2013

The Method of Lagrange Multipliers

William F. Trench, *Trinity University*

THE METHOD OF LAGRANGE MULTIPLIERS

William F. Trench

Andrew G. Cowles Distinguished Professor Emeritus
Department of Mathematics
Trinity University
San Antonio, Texas, USA
wtrench@trinity.edu

This is a supplement to the author's Introduction to Real Analysis. It has been judged to meet the evaluation criteria set by the Editorial Board of the American Institute of Mathematics in connection with the Institute's Open Textbook Initiative. A complete instructor's solution manual is available by email to wtrench@trinity.edu, subject to verification of the requestor's faculty status.

THE METHOD OF LAGRANGE MULTIPLIERS

William F. Trench

1 Foreword

This is a revised and extended version of Section 6.5 of my *Advanced Calculus* (Harper & Row, 1978). It is a supplement to my textbook *Introduction to Real Analysis*, which is referenced via hypertext links.

2 Introduction

To avoid repetition, it is to be understood throughout that f and g_1, g_2,..., g_m are continuously differentiable on an open set D in $\mathbb{R}^n$.

Suppose that $m < n$ and

$$g_1(\mathbf{X}) = g_2(\mathbf{X}) = \cdots = g_m(\mathbf{X}) = 0 \tag{1}$$

on a nonempty subset D_1 of D. If $\mathbf{X}_0 \in D_1$ and there is a neighborhood N of $\mathbf{X}_0$ such that

$$f(\mathbf{X}) \le f(\mathbf{X}_0) \tag{2}$$

for every $\mathbf{X}$ in $N \cap D_1$, then $\mathbf{X}_0$ is *a local maximum point of f subject to the constraints* (1). However, we will usually say "subject to" rather than "subject to the constraint(s)."

If (2) is replaced by

$$f(\mathbf{X}) \ge f(\mathbf{X}_0), \tag{3}$$

then "maximum" is replaced by "minimum." A local maximum or minimum of f subject to (1) is also called a *local extreme point of f subject to* (1). More briefly, we also speak of *constrained local maximum, minimum, or extreme points*. If (2) or (3) holds for all $\mathbf{X}$ in D_1, we omit "local."

Recall that $\mathbf{X}_0 = (x_{10}, x_{20}, \dots, x_{n0})$ is a *critical point* of a differentiable function $L = L(x_1, x_2, \dots, x_n)$ if

$$L_{x_i}(x_{10}, x_{20}, \dots, x_{n0}) = 0, \quad 1 \le i \le n.$$

Therefore, every local extreme point of L is a critical point of L; however, a critical point of L is not necessarily a local extreme point of L (pp. 334-5).

Suppose that the system (1) of simultaneous equations can be solved for x_1, ..., x_m in terms of the x_{m+1}, ..., x_n; thus,

$$x_j = h_j(x_{m+1}, \dots, x_n), \quad 1 \le j \le m. \tag{4}$$

Then a constrained extreme value of f is an unconstrained extreme value of

$$f(h_1(x_{m+1}, \dots, x_n), \dots, h_m(x_{m+1}, \dots, x_n), x_{m+1}, \dots, x_n). \tag{5}$$

However, it may be difficult or impossible to find explicit formulas for $h_1, h_2, \dots, h_m$, and, even if it is possible, the composite function (5) is almost always complicated. Fortunately, there is a better way to to find constrained extrema, which also requires the solvability assumption, but does not require an explicit formula as indicated in (4). It is based on the following theorem. Since the proof is complicated, we consider two special cases first.

Theorem 1 *Suppose that $n > m$. If $\mathbf{X}_0$ is a local extreme point of f subject to*

$$g_1(\mathbf{X}) = g_2(\mathbf{X}) = \cdots = g_m(\mathbf{X}) = 0$$

and

$$\left| \begin{array}{cccc} \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_{r_1}} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_{r_2}} & \cdots & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_{r_m}} \\ \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_{r_1}} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_{r_2}} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_{r_m}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_{r_1}} & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_{r_2}} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_{r_m}} \end{array} \right| \ne 0 \qquad (6)$$

for at least one choice of $r_1 < r_2 < \cdots < r_m$ in $\{1, 2, \dots, n\}$, then there are constants $\lambda_1, \lambda_2, \dots, \lambda_m$ such that $\mathbf{X}_0$ is a critical point of

$$f - \lambda_1 g_1 - \lambda_2 g_2 - \cdots - \lambda_m g_m;$$

that is,

$$\frac{\partial f(\mathbf{X}_0)}{\partial x_i} - \lambda_1 \frac{\partial g_1(\mathbf{X}_0)}{\partial x_i} - \lambda_2 \frac{\partial g_2(\mathbf{X}_0)}{\partial x_i} - \cdots - \lambda_m \frac{\partial g_m(\mathbf{X}_0)}{\partial x_i} = 0,$$

$1 \le i \le n$.

The following implementation of this theorem is the *method of Lagrange multipliers*.

(a) Find the critical points of

$$f - \lambda_1 g_1 - \lambda_2 g_2 - \cdots - \lambda_m g_m,$$

treating $\lambda_1, \lambda_2, \dots \lambda_m$ as unspecified constants.

(b) Find $\lambda_1, \lambda_2, \dots, \lambda_m$ so that the critical points obtained in (a) satisfy the constraints.

(c) Determine which of the critical points are constrained extreme points of f. This can usually be done by physical or intuitive arguments.

If a and $b_1, b_2, \dots, b_m$ are nonzero constants and c is an arbitrary constant, then the local extreme points of f subject to $g_1 = g_2 = \cdots = g_m = 0$ are the same as the local extreme points of $af - c$ subject to $b_1 g_1 = b_2 g_2 = \cdots = b_m g_m = 0$. Therefore, we can replace $f - \lambda_1 g_1 - \lambda_2 g_2 - \cdots - \lambda_m g_m$ by $af - \lambda_1 b_1 g_1 - \lambda_2 b_2 g_2 - \cdots - \lambda_m b_m g_m - c$ to simplify computations. (Usually, the "$-c$" indicates dropping additive constants.) We will denote the final form by L (for *Lagrangian*).

3 Extrema subject to one constraint

Here is Theorem 1 with $m = 1$.

Theorem 2 *Suppose that $n > 1$. If $\mathbf{X_0}$ is a local extreme point of f subject to $g(\mathbf{X}) = 0$ and $g_{x_r}(\mathbf{X_0}) \neq 0$ for some $r \in \{1, 2, \dots, n\}$, then there is a constant λ such that*

$$f_{x_i}(\mathbf{X_0}) - \lambda g_{x_i}(\mathbf{X_0}) = 0, \tag{7}$$

$1 \le i \le n$; thus, $\mathbf{X_0}$ is a critical point of $f - \lambda g$.

