Mathematics 170
MATHEMATICAL METHODS FOR OPTIMIZATION
Finite Dimensional Optimization

Spring 2021 version

Lawrence C. Evans
Department of Mathematics
University of California, Berkeley

Contents

PREFACE

We are extremely grateful to Kurt and Evelyn Riedel for a very generous contribution to the UC Berkeley Math Department that provided financial support for the redesign of Math 170.

A second course, Math 195, will extend the finite dimensional methods of Math 170 to the calculus of variations and optimal control theory.

I have used Inkscape and SageMath for the illustrations. Much of my presentation, especially in Chapter 2, is inspired by the outstanding book [**F1**] by J. Franklin. Thanks to Milind Hegde for typing the first draft of these notes, and for writing up the exercises.

Haotian Gu was a great course assistant during the Fall, 2020 semester, saving me from countless pedagogical, managerial and technological blunders when teaching the class online.

INTRODUCTION

General mathematical optimization theory comprises (at least) three major subareas:

A. Discrete optimization

B. Finite dimensional optimization

C. Infinite dimensional optimization

This class covers the rigorous mathematical theory of finite dimensional optimization methods, using techniques from calculus, linear algebra and geometry.

The big math ideas for Math 170 are

(i) First variations

(ii) Second variations

(iii) Lagrange multipliers

(iv) Separating hyperplanes

(v) Convex functions

(vi) Value functions

(vii) Duality.

Math 170 is (mostly) about how these concepts occur throughout optimization theory. It will not discuss discrete optimization, numerical methods, software packages, or much on algorithms; definitely take CS and IEOR courses for these. Notice that the word "programming" in this course means optimization theory, not computer programming.

Why are Lagrange multipliers important? Our most important theme will be understanding Lagrange multipliers, and especially understanding when they exist.

• Lagrange multipliers are useful for computations.

• Lagrange multipliers often have physical, economic or other interpretations.

• Lagrange multipliers appear in convex duality theory.

Important lessons. Some useful insights of mathematical philosophy are that

constraints cause Lagrange multipliers to appear

and

Lagrange multipliers contain useful information.

Another piece of wisdom is that

you can never know too much about convexity.

Chapter 1

VARIATIONS

1.1. Unconstrained minimizers

In this section we are given a function $f : \mathbb{R}^n \to \mathbb{R}$.

DEFINITION. A point $x_0 \in \mathbb{R}^n$ is called a **minimizer** of f if

$$f(x_0) \le f(x) \qquad \text{for all } x \in \mathbb{R}^n.$$

We then write

$$f(x_0) = \min_{x \in \mathbb{R}^n} f(x).$$

1.1.1. First variation.

We wish to find ways to characterize minimizers.

THEOREM 1.1.1 (First variation). Assume $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable and x_0 is a minimizer. Then

$$\boxed{\nabla f(x_0) = 0.} \tag{1.1}$$

Equivalently, we have

$$\frac{\partial f}{\partial x_i}(x_0) = 0 \qquad (i = 1, \dots, n). \tag{1.2}$$

Proof. Let $y \in \mathbb{R}^n$, $y = [y_1, \dots, y_n]^T$. For $t \in \mathbb{R}$, we define

$$\phi(t) := f(x_0 + ty).$$

Then

$$\phi(0) = f(x_0) \le f(x_0 + ty) = \phi(t)$$

for each $t \in \mathbb{R}$. So 0 is a minimizer of ϕ, and consequently the chain rule from multivariable calculus (see the Appendix) lets us compute

$$0 = \phi'(0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)\, y_i.$$

This means that

$$\nabla f(x_0) \cdot y = 0 \quad \text{for all } y \in \mathbb{R}^n.$$

Take $y = \nabla f(x_0)$ to deduce that $\nabla f(x_0) = 0$. □

DEFINITION. A point $x \in \mathbb{R}^n$ is called a **critical point** (or an **extremal**) for f if

$$\nabla f(x) = 0.$$

REMARK. So if x_0 is a minimizer, then x_0 is a critical point. Simple examples show that the converse is false. □

1.1.2. Second variation.

NOTATION. A symmetric $n \times n$ matrix A is **nonnegative definite** if

$$y^T A y = \sum_{i,j=1}^{n} a_{ij} y_i y_j \geq 0 \quad \text{for all } y \in \mathbb{R}^n,$$

in which case we will write

$$A \succeq 0.$$

□

INTERPRETATION. An important theorem of linear algebra says that a *symmetric* $n \times n$ matrix has all real eigenvalues (and a corresponding orthonormal basis of eigenvectors). For such a symmetric matrix A the condition $A \succeq 0$ means that all the eigenvalues are nonnegative.

See Appendix B for how to tell when a given symmetric matrix is nonnegative definite. □

THEOREM 1.1.2 (**Second variation**). Assume $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable and x_0 is a minimizer. Then

$$\boxed{\nabla^2 f(x_0) \succeq 0.}$$

Proof. Define $\phi(t) = f(x_0 + ty)$. Since $t = 0$ is a minimizer of ϕ, we again apply the chain rule from multivariable calculus:

$$0 \leq \phi''(0) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)\, y_i y_j.$$

□

REMARKS. (i) So if x_0 is a minimizer of f, then the symmetric matrix $\nabla^2 f(x_0)$ is nonnegative definite.

(ii) If x_0 is only a **local minimum** of f, we can still conclude that $\nabla f(x_0) = 0$ and $\nabla^2 f(x_0) \succeq 0$. □

1.2. Applications

The simple mathematical ideas in the previous section have lots of interesting applications. Following is a selection.

1.2.1. Refraction and reflection.

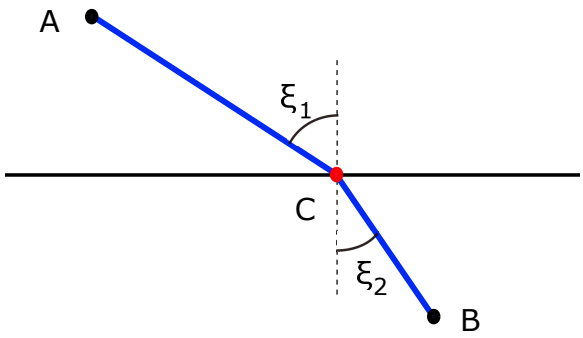
a. Refraction. Consider two points $A = (x_1, y_1)$ and $B = (x_2, y_2)$ lying in $\mathbb{R}^2$, with $y_2 < 0 < y_1$ and $x_1 < x_2$. Assume that the line $\{y = 0\}$ is an interface between two translucent materials, so that light moves with the speed v_1 in the upper half plane $\{y > 0\}$ and with speed v_2 in the lower half plane $\{y < 0\}$.

According to **Fermat's principle**, the light ray moves along a path that takes the least time to travel from A to B. We wish to describe this path, by introducing the point $C = (x, 0)$, as drawn, where the light ray hits the x-axis. Then if we set

$$d_1 = |AC| = ((x - x_1)^2 + y_1^2)^{\frac{1}{2}},\ d_2 = |BC| = ((x - x_2)^2 + y_2^2)^{\frac{1}{2}},$$

the time it takes to travel along the piecewise straight path from A to B via the point C is

$$f(x) = \frac{d_1}{v_1} + \frac{d_2}{v_2}.$$



Refraction

We calculate that

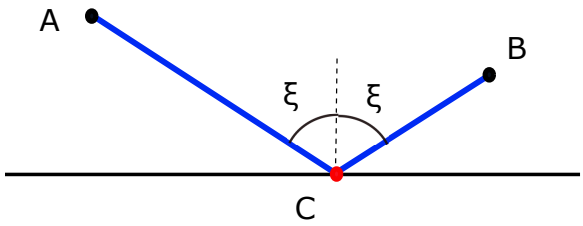
$$f'(x) = ((x-x_1)^2 + y_1^2)^{-\frac{1}{2}} \frac{x-x_1}{v_1} + ((x-x_2)^2 + y_2^2)^{-\frac{1}{2}} \frac{x-x_2}{v_2}$$
$$= \frac{\sin \xi_1}{v_1} - \frac{\sin \xi_2}{v_2},$$

for the angles ξ_1, ξ_2 as illustrated. The path of least time therefore passes through the point $C = (x_0, 0)$ for which $f'(x_0) = 0$ and therefore

$$\boxed{\frac{\sin \xi_1}{v_1} = \frac{\sin \xi_2}{v_2}.}$$

This is **Snell's Law** for the diffraction of a light ray at the interface of two materials within which light moves a different speeds.

b. Reflection. Similar considerations show that if a light ray goes between the points $A = (x_1, y_1)$ and $B = (x_2, y_2)$, where now $y_1, y_2 > 0$, by reflecting off the x-axis as illustrated, then the angles ξ of incidence and reflection agree. In this case we minimize the total length of the path from A to the x axis, and then back to B.



Reflection

□

1.2.2. Steiner trees.

Suppose we are given three non-colinear points a^1, a^2, a^3 lying in the plane $\mathbb{R}^2$. We wish find another point s, called the **Steiner point**, so that the sum of the lengths of the segments $[a^1, s], [a^2, s], [a^3, s]$ is as small as possible. What can we deduce about the geometry of this configuration?

LEMMA 1.2.1. If the point s is not one of the points a^1, a^2, a^3, then the angles between the line segments $[a^1, s], [a^2, s], [a^3, s]$ are all equal to $\frac{2\pi}{3}$.

Proof. 1. We may assume, upon relabelling the coordinates if necessary, that $s = 0$ and $a^1, a^2, a^3 \neq 0$. Thus $x_0 = 0$ minimizes the function

$$f(x) = |x - a^1| + |x - a^2| + |x - a^3|,$$

and therefore

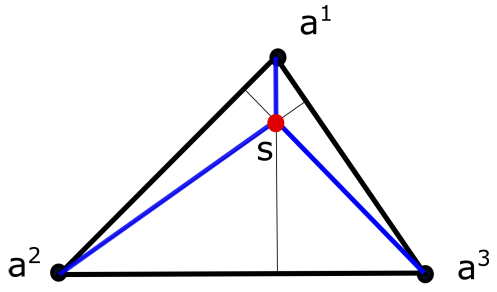
$$0 = \nabla f(0) = -\left(\frac{a^1}{|a^1|} + \frac{a^2}{|a^2|} + \frac{a^3}{|a^3|}\right).$$

Hence the three unit vectors

$$a = \frac{a^1}{|a^1|},\ b = \frac{a^2}{|a^2|},\ c = \frac{a^3}{|a^3|}$$

satisfy $a + b + c = 0$. We can therefore rearrange these unit vectors to form the sides of a equilateral triangle. It follows that the angle between each pair of the vectors a, b, c is $\frac{2\pi}{3}$. □

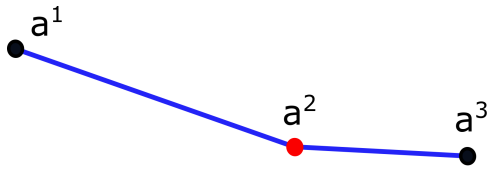
REMARK. Consider the triangle determined by a^1, a^2, a^3. If each angle is less than $\frac{2\pi}{3}$, a geometry argument shows that the Steiner point s is the intersection of the three line segments passing through each vertex and perpendicular to the opposite side.



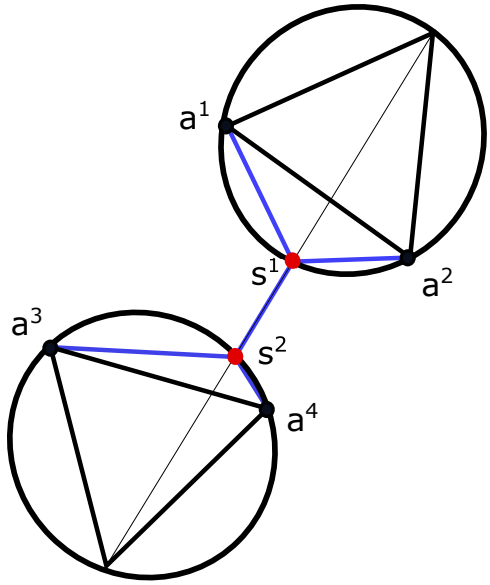
Shortest connection for 3 vertices

However, if some angle in the triangle is greater than or equal to $\frac{2\pi}{3}$, the Steiner point is the vertex of this angle, as drawn.

□



Steiner point at a vertex of angle $\geq \frac{2\pi}{3}$



Using equilateral triangles to construct a Steiner tree (m=4).

REMARK. Suppose next that we wish to interconnect a collection $\{a_k\}_{k=1}^m$ given points lying in the plane by a network of line segments of minimum total length. This is called a **Steiner tree**. The theorem above shows that at each triple junction the line segments meet at equal angles. Melzak [**M**, page 140] shows that at most $m-2$ additional vertices need to be added. Melzak also explains a geometric construction for the additional vertices needed to build the Steiner tree for the 4 points, as drawn. □

1.2.3. Electric circuits.

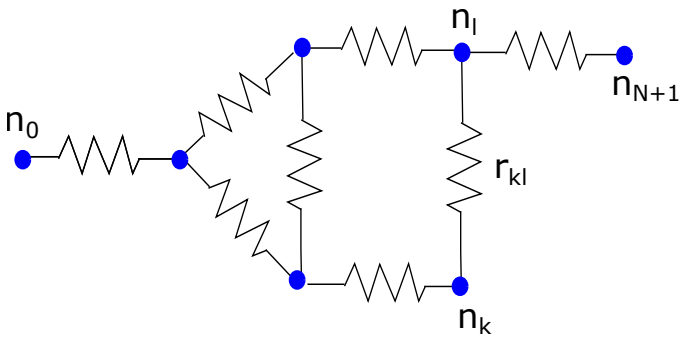
Consider an electric circuit comprising $N+2$ nodes $\{n_k\}_{k=0}^{N+1}$, some of which are connected by resistors. If there is a resistor connecting nodes n_k and n_l, we write

$$r_{kl} = r_{lk} > 0$$

for its **resistance**, and otherwise set $r_{kl} = \infty$. The corresponding **conductance** is

$$\sigma_{kl} = \frac{1}{r_{kl}} \geq 0.$$

We connect an external battery that provides a voltage difference of E across nodes n_0 and n_{N+1}. What then are the voltage differences between the other nodes? What are the currents across the resistors?



A network of resistors

This physical problem is **variational**, meaning that it can be cast as the optimization problem of minimizing an electrostatic energy function, defined for the voltages $v = [v_1, \ldots, v_N]^T$ as

$$e(v) = \frac{1}{2} \sum_{0 \leq k < l \leq N+1} \sigma_{kl}(v_k - v_l)^2, \tag{1.3}$$

where $v_0 = E$ and $v_{N+1} = 0$. We minimize the energy by setting

$$0 = \frac{\partial e}{\partial v_m}(v) = \sum_{k=0}^{N+1} \sigma_{km}(v_m - v_k) \qquad (m = 1, \ldots, N) \tag{1.4}$$

In accordance with **Ohm's law**, we define the **current** flowing from node m to node k to be

$$i_{mk} = \frac{v_m - v_k}{r_{km}} = \sigma_{km}(v_m - v_k). \tag{1.5}$$

Then (1.4) gives **Kirchhoff's law**

$$\sum_{k=0}^{N+1} i_{mk} = 0 \qquad (m = 1, \ldots, N),$$

asserting that the net current flow into and out of each internal node is 0. □

1.2.4. Least squares.

Let A denote an $m \times n$ matrix and assume $b \in \mathbb{R}^m$ is given. If the linear system

$$Ax = b$$

has no solution, we can nevertheless build an approximate solutions by finding $x_0 \in \mathbb{R}^n$ that solves the minimization problem

$$\min_{x \in \mathbb{R}^n} |Ax - b|^2. \tag{1.6}$$

THEOREM 1.2.1. If $A^T A$ is invertible, the unique solution of (1.6) is

$$\boxed{x_0 = (A^T A)^{-1} A^T b.} \tag{1.7}$$

Proof. Let $f(x) = |Ax - b|^2 = |Ax|^2 - 2Ax \cdot b + |b|^2$. Observe first that

$$\nabla(Ax \cdot b) = \nabla(x \cdot A^T b) = A^T b.$$

It is not hard to check that $\nabla(x \cdot Sx) = 2Sx$ if S is a symmetric $n \times n$ matrix. Observe also that the rule $(BC)^T = C^T B^T$ implies $A^T A$ is symmetric. Consequently,

$$\nabla|Ax|^2 = \nabla(Ax \cdot Ax) = \nabla(x \cdot A^T Ax) = 2A^T Ax.$$

Therefore

$$\nabla f(x) = 2(A^T Ax - A^T b);$$

and so $\nabla f(x_0) = 0$ implies (1.7).

We can use the Extreme Value Theorem (see the Appendix) to show that f attains its minimum at least one point, which then must be given by (1.7). □

We extend these ideas to study **linear approximation of data points**. Given **data points** $\{(x^k, y^k) \mid k = 1, \dots, N\}$, where each $x^k \in \mathbb{R}^n$, $y^k \in \mathbb{R}$, the **least squares** problem is to find $m_0 \in \mathbb{R}^n$, $b_0 \in \mathbb{R}$ so that the plane

$$y = m_0 \cdot x + b_0$$

minimizes the total **mean square error**

$$f(m, b) = \frac{1}{2} \sum_{k=1}^{N} (y^k - (m \cdot x^k + b))^2 = \frac{1}{2} \sum_{k=1}^{N} e_k^2, \tag{1.8}$$

where

$$e^k = y^k - (m \cdot x^k + b) \qquad (k = 1, \dots, N).$$

NOTATION. (i) We write

$$\bar{x} = \frac{1}{N}\sum_{k=1}^{N} x^k, \quad \bar{y} = \frac{1}{N}\sum_{k=1}^{N} y^k,$$

for the respective **averages** of $\{x^k\}_{k=1}^N$ and $\{y^k\}_{k=1}^N$.

(ii) The **covariance matrix** is the symmetric $n \times n$ matrix C whose (i,j)-th entry is

$$c_{ij} = \frac{1}{N}\sum_{k=1}^{N}(x_i^k - \bar{x}_i)(x_j^k - \bar{x}_j) \tag{1.9}$$

where $x^k = [x_1^k, \ldots, x_n^k]^T$ and $\bar{x} = [\bar{x}_1, \ldots, \bar{x}_n]^T$.

(iii) Define $d = [d_1, \ldots, d_n]^T$ by

$$d_i = \frac{1}{N}\sum_{k=1}^{N}(x_i^k - \bar{x}_i)(y^k - \bar{y}). \tag{1.10}$$

THEOREM 1.2.2. If C is invertible, then a minimizer (m_0, b_0) of f satisfies

$$\boxed{\begin{cases} m_0 = C^{-1}d \\ b_0 = \bar{y} - m_0 \cdot \bar{x}. \end{cases}} \tag{1.11}$$

Proof. 1. We search for a minimizer by first setting the partial derivatives of f to 0:

$$\begin{cases} 0 = \dfrac{\partial f}{\partial b}(m,b) & (1.12) \\ 0 = \dfrac{\partial f}{\partial m_i}(m,b) \qquad (i = 1 \ldots, n)\ . & (1.13) \end{cases}$$

These are $n+1$ equations for $n+1$ unknowns $b, m_1, \ldots, m_n$.

2. Using (1.12), we see that

$$0 = \frac{\partial f}{\partial b} = \sum_{k=1}^{N} e_k \frac{\partial e_k}{\partial b} = \sum_{k=1}^{N} e_k(-1).$$

Consequently,

$$\sum_{k=1}^{N} m \cdot x^k + b - y^k = 0. \tag{1.14}$$

Now look at (1.13):

$$0 = \frac{\partial f}{\partial m_i} = \sum_{k=1}^{N} e_k \frac{\partial e_k}{\partial m_i} = \sum_{k=1}^{N} e_k(-x_i^k),$$

since $e_k = y^k - (m \cdot x^k + b) = y^k - (\sum_{j=1}^n m_j x_j^k + b)$. So

$$\sum_{k=1}^{N} (m \cdot x^k + b - y^k) x_i^k = 0 \qquad (i = 1, \ldots, n). \tag{1.15}$$

3. Next we solve (1.14) and (1.15) for b and m. Equation (1.14) says

$$m \cdot \bar{x} + b = \bar{y}. \tag{1.16}$$

Using (1.16) in (1.15), we see that

$$\sum_{k=1}^{N} m \cdot (x^k - \bar{x}) x_i^k = \sum_{k=1}^{N} (y^k - \bar{y}) x_i^k \qquad (i = 1, \ldots, n).$$

Then

$$\sum_{k=1}^{N} m \cdot (x^k - \bar{x})(x_i^k - \bar{x}_i) = \sum_{k=1}^{N} (y^k - \bar{y})(x_i^k - \bar{x}_i) \qquad (i = 1, \ldots, n), \tag{1.17}$$

since $\sum_{k=1}^N m \cdot (x^k - \bar{x}) = \sum_{k=1}^N (y^k - \bar{y}) = 0$.

The identities (1.16) and (1.17) imply that a minimizer (m_0, b_0) solves $m_0 \cdot \bar{x} + b_0 = \bar{y}$ and $Cm_0 = d$. □

1.3. Equality constraints

We turn our attention now to optimization problems with constraints. Assume $f, g_1, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ are twice continuously differentiable functions.

NOTATION. We define $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$ by

$$\mathbf{g} = \begin{bmatrix} g_1 \\ \vdots \\ g_m \end{bmatrix}.$$

The gradient of $\mathbf{g}$ is

$$\nabla \mathbf{g} = \begin{bmatrix} (\nabla g_1)^T \\ \vdots \\ (\nabla g_m)^T \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix}.$$

This is an $m \times n$ matrix-valued function. □

Our constrained optimization problem is to find $x_0 \in \mathbb{R}^n$ to

(MIN) $\boxed{\text{minimize } f, \text{ subject to } \mathbf{g} = 0.}$

The equations $g_k(x) = 0$ for $k = 1, \ldots, m$ are **equality constraints**.

DEFINITION. A point $x \in \mathbb{R}^n$ is called **feasible** for (MIN) if $\mathbf{g}(x) = 0$; that is, if $g_k(x) = 0$ for $k = 1, \ldots, m$.

Hereafter we assume that x_0 solves (MIN); our goal is to find ways to characterize x_0.

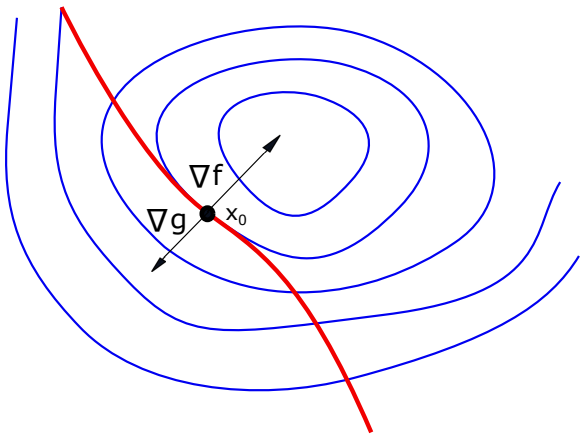
1.3.1. Lagrange multipliers.

The **method of Lagrange multipliers** reveals a linear relationship between the gradients of $f, g_1, \ldots, g_m$ at a minimizer x_0.

We discuss now three derivations, each informative, but none rigorous, for the case of a single constraint ($m = 1$):

a. Geometric interpretation. Imagine that we drive a car along a road and want to find a point of lowest elevation. The road is given by a curve, drawn as red in the picture, and the constant elevation contours are the curves drawn in blue.

We can drive back and forth along the road until we find the lowest point x_0. What is a geometric characterization of this location? The key observation is that *the red road is tangent to the blue elevation contour at* x_0. For if not, then by driving either forward or backwards a little bit on the road we could get to a lower elevation.



Highest elevation along a road

We convert this insight into a mathematical formula, by introducing the height function f. This means that $f(x)$ is the altitude of any point x and the blue curves are level curves for f. We can introduce as well a function g such that the red road is a level curve of g, say the curve $\{g = 0\}$. In mathematical terms, the point x_0 minimizes the function f, subject to the constraint $g = 0$.

Now a key insight from multivariable calculus is that *the gradient of a function at a point x is perpendicular to the level curve for the function passing through x.* Therefore, since the level curves of f and g are tangent at x_0, *the gradient vectors $\nabla f(x_0)$ and $\nabla g(x_0)$ must be parallel.* This in turn means that there exists a constant λ_0 such that

$$\nabla f(x_0) + \lambda_0 \nabla g(x_0) = 0. \tag{1.18}$$

We call λ_0 a **Lagrange multiplier**. □

b. Fewer variables interpretation. Another idea is to convert our constrained minimization problem (MIN) into an unconstrained problem in fewer variables. So assume that x_0 solves (MIN) and that near x_0 we can rewrite the constraint equation by solving for one of the variables, say x_n:

$$g(x) = 0 \iff x_n = \phi(x')$$

where $x' = [x_1, \dots, x_{n-1}]^T$ and x' is near x_0'. Thus

$$g(x', \phi(x')) = 0.$$

We differentiate this identity, to deduce

$$\frac{\partial g}{\partial x_i} + \frac{\partial g}{\partial x_n}\frac{\partial \phi}{\partial x_i} = 0 \qquad (i = 1, \dots, n-1). \tag{1.19}$$

Since x_0 solves (MIN), x_0' is an unconstrained minimizer of

$$f(x', \phi(x')).$$

Hence

$$\frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial x_n}\frac{\partial \phi}{\partial x_i} = 0 \qquad (i = 1, \dots, n-1) \tag{1.20}$$

at x_0'.

Now define

$$\lambda_0 = -\frac{\partial f}{\partial x_n}(x_0)\left(\frac{\partial g}{\partial x_n}(x_0)\right)^{-1}; \tag{1.21}$$

it then follows from (1.19) and (1.20) that the Lagrange multiplier formula (1.18) holds.

c. Value function interpretation. For each $a \in \mathbb{R}$ let us define the **value function**

$$v(a) = \min_{y \in \mathbb{R}^n} \{f(y) \mid g(y) = a\}.$$

Then for all $x \in \mathbb{R}^n$ we have

$$v(g(x)) \leq f(x),$$

and furthermore

$$v(g(x_0)) = v(0) = f(x_0).$$

Consequently, the function

$$h(x) = f(x) - v(g(x))$$

has an unconstrained minimum at x_0. Therefore

$$0 = \nabla h(x_0) = \nabla f(x_0) - v'(0)\nabla g(x_0).$$

This says that the Lagrange multiplier formula (1.18) holds for

$$\lambda_0 = -v'(0).$$

We see also that the Lagrange multiplier λ_0 is the negative of the derivative of the value function at $a = 0$. □

EXAMPLE. (Lagrange multipliers may not exist.) Unfortunately, it is easy to come up with examples for which all three heuristic derivations above completely fail.

For instance, let $n = 2$, $m = 1$,

$$f(x) = x_1 + x_2, \quad g(x) = (x_1)^2 + (x_2)^2.$$

The solution of (MIN) is clearly $x_0 = [0\,0]^T$, since this is the only feasible point. But $\nabla f(x_0) = [1\,1]^T$, $\nabla g(x_0) = 0$ and therefore

$$\nabla f(x_0) + \lambda \nabla g(x_0) \neq 0$$

for all real numbers λ. *There is no Lagrange multiplier!*

What went wrong? Why do the three derivations discussed above fail for this example?

a. The geometric derivation is not valid since the set of feasible x is a point and not a smooth curve. So it is not true that "the level curves of f and g are tangent at x_0".

b. The trick of converting to fewer variables similarly fails. In particular, we cannot define λ_0 by (1.21) since $\frac{\partial g}{\partial x_2}(x_0) = 0$.

c. The value function derivation also does not work, since we do not know that the value function v is differentiable at $a = 0$. In fact, v is not

even finite near 0. (There are no points $x \in \mathbb{R}^n$ satisfying $g(x) = a$ if $a < 0$, and the infimum of the empty set is $+\infty$.) □

1.3.2. Constrained first variation.

In light of the example above, we need to think much more carefully about when, and if, Lagrange multipliers exist.

THEOREM 1.3.1. Suppose that x_0 solves the constrained optimization problem (MIN).

Then there exist real numbers $\gamma_0, \lambda_0^1, \dots, \lambda_0^m$ (**not all equal to 0**) such that

$$\boxed{\gamma_0 \nabla f(x_0) + \sum_{k=1}^{m} \lambda_0^k \nabla g_k(x_0) = 0.} \tag{1.22}$$

REMARK. This is **F. John's form of the constrained first variation formula**, in which we interpret each λ_0^k as the Lagrange multiplier for the constraint $g_k(x_0) = 0$. We can also write (1.22) as

$$\gamma_0 \nabla f(x_0) + \nabla \mathbf{g}(x_0)^T \lambda_0 = 0$$

for the **vector of Lagrange multipliers**

$$\lambda_0 = [\lambda_0^1 \dots \lambda_0^m]^T.$$

When $\gamma_0 = 0$, we call it an **abnormal multipller**. If $\gamma_0 \neq 0$, it is a **normal multiplier**, in which case we can divide and convert to the case $\gamma_0 = 1$ (for possibly new constants $\lambda_0^1, \dots, \lambda_0^m$.) □

Proof. 1. Fix $\beta > 0$. For each $\alpha > 0$ define

$$F^\alpha(x) := f(x) + \frac{\alpha}{2}|\mathbf{g}(x)|^2 + \frac{\beta}{2}|x - x_0|^2. \tag{1.23}$$

We will later send $\alpha \to \infty$; this procedure is the **penalty method**.

Let $B = \{x \in \mathbb{R}^n \mid |x - x_0| \leq 1\}$ denote the closed ball of radius 1 and center x_0. Since the function $F^\alpha : B \to \mathbb{R}$ is continuous and B is closed and bounded, the Extreme Value Theorem (see the Appendix) tells us that there exists a point $x_\alpha \in B$ such that

$$F^\alpha(x_\alpha) = \min_{x \in B} F^\alpha(x).$$

Thus

$$f(x_\alpha) + \frac{\alpha}{2}|\mathbf{g}(x_\alpha)|^2 + \frac{\beta}{2}|x_\alpha - x_0|^2 = F^\alpha(x_\alpha) \leq F^\alpha(x_0) = f(x_0), \tag{1.24}$$

since $\mathbf{g}(x_0) = 0$. Therefore

$$\{\alpha|\mathbf{g}(x_\alpha)|^2\}_{\alpha>0}$$

is bounded, and consequently

$$\lim_{\alpha\to\infty} \mathbf{g}(x_\alpha) = 0. \tag{1.25}$$

2. Next, we use the Bolzano-Weierstrass Theorem (see Appendix) to select a convergent subsequence $\{x_{\alpha_j}\}_{j=1}^{\infty}$ of $\{x_\alpha\}_{\alpha>0} \subset B$ so that

$$x_{\alpha_j} \to \bar{x}$$

as $\alpha_j \to \infty$, for some $\bar{x} \in B$. Then (1.24) gives

$$f(\bar{x}) + \frac{\beta}{2}|\bar{x} - x_0|^2 \le f(x_0). \tag{1.26}$$

But (1.25) implies $\mathbf{g}(\bar{x}) = 0$ and therefore $\bar{x}$ is feasible. Hence $f(x_0) \le f(\bar{x})$ since x_0 solves (MIN). Thus (1.26) tells us that $|\bar{x} - x_0|^2 = 0$ and therefore $\bar{x} = x_0$. This is true for all convergent subsequences $x_{\alpha_j} \to \bar{x}$ and consequently

$$\lim_{\alpha\to\infty} x_\alpha = x_0. \tag{1.27}$$

3. So if α is large enough, F^α has a minimum over B at the point x_α and, in view of (1.27), x_α does not lie on the boundary of B. It follows that

$$0 = \nabla F^\alpha(x_\alpha) = \nabla f(x_\alpha) + \alpha \nabla \mathbf{g}(x_\alpha)^T \mathbf{g}(x_\alpha) + \beta(x_\alpha - x_0). \tag{1.28}$$

We used here the formula

$$\nabla\left(\frac{|\mathbf{g}|^2}{2}\right) = \nabla \mathbf{g}^T \mathbf{g}.$$

Then

$$0 = \gamma_\alpha \nabla f(x_\alpha) + \nabla \mathbf{g}(x_\alpha)^T \lambda_\alpha + \gamma_\alpha \beta(x_\alpha - x_0), \tag{1.29}$$

for

$$\gamma_\alpha = (1 + \alpha^2|\mathbf{g}(x_\alpha)|^2)^{-\frac{1}{2}},\ \lambda_\alpha = \gamma_\alpha \alpha \mathbf{g}(x_\alpha).$$

4. Now $\gamma_\alpha^2 + |\lambda_\alpha|^2 = 1$ and therefore $\{(\gamma_\alpha, \lambda_\alpha)\}_{\alpha>0}$ is bounded. Consequently, the Bolzano-Weierstrass Theorem asserts that there is a sequence $\alpha_j \to \infty$ such that

$$\gamma_{\alpha_j} \to \gamma_0 \quad \text{in } \mathbb{R}, \quad \lambda_{\alpha_j} \to \lambda_0 \quad \text{in } \mathbb{R}^m,$$

and $(\gamma_0, \lambda_0) \neq (0, 0)$. Let $\alpha = \alpha_j \to \infty$ in (1.29) and recall (1.27) to derive (1.22). □

REMARKS. (i) This proof uses ideas from McShane [**MS**]; see also Bertsekas [**B**] and Borwein–Lewis [**B-L**]. The term β played no role, but will be used in the later proof of Theorem 1.3.3.

(ii) The existence theory for Lagrange multipliers for **inequality constraints** of the form $h_j(x) \leq 0$ for $j = 1, \ldots, p$ is more complicated, and will be discussed later. □

1.3.3. Regular points.

In the abnormal case, the first variation formula (1.22) does not involve the function f we are minimizing and consequently may not be very useful. Throughout these notes we will therefore be interested in situations for which we have a normal multiplier $\gamma_0 = 1$.

DEFINITION. We say the point x_0 is **regular** if the vectors

$$\{\nabla g_k(x_0)\}_{k=1}^m \text{ are linearly independent in } \mathbb{R}^n.$$

THEOREM 1.3.2. Suppose that x_0 solves the constrained optimization problem (MIN) and furthermore that x_0 is regular.

Then there exist real numbers $\lambda_0^1, \ldots, \lambda_0^m$ such that

$$\boxed{\nabla f(x_0) + \sum_{k=1}^m \lambda_0^k \nabla g_k(x_0) = 0.} \tag{1.30}$$

REMARK. This is the usual form of the **constrained first variation formula**, which we can also write as

$$\boxed{\nabla f(x_0) + \nabla \mathbf{g}(x_0)^T \lambda_0 = 0.}$$

□

Proof. We know from (1.22) that

$$\gamma_0 \nabla f(x_0) + \sum_{k=1}^m \lambda_0^k \nabla g_k(x_0) = 0 \tag{1.31}$$

for constants γ_0 and $\lambda_0^1, \ldots, \lambda_0^m$ that are not all zero.

We claim that if x_0 is regular, then $\gamma_0 \neq 0$. To see this, suppose instead that $\gamma_0 = 0$; then

$$\sum_{k=1}^m \lambda_0^k \nabla g_k(x_0) = 0$$

and $[\lambda_0^1, \ldots, \lambda_0^m]^T \neq 0$. But this is impossible, since the vectors $\{\nabla g_k(x_0)\}_{k=1}^m$ are independent.

We can therefore divide (1.31) by $\gamma_0 \neq 0$, to obtain an expression of the form (1.30) (for possibly different constants $\lambda_0^1, \ldots, \lambda_0^m$). □

REMARK. (Linear equality constraints) For many important problems the equality constraints are linear or affine, and therefore of the form

$$\mathbf{g}(x) = Ax - b = 0 \tag{1.32}$$

where A is a non-zero $m \times n$ matrix and $b \in \mathbb{R}^m$. For such constraints, the regularity condition means that the rows of the matrix A are linearly independent.

This need not necessarily be so for a given non-zero matrix A, but *we can always convert to an equivalent problem with affine constraints for which the regularity condition holds.*

To see this, we apply elementary row operations and transform the linear system of constraints $Ax = b$ into an equivalent system of linear equations, having the form

$$\bar{A}x = \bar{b} \tag{1.33}$$

where $\bar{A}$ is a $l \times n$ matrix for some integer $l \in \{1, \ldots, m\}$, $\bar{b} \in \mathbb{R}^l$ and $\bar{A}$ has full rank l. (If we cannot convert (1.32) into this form, then there are no feasible x satisfying (1.32).) Now every point is regular for the equivalent problem of minimizing f subject to the constraints

$$\bar{\mathbf{g}}(x) = \bar{A}x - \bar{b} = 0.$$

Consequently, Theorem 1.3.2 guarantees the existence of Lagrange multipliers such that

$$\nabla f(x_0) + \sum_{j=1}^{l} \bar{\lambda}_0^j \bar{a}^j = 0.$$

Here $\{\bar{a}^j\}_{j=1}^l$ are the rows of $\bar{A}$. Since the rows of $\bar{A}$ are linear combinations of the rows of A, we can rewrite this as

$$\nabla f(x_0) + \sum_{k=1}^{m} \lambda_0^k a_k = 0.$$

for appropriate Lagrange multipliers and the rows $\{a_k\}_{k=1}^m$ of A. □

1.3.4. Constrained second variation.

We discuss next how to compute second variations when we have regular equality constraints.

LEMMA 1.3.1. If x_0 is regular, then the $m \times m$ matrix

$$\nabla \mathbf{g}(x_0) \nabla \mathbf{g}(x_0)^T$$

is nonsingular.

Proof. Suppose $y \in \mathbb{R}^m$ and

$$\nabla \mathbf{g}(x_0)\nabla \mathbf{g}(x_0)^T y = 0.$$

We must show $y = 0$. To do so, take the inner product with y:

$$0 = y \cdot \nabla \mathbf{g}(x_0)\nabla \mathbf{g}(x_0)^T y = |\nabla \mathbf{g}(x_0)^T y|^2.$$

(We used here the linear algebra formula $(Ax) \cdot y = x \cdot (A^T y)$.) Therefore

$$0 = \nabla \mathbf{g}(x_0)^T y = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial x_n} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \sum_{k=1}^m y_k \nabla g_k(x_0).$$

Since the vectors $\{\nabla g_k(x_0)\}_{k=1}^m$ are independent, it follows that $y = 0$. □

THEOREM 1.3.3. Suppose that x_0 solves the constrained optimization problem (MIN) and that x_0 is regular. Let $\lambda_0^1, \ldots, \lambda_0^m$ be corresponding Lagrange multipliers, satisfying the first variation formula (1.30).

Then

$$\boxed{y^T \left(\nabla^2 f(x_0) + \sum_{k=1}^m \lambda_0^k \nabla^2 g_k(x_0) \right) y \geq 0} \tag{1.34}$$

for all $y \in \mathbb{R}^n$ such that

$$\nabla \mathbf{g}(x_0) y = 0. \tag{1.35}$$

This is the **constrained second variation formula**.

Proof. 1. We return to the proof of Theorem 1.3.1 and extract more detailed information. Since x_0 is regular, we know from Theorem 1.3.2 that

$$\gamma_\alpha = (1 + \alpha^2 |\mathbf{g}(x_\alpha)|^2)^{-\frac{1}{2}} \to \gamma_0 \neq 0.$$

Hence $\{\alpha |\mathbf{g}(x_\alpha)|\}_{\alpha > 0}$ is bounded. We can therefore assume, passing if necessary to a further subsequence, that

$$\alpha_j \mathbf{g}(x_{\alpha_j}) \to \lambda_0. \tag{1.36}$$

2. Since x_α lies within the interior of the ball B for large α, we have

$$0 \leq z^T \nabla^2 F^\alpha(x_\alpha) z \quad \text{for all } z \in R^n.$$

Now

$$\nabla^2 F^\alpha(x) = \nabla^2 f(x) + \alpha \nabla \mathbf{g}(x)^T \nabla \mathbf{g}(x) + \alpha \sum_{k=1}^m g_k(x) \nabla^2 g_k(x) + \beta I;$$

hence

$$z^T\Bigg(\nabla^2 f(x_\alpha)+\alpha\nabla\mathbf{g}(x_\alpha)^T\nabla\mathbf{g}(x_\alpha) + \alpha\sum_{k=1}^{m} g_k(x_\alpha)\nabla^2 g_k(x_\alpha)+\beta I\Bigg)z\geq 0. \tag{1.37}$$

Remember from the Lemma that $\nabla\mathbf{g}(x_0)\nabla\mathbf{g}(x_0)^T$ is nonsingular; consequently $\nabla\mathbf{g}(x_\alpha)\nabla\mathbf{g}(x_\alpha)^T$ is invertible for large α. Given now $y\in\mathbb{R}^m$ with

$$\nabla\mathbf{g}(x_0)y=0, \tag{1.38}$$

define

$$z_\alpha := y-\nabla\mathbf{g}(x_\alpha)^T\left(\nabla\mathbf{g}(x_\alpha)\nabla\mathbf{g}(x_\alpha)^T\right)^{-1}\nabla\mathbf{g}(x_\alpha)y. \tag{1.39}$$

Then

$$\nabla\mathbf{g}(x_\alpha)z_\alpha=0.$$

Observe also that

$$z_\alpha\to y \quad \text{as } \alpha\to\infty.$$

This follows since $x_\alpha\to x_0$ and $\nabla\mathbf{g}(x_0)y=0$.

3. Now put $z=z_\alpha$ in (1.37):

$$0\leq (z_\alpha)^T\left(\nabla^2 f(x_\alpha)+\alpha\sum_{k=1}^{m} g_k(x_\alpha)\nabla^2 g_k(x_\alpha)+\beta I\right)z_\alpha. \tag{1.40}$$

Let $\alpha\to\infty$ in (1.40) and recall (1.36):

$$0\leq y^T\left(\nabla^2 f(x_0)+\sum_{k=1}^{m}\lambda_0^k\nabla^2 g_k(x_0)+\beta I\right)y.$$

To conclude, send $\beta\to 0$. □

1.4. Applications

We present next some interesting applications and interpretations of Lagrange multipliers, and many other examples appear later in these notes. Nahin's book [**N**] has lots of other fascinating applications of finding maxima and minima.

1.4.1. Fluid equilibria.

Lagrange multipliers often have physical interpretations, as we illustrate in this neat example from equilibrium fluid mechanics (M. Levi, SIAM News, May, 2020).

