Applied Mathematics 2

MATH2070 and MATH2970

Optimisation and Financial Mathematics

Lecture Notes

P. W. Buchen and D. J. Ivers

2014

The University of Sydney
School of Mathematics and Statistics

Contents

Chapter 1

Introduction

Example: Find the cylinder of given surface area S, for which the volume V is a maximum.

The problem can be formulated mathematically as follows: Let r and h be the radius and height of the cylinder. Then

$$V = \pi r^2 h\,, \qquad S = 2\pi r^2 + 2\pi r h\,.$$

The problem is then,

Maximise:	$V = \pi r^2 h$
subject to:	$2\pi r^2 + 2\pi r h = S \quad (= \text{const.})$
with:	$r, h \geq 0\,.$

The problem is thus to maximise the function $V = V(r, s)$ of two-independent variables r and h, for which S is given. V is the *objective function.* The condition and

$$2\pi r^2 + 2\pi r h = S = \text{const.}$$

is a *constraint.* The inequalities $r \geq 0$ and $h \geq 0$ are *natural constraints.*
We can simplify the form of the problem by changing to the new variables x, y and z defined by

$$r = x\sqrt{S/2\pi}\,, \qquad h = y\sqrt{S/2\pi}\,, \qquad V = z\sqrt{S^3/8\pi}\,.$$

The problem then takes the form

Maximise:	$z = x^2 y$
subject to:	$x^2 + xy = 1$
with:	$x, y \geq 0\,.$

By measuring length in units of $\sqrt{S/2\pi}$ and volume in units of $\sqrt{S^3/8\pi}$, we have *non-dimensionalised* the problem and removed a nuisance $\sqrt{\pi}$ factor.

We can reduce the above problem to a one-dimensional problem in the independent variable x. (Alternatively, we could use the more general method of Lagrange multipliers.) Thus

$$z = x(1 - x^2)\,.$$

This gives the equivalent problem,

Example: Maximise the function $z = f(x) = x(1 - x^2)$ on the interval $0 \leq x \leq 1$. The upper limit on x comes from $y \geq 0$. We can display this as:

Maximise $z = x(1 - x^2)$
with: $0 \leq x \leq 1\,.$

1 Optimising Differentiable Functions of One Variable

Since the objective function in Example 1.2 is differentiable, conditions for optimising $z = f(x)$ on an interval $[a, b]$, which rely upon $dz/dx = f'(x)$ and higher derivatives, can be used. These conditions, which should be familiar, are derived in Chapter 3. (If f is continuous on $[a, b]$, but $f'(x)$ does not exist everywhere on $[a, b]$, other methods must be used.)

First Derivative Test

If a maximum or minimum of $f(x)$ occurs at an interior point x_0 (that is, a point away from the boundary of the region determined by the constraints), then x_0 must be a *critical point*, i.e. a point where

$$f'(x_0) = 0\,.$$

This is a *necessary* condition for an extremum (a maximum or a minimum), but it is not sufficient. In general there is no analytic method of solving the equation $f'(x_0) = 0$ for x_0; instead some method must be used to approximate x_0. A critical point may be a local maximum or a local minimum, otherwise it is a *point of inflection*. A point x_0 is a *local maximum* (*local minimum*) of $f(x)$, if $f(x) \leq f(x_0)$ $(f(x) \geq f(x_0))$ for all x in some neighbourhood of x_0. In Example 1.2,

$$f'(x) = 1 - 3x^2\,.$$

Thus $f'(x) = 0$ implies

$$x = \pm\frac{1}{\sqrt{3}}\,.$$

Only the positive root satisfies $0 \leq x \leq 1$.

Second Derivative Test

The second derivative $f''(x)$ can be used to test for a local maximum or a local minimum:

1. If $f'(x_0) = 0$ and $f''(x_0) < 0$, then x_0 is a *local maximum*.

2. If $f'(x_0) = 0$ and $f''(x_0) > 0$, then x_0 is a *local minimum*.

3. If $f''(x_0) = 0$, then the test fails and higher derivatives must be examined or some other test for optimality must be used.

The tests for a local maximum or local minimum are *sufficient*. In Example 1.2, $f''(x) = -6x$. Since $f''(1/\sqrt{3}) = -2\sqrt{3} < 0$, $x = 1/\sqrt{3}$ is a local maximum of $f(x) = x(1 - x^2)$.
If the second derivative test fails, i.e. $f''(x_0) = 0$, higher derivatives must be considered:

1. If $f^{(1)}(x_0) = f^{(2)}(x_0) = \cdots = f^{(2m-1)}(x_0) = 0$ for some integer m and $f^{(2m)}(x_0) < 0$, then x_0 is a *local maximum*.

2. If $f^{(1)}(x_0) = f^{(2)}(x_0) = \cdots = f^{(2m-1)}(x_0) = 0$ for some integer m and $f^{(2m)}(x_0) > 0$, then x_0 is a *local minimum*.

3. If $f^{(1)}(x_0) = f^{(2)}(x_0) = \cdots = f^{(2m)}(x_0) = 0$ for some integer m and $f^{(2m+1)}(x_0) \neq 0$, then x_0 is a *point of inflection*.

It is possible for all the derivatives of a non-constant function to vanish at one point, e.g. all derivatives of

$$f(x) = \begin{cases} e^{-1/x^2}, & \text{if } |x| \neq 0; \\ 0, & \text{if } x = 0; \end{cases}$$

vanish at $x = 0$.

Global Extrema

Is $x = 1/\sqrt{3}$ a global maximum of $f(x) = x(1 - x^2)$? A point x_0 is a *global maximum* (*global minimum*) of $f(x)$ in the interval $[a, b]$, if $f(x) \leq f(x_0)$ ($f(x) \geq f(x_0)$) for all x in $[a, b]$. Global extrema can occur at critical points in the open interval (a, b), i.e. $a < x < b$, *or* at the endpoints a or b. (If $a = -\infty$ or $b = \infty$, the limits $x \to -\infty$ or $x \to \infty$, respectively, must be considered.) The easiest test for global extrema is to compare the values of f at the critical points and the endpoints. This will also tell which critical points are local maxima, which are local minima and which are points of inflection, and whether a and b are local maxima or local minima. In Example 1.2 ($a = 0$, $b = 1$) this gives

$$f(0) = 0\,, \qquad f(1/\sqrt{3}) = 2/3\sqrt{3}\,, \qquad f(1) = 0\,.$$

Hence $x = 1/\sqrt{3}$ is a (the) global maximum of $f(x) = x(1-x^2)$ in $[0, 1]$; $x = 0$ and $x = 1$ are global minima. See the graph of $z = x(1 - x^2)$ in Figure 1.1.
We can also use the derivative f' of f to test if a is a local maximum or minimum of f:

1. If $f'(a) < 0$, then a is a *local maximum*.

2. If $f'(a) > 0$, then a is a *local minimum*.

3. If $f'(a) = 0$, then the test fails and higher derivatives must be examined.

Similarly,

1. If $f'(b) > 0$, then b is a *local maximum*.

2. If $f'(b) < 0$, then b is a *local minimum*.

3. If $f'(b) = 0$, then the test fails and higher derivatives must be examined.

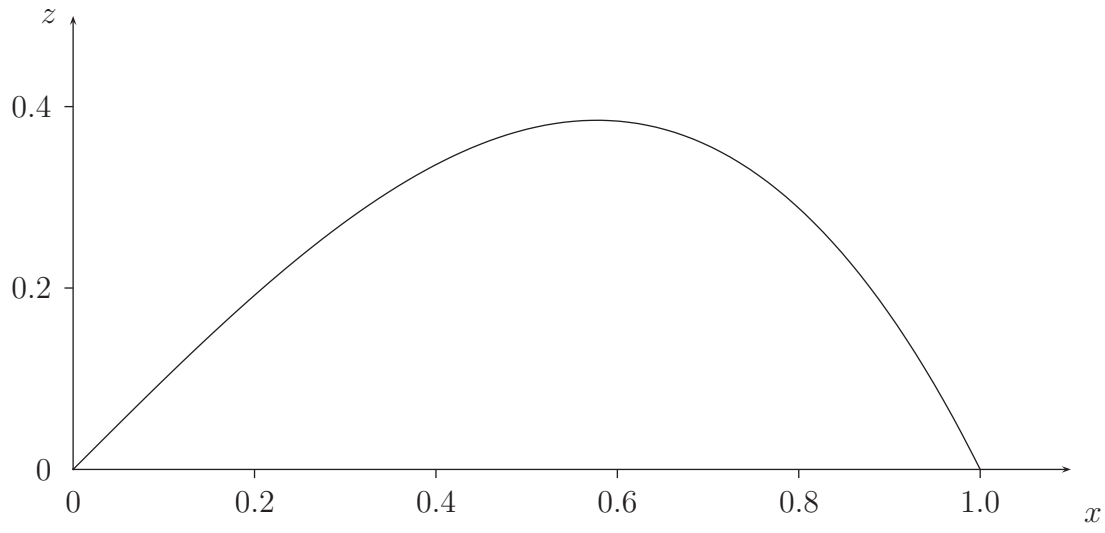


Figure 1.1: The graph of $z = x(1 - x^2)$.

Summary

To optimise a differentiable function $f(x)$:

1. Find the critical points $x_1 < x_2 < \cdots < x_m$ of f, where $f'(x_i) = 0$.
2. For $i = 1, \ldots, m$ test x_i for a local maximum, $f''(x_i) < 0$, or a local minimum, $f''(x_i) > 0$; if $f''(x_i) = 0$, use higher derivatives.
3. Compare $f(a), f(x_1), \ldots, f(x_m), f(b)$ (let $x_0 = a$, $x_{m+1} = b$):
   (a) If $f(x_i) < f(x_{i-1})$ and $f(x_i) < f(x_{i+1})$, then $f(x_i)$ is a local minimum.
   (b) If $f(x_i) > f(x_{i-1})$ and $f(x_i) > f(x_{i+1})$, then $f(x_i)$ is a local maximum.
   (c) If $f(x_{i-1}) < f(x_i) < f(x_{i+1})$ or $f(x_{i-1}) > f(x_i) > f(x_{i+1})$, then $f(x_i)$ is a point of inflection.

   The global maximum is
   $$\max\{f(a), f(x_1), \ldots, f(x_m), f(b)\}$$
   and the global minimum is
   $$\min\{f(a), f(x_1), \ldots, f(x_m), f(b)\}\,.$$

Step 2 is actually redundant if Step 3 is used.

2 Optimising Differentiable Functions of Several Variables

Optimising a function of several variables is usually much more difficult than optimising a function of one variable. Let $z = f(x_1, x_2, \ldots, x_n)$ be a function of n variables defined on a domain D in $\mathbb{R}^n$. The function f has a local extremum (maximum or minimum) at an interior point $\mathbf{x}_0 = (x_{10}, x_{20}, \ldots, x_{n0})$, only if $\mathbf{x}_0$ is a *critical point* of f, i.e. a point where

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \cdots = \frac{\partial f}{\partial x_n} = 0\,.$$

This condition, which is derived in Chapter 3, is only *necessary*, not sufficient. Sufficient conditions for local extrema are also considered later. The $n(n+1)/2$ second derivatives of f can be used to test whether a critical point is a local maximum or a local minimum. This test is complicated and will be treated later in the course. However, before we study extrema at interior points, we will consider optimisation problems, which have linear objective functions and domains with linear boundaries. These form the very important special class of *linear programming problems*, which we will examine extensively. As the following example shows, *linear objective functions must take their extrema on the boundary of their domain*. Even though linear objective functions are simple, such problems can be quite complicated to solve because the boundaries of domains in two or more dimensions can be very complicated.

Example: The function $z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$, where $c_1, c_2, \ldots, c_n$ are constants, defined on a domain D, cannot take a local maximum or a local minimum in the interior of D, unless $c_1 = c_2 = \cdots = c_n = 0$.
Here is an example linear programming problem:

Example: A pharmaceutical company produces two types of drug using three different processes. Each process requires the use of a chemical catalyst X which is in limited daily supply. Production data is listed in the table below:

	X usage in gm/unit		Maximum Daily Supply X (gm)
Process	Drug 1	Drug 2	
1	1	0	4
2	3	2	18
3	0	2	12
Profit/unit (k$)	3	5	

Thus Process 1 produces only Drug 1, uses 1 gm of X per unit of drug produced and has a maximum daily supply of 4 gm of X. Similar interpretations apply to Processes 2 and 3. Drug 1 sells at a unit price of $3000 while Drug 2 yields a profit of $5000 per unit.
Find the number of units of each drug which should be produced daily in order that the company maximizes its daily profit.

3 Gauss-Jordan Elimination

We review Gauss-Jordan elimination for systems of linear algebraic equations with more equations than unknowns.

3.1 Pivot Operations

Consider the following system of linear equations

$$\begin{aligned} x_1 - \ x_2 - \ x_3 &= -2 \\ 2x_1 + \ x_2 - \ x_3 &= \ 5 \\ -\ x_1 + 2x_2 + 3x_3 &= \ 0\,. \end{aligned} \tag{3.1}$$

We can perform elementary row operations on (3.1) without altering its solution set:

1. Multiply (or divide) any row by a non-zero constant;
2. Add a (non-zero) multiple of one row to any other row;
3. Interchange of the two rows.