Proof For notational convenience, let $r = 1$ and denote

$$\mathbf{U} = (x_2, x_3, \dots x_n) \text{ and } \mathbf{U_0} = (x_{20}, x_{30}, \dots x_{n0}).$$

Since $g_{x_1}(\mathbf{X_0}) \neq 0$, the Implicit Function Theorem (Corollary 6.4.2, p. 423) implies that there is a unique continuously differentiable function $h = h(\mathbf{U})$, defined on a neighborhood $N \subset \mathbb{R}^{n-1}$ of $\mathbf{U_0}$, such that $(h(\mathbf{U}), \mathbf{U}) \in D$ for all $\mathbf{U} \in N$, $h(\mathbf{U_0}) = x_{10}$, and

$$g(h(\mathbf{U}), \mathbf{U}) = 0, \quad \mathbf{U} \in N. \tag{8}$$

Now define

$$\lambda = \frac{f_{x_1}(\mathbf{X_0})}{g_{x_1}(\mathbf{X_0})}, \tag{9}$$

which is permissible, since $g_{x_1}(\mathbf{X_0}) \neq 0$. This implies (7) with $i = 1$. If $i > 1$, differentiating (8) with respect to x_i yields

$$\frac{\partial g(h(\mathbf{U}), \mathbf{U})}{\partial x_i} + \frac{\partial g(h(\mathbf{U}), \mathbf{U})}{\partial x_1} \frac{\partial h(\mathbf{U})}{\partial x_i} = 0, \quad \mathbf{U} \in N. \tag{10}$$

Also,

$$\frac{\partial f(h(\mathbf{U}), \mathbf{U}))}{\partial x_i} = \frac{\partial f(h(\mathbf{U}), \mathbf{U})}{\partial x_i} + \frac{\partial f(h(\mathbf{U}), \mathbf{U})}{\partial x_1} \frac{\partial h(\mathbf{U})}{\partial x_i}, \quad \mathbf{U} \in N. \tag{11}$$

Since $(h(\mathbf{U_0}), \mathbf{U_0}) = \mathbf{X_0}$, (10) implies that

$$\frac{\partial g(\mathbf{X_0})}{\partial x_i} + \frac{\partial g(\mathbf{X_0})}{\partial x_1} \frac{\partial h(\mathbf{U_0})}{\partial x_i} = 0. \tag{12}$$

If $\mathbf{X_0}$ is a local extreme point of f subject to $g(\mathbf{X}) = 0$, then $\mathbf{U_0}$ is an unconstrained local extreme point of $f(h(\mathbf{U}), \mathbf{U})$; therefore, (11) implies that

$$\frac{\partial f(\mathbf{X_0})}{\partial x_i} + \frac{\partial f(\mathbf{X_0})}{\partial x_1} \frac{\partial h(\mathbf{U_0})}{\partial x_i} = 0. \tag{13}$$

Since a linear homogeneous system

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

has a nontrivial solution if and only if

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = 0,$$

(Theorem 6.1.15, p. 376), (12) and (13) imply that

$$\begin{vmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} \\[2ex] \dfrac{\partial g(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_1} \end{vmatrix} = 0, \quad \text{so} \quad \begin{vmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_i} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_1} \end{vmatrix} = 0,$$

since the determinants of a matrix and its transpose are equal. Therefore, the system

$$\begin{bmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_i} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_1} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

has a nontrivial solution (Theorem 6.1.15, p. 376). Since $g_{x_1}(\mathbf{X}_0) \neq 0$, u must be nonzero in a nontrivial solution. Hence, we may assume that $u = 1$, so

$$\begin{bmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_i} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_1} \end{bmatrix} \begin{bmatrix} 1 \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{14}$$

In particular,

$$\frac{\partial f(\mathbf{X}_0)}{\partial x_1} + v\frac{\partial g(\mathbf{X}_0)}{\partial x_1} = 0, \quad \text{so} \quad -v = \frac{f_{x_1}(\mathbf{X}_0)}{g_{x_1}(\mathbf{X}_0)}.$$

Now (9) implies that $-v = \lambda$, and (14) becomes

$$\begin{bmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_i} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_1} \end{bmatrix} \begin{bmatrix} 1 \\ -\lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Computing the topmost entry of the vector on the left yields (7). ■

Example 1 Find the point (x_0, y_0) on the line

$$ax + by = d$$

closest to a given point (x_1, y_1).

Solution We must minimize $\sqrt{(x-x_1)^2+(y-y_1)^2}$ subject to the constraint. This is equivalent to minimizing $(x-x_1)^2+(y-y_1)^2$ subject to the constraint, which is simpler. For, this we could let

$$L = (x-x_1)^2 + (y-y_1)^2 - \lambda(ax+by-d);$$

however,

$$L = \frac{(x-x_1)^2+(y-y_1)^2}{2} - \lambda(ax+by)$$

is better. Since

$$L_x = x - x_1 - \lambda a \quad \text{and} \quad L_y = y - y_1 - \lambda b,$$

$(x_0, y_0) = (x_1+\lambda a, y_1+\lambda b)$, where we must choose λ so that $ax_0+by_0 = d$. Therefore,

$$ax_0 + by_0 = ax_1 + by_1 + \lambda(a^2+b^2) = d,$$

so

$$\lambda = \frac{d - ax_1 - by_1}{a^2+b^2},$$

$$x_0 = x_1 + \frac{(d-ax_1-by_1)a}{a^2+b^2}, \quad \text{and} \quad y_0 = y_1 + \frac{(d-ax_1-by_1)b}{a^2+b^2}.$$

The distance from (x_1, y_1) to the line is

$$\sqrt{(x_0-x_1)^2+(y_0-y_1)^2} = \frac{|d-ax_1-by_1|}{\sqrt{a^2+b^2}}.$$

Example 2 Find the extreme values of $f(x,y) = 2x+y$ subject to

$$x^2+y^2 = 4.$$

Solution Let

$$L = 2x + y - \frac{\lambda}{2}(x^2+y^2);$$

then

$$L_x = 2 - \lambda x \quad \text{and} \quad L_y = 1 - \lambda y,$$

so $(x_0, y_0) = (2/\lambda, 1/\lambda)$. Since $x_0^2+y_0^2 = 4$, $\lambda = \pm\sqrt{5}/2$. Hence, the constrained maximum is $2\sqrt{5}$, attained at $(4/\sqrt{5}, 2/\sqrt{5})$, and the constrained minimum is $-2\sqrt{5}$, attained at $(-4/\sqrt{5}, -2/\sqrt{5})$.

Example 3 Find the point in the plane

$$3x + 4y + z = 1 \tag{15}$$

closest to $(-1, 1, 1)$.

Solution We must minimize

$$f(x,y,z) = (x+1)^2 + (y-1)^2 + (z-1)^2$$

subject to (15). Let

$$L = \frac{(x+1)^2 + (y-1)^2 + (z-1)^2}{2} - \lambda(3x + 4y + z);$$

then

$$L_x = x + 1 - 3\lambda, \quad L_y = y - 1 - 4\lambda, \text{ and } L_z = z - 1 - \lambda,$$

so

$$x_0 = -1 + 3\lambda, \quad y_0 = 1 + 4\lambda, \quad z_0 = 1 + \lambda.$$

From (15),

$$3(-1 + 3\lambda) + 4(1 + 4\lambda) + (1 + \lambda) - 1 = 1 + 26\lambda = 0,$$

so $\lambda = -1/26$ and

$$(x_0, y_0, z_0) = \left(-\frac{29}{26}, \frac{22}{26}, \frac{25}{26}\right).$$

The distance from (x_0, y_0, z_0) to $(-1, 1, 1)$ is

$$\sqrt{(x_0+1)^2 + (y_0-1)^2 + (z_0-1)^2} = \frac{1}{\sqrt{26}}.$$

Example 4 Assume that $n \geq 2$ and $x_i \geq 0$, $1 \leq i \leq n$.

(a) Find the extreme values of $\sum_{i=1}^{n} x_i$ subject to $\sum_{i=1}^{n} x_i^2 = 1$.

(b) Find the minimum value of $\sum_{i=1}^{n} x_i^2$ subject to $\sum_{i=1}^{n} x_i = 1$.

Solution (a) Let

$$L = \sum_{i=1}^{n} x_i - \frac{\lambda}{2} \sum_{i=1}^{n} x_i^2;$$

then

$$L_{x_i} = 1 - \lambda x_i, \text{ so } x_{i0} = \frac{1}{\lambda}, \quad 1 \leq i \leq n.$$

Hence, $\sum_{i=1}^{n} x_{i0}^2 = n/\lambda^2$, so $\lambda = \pm\sqrt{n}$ and

$$(x_{10}, x_{20}, \dots, x_{n0}) = \pm\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right).$$

Therefore, the constrained maximum is $\sqrt{n}$ and the constrained minimum is $-\sqrt{n}$.