Suppose that we have sitting on a table n differently shaped containers, which we connect with small tubes at their bases, to allow water to flow freely between them. We then pour in a volume V of water, which freely flows among the containers. We will use Lagrange multipliers to demonstrate a version of **Pascal's principle**, that at equilibrium *the heights of the fluid in each container are all equal*, regardless of the shapes of the individual vessels. .

To see this, let us denote by $a_k(y) > 0$ the cross sectional area of the k-th container at distance y above the table. Then if x_k is the height of water in that vessel, the volume of water within is

$$\int_0^{x_k} a_k(y)\, dy$$

and the gravitational potential energy of the water in that container is proportional to

$$\int_0^{x_k} y a_k(y)\, dy.$$

The guiding physical principle is that the water will minimize the total potential energy

$$f(x_1, \dots, x_n) = \sum_{k=1}^{n} \int_0^{x_k} y a_k(y)\, dy,$$

subject to the volume constraint

$$g(x_1, \dots, x_n) = \sum_{k=1}^{n} \int_0^{x_k} a_k(y)\, dy - V = 0.$$

Since $\nabla g \neq 0$, the optimal heights $x_0 = [x_1^0, \dots, x_n^0]^T$ satisfy

$$\nabla f(x_0) + \lambda_0 \nabla g(x_0) = 0 \tag{1.41}$$

for an appropriate Lagrange multiplier. Since

$$\frac{\partial f}{\partial x_k} = x_k a_k(x_k), \quad \frac{\partial g}{\partial x_k} = a_k(x_k),$$

(1.41) tells us that

$$x_k^0 a_k(x_k^0) + \lambda_0 a_k(x_k^0) = 0 \quad (k = 1, \dots, n);$$

and so

$$x_k^0 = -\lambda_0 \quad (k = 1, \dots, n).$$

In particular, *the Lagrange multiplier is (minus) the common height of the water in the various containers.*

1.4.2. Maximizing entropy.

Statistical physics often studies questions about maximizing entropy or minimizing free energy. For the simplest such problems, we introduce the set

$$P = \left\{ p \in \mathbb{R}^m \mid p_i \geq 0 \ (i = 1, \ldots, m), \sum_{i=1}^m p_i = 1 \right\}$$

of probability distributions on the integers $\{1, \ldots, m\}$.

DEFINITION. For each $p \in P$, the corresponding **Shannon entropy** is

$$e(p) = -\sum_{i=1}^m p_i \log p_i, \tag{1.42}$$

with the convention that $0 \log 0 = 0$.

A fundamental question in statistical physics is to find which distributions $p_0 \in P$ maximize the entropy, subject to various constraints.

• **Uniform distribution.** First suppose we have no additional constraints beyond those in the definition of P. We assume also that we can ignore the inequality constraints $p \geq 0$ (which turn out to hold automatically for a minimizer). Then there exists a Lagrange multiplier λ so that

$$-\nabla e(p) + \lambda \nabla \left(\sum_{i=1}^m p_i \right) = 0.$$

This implies

$$\log p_i + 1 + \lambda = 0 \qquad (i = 1, \ldots, m).$$

So all the p_i are equal, and therefore the uniform distribution $p_0 = [\frac{1}{m}, \ldots, \frac{1}{m}]^T$ maximizes the entropy.

• **Gibbs distribution.** More interesting distributions appear if we add additional constraints. So assume we are given positive constants $\{E_i\}_{i=1}^m$ and interpret E_i as the "energy level" of the state $i \in \{1, \ldots, m\}$.

THEOREM 1.4.1. Let $p_0 = [p_0^1, \ldots, p_0^m]^T$ give the maximum of the entropy over P, subject to the additional constraint that the expected energy is given:

$$\sum_{i=1}^m p_i E_i = E. \tag{1.43}$$

Then there exists a constant β such

$$p_0^i = \frac{e^{-\beta E_i}}{Z} \qquad (i = 1, \dots, m) \tag{1.44}$$

for

$$Z = \sum_{i=1}^{m} e^{-\beta E_i}. \tag{1.45}$$

Physicists call (1.44) the **Gibbs** (or **Boltzmann**) **distribution** and (1.45) the **partition function**.

Proof. Since the constraints are linear, according to the Remark on page 19 there exist Lagrange multipliers λ and β such that

$$-\nabla e(p_0) + \lambda \nabla \left(\sum_{i=1}^{m} p_i^0 \right) + \beta \nabla \left(\sum_{i=1}^{m} p_i^0 E_i - E \right) = 0$$

for a maximizer p_0. Therefore

$$\log p_0^i + 1 + \lambda + \beta E_i = 0$$

and so (1.44) follows, provided we select the normalizing constant Z so that $\sum_{i=1}^{m} p_0^i = 1$. □

1.4.3. More on electric circuits.

We next return to the electric circuit example on page 8 and show how to recast the problem with the currents, and not the voltages, as the unknowns.

So given as before the $N+2$ nodes connected by resistors, we denote by i_{kl} the **current** flowing from node k to node l. Then

$$i_{kk} = 0, \quad i_{kl} = -i_{lk} \qquad (k, l = 0, \dots, N+1). \tag{1.46}$$

We furthermore assume that Kirchhoff's law holds at the internal nodes:

$$\sum_{l=0}^{N+1} i_{ml} = 0 \qquad (m = 1, \dots, N). \tag{1.47}$$

Let the total current flowing through the network (from node n_0 to node n_{N+1}) be I; then

$$\sum_{l=0}^{N+1} i_{0l} = I, \quad \sum_{k=0}^{N+1} i_{k,N+1} = I. \tag{1.48}$$

Hereafter, we introduce the variables

$$i = \{i_{kl} \mid 0 \le k < l \le N+1\}$$

and use (1.46) to define i_{kl} for $0 \le l \le k \le N+1$. Define also the functions

$$f(i) = \frac{1}{2} \sum_{0 \le k < l \le N+1} r_{kl} (i_{kl})^2$$

and

$$g_m(i) = \sum_{l=0}^{N+1} i_{ml} \qquad (m = 0, \dots, N+1).$$

Our new variational problem is

$$\begin{cases} \text{minimize } f(i), \text{ subject to} \\ g^0(i) = I, g^{N+1}(i) = -I, g_m(i) = 0 \quad (m = 1, \dots, N). \end{cases}$$

Assume that i is a minimizing selection of currents. Then, since our constraints are linear, there exist Lagrange multipliers $\lambda^0, \dots, \lambda^{N+1}$ such that

$$\nabla f(i) + \sum_{m=0}^{N+1} \lambda^m \nabla g_m(i) = 0. \tag{1.49}$$

PHYSICAL INTERPRETATION. What are the meanings of the Lagrange multipliers and of formula (1.49)? We compute that

$$\frac{\partial f}{\partial i_{kl}} = r_{kl} i_{kl} \qquad (0 \le k < l \le N+1).$$

Furthermore, since

$$g_m(i) = \sum_{l=0}^{N+1} i_{ml} = -\sum_{l=0}^{m} i_{lm} + \sum_{l=m}^{N+1} i_{ml},$$

we have for $0 \le k < l \le N+1$ that

$$\frac{\partial g_m}{\partial i_{kl}} = \begin{cases} 1 & \text{if } m = k < l \\ -1 & \text{if } k < l = m \\ 0 & \text{otherwise.} \end{cases}$$

Therefore (1.49) implies

$$r_{kl} i_{kl} + \lambda_k - \lambda^l = 0 \tag{1.50}$$

for the indices $0 \le k < l \le N+1$. Define now the **voltages**

$$v_k = -\lambda_k \qquad (k = 0, \dots, N+1);$$

then (1.50) says

$$i_{kl} = \frac{v_k - v_l}{r_{kl}}. \tag{1.51}$$

This is again Ohm's law, and our new variational principle shows that *the voltages are the Lagrange multipliers for the constraints imposed by Kirchhoff's law.* □

1.4.4. Roots of polynomials.

A novel application of Lagrange multipliers (from de Jong [**dJ**]) shows the existence of a root for a complex polynomial of degree $n \geq 1$:

$$f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 \qquad (z \in \mathbb{C}).$$

The coefficients $a_0, \ldots, a_{n-1}$ here are complex numbers.

If we substitute $z = x + iy$ and expand, we can write f in terms of its real and imaginary parts

$$f(z) = u(x, y) + iv(x, y) \qquad (x, y \in \mathbb{R}), \tag{1.52}$$

where $u, v : \mathbb{R}^2 \to \mathbb{R}$ are polynomials.

LEMMA 1.4.1. The functions u, v solve the **Cauchy-Riemann equations**

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \tag{1.53}$$

Proof. We first apply induction to $f(z) = z^n$. The case $n = 1$ is clear. Now write

$$z^n = u_n(x, y) + iv_n(x, y)$$

and assume (1.53) holds for u_n, v_n. Then

$$z^{n+1} = (x + iy)(u_n + iv_n) = (xu_n - yv_n) + i(xv_n + yu_n) = u_{n+1} + iv_{n+1}.$$

So the induction hypothesis lets us compute

$$\begin{aligned} \frac{\partial u_{n+1}}{\partial x} &= u_n + x\frac{\partial u_n}{\partial x} - y\frac{\partial v_n}{\partial x} \\ &= u_n + x\frac{\partial v_n}{\partial y} + y\frac{\partial v_n}{\partial y} = \frac{\partial v_{n+1}}{\partial y}. \end{aligned}$$

Likewise $\frac{\partial u_{n+1}}{\partial y} = -\frac{\partial v_{n+1}}{\partial x}$. This proves (1.53) for $f(z) = z^n$ and the proof for general polynomials follows by linearity. □

THEOREM 1.4.2 (**Fundamental Theorem of Algebra**)**.** There exists a point $z_0 \in \mathbb{C}$ for which

$$f(z_0) = 0. \tag{1.54}$$

Proof. 1. We introduce for $c \in \mathbb{R}$ the level sets

$$L_c = \{(x,y) \in \mathbb{R}^2 \mid u(x,y) = c\}, \quad M_c = \{(x,y) \in \mathbb{R}^2 \mid v(x,y) = c\}.$$

Since $u(x,0) = x^n +$ lower order terms, the function $u(x,0)$ takes on infinitely many values. It follows that the sets L_c are nonempty for infinitely many values of the parameter c.

We observe next that except for finitely many values of c, we have

$$\nabla u \neq 0 \quad \text{on } L_c, \quad \nabla v \neq 0 \quad \text{on } M_c. \tag{1.55}$$

To see this, note that f' is a polynomial, and consequently has at most finitely many zeros. But $f' = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}$, according to the Cauchy-Riemann equations (1.53), and therefore $\nabla u \neq 0$ except at finitely many points. This shows the first assertion of (1.55), and the second assertion has a similar proof.

2. Select a value of the parameter c so that $L_c \neq \emptyset$ and $\nabla u \neq 0$ on L_c. We introduce the constrained optimization problem

$$\begin{cases} \text{minimize } v^2, \\ \text{subject to } u = c. \end{cases} \tag{1.56}$$

Since $u^2 + v^2 = |f|^2 \to \infty$ as $|z| \to \infty$, we can use the Extreme Value Theorem to show that there exists a point $(x_c, y_c) \in L_c$ solving (1.56).

We claim that

$$v(x_c, y_c) = 0. \tag{1.57}$$

To see this, note that since $\nabla u \neq 0$ on L_c, (x_c, y_c) is a regular point. Hence Theorem 1.3.2 asserts that there exists a Lagrange multiplier λ such that

$$2v(x_c,y_c)\nabla v(x_c,y_c) + \lambda \nabla u(x_c,y_c) = 0.$$

But the Cauchy-Riemann equations (1.53) imply $\nabla u \cdot \nabla v = 0$ and $|\nabla u| = |\nabla v|$. Since $\nabla u(x_c, y_c) \neq 0$, it follows that $\lambda = 0$ and (1.57) holds.

3. In view of (1.57), we see that

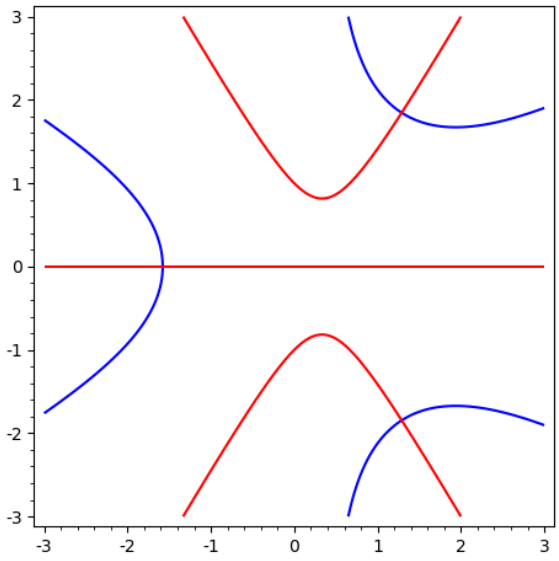
$$M_0 \neq \emptyset.$$

Select a sequence $c_k \to 0$, such that $M_{c_k} \neq \emptyset$ and $\nabla v \neq 0$ on M_{c_k}. Then the argument above (with the roles of u and v reversed) shows that there exist points $(x_k, y_k) \in M_{c_k}$ for which $u(x_k, y_k) = 0$. The sequence $\{(x_k, y_k)\}_{k=1}^{\infty}$ is bounded, and so, passing if necessary to a subsequence, we may assume

$$(x_k, y_k) \to (x_0, y_0) \in M_0.$$

Then $u(x_0, y_0) = v(x_0, y_0) = 0$, and therefore $f(z_0) = 0$ for $z_0 = x_0 + iy_0$. $\square$

EXAMPLE. If $f(z) = z^3 - z^2 + z + 8$, then $f = u + iv$ for $u = x^3 - 3xy^2 - x^2 + y^2 + x + 8$ and $v = 3x^2y - y^3 - 2xy + y$.



The three roots of $f(z) = 0$ are the intersections of the level sets $\{u = 0\}$ (drawn in blue) and $\{v = 0\}$ (drawn in red). □

Chapter 2

LINEAR OPTIMIZATION

Linear optimization theory, most commonly known as **linear programming**, concerns the minimization of linear functions, subject to affine equality and inequality constraints. We will follow Franklin [**F1, F2**] for much of the presentation.

2.1. Theory

2.1.1. Basic concepts.

NOTATION. (i) If $x = [x_1, \ldots, x_n]^T \in \mathbb{R}^n$, we write

$$x \geq 0$$

to mean that $x_i \geq 0$ for $i = 1, \ldots, n$. Similarly,

$$x > 0$$

means $x_i > 0$ $(i = 1, \ldots, n)$.

(ii) If $x, y \in \mathbb{R}^n$, we write

$$x \geq y$$

if $x_i \geq y_i$ $(i = 1, \ldots, n)$, and

$$x > y$$

if $x_i > y_i$ $(i = 1, \ldots, n)$.

DEFINITION. Let $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and assume A is an $m \times n$ matrix. The **canonical primal linear programming problem** is to find $x_0 \in \mathbb{R}^n$ to

$$\text{(P)} \qquad \boxed{\begin{cases} \text{minimize } c \cdot x, \\ \text{subject to } Ax = b, x \geq 0. \end{cases}}$$

DEFINITION. We say $x \in \mathbb{R}^n$ is **feasible** if $Ax = b$, $x \geq 0$, that is, if x satisfies the constraints in (P). We will often call a feasible x a **feasible solution**.

DEFINITION. The **canonical dual problem** is to find $y_0 \in \mathbb{R}^m$ to

$$\text{(D)} \qquad \boxed{\begin{cases} \text{maximize } b \cdot y \\ \text{subject to } A^T y \leq c. \end{cases}}$$

DEFINITION. We say $y \in \mathbb{R}^m$ is **feasible** for (D) if $A^T y \leq c$.

EXAMPLE. Consider the problem

$$\begin{cases} \text{minimize } x_1 + 2x_2 + 3x_3, \text{ subject to} \\ \quad x_1 - 2x_2 + x_3 = 4 \\ \quad -x_1 + 3x_2 \quad\quad = 5 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{cases}$$

Here $n = 3$ and $m = 2$. In the above notation,

$$c = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & -2 & 1 \\ -1 & 3 & 0 \end{bmatrix}.$$

The dual problem is

$$\begin{cases} \text{maximize } 4y_1 + 5y_2, \text{ subject to} \\ \quad y_1 - y_2 \leq 1 \\ -2y_1 + 3y_2 \leq 2 \\ \quad y_1 \quad\quad \leq 3. \end{cases}$$

There are no inequality constraints on y_1, y_2. □

The most important fact of linear programming is that *the primal and dual problems contain information about each other*:

THEOREM 2.1.1 (**Duality and optimality**).

(i) If x is feasible for (P) and y is feasible for (D), then

$$b \cdot y \leq c \cdot x \tag{2.1}$$

(ii) If x_0 is feasible for (P), y_0 is feasible for (D), and **if**

$$\boxed{b \cdot y_0 = c \cdot x_0,} \tag{2.2}$$

then x_0 solves (P) and y_0 solves (D).

Proof. 1. Let x, y be feasible. Then $Ax = b, x \geq 0$ and $A^T y \leq c$. Now $A^T y \leq c$ means $(A^T y)_i \leq c_i$ for $i = 1, \ldots, n$, and $x \geq 0$ means $x_i \geq 0$ for $i = 1, \ldots, n$. Consequently,

$$y \cdot b = y \cdot (Ax) = A^T y \cdot x = \sum_{i=1}^{n} x_i (A^T y)_i \leq \sum_{i=1}^{n} x_i c_i = x \cdot c.$$

2. Suppose x_0, y_0 are feasible and

$$b \cdot y_0 = c \cdot x_0.$$

By (2.1), $b \cdot y_0 \leq c \cdot x$ for all feasible x for (P). So $c \cdot x_0 \leq c \cdot x$ for all feasible x, and thus x_0 is optimal for (P). A similar argument works for y_0. □

EXAMPLE. Consider the problem (P) to

$$\begin{cases} \text{minimize } x_1 + 5x_2 + 2x_3 + 13x_4, \text{subject to} \\ \qquad 5x_1 - 6x_2 + 4x_3 - 2x_4 = 0 \\ \qquad\quad x_1 - x_2 + 6x_3 + 9x_4 = 16 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0. \end{cases}$$

Look at $x_0 = [0, 2, 3, 0]^T$. It is easy to check that $x_0 \geq 0$, $Ax_0 = b$, and we calculate that $c \cdot x_0 = 16$.

But is x_0 optimal? To answer, we look at the dual problem (D):

$$\begin{cases} \text{maximize } 0 \cdot y_1 + 16 \cdot y_2, \text{ subject to} \\ \quad 5y_1 + y_2 \leq 1 \\ -6y_1 - y_2 \leq 5 \\ \quad 4y_1 + 6y_2 \leq 2 \\ -2y_1 + 9y_2 \leq 13. \end{cases}$$

Let us guess that $y_0 = [-1, 1]^T$, and check $A^T y_0 \leq c$. We see that $c \cdot y_0 = 16$. Since $b \cdot x_0 = c \cdot y_0$, it follows that x_0 and y_0 are each optimal. □

OTHER FORMS OF LINEAR PROGRAMMING PROBLEMS

DEFINITION. The **standard linear programming problem** is to find $x_0 \in \mathbb{R}^n$ to

$$(P^*) \qquad \boxed{\begin{cases} \text{minimize } c \cdot x \\ \text{subject to } Ax \geq b, x \geq 0. \end{cases}}$$

Observe that both constraints are now inequalities.

DEFINITION. The **dual standard linear programming problem** is to find $y_0 \in \mathbb{R}^m$ to

$$(D^*) \qquad \boxed{\begin{cases} \text{maximize } b \cdot y \\ \text{subject to } A^T y \leq c, y \geq 0. \end{cases}}$$

Note carefully that we now have the additional sign constraint $y \geq 0$. The reader should check that the duality and optimality theorem (Theorem 2.1.1) applies to (P*) and (D*).

DEFINITION. The **general linear programming problem** is to find $x_0 \in \mathbb{R}^n$ to

$$(P^\circ) \qquad \begin{cases} \text{minimize } c \cdot x, \\ \text{subject to } \begin{cases} \sum_{j=1}^n a_{ij} x_j \geq b_i & (i \in I_1) \\ \sum_{j=1}^n a_{ij} x_j = b_i & (i \in I_2) \\ x_j \geq 0 & (j \in J_1), \end{cases} \end{cases}$$

where $I_1 \cup I_2 = I$, $I_1 \cap I_2 = \emptyset$, $J_1 \subseteq J$, $I = \{1, \ldots, m\}$ and $J = \{1, \ldots, n\}$. Here I_1, J_1 are the indices of the inequality constraints.

NOTATION. We write

$$[A, b, c, I_1, J_1]$$

to display the relevant information determining (P°). □

DEFINITION. The **dual of the general linear programming problem** is to find $y_0 \in \mathbb{R}^m$ to

$$(D^\circ) \qquad \begin{cases} \text{maximize } b \cdot y, \\ \text{subject to } \begin{cases} \sum_{i=1}^m y_i a_{ij} \leq c_j & (j \in J_1) \\ \sum_{i=1}^m y_i a_{ij} = c_j & (j \in J_2) \\ y_i \geq 0 & (i \in I_1), \end{cases} \end{cases}$$

where $J_1 \cup J_2 = J$, $J_1 \cap J_2 = \emptyset$, $I_1 \subseteq I$, $I = \{1, \ldots, m\}$ and $J = \{1, \ldots, n\}$.

REMARK. The canonical problem (P) and its dual (D) correspond to $J_1 = J$, $J_2 = \emptyset$, $I_2 = I$, $I_1 = \emptyset$.

The standard problem (P*) and its dual (D*) correspond to $J_1 = J$, $J_2 = \emptyset$, $I_1 = I$, $I_2 = \emptyset$. □

THEOREM 2.1.2. (Linear programming duality) The dual of (D$^\circ$) is (P$^\circ$)

Proof. The problem (D$^\circ$) is equivalent to minimizing $(-b) \cdot y$ subject to

$$\begin{cases} \sum_{i=1}^m y_i(-a_{ij}) \geq -c_j & (j \in J_1) \\ \sum_{i=1}^m y_i(-a_{ij}) = -c_j & (j \in J_2) \\ y_i \geq 0 & (i \in I_1). \end{cases}$$

This is $[-A^T, -c, -b, J_1, I_1]$. So duality converts

$$\underbrace{[A, b, c, I_1, J_1]}_{(\mathrm{P}^\circ)} \longrightarrow \underbrace{[-A^T, -c, -b, J_1, I_1]}_{(\mathrm{D}^\circ)}.$$

Hence the dual of (D$^\circ$) is

$$[-(-A^T)^T, -(-b), -(-c), I_1, J_1] = \underbrace{[A, b, c, I_1, J_1]}_{(\mathrm{P}^\circ)}. \quad \square$$

REMARK. (Changing linear programming problems) By adding new variables, we can in fact convert a general linear programming problem (P$^\circ$) into the canonical form (P).

Step 1: Define the **slack variables**

$$z_i = \sum_{j=1}^n a_{ij}x_j - b_i \geq 0 \quad (i \in I_1).$$

Then (P$^\circ$) becomes

$$\begin{cases} \sum_{j=1}^n a_{ij}x_j - z_i = b_i & (i \in I_1) \\ \sum_{j=1}^n a_{ij}x_j = b_i & (i \in I_2) \\ x_j \geq 0 & (j \in J_1) \\ z_i \geq 0 & (i \in I_1). \end{cases} \tag{2.3}$$

Step 2: For $j \in J_2 = J \setminus J_1$, we introduce the **surplus variables**

$$x_j = u_j - v_j \quad (j \in J_2)$$

where $u_j \geq 0, v_j \geq 0$. Now replace in (2.3) all occurrences of x_j for $j \in J_2$ with $u_j - v_j$. Thus

$$\begin{cases} x_j \geq 0 & (j \in J_1) \\ u_j \geq 0 & (j \in J_2) \\ v_j \geq 0 & (j \in J_2), \end{cases}$$

and we now have a problem of the canonical form (P).

As a consequence of these observations, *when we study the theory of linear programming, it is enough to consider the canonical problem (P) and its dual (D).* □

2.1.2. Equilibrium equations.

Recall the primal and dual problems

$$\text{(P)} \begin{cases} \min c \cdot x, \\ \text{subject to} \\ Ax = b, x \geq 0. \end{cases} \qquad \text{(D)} \begin{cases} \max b \cdot y, \\ \text{subject to} \\ A^T y \leq c. \end{cases}$$

THEOREM 2.1.3 (**Equilibrium equations**)**.** Suppose x is feasible for (P) and y is feasible for (D).

Then x and y are optimal if and only if they satisfy the **equilibrium equations**

$$\text{(E)} \qquad \boxed{\sum_{i=1}^{m} y_i a_{ij} = c_j \quad \text{if} \quad x_j > 0} \qquad (j = 1, \ldots, n).$$

REMARKS. Equivalently, (E) says

$$\sum_{i=1}^{m} y_i a_{ij} < c_j \quad \text{implies} \quad x_j = 0 \qquad (j = 1, \ldots, n).$$

The equilibrium equations are sometimes referred to as the **complementary slackness** conditions. This means that if the constraint $x_j \geq 0$ is slack (that is, if $x_j > 0$), then the complementary constraint $\sum_{i=1}^{m} y_i a_{ij} \leq c_j$ is tight (that is, $\sum_{i=1}^{m} y_i a_{ij} = c_j$). □

Proof. As before, compute

$$b \cdot y = Ax \cdot y = x_0 \cdot A^T y \leq c \cdot x.$$

When do we have equality in the last inequality?

Note that

$$x \cdot (A^T y - c) = \sum_{j=1}^{n} x_j \left((A^T y)_j - c_j \right) \begin{cases} = 0 & \text{if (E) holds} \\ < 0 & \text{if (E) fails,} \end{cases}$$

since $(A^T y)_j = \sum_{i=1}^{m} y_i a_{ij}$. Thus $b \cdot y = c \cdot x$ precisely when (E) holds, and consequently (E) is equivalent to the optimality of both x and y. □

REMARK. A similar result holds for the standard form linear programming problems

$$(\mathrm{P}^*) \begin{cases} \min\, c \cdot x, \\ \text{subject to} \\ Ax \geq b, x \geq 0. \end{cases} \qquad (\mathrm{D}^*) \begin{cases} \max\, b \cdot y, \\ \text{subject to} \\ y \geq 0, A^T y \leq c. \end{cases}$$

The equilibrium equations for these problems read

$$(\mathrm{E}^*) \qquad \boxed{\begin{cases} \text{(i)} \quad \sum_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{y_i a_{ij}} = \mathrm{c_j} \quad \text{if} \quad \mathrm{x_j} > 0 \qquad (\mathrm{j} = 1, \ldots, \mathrm{n}) \\ \text{(ii)} \quad \sum_{\mathrm{j}=1}^{\mathrm{n}} \mathrm{a_{ij} x_j} = \mathrm{b_i} \quad \text{if} \quad \mathrm{y_i} > 0 \qquad (\mathrm{i} = 1, \ldots, \mathrm{m}). \end{cases}}$$

So if x, y are feasible for (P*), (D*), then x, y are optimal if and only if (E*) holds. This is true since then

$$b \cdot y = Ax \cdot y = x \cdot A^T y = x \cdot c,$$

the first equality holding according to (ii) of (E*) and the last according to (i). □

EXAMPLE. (Continued from page 31) We have $x_0 = [0\,2\,3\,0]^T$ and $y_0 = [-1\,1]^T$. Observe that

$$A^T y_0 = \begin{bmatrix} -4 \\ 5 \\ 2 \\ 11 \end{bmatrix} \leq \begin{bmatrix} 1 \\ 5 \\ 2 \\ 13 \end{bmatrix}.$$

Then, as predicted by (E), we have $(A^T y_0)_j = c_j$ precisely when the corresponding entry of x^0 is positive, that is, for $j = 2, 3$. □

EXAMPLE. Let us use (E*) to find an optimal solution of

$$\begin{cases} \text{minimize } 7x_1 - 9x_2 - 16x_3, \\ \text{subject to } 2 \leq x_1 + 2x_2 + 9x_3 \leq 7 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{cases}$$

This has the form (P*) for

$$c = \begin{bmatrix} 7 \\ -9 \\ -16 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 2 & 9 \\ -1 & -2 & -9 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ -7 \end{bmatrix}.$$

The dual (D*) is

$$\begin{cases} \text{maximize } b \cdot y \\ \text{subject to } A^T y \leq c, y \geq 0. \end{cases}$$

Therefore

$$\begin{cases} y_1 - y_2 \leq 7 \\ 2y_1 - 2y_2 \leq -9 \\ 9y_1 - 9y_2 \leq -16. \end{cases}$$

Consequently, $y_1 - y_2 \leq 7, -\frac{9}{2}, -\frac{16}{9}$ and this implies $y_1 - y_2 \leq \min\{7, -\frac{9}{2}, -\frac{16}{9}\} = -\frac{9}{2}$. Let us look for y_1, y_2 with $y_1 - y_2 = -\frac{9}{2}$. Then

$$b \cdot y = 2y_1 - 7y_2 = 2y_1 - 7\left(y_1 + \frac{9}{2}\right) = -5y_1 - \frac{63}{2}.$$

We maximize this expression by taking $y_1 = 0, y_2 = \frac{9}{2}$. So we guess that $y_0 = [0, \frac{9}{2}]^T$. This gives $b \cdot y_0 = -\frac{63}{2}$.

How do we find x_0? Since

$$A^T y_0 = \begin{bmatrix} 1 & -1 \\ 2 & -2 \\ 9 & -9 \end{bmatrix} \begin{bmatrix} 0 \\ \frac{9}{2} \end{bmatrix} = \begin{bmatrix} -\frac{9}{2} \\ -9 \\ -\frac{81}{2} \end{bmatrix} \leq \begin{bmatrix} 7 \\ -9 \\ -16 \end{bmatrix},$$

it follows from (E*) that $x_0 = [0, x_2, 0]^T$. Consequently, $c \cdot x_0 = -9x_2$ and $2 \leq 2x_2 \leq 7$. So we select $x_2 = \frac{7}{2}$ to minimize $c \cdot x_0 = -\frac{63}{2} = b \cdot y_0$. Hence $x_0 = [0, \frac{7}{2}, 0]^T$ and $y_0 = [0, \frac{9}{2}]^T$ are optimal. □

2.1.3. Basic solutions.

We introduce next the concept of basic solutions to linear programming problems. These are solutions with the largest numbers of zero entries, which are consequently the easiest to study.

NOTATION. It is often useful to display the columns of A by writing

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} = \underbrace{\left[a^1 \mid a^2 \mid \cdots \mid a^n\right]}_{\text{columns}}.$$

Here

$$a^j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

is the j-th column vector of A, and so $a^j \in \mathbb{R}^m$ for $j = 1, \ldots, n$. □

LEMMA 2.1.1. If $Ax = b$, then b is a linear combination of the columns of A.

Proof. We can rewrite $Ax = b$ as

$$x_1 a^1 + \ldots + x_n a^n = b,$$

and this shows b to be a linear combination of the column vectors. □

DEFINITION. We say that $x \in \mathbb{R}^n$ is a **basic solution** of $Ax = b$ if the columns

$$\left\{a^j \mid x_j \neq 0, j = 1, \ldots, n\right\}$$

are linearly independent in $\mathbb{R}^m$.

Also, we say $x = 0$ is basic.

In particular, if $x \in \mathbb{R}^n$ is a basic solution of $Ax = b$, then x has at most m non-zero entries. This can be important, since often $m \ll n$ in linear programming.

THEOREM 2.1.4. For each $b \in \mathbb{R}^m$ the linear system of equations

$$Ax = b$$

has at most finitely many basic solutions.

Proof. Look at the columns $\{a^1, \ldots, a^n\}$ of A. There are only finitely many subsets $\{a^{j_1}, \ldots, a^{j_l}\} \subseteq \{a^1, \ldots, a^n\}$ which are independent in $\mathbb{R}^m$. Take such a subset $\{a^{j_1}, a^{j_2}, \ldots, a^{j_l}\}$.

We claim there is at most one solution of $Ax = b$ having the form $x = [0\, x_{j_1}\, 0\, x_{j_2} \cdots x_{j_l}\, 0]^T$. To see this, let $\hat{x} = [0\, \hat{x}_{j_1}\, 0\, \hat{x}_{j_2} \cdots \hat{x}_{j_l}\, 0]^T$ also solve $A\hat{x} = b$. Then

$$A(x - \hat{x}) = b - b = 0$$

and therefore

$$\sum_{k=1}^{l} (x_{j_k} - \hat{x}_{j_k}) a^{j_k} = 0.$$

Since the columns $\{a^{j_1}, \ldots, a^{j_l}\}$ are linearly independent, it follows that

$$x_{j_1} = \hat{x}_{j_1}, \ldots, \; x_{j_l} = \hat{x}_{j_l}.$$

Hence there is at most one basic solution of $Ax = b$ corresponding to each independent collection of columns.

□

Recall that we sometimes say that x is a **feasible solution** of (P) if $Ax = b, x \geq 0$. This is just another way of saying that x is feasible.

THEOREM 2.1.5 (**Basic solutions**).

(i) If there exists a feasible solution of (P), then there exists a basic feasible solution.

(ii) If there exists an optimal solution of (P), then there exists a basic optimal solution.

Proof. 1. Select a feasible solution x with fewest number of non-zero components. I will show it is a basic feasible solution.

If $x = 0$, we are done. If not, $x = [0\, x_{j_1}\, 0\, x_{j_2} \cdots x_{j_l}\, 0]^T$, with $x_{j_1}, \ldots, x_{j_l} > 0$ and

$$Ax = \sum_{k=1}^{l} x_{j_k} a^{j_k} = b. \tag{2.4}$$

Suppose x is not basic. Then there exist $\theta_{j_1}, \theta_{j_2}, \ldots, \theta_{j_l}$, not all equal to 0, such that

$$\sum_{k=1}^{l} \theta_{j_k} a^{j_k} = 0. \tag{2.5}$$

This means that $A\theta = 0$, where $\theta = [0\, \theta_{j_1}\, 0 \cdots \theta_{j_l}\, 0]^T$. Then (2.4) and (2.5) imply for any λ that

$$\sum_{k=1}^{l} (x_{j_k} - \lambda \theta_{j_k}) a^{j_k} = b.$$

We may assume $\theta_{j_p} > 0$ for some index j_p (if not, multiply θ by -1). Increase λ from $\lambda = 0$ to the first $\lambda^* > 0$ for which at least one of the values

$$x_{j_1} - \lambda^* \theta_{j_1}, \ \ldots, \ x_{j_l} - \lambda^* \theta_{j_l}$$

equals zero. Since $\theta_{j_p} > 0$, this must happen at a finite value of λ^*. Then

$$x^* = \begin{bmatrix} 0 \\ x_{j_1} - \lambda^* \theta_{j_1} \\ 0 \\ \vdots \\ x_{j_l} - \lambda^* \theta_{j_l} \\ 0 \end{bmatrix} = x - \lambda^* \theta$$

satisfies $x^* \geq 0$, $Ax^* = b$, and x^* has at least one fewer non-zero entry than x. This is a contradiction, and therefore x is indeed a basic feasible solution.

2. Now let x_0 be an optimal solution with the fewest number of non-zero components. I will show that x_0 is a basic optimal solution.

Suppose not. Then, as above, $x_0 = [0, x_{j_1}, 0, \ldots, x_{j_l}, 0]^T$ with $x_{j_1}, \ldots, x_{j_l} > 0$, and

$$\sum_{k=1}^{l} x_{j_k} a^{j_k} = b, \quad \sum_{k=1}^{l} \theta_{j_k} a^{j_k} = 0$$

for appropriate $\theta_{j_1}, \theta_{j_2}, \ldots, \theta_{j_l}$, not all equal to 0. Select λ^* as before and write

$$x_0^* = x_0 - \lambda^* \theta.$$

Then $Ax_0^* = b, x_0^* \geq 0$, and x_0^* has fewer non-zero components than x_0.

3. We now claim

$$c \cdot x_0^* = c \cdot x_0 = \min\{c \cdot x | Ax = b, x \geq 0\}. \tag{2.6}$$

To prove this, observe first that

$$c \cdot \theta = 0; \tag{2.7}$$

for otherwise, we could select a small value of λ so that

$$c \cdot (x_0 - \lambda\theta) < c \cdot x_0$$

($\lambda > 0$ if $c \cdot \theta > 0$, $\lambda < 0$ if $c \cdot \theta < 0$). This is a contradiction since $x_0 - \lambda\theta$ is feasible for small $|\lambda|$. Thus (2.7) holds and therefore

$$c \cdot (x_0^* - x_0) = -\lambda^* c \cdot \theta = 0.$$

This proves (2.6).

Thus x_0^* is optimal for (P), but has fewer non-zero components than x_0. And this is a contradiction: x_0 is a basic optimal solution.

□

REMARK. Our discussion of basic solutions leads to the very interesting realization that although linear programming problems are finite dimensional, with infinitely many feasible solutions, they are in effect finite optimization problems, with only finitely many basic solutions to consider and only finitely many optimal basic solutions.

The simplex algorithm, discussed next, builds upon this observation. □

2.2. Simplex algorithm

We describe next the simplex algorithm, developed by G. Dantzig. This algorithm is universally regarded as among the most important from the 20th century: see Cipra [**C**]. The actions of the simplex algorithm are somewhat analogous to elementary row operations in linear algebra.

The simplex algorithm comprises two procedures:

Phase I: Find a basic feasible solution of $Ax = b, x \geq 0$ (or show that none exists).

Phase II: Given a basic feasible solution, find a basic optimal solution (or show that none exists).

2.2.1. Nondegeneracy.

DEFINITION. The **nondegeneracy assumptions** are that

(i) $n > m$

(ii) the rows of A are linearly independent (and thus A has m columns which are independent)

(iii) b cannot be written as a linear combination of fewer than m columns of A.

REMARKS.

• Assumption (i) implies that there are more unknowns $(x_1, \ldots, x_n)$ than the m linear equality constraints in the linear system $Ax = b$.

• Assumption (ii) means

$$\text{rank}(A) = m = \dim(\text{column space}) = \dim(\text{row space}).$$

• Assumption (iii) says that if $Ax = b$, then x has at least m non-zero entries. □

REMARK. Under the nondegeneracy assumptions, *any basic, feasible solution of $Ax = b, x \geq 0$ has **precisely** m non-zero entries.* The next assertion shows that the converse is true as well. □

LEMMA 2.2.1 (On nondegeneracy). Assume the nondegeneracy conditions (i)-(iii) hold.

If $Ax = b, x \geq 0$ and x has precisely m non-zero entries, then x is a basic feasible solution.

Proof. Let $x = [0\, x_{j_1}\, 0 \ldots x_{j_m} \ldots 0]^T$, where $x_{j_1}, \ldots, x_{j_m} > 0$, and write $B = \{j_1, \ldots, j_m\}$. We will show that the columns $\{a^{j_1}, \ldots, a^{j_m}\}$ are independent.

We know

$$\sum_{j \in B} x_j a^j = b.$$

If the $\{a^{j_1}, \ldots, a^{j_m}\}$ were dependent, we could write some column as a linear combination of the others. That is, for some index j_k we have

$$a^{j_k} = \sum_{\substack{j \in B \\ j \neq j_k}} y_j a^j.$$

Then

$$b = x_{j_k} a^{j_k} + \sum_{\substack{j \in B \\ j \neq j_k}} x_j a^j = x_{j_k} \sum_{\substack{j \in B \\ j \neq j_k}} y_j a^j + \sum_{\substack{j \in B \\ j \neq j_k}} x_j a^j = \sum_{\substack{j \in B \\ j \neq j_k}} (x_{j_k} y_j + x_j) a^j.$$

Thus b is a linear combination of fewer than m columns of A, a contradiction to the nondegeneracy requirement (iii).

□

2.2.2. Phase II.

We discuss Phase II before Phase I (as the latter, somewhat surprisingly, will depend upon the former). The goal of Phase II is, given a basic feasible solution x, to find a basic optimal solution x_0, or show none exists. For this, we assume the nondegeneracy conditions (i), (ii), (iii).

So we are given

$$x = \begin{bmatrix} 0 \\ x_{j_1} \\ \vdots \\ x_{j_2} \\ \vdots \\ 0 \\ \vdots \\ x_{j_m} \\ 0 \end{bmatrix}$$

where $x_{j_1}, \ldots, x_{j_m} > 0$ are the m non-zero entries of x. We also have

$$Ax = b. \tag{2.8}$$

STEP 1: Use the dual problem to check for optimality.

We have a basic feasible solution x, and need to check if it is optimal or not.

DEFINITIONS.

(i) Write $B = \{j \mid x_j > 0\} = \{j_1, \ldots, j_m\}$. We call $\{a^j \mid j \in B\}$ the **basis** corresponding to x.

(ii) Define also the $m \times m$ matrix

$$M = \left[a^{j_1} \;\middle|\; a^{j_2} \;\middle|\; \ldots \;\middle|\; a^{j_m}\right]_{m \times m},$$

called the corresponding **basis matrix**.

(iii) If $c = \begin{bmatrix} c_1 & \ldots & c_n \end{bmatrix}^T$, define

$$\hat{c} = \begin{bmatrix} c_{j_1} \\ \vdots \\ c_{j_m} \end{bmatrix} \in \mathbb{R}^m.$$

The $m \times m$ matrix M is invertible, since its columns are independent. Thus there exists a unique $y \in \mathbb{R}^m$ solving

$$M^T y = \hat{c}. \tag{2.9}$$

Then

$$y = \left(M^T\right)^{-1} \hat{c} = \left(M^{-1}\right)^T \hat{c}.$$

Recall next that y is feasible for (D) **if**

$$A^T y \leq c; \tag{2.10}$$

this means

$$a^j \cdot y \leq c_j \qquad (j = 1, \ldots, n). \tag{2.11}$$

Note carefully: (2.10) may, or may not, be valid. But if so, we are done:

LEMMA 2.2.2. If (2.10) holds, then x is optimal for (P).

Proof. The equilibrium equations (E) say

$$a^j \cdot y = \sum_{i=1}^{m} y_i a_{ij} = c_j \quad \text{if } x_j > 0, \text{ that is, if } j \in B.$$

Now $M = \underbrace{\left[a^{j_1} \;\middle|\; \ldots \;\middle|\; a^{j_m}\right]}_{\text{columns}}$, and so

$$M^T = \left.\begin{bmatrix} (a^{j_1})^T \\ \vdots \\ (a^{j_m})^T \end{bmatrix}\right\} \text{rows}\,.$$

But (2.9) says $M^T y = \hat{c}$, which means that

$$a^j \cdot y = c_j \quad \text{for } j \in B.$$

These are precisely the equilibrium equations (E). So **if y is feasible for (D)**, it follows that x is optimal for (P), y is optimal for (D). □

Therefore we have two possibilities:

Case $\mathbf{A}_1$: ***y*** **defined by** (2.9) **satisfies** (2.10). Then STOP: $x_0 = x$ is optimal for (P).

