $A = U\Lambda U^\top$.

From Moon Lemma 6.2 the eigenvalues are real so we can order them as

$$\Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

with

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n.$$

Quadratic Forms

Lemma

Level curves of the quadratic form

$$Q_A(x - x_0) = (x - x_0)^\top A(x - x_0) = c$$

are hyper-ellipsoids with the length of the axes given by $\frac{1}{\sqrt{\lambda_i}}$.

Proof.

Let $z = U^\top y$ then

$$\begin{aligned} Q_A(y) &= y^\top Ay = y^\top U \Lambda U^\top y = z^\top \Lambda z \\ &= \begin{pmatrix} z_1 & \cdots & z_n \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \\ &= \lambda_1 z_1^2 + \cdots + \lambda_n z_n^2 \end{aligned}$$

Quadratic Forms

Note that in the variable z, the quadratic form is an ellipsoid:

$$Q_A(\mathbf{y}) = (\sqrt{\lambda_1})^2 z_1^2 + (\sqrt{\lambda_2})^2 z_2^2 + \cdots + (\sqrt{\lambda_n})^2 z_n^2 = 1$$

or

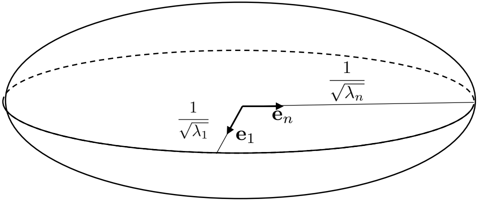
$$Q_A(\mathbf{y}) = \frac{z_1^2}{\left(\frac{1}{\sqrt{\lambda_1}}\right)^2} + \frac{z_2^2}{\left(\frac{1}{\sqrt{\lambda_2}}\right)^2} + \cdots + \frac{z_n^2}{\left(\frac{1}{\sqrt{\lambda_n}}\right)^2} = 1$$

Either of these are the general equation for an ellipsoid with minor axis $\mathbf{e}_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}^\top$ and major axis $\mathbf{e}_n = \begin{pmatrix} 0 & \cdots & 0 & 1 \end{pmatrix}^\top$

Quadratic Forms

Note that along $\mathbf{e}_1$, the stretching is

$$(\sqrt{\lambda_1})^2 z_1^2 = 1 \Rightarrow z_1 = \frac{1}{\sqrt{\lambda_1}}$$



In the original space, what is $\mathbf{e}_1$?

$$\mathbf{e}_1 = U^\top \mathbf{y} = \begin{pmatrix} \mathbf{u}_1^\top \\ \vdots \\ \mathbf{u}_n^\top \end{pmatrix} \mathbf{y}$$

$$= \begin{pmatrix} \mathbf{u}_1^\top \mathbf{y} \\ \vdots \\ \mathbf{u}_n^\top \mathbf{y} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Therefore $\mathbf{y} = \mathbf{u}_1$ since U is orthogonal.
i.e. $\mathbf{u}_i = U\mathbf{e}_i$.

Quadratic Forms

Therefore the major axis is given by the eigenvector associated with the smallest eigenvalue, and the minor axis is given by the eigenvector associated with the largest eigenvalue.

Question: What is the geometric picture associated with

$$(x - x_0)^\top A(x - x_0) = c$$

where c is a constant and A is symmetric and positive definite?

Answer: An ellipsoid of radius $\sqrt{c}$ centered at x_0 with axes along the eigenvectors of A and stretching along each axis given by $\dfrac{1}{\sqrt{\lambda_i}}$.

Quadratic Forms

Question: What if we would like to maximize

$$Q_A(y) = y^\top Ay \text{ where } \|y\| = 1.$$

Which axis provides the most bang-for-the-buck?

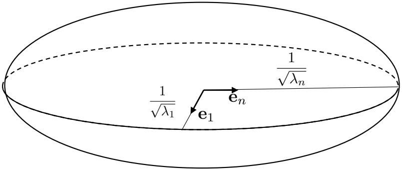
Answer: The <u>major</u> axis! i.e. the axis associated with the largest eigenvalue.