Solution (b) Let

$$L = \frac{1}{2} \sum_{i=1}^{n} x_i^2 - \lambda \sum_{i=1}^{n} x_i;$$

then

$$L_{x_i} = x_i - \lambda, \quad \text{so} \quad x_{i0} = \lambda, \qquad 1 \le i \le n.$$

Hence, $\displaystyle\sum_{i=1}^{n} x_{i0} = n\lambda = 1$, so $x_{i0} = \lambda = 1/n$ and the constrained minimum is

$$\sum_{i=1}^{n} x_{i0}^2 = \frac{1}{n}$$

There is no constrained maximum. (Why?)

Example 5 Show that

$$x^{1/p} y^{1/q} \le \frac{x}{p} + \frac{y}{q}, \qquad x, y \ge 0,$$

if

$$\frac{1}{p} + \frac{1}{q} = 1, \qquad p > 0, \quad \text{and} \quad q > 0. \tag{16}$$

Solution We first find the maximum of

$$f(x, y) = x^{1/p} y^{1/q}$$

subject to

$$\frac{x}{p} + \frac{y}{q} = \sigma, \qquad x \ge 0, \qquad y \ge 0, \tag{17}$$

where σ is a fixed but arbitrary positive number. Since f is continuous, it must assume a maximum at some point (x_0, y_0) on the line segment (17), and (x_0, y_0) cannot be an endpoint of the segment, since $f(p\sigma, 0) = f(0, q\sigma) = 0$. Therefore, (x_0, y_0) is in the open first quadrant.

Let

$$L = x^{1/p} y^{1/q} - \lambda \left(\frac{x}{p} + \frac{y}{q} \right).$$

Then

$$L_x = \frac{1}{px} f(x, y) - \frac{\lambda}{p} \quad \text{and} \quad L_y = \frac{1}{qy} f(x, y) - \frac{\lambda}{q} = 0,$$

so $x_0 = y_0 = f(x_0, y_0)/\lambda$. Now(16) and (17) imply that $x_0 = y_0 = \sigma$. Therefore,

$$f(x, y) \le f(\sigma, \sigma) = \sigma^{1/p} \sigma^{1/q} = \sigma = \frac{x}{p} + \frac{y}{q}.$$

This can be generalized (Exercise 53). It can also be used to generalize Schwarz's inequality (Exercise 54).

4 Constrained Extrema of Quadratic Forms

In this section it is convenient to write

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

An *eigenvalue* of a square matrix $\mathbf{A} = [a_{ij}]_{i,j=1}^n$ is a number λ such that the system

$$\mathbf{AX} = \lambda\mathbf{X},$$

or, equivalently,

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = \mathbf{0},$$

has a solution $\mathbf{X} \neq \mathbf{0}$. Such a solution is called an *eigenvector* of $\mathbf{A}$. You probably know from linear algebra that λ is an eigenvalue of $\mathbf{A}$ if and only if

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0.$$

Henceforth we assume that $\mathbf{A}$ is symmetric ($a_{ij} = a_{ji}$, $1 \le i, j \le n$). In this case,

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n),$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are real numbers.

The function

$$Q(\mathbf{X}) = \sum_{i,j=1}^n a_{ij}x_ix_j$$

is a *quadratic form*. To find its maximum or minimum subject to $\sum_{i=1}^n x_i^2 = 1$, we form the Lagrangian

$$L = Q(\mathbf{X}) - \lambda\sum_{i=1}^n x_i^2.$$

Then

$$L_{x_i} = 2\sum_{j=1}^n a_{ij}x_j - 2\lambda x_i = 0, \quad 1 \le i \le n,$$

so

$$\sum_{j=1}^n a_{ij}x_{j0} = \lambda x_{i0}, \quad 1 \le i \le n.$$

Therefore, $\mathbf{X_0}$ is a constrained critical point of Q subject to $\sum_{i=1}^n x_i^2 = 1$ if and only if $\mathbf{AX}_0 = \lambda\mathbf{X}_0$ for some λ; that is, if and only if λ is an eigenvalue and $\mathbf{X_0}$ is an associated unit eigenvector of $\mathbf{A}$. If $\mathbf{AX}_0 = \mathbf{X}_0$ and $\sum_i^n x_{i0}^2 = 1$, then

$$\begin{aligned} Q(\mathbf{X}_0) &= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} x_{j0} \right) x_{i0} = \sum_{i=1}^{n} (\lambda x_{i0}) x_{i0} \\ &= \lambda \sum_{i=1}^{n} x_{i0}^2 = \lambda; \end{aligned}$$

therefore, the largest and smallest eigenvalues of $\mathbf{A}$ are the maximum and minimum values of Q subject to $\sum_{i=1}^{n} x_i^2 = 1$.

Example 6 Find the maximum and minimum values

$$Q(\mathbf{X}) = x^2 + y^2 + 2z^2 - 2xy + 4xz + 4yz$$

subject to the constraint

$$x^2 + y^2 + z^2 = 1. \tag{18}$$

Solution The matrix of Q is

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & 2 \\ 2 & 2 & 2 \end{bmatrix}$$

and

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} 1-\lambda & -1 & 2 \\ -1 & 1-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{vmatrix} \\ &= -(\lambda + 2)(\lambda - 2)(\lambda - 4), \end{aligned}$$

so

$$\lambda_1 = 4, \quad \lambda_2 = 2, \quad \lambda_3 = -2$$

are the eigenvalues of $\mathbf{A}$. Hence, $\lambda_1 = 4$ and $\lambda_3 = -2$ are the maximum and minimum values of Q subject to (18).

To find the points (x_1, y_1, z_1) where Q attains its constrained maximum, we first find an eigenvector of $\mathbf{A}$ corresponding to $\lambda_1 = 4$. To do this, we find a nontrivial solution of the system

$$(\mathbf{A} - 4\mathbf{I}) \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} -3 & -1 & 2 \\ -1 & -3 & 2 \\ 2 & 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

All such solutions are multiples of $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$. Normalizing this to satisfy (18) yields

$$\mathbf{X}_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \pm \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

To find the points (x_3, y_3, z_3) where Q attains its constrained minimum, we first find an eigenvector of $\mathbf{A}$ corresponding to $\lambda_3 = -2$. To do this, we find a nontrivial solution of the system

$$(\mathbf{A} + 2\mathbf{I}) \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} = \begin{bmatrix} 3 & -1 & 2 \\ -1 & 3 & 2 \\ 2 & 2 & 4 \end{bmatrix} \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

All such solutions are multiples of $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$. Normalizing this to satisfy (18) yields

$$\mathbf{X}_3 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \pm \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

As for the eigenvalue $\lambda_2 = 2$, we leave it you to verify that the only unit vectors that satisfy $\mathbf{A}\mathbf{X}_2 = 2\mathbf{X}_2$ are

$$\mathbf{X}_2 = \pm \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

For more on this subject, see Theorem 4.

5 Extrema subject to two constraints

Here is Theorem 1 with $m = 2$.

Theorem 3 *Suppose that* $n > 2$*. If* $\mathbf{X}_0$ *is a local extreme point of* f *subject to* $g_1(\mathbf{X}) = g_2(\mathbf{X}) = 0$ *and*

$$\begin{vmatrix} \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_r} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_s} \\[2ex] \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_r} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_s} \end{vmatrix} \neq 0 \tag{19}$$

for some r *and* s *in* $\{1, 2, \dots, n\}$*, then there are constants* λ *and* μ *such that*

$$\frac{\partial f(\mathbf{X}_0)}{\partial x_i} - \lambda \frac{\partial g_1(\mathbf{X}_0)}{\partial x_i} - \mu \frac{\partial g_2(\mathbf{X}_0)}{\partial x_i} = 0, \tag{20}$$

$1 \le i \le n$.