We will not use the third row operation. For example, if we write the system (3.1) as the *tableau*,

$$T_1 = \begin{array}{|rrr|r|l} \multicolumn{3}{c|}{x_1 \quad x_2 \quad x_3} & \text{RHS} & \\ \hline 1 & -1 & 1 & -2 & \leftarrow R_1 \\ 2 & 1 & -1 & 5 & \leftarrow R_2 \\ -1 & 2 & 3 & 0 & \leftarrow R_3 \\ \hline \end{array}$$

then the row operations: $R_2 := R_2/2$ and $R_1 := R_1 - R_2$ transform T_1 to the equivalent tableau,

$$T_2 = \begin{array}{|rrr|r|} \multicolumn{3}{c|}{x_1 \quad x_2 \quad x_3} & \text{RHS} \\ \hline 0 & -3/2 & 3/2 & -9/2 \\ 1 & 1/2 & -1/2 & 5/2 \\ -1 & 2 & 3 & 0 \\ \hline \end{array}$$

T_2 is equivalent to T_1 in the sense that they have exactly the same solution set (x_1, x_2, x_3). A *pivot operation* is a combination of elementary row operations.
Let a_{ij} be any non-zero element (called the *pivot element* or simply pivot) of the coefficient matrix A of a given system of linear equations: i.e. a_{ij} is the element in row i and column j of the corresponding tableau. Then to pivot on $a_{ij} \neq 0$, denoted P_{ij},

1. Divide row i by a_{ij};
2. Transform to zero all elements a_{kj}, $k \neq i$ (i.e. the elements in the same column j as a_{ij} except row i) by adding suitable multiples of row i.

For example, let us pivot the tableau T_1 on the element $a_{21} = 2$ (row 2, column 1). We get

$$T_3 = P_{21}T_1 = \begin{array}{|rrr|r|} \multicolumn{3}{c|}{x_1 \quad x_2 \quad x_3} & \text{RHS} \\ \hline 0 & -3/2 & 3/2 & -9/2 \\ 1 & 1/2 & -1/2 & 5/2 \\ 0 & 5/2 & 5/2 & 5/2 \\ \hline \end{array}$$

T_3 is also equivalent to T_1 since the pivot operation does not change the solution set (x_1, x_2, x_3).

3.2 Gauss-Jordan Elimination

This is a procedure for obtaining the solution set of a system of linear equations, by performing a sequence of pivot operations. The usual method is to pivot in sequence, on the elements $a_{11}, a_{22}, a_{33}, \ldots$ but other sequences are also possible.
Let's see how this works for the system (3.1). The pivot elements are underlined.

	x_1	x_2	x_3	RHS
	$\underline{1}$	-1	1	-2
	2	1	-1	5
	-1	2	3	0
	1	-1	1	-2
P_{11} :	0	$\underline{3}$	-3	9
	0	1	4	-2
	1	0	0	1
P_{22} :	0	1	-1	3
	0	0	$\underline{5}$	-5
	1	0	0	1
P_{33} :	0	1	0	2
	0	0	1	-1

Notes:

1. Gauss-Jordan elimination has transformed the original coefficient matrix to the identity matrix.

2. A column of zeros and a single one in row k, is unaltered by a pivot operation on row i, $i \neq k$.

3. The solution set can be read off the last column: $(x_1, x_2, x_3) = (1, 2, -1)$.

3.3 More Variables Than Equations

Consider now the following system of three equations in five variables,

$$\begin{aligned} x_1 - x_2 + x_3 - x_4 &= -2 \\ 2x_1 + x_2 - x_3 + x_5 &= 5 \\ -x_1 + 2x_2 + 3x_3 + x_4 + 2x_5 &= 0 \end{aligned} \tag{3.2}$$

We can use Gauss-Jordan elimination to express any three of the variables in terms of the remaining two, e.g. (x_2, x_2, x_3) in terms of (x_4, x_5). The variables we solve for (x_1, x_2, x_3), say, are called the *basic variables*. They must be linearly independent. The remaining variables (x_4, x_5) are called *non-basic variables*: they may assume any value whatsoever.
Here are the tableaus for the system (3.2),

Basic			Non–Basic		
x_1	x_2	x_3	x_4	x_5	RHS
1	−1	1	−1	0	−2
2	1	−1	0	1	5
−1	2	3	1	2	0
1	−1	1	−1	0	−2
0	3	−3	2	1	9
0	1	4	0	2	−2
1	0	0	−1/3	1/3	1
0	1	−1	2/3	1/3	3
0	0	5	−2/3	5/3	−5
1	0	0	−1/3	1/3	1
0	1	0	8/15	2/3	2
0	0	1	−2/15	1/3	−1

The final tableau tells us that the solution is given by

$$\begin{aligned} x_1 &= 1 + \tfrac{1}{3}x_4 - \tfrac{1}{3}x_5 \\ x_2 &= 2 - \tfrac{8}{15}x_4 - \tfrac{2}{3}x_5 \\ x_3 &= -1 + \tfrac{2}{15}x_4 - \tfrac{1}{3}x_5 \end{aligned}$$

3.4 Transformation of Linear Functions by Gauss-Jordan Elimination

Let $Z = c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5 + c_0$ be a given linear function of 5 variables $(x_1, \ldots, x_5)$. Suppose also that these variables must satisfy the system (3.2). Then if (x_1, x_2, x_3) are basic variables, we can express Z entirely in terms of the non-basic variables (x_4, x_5); i.e.

$$Z = Ax_4 + Bx_5 + C \qquad \text{for some constant } A,\ B,\ C.$$

One way to do this is to solve system (3.2) as above, and then substitute for (x_1, x_2, x_3) in terms of (x_4, x_5). A better way is to regard the linear function as an extra equation and to introduce Z as an extra basic variable. For example, suppose we are given

$$Z = x_1 + x_2 - x_5 \ .$$

Then we can write this in the equivalent form

$$Z - x_1 - x_2 + 0x_3 + 0x_4 + x_5 = 0 \, .$$

This extra equation is added to the tableau and becomes row 1.

Z	x_1	x_2	x_3	x_4	x_5	RHS	
1	−1	−1	0	0	1	0	$\leftarrow R_1$
0	<u>1</u>	−1	1	−1	0	−2	$\leftarrow R_2$
0	2	<u>1</u>	−1	0	1	5	$\leftarrow R_3$
0	−1	2	<u>3</u>	1	2	0	$\leftarrow R_4$

Pivoting on the underlined elements in the usual way, we arrive at the final tableau:

Z	x_1	x_2	x_3	x_4	x_5	RHS	
1	0	0	0	$1/5$	2	3	$\leftarrow R_1$
0	1	0	0	$-1/3$	$1/3$	1	$\leftarrow R_2$
0	0	1	0	$8/15$	$2/3$	2	$\leftarrow R_3$
0	0	0	1	$-2/15$	$1/3$	-1	$\leftarrow R_4$

The last three rows of this tableau determine (x_1, x_2, x_3) in terms of (x_4, x_5) as in Section 3.2 above. Row 1 solves the transformation problem for Z and yields the result

$$Z = 3 - \tfrac{1}{5}x_4 - 2x_5 \, .$$

References

The following books cover aspects of this course and many other topics besides, hence are only recommended as references and for further reading:

Luenberger, D.G. (1998), *Investment Science*, Oxford University Press, ISBN 0-19-510809-4.

