Quadratic forms

Aim lecture: We use the spectral thm for self-adjoint operators to study solns to some multi-variable quadratic eqns.

In this lecture, we work over the real field $\mathbb{F} = \mathbb{R}$. We use the notn $\mathbf{x} = (x_1, \ldots, x_n), \mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$

Defn

A *(real) quadratic form* is a function $Q : \mathbb{R}^n \longrightarrow \mathbb{R}$ of the form $Q(\mathbf{x}) = \sum_{i \leq j} \beta_{ij} x_i x_j$ for some $\beta_{ij} \in \mathbb{R}$. A *quadric* in $\mathbb{R}^n$ is a non-empty set of the form

$$V(Q) = \{\mathbf{x} \in \mathbb{R}^n \,|\, Q(\mathbf{x}) = 1\}$$

for some given quadratic form Q i.e. $V(Q)$ is the set of solns to $Q(\mathbf{x}) = 1$.

Rem If Q is a quadratic form & $d \in \mathbb{R} - 0$ such that the solns V to $Q(\mathbf{x}) = d$ is non-empty, then V is the quadric $V(\frac{1}{d}Q)$.

E.g. For $a_{ij} \in \mathbb{R}$, consider the function $Q(\binom{x_1}{x_2}) = (x_1 \ x_2) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \binom{x_1}{x_2} =$

Quadratic forms via symmetric matrices

Quadratic forms can be studied via matrices because of the following

Prop-Defn

1. Let $A \in M_{nn}(\mathbb{R})$. Then $Q_A : \mathbb{R}^n \longrightarrow \mathbb{R} : \mathbf{x} \mapsto \mathbf{x}^T A\mathbf{x}$ is a quadratic form.
2. Let $Q : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a quadratic form. Then there is a unique symmetric matrix $A \in M_{nn}(\mathbb{R})$ such that $Q = Q_A$.

In fact if $A = (a_{ij})_{ij}$ then

$$Q_A(\mathbf{x}) = \sum_i a_{ii}x_i^2 + \sum_{i<j}(a_{ij} + a_{ji})x_i x_j.$$

Proof. Just generalise the calculation in the previous example.

E.g.

Non-degenerate forms

Recall that the e-values of a real symmetric matrix are all real (even when considered as a complex matrix).

Defn

Let $Q : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a quadratic form & $A \in M_{nn}(\mathbb{R})$ be the symmetric matrix with $Q = Q_A$. Let τ^+, τ^- be the number of positive resp negative e-values of A, counted with multiplicity (alg = geom). We say Q (or $V(Q)$) is *non-degenerate* if A is invertible. Otherwise we say it is *degenerate*. The *rank* of Q is $\text{rank}Q = \text{rank } A = \tau^+ + \tau^-$. The *signature* of Q is $\tau^+ - \tau^-$.

E.g. A degenerate quadric in $\mathbb{R}^2$ looks like

Conics

Defn

A non-degenerate quadric in $\mathbb{R}^2$ is called a *conic*.

Hopefully, you already know the following examples. Below we let $a_1, a_2 > 0$.

E.g. 1 If $D = (1) \oplus (1) = I_2$ then $V(Q_D)$ is the circle $x_1^2 + x_2^2 = 1$ & Q_D has signature 2.

E.g. 2 If $D = (\frac{1}{a_1^2}) \oplus (\frac{1}{a_2^2})$ then $V(Q_D)$ is the ellipse

$$\left(\frac{x_1}{a_1}\right)^2 + \left(\frac{x_2}{a_2}\right)^2 = 1$$

& Q_D has signature 2.

N.B. For $a_1 \neq a_2$, the x_1 and x_2-intercepts give the points closest and furthest from $(0,0)$.

Conics cont'd

E.g. 3 If $D = (\frac{1}{a_1^2}) \oplus (-\frac{1}{a_2^2})$ then $V(Q_D)$ is the hyperbola

$$\left(\frac{x_1}{a_1}\right)^2 - \left(\frac{x_2}{a_2}\right)^2 = 1$$

& Q_D has signature 0.

Note the following geometric features:
Closest points to (0,0):
Asymptotes:

Quadric surfaces

A *quadric surface* is a non-degenerate quadric in $\mathbb{R}^3$. Let $a_1, a_2, a_3 > 0$.