Proof For notational convenience, let $r = 1$ and $s = 2$. Denote

$$\mathbf{U} = (x_3, x_4, \dots x_n) \text{ and } \mathbf{U}_0 = (x_{30}, x_{30}, \dots x_{n0}).$$

Since

$$\begin{vmatrix} \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} \\[2ex] \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} \end{vmatrix} \neq 0, \tag{21}$$

the Implicit Function Theorem (Theorem 6.4.1, p. 420) implies that there are unique continuously differentiable functions

$$h_1 = h_1(x_3, x_4, \dots, x_n) \text{ and } h_2 = h_1(x_3, x_4, \dots, x_n),$$

defined on a neighborhood $N \subset \mathbb{R}^{n-2}$ of $\mathbf{U}_0$, such that $(h_1(\mathbf{U}), h_2(\mathbf{U}), \mathbf{U}) \in D$ for all $\mathbf{U} \in N$, $h_1(\mathbf{U}_0) = x_{10}$, $h_2(\mathbf{U}_0) = x_{20}$, and

$$g_1(h_1(\mathbf{U}), h_2(\mathbf{U}), \mathbf{U}) = g_2(h_1(\mathbf{U}), h_2(\mathbf{U}), \mathbf{U}) = 0, \quad \mathbf{U} \in N. \tag{22}$$

From (21), the system

$$\begin{bmatrix} \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} \\[2ex] \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} f_{x_1}(\mathbf{X}_0) \\ f_{x_2}(\mathbf{X}_0) \end{bmatrix} \tag{23}$$

has a unique solution (Theorem 6.1.13, p. 373). This implies (20) with $i = 1$ and $i = 2$. If $3 \le i \le n$, then differentiating (22) with respect to x_i and recalling that $(h_1(\mathbf{U}_0), h_2(\mathbf{U}_0), \mathbf{U}_0) = \mathbf{X}_0$ yields

$$\frac{\partial g_1(\mathbf{X}_0)}{\partial x_i} + \frac{\partial g_1(\mathbf{X}_0)}{\partial x_1}\frac{\partial h_1(\mathbf{U}_0)}{\partial x_i} + \frac{\partial g_1(\mathbf{X}_0)}{\partial x_2}\frac{\partial h_2(\mathbf{U}_0)}{\partial x_i} = 0$$

and

$$\frac{\partial g_2(\mathbf{X}_0)}{\partial x_i} + \frac{\partial g_2(\mathbf{X}_0)}{\partial x_1}\frac{\partial h_1(\mathbf{U}_0)}{\partial x_i} + \frac{\partial g_2(\mathbf{X}_0)}{\partial x_2}\frac{\partial h_2(\mathbf{U}_0)}{\partial x_i} = 0.$$

If $\mathbf{X}_0$ is a local extreme point of f subject to $g_1(\mathbf{X}) = g_2(\mathbf{X}) = 0$, then $\mathbf{U}_0$ is an unconstrained local extreme point of $f(h_1(\mathbf{U}), h_2(\mathbf{U}), \mathbf{U})$; therefore,

$$\frac{\partial f(\mathbf{X}_0)}{\partial x_i} + \frac{\partial f(\mathbf{X}_0)}{\partial x_1}\frac{\partial h_1(\mathbf{U}_0)}{\partial x_i} + \frac{\partial f(\mathbf{X}_0)}{\partial x_2}\frac{\partial h_2(\mathbf{U}_0)}{\partial x_i} = 0.$$

The last three equations imply that

$$\begin{vmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial f(\mathbf{X}_0)}{\partial x_2} \\[2ex] \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} \\[2ex] \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} \end{vmatrix} = 0,$$

$$\left|\begin{array}{ccc} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_i} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} \end{array}\right| = 0.$$

Therefore, there are constants c_1, c_2, c_3, not all zero, such that

$$\left[\begin{array}{ccc} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_i} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} \end{array}\right] \left[\begin{array}{c} c_1 \\ c_2 \\ c_3 \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ 0 \end{array}\right]. \tag{24}$$

If $c_1 = 0$, then

$$\left[\begin{array}{cc} \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} \\[2ex] \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} \end{array}\right] \left[\begin{array}{c} c_2 \\ c_3 \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \end{array}\right],$$

so (19) implies that $c_2 = c_3 = 0$; hence, we may assume that $c_1 = 1$ in a nontrivial solution of (24). Therefore,

$$\left[\begin{array}{ccc} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_i} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} \end{array}\right] \left[\begin{array}{c} 1 \\ c_2 \\ c_3 \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ 0 \end{array}\right], \tag{25}$$

which implies that

$$\left[\begin{array}{cc} \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} \\[2ex] \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} \end{array}\right] \left[\begin{array}{c} -c_2 \\ -c_3 \end{array}\right] = \left[\begin{array}{c} f_{x_1}(\mathbf{X}_0) \\ f_{x_2}(\mathbf{X}_0) \end{array}\right].$$

Since (23) has only one solution, this implies that $c_2 = -\lambda$ and $c_2 = -\mu$, so (25) becomes

$$\begin{bmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_i} \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} \end{bmatrix} \begin{bmatrix} 1 \\ -\lambda \\ -\mu \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Computing the topmost entry of the vector on the left yields (20). ■

Example 7 Minimize

$$f(x, y, z, w) = x^2 + y^2 + z^2 + w^2$$

subject to

$$x + y + z + w = 10 \quad \text{and} \quad x - y + z + 3w = 6. \tag{26}$$

Solution Let

$$L = \frac{x^2 + y^2 + z^2 + w^2}{2} - \lambda(x + y + z + w) - \mu(x - y + z + 3w);$$

then

$$\begin{aligned} L_x &= x - \lambda - \mu \\ L_y &= y - \lambda + \mu \\ L_z &= z - \lambda - \mu \\ L_w &= w - \lambda - 3\mu, \end{aligned}$$

so

$$x_0 = \lambda + \mu, \quad y_0 = \lambda - \mu, \quad z_0 = \lambda + \mu, \quad w_0 = \lambda + 3\mu. \tag{27}$$

This and (26) imply that

$$\begin{aligned} (\lambda + \mu) + (\lambda - \mu) + (\lambda + \mu) + (\lambda + 3\mu) &= 10 \\ (\lambda + \mu) - (\lambda - \mu) + (\lambda + \mu) + (3\lambda + 9\mu) &= 6. \end{aligned}$$

Therefore,

$$\begin{aligned} 4\lambda + 4\mu &= 10 \\ 4\lambda + 12\mu &= 6, \end{aligned}$$

so $\lambda = 3$ and $\mu = -1/2$. Now (27) implies that

$$(x_0, y_0, z_0, w_0) = \left(\frac{5}{2}, \frac{7}{2}, \frac{5}{2}\frac{3}{2}\right).$$

Since $f(x, y, z, w)$ is the square of the distance from (x, y, z, w) to the origin, it attains a minimum value (but not a maximum value) subject to the constraints; hence the constrained minimum value is

$$f\left(\frac{5}{2}, \frac{7}{2}, \frac{5}{2}, \frac{3}{2}\right) = 27.$$

Example 8 The distance between two curves in $\mathbb{R}^2$ is the minimum value of

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2},$$

where (x_1, y_1) is on one curve and (x_2, y_2) is on the other. Find the distance between the ellipse

$$x^2 + 2y^2 = 1$$

and the line

$$x + y = 4. \tag{28}$$

Solution We must minimize

$$d^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2$$

subject to

$$x_1^2 + 2y_1^2 = 1 \text{ and } x_2 + y_2 = 4.$$

Let

$$L = \frac{(x_1 - x_2)^2 + (y_1 - y_2)^2 - \lambda(x_1^2 + 2y_1^2)}{2} - \mu(x_2 + y_2);$$

then

$$\begin{aligned} L_{x_1} &= x_1 - x_2 - \lambda x_1 \\ L_{y_1} &= y_1 - y_2 - 2\lambda y_1 \\ L_{x_2} &= x_2 - x_1 - \mu \\ L_{y_2} &= y_2 - y_1 - \mu, \end{aligned}$$

so

$$\begin{aligned} x_{10} - x_{20} &= \lambda x_{10} && \text{(i)} \\ y_{10} - y_{20} &= 2\lambda y_{10} && \text{(ii)} \\ x_{20} - x_{10} &= \mu && \text{(iii)} \\ y_{20} - y_{10} &= \mu. && \text{(iv)} \end{aligned}$$