Case $\mathbf{A}_2$: ***y*** **defined by** (2.9) **does not satisfy** (2.10)**.** GO TO STEP 2.

STEP 2: Use a "wrong way" inequality to improve *x*.

When (2.10) fails, there exists some index $s \in \{1, \ldots, n\} \setminus B$ such that

$$\underbrace{a^s \cdot y > c_s}_{\text{"wrong way" inequality}}. \tag{2.12}$$

The key idea of the simplex algorithm is to use this fact to change the basis $\{a^j \mid j \in B\}$*, thereby constructing a new basic feasible solution* x^* *with a lower cost* $c \cdot x^*$. To do this, we first find

$$t = \begin{bmatrix} t_{j_1} \\ \vdots \\ t_{j_m} \end{bmatrix} \in \mathbb{R}^m$$

so that

$$Mt = a^s. \tag{2.13}$$

Since M is invertible, (2.13) has a unique solution. Then

$$a^s = \sum_{j \in B} t_j a^j,$$

and consequently for all λ, we have

$$\lambda a^s + \sum_{j \in B} (x_j - \lambda t_j) a^j = b. \tag{2.14}$$

Now define

$$\hat{x} = \begin{bmatrix} 0 \\ x_{j_1} - \lambda t_{j_1} \\ \vdots \\ x_{j_2} - \lambda t_{j_2} \\ \vdots \\ \lambda \\ \vdots \\ x_{j_m} - \lambda t_{j_m} \\ 0 \end{bmatrix} \in \mathbb{R}^n.$$

Here λ is in the s-th slot. According to (2.14),

$$A\hat{x} = b$$

and

$$\hat{x} \geq 0 \quad \text{(for small } \lambda > 0\text{)}.$$

So $\hat{x}$ is feasible for (P) for small $\lambda > 0$ (but has $m+1$ non-zero entries, and consequently is not basic).

How does replacing x by $\hat{x}$ affect the cost? The old cost is

$$c \cdot x = \sum_{j \in B} x_j c_j$$

and the new cost is

$$c \cdot \hat{x} = \lambda c_s + \sum_{j \in B} (x_j - \lambda t_j) c_j.$$

The change in cost is therefore

$$c \cdot \hat{x} - c \cdot x = \lambda c_s - \sum_{j \in B} \lambda t_j c_j = \lambda (c_s - z_s) \tag{2.15}$$

for

$$z_s = \sum_{j \in B} t_j c_j = \hat{c} \cdot t.$$

We next calculate using (2.9) that

$$z_s = \hat{c} \cdot t = \hat{c} \cdot (M^{-1} a^s) = \left(M^{-1}\right)^T \hat{c} \cdot a^s = y \cdot a^s.$$

We see that therefore the "wrong way" inequality (2.12) is equivalent to

$$z_s > c_s. \tag{2.16}$$

It follows that

$$c \cdot \hat{x} < c \cdot x$$

for $\lambda > 0$. Consequently, *we lower the cost by shifting to $\hat{x}$ from x.*

STEP 3: Change the basis, lower the cost

There are two possibilities as to how much we can lower the cost by increasing λ:

Case B_1: $t_j \leq 0$ for all $j \in B$.

Then $\hat{x}_j = x_j - \lambda t_j \geq x_j > 0$ for $j \in B$ and so $\hat{x}$ is feasible for all $\lambda > 0$. Thus (2.15) says

$$c \cdot \hat{x} = c \cdot x + \lambda(\underbrace{c_s - z_s}_{<0}) \to -\infty$$

as $\lambda \to \infty$. So we have learned that

$$\inf\{c \cdot x \mid Ax = b, x \geq 0\} = -\infty$$

and therefore STOP: (P) has no solution.

Case B_2: $t_j > 0$ for at least one index $j \in B$.

We increase λ, starting at 0 and stopping when $\lambda = \lambda^* > 0$ and $\hat{x}_{j_k} = x_{j_k} - \lambda^* t_{j_k} = 0$ for some index $j_k \in B$. Define

$$x^* := \begin{bmatrix} 0 \\ x_{j_1} - \lambda^* t_{j_1} \\ \vdots \\ 0 \\ \vdots \\ \lambda^* \\ \vdots \\ x_{j_m} - \lambda^* t_{j_m} \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow j_k\text{-th slot} \\ \\ \leftarrow s\text{-th slot} \\ \\ \\ \end{matrix}$$

Then x^* has no more than m non-zero entries; and, since $Ax^* = b$, the non-degeneracy conditions say that x^* has precisely m non-zero entries. According to Lemma 2.2.1, x^* is therefore a basic feasible solution. Furthermore, $c \cdot x^* < c \cdot x$.

Now define the new basis

$$B^* = \underbrace{\{j_1, \ldots, j_m\}}_{B} \setminus \{j_k\} \cup \{s\},$$

by removing the index j_k and adding the index s. Then GO TO STEP 1, with x^* replacing x and B^* replacing B.

Each time we cycle through STEP 1 $\rightarrow$ STEP 3, the cost strictly decreases. Thus the same collection of basis vectors B will never repeat. Hence *the simplex algorithm terminates in a finite number of steps.* This can only occur when Case A_1 hapens (showing we have reached an optimal solution) or Case B_1 happens (showing that none exists). $\square$

As a bonus, we can extract from our reasoning above the following

THEOREM 2.2.1 (**Simplex algorithm finds optimal solutions**)**.** Assume the nondegeneracy conditions (i)-(iii) hold, and that there exist feasible x for (P), feasible y for (D).

Then the simplex algorithm terminates in finitely many steps, and produces a basic optimal x_0 for (P) and an basic optimal y_0 for (D).

Proof. If y is feasible for (D),

$$\inf\{c \cdot x \mid Ax = b, x \geq 0\} \geq b \cdot y > -\infty$$

and so Case B_1 cannot occur. Consequently the simplex algorithm terminates at an optimal x_0 for (P). Furthermore $y_0 = y$ (defined by (2.9)) is optimal for (D). $\square$

EXAMPLE. We use the simplex algorithm to solve

$$\text{(P)} \quad \begin{cases} \min\ 3x_1 + 3x_2 + 2x_3, \\ \text{subject to} \\ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} x = \begin{bmatrix} 3 \\ 9 \end{bmatrix}, \quad x \geq 0. \end{cases}$$

Here $c = [3\,3\,2]^T$, $b = [3\,9]^T$.

First feasible solution: For the given basic feasible solution $x = [1\,1\,0]^T$ we have $B = \{1, 2\}$. Then $j_1 = 1, j_2 = 2$,

$$\hat{c} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \quad M = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}, \quad M^{-1} = \frac{1}{3}\begin{bmatrix} -5 & 2 \\ 4 & -1 \end{bmatrix}.$$

(Recall that the inverse of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.)

We must check for the "wrong way" inequality (for $s = 3$). To do this, we first find t solving $Mt = a^3$. Then

$$t = M^{-1}a^3 = \frac{1}{3}\begin{bmatrix} -5 & 2 \\ 4 & -1 \end{bmatrix}\begin{bmatrix} 3 \\ 6 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

We compute next that

$$z_3 = \hat{c} \cdot t = 3,$$

and note that $c_3 = 2$. Since $z_3 > c_3$, the wrong way inequality does indeed hold for $s = 3$.

So we must bring a^3 into the basis and remove one of the current basis vectors (or else show the infimum of the cost is $-\infty$). We have

$$\hat{x} = \begin{bmatrix} x_1 - \lambda t_1 \\ x_2 - \lambda t_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} 1 + \lambda \\ 1 - 2\lambda \\ \lambda \end{bmatrix}.$$

We increase λ from 0 to λ^*, when one of the first two entries of $\hat{x}$ hits 0. This happens for $\lambda^* = \frac{1}{2}$.

Our new basic feasible solution is therefore $x^* = [\frac{3}{2}\, 0\, \frac{1}{2}]^T$. Let us check that

$$\text{old cost} = c \cdot x = 6, \ \text{new cost} = c \cdot x^* = \tfrac{11}{6};$$

and so the cost dropped by $\frac{1}{2}$.

Second feasible solution: We now repeat the process starting with the new feasible solution $x = [\frac{3}{2}\, 0\, \frac{1}{2}]^T$ and $B = \{1, 3\}$. Then

$$M = \begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix}, \quad M^{-1} = \frac{1}{6}\begin{bmatrix} -6 & 3 \\ 4 & -1 \end{bmatrix}, \quad \hat{c} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

Solving $Mt = a^2$ gives $t = M^{-1}a^2 = [\frac{1}{2}\, \frac{1}{2}]^T$. So $z_2 = \hat{c} \cdot t = \frac{5}{2}$ and $c_2 = 3$. It is now *not* true that $z_2 > c_2$, and thus the wrong way inequality is false.

Therefore $x_0 = [\frac{3}{2}\, 0\, \frac{1}{2}]^T$ is optimal for (P). Furthermore

$$y_0 = (M^{-1})^T \hat{c} = \frac{1}{6}\begin{bmatrix} -6 & 4 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} -10 \\ 7 \end{bmatrix}$$

is optimal for (D). We check this conclusion by noting $y_0 \cdot b = \frac{11}{2} = x_0 \cdot c$. □

EXAMPLE. Use the simplex algorithm to solve the problem

$$\text{(P)} \quad \begin{cases} \min\ x_1 + x_2 - 3x_3, \\ \text{subject to} \\ \begin{bmatrix} 1 & 2 & -3 \\ 4 & 5 & -9 \end{bmatrix} x = \begin{bmatrix} 4 \\ 13 \end{bmatrix}, \quad x \geq 0. \end{cases}$$

First feasible solution: We employ Phase II, starting with basic feasible solution $x = [2\,1\,0]^T$. Then the basis is $B = \{1, 2\}$, so that $j_1 = 1$, $j_2 = 2$, and

$$M = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}, \quad M^{-1} = \frac{1}{3}\begin{bmatrix} -5 & 2 \\ 4 & -1 \end{bmatrix}.$$

Next, we check for "wrong way" inequality (for all indices in $\{1,\dots,n\}\setminus B$). We just need to consider $s=3$. We solve $Mt=a^3$, by putting

$$t = M^{-1}a^3 = \frac{1}{3}\begin{bmatrix}-5 & 2\\ 4 & -1\end{bmatrix}\begin{bmatrix}-3\\ -9\end{bmatrix} = \begin{bmatrix}-1\\ -1\end{bmatrix}.$$

We have $\hat{c} = [1\,1]^T$ and therefore

$$z_3 = \hat{c}\cdot t = -2.$$

Since $c_3 = -3$, we see that $z_3 > c_3$: the wrong way inequality (2.16) does indeed hold for $s=3$.

So we must bring a^3 into the basis, and remove one of the current basis vectors (or else show the infimum of the cost is $-\infty$). We have

$$\hat{x} = \begin{bmatrix}x_1-\lambda t_1\\ x_2-\lambda t_2\\ \lambda\end{bmatrix} = \begin{bmatrix}2+\lambda\\ 1+\lambda\\ \lambda\end{bmatrix}$$

and $c\cdot\hat{x} = 3-\lambda\to-\infty$ as $\lambda\to\infty$. Hence

$$\inf\{c\cdot x \mid Ax=b, x\geq 0\} = -\infty,$$

and we see that therefore (P) does not have an optimal solution.

We also know from our general theory that for this example there must be no feasible y for (D). This is easy to check directly. □

2.2.3. Phase I.

We now explain how to carry Phase I of the simplex algorithm.

REMARK. But to do so, we need to modify the third nondegeneracy condition, to become

(iii)′ b cannot be written as a linear combination of fewer than m columns of $\tilde{A} = [A\ I]$,

where I is the $m\times m$ identity matrix. Observe that (iii)′ implies $b_i\neq 0$ for $i=1,\dots,m$. □

We assume for this section that the nondegeneracy conditions (i), (ii), (iii)′ hold. The goal of Phase I is to find $x\geq 0$ solving $Ax=b$, that is,

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \quad (i=1,\dots,m).$$

We may assume $b_i>0$ for $i=1,\dots,m$: if not, multiply the i-th equation by -1.

Consider now the modified problem

$$(\widetilde{\mathrm{P}}) \qquad \boxed{\begin{cases} \min\, z_1 + \ldots z_m, \text{ subject to} \\ \sum_{j=1}^{n} a_{ij}x_j + z_i = b_i \quad (i = 1, \ldots, m) \\ x \geq 0, z \geq 0. \end{cases}}$$

This has the form

$$\begin{cases} \min\, \tilde{c} \cdot \tilde{x}, \\ \text{subject to } \tilde{x} \geq 0, \tilde{A}\tilde{x} = \tilde{b} \end{cases}$$

for

$$\tilde{x} = \begin{bmatrix} x \\ z \end{bmatrix} \in \mathbb{R}^{n+m}, \quad \tilde{A} = \begin{bmatrix} A & I \end{bmatrix}_{m\times(n+m)}, \quad \tilde{b} = b, \quad \tilde{c} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}0\\\vdots\\0\end{matrix}}\right\} n \\ \left.\vphantom{\begin{matrix}1\\\vdots\\1\end{matrix}}\right\} m \end{matrix}.$$

Since each $b_i > 0$, a basic feasible solution of $(\widetilde{\mathrm{P}})$ is

$$\tilde{x} = [0 \cdots 0\, b_1 \cdots b_m]^T.$$

Now apply Phase II to $(\widetilde{\mathrm{P}})$: we either produce a basic optimal solution $\tilde{x}_0$ of $(\widetilde{\mathrm{P}})$ or learn that none exists. Since $\tilde{c} \cdot \tilde{x} \geq 0$ for all feasible $\tilde{x}$, the latter cannot occur, as we will later see from the Duality Theorem 2.3.3 in the next section. Hence Phase II provides us with a basic optimal $\tilde{x}_0$ for $(\widetilde{\mathrm{P}})$, and we write

$$\tilde{x}_0 = \begin{bmatrix} x \\ z \end{bmatrix}$$

with $x \in \mathbb{R}^n$, $z \in \mathbb{R}^m$.

There are now two possibilities to consider:

Case 1: $\sum_{i=1}^{m} z_i = 0$. Then $z_1 = \ldots = z_m = 0$, and therefore $\tilde{A}\tilde{x}_0 = \tilde{b}$ implies

$$Ax = b, x \geq 0.$$

So we have found a basic feasible solution x for (P).

Case 2: $\sum_{i=1}^{m} z_i > 0$. In this situation (P) does not have any feasible solutions x. This is so, since if $Ax = b, x \geq 0$, then

$$\tilde{x}_0 = \begin{bmatrix} x \\ 0 \end{bmatrix}$$

would be optimal for $(\widetilde{\mathrm{P}})$, giving the cost $\tilde{c} \cdot \tilde{x} = \sum_{i=1}^{m} z_i = 0$. □

REMARKS. Our discussion of Phases I, II of the simplex algorithm illustrates how some basic mathematical ideas (concerning the dual problem and basic solutions) can be cleverly fashioned into a powerful computational procedure.

Franklin [**F1**], Bertsekas–Tsitsiklis [**B-T**] and many other texts show how to implement the algorithm for a moderate number of variables on a spreadsheet, the various steps appearing as elementary row operations. Current linear programming applications can entail hundreds of thousands of variables and often require different, more modern algorithms such as interior point methods: see Boyd–Vandenberghe [**B-V**] or Bertsekas–Tsitsiklis [**B-T**]. □

2.3. Duality Theorem

Next we return to theory and provide an analysis of the solvability of linear programming problems in standard form:

$$(P^*)\begin{cases}\min\, c\cdot x,\\ \text{subject to}\\ Ax\geq b, x\geq 0.\end{cases}\qquad (D^*)\begin{cases}\max\, b\cdot y,\\ \text{subject to}\\ A^Ty\leq c, y\geq 0.\end{cases}$$

We no longer need the nondegeneracy conditions from the previous section, but we do require this important assertion:

THEOREM 2.3.1 (**Variant of Farkas alternative**)**.** Either

$$\text{(i)}^*\quad Ax\leq b,\ x\geq 0 \text{ has a solution, or}$$
$$\text{(ii)}^*\quad A^Ty\geq 0,\ y\cdot b<0,\ y\geq 0 \text{ has a solution,}$$

but not both.

We will prove this later (on page 89), but for now ask readers just to accept it. Following is a major application, the most important theoretical assertion in linear programming, first proved by von Neumann and Gale–Kuhn–Tucker.

THEOREM 2.3.2 (**Duality Theorem for standard form problems**)**.** Precisely one of the following occurs:

(I) Both (P^*) and (D^*) have feasible solutions. In this case, both (P^*) and (D^*) have optimal solutions and

$$\boxed{\min\{c\cdot x\mid Ax\geq b, x\geq 0\}=\max\{b\cdot y\mid A^Ty\leq c, y\geq 0\}.}$$

(II) There are feasible solutions for (D^*), but not for (P^*). Then

$$\boxed{\sup\{b \cdot y \mid A^T y \leq c, y \geq 0\} = \infty.}$$

(III) There are feasible solutions for (P^*), but not for (D^*). Then

$$\boxed{\inf\{c \cdot x \mid Ax \geq b, x \geq 0\} = -\infty.}$$

(IV) Neither (P^*) nor (D^*) has feasible solutions.

INTERPRETATION. Statement I says that *there is no duality gap for linear programming.* □

Proof. We introduce the compound matrix

$$\hat{A} = \begin{bmatrix} -A & 0 \\ 0 & A^T \\ c^T & -b^T \end{bmatrix}_{(m+n+1)\times(n+m)}$$

and define also

$$\hat{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \hat{b} = \begin{bmatrix} -b \\ c \\ 0 \end{bmatrix}, \quad \hat{y} = \begin{bmatrix} v \\ u \\ \lambda \end{bmatrix}.$$

Here $x, u \in \mathbb{R}^n, b, v \in \mathbb{R}^m, \lambda \in \mathbb{R}$. The Farkas alternative (Theorem 2.3.1) says that either

$$\text{(i)}^* \quad \hat{A}\hat{x} \leq \hat{b},\ \hat{x} \geq 0 \text{ has a solution}$$

or

$$\text{(ii)}^* \quad \hat{A}^T\hat{y} \geq 0,\ \hat{y} \cdot \hat{b} < 0,\ \hat{y} \geq 0 \text{ has a solution,}$$

but not both.

2. If (i)* holds, we can solve $\hat{A}\hat{x} \leq \hat{b}, \hat{x} \geq 0$ and therefore there exist $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ so that $x \geq 0, y \geq 0$ and

$$\begin{bmatrix} -A & 0 \\ 0 & A^T \\ c^T & -b^T \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq \begin{bmatrix} -b \\ c \\ 0 \end{bmatrix}.$$

Thus

$$\begin{cases} -Ax \leq -b,\ x \geq 0 \\ A^T y \leq c,\ y \geq 0 \\ c \cdot x - b \cdot y \leq 0. \end{cases}$$

Then $Ax \geq b$ and so also $b \cdot y \leq Ax \cdot y = x \cdot A^T y \leq x \cdot c$. Hence

$$\text{(i)}^* \quad \begin{cases} Ax \geq b, \ x \geq 0 \\ A^T y \leq c, \ y \geq 0 \\ c \cdot x = b \cdot y. \end{cases}$$

Consequently, x is feasible for (P^*) and y is feasible for (D^*). Furthermore, according to Theorem 2.1.1 last line of (i)* implies that x is optimal for (P^*) and y is optimal for (D^*). This gives statement I of the Duality Theorem.

3. If instead (ii)* holds, then we can solve $\hat{A}^T \hat{y} \geq 0, \hat{y} \cdot \hat{b} < 0, \hat{y} \geq 0$. This means there exist $v \in \mathbb{R}^m$, $u \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ so that $u, v, \lambda \geq 0$,

$$\begin{bmatrix} -A^T & 0 & c \\ 0 & A & -b \end{bmatrix} \begin{bmatrix} v \\ u \\ \lambda \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

and

$$\hat{y} \cdot \hat{b} = -b \cdot v + c \cdot u < 0.$$

This is all equivalent to

$$\begin{cases} A^T v \leq \lambda c, \ v, \lambda \geq 0 \\ Au \geq \lambda b, \ u \geq 0 \\ c \cdot u < b \cdot v. \end{cases}$$

We assert next that

$$\lambda = 0.$$

To see this, observe that

$$\lambda(v \cdot b) \leq v \cdot Au = A^T v \cdot u \leq \lambda(c \cdot u).$$

This contradicts $c \cdot u < b \cdot v$, unless $\lambda \leq 0$. Since $\lambda \geq 0$, we must have $\lambda = 0$.

Therefore we have

$$\text{(ii)}^* \quad \begin{cases} A^T v \leq 0, \ v \geq 0 \\ Au \geq 0, \ u \geq 0 \\ c \cdot u < b \cdot v. \end{cases}$$

The existence of u, v satisfying (ii)* will lead us to statements II-IV of the Duality Theorem. We need to investigate various possibilities as to the sign of a certain term.

4. **Case A**: $\boldsymbol{c \cdot u < 0}$.

In this situation, I claim that (D^*) has no feasible solutions. To see this, suppose $A^T y \leq c, \ y \geq 0$. Then (ii)* implies

$$0 \leq y \cdot Au = A^T y \cdot u \leq c \cdot u < 0,$$

and this is a contradiction.

If in addition (P^*) has no feasible solution, we have statement IV of the Duality Theorem. If on the other hand (P^*) does have a feasible solution solving $Ax \geq b, x \geq 0$, then for all $\mu \geq 0$, (ii)* gives

$$\begin{cases} A(x + \mu u) = Ax + \mu Au \geq b \\ x + \mu u \geq 0. \end{cases}$$

Thus $x + \mu u$ is also feasible for (P^*). But then

$$c \cdot (x + \mu u) = c \cdot x + \mu(c \cdot u) \to -\infty \qquad \text{as } \mu \to \infty,$$

since $c \cdot u < 0$. This gives statement III of the Duality Theorem.

5. **Case B**: $\boldsymbol{c \cdot u \geq 0}$

Then (ii)* implies

$$b \cdot v > c \cdot u \geq 0.$$

I claim that now (P^*) has no feasible solutions. To see this, assume $Ax \geq b$, $x \geq 0$. It would then follow from (ii)* that

$$0 \geq x \cdot A^T v = Ax \cdot v \geq b \cdot v > 0,$$

a contradiction. If also (D^*) has no feasible solution, we have statement IV.

If (D^*) does have a feasible solution y satisfying $A^T y \leq c, y \geq 0$, then for $\mu \geq 0$ the inequalities (ii)* give

$$\begin{cases} A^T(y + \mu v) = A^T y + \mu A^T v \leq c \\ y + \mu v \geq 0. \end{cases}$$

Hence $y + \mu v$ is also feasible for (D^*). And then

$$b \cdot (y + \mu v) = b \cdot y + \mu b \cdot v \to \infty \qquad \text{as } \mu \to \infty,$$

since $b \cdot v > 0$. This establishes the remaining statement II of the Duality Theorem. □

Let us now return to the canonical forms of our primal and dual problems:

$$(P) \begin{cases} \min c \cdot x, \\ \text{subject to} \\ Ax = b, x \geq 0. \end{cases} \qquad (D) \begin{cases} \max b \cdot y, \\ \text{subject to} \\ A^T y \leq c. \end{cases}$$

Since we can use slack and surplus variables to convert between these and the standard form problems (P^*), (D^*), we likewise have a duality assertion for the canonical problems:

THEOREM 2.3.3 (Duality Theorem for canonical form problems). Precisely one of the following occurs:

(I) Both (P) and (D) have feasible solutions. In this case, both (P) and (D) have optimal solutions and

$$\boxed{\min\{c \cdot x \mid Ax = b, x \geq 0\} = \max\{b \cdot y \mid A^T y \leq c\}.}$$

(II) There are feasible solutions for (D), but not for (P). Then

$$\boxed{\sup\{b \cdot y \mid A^T y \leq c\} = \infty.}$$

(III) There are feasible solutions for (P), but not for (D). Then

$$\boxed{\inf\{c \cdot x \mid Ax = b, x \geq 0\} = -\infty.}$$

(IV) Neither (P) nor (D) has feasible solutions.

2.4. Applications

We discuss in the subsequent sections several interesting applications and extensions of linear programming.

2.4.1. Multiobjective linear programming.

Multiobjective optimization problems ask us to simultaneously minimize several cost functions at once, a task that is generally undefined and therefore impossible. Nevertheless, we will see that linear programming can sometimes let us "fairly" combine the individual cost functions into a single cost.

To be specific, assume $c^1, \ldots, c^N \in \mathbb{R}^n$ and suppose we wish to find $x \in \mathbb{R}^n$ to somehow

$$\begin{cases} \text{minimize } c^1 \cdot x, c^2 \cdot x, \ldots, c^N \cdot x \\ \text{subject to } Ax = b, x \geq 0. \end{cases}$$

ECONOMIC INTERPRETATION. Such a problem arises when a central authority must allocate, subject to constraints, resources to N different groups of clients, each of which wants to maximize their own payoff functions. Cohon's book [**Co**] on multiobjective programming and planning discusses many interesting applications. □

NOTATION. We introduce the $N \times n$ matrix

$$C = \begin{bmatrix} (c^1)^T \\ \vdots \\ (c^N)^T \end{bmatrix},$$

where each entry is a row vector.

Then we can symbolically rewrite the above as the **multiobjective linear programming** problem

$$\text{(M)} \qquad \boxed{\begin{cases} \text{minimize } Cx, \text{ subject to} \\ Ax = b, x \geq 0. \end{cases}}$$

But note carefully that *this does not have really a meaning, as there is no obvious mathematical way to define the "minimum" of an* $\mathbb{R}^N$*-valued function* for $N > 1$. We therefore borrow a concept from mathematical economics:

DEFINITION. A feasible x_0 is called a **(Pareto) efficient solution** of (M) if there does **not** exist another feasible x such that

$$\boxed{Cx \leq Cx_0, \quad Cx \neq Cx_0.} \tag{2.17}$$

INTERPRETATION. The idea is that if we have an efficient solution x_0, then there is no way to shift to another feasible solution, so that (i) at least one of the costs drops and (ii) none of the others goes up. In this limited sense, x_0 provides a "fair" solution to (M). □

The key insight is that we can use linear programming to characterize efficient solutions:

THEOREM 2.4.1. (Weights and efficient solutions) A vector x_0 is an efficient solution of (M) if and only if there exists $w \in \mathbb{R}^N$, with

$$w > 0,$$

such that x_0 is an optimal solution of

$$(P) \qquad \boxed{\begin{cases} \text{minimize } c \cdot x, \text{ subject to} \\ Ax = b, x \geq 0 \end{cases}}$$

for

$$\boxed{c = C^T w}.$$

INTERPRETATION. Since

$$c \cdot x = (C^T w) \cdot x = w \cdot Cx = \sum_{i=1}^{N} w_i (c^i \cdot x),$$

we can think of the *positive* entries of $w = [w_1 \cdots w_N]^T$ as **weights** that allow us to combine the different costs $c^i \cdot x$ $(i = 1, \ldots, N)$ into a single cost $c \cdot x$. Each individual cost is "fairly weighted", at least in the sense that optimal solutions x_0 of (P) are Pareto efficient for (M).

Note also that the theorem implies there are typically many efficient solutions to (M), since for each vector of weights $w > 0$ we can look for a corresponding solution of (P). □

Proof. 1. Suppose x_0 is optimal for (P), where $c = C^T w$ for some vector of weights $w > 0$. If x_0 is *not* efficient for (M), there exists a feasible x with

$$Cx \leq Cx_0, \ Cx \neq Cx_0.$$

Therefore $w \cdot Cx < w \cdot Cx_0$, since $w > 0$, and consequently

$$c \cdot x = C^T w \cdot x = w \cdot Cx < w \cdot Cx_0 = C^T w \cdot x_0 = c \cdot x_0.$$

This is a contradiction, since x_0 is optimal for (P).

2. Now suppose conversely that x_0 is efficient solution of (M). We want to find weights $w_0 = [w_1 \cdots w_N]^T > 0$ such that x_0 is optimal for the linear programming problem (P) with $c = C^T w_0$.

We introduce the new linear programming problem of finding

$$\tilde{x} = \begin{bmatrix} x \\ z \end{bmatrix} \tag{2.18}$$

for $x \in \mathbb{R}^n, z \in \mathbb{R}^N$, to

$$\begin{cases} \text{minimize } -\sum_{i=1}^{N} z_i, \text{ subject to} \\ Ax = b, \ Cx + z = Cx_0, \ x \geq 0, z \geq 0. \end{cases}$$

This has the form

$$(\widetilde{\mathrm{P}}) \qquad \begin{cases} \min \tilde{c} \cdot \tilde{x}, \text{ subject to} \\ \tilde{A}\tilde{x} = \tilde{b}, \tilde{x} \geq 0, \end{cases}$$

where

$$\tilde{A} = \begin{bmatrix} A & 0 \\ C & I \end{bmatrix}, \quad \tilde{b} = \begin{bmatrix} b \\ Cx_0 \end{bmatrix}, \quad \tilde{c} = \begin{bmatrix} 0 \\ -e \end{bmatrix}$$

for $e = [1 \cdots 1]^T \in \mathbb{R}^N$.

One feasible solution is

$$\widetilde{x}_0 = \begin{bmatrix} x_0 \\ 0 \end{bmatrix}, \tag{2.19}$$

and I claim that $\widetilde{x}_0$ is in fact optimal for $(\widetilde{\mathrm{P}})$. To see this, suppose not. Then there would exist some other feasible $\tilde{x}$ of the form (2.18) with $\sum_{i=1}^{N} z_i > 0$. But then $Cx \leq Cx_0, Cx \neq Cx_0$ and so x_0 would not be efficient for (M).

3. The dual problem to $(\widetilde{\mathrm{P}})$ is

$$(\widetilde{\mathrm{D}}) \qquad \begin{cases} \text{maximize } \tilde{b} \cdot \tilde{y}, \text{ subject to} \\ \tilde{A}^T \tilde{y} \leq \tilde{c}, \end{cases}$$

where

$$\tilde{y} = \begin{bmatrix} y \\ -w \end{bmatrix}$$

for $y \in \mathbb{R}^m$, $w \in \mathbb{R}^N$. The constraints say that

$$\underbrace{\begin{bmatrix} A^T & C^T \\ 0 & I \end{bmatrix}}_{\tilde{A}^T} \underbrace{\begin{bmatrix} y \\ -w \end{bmatrix}}_{\tilde{y}} \leq \underbrace{\begin{bmatrix} 0 \\ -e \end{bmatrix}}_{\tilde{c}}$$

and therefore

$$\begin{cases} A^T y \leq C^T w \\ 0 < e \leq w. \end{cases}$$

Since $(\widetilde{\mathrm{P}})$ has an optimal solution given by (2.19), Statement I of the Duality Theorem 2.3.2 implies that $(\widetilde{\mathrm{D}})$ likewise has an optimal solution

$$\tilde{y}_0 = \begin{bmatrix} y_0 \\ -w_0 \end{bmatrix}$$

with $\tilde{y}_0 \cdot \tilde{b} = \tilde{x}_0 \cdot \tilde{c} = 0$. Therefore $b \cdot y_0 - w_0 \cdot Cx_0 = 0$. Consequently,

$$Ax_0 = b, \ x_0 \geq 0, \ A^T y_0 \leq c = C^T w_0$$

and

$$b \cdot y_0 = w_0 \cdot Cx_0 = (C^T w_0) \cdot x_0 = c \cdot x_0.$$

Therefore x_0 is optimal for (P) with $c = C^T w_0$, and the entries of $w_0 > 0$ give the desired weights. □

REMARK. The second part of the proof illustrates very well the mathematical principle that *the dual problem often contains valuable information*, in this case providing the weights corresponding to an efficient solution x_0. It is remarkable that our knowing just that no feasible x satisfies (2.17) is enough for this. □

2.4.2. Two-person, zero-sum matrix games.

In a two-person, zero-sum game, we have two participants: player I (who wants to maximize some payoff) and player II (who wants to minimize this payoff). Each player selects his/her strategy without knowing what the other will do.

For a matrix game, the payoff is determined by a given $m \times n$ **payoff matrix** A:

$$\text{player I selects row} \left\{ \underbrace{\begin{bmatrix} a_{11} & \dots & & \dots & a_{1n} \\ & \ddots & & & \\ \vdots & & a_{ij} & & \vdots \\ & & & \ddots & \\ a_{m1} & \dots & & \dots & a_{mn} \end{bmatrix}}_{\text{player II selects column}} \right.$$

Player I selects a row index i, $i \in \{1, \dots, m\}$, and player II selects a column index j, $j \in \{1, \dots, n\}$. The **payoff** to player I is a_{ij} and the **loss** to player II is a_{ij}.

What are optimal strategies for the players?

DEFINITION. The (k, l)-th entry a_{kl} of the matrix A is a **saddle point** if

$$\boxed{\max_{1 \le i \le m} a_{il} = a_{kl} = \min_{1 \le j \le n} a_{kj}.} \tag{2.20}$$

Equivalently, a_{kl} is a saddle point if

$$a_{il} \le a_{kl} \le a_{kj} \tag{2.21}$$

for all $i = 1, \dots, m, j = 1, \dots, n$.

INTERPRETATION. If there exists a saddle point a_{kl}, then

- player I should always select $i = k$;
- player II should always select $j = l$.

These are called **pure strategies**. These choices are optimal in the sense that I's payoff is then at least a_{kl}, regardless of what II does. Likewise, II's loss is at most a_{kl}, irrespective of what I does. We say that then the game has **value**

$$\omega = a_{kl}.$$

□

So, *if* the matrix A has a saddle point, our matrix game is pretty simple. But when do saddle points exist?

THEOREM 2.4.2 (**Minimax and saddle points**)**.** The matrix A has a saddle point if and only if the **minimax condition**

$$\boxed{\min_{1\le j\le n}\max_{1\le i\le m} a_{ij} = \max_{1\le i\le m}\min_{1\le j\le n} a_{ij}} \tag{2.22}$$

holds.

REMARK. When (2.22) is valid, we will as above write

$$\omega = \min_{1\le j\le n}\max_{1\le i\le m} a_{ij} = \max_{1\le i\le m}\min_{1\le j\le n} a_{ij}.$$

It follows that if A has more than one saddle point, they all give the same value ω. □

Proof. 1. First, we show that for all matrices A

$$\max_{1\le i\le m}\min_{1\le j\le n} a_{ij} \le \min_{1\le j\le n}\max_{1\le i\le m} a_{ij}. \tag{2.23}$$

To see this, observe for each i and j that $a_{ij} \le \max_r a_{rj}$. Thus

$$\min_j a_{ij} \le \min_j \max_r a_{rj}.$$

This holds for all i and so (2.23) follows.

2. Now suppose a_{kl} is a saddle point, so that (2.20) holds. Then

$$\min_j \max_i a_{ij} \le \max_i a_{il} = a_{kl}$$
$$\max_i \min_j a_{ij} \ge \min_j a_{kj} = a_{kl}.$$

Consequently

$$\min_j \max_i a_{ij} \le \max_i \min_j a_{ij}.$$

This is the reverse inequality of (2.23), and therefore (2.22) is valid.

3. Next, assume (2.22) holds, so that

$$\min_j \max_i a_{ij} = \max_i \min_j a_{ij}.$$

By selecting a value of $j = l$ that gives the min on the left and selecting a value of $i = k$ that gives the max on the right, we get

$$\max_i a_{il} = \min_j a_{kj}.$$

Then

$$a_{kl} \le \max_i a_{il} = \min_j a_{kj} \le a_{kl};$$

and hence we must have equality. Consequently

$$\max_i a_{il} = a_{kl} = \min_j a_{kj}.$$

□

EXAMPLE. For the simple matrix

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

we have

$$\max_i \min_j a_{ij} = -1, \quad \min_j \max_i a_{ij} = 1.$$

Therefore the minimax condition fails (2.22) and hence A does not have a saddle point. □

In order to find optimal strategies even when A does not have a saddle point, we need to expand our notion of what options are available to our players. This we do by introducing "mixed strategies".

DEFINITION. The collection of **mixed strategies** for Player I is

$$P = \left\{ p \in \mathbb{R}^m \mid p_i \geq 0 \ (i = 1, \ldots, m), \sum_{i=1}^{m} p_i = 1 \right\}$$

and the collection of **mixed strategies** for Player II is

$$Q = \left\{ q \in \mathbb{R}^n \mid q_j \geq 0 \ (j = 1, \ldots, n), \sum_{j=1}^{n} q_j = 1 \right\}.$$

DEFINITION. Suppose now I selects a vector $p \in P$ and II selects a vector $q \in Q$. The **payoff** to Player I is

$$p \cdot Aq = \sum_{i,j} p_i a_{ij} q_j,$$

which is also the **loss** to Player II.

PROBABILISTIC INTERPRETATION. As in our earlier Section 1.4.2, we can interpret P as the collection of probability distributions on the integers $\{1, \ldots, m\}$ and Q as the collection of probability distributions on the integers $\{1, \ldots, n\}$. Now imagine that our matrix game is played repeatedly. When Player I follows the mixed strategy p, each time he independently selects the row i with probability p_i; and when Player II follows the mixed strategy q, she independently selects the column j with probability q_j. The payoff $p \cdot Aq$ is then the expected outcome of these repeated games, averaged over many trials. □

DEFINITIONS. (i) We call (p_0, q_0) a (**mixed strategy**) **saddle point** if

$$\boxed{\max_{p\in P}\{p \cdot Aq_0\} = p_0 \cdot (Aq_0) = \min_{q\in Q}\{p_0 \cdot Aq\}.} \tag{S}$$

(ii) If (p_0, q_0) is a mixed strategy saddle point, we call

$$\boxed{\omega = p_0 \cdot Aq_0}$$

the **value** of the game.

NOTATION. We write

$$p_0 = [p_1^0, \ldots, p_m^0]^T, \quad q_0 = [q_1^0, \ldots, q_n^0]^T.$$

LEMMA 2.4.1. The point (p_0, q_0) is a saddle point if and only if there exists $\omega \in \mathbb{R}$ such that

$$\sum_{j=1}^{n} a_{ij} q_j^0 \le \omega \le \sum_{i=1}^{m} p_i^0 a_{ij} \qquad (i = 1, \ldots, m, j = 1, \ldots, n). \tag{2.24}$$

Proof. 1. If (S) holds, let us define $\omega = p_0 \cdot Aq_0$. Then by taking $q = [0 \cdots 1 \cdots 0]^T$, with the 1 in the j-th slot, we see that

$$\omega \le \sum_{i=1}^{m} p_i^0 a_{ij}.$$

We similarly deduce that

$$\omega \ge \sum_{j=1}^{n} a_{ij} q_j^0.$$

This gives (2.24).

2. Now assume (2.24) holds. We multiply the first inequality by $p_i \ge 0$ and sum, and multiply the second inequality by $q_j \ge 0$ and sum. This gives

$$p \cdot Aq_0 \le \omega \le p_0 \cdot Aq.$$

Hence

$$p_0 \cdot Aq_0 \le \max_{p\in P} p \cdot Aq_0 \le \omega \le \min_{q\in Q} p_0 \cdot Aq \le p_0 \cdot Aq_0.$$

The condition (S) follows. □

LEMMA 2.4.2. The value ω of a game, if it exists, is unique.

Proof. Suppose also $\omega' = p' \cdot Aq'$. Then according to (2.24),

$$\sum_j a_{ij} q_j' \le \omega' \le \sum_i p_i' a_{ij}.$$

Therefore

$$\omega \leq \underbrace{\sum_{i,j} p_i a_{ij} q'_j}_{p \cdot Aq'} \leq \omega' \leq \underbrace{\sum_{i,j} p'_i a_{ij} q_j}_{p' \cdot Aq} \leq \omega.$$

□

A matrix A may fail to have a saddle point, but there always exist mixed strategy saddle points:

THEOREM 2.4.3 (**Mixed strategy saddle points from linear programming**)**.** Let A be an $m \times n$ matrix.

Then there exists a (mixed strategy) saddle point (p_0, q_0).

REMARK. The value $\omega = p_0 \cdot Aq_0$ is unique, but mixed strategies (p_0, q_0) giving the value ω need not be unique. □

Proof. 1. We may assume $a_{ij} > 0$. If not, add a large constant C to each entry of A, getting a new matrix $\tilde{A}$. Compute $\tilde{\omega}$ for $\tilde{A}$ as below; then $\omega = \tilde{\omega} - C$ is the value for A.