Taha, H.A. (1992), *Operations Research, An Introduction*, 5th ed., MacMillan Publishing, ISBN 0-02-418975-8.

Winston, W.L. (1995), *Introduction to Mathematical Programming, Applications and Algorithms*, 2nd ed., Duxbury Press/ Wadsworth Publishing, ISBN 0-534-23046-6.

Winston, W.L. (1991), *Operations Research, Applications and Algorithms*, 2nd ed., PWS-Kent Publishing, ISBN 0-534-92495-6.

More advanced references are:

Chvátal, V. (1983), *Linear Programming*, W.H. Freeman & Company, ISBN 0-7167-1587-2.

Fletcher, R. (1987), *Practical Methods of Optimization*, 2nd ed., John Wiley & Sons, ISBN 0-471-91547-5.

Luenberger, D.G. (2003), *Linear and Nonlinear Programming*, 2nd ed., Kluwer Academic Publishing, ISBN 1-4020-7593-6.

Software

We will use Matlab (Matrix Laboratory) in this course. A student edition (PC or Macintosh) is available,

Student Edition of Matlab, Version 7.0., Prentice-Hall, Englewood Cliffs, New Jersey.

Taha (1992) and Winston (1995) include disks with useful optimisation sotfware. Taha (1992) comes with PC software, called TORA, to perform linear programming, etc.; simulation software, called SIMNET II, is also available. Winston (1995) comes with student-edition PC software (Macintosh versions are available), called LINDO (Linear Interactive and Discrete Optimizer, Linus Schrage 1986), GINO and LINGO, to perform various optimisations.

Chapter 2

Linear Programming

1 The Standard LP Problem

Linear programming (LP) is the term which is used to describe a wide class of problems in the field of constrained optimization. Broadly speaking, LP seeks to solve the problem of sharing activities amongst limited resources in an optimal way. However, the key feature of LP problems is that all objectives and constraints occur as *linear* functions of their associated variables.

LP is one of the fundamental problems of Operations Research (OR). Techniques for solving LP problems were developed in military research facilities during World War II. The *simplex algorithm*, the basic method used for solving LP problems, was developed by George Dantzig in 1947.

1.1 A Model Linear Programming Problem

A pharmaceutical company can produce two types of drug by three different processes. Each process requires the use of a chemical catalyst X which is in very limited supply. Production data are listed in the table below:

	X usage/unit		X (gm)
	Drug 1	Drug 2	Availability
Process 1	1	0	4
2	3	2	18
3	0	2	12
Profit/unit	3	5	$\times$ \$1000

Thus process 1 produces only drug 1, uses 1 gm of X per unit produced and has only 4 gm of X available. Similar interpretations apply for processes 2 and 3. Drug 1 sells at a unit profit of \$3000 while drug 2 sells for a unit profit of \$5000. The problem is to find the number of units of each drug to produce in order to maximize the total profit.

Mathematical Formulation: Let x_1, x_2 be the number of units of each drug which are to be produced, and let Z be the total profit in units of \$1000. Then the LP formulation of the drug problem is as follows:

Model LP Problem

$$
\begin{array}{lrcrcrcr}
\text{Maximize:} & Z & = & 3x_1 & + & 5x_2 & & \\
\\
\text{subject to:} & & & x_1 & & & \leq & 4 \\
& & & 3x_1 & + & 2x_2 & \leq & 18 \\
& & & & & 2x_2 & \leq & 12 \\
\\
\text{with:} & & & x_1 \geq 0\,, & & x_2 \geq 0\,. & &
\end{array}
$$

We shall study this problem in great detail in the following sections. Rather than an abstract theoretical treatment, we adopt a more heuristic approach and use examples to illustrate the basic concepts of the *simplex algorithm*.

Before we begin, however, it is important to identify the generalization of this problem.

1.2 The Standard LP Problem

The generalization of the drug problem is given below:

Standard LP Problem

$$
\begin{array}{llrcccccccc}
\text{I.} & \text{Maximize:} & Z & = & c_1x_1 & + & c_2x_2 & + \cdots + & c_nx_n & & \\
\\
\text{II.} & \text{subject to:} & & & a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1n}x_n & \leq & b_1 \\
& & & & a_{21}x_1 & + & a_{22}x_2 & + \cdots + & a_{2n}x_n & \leq & b_2 \\
& & & & \vdots & & \vdots & & \vdots & & \vdots \\
& & & & a_{m1}x_1 & + & a_{m2}x_2 & + \cdots + & a_{mn}x_n & \leq & b_m \\
\\
\text{III.} & \text{with:} & & & x_1 \geq 0\,, & & x_2 \geq 0\,, & \cdots & x_n \geq 0 & &
\end{array}
$$

The standard LP problem consists of three parts:

I. Z is called the *objective function*. It is a linear function of the *decision variables* $(x_1, x_2, \cdots, x_n)$. The constants $(c_1, c_2, \cdots, c_n)$ are called *cost coefficients*. The increase in Z for unit increase in x_k is c_k.

II. This part states the *linear constraints* of the problem. The matrix of coefficients $A = (a_{ij})$ is called the *constraint matrix*. In the standard LP problem, all elements of the *resource vector* $(b_1, b_2, \cdots, b_n)$ are assumed to be non-negative; that is, $b_j \geq 0$ for $1 \leq j \leq m$.

III. The final part of the standard LP problem is the *positivity condition* (or non-negativity condition) of the decision vector $(x_1, x_2, \cdots, x_n)$.

Any (x_1, x_2) satisfying just II and III (i.e. the linear constraints and positivity conditions) of an LP problem is called a *feasible solution*. In general, all feasible solutions will lie in a closed region of decision-space [the (x_1, x_2)-plane for the drug problem, $(x_1, \ldots, x_n)$ generally], called the

feasible region. The feasible region is non-empty in standard LP problems, since the origin (i.e. $x_1 = x_2 = \cdots = x_n = 0$) is always a feasible solution. (In non-standard LP problems the feasible region may be empty.) Any (x_1, x_2) not in the feasible region is called an *infeasible solution*. A feasible solution which maximizes the objective function is called an *optimal solution*. We usually denote it by (x_1^*, x_2^*) with corresponding $Z^* = \max Z$. Since the objective function is linear, we know from Chapter 1 that, the maximum and minimum of Z must lie on the boundary of the feasible region.