From (i) and (iii), $\mu = -\lambda x_{10}$; from (ii) and (iv), $\mu = -2\lambda y_{10}$. Since the curves do not intersect, $\lambda \neq 0$, so $x_{10} = 2y_{10}$. Since $x_{10}^2 + 2y_{10}^2 = 1$ and (x_0, y_0) is in the first quadrant,

$$(x_{10}, y_{10}) = \left(\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right). \tag{29}$$

Now (iii), (iv), and (28) yield the simultaneous system

$$x_{20} - y_{20} = x_{10} - y_{10} = \frac{1}{\sqrt{6}}, \quad x_{20} + y_{20} = 4,$$

so

$$(x_{20}, y_{20}) = \left(2 + \frac{1}{2\sqrt{6}}, 2 - \frac{1}{2\sqrt{6}}\right).$$

From this and (29), the distance between the curves is

$$\left[\left(2 + \frac{1}{2\sqrt{6}} - \frac{2}{\sqrt{6}}\right)^2 + \left(2 - \frac{1}{2\sqrt{6}} - \frac{1}{\sqrt{6}}\right)^2\right]^{1/2} = \sqrt{2}\left(2 - \frac{3}{2\sqrt{6}}\right).$$

6 Proof of Theorem 1

Proof For notational convenience, let $r_\ell = \ell$, $1 \le \ell \le m$, so (6) becomes

$$\begin{vmatrix} \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_m} \\ \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{vmatrix} \ne 0 \tag{30}$$

Denote

$$\mathbf{U} = (x_{m+1}, x_{m+2}, \dots x_n) \ \text{ and } \ \mathbf{U}_0 = (x_{m+1,0}, x_{m+2,0}, \dots x_{n0}).$$

From (30), the Implicit Function Theorem implies that there are unique continuously differentiable functions $h_\ell = h_\ell(\mathbf{U})$, $1 \le \ell \le m$, defined on a neighborhood N of $\mathbf{U}_0$, such that

$$\begin{aligned} (h_1(\mathbf{U}), h_2(\mathbf{U}), \dots, h_m(\mathbf{U}), \mathbf{U}) \in D, \quad \text{for all } \ \mathbf{U} \in N, \\ (h_1(\mathbf{U}_0), h_2(\mathbf{U}_0), \dots, h_m(\mathbf{U}_0), \mathbf{U}_0) = \mathbf{X}_0, \end{aligned} \tag{31}$$

and

$$g_\ell(h_1(\mathbf{U}), h_2(\mathbf{U}), \dots, h_m(\mathbf{U}), \mathbf{U}) = 0, \quad \mathbf{U} \in N, \quad 1 \le \ell \le m. \tag{32}$$

Again from (30), the system

$$\begin{bmatrix} \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_m} \\ \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix} = \begin{bmatrix} f_{x_1}(\mathbf{X}_0) \\ f_{x_2}(\mathbf{X}_0) \\ \vdots \\ f_{x_m}(\mathbf{X}_0) \end{bmatrix} \tag{33}$$

has a unique solution. This implies that

$$\frac{\partial f(\mathbf{X}_0)}{\partial x_i} - \lambda_1 \frac{\partial g_1(\mathbf{X}_0)}{\partial x_i} - \lambda_2 \frac{\partial g_2(\mathbf{X}_0)}{\partial x_i} - \cdots - \lambda_m \frac{\partial g_m(\mathbf{X}_0)}{\partial x_i} = 0 \tag{34}$$

for $1 \le i \le m$.

If $m + 1 \le i \le n$, differentiating (32) with respect to x_i and recalling (31) yields

$$\frac{\partial g_\ell(\mathbf{X}_0)}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial g_\ell(\mathbf{X}_0)}{\partial x_j} \frac{\partial h_j(\mathbf{X}_0)}{\partial x_i} = 0, \quad 1 \le \ell \le m.$$

If $\mathbf{X}_0$ is local extreme point f subject to $g_1(\mathbf{X}) = g_2(\mathbf{X}) = \cdots = g_m(\mathbf{X}) = 0$, then $\mathbf{U}_0$ is an unconstrained local extreme point of $f(h_1(\mathbf{U}), h_2(\mathbf{U}), \dots h_m(\mathbf{U}), \mathbf{U})$; therefore,

$$\frac{\partial f(\mathbf{X}_0)}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial f(\mathbf{X}_0)}{\partial x_j} \frac{\partial h_j(\mathbf{X}_0)}{\partial x_i} = 0.$$

The last two equations imply that

$$\begin{vmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial f(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial f(\mathbf{X}_0)}{\partial x_m} \\ \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_m} \\ \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{vmatrix} = 0,$$

so

$$\begin{vmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_i} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_i} \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_1} \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_m} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_m} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_m} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{vmatrix} = 0.$$

Therefore, there are constant $c_0, c_1, \dots c_m$, not all zero, such that

$$\begin{bmatrix} \frac{\partial f(\mathbf{X}_0)}{\partial x_i} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_i} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_i} \\ \frac{\partial f(\mathbf{X}_0)}{\partial x_1} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_1} \\ \frac{\partial f(\mathbf{X}_0)}{\partial x_2} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f(\mathbf{X}_0)}{\partial x_m} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_m} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_m} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_3 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (35)$$

If $c_0 = 0$, then

$$\begin{bmatrix} \frac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \cdots & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_m} \\ \frac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m(\mathbf{X}_0)}{\partial x_1} & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_2} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and (30) implies that $c_1 = c_2 = \cdots = c_m = 0$; hence, we may assume that $c_0 = 1$ in a nontrivial solution of (35). Therefore,

$$\begin{bmatrix} \frac{\partial f(\mathbf{X}_0)}{\partial x_i} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_i} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_i} \\ \frac{\partial f(\mathbf{X}_0)}{\partial x_1} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_1} \\ \frac{\partial f(\mathbf{X}_0)}{\partial x_2} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f(\mathbf{X}_0)}{\partial x_m} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_m} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_m} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{bmatrix} \begin{bmatrix} 1 \\ c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (36)$$

which implies that

$$\begin{bmatrix} \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_m} \\ \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{bmatrix} \begin{bmatrix} -c_1 \\ -c_2 \\ \vdots \\ -c_m \end{bmatrix} = \begin{bmatrix} f_{x_1}(\mathbf{X}_0) \\ f_{x_2}(\mathbf{X}_0) \\ \vdots \\ f_{x_m}(\mathbf{X}_0) \end{bmatrix}$$

Since (33) has only one solution, this implies that $c_j = -\lambda_j$, $1 \le j \le n$, so (36) becomes

$$\begin{bmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_i} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_i} \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_1} \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_m} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_m} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_m} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{bmatrix} \begin{bmatrix} 1 \\ -\lambda_1 \\ -\lambda_2 \\ \vdots \\ -\lambda_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Computing the topmost entry of the vector on the left yields yields (34), which completes the proof. □

Example 9 Minimize $\displaystyle\sum_{i=1}^{n} x_i^2$ subject to

$$\sum_{i=1}^{n} a_{ri} x_i = c_r, \quad 1 \le r \le m, \tag{37}$$

where

$$\sum_{i=1}^{n} a_{ri} a_{si} = \begin{cases} 1 & \text{if } r = s, \\ 0 & \text{if } r \ne s. \end{cases} \tag{38}$$