2. We will find $p_0, q_0, \omega > 0$ solving (2.24). Define $u_i = \frac{p_i}{\omega}$ $(i = 1, \ldots, m)$ and $v_j = \frac{q_j}{\omega}$ $(j = 1, \ldots, n)$; so that

$$p = \omega u, \quad q = \omega v.$$

Then (2.24) holds if and only if

$$\begin{cases} \sum_{i=1}^m u_i a_{ij} \geq 1 \\ \sum_{j=1}^n a_{ij} v_j \leq 1 \\ \sum_i u_i = \sum_j v_j = \frac{1}{\omega} \\ u, v \geq 0. \end{cases}$$

Let $c = [1 \cdots 1]^T \in \mathbb{R}^m$, $b = [1 \cdots 1]^T \in \mathbb{R}^n$. We introduce the dual standard linear programming problems

$$\begin{aligned} &(\mathrm{P}^*) \qquad \min c \cdot u, \text{ subject to } A^T u \geq b, u \geq 0 \\ &(\mathrm{D}^*) \qquad \max b \cdot v, \text{ subject to } Av \leq c, v \geq 0. \end{aligned}$$

3. Note that $u = [M\, M \cdots M]^T$ is feasible for (P^*) if $M > 1$ is large enough (since all the entries of the matrix A are positive) and $v = [0\, 0 \cdots 0]^T$ is feasible for (D^*). We are thus in Statement I of the Duality Theorem. Hence there exist optimal u_0 for (P^*) and v_0 for (D^*) with $u_0 \neq 0$ and

$$u_0 \cdot c = v_0 \cdot b.$$

This says

$$\sum_i u_i^0 = \sum_j v_j^0.$$

Define

$$\omega = \frac{1}{\sum_i u_i^0} = \frac{1}{\sum_j v_j^0}, \quad p_0 = \omega u_0, \quad q_0 = \omega v_0.$$

Then $p_0 \in P, q_0 \in Q$ and p_0, q_0, ω solve (2.24). □

We could also invoke the general Minimax Theorem 5.5.1, which we will establish later, for a completely different proof. But the linear programming approach has the advantage that it also provides ways to compute mixed strategy saddle points:

EXAMPLE. Find optimal mixed strategies p_0, q_0 for the matrix

$$A = \begin{bmatrix} 5 & 4 & 2 \\ 2 & 1 & 6 \end{bmatrix}.$$

Here $m = 3, n = 2$. Set $b = [1\,1\,1]^T$, $c = [1\,1]^T$. Our linear programming problem is

$$(\mathrm{P}^*) \qquad \begin{cases} \min u_1 + u_2 \\ \text{subject to } u \geq 0, A^T u \geq b. \end{cases}$$

The constraints therefore say

$$\begin{bmatrix} 5 & 2 \\ 4 & 1 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 5u_1 + 2u_2 \\ 4u_1 + u_2 \\ 2u_1 + 6u_2 \end{bmatrix} \geq \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Let us guess that we have equalities for the second and third equation, so that $4u_1 + u_2 = 1$ and $2u_1 + 6u_2 = 1$. Then $5u_1 + 2u_2 \geq 1$. The solution is

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \frac{5}{22} \\ \frac{2}{22} \end{bmatrix}.$$

The corresponding dual problem is

$$(\mathrm{D}^*) \qquad \begin{cases} \max v_1 + v_2 + v_3 \\ \text{subject to } v \geq 0, Av \leq c. \end{cases}$$

Hence

$$\begin{bmatrix} 5 & 4 & 2 \\ 2 & 1 & 6 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \leq \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Since $5u_1 + 2u_2 > 1$, we must have $v_1 = 0$, according to the equilibrium equations. Hence

$$\begin{bmatrix} 4 & 2 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

and so

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{4}{22} \\ \frac{3}{22} \end{bmatrix}.$$

Thus $u_1 + u_2 = v_1 + v_2 + v_3 = \frac{7}{22}$, and then

$$\omega = \frac{1}{\sum_i u_i} = \frac{1}{\sum_j v_j} = \frac{22}{7},$$

$$p_0 = \omega u = \begin{bmatrix} \frac{5}{7} \\ \frac{2}{7} \end{bmatrix}, \quad q_0 = \omega v = \begin{bmatrix} 0 \\ \frac{4}{7} \\ \frac{3}{7} \end{bmatrix}.$$

□

2.4.3. Network flows.

Our next example illustrates that sometimes we can exploit particular features of a linear programming problem to get useful information directly, without employing the full power of our general theory.

DEFINITION. A **flow network** is a pair (N, k), where

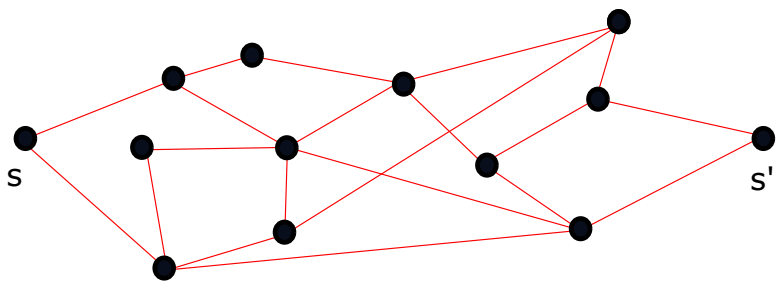
(i) $N = \{s, a, b, \ldots, s'\}$ is a finite set of points, called **nodes**. We call s the **source** and s' the **sink**.

(ii) $k : N \times N \to [0, \infty)$ is a function such that $k(x, x) = 0$ and

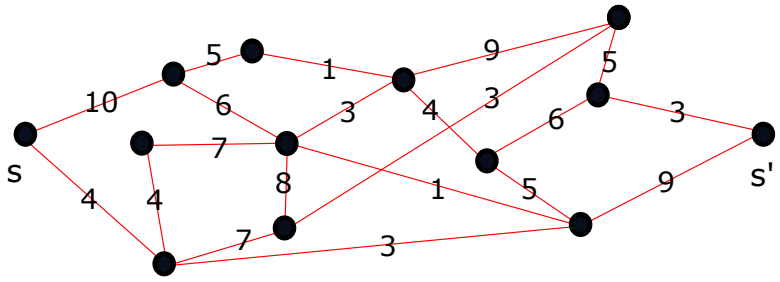
$$k(x, y) = k(y, x) \qquad (x, y \in N). \tag{2.25}$$

We interpret $k(x, y)$ as the **capacity** of the edge from x to y.

We draw each edge for which $k(x, y) > 0$.



A network of connected nodes



A network with flow capacities

REMARK. The symmetry condition (2.25) says that the capacity is the same regardless of the direction of the flow along an edge. We do not really need to require this, but it simplifies our illustrations. □

DEFINITION. A **flow** in (N, k) is a function $f : N \times N \to \mathbb{R}$ such that

(i) $f(x, y) = -f(y, x)$

(ii) $f(x, y) \leq k(x, y)$

(iii) $\sum_{y \in N} f(x, y) = 0$ if $x \neq s, s'$.

(iv) $f(s, x) \geq 0, f(x, s') \geq 0$ for all $x \in N$.

Note that (i) implies that $f(x, x) = 0$

NOTATION. If A, B are subsets of N, we write

$$k(A, B) = \sum_{x \in A} \sum_{y \in B} k(x, y),$$

$$f(A, B) = \sum_{x \in A} \sum_{y \in B} f(x, y).$$

□

DEFINITIONS.

(i) The **value** of a flow is

$$v(f) = f(s, N) = \sum_{x \in N} f(s, x).$$

(ii) A flow f_0 is called **maximal** if

$$v(f_0) \geq v(f)$$

for all flows f on our network.

We wish to study maximal flows and in particular to find some sort of mathematical characterization. Now this is in fact a linear programming problem, and we can write the foregoing explicitly as a canonical problem (P). And this would then lead us to some characterization of a maximal flow in terms of the minimum of an appropriate dual problem (D). See Bertsekas–Tsitsiklis [**B-T**] for more insight on this.

But it is more interesting to search directly for a dual problem, defined in terms of the network geometry:

DEFINITIONS.

(i) Let $C \subseteq N$ such that $s \in C$, $s' \notin C$, and let

$$C' = N \setminus C.$$

The pair (C, C') is called a **cut**.

(ii) The **capacity** of the cut is

$$k(C, C').$$

(iii) A cut C_0 is **minimal** if

$$k(C_0, C_0') \leq k(C, C')$$

for all cuts C.

LEMMA 2.4.3. (i) If f is any flow and (C, C') is any cut, we have

$$f(C, C') \leq k(C, C'). \tag{2.26}$$

and

$$v(f) = f(C, C'). \tag{2.27}$$

(ii) Therefore

$$\max_f v(f) \leq \min_{(C,C')} k(C, C'). \tag{2.28}$$

REMARK. The inequality (2.28) asserts a "weak duality" between flows and cuts, and is an analog of our general inequality (2.1) for linear programming. □

Proof. 1. We know $f(x, y) \leq k(x, y)$. Sum over $x \in C$ and $y \in C'$, to derive (2.26).

2. We also have $f(C, C) = \sum_{x \in C, y \in C} f(x, y) = 0$, since $f(x, y) = -f(y, x)$. So

$$\begin{aligned} f(C, C') &= f(C, C') + f(C, C) \\ &= f(C, N) \end{aligned}$$

$$= \sum_{x \in C} f(x, N) = f(s, N) = v(f),$$

since $f(x, N) = 0$ for all $x \neq s, s'$. This gives (2.27), which with (2.26) implies (2.28). □

By analogy with our basic linear programming theory, we expect to have a "strong duality" between maximal flows and minimal cuts. This is indeed so:

THEOREM 2.4.4 (Max Flow, Min Cut Theorem). The flow f_0 is maximal if and only if there exists a minimal cut (C_0, C_0') such that

$$v(f_0) = k(C_0, C_0'). \tag{2.29}$$

Therefore

$$\boxed{\max_f v(f) = \min_{(C,C')} k(C, C').} \tag{2.30}$$

Proof. 1. Suppose f_0 is a maximal flow. We will prove that there exists a minimal cut (C_0, C_0') with $v(f_0) = k(C_0, C_0')$.

Let us first introduce some terminology. We say an edge (x, y) is *unsaturated* if $f_0(x, y) < k(x, y)$. Likewise, a path $x_1, \ldots, x_k$ of nodes is *unsaturated* if each edge (x_{i-1}, x_i) along this path is unsaturated. Let us also define

$$C_0 = \{x \in N \mid \text{there exists an unsaturated path from } s \text{ to } x\}.$$

Thus $s \in C_0$, $s' \notin C_0$, since, if not, we could increase flow along an unsaturated path from s to s', contradicting that f_0 is a maximal flow. Therefore (C_0, C_0') is a cut.

2. We next claim that

$$f_0(x, x') = k(x, x') \quad \text{if } x \in C_0, x' \in C_0'. \tag{2.31}$$

To prove this, let us suppose instead that $f_0(x, x') < k(x, x')$. We could then add the edge (x, x') to unsaturated path from s to x, thereby finding an unsaturated path from s to x'. But this is impossible, since $x' \in C_0'$.

Now sum the equality (2.31) over $x \in C_0$, $x' \in C_0'$ and recall (2.27), to learn that

$$v(f_0) = f_0(C_0, C_0') = k(C_0, C_0').$$

Recall also from (2.26) that $v(f_0) = f_0(C, C') \leq k(C, C')$ for any cut C. It follows that

$$k(C_0, C_0') \leq k(C, C')$$

for any cut C; and consequently (C_0, C_0') is a minimal cut.

3. Conversely, suppose now that $f_0(C_0, C_0') = k(C_0, C_0')$. We will show f_0 is a maximal flow and (C_0, C_0') is a minimal cut.

Let f be any flow. Then

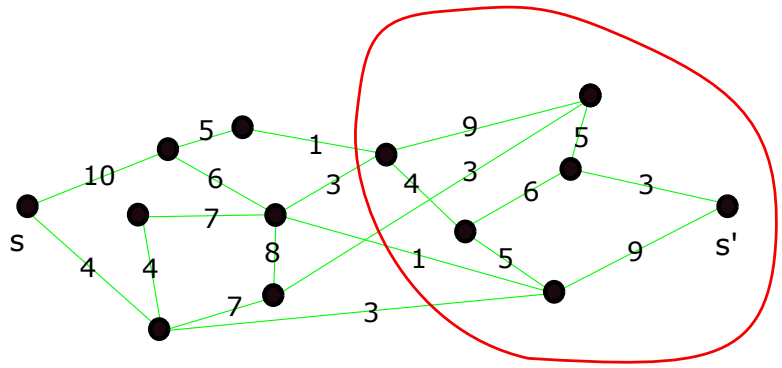
$$v(f) = f(C_0, C_0') \leq k(C_0, C_0') = f(C_0, C_0') = v(f_0);$$

it follows that f_0 is maximal. Finally, let C be any cut.Then

$$k(C, C') \geq f_0(C, C') = v(f_0) = k(C_0, C_0'),$$

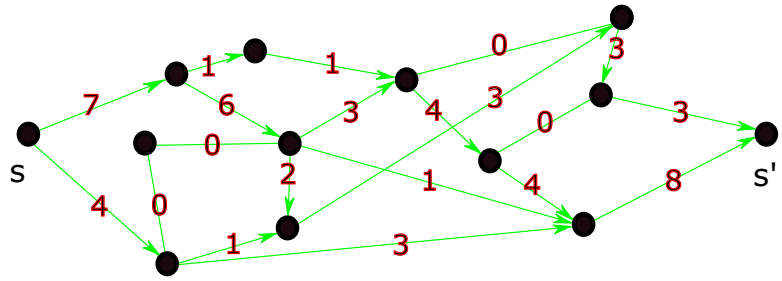
and so (C_0, C_0') is a minimal cut. □

EXAMPLE. For our previous example, we identify a minimal cut as drawn:



A minimal cut

The capacity of this cut, which is the total capacity of the 5 edges crossing the red curve, is $1+3+3+1+3 = 11$. Here is a flow that saturates the flow capacity across this cut, and is therefore maximal:

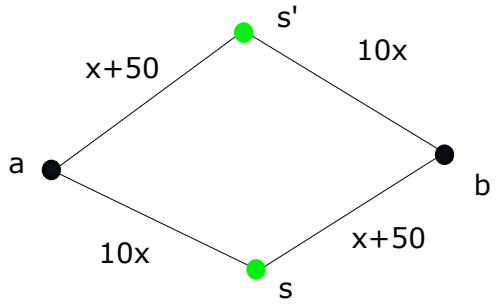


A maximal flow

□

EXAMPLE. (BRAESS' PARADOX) We discuss next an interesting, and somewhat related, example of a game theory problem for network flow.

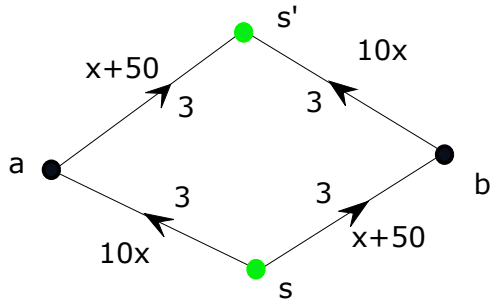
The picture illustrates a simple highway network, connecting the 2 cites s and s' by two routes, one through the village a and the other through the village b. If we let x denote the number of cars using a given road segment, the "cost" to each driver is the time required to drive that segment, which we assume is a function of the traffic density x, as marked in the picture.



Transit times on a road network

Driving a road network. Suppose now that 6 cars enter the road network at s and exit at s'. Each driver wants to minimize his/her total driving time. How should each driver go from s to s'?

Consider the possibility that 3 drivers go through village a, and 3 go through village b. Then the cost for every driver is $(x + 50) + 10x = 83$, since $x = 3$.



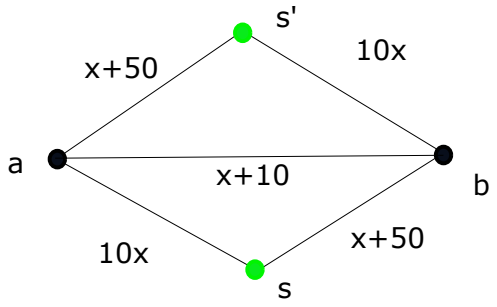
A stable traffic pattern

Is this allocation of cars on the road network stable? In other words, can any driver lower her/his cost by changing their route, assuming that the other drivers do not change their driving patterns?

To decide this, suppose one driver decides to change from going through village a and instead drives through village b. The cost to the 4 drivers who now go through village b is $(x + 50) + 10x = 94$ for $x = 4$. Since this is greater than the original cost, no driver has an incentive to change.

Consequently the original configuration of 3 cars going through each of the villages is stable.

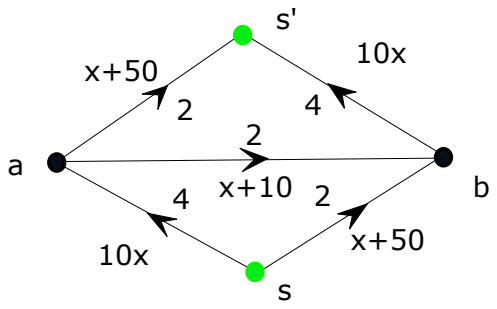
Building a new roadway. But suppose now that we build a new (one-way) road from a to b, along which the cost is $x + 10$. How does this new road change the incentives of our drivers?



A new road

Suppose a particular driver, who had been driving $[s, b, s']$, decides instead to take the new route $[s, a, b, s']$. Her driving time is now $(10 \times 4) + (1 + 10) + (10 \times 3) = 81$, which is less than her original travel time of 83. Consequently she has an inventive to change his route, and her change may in turn cause the others to change their routes.

What if drivers reach the following pattern?



A new driving pattern

We calculate that the cost for the route $[s, a, s']$ is now $(10\times4)+(2+50) = 92$; the cost for $[s, a, b, s']$ is $(10 \times 4) + (2 + 10) + (10 \times 4) = 92$; and the cost for $[s, b, s']$ is $(2 + 50) + (4 \times 10) = 92$. This new driving pattern is worse for everyone. In addition, *we have the paradoxical fact that this worse pattern is also stable*, meaning that no one can lower his/her cost if everyone else continues driving as before. To see this, we compute the costs if one driver alone changes from the pattern above:

• If a driver changes her route from $[s, a, b, s']$ to $[s, a, s']$, her cost changes from 92 to $(10 \times 4) + (3 + 50) = 93$.

• If instead a driver changes from $[s, b, s']$ to $[s, a, s']$, his cost changes from 92 to $(10 \times 5) + (3 + 50) = 103$.

• If a driver changes from $[s, b, s']$ to $[s, a, b, s']$, her cost changes from 92 to $(10 \times 5) + (3 + 10) + (10 \times 4) = 103$.

So no one has an incentive to change, and thus *building the new road has caused everyone's driving time to increase.* (See Körner [**Ko**] for more, and consult also Cohen–Horowitz [**C-H**] for further examples of related optimization paradoxes.)

□

2.4.4. Transportation problem.

We discuss next the famous linear programing transportation problem, also known as the discrete **Monge-Kantorovich** problem.

ECONOMIC INTERPRETATION. We have M factories, from which we will ship their output of some product to N different customers. Let $s_i \geq 0$ denote the supply available at a factory i $(i = 1, \dots, M)$ and $d_j \geq 0$ denote the demand of customer j $(j = 1, \dots, N)$. Our problem is to decide how much of the output of each factory to send to each customer, so as to minimize the total shipping costs.

We write $x_{ij} \geq 0$ for the amount shipped from factory i to customer j and $c_{ij} \geq 0$ for the unit cost of transportation from i to j. The constraint $\sum_{j=1}^{N} x_{ij} = s_i$ means that everything from factory i is shipped away. Likewise, the requirement $\sum_{i=1}^{M} x_{ij} = d_j$ means that each customer receives precisely enough to meet his/her requirements. □

In mathematical terms, we wish to find x_{ij}^0 to

$$\text{(MK)} \qquad \begin{cases} \text{minimize} \quad \sum_{i,j} c_{ij} x_{ij} \\ \text{subject to} \quad x_{ij} \geq 0, \\ \sum_{j=1}^{N} x_{ij} = s_i \qquad (i = 1, \dots, M) \\ \sum_{i=1}^{M} x_{ij} = d_j \qquad (j = 1, \dots, N). \end{cases}$$

We assume

$$\sum_{i=1}^{M} s_i = \sum_{j=1}^{N} d_j, \tag{2.32}$$

so that the total supply equals the total demand.

NOTATION. We will usually regard X as an $M \times N$ matrix, and write

$$X = \begin{bmatrix} x_{11} & \dots & x_{1N} \\ \vdots & \ddots & \vdots \\ x_{M1} & \dots & x_{MN} \end{bmatrix}, \quad C = \begin{bmatrix} c_{11} & \dots & c_{1N} \\ \vdots & \ddots & \vdots \\ c_{M1} & \dots & c_{MN} \end{bmatrix},$$

$$C \cdot X = \sum_{i,j} c_{ij} x_{ij}, \quad s = \begin{bmatrix} s_1 \\ \vdots \\ s_M \end{bmatrix}, \quad d = \begin{bmatrix} d_1 \\ \vdots \\ d_N \end{bmatrix}.$$

□

THEOREM 2.4.5.

(i) There exists a feasible solution for (MK) if and only if (2.32) holds.

(ii) If (2.32) holds, there exists an optimal solution X_0 of (MK).

Proof. 1. To prove (i), note that if X is feasible, then

$$\sum_j x_{ij} = s_i, \ \sum_i x_{ij} = d_j,$$

and therefore $\sum_i s_i = \sum_{ij} x_{ij} = \sum_{j,i} x_{ij} = \sum_j d_j$. Conversely, if (2.32) holds, we can design a feasible X as follows:

- Factory 1 sends as much as possible to customer 1, then to customer 2 (if $s_1 > d_1$), then to customer 3 (if $s_1 > d_1 + d_2$), etc.
- Then factory 2 sends as much as possible to the last customer not fully satisfied with the shipment from factory 1, etc.
- Continue, until factory M ships the last of its supply to customer N.

Since the total supply equals the total demand, this produces a feasible shipping plan X.

2. According to (i), (MK) has feasible solutions if the supply and demand balance condition (2.32) is valid. Furthermore, since $c_{ij} \geq 0$,

$$\inf \{C \cdot X \mid X \text{ is feasible}\} \geq 0.$$

Thus Statement I of the Duality Theorem applies and so (MK) has an optimal solution. □

We wish now to study the structure of an optimal transportation plan X_0. As usual, the dual problem and equilibrium equations contain interesting information:

THEOREM 2.4.6. (i) The dual problem for (MK) is to find $u \in \mathbb{R}^M, v \in \mathbb{R}^N$ to

$$\text{(MK}^*\text{)} \qquad \begin{cases} \text{maximize } v \cdot d - u \cdot s, \text{ subject to} \\ v_j - u_i \le c_{ij} \qquad (i = 1, \dots, M; j = 1, \dots, N). \end{cases}$$

(ii) The equilibrium conditions are

$$\text{(E)} \qquad v_j - u_i = c_{ij} \quad \text{if } x_{ij} > 0.$$

Proof. 1. The derivation of the dual problem is a bit tricky, since the variable X has two indices. Remember that we usually think of X as an $M \times N$ matrix, but sometimes as a vector in $\mathbb{R}^{MN}$. Then the equality constraints in (MK) become

$$AX = \begin{bmatrix} \sum_j x_{1j} \\ \vdots \\ \sum_j x_{Mj} \\ \sum_i x_{i1} \\ \vdots \\ \sum_i x_{iN} \end{bmatrix} = \begin{bmatrix} s_1 \\ \vdots \\ s_M \\ d_1 \\ \vdots \\ d_N \end{bmatrix} = \begin{bmatrix} s \\ d \end{bmatrix} = b.$$

Now let $y = \begin{bmatrix} -u \\ v \end{bmatrix} \in \mathbb{R}^{M+N}$; so that

$$y \cdot b = v \cdot d - u \cdot s.$$

For any $M \times N$ matrix Z, we have

$$\begin{aligned} A^T y \cdot Z = y \cdot AZ &= -\sum_{i=1}^{M} u_i \left(\sum_{j=1}^{N} z_{ij} \right) + \sum_{j=1}^{N} v_j \left(\sum_{i=1}^{M} z_{ij} \right) \\ &= \sum_{ij} (v_j - u_i) z_{ij}. \end{aligned} \tag{2.33}$$

Select $k \in \{1, \dots, M\}, l \in \{1, \dots, N\}$, and put

$$z_{ij} = \begin{cases} 1 & \text{if } i = k, j = l \\ 0 & \text{otherwise.} \end{cases}$$

Then (2.33) tells us

$$(A^T y)_{kl} = v_l - u_k.$$

Hence the duality condition $A^T y \le C$ is equivalent to

$$v_l - u_k \le c_{kl} \qquad (k = 1, \dots, M; l = 1, \dots, N),$$

and the equilibrium equations (E) now follow from our general theory. $\square$

ECONOMIC INTERPRETATION. Let u_0, v_0 be optimal for the dual problem, and observe that, if necessary, we can add constants to ensure

$$u_i^0 \geq 0, v_j^0 \geq 0.$$

Let us interpret

$$u_i^0 = \text{“cost” to produce one unit at factory } i,$$

$$v_j^0 = \text{“payment” when we sell one unit to customer } j$$

The equilibrium conditions (E) say $v_j^0 = u_i^0 + c_{ij}$ if $x_{ij}^0 > 0$. This means that we ship from i to j only if “payment = cost of production + cost of shipping”.

So if $v_j^0 < u_i^0 + c_{ij}$, we should not ship anything from i to j. The point is that we can get useful information about the MN entries x_{ij}^0 for an optimal shipping plan, in terms of the far fewer $M + N$ numbers u_i^0, v_j^0. □

EXAMPLE. Suppose our supplies, demands and transport costs are

$$s = \begin{bmatrix} 8 \\ 3 \end{bmatrix}, \; d = \begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix}, \; C = \begin{bmatrix} 9 & 7 & 1 \\ 5 & 4 & 0 \end{bmatrix}.$$

Let us show that

$$X_0 = \begin{bmatrix} 1 & 2 & 5 \\ 3 & 0 & 0 \end{bmatrix}$$

is optimal. To confirm this, we note that X_0 has the required row and column sums, and compute that $C \cdot X_0 = \sum_{i,j} c_{ij} x_{ij}^0 = 43$.

We turn next to the dual problem, and recall that the constraints and equilibrium equations read

$$v_j - u_i \leq c_{ij}, \; v_j - u_i = c_{ij} \quad \text{if } x_{ij} > 0.$$

Let us show that

$$u_0 = \begin{bmatrix} -1 \\ 3 \end{bmatrix} \qquad v_0 = \begin{bmatrix} 8 \\ 6 \\ 0 \end{bmatrix}.$$

are optimal for the dual problem. We have

$$\begin{bmatrix} v_1^0 - u_1^0 & v_2^0 - u_1^0 & v_3^0 - u_1^0 \\ v_1^0 - u_2^0 & v_2^0 - u_2^0 & v_3^0 - u_2^0 \end{bmatrix} = \begin{bmatrix} 9 & 7 & 1 \\ 5 & 3 & -3 \end{bmatrix} \leq \begin{bmatrix} 9 & 7 & 1 \\ 5 & 4 & 0 \end{bmatrix}$$

and thus u_0, v_0 are feasible. Since

$$\sum_{j=1}^{3} v_j^0 d_j - \sum_{i=1}^{2} u_i^0 s_i = 44 - 1 = 43 = C \cdot X_0,$$

u_0, v_0 are optimal for the dual, and our transportation plan X_0 is indeed optimal for (MK). □

We will see in the next section that it is no accident that when the supplies and demands are all integers, so also are the entries of X_0.

2.4.5. Integer-valued solutions.

We will show in this section that certain linear programming problems, including the transportation problem, have integer solutions. (We will freely use various definitions and theorems concerning convex sets that will be developed in the next chapter. Readers may consequently wish to skip this section until later.)

We continue our discussion of the transportation problem and now make the additional assumption that

(2.34) $\{s_1, \dots, s_M\}$ and $\{d_1, \dots, d_N\}$ are nonnegative integers.

We will prove that there then exists an optimal solution X_0 such that each x_{ij}^0 is a nonnegative integer.

DEFINITION. The set of **transportation matrices** is

$$\mathcal{X} = \left\{ X = \begin{bmatrix} x_{11} & \dots & x_{1N} \\ \vdots & \ddots & \vdots \\ x_{M1} & \dots & x_{MN} \end{bmatrix} \mid x_{ij} \geq 0, \sum_{j=1}^{N} x_{ij} = s_i, \sum_{i=1}^{M} x_{ij} = d_j \right\}.$$

It is easy to check that $\mathcal{X}$ is a convex subset of the space of all real $M \times N$ matrices.

LEMMA 2.4.4. If the matrix X is an extreme point of $\mathcal{X}$, then

(2.35) X has at most $M + N - 1$ non-zero entries.

Proof. 1. Suppose instead that X has at least $M + N$ positive entries. Select any subset of precisely $M + N$ positive entries:

(2.36) $$\{x_{ij} \mid (i, j) \in K\}$$

Here the index set $K \subset \{1, \dots, M\} \times \{1, \dots, N\}$ has cardinality $|K| = M+N$ and

(2.37) $$x_{ij} > 0 \quad \text{if } (i, j) \in K.$$

2. Consider now the system of linear equations

$$AZ = \begin{bmatrix} \sum_j z_{1j} \\ \vdots \\ \sum_j z_{Mj} \\ \sum_i z_{i1} \\ \vdots \\ \sum_i z_{iN} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0, \tag{2.38}$$

for the unknowns

$$Z = \begin{bmatrix} z_{11} & \cdots & z_{1N} \\ \vdots & \ddots & \vdots \\ z_{M1} & \cdots & z_{MN} \end{bmatrix}.$$

Note that the range of the matrix A is not all of $\mathbb{R}^{M+N}$, since if

$$AX = \begin{bmatrix} \sum_j x_{1j} \\ \vdots \\ \sum_j x_{Mj} \\ \sum_i x_{i1} \\ \vdots \\ \sum_i x_{iN} \end{bmatrix} = \begin{bmatrix} e_1 \\ \vdots \\ e_M \\ f_1 \\ \vdots \\ f_N \end{bmatrix} = \begin{bmatrix} e \\ f \end{bmatrix},$$

then necessarily

$$\sum_{i=1}^{M} e_i = \sum_{j=1}^{N} f_j.$$

3. Our goal is to find a non-zero solution Z of (2.38) such that

$$z_{ij} = 0 \quad \text{if } (i,j) \notin K, \tag{2.39}$$

where K is the index set from (2.36). Now if we enforce (2.39), then (2.38) becomes a system of $M+N$ equations for the $M+N$ unknowns

$$\hat{Z} = \{z_{ij} \mid (i,j) \in K\}.$$

We can write this as $\hat{A}\hat{Z} = 0$ for an appropriate $(M+N) \times (M+N)$ matrix $\hat{A}$. But the matrix $\hat{A}$ does not have full rank $M+N$, since, as noted above, the matrix A has rank strictly less than $M+N$. Hence there exists $Z \neq 0$ solving (2.38) and (2.39)

4. Now set

$$X^+ = X + \varepsilon Z, \; X^- = X - \varepsilon Z.$$

For $\varepsilon > 0$ small enough, we have $X^{\pm} \in \mathcal{X}$ and also

$$X = \frac{1}{2}X^{+} + \frac{1}{2}X^{-}.$$

But this is impossible, since X is an extreme point of $\mathcal{X}$. This contradiction shows that (2.35) must be valid. □

This proof, and the next, are based upon ideas in Karlin [**K**, Section 5.8].

THEOREM 2.4.7. If X is an extreme point of $\mathcal{X}$, then

(2.40) each x_{ij} is a nonnegative integer.

Proof. 1. Assume first that $M \leq N$. If each of the N columns of X contained two or more nonzero entries, the total number k of nonzero entries would satisfy

$$k \geq 2N \geq M + N.$$

But this is impossible, since (2.35) says $k \leq M + N - 1$. Therefore X has at least one column with at most one nonzero entry. Since $\sum_{j=1}^{N} x_{ij} = s_i$ is a nonnegative integer, this column has at most one positive entry, which is an integer.

Next, remove this column from X, creating a new $M \times (N-1)$ matrix $\bar{X}$, which is an extreme point of the convex set

$$\left\{ \bar{X} \mid \bar{x}_{ij} \geq 0, \sum_{j=1}^{N-1} \bar{x}_{ij} = \bar{s}_i, \sum_{i=1}^{M} \bar{x}_{ij} = \bar{d}_j \right\}$$

for appropriate integers $\bar{s}_i \geq 0, \bar{d}_i \geq 0$.

2. If $N \leq M$, a similar argument shows that X has a least one row, with at most one positive entry, which is an integer. We remove that row, and obtain a new matrix $\bar{X}$ that is an extreme point of an appropriate convex set of $(M-1) \times N$ matrices, with nonnegative integer row and column sums.

3. We now repeat the foregoing argument, removing at each step either a column or row containing at most one nonzero entry, which is an integer. The process stops after $M + N$ steps. □

THEOREM 2.4.8. Under the assumption (2.34) that the individual supplies and demands are positive integers, the transportation problem (MK) has a solution X_0 such that

(2.41) x_{ij}^{0} is a nonnegative integer

for $i = 1, \ldots, M$ and $j = 1, \ldots, N$.

Proof. 1. Let $\mathcal{Y} \subseteq \mathcal{X}$ denote the collection of transportation matrices with integer entries. Then clearly

$$\min_{X \in \mathcal{X}} C \cdot X \leq \min_{Y \in \mathcal{Y}} C \cdot Y.$$

Suppose next that $X_0 \in \mathcal{X}$ is optimal. Then according to Theorem 2.4.7, we can write X_0 as a weighted sum of matrices with integer entries:

$$X_0 = \sum_{k=1}^{m} \theta_k Y^k, \ \sum_{k=1}^{m} \theta_k = 1, \ 0 < \theta_k \leq 1, \ Y^k \in \mathcal{Y}. \tag{2.42}$$

Then

$$\begin{aligned} \min_{X \in \mathcal{X}} C \cdot X &= C \cdot X_0 \\ &= \sum_{k=1}^{m} \theta_k \left(C \cdot Y^k \right) \\ &\geq \sum_{k=1}^{m} \theta_k \left(\min_{Y \in \mathcal{Y}} C \cdot Y \right) \\ &= \min_{Y \in \mathcal{Y}} C \cdot Y. \end{aligned}$$

This shows each matrix Y^k in (2.42) must be optimal for the transportation problem, since otherwise we would have a strict inequality in the last calculation above. □

Chapter 3

CONVEXITY

This chapter introduces concept of **convexity**, in both its geometric and functional guises. It is impossible to overstate the importance of convexity in pure and applied mathematics.

3.1. Convex geometry

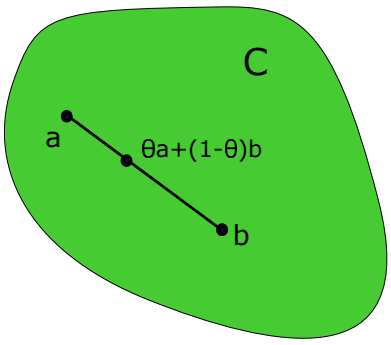
We have so far introduced calculus tools for optimization in Chapter 1, and linear algebra tools in Chapter 2. We now add in geometric insights.

3.1.1. Convex sets.

DEFINITION. A set $C \subseteq \mathbb{R}^n$ is **convex** if for all $a, b \in C$ and $0 \leq \theta \leq 1$, we have

$$\theta a + (1 - \theta)b \in C.$$

GEOMETRIC INTERPRETATION. So if C is convex, then for all choices of $a, b \in C$, the line segment connecting a and b also lies in C. □



DEFINITIONS. Let $\{a^1, \ldots, a^p\} \subset \mathbb{R}^n$.

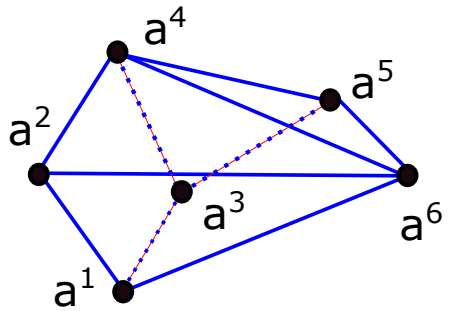
(i) If $\theta_k \geq 0$ and $\sum_{k=1}^p \theta_k = 1$, we call

$$\sum_{k=1}^{p} \theta_k a_k \tag{3.1}$$

a **convex combination** of $a^1, \ldots, a^p$.

(ii) The **convex polytope** generated by $a^1, \ldots, a^p$ is

$$\langle a^1, \ldots, a^p \rangle = \left\{ \sum_{k=1}^{p} \theta_k a_k \mid \theta_k \geq 0, \sum_{k=1}^{p} \theta_k = 1 \right\}.$$



A convex polytope

REMARK. It is straightforward to prove by induction that if C is convex and $\{a^1, \ldots, a^p\} \subset C$, then each convex combination of the form (3.1) also belongs to C.

It is also easy to see that the convex polytope $\langle a^1, \ldots, a^p \rangle$ is convex and is the smallest convex set containing the points $\{a^1, \ldots, a^p\}$. □

Next we use ideas from linear programming to prove a famous theorem about convex polytopes.

THEOREM 3.1.1 (Caratheodory's Theorem). Let b belong to the convex polytope $\langle a^1, \ldots, a^p \rangle \subset \mathbb{R}^n$.

Then we can write

$$b = \sum_{k=1}^{n+1} \theta_k a^{j_k} \tag{3.2}$$

where $1 \leq j_1 < \cdots < j_{n+1} \leq p$, $\theta_k \geq 0$, $\sum_{k=1}^{n+1} \theta_k = 1$.

REMARK. So even if p is very large, *we can write any point in the polytope* $\langle a^1, \ldots, a^p \rangle$ *as a convex combination of at most* $n+1$ *of the* $\{a^1, \ldots, a^p\}$. □

Proof. Since $b \in \langle a^1, \ldots, a^p \rangle$, there exists a solution $x \in \mathbb{R}^p$ of

$$Ax = \begin{bmatrix} b \\ 1 \end{bmatrix}, \quad x \geq 0 \tag{3.3}$$

for the $(n+1) \times p$ matrix

$$A = \left[\begin{array}{c|c|c|c} a^1 & a^2 & \cdots & a^p \\ 1 & 1 & & 1 \end{array}\right].$$

But according to Theorem 2.1.5, there exists also a *basic* solution $x^* \in \mathbb{R}^p$ of

$$Ax^* = b,\ x^* \geq 0.$$

This means that x^* has at most $m+1$ non-zero entries $\{x_{j_1}, \ldots, x_{j_m}\}$, corresponding to independent columns of A. Then $m \leq n+1$ and we can relabel $\theta_k = x_{j_k}$ for $k = 1, \ldots, n$ to obtain (3.2). □

REMARK (**Extreme points**). Let C be a convex set. We say that $e \in C$ is an **extreme point** of C if there do **not** exist $x, y \in C$ so that

$$e = \theta x + (1-\theta)y, \quad \text{with } 0 < \theta < 1. \tag{3.4}$$

In other words, $e \in C$ is an extreme point if it cannot be written as a nontrivial convex combination of two other points in C.

An important theorem, the proof of which we omit, states that if $C \subset \mathbb{R}^n$ is closed, bounded and convex, then

$$C = \left\{ \sum_{k=1}^{n+1} \theta_k e^k \mid 0 \leq \theta_k \leq 1, \sum_{k=1}^{n+1} \theta_k = 1, e^k \in E \right\},$$

where E denotes the extreme points of C. In other words, *each point in a convex set is a convex combination of its extreme points.* □

3.1.2. Separating hyperplanes.

We discuss now the geometry of convex sets and of hyperplanes. The reader should first review as necessary the notion of a *closed* set, discussed in Appendix D.

LEMMA 3.1.1. Let C be a non-empty, closed, convex subset of $\mathbb{R}^n$ and suppose $0 \notin C$.

(i) Then there exists a unique point $x_0 \in C$ such that

$$|x_0| = \min\{|x| \mid x \in C\} > 0. \tag{3.5}$$

(ii) Furthermore,

$$0 \leq x_0 \cdot (x - x_0) \quad \text{for all } x \in C. \tag{3.6}$$

REMARK. We call (3.6) a **variational inequality**, which characterizes our solution x_0 of the minimization problem (3.5). We will learn more about variational inequalities in Chapter 5. □

Proof. 1. Let

$$\delta = \inf\{|x| \mid x \in C\} \geq 0.$$

Select $\{x^k\}_{k=1}^{\infty} \subset C$ with $\delta = \lim_{k\to\infty} |x^k|$. According to the Bolzano-Weierstrass Theorem, there is a convergent subsequence

$$\lim_{j\to\infty} x^{k_j} = x_0.$$

Consequently, $|x_0| = \lim_{j\to\infty} |x^{k_j}| = \delta$. Since C is closed, $x_0 \in C$; and since $0 \notin C, \delta > 0$.

2. We claim next that x_0 is the unique point in C with $|x_0| = \delta$. To see this, suppose $x_1 \in C$ also satisfies $|x_1| = \delta$. Then $(x_1 + x_0)/2 \in C$ and therefore

$$\left|\frac{x_1 + x_0}{2}\right| \geq \delta.$$

But

$$\underbrace{|x_1 - x_0|^2}_{\geq 0} + \underbrace{|x_1 + x_0|^2}_{\geq 4\delta^2} = \underbrace{2(|x_1|^2 + |x_0|^2)}_{4\delta^2},$$

and thus $x_1 = x_0$.

3. To prove (3.6), let x be any point in C. Then also $(1-\theta)x_0 + \theta x \in C$ if $0 < \theta \leq 1$. Therefore

$$\begin{aligned} |x_0|^2 &\leq |(1-\theta)x_0 + \theta x|^2 = |x_0 + \theta(x - x_0)|^2 \\ &= |x_0|^2 + 2\theta x_0 \cdot (x - x_0) + \theta^2 |x - x_0|^2. \end{aligned}$$

Consequently,

$$0 \leq 2x_0 \cdot (x - x_0) + \theta |x - x_0|^2.$$

Send $\theta \to 0$, to derive the variational inequality (3.6).

□

DEFINITION. Let $a \in \mathbb{R}^n$, $b \in \mathbb{R}$. An expression of the form

$$\boxed{a \cdot x + b = 0}$$

determines a **hyperplane** in $\mathbb{R}^n$.

More precisely, the hyperplane comprises all the points $x \in \mathbb{R}^n$ that satisfy the equation $a \cdot x + b = 0$. It is an $(n-1)$-dimensional affine subspace and passes through the origin if and only if $b = 0$.

DEFINITION. Let S_1, S_2 be two subsets of $\mathbb{R}^n$.

(i) We say the hyperplane $a \cdot x + b$ **separates** S_1 and S_2 if

$$a \cdot x + b \geq 0 \quad \text{for all } x \in S_1,$$
$$a \cdot x + b \leq 0 \quad \text{for all } x \in S_2.$$

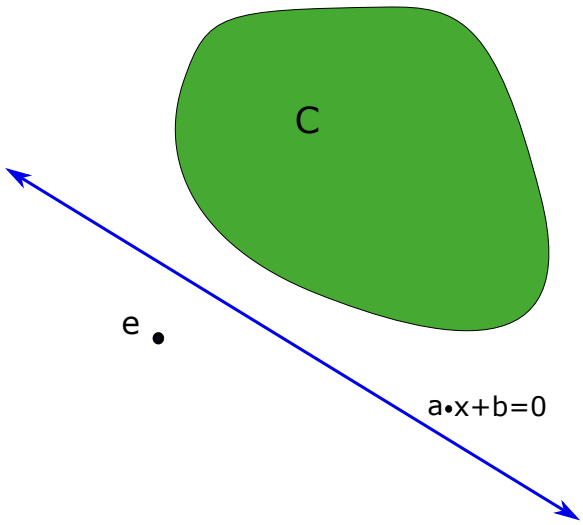
(ii) We say that $a \cdot x + b$ **strictly separates** S_1 and S_2 if

$$a \cdot x + b > 0 \quad \text{for all } x \in S_1,$$
$$a \cdot x + b < 0 \quad \text{for all } x \in S_2.$$

THEOREM 3.1.2 (Separating Hyperplane Theorem). Let $C \subset \mathbb{R}^n$ be convex, closed and non-empty, and suppose $e \notin C$.