Needless to say there are many non-standard LP problems which deviate in one or more ways from the standard problem defined above. We shall discuss procedures for handling these non-standard problems in a later section. However, note that the drug problem is in standard form.

2 Graphical Solution of LP Problems

2.1 Graphical Solution of the Drug Problem

LP problems with two decision variables can be solved graphically. For the drug problem the feasible region is bounded by the five straight lines:

$$x_1 = 0\,, \quad x_2 = 0\,, \quad x_1 = 4\,, \quad 3x_1 + 2x_2 = 18\,, \quad 2x_2 = 12\,.$$

This region is shown in Figure 2.1.

The variation of Z over the feasible region can be seen by sketching a few contours of Z, $3x_1 + 5x_2 =$ constant. See Figure 2.2. The maximum value $Z^* = 36$ is given by the corner point $(2, 6)$ of the feasible region. Thus the optimal solution is $Z^* = 36$, $x_1^* = 2$, $x_2^* = 6$. We can confirm that $(2, 6)$ by comparing the slopes $0, -3/2$ of the boundary lines through $(2, 6)$ with the slope $-3/5$ of the level contours of Z, $0 > -3/5 > -3/2$.

An alternative method to find the optimal solution is to compare the values of Z at the corner points of the feasible region,

	FCP(x_1, x_2)	$Z = 3x_1 + 5x_2$
F_1	(0,0)	0
F_2	(4,0)	12
F_3	(4,3)	27
F_4^*	$(2, 6)^*$	36^*
F_5	(0,6)	30

2.2 Observations and Generalizations

The feasible region is a bounded convex polygon and the optimal solution is a *vertex* or *corner point* of the feasible region, i.e. a *feasible corner point*. A region R is *convex* if, for any points $\mathbf{x}_1$ and $\mathbf{x}_2$ in R, and scalar λ, $0 \leq \lambda \leq 1$, the point $\lambda\mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2$ lies in R. Graphically, the straight line segment joining $\mathbf{x}_1$ and $\mathbf{x}_2$ lies in the region R. As one might expect there are exceptions to this rule.

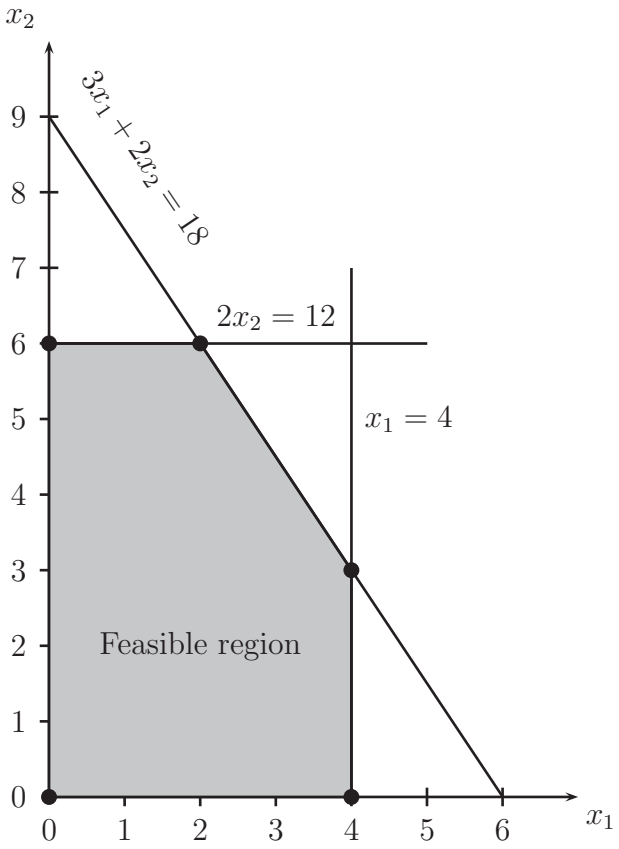


Figure 2.1: Feasible region of the drug problem.

1. The feasible region is non-empty in the standard LP problem (since $b_j \geq 0$ for $0 \leq j \leq m$) but may be empty in the non-standard problem. The problem is then *infeasible*. This occurs when the linear constraints are inconsistent. The problem is badly defined and there are no feasible solutions.

2. The feasible region may, in some circumstances, be unbounded. For example, consider the linear constraints:

$$\begin{aligned} -x_1 &+ x_2 \leq 4 \\ x_1 &- x_2 \leq 2 \end{aligned}$$

$$\text{with:} \quad x_1 \geq 0\,, \quad x_2 \geq 0\,.$$

   See Figure 2.3. If the feasible region is unbounded, then Z will also be unbounded (if the cost coefficients c_j are positive), i.e. $\max Z = \infty$. However, $\min Z$ may be finite.

3. There may be more than one optimal solution. For example, consider the drug problem with a new objective function: $Z = 6x_1 + 4x_2$. See Figure 2.4. The feasible corner points are

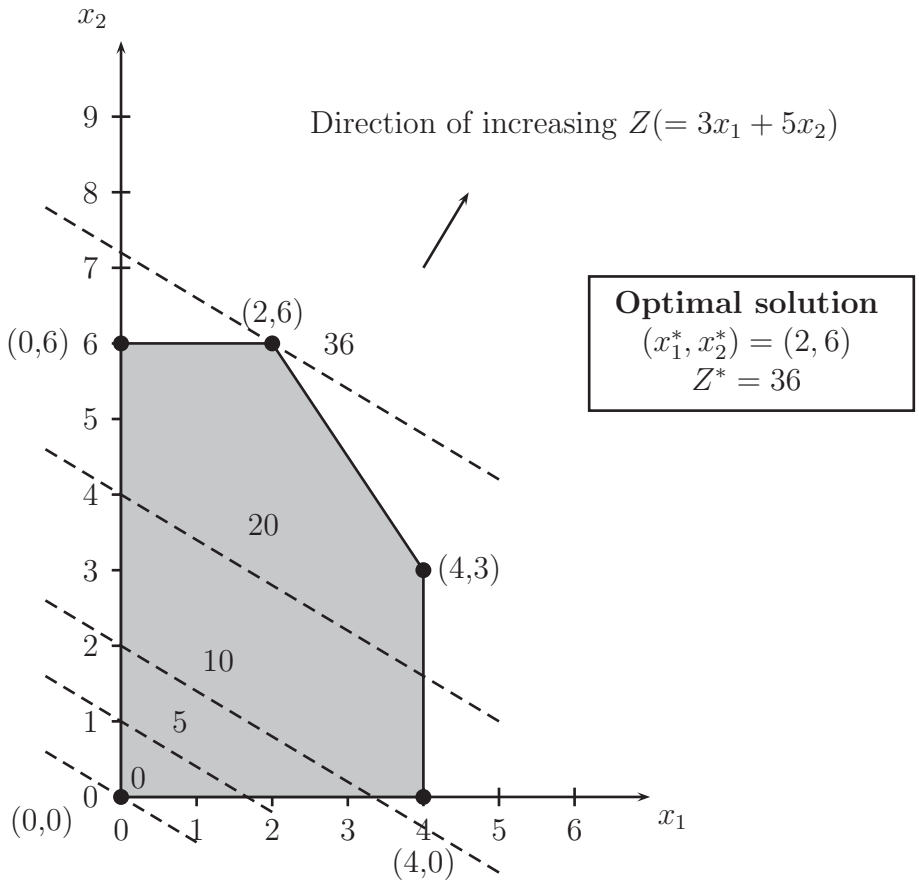


Figure 2.2: Level contours of the objective function Z in the drug problem.

