Video Lecture F9: Quadratic Forms & Principal Axes Theorem

Tom Roby

Outline & Objectives

- Quadratic forms generalize taking the inner (dot, scalar) product of a vector with itself. They come up in engineering applications involving optimization and signal processing, utility functions in economics, confidence ellipsoids in statistics, etc.
- Analyze how a symmetric matrix A defines a *quadratic form* $Q(\vec{x}) = \vec{x}^T A\vec{x}$.
- Leverage the orthogonal diagonalization, $A = PDP^T$, to compute a change of basis matrix P so that $\vec{x} = P\vec{y}$ transforms $Q(\vec{x}) = \vec{x}^T A\vec{x}$ to $Q(\vec{y}) = \vec{y}^T D\vec{y}$, with no cross terms.

Definition

A *quadratic form* $\mathcal{Q}$ *on* $\mathbb{R}^n$ *is a function* $\mathcal{Q} : \mathbb{R}^n \to \mathbb{R}$ *of the form* $\mathcal{Q}(\vec{x}) = \vec{x}^\top A\vec{x}$, *where* $A \in \mathbb{R}^{n\times n}$ *is* *symmetric* *and called the* *matrix of the quadratic form*.

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ A = \begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix},\ B = \begin{bmatrix} 7 & -3 \\ -3 & 4 \end{bmatrix},\ C = \begin{bmatrix} 4 & -3 \\ -5 & 4 \end{bmatrix}.$$

Find A for $\mathcal{Q}(\vec{x}) = -4x_1^2 + 7x_2^2 - 5x_3^2 - 6x_1x_2 + 3x_2x_3$.

Any *invertible* matrix $P \in \mathbb{R}^{n\times n}$ represents a change of basis from $\mathcal{E}$-coords to $\mathcal{B}$-coords: $\vec{x} = P\vec{y} \iff \vec{y} = P^{-1}\vec{x} \iff [\vec{y}]_{\mathcal{B}} = [\vec{x}]_{\mathcal{E}}$

KEY IDEA: We can change variables to make QF simpler!

$$\vec{x}^\top A\vec{x} = (P\vec{y})^\top A(P\vec{y}) = \vec{y}^\top P^\top AP\vec{y} = \vec{y}^\top(P^\top AP)\vec{y} = \vec{y}^\top(D)\vec{y}$$

Principal Axes Theorem

Use change of variable to simplify: $\mathcal{Q}(\vec{x}) = 2x_1^2 - 4x_1x_2 + 5x_2^2$.

Theorem (Principal Axes Theorem)

Let $\mathcal{Q}$ be a QF on $\mathbb{R}^n$ corr. to (symm) $A \in \mathbb{R}^{n\times n}$. Then we can find an orthogonal $P \in \mathbb{R}^{n\times n}$ such that $\vec{x} = P\vec{y}$, transforming $\mathcal{Q}(\vec{x}) = \vec{x}^\top A\vec{x}$ to $\mathcal{Q}(\vec{y}) = \vec{y}^\top D\vec{y}$, with no cross term. ($D$ is diag.)

The columns of P are called the *principal axes* of the quadratic form $\mathcal{Q}$.

We'll see soon how this helps us classify QF and handle optimization problems.