Solution Let

$$L = \frac{1}{2} \sum_{i=1}^{n} x_i^2 - \sum_{s=1}^{m} \lambda_s \sum_{i=1}^{n} a_{si} x_i.$$

Then

$$L_{x_i} = x_i - \sum_{s=1}^{m} \lambda_s a_{si}, \quad 1 \le i \le n,$$

so

$$x_{i0} = \sum_{s=1}^{m} \lambda_s a_{si} \quad 1 \le i \le n, \tag{39}$$

and

$$a_{ri} x_{i0} = \sum_{s=1}^{m} \lambda_s a_{ri} a_{si}.$$

Now (38) implies that

$$\sum_{i=1}^{n} a_{ri} x_{i0} = \sum_{s=1}^{m} \lambda_s \sum_{i=1}^{n} a_{ri} a_{si} = \lambda_r.$$

From this and (37), $\lambda_r = c_r$, $1 \le r \le m$, and (39) implies that

$$x_{i0} = \sum_{s=1}^{m} c_s a_{si}, \quad 1 \le i \le n.$$

Therefore,

$$x_{i0}^2 = \sum_{r,s=1}^{m} c_r c_s a_{ri} a_{si}, \quad 1 \le i \le n,$$

and (38) implies that

$$\sum_{i=1}^{n} x_{i0}^2 = \sum_{r,s=1}^{m} c_r c_s \sum_{i=1}^{n} a_{ri} a_{si} = \sum_{r=1}^{m} c_r^2.$$

The next theorem provides further information on the relationship between the eigenvalues of a symmetric matrix and constrained extrema of its quadratic form. It can be proved by successive applications of Theorem 1; however, we omit the proof.

Theorem 4 *Suppose that* $\mathbf{A} = [a_{rs}]_{r,s=1}^{n} \in \mathbb{R}^{n \times n}$ *is symmetric and let*

$$Q(\mathbf{x}) = \sum_{r,s=1}^{n} a_{rs} x_r x_s.$$

Suppose also that

$$\mathbf{x}_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix}$$

minimizes Q *subject to* $\sum_{i=1}^{n} x_i^2$*. For* $2 \le r \le n$*, suppose that*

$$\mathbf{x}_r = \begin{bmatrix} x_{1r} \\ x_{2r} \\ \vdots \\ x_{nr} \end{bmatrix},$$

minimizes Q subject to

$$\sum_{i=1}^{n} x_i^2 = 1 \ \ \textit{and} \ \ \sum_{i=1}^{n} x_{is} x_i = 0, \quad 1 \le s \le r-1.$$

Denote

$$\lambda_r = \sum_{i,j=1}^{n} a_{ij} x_{ir} x_{jr}, \quad 1 \le r \le n.$$

Then

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n \ \ \textit{and} \ \ Ax_r = \lambda_r x_r, \quad 1 \le r \le n.$$

7 Exercises

1. Find the point on the plane $2x + 3y + z = 7$ closest to $(1, -2, 3)$.

2. Find the extreme values of $f(x, y) = 2x + y$ subject to $x^2 + y^2 = 5$.

3. Suppose that $a, b > 0$ and $a\alpha^2 + b\beta^2 = 1$. Find the extreme values of $f(x, y) = \beta x + \alpha y$ subject to $ax^2 + by^2 = 1$.

4. Find the points on the circle $x^2 + y^2 = 320$ closest to and farthest from $(2, 4)$.

5. Find the extreme values of

$$f(x, y, z) = 2x + 3y + z \quad \text{subject to} \quad x^2 + 2y^2 + 3z^2 = 1.$$

6. Find the maximum value of $f(x, y) = xy$ on the line $ax + by = 1$, where $a, b > 0$.

7. A rectangle has perimeter p. Find its largest possible area.

8. A rectangle has area A. Find its smallest possible perimeter.

9. A closed rectangular box has surface area A. Find it largest possible volume.

10. The sides and bottom of a rectangular box have total area A. Find its largest possible volume.

11. A rectangular box with no top has volume V. Find its smallest possible surface area.

12. Maximize $f(x, y, z) = xyz$ subject to

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1,$$

where $a, b, c > 0$.

13. Two vertices of a triangle are $(-a, 0)$ and $(a, 0)$, and the third is on the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Find its largest possible area.

14. Show that the triangle with the greatest possible area for a given perimeter is equilateral, given that the area of a triangle with sides x, y, z and perimeter s is

$$A = \sqrt{s(s-x)(s-y)(s-z)}.$$

15. A box with sides parallel to the coordinate planes has its vertices on the ellipsoid

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1.$$

Find its largest possible volume.

16. Derive a formula for the distance from (x_1, y_1, z_1) to the plane

$$ax+by+cz=\sigma.$$

17. Let $\mathbf{X}_i = (x_i, y_i, z_i)$, $1 \le i \le n$. Find the point in the plane

$$ax+by+cz=\sigma$$

for which $\sum_{i=1}^{n} |\mathbf{X}-\mathbf{X}_i|^2$ is a minimum. Assume that none of the $\mathbf{X}_i$ are in the plane.

18. Find the extreme values of $f(\mathbf{X}) = \displaystyle\sum_{i=1}^{n}(x_i - c_i)^2$ subject to $\displaystyle\sum_{i=1}^{n} x_i^2 = 1$.

19. Find the extreme values of

$$f(x,y,z) = 2xy + 2xz + 2yz \quad \text{subject to} \quad x^2+y^2+z^2=1.$$

20. Find the extreme values of

$$f(x,y,z) = 3x^2 + 2y^2 + 3z^2 + 2xz \quad \text{subject to} \quad x^2+y^2+z^2=1.$$

21. Find the extreme values of

$$f(x,y) = x^2 + 8xy + 4y^2 \quad \text{subject to} \quad x^2+2xy+4y^2=1.$$

22. Find the extreme value of $f(x,y) = \alpha + \beta xy$ subject to $(ax+by)^2 = 1$. Assume that $ab \ne 0$.

23. Find the extreme values of $f(x,y,z) = x + y^2 + 2z$ subject to

$$4x^2 + 9y^2 - 36z^2 = 36.$$

24. Find the extreme values of $f(x,y,z,w) = (x+z)(y+w)$ subject to

$$x^2+y^2+z^2+w^2=1.$$

25. Find the extreme values of $f(x,y,z,w) = (x+z)(y+w)$ subject to

$$x^2+y^2=1 \text{ and } z^2+w^2=1.$$

26. Find the extreme values of $f(x, y, z, w) = (x + z)(y + w)$ subject to

$$x^2 + z^2 = 1 \text{ and } \quad y^2 + w^2 = 1.$$

27. Find the distance between the circle $x^2 + y^2 = 1$ the hyperbola $xy = 1$.

28. Minimize $f(x, y, x) = \dfrac{x^2}{\alpha^2} + \dfrac{y^2}{\beta^2} + \dfrac{z^2}{\gamma^2}$ subject to $ax + by + cz = d$ and x, $y, z > 0$.

29. Find the distance from $(c_1, c_2, \ldots, c_n)$ to the plane

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = d.$$

30. Find the maximum value of $f(\mathbf{X}) = \sum_{i=1}^{n} a_i x_i^2$ subject to $\sum_{i=1}^{n} b_i x_i^4 = 1$, where $p, q > 0$ and $a_i, b_i\ x_i > 0, 1 \leq i \leq n$.

31. Find the extreme value of $f(\mathbf{X}) = \sum_{i=1}^{n} a_i x_i^p$ subject to $\sum_{i=1}^{n} b_i x_i^q = 1$, where p, q>0 and $a_i, b_i, x_i > 0, 1 \leq i \leq n$.