	FCP(x_1, x_2)	$Z = 3x_1 + 5x_2$
F_1	(0,0)	0
F_2	(4,0)	24
F_3	(4,3)	36^*
F_4^*	$(2,6)^*$	36^*
F_5	(0,6)	24

All points of the line segment C_1C_2 are optimal with the optimal objective function value $Z^* = 36$. Thus, if there is more than one optimal solution there will be an infinite number of them. Geometrically, this condition can occur when the level contours of the objective function are parallel to a boundary line of the feasible region.

To represent the solution mathematically let $x_1 = t \geq 0$. The boundary line is given by equality in the second constraint $3x_1 + 2x_2 \leq 18$, which implies $x_2 = 9 - \frac{3}{2}x_1 = 9 - \frac{3}{2}t$. Since $x_2 \geq 0$ we must have $t \leq 6$. The first constraint, $x_1 \leq 4$, implies $t \leq 4$ and the third constraint, $2x_2 \leq 12$, implies $x_2 \leq 6$ or $t \geq 2$. Thus $x_1^* = t$, $x_2^* = 9 - \frac{3}{2}t$, where $2 \leq 6 \leq 4$.

Since an optimal solution is a corner point of the feasible region, it is clear that an algorithm to solve the standard LP problem need only consider *feasible corner points* (FCP's). Observe that

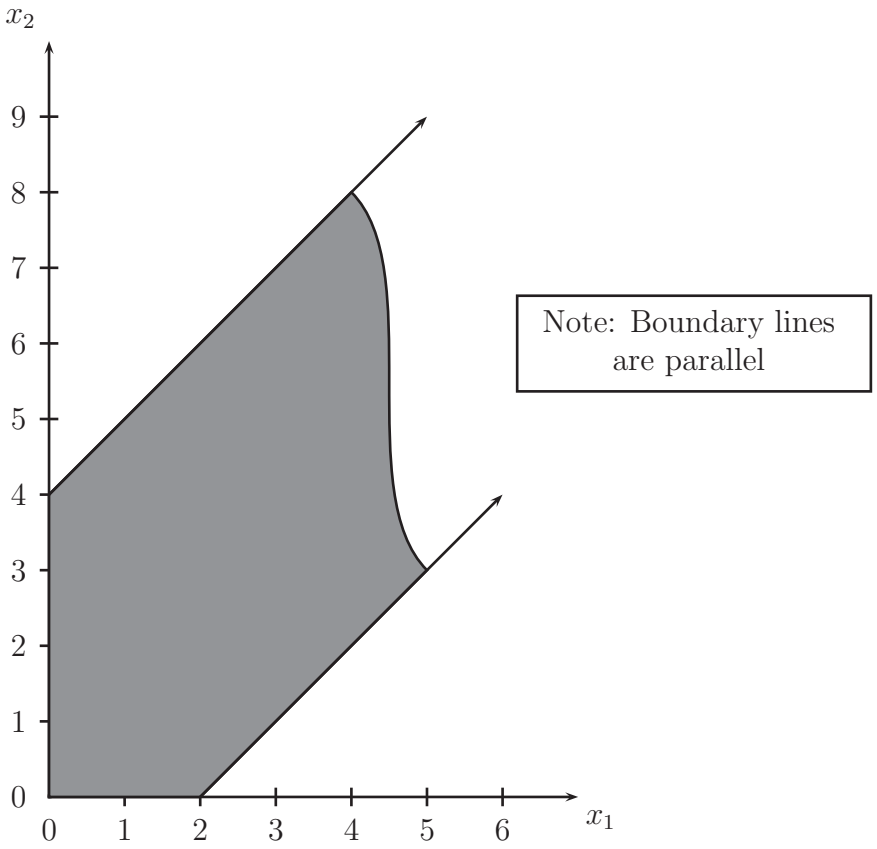


Figure 2.3: An unbounded feasible region.

in Figure 2.5 each FCP has exactly $n = 2$ adjacent FCP's. F_4^* has the unique property that its value of Z is bigger than the Z values of *all* its adjacent neighbours. Thus a procedure to solve the LP problem need not even consider all FCP's; it need only search for a FCP with the property that its value of Z is greater than the values of Z at the neighbouring FCP's. This FCP must then be the optimal solution. These observations are fundamental in the construction of the Simplex Algorithm.

3 The Simplex Algorithm

The graphical method discussed in §2.2 is fine for the case of $n = 2$ decision variables, but is not very useful if $n > 2$. The simplex algorithm is an algebraic method of solution which simulates a particular form of the graphical procedure for $n = 2$ and generalizes it in a simple way to $n > 2$.

In its broadest terms, the algorithm can be summarized as follows:

The Simplex Algorithm

I. *Initialization*: Start at an FCP solution.

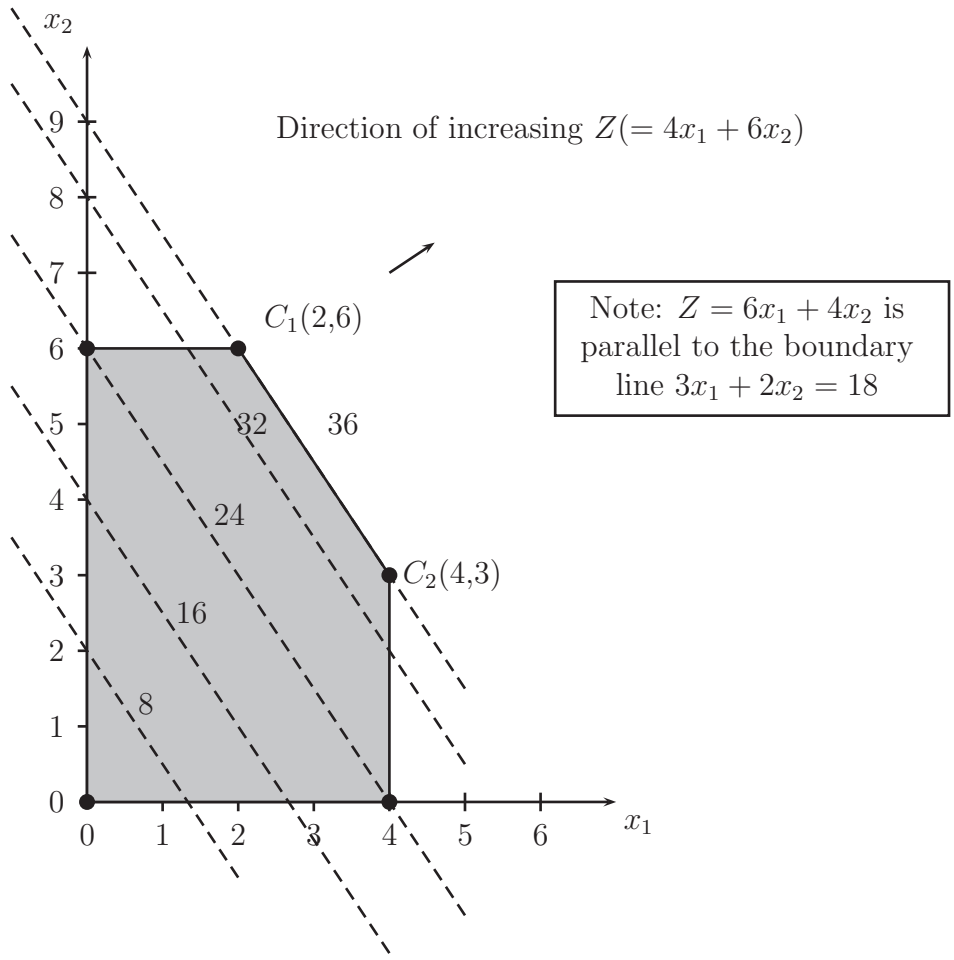


Figure 2.4: Multiple optimal solutions.

II. *Iteration Steps*: Move to an adjacent FCP with the best potential Z increase.

III. *Stopping Rule*: Stop at FCP* if its Z^* is greater than the Z values of all its adjacent FCP's.

For the drug problem the simplex algorithm might proceed as follows:

1. Start at FCP $F_1(0,0)$ with $Z = 0$. See Figure 2.5.

2. Move to $F_5(0,6)$ since Z increases at a faster rate along $x_1 = 0$ than along $x_2 = 0$ [towards $F_2(4,0)$]. Move to $F_4(2,6)$ with $Z = 36$ for the same reason (Z actually decreases going back to F_1).

3. Stop at F_4^* since $Z_5 < Z_4^* > Z_3$. $(Z_5 = 30\,, Z_3 = 27)$.

We now investigate the algebra of the simplex algorithm. In particular, how do we characterize FCP's algebraically?

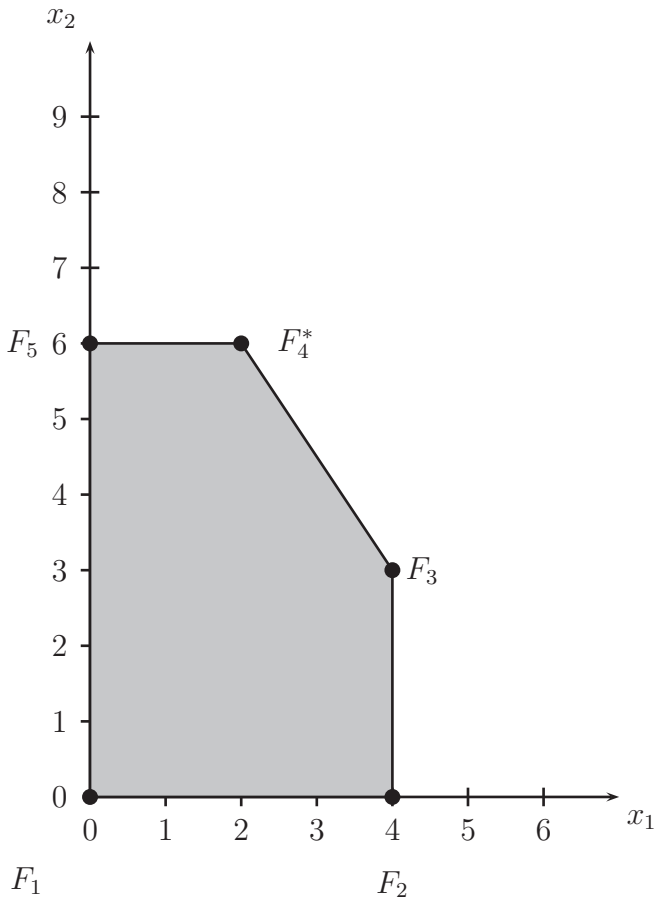


Figure 2.5: Feasible corner points of the drug problem.

3.1 Algebraic Representation of Corner Points

Consider the model drug problem. For each constraint introduce a new variable equal to the difference between the RHS and LHS of the constraint. Thus:

$$\begin{array}{llll} \text{for} & x_1 \leq 4 & \text{let} & x_3 = 4 - x_1 \quad ; \\ \text{for} & 3x_1 + 2x_2 \leq 18 & \text{let} & x_4 = 18 - 3x_1 - 2x_2 \quad ; \\ \text{for} & 2x_2 \leq 12 & \text{let} & x_5 = 12 - 2x_2 \quad . \end{array}$$

The constraints are then satisfied if, in terms of these new variables,

$$x_3 \geq 0\,, \qquad x_4 \geq 0\,, \qquad x_5 \geq 0\,.$$

The new variables (x_3, x_4, x_5) are called *slack variables*. They literally take up the slack in their corresponding inequalities.

The boundary lines of the feasible region are now given by the five equations

$$x_1 = 0\,, \quad x_2 = 0\,, \quad x_3 = 0\,, \quad x_4 = 0\,, \quad x_5 = 0\,.$$

A corner point is generated by the intersection of any pair of these boundary lines. The number of ways of selecting two of these five equations is ${}^5C_2 = 10$ and so there are a maximium of 10 corner

points for the drug problem. We say "maximum" because some may be degenerate, when three or more lines meet at the same point, and some may not strictly exist, when a pair of parallel lines "meet at infinity".