Then there exists a hyperplane $a \cdot x + b$ that strictly separates C and e.

REMARK. It is important for subsequent applications that we do not require that C be bounded. □



A point separated from a convex set

Proof. Upon shifting the coordinates if necessary, we may assume $e = 0$. According to Lemma 3.1.1, there exists $x_0 \in C$ such that

$$0 < \delta = |x_0| = \min\{|x| \mid x \in C\}.$$

Let $m = \frac{1}{2}x_0$ and $a = x_0/\delta$, so that $|a| = 1$. Consider the hyperplane $a \cdot (x - m) = 0$; that is,

$$a \cdot x + b = 0$$

where $b = -a \cdot m$.

2. For $e = 0$, we have

$$a \cdot e + b = b = -a \cdot m = -\frac{x_0}{\delta} \cdot \frac{x_0}{2} = -\frac{\delta}{2} < 0 \tag{3.7}$$

3. Next, assume $x \in C$. The variational inequality (3.6) says $0 \leq x_0 \cdot (x - x_0)$, and therefore

$$0 \leq (\delta a) \cdot (x - 2m).$$

We cancel $\delta > 0$, to learn that

$$0 \leq a \cdot (x - m) - a \cdot m = a \cdot (x - m) - \frac{\delta}{2}.$$

Hence for all $x \in C$,

$$a \cdot x + b = a \cdot (x - m) \geq \frac{\delta}{2} > 0. \tag{3.8}$$ □

3.1.3. Dual convex sets.

As a first application of separating hyperplanes, we discuss next a geometric form of convex duality.

DEFINITION. Let $C \subset \mathbb{R}^n$ be closed and convex, with $0 \in C$. Its **polar dual** is the set

$$\boxed{C^0 = \{y \in \mathbb{R}^n \mid x \cdot y \leq 1 \text{ for all } x \in C\}.}$$

THEOREM 3.1.3 (Dual convex sets).

(i) C^0 is closed, convex, $0 \in C^0$.

(ii) We have the duality assertion

$$\boxed{(C^0)^0 = C.}$$

Proof. Statement (i) is easy. To prove (ii), note that

$$(C^0)^0 = \left\{z \in \mathbb{R}^n \mid y \cdot z \leq 1 \text{ for all } y \in C^0\right\}.$$

Let $x \in C$. Then $y \cdot x \leq 1$ for all $y \in C^0$ and thus $x \in (C^0)^0$. Consequently $C \subseteq (C^0)^0$.

If $z \in (C^0)^0 \setminus C$, then since C is closed, the Supporting Hyperplane Theorem says there exist $a \in \mathbb{R}^n, b \in \mathbb{R}$ such that

(a) $a \cdot z + b < 0$,

(b) $a \cdot x + b > 0$ for all $x \in C$.

Since $0 \in C$, (b) implies $b > 0$. So if we write

$$y = -\frac{a}{b},$$

(b) says also that $y \cdot x < 1$ for all $x \in C$. Hence $y \in C^0$. Since $z \in (C^0)^0$, it follows that $y \cdot z \leq 1$; and therefore

$$a \cdot z + b \geq 0.$$

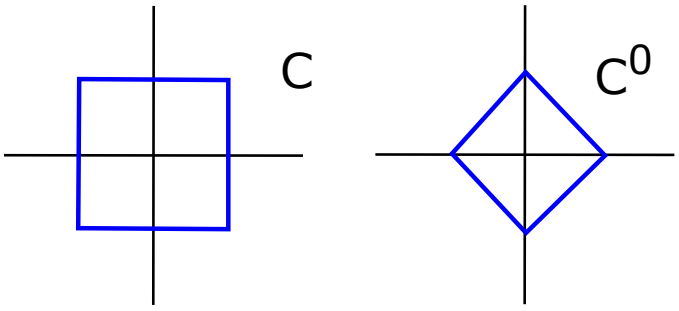
But this contradicts (a). Thus $(C^0)^0 \setminus C$ is empty and hence $(C^0)^0 = C$. □

REMARK. The proof illustrates an interesting mathematical fact, that *separating hyperplanes imply convex duality.* We will see more of this later. □

EXAMPLE. Let $C = \left\{x \in \mathbb{R}^2 \mid |x_1| \leq \frac{R}{2}, |x_2| \leq \frac{R}{2}\right\}$ be the square with center 0 and sides of length R parallel to the coordinate axes. Its polar is

$$\begin{aligned} C^0 &= \left\{y \in \mathbb{R}^2 \mid x \cdot y \leq 1 \text{ for all } x \in C\right\} \\ &= \left\{y \in \mathbb{R}^2 \mid |y_1| + |y_2| \leq \tfrac{2}{R}\right\}. \end{aligned}$$

This is a square with sides of length $R^0 = \frac{2\sqrt{2}}{R}$, rotated by $\frac{\pi}{4}$. Observe that then $(C^0)^0 = C$ is the original square, with side length $\frac{2\sqrt{2}}{R^0} = R$.



Look online for more exciting pictures of convex polyhedra and their duals. □

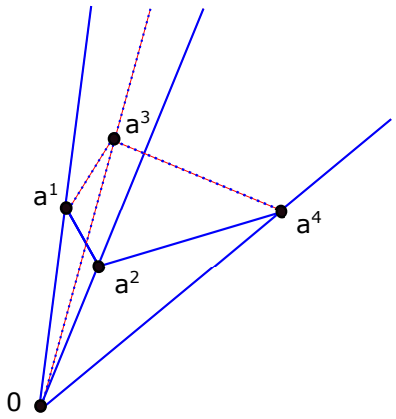
3.1.4. Farkas alternative.

Our next goal is the Farkas alternative, a statement about solving vector inequalities. This turns out to have a surprising geometric interpretation involving separating hyperplanes for certain convex cones.

DEFINITION. Let $\{a^1, \ldots, a^n\} \subset \mathbb{R}^m$. The set

$$C = \left\{ \sum_{i=1}^{n} x_i a^i \mid x_i \geq 0 \ (i = 1, \ldots, n) \right\}$$

is called the **finite cone** generated by $\{a^1, \ldots, a^n\}$.



A finite cone

REMARK. Observe that $b \in C$ precisely when we can solve $Ax = b, x \geq 0$, when

$$A = [a^1 \mid \ldots \mid a^n]$$

is the $m \times n$ matrix whose columns are $\{a^1, \ldots, a^n\}$. This is because $Ax = \sum_{i=1}^{n} x_i a^i$. □

DEFINITION. If $\{a^1, \ldots, a_k\}$ are independent, we call the finite cone they generate a **basic cone**.

LEMMA 3.1.2. Suppose $\{a^1, \ldots, a^n\}$ generate the finite cone C. Let $C_1, \ldots, C_q$ be the basic cones generated by all linearly independent subsets of $\{a^1, \ldots, a^n\}$. Then

$$C = \bigcup_{i=1}^{q} C_i. \tag{3.9}$$

Proof. Obviously $C_i \subseteq C$ $(i = 1, \ldots, q)$ and so $\bigcup_{i=1}^{q} C_i \subseteq C$.

Now select $b \in C = \{Ax \mid x \geq 0\}$. There then exists a solution of

$$Ax = b, \ x \geq 0;$$

and according to our earlier Theorem 2.1.5, there in fact exists a *basic* solution:

$$Ax^* = b,\ x^* \geq 0.$$

This means that the columns $\{a^{j_1}, \ldots, a^{j_m}\}$ of A corresponding to the nonzero entries of x^* are independent. So b belongs to the basic cone generated by $\{a^{j_1}, \ldots, a^{j_m}\}$ and thus $b \in \bigcup_{i=1}^q C_i$. This is true for all $b \in C$; consequently, $C \subseteq \bigcup_{i=1}^q C_i$. □

THEOREM 3.1.4. Let C be the finite cone generated by $\{a^1, \ldots, a^n\} \subset \mathbb{R}^m$. Then C is convex and closed.

Proof. 1. Let $b^1, b^2 \in C$, $0 \leq \theta \leq 1$. Then there exist x^1, x^2 such that $b^1 = Ax^1$, $x^1 \geq 0$ and $b^2 = Ax^2$, $x^2 \geq 0$. Therefore

$$(1-\theta)b^1 + \theta b^2 = A\left((1-\theta)x^1 + \theta x^2\right)$$

for $x = (1-\theta)x^1 + \theta x^2 \geq 0$. Thus $(1-\theta)b^1 + \theta b^2 \in C$, and so C is convex.

2. Let C_i be a basic cone, generated by an independent set

$$\{a^{j_1}, \ldots, a^{j_l}\} \subseteq \{a^1, \ldots, a^n\}.$$

Assume $\{b^k\}_{k=1}^\infty \subset C_i$, with $\lim_{k\to\infty} b^k = b^0$. I claim

$$b^0 \in C_i, \tag{3.10}$$

and this will show that C_i is closed.

First, let us write $B = \{j_1, \ldots, j_l\}$. Since the vectors $\{a^j \mid j \in B\}$ are independent, if $u = [u_{j_1} \ \ldots \ u_{j_l}]^T \in \mathbb{R}^l$ and $\sum_{j\in B} u_j a^j = 0$, it follows that $u = 0$. Therefore for all $u \in \mathbb{R}^l$ with $|u| = 1$, $\sum_{j\in B} u_j a^j \neq 0$. Hence the Extreme Value Theorem implies that there exists $\varepsilon > 0$ such that

$$\min\left\{ \left| \sum_{j\in B} u_j a^j \right| \mid |u| = 1 \right\} = \varepsilon > 0.$$

Thus if $v \in \mathbb{R}^l$,

$$\left| \sum_{j\in B} v_j a^j \right| \geq \varepsilon |v|. \tag{3.11}$$

We turn now to the proof of (3.10). Observe that we can write $b^k = Ax^k$, where $x^k \geq 0$, $x^k = [0\, x^k_{j_1}\, 0 \ \ldots \ 0\, x^k_{j_l}\, 0]^T$. Then

$$b^k = \sum_{j\in B} x^k_j a^j,$$

and therefore (3.11) implies

$$|x^k| \le \frac{1}{\varepsilon}|b^k| \qquad (k = 1, \dots). \tag{3.12}$$

The sequence $\{x^k\}_{k=1}^{\infty}$ is therefore bounded, and so we can apply the Bolzano-Weierstrass Theorem to extract a convergence subsequence:

$$\lim_{j\to\infty} x^{k_j} = x^0.$$

Then $x^0 \ge 0$, and $Ax^0 = \lim_{j\to\infty} Ax^{k_j} = \lim_{j\to\infty} b^{k_j} = b^0$. Furthermore, $x_j^0 = 0$ except possibly for the indices $j \in B$. Hence $b \in C_i$.

3. So each basic cone C_i is closed. As we show in Appendix D, the finite union of closed sets is closed, and hence $C = \bigcup_{i=1}^{q} C_i$ is closed. □

REMARK. This proof that a finite cone is closed is surprisingly tricky, and the key is the estimate (3.12) within a basic cone. □

In view of the previous theorem, we can apply the Separating Hyperplane Theorem to a finite cone. This has the following major payoff:

THEOREM 3.1.5 (Farkas alternative). Let A be an $m \times n$ matrix, $b \in \mathbb{R}^m$. Then **either**

(i)	$Ax = b,\ x \ge 0$ has a solution $x \in \mathbb{R}^n$, **or**
(ii)	$A^T y \ge 0,\ y \cdot b < 0$ has a solution $y \in \mathbb{R}^m$,

but not both.

Proof. 1. Assume x solves (i), y solves (ii). Then

$$0 \le x \cdot (A^T y) = Ax \cdot y = b \cdot y < 0,$$

which is a contradiction. So (i) and (ii) cannot both be true.

2. Suppose (i) fails. We will show that then (ii) holds. Now the failure of (i) means

$$b \notin C = \{Ax \mid x \ge 0\}.$$

We know that C is closed and convex. Hence the Separating Hyperplane Theorem asserts that there exist $a \in \mathbb{R}^m, c \in \mathbb{R}$ such that

$$a \cdot z + c > 0 \qquad (z \in C) \tag{3.13}$$

and

$$a \cdot b + c < 0. \tag{3.14}$$

Let $x \geq 0, \mu \geq 0$. Set $z = \mu Ax = A(\mu x) \in C$. According to (3.13),

$$a \cdot (\mu Ax) + c > 0.$$

Dividing by $\mu > 0$ and letting $\mu \to \infty$, we see that

$$a \cdot Ax \geq 0.$$

So

$$(A^T a) \cdot x \geq 0$$

for all $x \geq 0$. Thus $A^T a \geq 0$.

Let $y = a$; then $A^T y \geq 0$. Put $z = 0$ in (3.13), to deduce that $c > 0$. Then (3.14) says

$$b \cdot y = a \cdot b < -c < 0.$$

So $b \cdot y < 0$. □

3.1.5. Applications.

This section explores several interesting uses of the Farkas alternative.

1. Linear programming

In our proof of the Duality Theorem for linear programming in Chapter 2, we invoked without proof Theorem 2.3.1, a variant of the Farkas alternative stating that either

$$\text{(i)}^* \quad Ax \leq b, x \geq 0 \quad \text{has a solution } x \in \mathbb{R}^n, \text{ or}$$
$$\text{(ii)}^* \quad A^T y \geq 0, y \cdot b < 0, y \geq 0 \quad \text{has a solution } y \in \mathbb{R}^m,$$

but not both.

Proof. Note $Ax \leq b$ if and only if $Ax + z = b$ for some $z \geq 0$. So (i)* says

$$\text{(i)} \qquad \underbrace{\begin{bmatrix} A & I \end{bmatrix}}_{\tilde{A}} \begin{bmatrix} x \\ z \end{bmatrix} = b \ \text{ has a solution } \tilde{x} = \begin{bmatrix} x \\ z \end{bmatrix} \geq 0.$$

The Farkas alternative of (i) is

$$\text{(ii)} \qquad \tilde{A}^T y = \begin{bmatrix} A^T y \\ y \end{bmatrix} \geq 0, \ y \cdot b < 0 \text{ has a solution.}$$

This is (ii)*. □

With this, we have finally completed the full proof of the Duality Theorem for linear programming.

REMARK. Recall our dual canonical linear programming problems

$$\text{(P)} \begin{cases} \min c \cdot x, \\ \text{subject to} \\ Ax = b, x \geq 0. \end{cases} \qquad \text{(D)} \begin{cases} \max b \cdot y, \\ \text{subject to} \\ A^T y \leq c. \end{cases}$$

Case II of the Duality Theorem 2.3.3 tells us that if (P) has no feasible solutions, but (D) does, then $\sup\{b \cdot y \mid A^T y \leq c\} = +\infty$.

It is informative to see that the Farkas alternative provides a quick proof of this. Let y be feasible for (D), so that $A^T y \leq c$. If there are no feasible solutions x for (P), then Farkas (i) fails. Thus Farkas (ii) holds: there exists $z \in \mathbb{R}^m$ with

$$A^T z \geq 0, \ z \cdot b < 0.$$

But then $y - \mu z$ is feasible for (P) if $\mu > 0$, since

$$A^T(y - \mu z) = A^T y - \mu A^T z \leq A^T y \leq c.$$

Also, $b \cdot (y - \mu z) = b \cdot y - \mu(\underbrace{b \cdot z}_{<0}) \to \infty$ as $\mu \to \infty$. □

2. The Fredholm alternative of linear algebra

Next we utilize the Farkas Alternative for an unusual proof of the basic duality assertion for linear mappings. That we can do this is not surprising, since linear algebra is, strictly speaking, a subarea of linear programming.

NOTATION. Let A be an $m \times n$ matrix. Its **null space** is

$$N(A) = \{x \mid Ax = 0\} \subseteq \mathbb{R}^n$$

and its **range** is

$$R(A) = \{Ax \mid x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m.$$

□

DEFINITION. If S is a subspace of $\mathbb{R}^n$, its **dual subspace** is

$$S^\perp = \{x \in \mathbb{R}^n \mid x \cdot s = 0 \text{ for all } s \in S\}.$$

THEOREM 3.1.6. (**Duality for linear mappings**)

(3.15) $$Ax = b \text{ has a solution } x$$

if and only if

(3.16) $$b \in N(A^T)^\perp.$$

GEOMETRIC INTERPRETATION. This assertion and the corresponding statement for A^T mean geometrically that

$$\boxed{R(A) = N(A^T)^\perp, \quad R(A^T) = N(A)^\perp.}$$

□

Proof. $Ax = b$ has a solution x if and only if $A(u-v) = b, u \geq 0, v \geq 0$ has a solution. Then

$$\underbrace{\begin{bmatrix} A & -A \end{bmatrix}}_{\tilde{A}} \begin{bmatrix} u \\ v \end{bmatrix} = b.$$

This is Farkas (i). Farkas (ii) says

$$\tilde{A}^T y \geq 0, y \cdot b < 0 \quad \text{has a solution.}$$

That is,

$$\begin{bmatrix} A^T \\ -A^T \end{bmatrix} y \geq \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Then $A^T y \geq 0$ and $-A^T y \geq 0$, and so $A^T y = 0$. Consequently (ii) says

$$A^T y = 0, y \cdot b < 0 \text{ has a solution.}$$

So **either** (i) $Ax = b$ has a solution x, **or** (ii) $A^T y = 0, y \cdot b < 0$ has a solution y. Observe that (ii) is false if and only if $b \in N(A^T)^\perp$. □

3. Equilibria for Markov chains

A **Markov matrix** has the form

$$P = \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix},$$

where $p_{ij} \geq 0$ and

$$\sum_{i=1}^{n} p_{ij} = 1 \quad (j = 1, \ldots, n). \tag{3.17}$$

PROBABILISTIC INTERPRETATION. Consider a particle that jumps at each time $k = 1, 2, \ldots$ from one "box" or "state" to another. Assume there are n such states and that p_{ij} is the probability of a jump from state j to state i. The condition (3.17) means the sum of the probabilities over all possible jumps is 1. We assume these probabilities are the same at each time step, and that the jumps are independent.

Now suppose at some time k, x_j is the probability that particle is in state j for $j = 1, \ldots, n$. We write

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \text{ state vector at time } k$$

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \text{ state vector at time } k+1$$

Then $y_i = \sum_{j=1}^{n} p_{ij} x_j$; that is, $y = Px$. □

DEFINITION. A **steady state** is a vector $x \in \mathbb{R}^n$ such that

$$\boxed{x = Px,\ x \geq 0,\ \sum_{i=1}^{n} x_i = 1.}$$

THEOREM 3.1.7. If P is a Markov matrix, there exists a steady state.

REMARK. In particular, P has the real eigenvalue 1 and there is a corresponding eigenvector x *with nonnegative entries.* □

Proof. We write

$$A = \begin{bmatrix} P - I \\ e^T \end{bmatrix}_{(n+1)\times n}, \qquad e = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^n$$

$$I = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}_{n\times n}, \quad b = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^{n+1}.$$

We first observe that P has a steady state if and only if

(i) $Ax = b, x \geq 0$ has a solution.

To see this, note that $Ax = b, x \geq 0$ holds if and only if $Px = x, x \geq 0, e \cdot x = 1$, since $e \cdot x = \sum_{i=1}^{n} x_i$.) We will show (i) is true, by showing Farkas (ii) fails:

(ii) $A^T y \geq 0, y \cdot b < 0$ has a solution.

Here

$$A^T = \left[P^T - I \mid e\right]_{n\times(n+1)}, \quad y = \begin{bmatrix} z_1 \\ \vdots \\ z_n \\ -\mu \end{bmatrix}, \quad z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}.$$

We have

$$\text{(a) } 0 \le A^T y = \begin{bmatrix} P^T - I \mid e \end{bmatrix} \begin{bmatrix} z \\ -\mu \end{bmatrix} = P^T z - z - \mu e$$

$$\text{(b) } 0 > y \cdot b = [z \ -\mu]^T \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\mu.$$

Now (b) says $\mu > 0$, and (a) says

$$P^T z - z \ge \mu e.$$

That is,

$$\sum_{i=1}^{n} z_i p_{ij} \ge \mu + z_j \quad (j = 1, \dots, n).$$

Assume $z_m = \max_{1 \le i \le n} z_i$, and let $j = m$ above:

$$\mu + z_m \le \sum_{i=1}^{n} z_i p_{im} \le z_m \sum_{i=1}^{n} p_{im} = z_m.$$

This however is a contradiction, since $\mu > 0$. Consequently, Farkas (ii) fails, and thus Farkas (i) holds. □

3.2. Convex functions

A convex function is a real-valued function such that the region above its graph is a convex set. Convex functions therefore inherit many useful properties from convex sets.

3.2.1. Convex functions of one variable.

DEFINITION. (i) A function $f : \mathbb{R} \to \mathbb{R}$ is called **convex** if

$$(C_1) \qquad \boxed{f(\theta x + (1-\theta)\hat{x}) \le \theta f(x) + (1-\theta) f(\hat{x})}$$

for all $x, \hat{x} \in \mathbb{R}$, $0 \le \theta \le 1$.

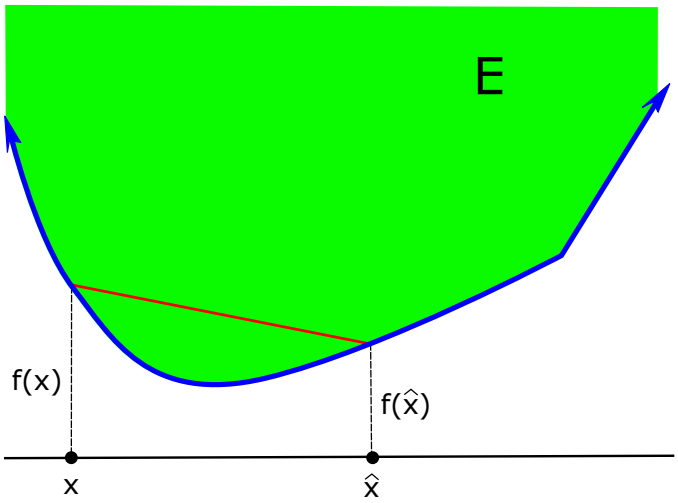
(ii) A function $g : \mathbb{R} \to \mathbb{R}$ is called **concave** if $-g$ is convex.

GEOMETRIC INTERPRETATION. If f is convex, then for all points $x, \hat{x}$ the graph of f lies **below** the line segment connecting $[x\, f(x)]^T$ and $[\hat{x}\, f(\hat{x})]^T$.

It is easy to see that $f : \mathbb{R} \to \mathbb{R}$ is a convex function if and only if its **epigraph**

$$E = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid y \ge f(x), x \in \mathbb{R} \right\} \subset \mathbb{R}^2$$

is a convex set. □



The graph of a convex function

REMARK. It follows by induction that if $f : \mathbb{R} \to \mathbb{R}$ is convex, then

$$\boxed{f\left(\sum_{i=1}^{m} \theta_i x_i\right) \leq \sum_{i=1}^{m} \theta_i f(x_i)} \tag{3.18}$$

for all positive integers m, all $x_1, \ldots, x_m \in \mathbb{R}$, and all $\theta_1, \ldots, \theta_m \geq 0$ such that $\sum_{i=1}^{m} \theta_i = 1$. □

THEOREM 3.2.1 (**Equivalent characterizations of convexity**).

(i) If $f : \mathbb{R} \to \mathbb{R}$ is continuously differentiable, then f is convex if and only if

$$\boxed{f(x) + f'(x)(\hat{x} - x) \leq f(\hat{x})} \tag{C_2}$$

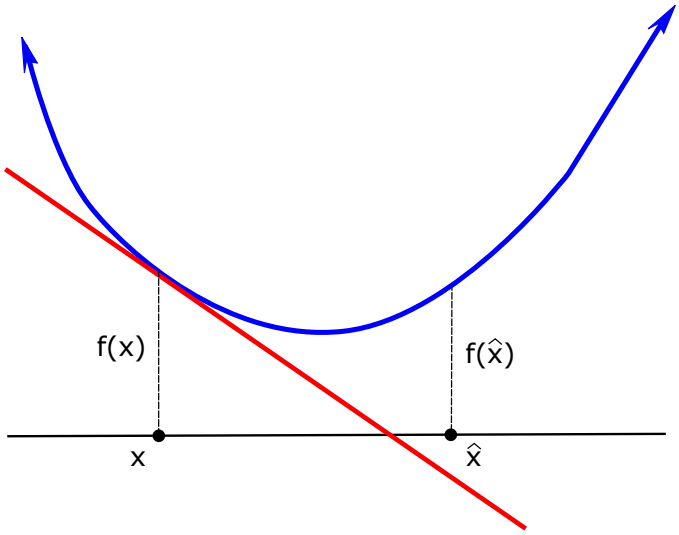
for all $x, \hat{x} \in \mathbb{R}$.

(ii) If f is twice continuously differentiable, then f is convex if and only if

$$\boxed{f''(x) \geq 0} \tag{C_3}$$

for all $x \in \mathbb{R}$.

GEOMETRIC INTERPRETATION. The condition (C_2) means that the graph of the convex function f lies above each of its tangent lines. □



The graph lies above tangent lines

Proof. 1. Assume f is continuously differentiable, and let us show (C_1) holds if and only if (C_2) holds. So suppose (C_1). Then

$$f(\theta\hat{x} + (1-\theta)x) \leq \theta f(\hat{x}) + (1-\theta)f(x),$$

and thus

$$\frac{f(x + \theta(\hat{x} - x)) - f(x)}{\theta} \leq f(\hat{x}) - f(x).$$

Let $\theta \to 0$, to deduce that

$$f'(x) \cdot (\hat{x} - x) \leq f(\hat{x}) - f(x).$$

This is (C_2).

Now assume (C_2). Then if $w = \theta x + (1-\theta)\hat{x}$, we have

$$f(x) \geq f(w) + f'(w)(x - w)$$

and

$$f(\hat{x}) \geq f(w) + f'(w)(\hat{x} - w).$$

So

$$\theta f(x) + (1-\theta)f(\hat{x}) \geq f(w) + f'(w)(\theta(x-w) + (1-\theta)(\hat{x} - w)).$$

But a calculations shows that

$$\theta(x-w) + (1-\theta)(\hat{x} - w) = 0.$$

Thus $\theta f(x) + (1-\theta)f(\hat{x}) \geq f(w) = f(\theta x + (1-\theta)\hat{x})$. This is (C_1).

2. Suppose now that f is twice continuously differentiable. We will show (C_2) holds if and only if (C_3) holds. Assume (C_2). Then

$$\begin{cases} f(x) + f'(x)(w-x) \le f(w) \\ f(w) + f'(w)(x-w) \le f(x). \end{cases}$$

Add, to get

$$(f'(w) - f'(x))(w-x) \ge 0$$

for all $x, w \in \mathbb{R}$. Now put $w = x + \theta y$ for $\theta > 0$:

$$\left(\frac{f'(x+\theta y) - f(x)}{\theta} \right) y \ge 0.$$

Send $\theta \to 0$, to deduce that $f''(x)y^2 \ge 0$ for all $y \in \mathbb{R}$ and thus $f''(x) \ge 0$. This is (C_3).

Now assume (C_3). Then for all x, y

$$\begin{aligned} f(y) &= f(x) + \int_x^y f'(t)\,dt \\ &= f(x) - \int_x^y (y-t)' f'(t)\,dt \\ &= f(x) + f'(x)(y-x) + \int_x^y (y-t) f''(t)\,dt. \end{aligned}$$

If $y > x$, $\int_x^y (y-t)f''(t)\,dt \ge 0$ as $f''(t) > 0$. If $y < x$, then we have $\int_x^y (y-t)f''(t)\,dt = \int_y^x (t-y)f''(t)\,dt \ge 0$, since both the factors of the last integrand are positive. So in both cases, $f(y) \ge f(x) + f'(x)(y-x)$. This is (C_2). □

The condition (C_3) is especially convenient for checking if a given function is convex or not. But the graphs of convex functions can have corners, and so *convex functions need not be twice, or even once, continuously differentiable*. However, they are always continuous:

THEOREM 3.2.2 (Convex functions are continuous). If $f : \mathbb{R} \to \mathbb{R}$ is convex, then

$$f \text{ is continuous.}$$

REMARK. We will later extend our definition of convexity to allow for functions $f : \mathbb{R} \to (-\infty, \infty]$, which may take on the value ∞. Simple examples show that such convex functions need not be continuous. So perhaps a better title for Theorem 3.2.2 is "*finite-valued convex functions are continuous*". □

Proof. 1. First we show that f is bounded on each interval $[a,b]$ of finite length. For each $a \le x \le b$, we can write $x = \theta a + (1-\theta)b$ where $0 \le \theta \le 1$. Therefore

$$f(x) \le \theta f(a) + (1-\theta) f(b) \le \max\{|f(a)|, |f(b)|\}.$$

Now if $\frac{a+b}{2} \le x \le b$, then $\frac{a+b}{2} = \theta x + (1-\theta)a$ for $\theta = \frac{b-a}{2(x-a)}$. Then $\frac{1}{2} \le \theta \le 1$ and convexity implies

$$f(\tfrac{a+b}{2}) \le \theta f(x) + (1-\theta) f(a).$$

Hence

$$f(x) \ge \frac{1}{\theta}\Big(f(\tfrac{a+b}{2}) - (1-\theta) f(a)\Big) \ge -2\Big(|f(\tfrac{a+b}{2})| + |f(a)|\Big),$$

since $\frac{1}{\theta} \le 2$. Likewise, if $a \le x \le \frac{a+b}{2}$, we have

$$f(x) \ge -2\Big(|f(\tfrac{a+b}{2})| + |f(b)|\Big).$$

Therefore

$$\sup_{[a,b]} |f| \le 4 \max\{|f(a)|, |f(\tfrac{a+b}{2})|, |f(b)|\} < \infty. \tag{3.19}$$

This proves assertion (i).

2. Now assume $[a,b] = [-1,1]$ and let $-1 \le x < y \le 1$. Then $y = \theta x + (1-\theta)2$ for $\theta = \frac{2-y}{2-x}$. Thus

$$f(y) \le \theta f(x) + (1-\theta) f(2),$$

and so

$$f(y) - f(x) \le (1-\theta)(f(2) - f(x)) = (y-x)\frac{f(2) - f(x)}{2-x} \le 2|y-x| \sup_{[-2,2]} |f|.$$

Similarly, we have $x = \theta y + (1-\theta)(-2)$ for $\theta = \frac{x+2}{y+2}$. Consequently,

$$f(x) \le \theta f(y) + (1-\theta) f(-2);$$

whence

$$f(x) - f(y) \le (1-\theta)(f(-2) - f(y)) = (y-x)\frac{f(-2) - f(y)}{2+y} \le 2|y-x| \sup_{[-2,2]} |f|.$$

Putting together the above estimates, we see that

$$|f(y) - f(x)| \le 2|y-x| \sup_{[-2,2]} |f| \qquad (x, y \in [-1,1]). \tag{3.20}$$

Since $\sup_{[-2,2]} |f| < \infty$, this inequality implies f is continuous on $[-1,1]$.

Now consider any finite interval $[a,b]$ and set $\hat{f}(x) = f\left(\frac{b-a}{2}x + \frac{a+b}{2}\right)$. Applying estimate (3.20) to the convex function $\hat{f}$ on $[0,1]$, we conclude that f is continuous on $[a,b]$. $\square$

3.2.2. Convex functions of more variables.

DEFINITION. $f : \mathbb{R}^n \to \mathbb{R}$ is called **convex** if

$$(C_1) \qquad \boxed{f(\theta x + (1-\theta)\hat{x}) \leq \theta f(x) + (1-\theta) f(\hat{x}).}$$

for all $x, \hat{x} \in \mathbb{R}^n$, $0 \leq \theta \leq 1$.

REMARKS. We see that f is a convex function if and only if its epigraph

$$E = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid y \geq f(x), x \in \mathbb{R}^n \right\} \subset \mathbb{R}^{n+1}$$

is a convex set. And, as before, it follows by induction that if $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then

$$(3.21) \qquad f\left(\sum_{i=1}^{m} \theta_i x_i \right) \leq \sum_{i=1}^{m} \theta_i f(x_i)$$

for all positive integers m, all $x_1, \ldots, x_m \in \mathbb{R}^n$, and all $\theta_1, \ldots, \theta_m \geq 0$ such that $\sum_{i=1}^{m} \theta_i = 1$. □

Our calculations in the previous section let us fairly easily derive various useful characterizations of multivariable convex functions.

THEOREM 3.2.3 (Equivalent characterization of multivariable convexity).

(i) If $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable, then f is convex if and only if

$$(C_2) \qquad \boxed{f(x) + \nabla f(x) \cdot (\hat{x} - x) \leq f(\hat{x})}$$

for all $x, \hat{x} \in \mathbb{R}^n$.

(ii) If $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable, then f is convex if and only if

$$(C_3) \qquad \boxed{\nabla^2 f(x) \succeq 0}$$

for all $x \in \mathbb{R}^n$.

REMARK. Recall from page 4 that $\nabla^2 f(x) \succeq 0$ means

$$y^T \nabla^2 f(x) y = \sum_{i,j=1}^{n} \frac{\partial^2 f(x)}{\partial x_i \partial x_j} y_i y_j \geq 0.$$

for all $y \in \mathbb{R}^n$. □

GEOMETRIC INTERPRETATION. Now the condition (C_2) means that the graph of the convex function lies above each of its tangent hyperplanes. □

Proof. 1. We first claim that $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if

$$\phi(t) = f(x + ty)$$

is a convex function of t for all $x, y \in \mathbb{R}^n$.

To see this, assume f is convex. Select $x, y \in \mathbb{R}^n$, $0 \leq \theta \leq 1$, $t, \hat{t} \in \mathbb{R}$ and let

$$z = x + ty, \ \hat{z} = x + \hat{t}y.$$

Then

$$f(\theta z + (1-\theta)\hat{z}) \leq \theta \underbrace{f(z)}_{\phi(t)} + (1-\theta) \underbrace{f(\hat{z})}_{\phi(\hat{t})}$$

Observe also that

$$f(\theta z + (1-\theta)\hat{z}) = f\left(x + (\theta t + (1-\theta)\hat{t})y\right) = \phi(\theta t + (1-\theta)\hat{t}).$$

Thus ϕ is convex.

Conversely, let $0 \leq \theta \leq 1$, $x, \hat{x} \in \mathbb{R}^n$. Let $\phi(t) = f(\hat{x} + t(x - \hat{x}))$. If ϕ is convex, then

$$\begin{aligned} f(\theta x + (1-\theta)\hat{x}) = \phi(\theta) &= \phi\left((1-\theta)0 + \theta \cdot 1\right) \\ &\leq (1-\theta)\phi(0) + \theta\phi(1) \\ &= (1-\theta)f(\hat{x}) + \theta f(x). \end{aligned}$$

So if we know that $t \mapsto \phi(t) = f(x + ty)$ is convex for all x, y, then f is convex.

2. Now the one dimensional version of (C_2) implies that

$$\phi(t) + \phi'(t)(\hat{t} - t) \leq \phi(\hat{t}) \quad \text{for all } t, \hat{t} \in \mathbb{R},$$

for the convex function $\phi(t) = f(x + ty)$. Also $\phi'(t) = \nabla f(x + ty) \cdot y$. Let $t = 0$, $\hat{t} = 1$ above, to get

$$f(x) + \nabla f(x) \cdot y \leq f(x + y).$$

Now put $\hat{x} = x + y$; then

$$f(x) + \nabla f(x) \cdot (\hat{x} - x) \leq f(\hat{x}).$$

This is (C_2) for the function f.

3. The one-dimensional version of (C_3) for convex ϕ says

$$\phi''(t) \geq 0,$$

where $\phi(t) = f(x+ty)$. Then

$$\phi'(t) = \nabla f(x+ty)\cdot y, \quad \phi''(t) = y^T \nabla^2 f(x+ty) y.$$

Let $t = 0$, so that

$$y^T \nabla^2 f(x) y \geq 0 \quad \text{for all } y \in \mathbb{R}^n.$$

This is (C_3) for f. □

Checking (C_3) is usually the best way of determining if a given twice-differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is convex or not. But the graph of a general convex function f can have corners, edges, etc and so f need not be differentiable. However, as in one dimension, a real-valued convex function is always continuous:

THEOREM 3.2.4 (Multivariable convex functions are continuous). If $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then

$$f \text{ is continuous.}$$

Proof. 1. Write

$$Q(k) = [-k,k] \times \cdots \times [-k,k] \subset \mathbb{R}^n$$

for the cube centered at the origin, with sides of length $2k$ parallel to the coordinate axes.

We will first show that f restricted to $Q(k)$ is bounded. This follows by induction, the case $n = 1$ being our earlier Theorem 3.2.2. Then for $n \geq 2$ we see from the induction hypothesis that f is bounded on each of the three $n-1$ dimensional boxes

$$Q_i = [-k,k] \times \cdots \times [-k,k] \times \{x_n = c_i\}$$

for $c_1 = -k, c_2 = 0, c_3 = k$.

Now for each point $x' \in [-k,k] \times \cdots \times [-k,k] \subset \mathbb{R}^{n-1}$, the function $g(x_n) = f(x', x_n)$ is convex. Thus estimate (3.19) from the proof of Theorem 3.2.2 shows that

$$\sup_{Q(k)} |f| \leq 4 \max \left\{ \sup_{Q_1} |f|, \sup_{Q_2} |f|, \sup_{Q_3} |f| \right\} < \infty.$$

2. Fix any two points $x, y \in \mathbb{R}^n$, with $0 < |x-y| \leq 1$. Define the convex function $\phi(t) = f(x+tz)$ for $z = \frac{y-x}{|y-x|}$ According to (3.20) from the proof of Theorem 3.2.2, we have the estimate

$$\frac{|\phi(t) - \phi(0)|}{|t|} \leq 2 \max_{|s| \leq 2} |\phi(s)|.$$

Let $t = |y - x|$; then the foregoing implies

$$\frac{|f(y) - f(x)|}{|y - x|} \leq 2 \max_{|z-x|\leq 2} |f(z)|.$$

Consequently, f is continuous at x. □

3.2.3. Subdifferentials.

Our calculations above show also that if $f : \mathbb{R}^n \to \mathbb{R}$ is convex and **if** f is differentiable at x, then

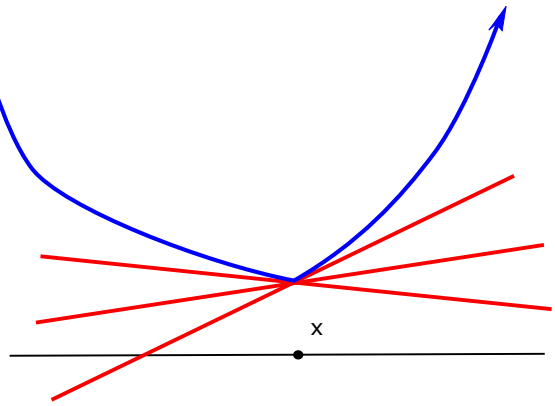
$$f(x) + \nabla f(x) \cdot (\hat{x} - x) \leq f(\hat{x})$$

for all $\hat{x} \in \mathbb{R}^n$. We now extend this inequality, even if f is not differentiable at x:

DEFINITION. Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. For each $x \in \mathbb{R}^n$, we define

$$\boxed{\partial f(x) = \{r \in \mathbb{R}^n \mid f(x) + r \cdot (\hat{x} - x) \leq f(\hat{x}) \quad \text{for all } \hat{x} \in \mathbb{R}^n\}\,.} \tag{3.22}$$

This set is called the **subdifferential** of f at x.



A multivalued subdifferential

GEOMETRIC INTERPRETATION. The affine function

$$g(y) = a \cdot y + b$$

is called a **supporting hyperplane** to the graph of f at x provided $g \leq f$ on $\mathbb{R}^n$ and $g(x) = f(x)$.

The subdifferential $\partial f(x)$ therefore records the "slopes" of all the supporting hyperplanes to the graph of f at x. If f is differentiable at x, then the unique supporting hyperplane is the tangent hyperplane and

$$\partial f(x) = \{\nabla f(x)\}.$$

But in general $\partial f(x)$ may contain more than one element, since the graph of f may have a corner or edge; and thus there may be more than one supporting hyperplane at x. □

EXAMPLES. (i) Let $n = 1$ and $f(x) = |x|$. Then

$$\partial f(x) = \begin{cases} \{1\} & \text{if } x > 0 \\ [-1, 1] & \text{if } x = 0 \\ \{-1\} & \text{if } x < 0. \end{cases}$$

(ii) Now assume $n > 1$ and $f(x) = |x|$. Then

$$\partial f(x) = \begin{cases} \{\frac{x}{|x|}\} & \text{if } x \neq 0 \\ B(0, 1) & \text{if } x = 0. \end{cases}$$

To verify the last statement, observe that $r \in \partial f(0)$ means

$$|0| + r \cdot (x - 0) \leq |x| \quad \text{for all } x \in \mathbb{R}^n.$$

This is valid precisely when $|r| \leq 1$. □

THEOREM 3.2.5. Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. Then for each $x \in \mathbb{R}^n$,

(3.23) $\partial f(x)$ is a closed and convex set.

Proof. 1. The set $\partial f(x)$ is clearly convex. Assume now

$$\{r^k\}_{k=1}^{\infty} \subseteq \partial f(x) \text{ and } r^0 = \lim_{k \to \infty} r^k.$$

Then for each k and each $\hat{x}$

$$f(x) + r^k \cdot (\hat{x} - x) \leq f(\hat{x}).$$

Let $k \to \infty$ to deduce that

$$f(x) + r^0 \cdot (\hat{x} - x) \leq f(\hat{x})$$

for each $\hat{x} \in \mathbb{R}^n$, and hence $r^0 \in \partial f(x)$. Consequently, $\partial f(x)$ is closed. □

Next is the important assertion that a finite-valued convex function on $\mathbb{R}^n$ has a non-empty subdifferential at every point.

THEOREM 3.2.6. Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. Then for each $x \in \mathbb{R}^n$,

(3.24) $\partial f(x)$ is non-empty.

Proof. 1. Select any point $x \in \mathbb{R}^n$. We will show $\partial f(x) \neq \emptyset$.