32. Find the minimum value of

$$f(x, y, z, w) = x^2 + 2y^2 + z^2 + w^2$$

subject to

$$\begin{aligned} x + y + \ \ z + 3w &= 1 \\ x + y + 2z + \ \ w &= 2. \end{aligned}$$

33. Find the minimum value of

$$f(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}$$

subject to $p_1x + p_2y + p_3z = d$, assuming that at least one of p_1, p_2, p_3 is nonzero.

34. Find the extreme values of $f(x, y, z) = p_1x + p_2y + p_3z$ subject to

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

assuming that at least one of p_1, p_2, p_3 is nonzero.

35. Find the distance from $(-1, 2, 3)$ to the intersection of the planes $x + 2y - 3z = 4$ and $2x - y + 2z = 5$.

36. Find the extreme values of $f(x, y, z) = 2x + y + 2z$ subject to $x^2 + y^2 = 4$ and $x + z = 2$.

37. Find the distance between the parabola $y = 1 + x^2$ and the line $x + y = -1$.

38. Find the distance between the ellipsoid

$$3x^2 + 9y^2 + 6z^2 = 10$$

and the plane

$$3x + 3y + 6z = 70.$$

39. Show that the extreme values of $f(x, y, z) = xy + yz + zx$ subject to

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

are the largest and smallest eigenvalues of the matrix

$$\begin{bmatrix} 0 & a^2 & a^2 \\ b^2 & 0 & b^2 \\ c^2 & c^2 & 0 \end{bmatrix}.$$

40. Show that the extreme values of $f(x, y, z) = xy + 2yz + 2zx$ subject to

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

are the largest and smallest eigenvalues of the matrix

$$\begin{bmatrix} 0 & a^2/2 & a^2 \\ b^2/2 & 0 & b^2 \\ c^2 & c^2 & 0 \end{bmatrix}.$$

41. Find the extreme values of $x(y + z)$ subject to

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

42. Let $a, b, c, p, q, r, \alpha, \beta$, and γ be positive constants. Find the maximum value of $f(x, y, z) = x^\alpha y^\beta z^\gamma$ subject to

$$ax^p + by^q + cz^r = 1 \quad \text{and} \quad x, y, z > 0.$$

43. Find the extreme values of

$$f(x, y, z, w) = xw - yz \quad \text{subject to} \quad x^2 + 2y^2 = 4 \quad \text{and} \quad 2z^2 + w^2 = 9.$$

44. Let a, b, c,and d be positive. Find the extreme values of

$$f(x,y,z,w) = xw - yz$$

subject to

$$ax^2 + by^2 = 1, \quad cz^2 + dw^2 = 1,$$

if **(a)** $ad \neq bc$; **(b)** $ad = bc$.

45. Minimize $f(x,y,z) = \alpha x^2 + \beta y^2 + \gamma z^2$ subject to

$$a_1x + a_2y + a_3z = c \text{ and } b_1x + b_2y + b_3z = d.$$

Assume that

$$\alpha, \beta, \gamma > 0, \quad a_1^2 + a_2^2 + a_3^2 \neq 0, \text{ and } b_1^2 + b_2^2 + b_3^2 \neq 0.$$

Formulate and apply a required additional assumption.

46. Minimize $f(\mathbf{X}, \mathbf{Y}) = \sum_{i=1}^{n} (x_i - \alpha_i)^2$ subject to

$$\sum_{i=1}^{n} a_i x_i = c \text{ and } \sum_{i=1}^{n} b_i x_i = d,$$

where

$$\sum_{i=1}^{n} a_i^2 = \sum_{i=1}^{n} b_i^2 = 1 \text{ and } \sum_{i=1}^{n} a_i b_i = 0.$$

47. Find $(x_{10,x_{20}}, \dots, x_{n0})$ to minimize

$$Q(\mathbf{X}) = \sum_{i=1}^{n} x_i^2$$

subject to

$$\sum_{i=1}^{n} x_i = 1 \quad \text{and} \quad \sum_{i=1}^{n} i x_i = 0.$$

Prove explicitly that if

$$\sum_{j=1}^{n} y_i = 1, \quad \sum_{i=1}^{n} i y_i = 0$$

and $y_i \neq x_{i0}$ for some $i \in \{1, 2, \dots, n\}$, then

$$\sum_{i=1}^{n} y_i^2 > \sum_{i=1}^{n} x_{i0}^2.$$

48. Let $p_1, p_2, \ldots, p_n$ and s be positive numbers. Maximize

$$f(\mathbf{X}) = (s - x_1)^{p_1}(s - x_2)^{p_2} \cdots (s - x_n)^{p_n}$$

subject to $x_1 + x_2 + \cdots + x_n = s$.

49. Maximize $f(\mathbf{X}) = x_1^{p_1} x_2^{p_2} \cdots x_n^{p_n}$ subject to $x_i > 0$, $1 \le i \le n$, and

$$\sum_{i=1}^{n} \frac{x_i}{\sigma_i} = S,$$

where $p_1, p_2, \ldots, p_n, \sigma_1, \sigma_2, \ldots, \sigma_n$, and V are given positive numbers.

50. Maximize

$$f(\mathbf{X}) = \sum_{i=1}^{n} \frac{x_i}{\sigma_i}$$

subject to $x_i > 0$, $1 \le i \le n$, and

$$x_1^{p_1} x_2^{p_2} \cdots x_n^{p_n} = V,$$

where $p_1, p_2, \ldots, p_n, \sigma_1, \sigma_2, \ldots, \sigma_n$, and S are given positive numbers.

51. Suppose that $\alpha_1, \alpha_2, \ldots \alpha_n$ are positive and at least one of $a_1, a_2, \ldots, a_n$ is nonzero. Let $(c_1, c_2, \ldots, c_n)$ be given. Minimize

$$Q(\mathbf{X}) = \sum_{i=1}^{n} \frac{(x_i - c_i)^2}{\alpha_i}$$

subject to

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = d.$$

52. Schwarz's inequality says that $(a_1, a_2, \ldots, a_n)$ and $(x_1, x_2, \ldots, x_n)$ are arbitrary n-tuples of real numbers, then

$$|a_1 x_1 + a_2 x_2 + \cdots + a_n x_n| \le (a_1^2 + a_2^2 + \cdots + a_n^2)^{1/2}(x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}.$$

Prove this by finding the extreme values of $f(\mathbf{X}) = \sum_{i=1}^{n} a_i x_i$ subject to $\sum_{i=1}^{n} x_i^2 = \sigma^2$.

53. Let $x_1, x_2, \ldots, x_m, r_1, r_2, \ldots, r_m$ be positive and

$$r_1 + r_2 + \cdots + r_m = r.$$

Show that

$$\left(x_1^{r_1} x_2^{r_2} \cdots x_m^{r_m}\right)^{1/r} \le \frac{r_1 x_1 + r_2 x_2 + \cdots r_m x_m}{r},$$

and give necessary and sufficient conditions for equality. (Hint: Maximize $x_1^{r_1} x_2^{r_2} \cdots x_m^{r_m}$ subject to $\sum_{j=1}^{m} r_j x_j = \sigma > 0$, $x_1 > 0$, $x_2 > 0, \ldots, x_m > 0$.)

54. Let $\mathbf{A} = [a_{ij}]$ be an $m \times n$ matrix. Suppose that $p_1, p_2, \dots, p_m > 0$ and

$$\sum_{j=1}^{m} \frac{1}{p_j} = 1,$$

and define

$$\sigma_i = \sum_{j=1}^{n} |a_{ij}|^{p_i}, \quad 1 \le i \le m.$$

Use Exercise 53 to show that

$$\left| \sum_{j=1}^{n} a_{ij} a_{2j} \cdots a_{mj} \right| \le \sigma_1^{1/p_1} \sigma_2^{1/p_2} \cdots \sigma_m^{1/p_m}.$$

(With $m = 2$ this is *Hölder's inequality*, which reduces to Schwarz's inequality if $p_1 = p_2 = 2$.)

