7.3 Constrained Optimization

In optimization problems, we seek an optimum solution, such as the maximum or minimum value of an expression. One such expression is $\mathbf{x}^T A\mathbf{x}$, in which A is a symmetric matrix. This form is called the quadratic form. The simplest example of a nonzero quadratic form is $\mathbf{x}^T I\mathbf{x} = \|\mathbf{x}\|^2$. Typically, we may arrange the optimization problem so that $\mathbf{x}$ varies over the set of unit vectors. This constrained optimization problem has an interesting and elegant solution.

If $\mathbf{x}$ represents a variable vector in $\mathbb{R}^n$, then a *change of variable* is an equation of the form

$$\mathbf{x} = P\mathbf{y} \tag{1}$$

where P is an invertible matrix and $\mathbf{y}$ is a new variable vector in $\mathbb{R}^n$. Here $\mathbf{y}$ is the coordinate vector of $\mathbf{x}$ relative to the basis of $\mathbb{R}^n$ determined by the columns of P. Then

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T(P^T AP)\mathbf{y} \tag{2}$$

and the new matrix of the quadratic form is $P^T AP$. Since A is symmetric, Theorem 2 guarantees that there is an orthogonal matrix P such that $P^T AP$ is a diagonal matrix D, and the quadratic form in (2) becomes $\mathbf{y}^T D\mathbf{y}$.

Example 1. *Find the change of variable $\mathbf{x} = P\mathbf{y}$ that transforms the quadratic form $\mathbf{x}^T A\mathbf{x}$ into $\mathbf{y}^T D\mathbf{y}$ as shown.*

$$3x_1^2 + 3x_2^2 + 5x_3^2 + 6x_1x_2 + 2x_1x_3 + 2x_2x_3 = 7y_1^2 + 4y_2^2.$$

Theorem 6. *Let A be a symmetric matrix, and define m and M as*

$$m = \min\{\mathbf{x}^T A\mathbf{x} : \|\mathbf{x}\| = 1\}, \quad M = \max\{\mathbf{x}^T A\mathbf{x} : \|\mathbf{x}\| = 1\}. \tag{3}$$

Then M is the greatest eigenvalue λ_1 of A and m is the least eigenvalue of A. The value of $\mathbf{x}^T A\mathbf{x}$ is M when $\mathbf{x}$ is a unit eigenvector $\mathbf{u}_1$ corresponding to M. The value of $\mathbf{x}^T A\mathbf{x}$ is m when $\mathbf{x}$ is a unit eigenvector corresponding to m.

Theorem 7. *Let $A, \lambda_1,$ and $\mathbf{u}_1$ be as in Theorem 6. Then the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints*

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0$$

is the second greatest eigenvalue, λ_2, and this maximum is attained when $\mathbf{x}$ is an eigenvector $\mathbf{u}_2$ corresponding to λ_2.

Example 2. *Find (a) the maximum value of $Q(\mathbf{x})$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$, (b) a unit vector $\mathbf{u}$ where this maximum is attained, and (c) the maximum of $Q(\mathbf{x})$ subject to the constraints $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^T\mathbf{u} = 0$.*

$$Q(\mathbf{x}) = 3x_1^2 + 3x_2^2 + 5x_3^2 + 6x_1x_2 + 2x_1x_3 + 2x_2x_3$$

Theorem 8. *Let A be a symmetric $n \times n$ matrix with an orthogonal diagonalization $A = PDP^{-1}$, where the entries on the diagonal of D are arranged so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and where the columns of P are corresponding unit eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Then for $k = 2, \ldots, n$, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints*

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

is the eigenvalue λ_k, and this maximum is attained at $\mathbf{x} = \mathbf{u}_k$.

Example 3. *Suppose $Q(\mathbf{x}) = 3x_1^2 + 9x_2^2 + 8x_1x_2$.*

a) Find the maximum value of $Q(\mathbf{x})$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$.

b) Find a unit vector $\mathbf{u}$ where this maximum is attained.

c) Find the maximum of $Q(\mathbf{x})$ subject to the constraints $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^T\mathbf{u} = 0$.