The epigraph E is convex and closed (since f is continuous). Let k be a positive integer. Then

$$e^k = \begin{bmatrix} x \\ f(x) - \frac{1}{k} \end{bmatrix} \notin E.$$

According then to the Separating Hyperplane Theorem, there exists a hyperplane

$$a_k \cdot z + d^k$$

in $\mathbb{R}^{n+1}$ that strictly separates e^k and E. So there exist

$$a_k = \begin{bmatrix} b^k \\ c^k \end{bmatrix},$$

with $b^k \in \mathbb{R}^n$ and $c^k \in \mathbb{R}$, and $d^k \in \mathbb{R}$, such that

$$a_k \cdot e^k + d^k < 0, \tag{3.25}$$

$$a_k \cdot z + d^k > 0 \qquad (z \in E). \tag{3.26}$$

2. Now (3.25) says

$$b^k \cdot x + c^k \left(f(x) - \frac{1}{k} \right) + d^k < 0. \tag{3.27}$$

Also, for any $\hat{x} \in \mathbb{R}^n$ we have

$$z = \begin{bmatrix} \hat{x} \\ f(\hat{x}) \end{bmatrix} \in E,$$

and therefore (3.26) implies

$$b^k \cdot \hat{x} + c^k f(\hat{x}) + d^k > 0. \tag{3.28}$$

Multiply (3.27) by -1 and add to (3.28):

$$b^k \cdot (\hat{x} - x) + c^k \left(f(\hat{x}) - f(x) + \frac{1}{k} \right) > 0 \tag{3.29}$$

for $\hat{x} \in \mathbb{R}^n$. If we let $\hat{x} = x$, (3.29) says $\frac{c^k}{k} > 0$ and thus $c^k > 0$. Therefore (3.29) implies

$$f(\hat{x}) + \frac{1}{k} \geq r^k \cdot (\hat{x} - x) + f(x) \tag{3.30}$$

for all $\hat{x} \in \mathbb{R}^n$, where

$$r^k = -\frac{1}{c^k} b^k.$$

3. We now claim that the sequence $\{r^k\}_{k=1}^{\infty} \subset \mathbb{R}^n$ is bounded. To see this, let $\hat{x} = x + \frac{r^k}{|r^k|}$ in (3.30) (if $r^k \neq 0$):

$$f\left(x + \frac{r^k}{|r^k|}\right) + \frac{1}{k} \geq |r^k| + f(x).$$

Hence

$$|r^k| \leq 1 + \max_{y \in B(x,1)} |f(y)| + |f(x)| = M$$

for all $k = 1, 2, \ldots$.

According now to the Bolzano-Weierstrass Theorem, there then exists a subsequence such that

$$r = \lim_{j \to \infty} r^{k_j} \quad \text{exists.}$$

Now let $k = k_j \to \infty$ in (3.30):

$$f(\hat{x}) \geq r \cdot (\hat{x} - x) + f(x)$$

for all $\hat{x} \in \mathbb{R}^n$. This says that $r \in \partial f(x)$. □

REMARK (**Infinite-valued convex functions**). It is often useful to allow convex functions f to take the value ∞. We say

(i) $f : \mathbb{R}^n \to (-\infty, \infty]$ is **convex** if

$$f\left(\theta x + (1-\theta)\hat{x}\right) \leq \theta f(x) + (1-\theta) f(\hat{x});$$

(ii) $f : \mathbb{R}^n \to (-\infty, \infty]$ is **lower semicontinuous** if

$$\lim_{k \to \infty} x^k = x^0 \quad \text{implies} \quad f(x^0) \leq \liminf_{k \to \infty} f(x^k).$$

Observe that if $f : \mathbb{R}^n \to (-\infty, +\infty]$ is convex and lower semicontinuous, its epigraph

$$E = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid y \geq f(x), f(x) < \infty \right\}$$

is convex and closed. □

IMPORTANT REMARK. *If $f : \mathbb{R}^n \to (-\infty, \infty]$ is convex and lower semicontinuous, it is possible that $\partial f(x) = \emptyset$, even if $f(x) < \infty$.*

Here is an example:

$$f(x) = \begin{cases} \infty & (x < -1) \\ -(1 - x^2)^{1/2} & (-1 \leq x \leq 1) \\ \infty & (x > 1), \end{cases}$$

for which $\partial f(\pm 1) = \emptyset$. □

REMARK. However, the proof of Theorem 3.2.6 can be modified to show that *if $f : \mathbb{R}^n \to (-\infty, \infty]$ is convex and lower semicontinuous, and if f is finite-valued on an open, convex set $U \subset \mathbb{R}^n$, then*

$$\partial f(x) \neq \emptyset \qquad (x \in U). \tag{3.31}$$

We omit a full discussion of this observation, other than to note that if the ball $B(x, 2r)$ lies in U, then estimates from earlier proofs show that f restricted to the smaller ball $B(x, r)$ is bounded and continuous. □

3.2.4. Dual convex functions.

For this section, assume $f : \mathbb{R}^n \to \mathbb{R}$ is convex, with

$$\lim_{|x|\to\infty} \frac{f(x)}{|x|} = +\infty. \tag{3.32}$$

This is called a "super-linear growth" condition.

DEFINITION. For $y \in \mathbb{R}^n$, we define

$$\boxed{f^*(y) = \max_{x\in\mathbb{R}^n} \{x \cdot y - f(x)\}.} \tag{3.33}$$

We call f^* the **convex dual function** (or **Legendre transform**) of f.

EXAMPLES. (i) Let $f(x) = x^2/2$ for $x \in \mathbb{R}$. Then

$$f^*(y) = \max_x \left(xy - \frac{x^2}{2} \right) = \frac{y^2}{2}.$$

(ii) Let $f(x) = |x|^p/p$, where $1 < p < \infty$. We claim that

$$f^*(y) = \frac{|y|^q}{q} \quad \text{where} \quad \frac{1}{p} + \frac{1}{q} = 1.$$

To confirm this, we write

$$f^*(y) = \max_x \left(xy - \frac{|x|^p}{p} \right)$$

For fixed y, let $g(x) = xy - |x|^p/p$. Then $g'(x) = y - |x|^{p-1}\operatorname{sgn}(x)$. So $0 = g'(x)$ when $y = |x|^{p-1}\operatorname{sgn}(x)$, or equivalently $x = |y|^{1/(p-1)}\operatorname{sgn}(y)$. Consequently

$$\begin{aligned} f^*(y) &= \left(|y|^{1/(p-1)}\operatorname{sgn}(y) \right) y - \frac{\left| |y|^{1/(p-1)}\operatorname{sgn}(y) \right|^p}{p} \\ &= |y|^{\frac{1}{p-1}+1} - \frac{|y|^{\frac{p}{p-1}}}{p} = \left(1 - \frac{1}{p}\right) |y|^{p/(p-1)} \end{aligned}$$

$$= \frac{|y|^q}{q},$$

where $q = \frac{p}{p-1}$ (and so $\frac{1}{p} + \frac{1}{q} = 1$). □

LEMMA 3.2.1. For all $x, y \in \mathbb{R}^n$ we have

$$\boxed{x \cdot y \leq f(x) + f^*(y).} \tag{3.34}$$

This is the **Fenchel–Young inequality**.

THEOREM 3.2.7 (Dual convex functions).

(i) The function $f^* : \mathbb{R}^n \to \mathbb{R}$ is convex.

(ii) $\lim_{|y|\to\infty} \frac{f^*(y)}{|y|} = +\infty$

(iii) Furthermore

$$\boxed{f^{**} = f.} \tag{3.35}$$

Proof. 1. We have

$$\begin{aligned} f^*\left(\theta y + (1-\theta)\hat{y}\right) &= \max_x \left((\theta y + (1-\theta)\hat{y}) \cdot x - f(x)\right) \\ &= \max_x \left(\theta(y \cdot x - f(x)) + (1-\theta)(\hat{y} \cdot x - f(x))\right) \\ &\leq \theta \max_x \left(y \cdot x - f(x)\right) + (1-\theta) \max_x \left(\hat{y} \cdot x - f(x)\right) \\ &= \theta f^*(y) + (1-\theta) f^*(\hat{y}). \end{aligned}$$

2. Recall that $f(x) + f^*(y) \geq x \cdot y$ for all x, y. Fix $y \neq 0, \mu > 0$ and let $x = \mu y/|y|$. Then

$$f^*(y) \geq \left(\mu \frac{y}{|y|}\right) \cdot y - f\left(\mu \frac{y}{|y|}\right) \geq \mu|y| - \max_{B(0,\mu)} f.$$

So

$$\frac{f^*(y)}{|y|} \geq \mu - \frac{1}{|y|} \max_{B(0,\mu)} f,$$

and thus

$$\liminf_{|y|\to\infty} \frac{f^*(y)}{|y|} \geq \mu$$

for all $\mu > 0$.

3. Since $f^*(y) + f(x) \geq x \cdot y$, we have

$$f(x) \geq \max_y \left(x \cdot y - f^*(y)\right) = f^{**}(x). \tag{3.36}$$

Conversely, recall from Theorem 3.2.6 that $\partial f(x) \neq \emptyset$. Select $r \in \partial f(x)$; then

$$f(z) \geq f(x) + r \cdot (z - x)$$

for all $z \in \mathbb{R}^n$. Consequently,

$$r \cdot x - f(x) = \max_z (r \cdot z - f(z)) = f^*(r);$$

and so

$$f^{**}(x) = \max_y (x \cdot y - f^*(y)) \geq x \cdot r - f^*(r) = f(x).$$

This and (3.36) imply $f^{**} = f$. □

REMARK. This proof depends upon Theorem 3.2.6, which in turn depends upon the Separating Hyperplane Theorem. So again, as on page 85, we see that *separating hyperplanes imply convex duality.* □

THEOREM 3.2.8 (Subdifferentials and dual functions). For all points $x, y \in \mathbb{R}^n$, we have

$$x \cdot y = f(x) + f^*(y)$$

if and only if

$$y \in \partial f(x)$$

if and only if

$$x \in \partial f^*(y).$$

REMARK. This means in particular that if both $\nabla f(x)$ and $\nabla f^*(y)$ exist, then

$$y = \nabla f(x) \text{ if and only if } x = \nabla f^*(y).$$

Thus $(\nabla f)^{-1} = \nabla f^*$. □

Proof. 1. Assume $x \cdot y = f(x) + f^*(y)$. Then

$$x \cdot y - f(x) = f^*(y) \geq y \cdot z - f(z)$$

for all z. Hence

$$f(z) \geq y \cdot (z - x) + f(x)$$

for all $z \in \mathbb{R}^n$. This implies $y \in \partial f(x)$.

2. Conversely, if $y \in \partial f(x)$, we have

$$f(z) \geq y \cdot (z - x) + f(x)$$

for all z and so

$$x \cdot y - f(x) \geq y \cdot z - f(z)$$

for all z. Maximizing over z gives

$$x \cdot y \geq f(x) + f^*(y).$$

Since we always have $x \cdot y \leq f(x) + f^*(y)$, we have equality.

3. The proof that $x \cdot y = f(x) + f^*(y)$ if and only if $x \in \partial f^*(y)$ is similar. □

REMARK. If $f : \mathbb{R}^n \to (-\infty, \infty]$ is convex, lower semicontinuous, $f \not\equiv \infty$, we can still define the dual function

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \{x \cdot y - f(x)\}.$$

This definition also makes sense if f is finite-valued, but does not satisfy the superlinear growth condition (3.32). For example, let $n = 1$ and $f(x) = |x|$. Then

$$f^*(y) = \begin{cases} 0 & -1 \le y \le 1 \\ +\infty & \text{otherwise.} \end{cases}$$

□

3.2.5. Applications.

a. Gradient flows. Suppose we are given a function $\Phi : \mathbb{R}^n \to \mathbb{R}$. Consider next the system of ODE

$$\begin{cases} \dot{\mathbf{x}} = -\nabla \Phi(\mathbf{x}) & (t \ge 0) \\ \mathbf{x}(0) = x_0 \end{cases} \tag{3.37}$$

which describes a "downhill" gradient flow.

THEOREM 3.2.9. Assume $\Phi : \mathbb{R}^n \to \mathbb{R}$ is convex and $\mathbf{x}$ solves (3.37).

(i) Then

$$\phi(t) = \Phi(\mathbf{x}(t)) \text{ is nonincreasing and convex}$$

(ii) If $\mathbf{y}$ solves

$$\begin{cases} \dot{\mathbf{y}} = -\nabla \Phi(\mathbf{y}) & (t \ge 0) \\ \mathbf{y}(0) = y_0, \end{cases} \tag{3.38}$$

then

$$|\mathbf{x}(t) - \mathbf{y}(t)| \le |\mathbf{x}(s) - \mathbf{y}(s)| \qquad \text{for all } 0 \le s < t. \tag{3.39}$$

Proof. 1. We calculate

$$\dot{\phi} = \nabla \Phi(\mathbf{x}) \cdot \dot{\mathbf{x}} = -|\nabla \Phi(\mathbf{x})|^2 \le 0;$$

and

$$\ddot{\phi} = -\frac{d}{dt}|\nabla \Phi(\mathbf{x})|^2 = -2\nabla \Phi(\mathbf{x}) \cdot \nabla^2 \Phi(\mathbf{x}) \dot{\mathbf{x}}$$
$$= 2\nabla \Phi(\mathbf{x})^T \nabla^2 \Phi(\mathbf{x}) \nabla \Phi(\mathbf{x}) \ge 0.$$

2. If $\mathbf{y}$ solves (3.38), we have

$$\frac{d}{dt}|\mathbf{x}-\mathbf{y}|^2 = 2(\mathbf{x}-\mathbf{y})\cdot(\dot{\mathbf{x}}-\dot{\mathbf{y}}) = -2(\mathbf{x}-\mathbf{y})\cdot(\nabla\Phi(\mathbf{x})-\nabla\Phi(\mathbf{y})) \le 0.$$

□

b. Inequalities. Many useful inequalities in mathematics are consequences of convexity:

THEOREM 3.2.10. (Jensen's inequality) Assume $f:\mathbb{R}\to\mathbb{R}$ is convex. Then for all $-\infty < a < b < \infty$ and all integrable functions $g:[a,b]\to\mathbb{R}$ we have the inequality

$$\boxed{f\left(\frac{1}{b-a}\int_a^b g(x)\,dx\right) \le \frac{1}{b-a}\int_a^b f(g(x))\,dx.} \tag{3.40}$$

Proof. Let $y = \frac{1}{b-a}\int_a^b g(x)\,dx$ and select $r\in\partial f(y)$. Then

$$f(y) + r(g(x)-y) \le f(g(x)) \qquad (a\le x\le b).$$

Now integrate in x over the interval $[a,b]$ and divide by $b-a$. □

EXAMPLE. If $a_1, a_2, \dots, a_m > 0$, then

$$\boxed{(a_1a_2\cdots a_m)^{\frac{1}{m}} \le \frac{a_1+a_2+\cdots+a_m}{m}.} \tag{3.41}$$

This is the **inequality of the geometric and arithmetic means**.

To prove this, we take $f(x) = e^x$ and

$$g(x) = \begin{cases} \log a_1 & (0\le x < \frac{1}{m}) \\ \log a_2 & (\frac{1}{m}\le x < \frac{2}{m}) \\ \quad\vdots & \\ \log a_m & (\frac{m-1}{m}\le x\le 1). \end{cases}$$

Then Jensen's inequality implies

$$\begin{aligned}(a_1a_2\cdots a_m)^{\frac{1}{m}} = e^{\frac{\log a_1+\cdots+\log a_m}{m}} &= f\left(\int_0^1 g\,dx\right) \\ &\le \int_0^1 f(g)\,dx = \frac{a_1+a_2+\cdots+a_m}{m}.\end{aligned}$$

□

EXAMPLE. If $p, q > 1$ satisfy

$$\frac{1}{p} + \frac{1}{q} = 1,$$

then

$$\boxed{ab \le \frac{a^p}{p} + \frac{b^q}{q}.} \tag{3.42}$$

for all $a, b > 0$. This is **Young's inequality**.

To see this, we again take $f(x) = e^x$ and now

$$g(x) = \begin{cases} p \log a & (0 \le x < \frac{1}{p}) \\ q \log b & (\frac{1}{p} \le x \le 1). \end{cases}$$

Then Jensen's inequality implies

$$ab = e^{\frac{p \log a}{p} + \frac{q \log b}{q}} = f\left(\int_0^1 g \, dx \right) \le \int_0^1 f(g) \, dx = \frac{a^p}{p} + \frac{b^q}{q}.$$

This is also a special case of the general Fenchel–Young inequality. □

Chapter 4

NONLINEAR OPTIMIZATION

In this chapter we examine minimization problems with *inequality* constraints and study when and how Lagrange multipliers can be used to characterize minimizers.

4.1. Inequality constraints

Assume $f, h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable.

NOTATION. As usual, we write

$$\mathbf{h} = \begin{bmatrix} h_1 \\ \vdots \\ h_p \end{bmatrix}$$

and

$$\nabla \mathbf{h} = \begin{bmatrix} (\nabla h_1)^T \\ \vdots \\ (\nabla h_p)^T \end{bmatrix} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_p}{\partial x_1} & \cdots & \frac{\partial h_p}{\partial x_n} \end{bmatrix}. \quad \square$$

We study in this section the constrained optimization problem to find $x_0 \in \mathbb{R}^n$ to

$$(MIN^*) \qquad \boxed{\begin{cases} \text{minimize } f(x), \\ \text{subject to } \mathbf{h}(x) \leq 0. \end{cases}}$$

The requirements that $h_j(x) \leq 0$ for $j = 1, \ldots, p$ are **inequality constraints**; the j-th constraint is **active** if $h_j(x) = 0$. A point x is **feasible** for (MIN*) if $\mathbf{h}(x) \leq 0$.

A basic question is how to characterize x_0 solving (MIN*).

4.1.1. Constraint qualification.

Suppose hereafter x_0 solves (MIN*). Our plan is to make a first variation calculation, but for this we need to be careful in designing an appropriate curve of variations staying within the feasible region.

NOTATION. Write

$$J = \{j \in \{1, \ldots, p\} \mid h_j(x_0) = 0\}.$$

These are the indices of the **active constraints** for x_0.

NOTATION. Below we write "$o(t)$" to denote any vector function $\mathbf{r}(t)$ such that

$$\lim_{t \to 0^+} \frac{|\mathbf{r}(t)|}{t} = 0.$$

DEFINITION. We say the **constraint qualification condition (CQ)** holds at x_0 if for each vector $y \in \mathbb{R}^n$ satisfying

$$y \cdot \nabla h_j(x_0) \leq 0 \quad (j \in J), \tag{4.1}$$

there exists a continuous curve $\{\mathbf{x}(t) \mid 0 \leq t < t_0\}$ for some $t_0 > 0$ such that

$$\mathbf{h}(\mathbf{x}(t)) \leq 0 \qquad (0 \leq t < t_0); \tag{4.2}$$

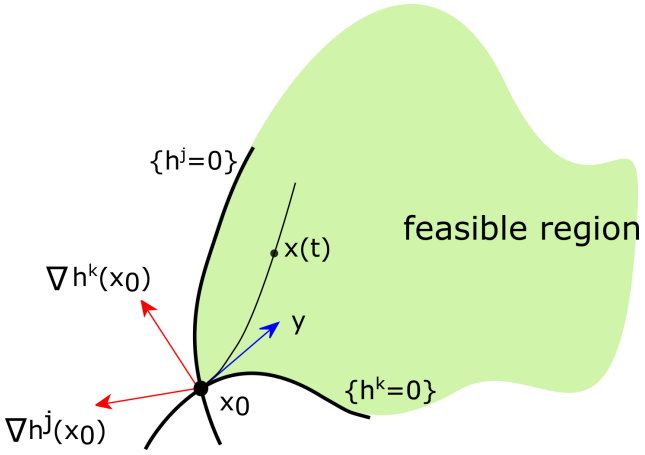
and

$$\mathbf{x}(t) = x_0 + ty + o(t) \quad \text{as } t \to 0^+. \tag{4.3}$$

The condition (4.2) says that $\mathbf{x}(t)$ is feasible for all $0 \leq t < t_0$. And (4.3) means that the right hand derivative $\mathbf{x}'(0)$ exists, with

$$\mathbf{x}'(0) = y.$$

GEOMETRIC INTERPRETATION. We must be careful not to leave the feasible set when we take our variations; and this is potentially problematic if the j-th constraint is active, so that $h_j(x_0) = 0$. But if $y \cdot \nabla h_j(x_0) \leq 0$, then our moving away from x_0 along the curve $\mathbf{x}(t)$ in the direction y will not increase $h_j(\mathbf{x}(t))$, at least to first-order.



So (CQ) is the reasonable requirement that if $y \cdot \nabla h_j(x_0) \leq 0$ for all the active constraints at x_0, there is indeed a curve $\mathbf{x}(t)$ of feasible variations with $\mathbf{x}'(0) = y$. (But see Franklin's book [**F1**] for an example showing that (CQ) can fail.) □

4.1.2. Karush-Kuhn-Tucker conditions.

THEOREM 4.1.1. Let x_0 solve (MIN*) and suppose the constraint qualification condition (CQ) holds at x_0.

Then there exist real numbers $\mu_0^1, \ldots, \mu_0^p$ such that

$$\boxed{\nabla f(x_0) + \sum_{j=1}^{p} \mu_0^j \nabla h_j(x_0) = 0}. \tag{4.4}$$

Furthermore, the vector $\mu_0 = [\mu_0^1, \ldots, \mu_0^p]^T$ satisfies

$$\boxed{\mu_0 \geq 0, \quad \mu_0 \cdot \mathbf{h}(x_0) = 0.} \tag{4.5}$$

REMARK. We call (4.4) and (4.5) the **Karush-Kuhn-Tucker (KKT) conditions**. Observe that we can also write (4.4) as

$$\nabla f(x_0) + \nabla \mathbf{h}(x_0)^T \mu_0 = 0.$$

□

INTERPRETATION. We interpret μ_0^j as the Lagrange multiplier for the constraint $h_j(x_0) \leq 0$. So for our inequality constrained problem (MIN*) we are asserting both that Lagrange multipliers exist and that they are nonnegative.

In addition, if $h_j(x_0) < 0$ for some index j, that constraint is inactive and so the corresponding Lagrange multiplier μ_0^j equals zero. This is a **complementary slackness** condition. □

Proof. 1. Assume the vector y satisfies (4.1). Let $\{\mathbf{x}(t) \mid 0 \le t < t_0\}$ be the corresponding curve, whose existence is assured according to (CQ).

Write $\phi(t) = f(\mathbf{x}(t))$. Then

$$\phi(0) = f(x_0) \le f(\mathbf{x}(t)) = \phi(t) \quad (0 \le t \le t_0),$$

since x_0 solves (MIN*). Thus ϕ has a minimum at $t = 0$ and hence

$$\phi'(0) \ge 0.$$

Now

$$\phi'(0) = \nabla f(x_0) \cdot \mathbf{x}'(0) = \nabla f(x_0) \cdot y,$$

and therefore

$$\nabla f(x_0) \cdot y \ge 0$$

for all y satisfying (4.1).

So we have shown that

$$y \cdot \nabla h_j(x_0) \le 0 \ \ (j \in J) \quad \text{implies} \quad y \cdot \nabla f(x_0) \ge 0. \tag{4.6}$$

2. Now recall the Farkas alternative:

(i) $Ax = b, x \ge 0$ has a solution x, or

(ii) $A^T y \ge 0, y \cdot b < 0$ has a solution y,

but not both. We apply this to

$$A = -\underbrace{\left[\nabla h^{j_1}(x_0) \;\middle|\; \nabla h^{j_2}(x_0) \;\middle|\; \dots \;\middle|\; \nabla h^{j_k}(x_0)\right]}_{\text{columns}}, \ b = \nabla f(x_0),$$

where $J = \{j_1, j_2, \dots, j_k\}$. Then assertion (4.6) says

$$A^T y \ge 0 \quad \text{implies} \quad y \cdot b \ge 0$$

and therefore Farkas (ii) fails.

Consequently Farkas (i) holds: there exist $\sigma_j \ge 0$ $(j \in J)$ such that

$$-\sum_{j \in J} \sigma_j \nabla h_j(x_0) = \nabla f(x_0). \tag{4.7}$$

Define $\mu_0 \in \mathbb{R}^p$ by

$$\mu_0^j = \begin{cases} \sigma_j & j \in J \\ 0 & j \notin J; \end{cases}$$

then $\mu_0 \geq 0$, $\mu_0 \cdot \mathbf{h}(x_0) = 0$ and

$$\nabla f(x_0) + \sum_{j=1}^{p} \mu_0^j \nabla h_j(x_0) = 0.$$

□

4.1.3. When does (CQ) hold?

The above proof is elegant, but it may be far from clear for particular problems if the constraint qualification condition is valid. We discuss next two important cases.

a. Linear inequality and equality constraints

THEOREM 4.1.2. If the functions $\{h_j\}_{j=1}^p$ are linear (or affine) functions of x, then (CQ) holds for each point x_0.

Proof. Let y satisfy the condition (4.1) in (CQ). We will use the straight line $\mathbf{x}(t) = x_0 + ty$ for $0 \leq t \leq t_0$. So if $h_j(x) = a \cdot x + b$, then

$$h_j(\mathbf{x}(t)) = a \cdot \mathbf{x}(t) + b = a \cdot (x_0 + ty) + b.$$

Now $\nabla h_j = a$, and thus the condition (4.1) in (CQ) says

$$y \cdot a \leq 0.$$

So if $h_j(x_0) = 0$ (that is, if $j \in J$), then

$$h_j(\mathbf{x}(t)) = \underbrace{a \cdot x_0 + b}_{h_j(x_0)=0} + t \underbrace{a \cdot y}_{\leq 0} \leq 0$$

for all $t \geq 0$. If on the other hand $h_j(x_0) < 0$, then $h_j(\mathbf{x}(t)) < 0$ for small $t > 0$, by continuity. □

EXAMPLE. (Linear programming redux) Remember the standard linear programming problems:

$$(\mathrm{P}^*) \begin{cases} \min\, c \cdot x, \\ \text{subject to} \\ Ax \geq b,\ x \geq 0. \end{cases} \qquad (\mathrm{D}^*) \begin{cases} \max\, b \cdot y, \\ \text{subject to} \\ A^T y \leq c,\ y \geq 0. \end{cases}$$

Now (P*) is equivalent to (MIN*) for

$$f(x) = c \cdot x, \quad \mathbf{h} = \begin{bmatrix} -x \\ b - Ax \end{bmatrix}.$$

Since the inequality constraints are linear, (CQ) holds according to Theorem 4.1.2. Hence we can find

$$\mu_0 = \begin{bmatrix} \lambda_0 \\ y_0 \end{bmatrix} \in \mathbb{R}^{n+m}$$

such that

$$\nabla f + \nabla(\lambda_0 \cdot (-x)) + \nabla(y_0 \cdot (b - Ax)) = 0 \tag{4.8}$$

at x_0 and

$$\lambda_0, y_0 \geq 0, \quad \lambda_i^0 = 0 \ \text{ if } x_i^0 > 0, \quad y_j^0 = 0 \ \text{ if } (Ax_0 - b)_j > 0 \tag{4.9}$$

for $i = 1, \ldots, n$, $j = 1, \ldots, m$.

Now (4.8) says

$$c = \lambda_0 + A^T y_0.$$

Then $A^T y_0 \leq c$, as $\lambda_0 \geq 0$. Since also $y_0 \geq 0$, we see that y_0 is feasible for the dual problem (D*). In addition, (4.9) tells us that

$$y_0 \cdot (Ax_0 - b) = 0, \quad x_0 \cdot (A^T y_0 - c) = -x_0 \cdot \lambda_0 = 0.$$

Therefore

$$b \cdot y_0 = y_0 \cdot Ax_0 = A^T y_0 \cdot x_0 = x_0 \cdot c$$

and so y_0 is an optimal solution of (D*).

We can thus interpret the components of an optimal solution y_0 of the dual problem (D*) *as Lagrange multipliers for the primal problem* (P*). More precisely, the entries of y_0 are Lagrange multipliers corresponding to the constraints $Ax \geq b$. □

b. Regular equality constraints

Recall that

$$J = \{j \in \{1, \ldots, p\} \mid h_j(x_0) = 0\}$$

are the indices for the active constraints at x_0.

DEFINITION. We say that x_0 is **regular** for (MIN*) if the vectors

$$\{\nabla h_j(x_0)\}_{j \in J} \text{ are linearly independent in } \mathbb{R}^n.$$

THEOREM 4.1.3. If x_0 is regular for (MIN*), then (CQ) holds at x_0.

Proof. 1. Upon reindexing if necessary, we may assume that $J = \{1, \ldots, k\}$ where $k \leq \min\{p, n\}$. Since x_0 is regular, we can select vectors $\{a^{k+1}, \ldots, an\}$

so that the $n \times n$ matrix

$$A(x) = \begin{bmatrix} (\nabla h_1)^T \\ \vdots \\ (\nabla h_k)^T \\ (a^{k+1})^T \\ \vdots \\ (a^n)^T \end{bmatrix} \tag{4.10}$$

is nonsingular for all x sufficiently close to x_0.

2. Let y satisfy the condition (4.1) in (CQ); so that

$$y \cdot \nabla h_j(x_0) \leq 0 \quad (j = 1, \ldots, k).$$

Define

$$e = A(x_0)y;$$

then

$$e_j = \nabla h_j(x_0) \cdot y \leq 0 \quad (j = 1, \ldots, k). \tag{4.11}$$

3. Next define the vector field

$$\mathbf{w}(x) = A^{-1}(x)A(x_0)y = A^{-1}(x)e$$

for x near x_0; and solve the system of differential equations

$$\begin{cases} \mathbf{x}'(t) = \mathbf{w}(\mathbf{x}(t)) \quad (0 \leq t < t_1) \\ \mathbf{x}(0) = x_0. \end{cases}$$

According to standard ODE theory, there exists a unique solution.

We show next that this curve meets the requirements of the (CQ) condition. Note first that

$$\mathbf{x}'(0) = \mathbf{w}(\mathbf{x}(0)) = \mathbf{w}(\mathbf{x}_0) = A^{-1}(x_0)A(x_0)y = y;$$

consequently $\mathbf{x}(t) = x_0 + ty + o(t)$ as $t \to 0$. We assert also that $\mathbf{h}(\mathbf{x}(t)) \leq 0$ for all small enough times t. To see this, first let $j \in J = \{1, \ldots, k\}$. Then $h_j(x_0) = 0$ and

$$\begin{aligned} h_j(\mathbf{x}(t))' &= \nabla h_j(\mathbf{x}(t)) \cdot \mathbf{x}'(t) \\ &= \nabla h_j(\mathbf{x}(t)) \cdot \mathbf{w}(\mathbf{x}(t)) \\ &= \nabla h_j(\mathbf{x}(t)) \cdot A^{-1}(\mathbf{x}(t))e \\ &= e_j \leq 0, \end{aligned}$$

according to (4.10) and (4.11). Therefore $h_j(\mathbf{x}(t)) \leq 0$ for times $0 \leq t < t_1$ if $j \in J$. Furthermore, for indices $j \notin J$, we have $h_j(x_0) < 0$; and consequently,

by continuity, $g^j(\mathbf{x}(t)) < 0$ for $0 \leq t < t_0$, provided we select $0 < t_0 \leq t_1$ small enough. □

4.2. More on Lagrange multipliers

Now we introduce a general minimization problem, with both inequality and equality constraints, and show how to use the penalty method (introduced in Chapter 1) to build Lagrange multipliers. This is an alternative to the techniques introduced in the previous section.

NOTATION. Given $f, g_1, \ldots, g_m, h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$, we write

$$\mathbf{g} = \begin{bmatrix} g_1 \\ \vdots \\ g_m \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} h_1 \\ \vdots \\ h_p \end{bmatrix}.$$

Our minimization problem for this section is to find $x_0 \in \mathbb{R}^n$ to

$$(MIN^{**}) \qquad \boxed{\begin{cases} \text{minimize } f(x), \text{ subject to} \\ \mathbf{g}(x) = 0, \ \mathbf{h}(x) \leq 0. \end{cases}}$$

So there are m equality constraints and p inequality constraints.

4.2.1. F. John's formulation.

THEOREM 4.2.1. Suppose that x_0 solves the constrained optimization problem (MIN**).

Then there exist real numbers $\gamma_0, \lambda_0^1, \ldots, \lambda_0^m, \mu_0^1, \ldots, \mu_0^p$, **not all equal to zero**, such that

$$\boxed{\gamma_0 \nabla f(x_0) + \sum_{k=1}^{m} \lambda_0^k \nabla g_k(x_0) + \sum_{j=1}^{p} \mu_0^j \nabla h_j(x_0) = 0} \tag{4.12}$$

and

$$\boxed{\gamma_0 \geq 0, \quad \mu_0 \geq 0, \quad \mu_0 \cdot \mathbf{h}(x_0) = 0.} \tag{4.13}$$

NOTATION. Here

$$\lambda_0 = [\lambda_0^1 \ \ldots \ \lambda_0^m]^T, \quad \mu_0 = [\mu_0^1 \ \ldots \ \mu_0^p]^T.$$

We can also write (4.12) as

$$\gamma_0 \nabla f(x_0) + \nabla \mathbf{g}(x_0)^T \lambda_0 + \nabla \mathbf{h}(x_0)^T \mu_0 = 0.$$

□

TERMINOLOGY. (i) The formulas (4.12) and (4.13) are **F. John's conditions.**

(ii) When $\gamma_0 = 1$, (4.12) and (4.13) become

$$\boxed{\nabla f(x_0) + \sum_{k=1}^{m} \lambda_0^k \nabla g_k(x_0) + \sum_{j=1}^{p} \mu_0^j \nabla h_j(x_0) = 0} \tag{4.14}$$

and

$$\boxed{\mu_0 \geq 0, \quad \mu_0 \cdot \mathbf{h}(x_0) = 0.} \tag{4.15}$$

To be consistent with previous notation, we will call (4.14) and (4.15) the **Karush-Kuhn-Tucker (KKT) conditions**.

(iii) If $\gamma_0 \neq 0$, we can divide and convert (4.12) into the KKT form (4.14). If instead $\gamma_0 = 0$, we call it an **abnormal multiplier**. □

Proof. 1. For each $\alpha > 0$ define

$$F^\alpha(x) := f(x) + \frac{\alpha}{2}(|\mathbf{g}(x)|^2 + |\mathbf{h}_+(x)|^2) + \frac{1}{2}|x - x_0|^2, \tag{4.16}$$

where for $j = 1, \ldots, p$ we define

$$h_j^+(x) = \begin{cases} h_j(x) & \text{if } \ h_j(x) \geq 0 \\ 0 & \text{if } \ h_j(x) \leq 0. \end{cases}$$

Let B be the closed ball with center x_0 and radius 1. According to the Extreme Value Theorem there exists a point $x_\alpha \in B$ such that

$$F^\alpha(x_\alpha) = \min_{x \in B} F^\alpha(x).$$

Thus $F^\alpha(x_\alpha) \leq F^\alpha(x_0)$ and hence

$$f(x_\alpha) + \frac{\alpha}{2}(|\mathbf{g}(x_\alpha)|^2 + |\mathbf{h}_+(x_\alpha)|^2) + \frac{1}{2}|x_\alpha - x_0|^2 \leq f(x_0), \tag{4.17}$$

since $\mathbf{h}(x_0) \leq 0, \mathbf{g}(x_0) = 0$. Therefore $\{\alpha|\mathbf{g}(x_\alpha)|^2\}_{\alpha>0}$ and $\{\alpha|\mathbf{h}_+(x_\alpha)|^2\}_{\alpha>0}$ are bounded, and consequently

$$\lim_{\alpha \to \infty} \mathbf{g}(x_\alpha) = 0, \ \lim_{\alpha \to \infty} \mathbf{h}_+(x_\alpha) = 0. \tag{4.18}$$

2. Next, use the Bolzano-Weierstrass Theorem to select a convergent subsequence $\{x_{\alpha_j}\}_{j=1}^{\infty}$ of $\{x_\alpha\}_{\alpha>0} \subset B$ so that

$$x_{\alpha_j} \to \bar{x}$$

as $\alpha_j \to \infty$, for some $\bar{x} \in B$. Then (4.17) implies

$$f(x_\alpha) + \frac{1}{2}|x_\alpha - x_0|^2 \leq f(x_0)$$

and so

$$f(\bar{x}) + \frac{1}{2}|\bar{x} - x_0|^2 \leq f(x_0).$$

But (4.18) gives $\mathbf{h}(\bar{x}) \leq 0, \mathbf{g}(\bar{x}) = 0$ and therefore $\bar{x}$ is feasible. Hence

$$f(x_0) \leq f(\bar{x})$$

since x_0 solves (MIN**). Therefore $\bar{x} = x_0$. This is true for all convergent subsequences $x_{\alpha_j} \to \bar{x}$ and thus

$$\lim_{\alpha \to \infty} x_\alpha = x_0. \tag{4.19}$$

3. Then for k large enough x_α does not lie on the boundary of B . Thus $x \mapsto F^\alpha(x)$ has an unconstrained local minimum at x_α and hence

$$\begin{aligned} 0 &= \nabla F^\alpha(x_\alpha) \\ &= \nabla f(x_\alpha) + \alpha(\nabla \mathbf{g}(x_\alpha)^T \mathbf{g}(x_\alpha) + \nabla \mathbf{h}_+(x_\alpha)^T \mathbf{h}_+(x_\alpha)) + x_\alpha - x_0 \\ &= \nabla f(x_\alpha) + \alpha(\nabla \mathbf{g}(x_\alpha)^T \mathbf{g}(x_\alpha) + \nabla \mathbf{h}(x_\alpha)^T \mathbf{h}_+(x_\alpha)) + x_\alpha - x_0, \end{aligned}$$

since $\nabla h_j^+ = \nabla h_j$ if $h_j^+ \neq 0$.

We next multiply this identity by the constant

$$\gamma_\alpha = (1 + \alpha^2|\mathbf{g}(x_\alpha)|^2 + \alpha^2|\mathbf{h}_+(x_\alpha)|^2)^{-\frac{1}{2}} > 0.$$

This gives

$$0 = \gamma_\alpha \nabla f(x_\alpha) + \nabla \mathbf{g}(x_\alpha)^T \lambda_\alpha + \nabla \mathbf{h}(x_\alpha)^T \mu_\alpha + \gamma_\alpha(x_\alpha - x_0) \tag{4.20}$$

for

$$\lambda_\alpha = \gamma_\alpha \alpha \mathbf{g}(x_\alpha), \ \mu_\alpha = \gamma_\alpha \alpha \mathbf{h}_+(x_\alpha) \geq 0.$$

4. Observe that

$$(\gamma_\alpha)^2 + |\lambda_\alpha|^2 + |\mu_\alpha|^2 = 1,$$

and therefore $\{(\gamma_\alpha, \lambda_\alpha, \mu_\alpha)\}_{\alpha > 0}$ is bounded. Hence there is a sequence $\alpha_j \to \infty$ such that

$$\gamma_{\alpha_j} \to \gamma_0 \geq 0 \text{ in } \mathbb{R}, \quad \lambda_{\alpha_j} \to \lambda_0 \text{ in } \mathbb{R}^m, \quad \mu_{\alpha_j} \to \mu_0 \geq 0 \text{ in } \mathbb{R}^p.$$

Then $(\gamma_0)^2 + |\lambda_0|^2 + |\mu_0|^2 = 1$, and consequently

$$(\gamma_0, \lambda_0, \mu_0) \neq (0, 0, 0).$$

Now let $\alpha = \alpha_j \to \infty$ in (4.20), and recall (4.19) to prove (4.12).

5. Note next that $\mathbf{h}_+(x_\alpha) \geq 0$ and therefore $\mu_0 \geq 0$. Furthermore, if for some index j we have

$$h_j(x_0) < 0,$$

then also

$$h_j(x_\alpha) < 0$$

if α is large enough. Consequently $h_j^+(x_\alpha) = 0$, and so $\mu_\alpha^j = 0$ for all large α. Thus $\mu_0^j = 0$, and therefore $\mu_0 \cdot \mathbf{h}(x_0) = 0$. This gives (4.13). □

REMARK. Like the earlier Theorem 1.3.1, this proof is based upon McShane [**MS**]. □

4.2.2. More on constraint qualification.

The John conditions (4.12) and (4.13) are not particularly useful for an abnormal multiplier $\gamma_0 = 0$, since the resulting formula no longer involves the function f we are minimizing.

*We therefore are interested in finding various addition circumstances (usually called **constraint qualifications**) that allow us to conclude that* $\gamma_0 > 0$, in which case we can divide, convert to the case $\gamma_0 = 1$, and thereby establish the KKT conditions.

One circumstance for which we can take $\gamma_0 = 1$ is when x_0 is regular. As before, let us write

$$J = \{j \in \{1, \ldots, p\} \mid h_j(x_0) = 0\}.$$

for the indices of the active inequality constraints at x_0.

DEFINITION. We say that x_0 is **regular** for (MIN**) if the vectors

$$\{\nabla h_j(x_0)\}_{j \in J} \cup \{\nabla g_k(x_0)\}_{k=1}^m \text{ are linearly independent in } \mathbb{R}^n.$$

THEOREM 4.2.2. Suppose that x_0 solves the constrained optimization problem (MIN**) and that x_0 is regular.

Then there exist λ_0 and μ_0 such that the KKT conditions (4.14) and (4.15) hold.

Proof. According to (4.12), we have

$$0 = \gamma_0 \nabla f(x_0) + \nabla \mathbf{g}(x_0)^T \lambda_0 + \nabla \mathbf{h}(x_0)^T \mu_0 \tag{4.21}$$

for appropriate $(\gamma_0, \lambda_0, \mu_0) \neq (0, 0, 0)$. If $\gamma_0 = 0$, then

$$\nabla \mathbf{g}(x_0)^T \lambda_0 + \nabla \mathbf{h}(x_0)^T \mu_0 = 0;$$

and hence $(\lambda_0, \mu_0) = (0, 0)$ since x_0 is regular. But this is a contradiction, and thus $\gamma_0 > 0$. We can consequently divide the equation (4.21) by γ_0, to convert it into the form (4.14). □

Recall that if we drop the equality constraints, our problem is

$$(MIN^*) \qquad \begin{cases} \text{minimize } f(x), \\ \text{subject to } \mathbf{h}(x) \leq 0. \end{cases}$$

DEFINITION. We say that x_0 satisfies the **Fromovitz–Mangasarian constraint qualification condition** for (MIN*) if there exists a vector $p \in \mathbb{R}^n$ such that

$$p \cdot \nabla h_j(x_0) < 0 \quad (j \in J), \tag{4.22}$$

J as usual denoting the indices of the active constraints.

THEOREM 4.2.3. Suppose that x_0 solves the constrained optimization problem (MIN*) and x_0 also satisfies the Fromovitz–Mangasarian condition.

Then there exists μ_0 such that the KKT conditions (4.14) and (4.15) hold (for $\mathbf{g} = 0$).