Quadratic Forms

Rather than drawing

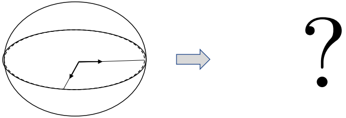
$$\lambda_1 z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_n z_n^2 = 1$$

which is



lets draw the mapping of the unit circle through $\mathbf{y}^\top A\mathbf{y}$ i.e.

$$\{\|\mathbf{y}\| = 1\} \xrightarrow{Q_A(\mathbf{y})} \{\mathbf{y}^\top A\mathbf{y}\}$$



Quadratic Forms

If $A = A^\top$ then $A = U\Lambda U^\top$ where U is orthogonal, i.e., $UU^\top = U^\top U = I$. Then

$$\max_{\|\mathbf{y}\|=1} \mathbf{y}^\top A\mathbf{y} = \max_{\|\mathbf{y}\|=1} \mathbf{y}^\top U\Lambda U^\top \mathbf{y}.$$

Let $\mathbf{z} = U^\top \mathbf{y}$ and note that $\|\mathbf{z}\| = \left\|U^\top \mathbf{y}\right\| = \|\mathbf{y}\|$ since U is orthogonal. Then

$$\begin{aligned} \max_{\|\mathbf{y}\|=1} \mathbf{y}^\top U\Lambda U^\top \mathbf{y} &= \max_{\|\mathbf{z}\|=1} \mathbf{z}^\top \Lambda \mathbf{z} \\ &= \max_{\|\mathbf{z}\|=1} \left(\lambda_1 z_1^1 + \lambda_2 z_2^2 + \cdots + \lambda_n z_n^2\right) \end{aligned}$$

where Λ is arranged such that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n.$$

Quadratic Forms

The maximum is therefore

$$\mathbf{z}^* = \begin{pmatrix} 1 & 0 & \vdots & 0 \end{pmatrix}^\top$$

where it is clear that $\|\mathbf{z}\| = 1$.
Furthermore

$$\max_{\|\mathbf{z}\|=1} \mathbf{z}^\top \Lambda \mathbf{z} = \lambda_1,$$

which implies that

$$\mathbf{y}^* = U\mathbf{z}^* = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{u}_1.$$

Quadratic Forms

This mapping also forms an ellipsoid but with a different effect.
Let $\mathbf{y} = \mathbf{u}_1 \implies \|\mathbf{y}\| = 1$ to get

$$Q_A(\mathbf{y}) = \lambda_1$$

Question: Is it possible to pick a $\hat{\mathbf{y}}$ where $\|\hat{\mathbf{y}}\| = 1$ such that

$$Q_A(\hat{\mathbf{y}}) > Q_A(\mathbf{u}_1)?$$

Answer: No.

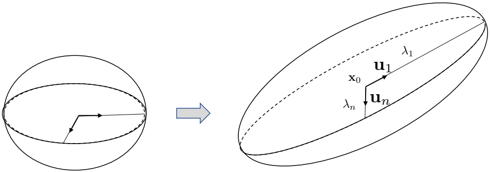
Quadratic Forms

Explanation: Recall that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ and $\|\hat{\mathbf{y}}\| = y_1^2 + \cdots + y_n^2 = 1$.

Therefore

$$\begin{aligned} Q_A(\hat{\mathbf{y}}) &= \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \\ &\leq \lambda_1 y_1^2 + \lambda_1 y_2^2 + \cdots + \lambda_1 y_n^2 \\ &= \lambda_1 \|\hat{\mathbf{y}}\|^2 \\ &= Q_A(\mathbf{u}_1) \end{aligned}$$

So the mapping of the unit circle looks like



Quadratic Forms

We have essentially proved the following theorem:

Theorem (Moon Theorem 6.5)

For a positive semi-definite Hermitian matrix A, the maximum

$$\lambda_1 = \max_{\|\mathbf{x}\|_2=1} \mathbf{x}^H A\mathbf{x}$$

where λ_1 is the largest eigenvalue of A, and the maximizing $\mathbf{x}$ is $\mathbf{x} = \mathbf{u}_1$, the associated eigenvector.
Furthermore if we maximize $\mathbf{x}^H A\mathbf{x}$ subject to the constraints

$$\begin{aligned} \langle \mathbf{x}, \mathbf{u}_i \rangle &= 0 \qquad i = 1, \ldots, k-1, \\ \|\mathbf{x}\|_2 &= 1 \end{aligned}$$

then the maximum is λ_k and $\mathbf{x}_{max} = \mathbf{u}_k$.

Quadratic Forms

Note that if A is positive semi-definite Hermitian then

$$\|A\|_2 = \sup_{\|\mathbf{x}\|_2 \neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \max_{\|\mathbf{x}\|_2 = 1} \sqrt{\mathbf{x}^H A^H A\mathbf{x}} = \sqrt{\lambda_1 \mathbf{u}_1^H \mathbf{u}_1} = \sqrt{\lambda_1}$$

where λ_1 is the largest eigenvalue of $A^H A$.

More generally,

$$R(x) = \frac{\mathbf{x}^\top A\mathbf{x}}{\mathbf{x}^\top \mathbf{x}}$$

is called a Rayleigh quotient and

$$\max_{\|\mathbf{x}\| \neq 0} R(\mathbf{x}) = \lambda_1$$

$$\min_{\|\mathbf{x}\| \neq 0} R(\mathbf{x}) = \lambda_n.$$