Prop

If $D = (\frac{1}{a_1^2}) \oplus (\frac{1}{a_2^2}) \oplus (\frac{1}{a_3^2})$ then $V(Q_D)$ is the *ellipsoid*

$$\left(\frac{x_1}{a_1}\right)^2 + \left(\frac{x_2}{a_2}\right)^2 + \left(\frac{x_3}{a_3}\right)^2 = 1.$$

The points closest (resp furthest) from $(0,0)$ found by setting $x_i = 0$ for all non-minimal (resp non-maximal) a_i. The signature is 3.

Prop

If $D = (\frac{1}{a_1^2}) \oplus (\frac{1}{a_2^2}) \oplus (-\frac{1}{a_3^2})$ then $V(Q_D)$ is the *hyperboloid of 1-sheet*

$$\left(\frac{x_1}{a_1}\right)^2 + \left(\frac{x_2}{a_2}\right)^2 - \left(\frac{x_3}{a_3}\right)^2 = 1.$$

The pts closest to $(0,0)$ must lie on the ellipse $x_3 = 0$. The signature $= 1$.

Quadric surfaces cont'd

A *quadric surface* is a non-degenerate quadric in $\mathbb{R}^3$. Let $a_1, a_2, a_3 > 0$.

Prop

If $D = (\frac{1}{a_1^2}) \oplus (-\frac{1}{a_2^2}) \oplus (-\frac{1}{a_3^2})$ then $V(Q_D)$ is the *hyperboloid of 2-sheets*

$$\left(\frac{x_1}{a_1}\right)^2 - \left(\frac{x_2}{a_2}\right)^2 - \left(\frac{x_3}{a_3}\right)^2 = 1.$$

The points closest to $(0,0)$ are the x_1-intercepts. The signature is -1.

Sketch Let's re-scale axes by changing variables to $y_i = \frac{x_i}{a_i}$. Then in $\mathbb{R}^3_y$-space the surface is just the surface of revolution of

Principal axis theorem

Let $A \in M_{nn}(\mathbb{R})$ be a symm matrix. We apply the spectral thm for self-adjoint operators to orthog diagonalise $D = U^T AU$ where $U \in O_n$ & $D = \oplus_i(\lambda_i)$. Let $\mathbf{v}_i$ be the i-th column of U, a unit norm e-vector of A with e-value λ_i.

Consider now the orthonormal change of co-ord system $U : \mathbb{R}^n_y \longrightarrow \mathbb{R}^n_x$ & introduce new co-ords $\mathbf{y} = U^{-1}\mathbf{x} = U^T\mathbf{x}$. In our new co-ords, our quadratic form Q_A becomes

$$Q_A(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = (U\mathbf{y})^T(UDU^T)(U\mathbf{y}) = \mathbf{y}^T U^T UDU^T U\mathbf{y} = \mathbf{y}^T D\mathbf{y} = Q_D(\mathbf{y}).$$

We obtain

Theorem (Principal axis)

There is an orthonormal change of co-ords such that the quadric $V(Q_A)$ has the form $V(Q_D)$ for D the diagonal matrix of e-values. The new co-ord axes (called principal axes *go through the corresponding e-vectors $\mathbf{v}_i$.*

In particular, any conic is either an ellipse or an hyperbola depending on whether the signature of A is 2 or 0. Any quadric surface is either an ellipsoid, an hyperboloid of 1-sheet, or an hyperboloid of 2-sheets, depending on whether the signature is 3,1 or -1.

Example

E.g. Describe the quadric

$$-2x_1^2 + x_2^2 + x_3^2 + 4x_1x_2 + 4x_1x_3 + 8x_2x_3 = 3$$

Find the closest point on the quadric to the $(0, 0, 0)^T$.

A We may re-write the quadric as $\mathbf{x}^T A\mathbf{x} = 3$ where

$$A = \begin{pmatrix} -2 & 2 & 2 \\ 2 & 1 & 4 \\ 2 & 4 & 1 \end{pmatrix}$$

We orthogonally diagonalised this last lecture to see $A = UDU^T$ where $D = (6) \oplus (-3) \oplus (-3)$ & $U \in O_3$. Also the e-spaces were $E_6 = \text{Span}(1, 2, 2)^T, E_{-3} = E_6^\perp$. Hence there's an orthonormal change of co-ords $U : \mathbb{R}^3_y \longrightarrow \mathbb{R}^3_x$ so that

Example cont'd