Proof. Since we have no equality constraints, the John condition (4.12) says

$$0 = \gamma_0 \nabla f(x_0) + \nabla \mathbf{h}(x_0)^T \mu_0 \tag{4.23}$$

for $(\gamma_0, \mu_0) \neq (0, 0)$. So if $\gamma_0 = 0$, then

$$\sum_{j=1}^{p} \mu_0^j \nabla h_j(x_0) = 0.$$

But taking the inner product with p then gives the contradiction

$$0 = \sum_{j=1}^{p} \mu_0^j \nabla h_j(x_0) \cdot p < 0,$$

the last inequality following from the Fromovitz–Mangasarian condition since $\mu_0 \geq 0, \mu_0 \neq 0$. Hence $\gamma_0 > 0$, and we can therefore convert into the KKT form. □

4.3. Quadratic programming

The **quadratic programming program** is finding $x_0 \in \mathbb{R}^n$ to

$$\text{(Q)} \qquad \boxed{\begin{cases} \text{minimize} \ \ \frac{1}{2} x^T C x + c \cdot x \\ \text{subject to } Ax = b, \ x \geq 0, \end{cases}}$$

where C is a symmetric $n \times n$ matrix.

The problem (Q) is of the form (MIN**) for

$$f(x) = \tfrac{1}{2} x^T C x + c \cdot x, \ \mathbf{g} = Ax - b, \ \mathbf{h} = -x.$$

Because the constraints are linear, this problem is normal. Consequently there exist $\lambda_0 \in \mathbb{R}^m, \mu_0 \in \mathbb{R}^n$ such that

$$Cx + c + \nabla(\lambda_0 \cdot (Ax - b)) + \nabla(\mu_0 \cdot (-x)) = 0 \tag{4.24}$$

at x_0 and

$$\mu_0 \geq 0, \quad \mu_0 \cdot x_0 = 0.$$

Therefore

$$Cx_0 + c + A^T\lambda_0 - \mu_0 = 0.$$

Summarizing all the information above, we find that x_0, λ_0, μ_0 satisfy

$$\begin{cases} Ax_0 = b, \ A^T\lambda_0 = -Cx_0 - c + \mu_0 \\ x_0 \geq 0, \ \mu_0 \geq 0 \\ \mu_0 \cdot x_0 = 0. \end{cases} \tag{4.25}$$

REMARK. These are all linear conditions, except for the last line. It turns out that it is therefore possible to modify the simplex algorithm to handle quadratic programming problems: see Franklin [**F1**] for more. □

Application: non-zero sum matrix games. A remarkable application of quadratic programming is to two-person, *non-zero sum* matrix games. For this, we are given two payoff matrices

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mn} \end{bmatrix}.$$

As in our earlier discussion of games (see page 60), Player I selects his mixed strategy from the set

$$P = \left\{ p \in \mathbb{R}^m \mid p_i \geq 0, \sum_{i=1}^{m} p_i = 1 \right\}$$

and Player II selects hers from

$$Q = \left\{ q \in \mathbb{R}^n \mid q_j \geq 0, \sum_{j=1}^{n} q_j = 1 \right\}.$$

The corresponding respective payoffs to Player I and Player II are

$$p \cdot Aq = \sum_{i,j} p_i a_{ij} q_j, \ p \cdot Bq = \sum_{i,j} p_i b_{ij} q_j. \tag{4.26}$$

Each player wants to maximize his/her payoff.

DEFINITION. We say (p_0, q_0) is a mixed-strategy **Nash equilibrium** if

$$\begin{cases} \max_{p \in P} \{p \cdot Aq_0\} = p_0 \cdot Aq_0 \\ \max_{q \in Q} \{p_0 \cdot Bq\} = p_0 \cdot Bq_0. \end{cases} \tag{N}$$

LEMMA 4.3.1. The pair (p_0, q_0) satisfies (N) if and only if

$$(p_0 \cdot Aq_0)e \geq Aq_0 \tag{4.27}$$

and

$$(p_0 \cdot Bq_0)e \geq B^T p_0, \tag{4.28}$$

where $e = [1 \ldots 1]^T \in \mathbb{R}^m$ in (4.27) and $e = [1 \ldots 1]^T \in \mathbb{R}^n$ in (4.28) .

Proof. 1. Assume (p_0, q_0) satisfies (N) and write $p = [0 \ldots 1 \ldots 0]^T$, the 1 in the i-th slot. Then

$$p_0 \cdot Aq_0 \geq p \cdot Aq_0 = (Aq_0)_i \quad (i = 1, \ldots, m).$$

This says $(p_0 \cdot Aq_0)e \geq Aq_0$. The proof that $(p_0 \cdot Bq_0)e \geq B^T p_0$ is similar.

2. Conversely, assume (4.27) and (4.28) are valid. We take the dot product of (4.27) with any $p \in P$, to deduce that $p_0 \cdot Aq_0 \geq p \cdot Aq_0$. This gives the first line of (N) and the second line follows similarly. □

We introduce now the quadratic programming problem of finding (p_0, q_0, x_0, y_0) to

$$\begin{cases} \text{maximize } p \cdot (A+B)q - (x+y), \text{ subject to} \\ Aq \leq xe, \ B^T p \leq ye, \\ p \cdot e = 1, \ q \cdot e = 1, \\ p \geq 0, \ q \geq 0. \end{cases} \tag{4.29}$$

THEOREM 4.3.1. The pair (p_0, q_0) is a Nash equilibrium if and only if (p_0, q_0, x_0, y_0) solves the maximization problem (4.29), where

$$x_0 = p_0 \cdot Aq_0, \ y_0 = p_0 \cdot Bq_0. \tag{4.30}$$

Proof. 1. Suppose first (p_0, q_0) satisfies (N), and define x_0, y_0 by (4.30). According to the Lemma, $Aq_0 \leq x_0 e$ and $B^T p_0 \leq y_0 e$; and consequently (p_0, q_0, x_0, y_0) is feasible for the maximization problem (4.29). Now let (p, q, x, y) be any other feasible solutions for (4.29). Then

$$p \cdot Aq \leq x(e \cdot p) = x, \ q \cdot B^T p \leq y(e \cdot q) = y.$$

We add these inequalities, to learn that

$$p \cdot (A+B)q \leq x + y.$$

But then

$$p \cdot (A+B)q - (x+y) \leq 0 = p_0 \cdot (A+B)q_0 - (x_0 + y_0).$$

Hence (p_0, q_0, x_0, y_0) is optimal for (4.29).

2. Conversely, let (p_0, q_0, x_0, y_0) solve the quadratic maximization problem (4.29). Assume (p_1, q_1) satisfies (N) and define

$$x_1 = p_1 \cdot Aq_1,\ y_1 = p_1 \cdot Bq_1.$$

Then

$$p_1 \cdot (A + B)q_1 - (x_1 + y_1) = 0.$$

Since (p_0, q_0, x_0, y_0) is optimal for (4.29), it must therefore be that

$$0 \leq p_0 \cdot (A + B)q_0 - (x_0 + y_0). \tag{4.31}$$

However, $p_0 \cdot Aq_0 \leq x_0$ and $p_0 \cdot Bq_0 \leq y_0$, and so the right hand side of (4.31) is non-positive. Therefore

$$p_0 \cdot Aq_0 = x_0,\ p_0 \cdot Bq_0 = y_0.$$

Hence the constraints for (4.29) give

$$(p_0 \cdot Aq_0)e \geq Aq_0,\ (p_0 \cdot Bq_0)e \geq B^T p_0.$$

The Lemma now implies that (p_0, q_0) is a Nash equilibrium. □

REMARK. This proof is from Barron [**Ba**] and is based upon ideas in Lemke-Howson [**L-H**]. □

Chapter 5

CONVEX OPTIMIZATION

We now make additional convexity assumptions, which will let us greatly strengthen the theory from the previous chapter.

5.1. Variational inequalities

We start with a simple situation that clearly illustrates how convexity lets us deduce global minimality from local, first variational information.

Let $C \subset \mathbb{R}^n$ be a convex set. We begin with the basic optimization problem of finding x_0 to

(C) $$\boxed{\text{minimize } f(x), \text{ subject to } x \in C.}$$

THEOREM 5.1.1. (i) If $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and x_0 solves (C), then

(VI) $$\boxed{\nabla f(x_0) \cdot (x - x_0) \geq 0 \quad \text{for all } x \in C.}$$

(ii) Suppose also that f is convex. Then if $x_0 \in C$ satisfies (VI), x_0 solves the minimization problem (C).

.

REMARK. We call (VI) a **variational inequality**, which is a form of first variation for the constrained optimization problem (C).

Note carefully that assertion (ii) provides us with a first-order *sufficient condition* that x_0 be a solution of (C). □

Proof. 1. Let $x \in C$. Then $x_0+t(x-x_0) = tx+(1-t)x_0 \in C$ for $0 \le t \le 1$. Thus if x_0 solves (C), it follows that

$$\phi(t) = f(x_0 + t(x - x_0))$$

has its minimum for $0 \le t \le 1$ at $t = 0$. Therefore $\phi'(0) \ge 0$. But $\phi'(t) = \nabla f(x_0 + t(x - x_0)) \cdot (x - x_0)$ and hence

$$\phi'(0) = \nabla f(x_0) \cdot (x - x_0) \ge 0.$$

2. If f is convex and differentiable at x_0, then

$$f(x) \ge f(x_0) + \nabla f(x_0) \cdot (x - x_0)$$

for all $x \in \mathbb{R}^n$. Since $\nabla f(x_0) \cdot (x - x_0) \ge 0$ if $x \in C$ according to (VI), we have

$$f(x) \ge f(x_0)$$

for all $x \in C$. So x_0 solves (C). □

INTERPRETATION. Note that *Lagrange multipliers do not appear in (VI)*, in spite of the general principle that "constraints cause Lagrange multipliers to appear." This is sometimes an advantage, since it may be simpler to study directly the inequalities in (VI). □

EXAMPLE. We can sometimes deduce the existence of Lagrange multipliers directly from the variational inequality. Consider for example the problem

$$\begin{cases} \text{minimize } f(x), \\ \text{subject to } Ax = b, \end{cases}$$

where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$.

If x_0 is a minimizer, then (VI) says $\nabla f(x_0) \cdot (x - x_0) \ge 0$ for all x such that $Ax = b$. Any such x has the form $x = x_0 + w$ for $w \in N(A)$. Consequently, $\nabla f(x_0) \cdot w \ge 0$ for all $w \in N(A)$). Replacing w by $-w$, we see that in fact

$$\nabla f(x_0) \cdot w = 0 \qquad (w \in N(A)).$$

Therefore

$$\nabla f(x_0) \in N(A)^\perp = R(A^T),$$

the second equality following from Theorem 3.1.6. So we can write

$$\nabla f(x_0) + A^T \lambda_0 = 0$$

for some $\lambda_0 \in \mathbb{R}^m$. Then $\lambda_0 = [\lambda_0^1 \cdots \lambda_0^m]^T$ is the vector of Lagrange multipliers for the constraint $Ax = b$ (in accordance with Theorem 4.2.1). □

5.2. Convexity and Lagrange multipliers

The general theory in Chapter 4 provides the existence of Lagrange multipliers for the problem of finding $x_0 \in \mathbb{R}^n$ to

$$(MIN^*) \qquad \begin{cases} \text{minimize } f(x), \\ \text{subject to } \mathbf{h}(x) \leq 0, \end{cases}$$

with no convexity assumptions. However, the various constraint qualification conditions introduced in Chapter 4 are often difficult to check in practice. This section presents an alternative approach, under the additional (and strong) assumption that

$$f, h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R} \text{ are convex functions.} \tag{5.1}$$

5.2.1. Sufficient condition for minimality. First we show that for convex functions, the KKT conditions are sufficient for optimality:

THEOREM 5.2.1. Assume $f, h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$ are convex. Suppose also that

$$\mathbf{h}(x_0) \leq 0,$$

and that there exists $\mu_0 \in \mathbb{R}^p$ such that the KKT conditions hold:

$$\nabla f(x_0) + \sum_{j=1}^{p} \mu_0^j \nabla h_j(x_0) = 0, \tag{5.2}$$

$$\mu_0 \geq 0, \quad \mu_0 \cdot \mathbf{h}(x_0) = 0. \tag{5.3}$$

Then x_0 solves (MIN*).

Proof. Let $C = \{x \in \mathbb{R}^n \mid \mathbf{h}(x) \leq 0\}$ denote the feasible set, which is convex since the functions $h_1, \ldots, h_p$ are convex.

Now (5.2) implies

$$\nabla f(x_0) \cdot (x - x_0) + \sum_{j=1}^{p} \mu_0^j \nabla h_j(x_0) \cdot (x - x_0) = 0 \tag{5.4}$$

for all x. Furthermore, if $x \in C$, we have

$$h_j(x_0) + \nabla h_j(x_0) \cdot (x - x_0) \leq h_j(x) \leq 0.$$

We multiply by $\mu_0^j \geq 0$ and sum, to deduce from (5.3) that

$$\sum_{j=1}^{p} \mu_0^j \nabla h_j(x_0) \cdot (x - x_0) \leq 0.$$

Then (5.4) implies $\nabla f(x_0) \cdot (x - x_0) \geq 0$ for each $x \in C$. Consequently, the convexity of f implies

$$f(x) \geq f(x_0) + \nabla f(x_0) \cdot (x - x_0) \geq f(x_0),$$

and hence x_0 is optimal for (MIN*). □

5.2.2. Slater's condition.

We introduce next a new constraint qualification condition:

DEFINITION. We say that **Slater's condition** for (MIN*) holds provided

(S) $\boxed{\text{there exists a point } \bar{x} \in \mathbb{R}^n \text{ such that } \mathbf{h}(\bar{x}) < 0.}$

THEOREM 5.2.2. Let x_0 solve (MIN*) and that $f, h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$ are convex. Assume further that Slater's condition (S) holds.

(i) Then there exists $\mu_0 \in \mathbb{R}^p$ such that the KKT conditions (5.2) and (5.2) hold.

(ii) Furthermore,

(5.5) $$x \mapsto f(x) + \mu_0 \cdot \mathbf{h}(x) \text{ has a minimum at } x_0.$$

Proof. F. John's formulation from Theorem 4.2.1 tells us that there exist nonnegative numbers $\gamma_0, \mu_0^1, \ldots, \mu_0^p$, not all equal to zero, such that

(5.6) $$\gamma_0 \nabla f(x_0) + \sum_{j=1}^{p} \mu_0^j \nabla h_j(x_0) = 0$$

and

(5.7) $$\gamma_0 \geq 0,\ \mu_0 \geq 0,\ \mu_0 \cdot \mathbf{h}(x_0) = 0.$$

We claim that

(5.8) $$\gamma_0 \neq 0.$$

To see this, suppose otherwise. Then $\mu_0 \neq 0$ and

$$\sum_{j=1}^{p} \mu_0^j \nabla h_j(x_0) = 0.$$

Hence

$$\sum_{j=1}^{p} \mu_0^j \nabla h_j(x_0)(\bar{x} - x_0) = 0,$$

where $\bar{x}$ satisfies (S). But convexity implies for $j = 1, \ldots, p$ that

$$h_j(x_0) + \nabla h_j(x_0) \cdot (\bar{x} - x_0) \leq h_j(\bar{x}).$$

Therefore

$$\mu_0 \cdot \mathbf{h}(x_0) \leq \mu_0 \cdot \mathbf{h}(\bar{x}) < 0$$

according to Slater's condition, since $\mu_0 \geq 0$, $\mu_0 \neq 0$. This contradicts (5.7), and thereby proves (5.8).

We as usual now divide (5.6) by $\gamma_0 > 0$ to establish the KKT statements, with possibly new constants $\mu_0^1, \ldots, \mu_0^p$. The assertion (5.5) follows from convexity. □

5.2.3. Value functions.

This section provides another interpretation of Slater's condition and its relation to the value function, defined as follows:

DEFINITION. The **value function** associated with (MIN*) is defined for $a \in \mathbb{R}^p$ by

(5.9)

$$v(a) = \begin{cases} \inf_{x \in \mathbb{R}^n} \{f(x) \mid \mathbf{h}(x) \leq a\} & \text{if } \{x \in \mathbb{R}^n \mid \mathbf{h}(x) \leq a\} \neq \emptyset \\ \infty & \text{if } \{x \in \mathbb{R}^n \mid \mathbf{h}(x) \leq a\} = \emptyset. \end{cases}$$

Henceforth, we assume the superlinear growth condition

$$\lim_{|x| \to \infty} \frac{f(x)}{|x|} = \infty; \tag{5.10}$$

so that in particular

$$\inf_{a \in \mathbb{R}^n} v(a) > -\infty. \tag{5.11}$$

LEMMA 5.2.1. Assume that the functions $f, h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$ are convex. Then the value function

$$v : \mathbb{R}^p \to (-\infty, \infty]$$

is convex and lower semicontinuous.

Proof. 1. Select $a, \hat{a} \in \mathbb{R}^p$ and $0 < \theta < 1$. We will show that

$$v(\theta a + (1-\theta)\hat{a}) \leq \theta v(a) + (1-\theta) v(\hat{a}).$$

This holds if either $v(a) = \infty$ or $v(\hat{a}) = \infty$; and so we may assume $v(a), v(\hat{a}) < \infty$. Select $\varepsilon > 0$ and then choose $x, \hat{x} \in \mathbb{R}^n$ so that

$$\mathbf{h}(x) \leq a, \ \mathbf{h}(\hat{x}) \leq \hat{a}$$

and

$$f(x) \leq v(a) + \varepsilon, \ f(\hat{x}) \leq v(\hat{a}) + \varepsilon.$$

Then, since f is convex,

$$f(\theta x + (1-\theta)\hat{x}) \leq \theta f(x) + (1-\theta) f(\hat{x}) \leq \theta v(a) + (1-\theta) v(\hat{a}) + \varepsilon.$$

Since

$$\mathbf{h}(\theta x + (1-\theta)\hat{x}) \le \theta \mathbf{h}(x) + (1-\theta)\mathbf{h}(\hat{x}) \le \theta a + (1-\theta)\hat{a},$$

it follows that

$$v(\theta a + (1-\theta)\hat{a}) \le \theta v(a) + (1-\theta)v(\hat{a}) + \varepsilon.$$

Let $\varepsilon \to 0$. This proves v is convex.

2. To show v is lower semicontinuous, assume $\lim_{k\to\infty} a_k = a^0$, with $\liminf_{k\to\infty} v(a_k) < \infty$. Fix $\varepsilon > 0$ and then select points x^k such that

$$\mathbf{h}(x^k) \le a_k, \ f(x^k) \le v(a_k) + \varepsilon \qquad (k = 1, \dots). \tag{5.12}$$

Using the superlinear growth condition (5.10) and the Bolzano-Weierstrass Theorem, we can extract a subsequence such that the limit

$$\lim_{j\to\infty} x^{k_j} = x^0$$

exists and $\lim_{j\to\infty} v(a^{k_j}) = \liminf_{k\to\infty} v(a_k)$. Then (5.12) gives

$$\mathbf{h}(x^0) \le a, \ f(x^0) \le \liminf_{k\to\infty} v(a_k) + \varepsilon.$$

Hence $v(a) \le \liminf_{k\to\infty} v(a_k) + \varepsilon$ for each $\varepsilon > 0$. □

The key point is now to understand when the subdifferential of v at 0 is nonempty.

THEOREM 5.2.3. Let x_0 solve (MIN*) and that $f, h_1, \dots, h_p : \mathbb{R}^n \to \mathbb{R}$ are convex. Assume further that Slater's condition (S) holds.

(i) Then

$$\partial v(0) \neq \emptyset.$$

(ii) Consequently, there exists $\mu_0 \in \mathbb{R}^p$ such that the KKT conditions (5.2) and (5.3) are valid.

INTERPRETATION. If the value function v is differentiable at 0, the proof below shows that the Lagrange multiplier is

$$\boxed{\mu_0 = -\nabla v(0).}$$

So *the Lagrange multiplier for the constraint* $\mathbf{h} \le 0$ *is the negative of the gradient of the value function at* $a = 0$. (Recall our earlier introduction of the value function on page 15.)

In general $\nabla v(0)$ need not exist, but Slater's condition ensures that $\partial v(0) \neq \emptyset$. This discussion and the following proof are based upon Borwein-Lewis [**B-L**]. □

Proof. 1. In view of (S), the sets $\{x \in \mathbb{R}^n \mid \mathbf{h}(x) \leq a\}$ are nonempty for all $a \in \mathbb{R}^m$ with $|a|$ sufficiently small. This implies that v does not take on the value ∞ in some ball centered at 0. Hence

$$\partial v(0) \neq \emptyset$$

according to the Remark on page 105.

2. Select

$$-\mu_0 \in \partial v(0).$$

Then

$$v(a) \geq v(0) - \mu_0 \cdot a = f(x_0) - \mu_0 \cdot a \quad (a \in \mathbb{R}^p). \tag{5.13}$$

Now if $a \geq 0$, then clearly $\mathbf{h}(x_0) \leq 0 \leq a$. Thus

$$f(x_0) \geq v(a),$$

and hence (5.13) implies

$$f(x_0) \geq f(x_0) - \mu_0 \cdot a.$$

This inequality is valid for all $a \geq 0$; whence $\mu_0 \geq 0$.

3. Observe furthermore from the definition of the value function that

$$f(x) \geq v(\mathbf{h}(x)) \quad (x \in \mathbb{R}^n).$$

Consequently (5.13) implies

$$f(x) \geq v(\mathbf{h}(x)) \geq f(x_0) - \mu_0 \cdot \mathbf{h}(x).$$

Therefore

$$f(x) + \mu_0 \cdot \mathbf{h}(x) \geq f(x_0) \geq f(x_0) + \mu_0 \cdot \mathbf{h}(x_0) \quad (x \in \mathbb{R}^n), \tag{5.14}$$

since $\mu_0 \cdot \mathbf{h}(x_0) \leq 0$. So the function $x \mapsto f(x) + \mu_0 \cdot \mathbf{h}(x)$ has an unconstrained minimum at x_0 and thus

$$\nabla f(x_0) + \sum_{j=1}^{m} \mu_0^j \nabla h_j(x_0) = 0.$$

Finally, putting $x = x_0$ in (5.14) shows that $\mu_0 \cdot \mathbf{h}(x_0) = 0$. □

5.3. Convex duality I

In this section we reinterpret and extend the theory above to provide a duality theory for convex optimization problems. Our references are Boyd–Vandenberghe [**B-V**] and Calafiore–El Ghaoui [**C-EG**].

5.3.1. Dual problems.

NOTATION. Given $f, g_1, \ldots, g_m, h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$, we as usual write

$$\mathbf{g} = \begin{bmatrix} g_1 \\ \vdots \\ g_m \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} h_1 \\ \vdots \\ h_p \end{bmatrix}.$$

We return to our general minimization problem, with both inequality and equality constraints, which we for this section write as finding $x_0 \in \mathbb{R}^n$ to

$$\text{(P)} \qquad \boxed{\begin{cases} \text{minimize } f(x), \text{ subject to} \\ \mathbf{g}(x) = 0, \ \mathbf{h}(x) \leq 0. \end{cases}}$$

DEFINITION. The associated **Lagrangian function** is

$$\boxed{L(x, \lambda, \mu) = f(x) + \lambda \cdot \mathbf{g}(x) + \mu \cdot \mathbf{h}(x)}$$

for $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}^p$. That is,

$$L(x, \lambda, \mu) = f(x) + \sum_{k=1}^{m} \lambda_k g_k(x) + \sum_{j=1}^{p} \mu_j h_j(x).$$

DEFINITION. The corresponding **dual function** is

$$\boxed{g(\lambda, \mu) = \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu) = \inf_{x \in \mathbb{R}^n} \{f(x) + \lambda \cdot \mathbf{g}(x) + \mu \cdot \mathbf{h}(x)\}}$$

EXAMPLE. Assume f is convex, and consider the problem

$$\text{(P)} \qquad \begin{cases} \text{minimize } f(x), \text{ subject to} \\ Ax = b, x \geq 0. \end{cases}$$

So $\mathbf{h}(x) = -x$, $\mathbf{g}(x) = Ax - b$. The dual function is

$$\begin{aligned} g(\lambda, \mu) &= \min_x \{f(x) + \lambda \cdot \mathbf{g}(x) + \mu \cdot \mathbf{h}(x)\} \\ &= \min_x \{f(x) + \lambda \cdot (Ax - b) + \mu \cdot (-x)\} \\ &= -\lambda \cdot b + \min_x \left\{f(x) + (A^T\lambda - \mu) \cdot x\right\} \\ &= -\lambda \cdot b - \max_x \left\{(\mu - A^T\lambda) \cdot x - f(x)\right\}. \end{aligned}$$

Therefore

$$g(\lambda, \mu) = -\lambda \cdot b - f^*(\mu - A^T\lambda)$$

where $f^*(y) = \max_x\{x \cdot y - f(x)\}$ is the dual convex function to f. □

DEFINITION. The **dual problem** to (P) is to find $\lambda_0 \in \mathbb{R}^m, \mu_0 \in \mathbb{R}^p$ to

$$\text{(D)} \qquad \boxed{\begin{cases}\text{maximize } g(\lambda,\mu),\\ \text{subject to } \mu \geq 0.\end{cases}}$$

EXAMPLE (**Linear programming yet again**)**.** If we set $f(x) = c \cdot x$ in our previous example, we once again encounter the canonical linear programming progrem

$$\text{(P)} \qquad \begin{cases}\text{minimize } x \cdot c, \text{ subject to}\\ Ax = b, x \geq 0.\end{cases}$$

Now

$$f^*(y) = \max_x\{(y-c)\cdot x\} = \begin{cases}0 & \text{if } y = c\\ \infty & \text{otherwise;}\end{cases}$$

therefore

$$g(\lambda,\mu) = \begin{cases}-\lambda\cdot b & \text{if } \mu - A^T\lambda = c\\ -\infty & \text{otherwise.}\end{cases} .$$

The dual problem is then

$$\text{(D)} \qquad \begin{cases}\text{maximize } -\lambda\cdot b, \text{ subject to}\\ \mu - A^T\lambda = c,\ \mu \geq 0.\end{cases}$$

We write $y = -\lambda$ to recover the usual dual canonical problem.

This example provides a belated answer to a question left over from Chapter 2: where did the linear programming dual problem come from? □

THEOREM 5.3.1. We have the "weak duality" inequality

$$\text{(5.15)} \qquad \sup_{\substack{\mu,\lambda\\ \mu\geq 0}} g(\lambda,\mu) \leq \inf_{\substack{\mathbf{g}(x)=0\\ \mathbf{h}(x)\leq 0}} f(x).$$

If this is a strict inequality, we say there is a **duality gap**.

Proof. Assume $\mu \geq 0$ and $\mathbf{g}(x) = 0$, $\mathbf{h}(x) \leq 0$. Then

$$L(x,\lambda,\mu) = f(x) + \underbrace{\mu}_{\geq 0}\cdot\underbrace{\mathbf{h}(x)}_{\leq 0} + \lambda\cdot\underbrace{\mathbf{g}(x)}_{=0} \leq f(x).$$

Since

$$g(\lambda,\mu) = \inf_{z\in\mathbb{R}^n} L(z,\lambda,\mu) \leq L(x,\lambda,\mu),$$

we deduce that $g(\lambda,\mu) \leq f(x)$. This implies (5.15) . □

5.3.2. Slater's condition again.

Hereafter we will need to consider linear equality constraints, and so assume that

$$\boxed{\mathbf{g}(x) = Ax - b,}$$

where A is an $m \times n$ matrix and $m \leq n$. We will suppose also the nondegeneracy condition that

$$\text{the rank of } A \text{ is } m. \tag{5.16}$$

(Recall from the remark on page 19 that we can always reduce to this situation.)

Thus our primal problem becomes finding $x_0 \in \mathbb{R}^n$ to

$$(\widehat{\mathrm{P}}) \qquad \boxed{\begin{cases} \text{minimize } f(x), \text{ subject to} \\ Ax = b, \ \mathbf{h}(x) \leq 0. \end{cases}}$$

and the dual problem is finding $\mu_0 \in \mathbb{R}^m, \lambda_0 \in \mathbb{R}^p$ to

$$(\widehat{\mathrm{D}}) \qquad \boxed{\begin{cases} \text{maximize } g(\lambda, \mu), \\ \text{subject to } \mu \geq 0, \end{cases}}$$

where $g(\lambda, \mu) = \inf_x L(x, \lambda, \mu)$ for the Lagrangian function

$$L(x, \lambda, \mu) = f(x) + \lambda \cdot (Ax - b) + \mu \cdot \mathbf{h}(x).$$

THEOREM 5.3.2. Assume $f, h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$ are convex. Suppose also this modified Slater's condition holds:

$$(\widehat{\mathrm{S}}) \qquad \boxed{\text{there exists a point } \bar{x} \in \mathbb{R}^n \text{ with } \mathbf{h}(\bar{x}) < 0, \ A\bar{x} = b.}$$

(i) Then if x_0 solves $(\widehat{\mathrm{P}})$, there exists a pair (λ_0, μ_0) solving $(\widehat{\mathrm{D}})$.

(ii) Furthermore,

$$g(\lambda_0, \mu_0) = f(x_0). \tag{5.17}$$

Hence we have strong duality:

$$\boxed{\max_{\substack{\mu,\lambda \\ \mu \geq 0}} g(\lambda, \mu) = \min_{\substack{\mathbf{g}(x)=0 \\ \mathbf{h}(x) \leq 0}} f(x).} \tag{5.18}$$

Proof. 1. Theorem 4.2.1 provides us with $(\gamma_0, \lambda_0, \mu_0) \neq (0, 0, 0)$ satisfying the F. John conditions that

$$\gamma_0 \nabla f(x_0) + A^T \lambda_0 + \nabla \mathbf{h}(x_0)^T \mu_0 = 0$$

and

$$\gamma_0 \geq 0, \quad \mu_0 \geq 0, \quad \mu_0 \cdot \mathbf{h}(x_0) = 0.$$

We once again claim that

$$\gamma_0 \neq 0. \tag{5.19}$$

To see this, assume otherwise. Then

$$A^T\lambda_0 + \nabla \mathbf{h}(x_0)^T\mu_0 = 0. \tag{5.20}$$

2. Take the inner product with $\bar{x} - x_0$, where $\bar{x}$ satisfies $(\widehat{\mathrm{S}})$, to deduce

$$\begin{aligned} 0 &= A^T\lambda_0 \cdot (\bar{x} - x_0) + \nabla \mathbf{h}(x_0)^T\mu_0 \cdot (\bar{x} - x_0) \\ &= \lambda_0 \cdot (A\bar{x} - Ax_0) + \mu_0 \cdot \nabla \mathbf{h}(x_0)(\bar{x} - x_0). \end{aligned}$$

But $A\bar{x} - Ax_0 = b - b = 0$, and so

$$\mu_0 \cdot \nabla \mathbf{h}(x_0)(\bar{x} - x_0) = 0.$$

Furthermore, convexity implies for $j = 1, \dots, p$ that

$$h_j(x_0) + \nabla h_j(x_0) \cdot (\bar{x} - x_0) \leq h_j(\bar{x}).$$

Therefore

$$\mu_0 \cdot \mathbf{h}(x_0) \leq \mu_0 \cdot \mathbf{h}(\bar{x}).$$

Now if $\mu_0 \neq 0$, our modified Slater's condition $(\widehat{\mathrm{S}})$ implies $\mu_0 \cdot \mathbf{h}(\bar{x}) < 0$ and this gives the contradiction $\mu_0 \cdot \mathbf{h}(x_0) < 0$.

3. Hence $\mu_0 = 0$ and so (5.20) becomes

$$A^T\lambda_0 = 0.$$

But recalling our nondegeneracy assumption (5.16), we see that therefore $\lambda_0 = 0$.

In summary, we have shown that if $\gamma_0 = 0$, then also $\mu_0 = \lambda_0 = 0$. However this is impossible, since not all of these equal zero.

4. As usual, (5.19) implies that we have the KKT condition

$$\nabla f(x_0) + A^T\lambda_0 + \nabla \mathbf{h}(x_0)^T\mu_0 = 0, \tag{5.21}$$

for possibly new choices of μ_0, λ_0. We next demonstrate that this identity implies (5.17). To do so, we use the convexity of f and $\mathbf{h}$ to compute for any $x \in \mathbb{R}^n$ that

$$\begin{aligned} f(x_0) &= f(x_0) + \lambda_0 \cdot (Ax_0 - b) + \mu_0 \cdot \mathbf{h}(x_0) \\ &\leq f(x) + \nabla f(x_0) \cdot (x_0 - x) \\ &\qquad + \lambda_0 \cdot [(Ax_0 - Ax) + (Ax - b)] \\ &\qquad + \mu_0 \cdot [\mathbf{h}(x) + \nabla \mathbf{h}(x_0)(x_0 - x)] \\ &= f(x) + \lambda_0 \cdot (Ax - b) + \mu_0 \cdot \mathbf{h}(x) \end{aligned}$$

$$+ [\nabla f(x_0) + A^T\lambda_0 + \nabla \mathbf{h}(x_0)^T\mu_0] \cdot (x_0 - x)$$
$$= f(x) + \lambda_0 \cdot (Ax - b) + \mu_0 \cdot \mathbf{h}(x),$$

according to (5.21). Taking the infimum over x, we deduce that therefore

$$f(x_0) \leq g(\lambda_0, \mu_0).$$

The opposite inequality follows from (5.15). □

REMARK. A lesson from this proof is that *when we have convexity, the existence of the KKT Lagrange multipliers implies duality*, meaning that that there is no duality gap between the primal and dual problems. □

5.4. Convex duality II

We explain next an alternate approach to convex duality, following Borwein–Lewis [**B-L**].

5.4.1. Fenchel duality.

Suppose that we are given two convex, proper, lower semicontinuous functions

$$f : \mathbb{R}^n \to (-\infty, \infty], \quad g : \mathbb{R}^m \to (-\infty, \infty]$$

and an $m \times n$ matrix A.

DEFINITION. The **domain** of f is

$$\operatorname{dom} f = \{x \in \mathbb{R}^n \mid f(x) < \infty\},$$

and the domain of g is

$$\operatorname{dom} g = \{y \in \mathbb{R}^m \mid g(y) < \infty\}.$$

□

Now if $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, we have

$$f(x) + f^*(A^T y) \geq x \cdot A^T y$$

and

$$g(Ax) + g^*(-y) \geq -Ax \cdot y.$$

Adding these inequalities and rewriting, we see that

$$f(x) + g(Ax) \geq -f^*(A^T y) - g^*(-y) \qquad (x \in \mathbb{R}^n, y \in \mathbb{R}^m). \tag{5.22}$$

This suggests that we introduce the **primal problem**

$$\boxed{\text{minimize } f(x) + g(Ax)} \qquad (x \in \mathbb{R}^n) \tag{P}$$

and the **dual problem**

$$\text{(D)} \qquad \boxed{\text{maximize } -f^*(A^T y) - g^*(-y)} \qquad (y \in \mathbb{R}^m).$$

Then (5.22) establishes weak duality:

THEOREM 5.4.1. We have

$$\sup_{y \in \mathbb{R}^m} \left\{ -f^*(A^T y) - g^*(-y) \right\} \leq \inf_{x \in \mathbb{R}^n} \left\{ f(x) + g(Ax) \right\}. \tag{5.23}$$

Notice that we allow for the possibility that f, g and their dual functions take the value ∞.

When do we have equality in (5.23)? The idea, as in Section 5.2.3 above, will be to introduce a value function and to study its subdifferential.

DEFINITION. The **value function** is

$$\boxed{v(a) = \inf_{x \in \mathbb{R}^n} \left\{ f(x) + g(Ax + a) \right\}}$$

for $a \in \mathbb{R}^m$.

We assume v never takes on the value $-\infty$.

LEMMA 5.4.1. The value function $v : \mathbb{R}^m \to (-\infty, \infty]$ is convex and lower semicontinuous.

DEFINITION. We say that f, g and A satisfy the **duality condition** if

$$\textit{the value function } v \textit{ does not take on the value } \infty \textit{ near } 0. \tag{5.24}$$

REMARK. This means that there exists a small number $\delta > 0$ such that if $a \in \mathbb{R}^m$ and $|a| \leq \delta$, then we can write

$$a = y - Ax \qquad \text{for some } y \in \operatorname{dom} g, x \in \operatorname{dom} f. \tag{5.25}$$

Thus

$$v(a) \leq f(x) + g(Ax + a) = f(x) + g(y) < \infty.$$

□

THEOREM 5.4.2. Assume that $x_0 \in \mathbb{R}^n$ solves the primal minimization problem (P), and suppose also that f, g, A satisfy the duality condition (5.24).

(i) Then there exists $y_0 \in \mathbb{R}^m$ such that

$$f(x_0) + g(Ax_0) = -f^*(A^T y_0) - g^*(-y_0). \tag{5.26}$$

(ii) We thus have strong duality:

$$\boxed{\min_{x\in\mathbb{R}^n}\{f(x)+g(Ax)\}=\max_{y\in\mathbb{R}^m}\left\{-f^*(A^Ty)-g^*(-y)\right\}.} \tag{5.27}$$

Proof. As noted above, the duality condition (5.24) implies that the value function v is finite-valued near zero and consequently $\partial v(0)\neq\emptyset$. Select

$$-y_0\in\partial v(0).$$

Then for any $x\in\mathbb{R}^n$ and $a\in\mathbb{R}^m$ we have

$$\begin{aligned} f(x_0)+g(Ax_0) &= v(0)\\ &\leq v(a)+a\cdot y_0\\ &\leq f(x)+g(Ax+a)+a\cdot y_0\\ &= -(A^Ty_0\cdot x-f(x))-(-y_0\cdot(Ax+a)-g(Ax+a)). \end{aligned}$$

Take the infimum over a:

$$f(x_0)+g(Ax_0)\leq -(A^Ty_0\cdot x-f(x))-g^*(-y_0);$$

and then take the infimum over x:

$$f(x_0)+g(Ax_0)\leq -f^*(A^Ty_0)-g^*(-y_0).$$

The reverse inequality follows from (5.23). □

5.4.2. Semidefinite programming. A natural generalization of standard linear programming to symmetric matrices is called **semidefinite programming**.

NOTATION. (i) We will write $\mathbb{S}^n$ to denote the linear space of $n\times n$ symmetric matrices and

$$\mathbb{S}^n_+=\{A\in\mathbb{S}^n\mid A\succeq 0\}$$

to denote the nonnegative definite symmetric matrices.

(ii) The inner product of two matrices $A,B\in\mathbb{S}^n$ is

$$A\cdot B=\sum_{i,j=1}^n a_{ij}b_{ij}.$$

□

Given now matrices $\{A_k\}_{k=1}^m, C$ in $\mathbb{S}^n$ and $b \in \mathbb{R}^m$, we introduce the **semidefinite primal problem**

$$\text{(P)} \qquad \boxed{\begin{cases} \text{minimize } C \cdot X, \quad \text{subject to} \\ A_1 \cdot X = b_1, \cdots, A_m \cdot X = b_m \\ X \succeq 0. \end{cases}}$$

This is a matrix version of the canonical linear programming problem, for symmetric matrices X.

Duality. What is the dual problem? To figure this out, set

$$AX = [A_1 \cdot X, \cdots, A_m \cdot X]^T \tag{5.28}$$

and rewrite (P) as minimizing $f(X) + g(AX)$ for

$$f(X) = \begin{cases} C \cdot X & \text{if } X \succeq 0 \\ \infty & \text{otherwise} \end{cases}$$

and

$$g(y) = \begin{cases} 0 & \text{if } y = b \\ \infty & \text{otherwise.} \end{cases}$$

Then $g^*(z) = z \cdot b$. Furthermore, if $W \in \mathbb{S}^n$ we have

$$\begin{aligned} f^*(W) &= \sup_{X \in \mathbb{S}^n} \{W \cdot X - f(X)\} \\ &= \sup_{X \succeq 0} \{(W - C) \cdot X\} \\ &= \begin{cases} 0 & \text{if } W \preceq C \\ \infty & \text{otherwise.} \end{cases} \end{aligned}$$

Note carefully: the matrix inequality "$W \preceq C$" means that $C - W \in \mathbb{S}^n_+$. We can also calculate that

$$A^T y = \sum_{k=1}^m y_k A_k.$$

Therefore the Finchel dual problem to maximize $-f^*(A^T y) - g^*(-y)$ becomes the **semidefinite dual problem**

$$\text{(D)} \qquad \boxed{\begin{cases} \text{maximize } y \cdot b, \\ \text{subject to } \sum_{k=1}^m y_k A_k \preceq C. \end{cases}}$$

This is an analog of the linear programming canonical dual problem, except that the constraint is now an inequality for symmetric matrices. Note

carefully that in (P) the unknown is a symmetric matrix X, but in (D) the unknown is a vector $y \in \mathbb{R}^m$.

THEOREM 5.4.3. (**Duality for semidefinite programming**) Suppose that there exists a feasible matrix $\bar{X}$ for (P) with

$$\bar{X} > 0. \tag{5.29}$$

Assume also that $A : \mathbb{S}^n \to \mathbb{R}^m$ is onto.

Then if $X_0 \in \mathbb{S}^n_+$ solves the primal semidefinite programming problem (P), there exists $y_0 \in \mathbb{R}^m$ that is feasible for (D) and satisfies

$$C \cdot X_0 = b \cdot y_0.$$

Proof. In view of (5.29), there exists $\lambda > 0$ so small that

$$X = \bar{X} + Y \in \mathbb{S}^n_+$$

if $Y \in B(0, \lambda)$. Since $A : \mathbb{S}^n \to \mathbb{R}^m$ is onto, there further exists $\eta > 0$ such that $A : B(0,1) \supseteq B(0,\eta)$ and therefore $A : B(0,\lambda) \supseteq B(0,\lambda\eta)$. Hence for each $a \in B(0,\lambda\eta)$, there exists X as above such that $X \in \mathbb{S}^n_+$ and $AX = A\bar{X} + AY = b - a$.

For f, g as above, we have $\operatorname{dom} f = \mathbb{S}^n_+$, $\operatorname{dom} g = \{b\}$. Then if $\delta = \lambda\eta$ and $|a| \leq \delta$, we can write

$$a = b - AX$$

as required by (5.25). Consequently, the duality condition (5.24) holds; and the rest now follows from Theorem 5.4.2. □

EXAMPLE. The full theory for linear programming does not apply for semidefinite programming. For instance, let $n = 2$, $m = 1$ and

$$b = 0, \quad C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Then (P) has the optimal solution $X = 0$. However for all $y \in \mathbb{R}$

$$C - yA = \begin{bmatrix} -y & 1 \\ 1 & 0 \end{bmatrix} \notin \mathbb{S}^n_+.$$

Hence there are no feasible solutions for (D). □

5.5. Minimax and duality

In this final section we discuss the connections between convex duality theory and minimax conditions, which we encountered earlier in our discussion of two-person, zero-sum game theory (see page 58).

