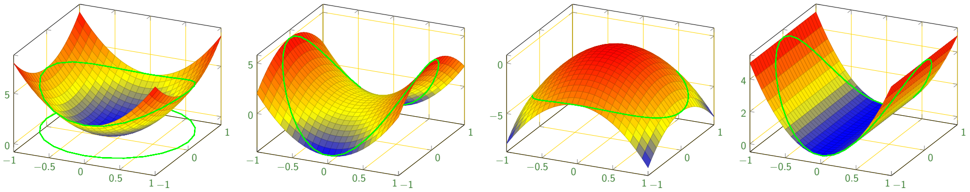
Video Lecture F11: Constrained Optimization of Quadratic Forms

Tom Roby

Outline & Objectives



- Maximizing and minimizing is a major application of mathematical modeling. Here we solve problems in constrained optimization that can be recast as finding the max or min value attained by a QF on the set of unit vectors.
- More concretely we analyze how to maximize a quadratic form $\mathcal{Q} : \mathbb{R}^n \to \mathbb{R}$ subject to the constraint that $\|\vec{x}\| = 1 \iff \|\vec{x}\|^2 = \vec{x}^T\vec{x} = 1 \iff x_1^2 + \cdots + x_n^2 = 1$.
- *Key idea:* It's easy to maximize a QF when there are no cross terms, so we will leverage the Principal Axes Theorem. The max constrained value will be the largest eigenvector of the matrix A corresponding to $\mathcal{Q}$.
- Smaller eigenvalues of A are max value of $\mathcal{Q}$ with additional constraints.

Constrained Optimization of Quadratic Forms

$\mathcal{Q}(\vec{x}) = 7x_1^2 - 5x_2^2 + 4x_3^2$, subject to $\vec{x}^T\vec{x} = 1$.

Theorem (Optimizing a QF on unit vectors)

Let $A \in \mathbb{R}^{n\times n}$ be symmetric with corresp QF $\mathcal{Q}(\vec{x}) := \vec{x}^T A\vec{x}$. Set $M = \max\{\vec{x}^T A\vec{x} : \|\vec{x}\| = 1\}$ and $m = \min\{\vec{x}^T A\vec{x} : \|\vec{x}\| = 1\}$. Then M is the greatest eigenvalue λ_1 and m is the least eigenvalue λ_n of A. These values are obtained by setting $\vec{x}$ to be the corresponding unit eigenvector.

Proof: Orthogonally diagonalize $A = PDP^{-1}$, and use P as the change-of-variable matrix that eliminates the cross-terms in $\mathcal{Q}$, $\vec{x} = P\vec{y} \implies \vec{x}^T A\vec{x} = \vec{y}^T D\vec{y}$. NB: $\|\vec{x}\|^2 = \vec{x}^T\vec{x} = (P\vec{y})^T(P\vec{y})\cdots$ ∎

Find the max and min of $\mathcal{Q}(\vec{x}) = x_1^2 + 3x_2^2 + x_3^2 + 2x_1x_2 + 6x_1x_3 + 2x_2x_3$, subject to $\|\vec{x}\| = 1$.

Additional Constraints

We can interpret the other eigenvalues of (the matrix of) a QF $\mathcal{Q}$ as maximizing $\mathcal{Q}$ with respect to additional constraints.

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be symmetric with corresponding QF $\mathcal{Q}(\vec{x}) := \vec{x}^T A \vec{x}$ and with orthogonal diagonalization $A = PDP^{-1}$. WLOG diagonal entries of D are in decreasing order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, so $P = [\vec{u}_1 \cdots \vec{u}_n]$, where $\vec{u}_i$ is the eigenvector corresponding to λ_i. Then for $k = 2, 3, \ldots n$, the max of $\mathcal{Q}(\vec{x}) := \vec{x}^T A \vec{x}$ subject to

$$\vec{x}^T \vec{x} = 1, \quad \vec{x}^T \vec{u}_1 = 0, \quad \ldots, \quad \vec{x}^T \vec{u}_{k-1} = 0$$

is the eigenvalue λ_k, attained when $\vec{x} = \vec{u}_k$.

$\mathcal{Q}(\vec{x}) = 7x_1^2 - 5x_2^2 + 4x_3^2$, subject to $\vec{x}^T \vec{x} = 1$ and $\vec{x}^T \vec{u}_1 = 0$.

Maximize $\mathcal{Q} = x_1^2 + 3x_2^2 + x_3^2 + 2x_1x_2 + 6x_1x_3 + 2x_2x_3$, subject to $\|\vec{x}\| = 1$, and $\vec{x}^T \vec{u}_1 = 0$.