55. Let $c_0, c_1, \dots, c_m$ be given constants and $n \ge m + 1$. Show that the minimum value of

$$Q(\mathbf{X}) = \sum_{r=0}^{n} x_r^2$$

subject to

$$\sum_{r=0}^{n} x_r r^s = c_s, \quad 0 \le s \le m,$$

is attained when

$$x_r = \sum_{s=0}^{m} \lambda_s r^s, \quad 0 \le r \le n,$$

where

$$\sum_{\ell=0}^{m} \sigma_{s+\ell} \lambda_\ell = c_s \quad \text{and} \quad \sigma_s = \sum_{r=0}^{n} r^s, \quad 0 \le s \le m.$$

Show that if $\{x_r\}_{r=0}^{n}$ satisfies the constraints and $x_r \ne x_{r0}$ for some r, then

$$\sum_{r=0}^{n} x_r^2 > \sum_{r=0}^{n} x_{r0}^2.$$

56. Suppose that $n > 2k$. Show that the minimum value of $f(\mathbf{W}) = \sum_{i=-n}^{n} w_i^2$, subject to the constraint

$$\sum_{i=-n}^{n} w_i P(r - i) = P(r)$$

whenever r is an integer and P is a polynomial of degree $\le 2k$, is attained with

$$w_{i0} = \sum_{r=0}^{2k} \lambda_r i^r, \quad 1 \le i \le n,$$

where

$$\sum_{r=0}^{2k} \lambda_r \sigma_{r+s} = \begin{cases} 1 & \text{if } s = 0, \\ 0 & \text{if } 1 \le s \le 2k, \end{cases} \quad \text{and} \quad \sigma_s = \sum_{j=-n}^{n} j^s.$$

Show that if $\{w_i\}_{i=-n}^{n}$ satisfies the constraint and $w_i \ne w_{i0}$ for some i, then

$$\sum_{i=-n}^{n} w_i^2 > \sum_{i=-n}^{n} w_{i0}^2.$$

57. Suppose that $n \ge k$. Show that the minimum value of $f \sum_{i=0}^{n} w_i^2$, subject to the constraint

$$\sum_{i=0}^{n} w_i P(r-i) = P(r+1)$$

whenever r is an integer and P is a polynomial of degree $\le k$, is attained with

$$w_{i0} = \sum_{r=0}^{k} \lambda_r i^r, \quad 0 \le i \le n,$$

where

$$\sum_{r=0}^{k} \sigma_{r+s} \lambda_r = (-1)^s, \quad 0 \le s \le k, \quad \text{and} \quad \sigma_\ell = \sum_{i=0}^{n} i^\ell, \quad 0 \le \ell \le 2k.$$

Show that if

$$\sum_{i=0}^{n} u_i P(r-i) = P(r+1)$$

whenever r is an integer and P is a polynomial of degree $\le k$, and $u_i \ne w_{i0}$ for some i, then

$$\sum_{i=0}^{n} u_i^2 > \sum_{i=0}^{n} w_{i0}^2.$$

58. Minimize

$$f(\mathbf{X}) = \sum_{i=1}^{n} \frac{(x_i - c_i)^2}{\alpha_i}$$

subject to

$$\sum_{i=1}^{n} a_{ir} x_i = d_r, \quad 1 \le r \le m$$

Assume that $m > 1, \alpha_1, \alpha_2, \ldots \alpha_m > 0$, and

$$\sum_{i=1}^{n} \alpha_i a_{ir} a_{is} = \begin{cases} 1 & \text{if } r = s, \\ 0 & \text{if } r \neq s. \end{cases}$$

8 Answers to selected exercises

1. $\left(\frac{15}{7} - \frac{2}{7}, \frac{25}{7}\right)$ **2.** ± 5 **3.** $1/\sqrt{ab}, -1/\sqrt{ab}$

4. (8, 16) is closest, (−8, −16) is farthest. **5.** $\pm\sqrt{53/6}$ **6.** $1/4ab$ **7.** $p^2/4$

8. $4\sqrt{A}$ **9.** $A^{3/2}/6\sqrt{6}$ **10.** $A^{3/2}/6\sqrt{3}$ **11.** $3(2V)^{2/3}$ **12.** $abc/27$

13. ab **15.** $8abc/3\sqrt{3}$

18. $(1-\mu)^2$ and $(1+\mu)^2$, where $\mu = \left(\sum_{j=1}^{n} c_j^2\right)^{1/2}$ **19.** $-1, 2$ **20.** $2, 4$

21. $-2/3, 2$ **22.** $\alpha \pm |\beta|/4|ab|$ **23.** $-\sqrt{5}, 73/16$ **24.** ± 1 **25.** ± 2

26. ± 2 **27.** $\sqrt{2} - 1$ **28.** $\dfrac{d^2}{(a\alpha)^2 + (b\beta^2) + (c\gamma)^2}$

29. $\dfrac{|d - a_1c_1 - a_2c_2 - \cdots - a_nc_n)a_i|}{\sqrt{a_1^2 + a_2^2 + \cdots a_n^2}}$ **30.** $\left(\sum_{i=1}^{n} \dfrac{a_i^2}{b_i}\right)^{1/2}$

31. $\left(\sum_{i=1}^{n} a_i^{q/(q-p)} b_i^{p/(p-q)}\right)^{1-p/q}$ is a constrained maximum if $p < q$, a constrained minimum if $p > q$

32. $689/845$ **33.** $\dfrac{d^2}{p_1^2a^2 + p_2^2b^2 + p_3^2c^2}$ **34.** $\pm(p_1^2a^2 + p_2^2b^2 + p_3^2c^2)^{1/2}$

35. $\sqrt{693/45}$ **36.** $2, 6$ **37.** $7/4\sqrt{2}$ **38.** $10\sqrt{6}/3$ **41.** $\pm|c|\sqrt{a^2+b^2}/2$

42. $\dfrac{\alpha\beta\gamma}{pqr}\left(\dfrac{\alpha}{p} + \dfrac{\beta}{q} + \dfrac{\gamma}{r}\right)^{-3}$ **43.** ± 3 **44.** **(a)** $\pm 1/\sqrt{bc}$ **(b)** $\pm 1/\sqrt{ad} = \pm 1/\sqrt{bc}$

46. $\left(c - \sum_{i=1}^{n} a_i\alpha_i\right)^2 + \left(d - \sum_{i=1}^{n} b_i\alpha_i\right)^2$ **47.** $x_{i0} = (4n + 2 - 6i)/n(n-1)$

48. $\left[\frac{(n-1)s}{P}\right]^P p_1^{p_1} p_2^{p_2} \cdots p_n^{p_n}$

49. $\left(\dfrac{S}{p_1 + p_2 + \cdots + p_n}\right)^{p_1+p_2+\cdots+p_n} (p_1\sigma_1)^{p_1}(p_2\sigma_2)^{p_2}\cdots(p_n\sigma_n)^{p_n}$

50. $(p_1 + p_2 + \cdots + p_n)\left(\dfrac{V}{(\sigma_1 p_1)^{p_1}(\sigma_2 p_2)^{p_2}\cdots(\sigma_n p_n)^{p_n}}\right)^{\frac{1}{p_1+p_2+\cdots+p_n}}$

51. $\left(d - \sum_{i=1}^{n} a_ic_i\right))^2 / \left(\sum_{i=1}^{n} a_i^2\alpha_i\right)$ **52.** $\pm\left(\sum_{i=1}^{n} a_i^2\right)^{1/2}\left(\sum_{i=1}^{n} x_{i0}^2\right)^{1/2}$

58. $\sum_{r=1}^{m}\left(d_r - \sum_{i=1}^{n} a_{ir}c_i\right)^2$