REMARK. Returning to the duality theory set forth above in Section 5.3.1, we recall that the primal problem (P) there is to solve

$$\begin{aligned}\min_{\substack{\mathbf{g}(x)=0\\ \mathbf{h}(x)\le 0}} f(x) &= \min_x \max_{\substack{\mu,\lambda\\ \mu\ge 0}} \{f(x) + \lambda\cdot\mathbf{g}(x) + \mu\cdot\mathbf{h}(x)\}\\ &= \min_x \max_{\substack{\mu,\lambda\\ \mu\ge 0}} L(x,\lambda,\mu);\end{aligned}$$

and the dual problem (D) is to find

$$\max_{\substack{\mu,\lambda\\ \mu\ge 0}} g(\lambda,\mu) = \max_{\substack{\mu,\lambda\\ \mu\ge 0}} \min_x L(x,\lambda,\mu).$$

Writing the problems this way, we see that strong duality holds **if** we can exchange the min and max in these formulas. This demonstrates that convex duality theory can be written in terms of minimax conditions. □

Continuing this theme, we prove next a general minimax theorem. We henceforth assume that $X \subset \mathbb{R}^n, Y \subset \mathbb{R}^m$ are closed, bounded, convex sets and we are given a continuous function

$$L : X \times Y \to \mathbb{R}.$$

We write $L = L(x,y)$.

DEFINITION. A **saddle point** $(x_0,y_0) \in X\times Y$ for L is a point for which

$$L(x_0,y) \le L(x_0,y_0) \le L(x,y_0) \quad \text{for all } x\in X, y\in Y. \tag{5.30}$$

REMARK. If L has a saddle point, we have strong duality between the problems

$$(P)\ \begin{cases}\min\ f(x)\\ \text{for } x\in X\end{cases} \qquad (D)\ \begin{cases}\max\ g(y)\\ \text{for } y\in Y,\end{cases}$$

where

$$f(x) = \max_{y\in Y} L(x,y), \qquad g(y) = \min_{x\in X} L(x,y).$$

□

THEOREM 5.5.1. (**Minimax Theorem**) Assume that $X \subset \mathbb{R}^n, Y\subset\mathbb{R}^m$ are non-empty, closed, convex and bounded sets.

Suppose also that

$$\begin{cases} x \mapsto L(x,y) \text{ is convex for each } y\in Y,\\ y\mapsto L(x,y) \text{ is concave for each } x\in X.\end{cases} \tag{5.31}$$

Then L has a saddle point $(x_0, y_0) \in X \times Y$, and we therefore have the minimax condition

$$\boxed{\min_{x\in X}\max_{y\in Y} L(x,y) = \max_{y\in Y}\min_{x\in X} L(x,y).} \tag{5.32}$$

Proof. 1. Let $\varepsilon > 0$ and define $L_\varepsilon(x,y) = L(x,y) + \varepsilon|x|^2$. Define also

$$f_\varepsilon(y) = \min_{x\in X} L_\varepsilon(x,y) \qquad (y \in Y). \tag{5.33}$$

Since $x \mapsto L_\varepsilon(x,y)$ is strictly convex, there exists for each $y \in Y$ a *unique* point $\mathbf{e}(y) \in X$ for which

$$f_\varepsilon(y) = L_\varepsilon(\mathbf{e}(y), y). \tag{5.34}$$

2. We show now that

$$\mathbf{e} : Y \to X \text{ is continuous.}$$

To see this, let $\{y_k\}_{k=1}^\infty$ be any sequence in Y such that $\lim_{k\to\infty} y_k = y_0$. Let $x_k = \mathbf{e}(y_k)$ and consider any convergent subsequence: $\lim_{j\to\infty} x_{k_j} = x_0$. We have for all $x \in X$ that

$$L_\varepsilon(x_{k_j}, y_{k_j}) = L_\varepsilon(\mathbf{e}(y_{k_j}), y_{k_j}) \le L_\varepsilon(x, y_{k_j}).$$

Therefore

$$L_\varepsilon(x_0, y_0) \le L_\varepsilon(x, y_0) \quad (x \in X).$$

As a minimizing point is unique, this implies $x_0 = \mathbf{e}(y_0)$.

3. The function $f_\varepsilon(y) = L_\varepsilon(\mathbf{e}(y), y)$ is continuous and so we can select a point $y_\varepsilon \in Y$ that maximizes f_ε:

$$f_\varepsilon(y_\varepsilon) = \max_{y\in Y} f_\varepsilon(y). \tag{5.35}$$

Define then

$$x_\varepsilon = \mathbf{e}(y_\varepsilon). \tag{5.36}$$

We claim that

$$(x_\varepsilon, y_\varepsilon) \text{ is a saddle point for } L_\varepsilon.$$

To see this, note first that (5.36), (5.34) and (5.33) imply for each x that

$$L_\varepsilon(x_\varepsilon, y_\varepsilon) = L_\varepsilon(\mathbf{e}(y_\varepsilon), y_\varepsilon) = f_\varepsilon(y_\varepsilon) \le L_\varepsilon(x, y_\varepsilon).$$

Hence

$$L_\varepsilon(x_\varepsilon, y_\varepsilon) \le L_\varepsilon(x, y_\varepsilon) \qquad (x \in X).$$

4. The other saddle point inequality is trickier and uses the concavity of L in y. Now for each $x \in X, y \in Y$ and $0 < \theta < 1$, we have

$$\begin{aligned} L_\varepsilon(x, (1-\theta)y_\varepsilon + \theta y) &\geq (1-\theta)L_\varepsilon(x, y_\varepsilon) + \theta L_\varepsilon(x, y) \\ &\geq (1-\theta)f_\varepsilon(y_\varepsilon) + \theta L_\varepsilon(x, y), \end{aligned}$$

the second inequality following from (5.33). Put

$$x = \mathbf{e}((1-\theta)y_\varepsilon + \theta y)$$

and recall (5.34), to deduce

$$\begin{aligned} f_\varepsilon((1-\theta)y_\varepsilon + \theta y) &= L_\varepsilon(\mathbf{e}((1-\theta)y_\varepsilon + \theta y), (1-\theta)y_\varepsilon + \theta y) \\ &\geq (1-\theta)f_\varepsilon(y_\varepsilon) + \theta L_\varepsilon(\mathbf{e}((1-\theta)y_\varepsilon + \theta y), y), \end{aligned}$$

Then (5.35) gives

$$\begin{aligned} f_\varepsilon(y_\varepsilon) &\geq f_\varepsilon((1-\theta)y_\varepsilon + \theta y) \\ &\geq (1-\theta)f_\varepsilon(y_\varepsilon) + \theta L_\varepsilon(\mathbf{e}((1-\theta)y_\varepsilon + \theta y), y); \end{aligned}$$

and consequently

$$f_\varepsilon(y_\varepsilon) \geq L_\varepsilon(\mathbf{e}((1-\theta)y_\varepsilon + \theta y), y).$$

Let $\theta \to 0$, to conclude that

$$L_\varepsilon(x_\varepsilon, y_\varepsilon) = f_\varepsilon(y_\varepsilon) \geq L_\varepsilon(\mathbf{e}(y_\varepsilon), y) = L_\varepsilon(x_\varepsilon, y).$$

This is the other saddle point inequality.

5. We have shown for all $x \in X, y \in Y$ that

$$L_\varepsilon(x_\varepsilon, y) \leq L_\varepsilon(x_\varepsilon, y_\varepsilon) \leq L_\varepsilon(x, y_\varepsilon).$$

Select a subsequence $\varepsilon_k \to 0$ and points x_0, y_0 such that

$$x_{\varepsilon_k} \to x_0, \ y_{\varepsilon_k} \to y_0.$$

Then

$$L(x_0, y) \leq L(x_0, y_0) \leq L(x, y_0).$$

This proof is based upon Karlin [**K**]. □

EXAMPLE. (**Zero-sum matrix games again**) The Minimax Theorem provides a quick new proof of the existence of mixed strategies for two-person, zero-sum matrix games, discussed earlier in Section 2.4.2.

Remember that the collection of mixed strategies for Player I is

$$P = \left\{ p \in \mathbb{R}^m \mid p_i \geq 0 \ (i = 1, \ldots, m), \sum_{i=1}^m p_i = 1 \right\}$$

and the collection of mixed strategies for Player II is

$$Q = \left\{ q \in \mathbb{R}^n \mid q_j \geq 0 \ (j = 1, \ldots, n), \sum_{j=1}^{n} q_j = 1 \right\}.$$

Define

$$L(q, p) = p \cdot Aq \qquad (p \in P, q \in Q).$$

Then

$$\begin{cases} q \mapsto L(p, q) \text{ is linear (and thus convex) for each } p, \\ p \mapsto L(p, q) \text{ is linear (and thus concave) for each } q. \end{cases}$$

Hence there exists a mixed strategy saddle point (p_0, q_0):

$$\max_{p \in P} \{p \cdot Aq_0\} = p_0 \cdot (Aq_0) = \min_{q \in Q} \{p_0 \cdot Aq\}.$$

□

APPENDIX

A. Notation

$\mathbb{R}^n$ denotes n-dimensional Euclidean space, a typical point of which is the column vector

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

To save space we will often write the corresponding row vector

$$x = [x_1 \cdots x_n]^T.$$

If $x, y \in \mathbb{R}^n$, we define

$$x \cdot y = \sum_{i=1}^{n} x_i y_i = x^T y, \qquad |x| = (x \cdot x)^{\frac{1}{2}} = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$$

We have the **Cauchy-Schwarz inequality**

$$|x \cdot y| \leq |x||y|$$

and the **parallelogram law**

$$|x - y|^2 + |x + y|^2 = 2|x|^2 + 2|y|^2.$$

We will also write

(i) $B(x, r) = \{y \in \mathbb{R}^n \mid |x - y| \leq r\}$ for the **closed ball** with center x and radius $r > 0$,

(ii) $B^0(x, r) = \{y \in \mathbb{R}^n \mid |x - y| < r\}$ for the **open ball** with center x and radius $r > 0$, and

(iii) $\partial B(x,r) = \{y \in \mathbb{R}^n \mid |x-y| = r\}$ for the **boundary** of $B(x,r)$.

B. Linear algebra

Throughout these notes A denotes an $m \times n$ matrix and A^T denotes its transpose:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \quad A^T = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \dots & a_{mn} \end{bmatrix}.$$

We can interpret A and A^T as linear mappings:

$$A : \mathbb{R}^n \to \mathbb{R}^m, \ A^T : \mathbb{R}^m \to \mathbb{R}^n.$$

The vector equation

$$Ax = b$$

is the system of m equations

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

LEMMA. (Transpose formula) If A is an $m \times n$ matrix, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, then

$$\boxed{(Ax) \cdot y = x \cdot (A^T y).}$$

Proof. The ith entry of Ax is $(Ax)_i = \sum_{j=1}^n a_{ij} x_j$ $(i = 1, \dots, m)$; and the jth entry of $A^T y$ is $(A^T y)_j = \sum_{i=1}^m y_i a_{ij}$ $(j = 1, \dots, n)$. So

$$(Ax) \cdot y = \sum_{i=1}^n (Ax)_i y_i = \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_j y_i = \sum_{j=1}^n (A^T y)_j x_j = x \cdot (A^T y).$$

□

An $n \times n$ matrix A is **symmetric** if $A = A^T$, and a symmetric matrix A is **nonnegative definite** if

$$y^T A y = \sum_{i,j=1}^n a_{ij} y_i y_j \geq 0 \quad \text{for all } y \in \mathbb{R}^n.$$

We then write $A \succeq 0$. We say A is **positive definite** if

$$y^T A y = \sum_{i,j=1}^n a_{ij} y_i y_j > 0 \quad \text{for all } y \in \mathbb{R}^n;$$

and write $A \succ 0$.

C. Multivariable chain rule

Let $f : \mathbb{R}^n \to \mathbb{R}$, $f = f(x) = f(x_1, \ldots, x_n)$. Then we define for $k = 1, \ldots, n$ the k-th partial derivative

$$\frac{\partial f}{\partial x_k}(x) = \lim_{h \to 0} \frac{f(x_1, \ldots, x_k + h, \ldots, x_n) - f(x_1, \ldots, x_k, \ldots, x_n)}{h},$$

provided this limit exists. We likewise define

$$\frac{\partial^2 f}{\partial x_k \partial x_l} = \frac{\partial}{\partial x_l}\left(\frac{\partial f}{\partial x_k}\right) \qquad (k, l = 1, \ldots, n),$$

and recall that $\frac{\partial^2 f}{\partial x_k \partial x_l} = \frac{\partial^2 f}{\partial x_l \partial x_k}$ if f is twice continuously differentiable.

The **gradient** ∇f is the vector

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

and the **Hessian matrix** of second partial derivatives $\nabla^2 f$ is the symmetric $n \times n$ matrix

$$\nabla^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}.$$

The chain rule tells us how to compute the partial derivatives of composite functions, made from simpler functions. For this, assume that we are given a function

$$f : \mathbb{R}^n \to \mathbb{R},$$

which we write as $f(x) = f(x_1, \ldots, x_n)$. Suppose also we have functions

$$g_1, \ldots, g_n : \mathbb{R}^m \to \mathbb{R}$$

so that $g_i(y) = g_i(y_1, \ldots, y_m)$ for $i = 1, \ldots, n$. We then build the composite function $h : \mathbb{R}^m \to \mathbb{R}$ by setting $x_i = g_i(y)$ in the definition of f; that is, we define

$$h(y) = f(g_1(y), g_2(y), \ldots, g_n(y)) = f(\mathbf{g}(y))$$

for $\mathbf{g} = [g_1\, g_2\, \ldots\, g_n]^T$.

THEOREM. (Multivariable chain rule) We have

$$\boxed{\frac{\partial h}{\partial y_k}(y) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{g}(y)) \frac{\partial g_i}{\partial y_k}(y)} \qquad (k = 1, \ldots, m).$$

D. Open and closed sets

DEFINITION. If $\{x^k\}_{k=1}^{\infty}$ is a sequence in $\mathbb{R}^n$ we say

$$\lim_{k \to \infty} x^k = x^0$$

if

$$\lim_{k \to \infty} |x^k - x^0| = 0.$$

DEFINITION. A set $F \subseteq \mathbb{R}^n$ is called **closed** if for all sequences $\{x^k\}_{k=1}^{\infty} \subseteq F$ such that $\lim_{k \to \infty} x^k = x^0$, then $x^0 \in F$.

So a set $F \subseteq \mathbb{R}^n$ is closed when every point in $\mathbb{R}^n$ that is a limit of points in F also belongs to F.

DEFINITION. A set $U \subseteq \mathbb{R}^n$ is called **open** if for each point $x \in U$, there exists $r > 0$ such that $B(x, r) \subseteq U$.

NOTATION. $F^c =$ **complement** of F in $\mathbb{R}^n = \{x \in \mathbb{R}^n \mid x \notin F\}$.

LEMMA.

(i) $F \subseteq \mathbb{R}^n$ is closed if and only if $U = F^c$ is open.

(ii) Let $F_1, \ldots, F_p$ be closed subsets of $\mathbb{R}^n$. Then

$$F = \bigcup_{i=1}^{p} F_i \quad \text{is also closed.}$$

Therefore a finite union of closed sets is closed.

Proof. Let $\{x^k\}_{k=1}^{\infty} \subseteq F$ and suppose the limit

$$\lim_{k \to \infty} x^k = x^0$$

exists. We must show $x^0 \in F$.

There exists $m \in \{1, \ldots, p\}$ and a subsequence $1 \leq k_1 < k_2 < \ldots < k_j \to \infty$ such that $x^{k_j} \in F_m$ for all k_j. Since $x^0 = \lim_{j \to \infty} x^{k_j}$ and F_m is closed, $x^0 \in F_m$. Hence $x^0 \in F = \bigcup_{i=1}^{p} F_i$. $\square$

E. Convergent subsequences

DEFINITION. A sequence $\{x^k\}_{k=1}^{\infty}$ in $\mathbb{R}^n$ is **bounded** if there exists a constant M such that

$$|x^k| \leq M \quad (k = 1, 2, \ldots).$$

THEOREM (**Bolzano-Weierstrass Theorem**)**.** Every bounded sequence in $\mathbb{R}^n$ contains a convergent subsequence.

So if $\{x^k\}_{k=1}^{\infty} \subset \mathbb{R}^n$ is bounded, there exists a sequence $1 \leq k_1 < k_2 < \ldots < k_j \to \infty$ and a point $x^0 \in \mathbb{R}^n$ so that

$$\lim_{j\to\infty} x^{k_j} = x^0.$$

F. Extreme values

DEFINITION. A set $E \subseteq \mathbb{R}^n$ is **bounded** if there exists a constant M such that

$$|x| \leq M \quad \text{for all } x \in E.$$

THEOREM (**Extreme Value Theorem**)**.** Let $C \subseteq \mathbb{R}^n$ be closed and bounded. If $f : C \to \mathbb{R}$ is continuous, then there exists $y \in C$ with

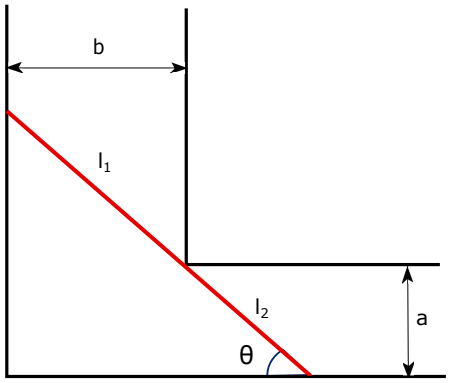
$$f(y) = \min_{x\in C} f(x)$$

Consequently a continuous function f attains its minimum (and maximum) on any closed, bounded set.

EXERCISES

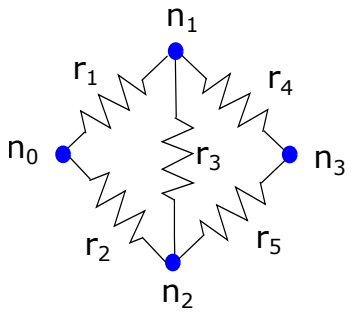
Some of the following problems are taken from Franklin [**F1**].

1. What is the maximum value attained by the function $f(x) = x^{\frac{1}{x}}$ on $(0, \infty)$?
2. In the two-dimensional world of Flatland a corridor of width $a > 0$ meets at right angles another corridor of width $b > 0$, as drawn. What is the length l of the longest (one-dimensional) pipe that can be moved from one corridor to the other?



3. Prove that, as stated on page 6, the angles ξ of incidence and reflection agree for a light ray going from the point A to the point B by reflecting off the x-axis.
4. Four houses are located at the corners of a square with side length 1 mile. Draw, and calculate the length of, the shortest network of roads that can be built to connect all four houses.

5. The electric circuit drawn below is called a Wheatstone bridge.



Suppose a battery is connected across the nodes n_0 and n_3. Show that if the values of the resistors satisfy

$$\frac{r_1}{r_2} = \frac{r_4}{r_5},$$

then no current flows between nodes n_1 and n_2.

6. (i) What point on the line determined by the intersection of the planes $x + 2y + 3z = 0$ and $2x + 3y + z = 4$ is closest to the origin? (ii) For which radius r and height h does a cylindrical tin can of given volume V have the least surface area?

7. Let Γ be a smooth curve in $\mathbb{R}^3$ parameterized by

$$\mathbf{x}(t) = [x^1(t), x^2(t), x^3(t)]^T$$

for $t \in \mathbb{R}$. Suppose that $h(t) = f(\mathbf{x}(t))$ has a minimum at $t = 0$. Show that $\dot{\mathbf{x}}(0)$ is perpendicular to $\nabla f(x_0)$, where $x_0 = \mathbf{x}(0)$.

8. (Continued) Let the intersection of the level sets $\{g_1 = 0\}$ and $\{g^2 = 0\}$ in $\mathbb{R}^3$ be a smooth curve Γ along which $\nabla g_1 \times \nabla g^2 \neq 0$. Assume that $x_0 \in \Gamma$ solves the problem of minimizing f, subject to $g_1, g^2 = 0$.

Explain geometrically why there exist Lagrange multipliers λ_0^1, λ_0^2 such that

$$\nabla f(x_0) + \lambda_0^1 \nabla g_1(x_0) + \lambda_0^2 \nabla g^2(x_0) = 0.$$

9. Rewrite as a canonical linear programming problem:

$$\begin{cases} \text{maximize } x_2, \text{ subject to} \\ 3x_1 - 4x_2 = 7, 4x_1 + 5x_3 = 3 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{cases}$$

State the dual problem.

10. (i) Consider the linear system

$$Ax = b, \quad x \geq 0$$

with nothing to be minimized. Show how to select c so that this is a canonical minimum problem.
(ii) Write the linear system $Ax = b$ as a canonical minimum problem. What is the corresponding dual problem?

11. Find the dual of the problem

$$\begin{cases} \text{minimize } c \cdot x, \text{ subject to} \\ -d \leq Ax - b \leq d, x \geq 0. \end{cases}$$

12. Show how to convert the canonical problem (P) into a problem of the standard form (P^*). Find the corresponding dual problem (D^*) and show how to convert this into the dual (D) of the original problem.

13. Suppose A is a symmetric $n \times n$ matrix. Assume $b \in \mathbb{R}^n$ and consider the linear programming problem:

$$\begin{cases} \text{minimize } b \cdot x, \text{ subject to} \\ Ax = b, \ x \geq 0. \end{cases}$$

Show that any feasible x is optimal.

14. Find the dual problem of

$$\begin{cases} \text{minimize } x_1 + x_2 + x_3, \text{ subject to} \\ x_1 + 2x_2 + 3x_3 = 1, \\ 6x_1 + 5x_2 + 4x_3 = 2, \\ x \geq 0. \end{cases}$$

Write the equilibrium conditions.

15. Consider the canonical problem

$$(P) \qquad \begin{cases} \text{minimize } c \cdot x, \text{ subject to} \\ Ax = b, \ x \geq 0. \end{cases}$$

Suppose x is feasible for (P), and there exist vectors y, z such that

$$A^T y + z = c, \ z \cdot x = 0, \ z \geq 0.$$

Show that x is optimal for (P) and y is optimal for the dual problem (D).

16. The vector $x_0 = [\frac{4}{3}\ \frac{10}{3}\ 0\ 0]^T$ is an optimal solution of

$$(P) \qquad \begin{cases} \text{minimize } 3x_1 + 2x_2, \text{ subject to} \\ 2x_1 + x_2 - x_3 = 6, \\ x_1 + 2x_2 - x_4 = 8, \\ x \geq 0. \end{cases}$$

Use the equilibrium equations to find an optimal solution $y_0 = [y_1, y_2]^T$ of the dual problem (D).

17. Use the equilibrium equations to solve both the primal problem

$$\begin{cases} \text{minimize } 3x_1 + x_2 + 4x_3 + 6x_4, \text{ subject to} \\ x_1 + x_3 = 2,\ x_1 + x_2 + x_3 = 4, x \geq 0 \end{cases}$$

and its dual problem.

18. Derive the equilibrium equations for the standard linear problem

$$\begin{cases} \text{minimize } c \cdot x, \text{ subject to} \\ Ax \geq b,\ x \geq 0. \end{cases}$$

19. Find the equilibirum conditions for the Chebyshev approximation problem:

$$\begin{cases} \text{minimize } x_0, \text{ subject to} \\ -x_0 \leq \sum_{j=1}^{n} a_{ij}x_j - b_i \leq x_0 \quad (i = 1, \ldots, m) \end{cases}$$

20. Solve this problem:

$$\begin{cases} \text{minimize } 5x_1 + 6x_2, \text{ subject to} \\ 3x_1 + 4x_2 \leq 12 \\ 4x_1 + 5x_2 \leq 20 \\ x \geq 0. \end{cases}$$

21. Find all the basic solutions of

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} x = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}.$$

22. Find all the basic solutions of

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}.$$

23. Find the optimal basic solution for

$$\begin{cases} \text{minimize } x_2 - x_3, \text{ subject to} \\ \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix} x = \begin{bmatrix} 0 \\ 7 \end{bmatrix}, \ x \geq 0. \end{cases}$$

24. Guess the optimal solution of

$$\begin{cases} \text{minimize } 2x_1 + x_2 + x_3, \text{ subject to} \\ \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix} x = \begin{bmatrix} 5 \\ 7 \end{bmatrix}, \quad x \geq 0. \end{cases}$$

Then solve the two equilibrium equations for the optimal dual vector y.

25. Starting with the basic solution that depends on the first two columns, apply Phase II to this program:

$$\begin{cases} \text{minimize } x_2, \text{subject to} \\ \begin{bmatrix} -1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} x = \begin{bmatrix} 1 \\ 5 \end{bmatrix}, \quad x \geq 0. \end{cases}$$

26. Use Phase I to compute one of the basic feasible solutions of

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix} x = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad x \geq 0.$$

Remember to multiply the first equation by -1; then start with $x_1 = x_2 = x_3 = 0, z_1 = z_2 = 1$.

27. Apply Phase I to this problem:

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix} x = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad x \geq 0.$$

28. Use the simplex algorithm to solve

$$\begin{cases} \text{minimize } x_1 + x_3, \text{ subject to} \\ \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix} x = \begin{bmatrix} 4 \\ 6 \end{bmatrix}, \quad x \geq 0. \end{cases}$$

29. Consider the canonical problem

$$(P) \qquad \begin{cases} \text{minimize } c \cdot x, \text{ subject to} \\ Ax = b, x \geq 0. \end{cases}$$

for the matrix

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Identify those pairs of vectors b, c that produce each of the four cases in the Duality Theorem.

30. Recall that the following primal problems are equivalent:

$$(\mathrm{P})\begin{cases}\text{minimize } c\cdot x,\\ \text{subject to}\\ Ax=b, x\geq 0\end{cases}\qquad (\mathrm{P}^*)\begin{cases}\text{mininimize } c\cdot x,\\ \text{subject to}\\ \tilde{A}x\geq \tilde{b}, x\geq 0,\end{cases}$$

for

$$\tilde{A}=\begin{bmatrix}A\\ -A\end{bmatrix},\ \tilde{b}=\begin{bmatrix}b\\ -b\end{bmatrix}.$$

Write down the corresponding dual problems (D),(D*) and show that (D) has a feasible solution if and only if (D*) has a feasible solution.

31. Show that if a canonical problem (P) has a feasible solution x and if $c\geq 0$, then it has an optimal solution.

32. Draw the region in $\mathbb{R}^2$ determined by solutions of these inequalities:

$$\begin{cases}3x_1+x_2\geq 6, & x_1+x_2\geq 4\\ x_1+3x_2\geq 6, & x\geq 0.\end{cases}$$

Suppose

$$Cx=\begin{bmatrix}5x_1+x_2\\ x_1+2x_2\end{bmatrix}.$$

Show that the efficient points for the corresponding multiobjective problem lie on the line segments $[p,q]$ and $[q,r]$, where

$$p=\begin{bmatrix}0\\6\end{bmatrix},\quad q=\begin{bmatrix}1\\3\end{bmatrix},\quad r=\begin{bmatrix}3\\1\end{bmatrix}.$$

33. Prove that if the multiobjective linear program

$$\begin{cases}\text{minimize } Cx, \text{ subject to}\\ Ax=b,\ x\geq 0\end{cases}$$

has an efficient solution, then it has a basic efficient solution.

34. A two-person, zero-sum game has the payoff matrix

$$\begin{bmatrix}-3 & 0 & 5\\ 0 & 3 & 8\\ 1 & 4 & 9\end{bmatrix}.$$

Show that this matrix has a saddle point, and find the optimal strategies for each player.

35. Use linear programming to solve the game with payoff matrix

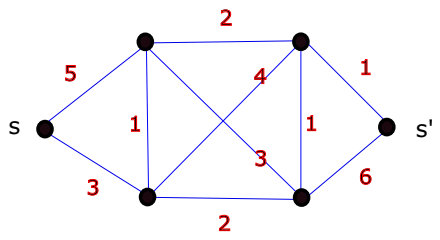
$$A = \begin{bmatrix} 5 & -7 \\ -9 & 4 \end{bmatrix}.$$

To "solve the game" means to find the value ω and optimal mixed strategies p_0, q_0.

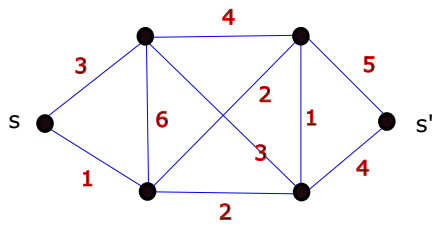
36. Use linear programming to solve the game with payoff matrix

$$A = \begin{bmatrix} -2 & 3 & -1 \\ 1 & -1 & 2 \end{bmatrix}.$$

37. Find a maximal flow and minimum cut for this network:



38. Find a maximal flow and minimum cut for this network, with the same geometry as the previous problem, but different flow capacities:



39. Solve the optimal transport problem for

$$s = \begin{bmatrix} 6 \\ 3 \end{bmatrix}, \; d = \begin{bmatrix} 1 \\ 1 \\ 7 \end{bmatrix}, \; C = \begin{bmatrix} 4 & 3 & 1 \\ 2 & 1 & 3 \end{bmatrix}.$$

(Hint: The answer X_0 has integer entries. Try to guess X_0 and confirm you are correct by computing the corresponding u_0, v_0.)

40. Prove by induction that if C is convex and $\{a^1, \ldots, a^m\} \subset C$, then $\sum_{i=1}^m \theta_i a^i \in C$ for all choices of $\theta_i \geq 0$ such that $\sum_{i=1}^m \theta_i = 1$.

41. (i) Show that the half-plane $3x_1 - 5x_2 < 7$ is convex, but not closed.
(ii) Show that the annulus $1 \leq |x| \leq 2$ is closed, but not convex.

42. Assume $x \in \mathbb{R}^n$ and $r > 0$. Show that the ball $B(x, r) = \{y \in \mathbb{R}^n \mid |x - y| \leq r\}$ is closed and convex.
43. Prove that a convex polytope is closed and convex.
44. Show that the polar dual of $C = B(0, R)$ is $C^0 = B(0, \frac{1}{R})$.
45. Let C_1 and C_2 be convex, disjoint, and closed sets, and suppose C_1 is bounded. Show that the distance between the sets is positive, and prove there is a strictly separating plane.
46. Draw the cone $C = \{Ax \mid x \geq 0\}$ for the matrix
$$A = \begin{bmatrix} 4 & 1 & -2 \\ 1 & 0 & 5 \end{bmatrix}$$
47. A cone in $\mathbb{R}^N$ is a set C such that if $x \in C$, then $\mu x \in C$ for all scalars $\mu \geq 0$. A finite cone has the form
$$C = \{x_1 a^1 + x_2 a^2 + \ldots + x_n a^n \mid x_j \geq 0\},$$
where $a^1 \ldots, a^n \in \mathbb{R}^N$. Find a cone in $\mathbb{R}^N$ that is not a finite cone.
48. Note that
(i) $Ax = 0$, $x \geq 0$ has a non-zero solution x
if and only if $Ax = 0$, $e \cdot x = 1$ has a solution $x \geq 0$, where $e^T = [1 \ldots 1]^T$. Prove that the Farkas alternative of this is
(ii) $\mathrm{A^T y} > 0$ has a solution y.
49. Find the Farkas alternative of this assertion: The system
$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} x = b$$
has a solution x with $x_1 \geq 0$ and $x_3 \geq 0$. (Hint: Set $x_2 = u_2 - v_2$, with $u_2 \geq 0$, $v_2 \geq 0$.)
50. Find the Farkas alternative to the assertion
(i)* $Ax \leq b$ has a solution x.
(Hint: First write down statements (i) and (ii) of the usual Farkas alternative. Introduce slack variables to rewrite (i)* into the form (i).)
51. Prove by induction that if $f : \mathbb{R} \to \mathbb{R}$ is convex, then
$$f\left(\sum_{i=1}^{m} \theta_i x_i\right) \leq \sum_{i=1}^{m} \theta_i f(x_i)$$

for all positive integers m, all $x_1, \dots, x_m \in \mathbb{R}$, and all $\theta_1, \dots, \theta_m \geq 0$ such that $\sum_{i=1}^m \theta_i = 1$.

52. Determine which of the following functions $f : \mathbb{R} \to \mathbb{R}$ are convex:
    (a) $f(x) = e^{-x}$
    (b) $f(x) = e^{-x^2}$
    (c) $f(x) = |x|^5$
    (d) $f(x) = (x^2 - 1)^2$.
53. Determine which of the following functions $f : \mathbb{R}^n \to \mathbb{R}$ are convex:
    (a) $f(x) = |x|^4$
    (b) $f(x) = e^{-|x|^2}$
    (c) $f(x) = e^{a \cdot x}$, where $a \in \mathbb{R}^n$.
54. Suppose that $f^1, \dots, f^m : \mathbb{R}^n \to \mathbb{R}$ are convex functions. Show that the function
$$g(x) = \max\{f^k(x) \mid k = 1, \dots, m\}$$
is also convex.
55. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Prove that $x_0 \in \mathbb{R}^n$ satisfies $f(x_0) = \min_{x \in \mathbb{R}^n} f(x)$ if and only if $0 \in \partial f(x_0)$.
56. Compute ∂f for these convex functions $f : \mathbb{R} \to \mathbb{R}$:
    (a) $f(x) = \max\{x^2 + x, x^2 - x\}$
    (b) $f(x) = |x| + |x - 1|$.
57. Compute f^* for the following functions $f : \mathbb{R} \to \mathbb{R}$:
    (a) $f(x) = \frac{|x|^p}{p} \quad (1 < p < \infty)$
    (b) $f(x) = e^x$
    (c) $f(x) = |x|$.
58. Use Jensen's inequality to prove that if $a_1, \dots, a_m > 0$ and $\theta_1, \dots, \theta_m > 0$ with $\sum \theta_i = 1$, then
$$a_1^{\theta_1} \cdots a_m^{\theta_m} \leq \theta_1 a_1 + \cdots + \theta_m a_m.$$
59. Show that if f is convex, then for each interval $(a, b) \subset \mathbb{R}$ we have
$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a} \int_a^b f(x)\, dx \leq \frac{f(a) + f(b)}{2}.$$
60. Prove Hölder's inequality: If $\frac{1}{p} + \frac{1}{q} = 1$, then
$$\int_a^b fg\, dx \leq \left(\int_a^b |f|^p\, dx\right)^{\frac{1}{p}} \left(\int_a^b |g|^q\, dx\right)^{\frac{1}{q}}.$$
(Hint: By Young's Inequality, $fg = \frac{f}{\mu}\, \mu g \leq \frac{|f|^p}{\mu^p p} + \frac{\mu^q |g|^q}{q}$. Integrate and then select $\mu > 0$ to minimize the right hand side.)

61. Solve the problem

$$\begin{cases} \text{maximize } (x_1 - 3)^6 + (x_2 - 4)^6 \\ \text{subject to } x_1^2 + x_2^2 \leq 25, \\ x_1 + x_2 \geq 7,\ x \geq 0. \end{cases}$$

Draw a relevant picture. Show that (CQ) holds, and verify the KKT conditions.

62. Consider the problem

$$\begin{cases} \text{maximize } (x_1 - 3)^2 + (x_2 - 4)^2 \\ \text{subject to } x_1^2 + x_2^2 \leq 25,\ x \geq 0. \end{cases}$$

Guess the solution x_0, and verify the KKT conditions.

63. Consider in two dimensions the problem (MIN*) for minimizing a function f subject to the inequality constraints $h_1 \leq 0, h^2 \leq 0$. Suppose that x_0 solves (MIN*) and that $h_1(x_0) = h^2(x_0) = 0$, with $\{\nabla h_1(x_0), \nabla h^2(x_0)\}$ linearly independent.

Draw a picture and explain geometrically why the Lagrange multipliers μ_0^1, μ_0^2 are nonnegative.

64. Recall that an $n \times n$ symmetric matrix A has real eigenvalues. For such a matrix A, let x_0 be a solution of

$$\begin{cases} \text{minimize } x \cdot Ax, \\ \text{subject to } |x|^2 = 1. \end{cases}$$

Use the KKT conditions to show that x_0 is an eigenvector of A, corresponding to the smallest eigenvalue.

65. Find the solution x_0 of

$$\begin{cases} \text{maximize } c \cdot x, \\ \text{subject to } x^T Ax \leq 1, \end{cases}$$

where A is an $n \times n$ symmetric, positive definite matrix. What is the Lagrange multiplier μ_0?

66. Let x_0 minimize the function f over the set $C = \{x \in \mathbb{R}^n \mid x \geq 0\}$. Use the variational inequality (VI) to show that

$$\nabla f(x_0) \cdot x_0 = 0.$$

Explain the meaning of this, by considering the the various possibilities for the location of x_0 within C.

Assume that the functions $h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$ are convex and continuously differentiable.

67. Show that if $h_1, \dots, h_p$ are convex, then Slater's condition implies the Fromovitz–Mangasarian condition for each x_0.

68. Consider the problem

$$\min_{x \in \mathbb{R}^n} \tfrac{1}{2}|Ax - b|^2.$$

(i) Rewrite this into the form

(P) minimize $\tfrac{1}{2}|z|^2$, subject to $Ax - b = z$.

(ii) Show the dual problem is

(D) maximize $\{b \cdot y - \tfrac{1}{2}|y|^2\}$, subject to $A^T y = 0$.

69. Solving the previous problem is equivalent to solving

$$\min_{x \in \mathbb{R}^n} |Ax - b|.$$

(i) Rewrite this into the form

(P) minimize $|z|$, subject to $Ax - b = z$.

(ii) Show the dual problem is

(D) maximize $b \cdot y$, subject to $|y| \le 1, A^T y = 0$.

70. Suppose that $x_0 \in \mathbb{R}^n$ solves the problem

$$\begin{cases} \text{minimize } f(x), \text{ subject to} \\ \mathbf{h}(x) \le 0, x \ge 0, \end{cases}$$

where $\mathbf{h} = [h_1 \ \cdots \ h_p]^T$. Let

$$L(x, \mu) = f(x) + \mu \cdot \mathbf{h}(x).$$

Assume the KKT conditions hold and use them to show that there exists $\mu_0 \in \mathbb{R}^p$ such that

$$\begin{cases} \nabla_x L(x_0, \mu_0) \ge 0, \ \nabla_x L(x_0, \mu_0) \cdot x_0 = 0, \ x_0 \ge 0 \\ \nabla_\mu L(x_0, \mu_0) \le 0, \ \nabla_\mu L(x_0, \mu_0) \cdot \mu_0 = 0, \ \mu_0 \ge 0. \end{cases}$$

(Notation: ∇_x means the gradient with respect to $x \in \mathbb{R}^n$; ∇_μ means the gradient with respect to $\mu \in \mathbb{R}^p$.)

71. Select $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and define

$$f(x) = \begin{cases} x \cdot c & \text{if } x \ge 0 \\ \infty & \text{otherwise,} \end{cases} \qquad g(y) = \begin{cases} 0 & \text{if } y = b \\ \infty & \text{otherwise.} \end{cases}$$

Let A be an $m \times n$ matrix. Interpret the problems of minimizing $f(x) + g(Ax)$ and maximizing $-f^*(A^T y) - g^*(-y)$ in terms of linear programming.

72. Show that the function $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x, y) = x \cdot y$$

is not convex, but

$$\begin{cases} x \mapsto L(x, y) & \text{is convex for each } y, \\ y \mapsto L(x, y) & \text{is convex for each } x. \end{cases}$$

73. Define

$$L(x, y) = x \cdot c + y \cdot (b - Ax).$$

and suppose that (x_0, y_0) satisfies the saddle point condition

$$L(x_0, y) \leq L(x_0, y_0) \leq L(x, y_0)$$

for all $y \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$ with $x \geq 0$. Show that x_0 solves the canonical linear programming problem problem (P) and y_0 solves the dual problem (D).

(Hint: Show that $Ax_0 = b$, $A^T y_0 \leq c$ and $x_0 \cdot (A^T y_0 - c) = 0$.)

Bibliography

[Ba] E. N. Barron, *Game Theory, An Introduction* (2nd ed), Wiley, 2013

[B] D. Bertsekas, *Nonlinear Programming* (3rd ed), Athena Scientific

[B-T] D. Bertsekas and J. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific, 1997

[B-L] J. Borwein and A. Lewis, *Convex Analysis and Nonlinear Optimization: Theory and Examples* (CMS Books in Mathematics) 2nd ed, Springer, 2006

[B-V] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge U Press, 2004

[B] H. Brezis, Remarks on the Monge-Kantorovich problem in the discrete setting, C. R. Acad. Paris Ser I (1356) 2018, 207–213

[C-EG] G. Calafiore and L. El Ghaoui, *Optimization Models*, Cambridge U Press, 2014

[C] B. Cipra, The best of the 20th century: Editors name top 10 algorithms, SIAM News (33) 2000

[C-H] J. Cohen and P. Horowitz, Paradoxical behavior of mechanical and electrical networks, Nature (352) 1991, 699–701

[Co] J. Cohon, *Multiobjective Programming and Planning*, Dover, 2004

[C-G] H. S. M. Coxeter and S. L. Greitzer, *Geometry Revisited*, A. Lax New Mathematics Library, Vol 19, Math Association America, 1967

[dJ] T. de Jong, Lagrange multipliers and the fundamental theorem of algebra, American Math Monthly (116) 2009, 828–830

[F1] J. Franklin, *Methods of Mathematical Economics: Linear and Nonlinear Programming, Fixed-Point Theorems*, Springer, 1980

[F2] J. Franklin, Mathematical methods of economics, Amer. Math. Monthly (90) 1983, 229–244.

[K] S. Karlin, *Mathematical Methods and Theory in Games, Programming, and Economics*, Volume I, Addison-Wesley, 1959

[Ko] T. W. Körner, *The Pleasures of Counting*, Cambridge, 1996

[L-H] C. E. Lemke and J. T. Howson Jr., Equilibrium points of bimatrix games, J. Soc. Indust. Appl. Math (12) 1964, 413–423

[M] Z. A. Melzak, *Companion to Concrete Mathematics*, Wiley–Interscience, 1973

[MS] E. J. McShane, The Lagrange multiplier rule, American Math Monthly (80) 1973, 922–925

[N] P. J. Nahin, *When Least is Best*, Princeton U Press, 2004

[S] G. Strang, Duality in the classroom, SIAM Review, 1984