For the general standard LP problem, which has m constraints in n decision variables with n positivity contraints, a corner point lies at the intersection of n constraint boundaries. Introduce m slack variables $x_{n+1}, \ldots, x_{n+m}$, one for each of the m linear constraints. In terms of the decision variables and slack variables, the boundaries of the feasible region are given by the $n+m$ equations,

$$x_1 = 0\,, \quad \ldots\,, \quad x_n = 0\,, \quad x_{n+1} = 0\,, \quad \ldots\,, \quad x_{n+m} = 0\,.$$

Thus the maximum number of corner points is the number of ways of choosing n of these $n+m$ equations, i.e. ${}^{n+m}C_n$. This number increases vary rapidly with n and m; e.g. for 20 decision variables subject to 20 linear constraints ($n = m = 20$ — not a large problem) the maximum number of corner points is ${}^{40}C_{20} \approx 1.4 \times 10^{11}$.

The number of distinct corner points could actually be less than ${}^{n+m}C_n$ due to degeneracy. Furthermore, not all corner points are feasible. That is, some corner points may lie outside the feasible region. These are called *infeasible corner points* (ICP's). For the drug problem there are five FCP's, three ICP's and the remaining two ICP's do not strictly exist. The complete list is given in the following table:

	Corner Point	Defining Equations	$(x_1, x_2, x_3, x_4, x_5)$	
Feasible	F_1	$x_1 = x_2 = 0$	(0,0,4,18,12)	↑
	F_2	$x_2 = x_3 = 0$	(4,0,0,6,12)	
	F_3	$x_3 = x_4 = 0$	(4,3,0,0,6)	All $x_i \geq 0$
	F_4	$x_4 = x_5 = 0$	(2,6,2,0,0)	
	F_5	$x_5 = x_1 = 0$	(0,6,4,6,0)	↓
Infeasible	I_1	$x_1 = x_4 = 0$	$(0, 9, 4, 0, -6)$	↑
	I_2	$x_3 = x_5 = 0$	$(4, 6, 0, -6, 0)$	
	I_3	$x_2 = x_4 = 0$	$(6, 0, -2, 0, 12)$	Not all $x_i \geq 0$
	I_4	$x_1 = x_3 = 0$	No solution	
	I_5	$x_2 = x_5 = 0$	No solution	↓

The results in this table were obtained by simply solving the constraint equations of the drug problem for the remaining three non-zero variables, when two variables are selected to be zero. Using slack variables, the original constraint equations can be written as

$$\begin{aligned} x_1 \qquad + x_3 \qquad\qquad\quad &= \quad 4 \\ 3x_1 + 2x_2 \qquad + x_4 \qquad &= \quad 18 \\ 2x_2 \qquad\qquad + x_5 \quad &= \quad 12 \end{aligned}$$

Their geometrical interpretation is shown in the Figure 2.6.

For each corner point (feasible or infeasible), the two ($n = 2$) zero variables are called *non-basic variables*. The remaining three non-zero variables for each corner point are called *basic variables*. This is precisely the terminology used in Gauss-Jordan elimination. E.g. for F_3, (x_1, x_2, x_5) are basic variables, while (x_3, x_4) are non-basic variables.

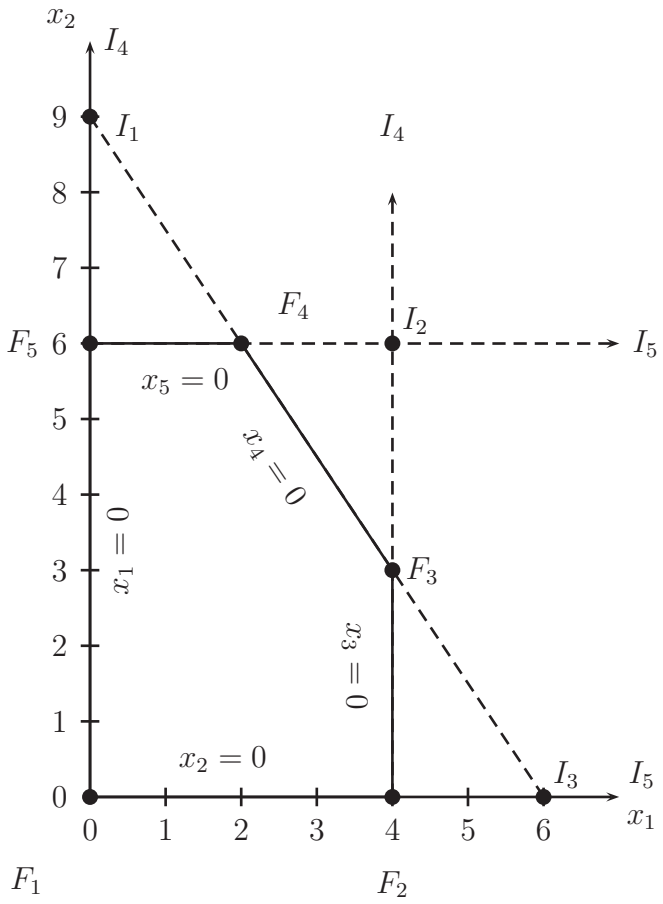


Figure 2.6: Geometric interpretation of corner points in the drug problem.

We introduce the important definition:

Adjacent Corner Points: Two corner points are *adjacent*, if they differ in precisely one basic variable.

For example, in the drug problem, F_5 has 6 adjacent corner points as seen by studying their basic variables:

$$F_5(x_2, x_3, x_4) \begin{cases} F_1(x_3, x_4, x_5) & \\ F_4(x_1, x_2, x_3) & \\ I_1(x_2, x_3, x_5) & \\ I_4(x_2, x_4, x_5) & \text{— at infinity} \\ I_5(x_1, x_3, x_4) & \text{— at infinity} \\ I_2(x_1, x_2, x_4) & \end{cases}$$

Only two of the adjacent corner points are feasible. Geometrically, a corner point C is adjacent to all corner points on the boundaries which pass through C.

Algebraically, we move from one corner point to an adjacent one by simply replacing one of its basic variables. The new variable in the basis is said *to enter the basis*. The old variable it replaces

is said *to leave the basis*. Thus in the drug problem,

$$F_1(0,0,4,18,12) \to F_5(0,6,4,6,0)\,,$$

i.e. x_2 enters the basis and x_5 leaves the basis;

$$F_5(0,6,4,6,0) \to F_4(2,6,2,0,0)\,,$$

i.e. x_1 enters the basis and x_4 leaves the basis;

$$F_4(2,6,2,0,0) \to F_3(4,3,0,0,6)\,,$$

i.e. x_5 enters the basis and x_3 leaves the basis. We would have stopped at F_4, where the maximum profit occurs, because the Z value at F_4 is greater than the Z values at the adjacent feasible corner points (F_5, F_3). How does Z change between FCP's?

3.2 Implementing the Algorithm

The above approach of moving from a FCP to an adjacent FCP with the largest potential increase in Z can now be translated into a computational algorithm. We consider each of the three main steps in turn.

I. Initialisation: Choose a feasible solution. This is easy for a problem in standard form: simply set the decision variables $(x_1, x_2, \ldots, x_n)$ to zero,

$$x_1 = 0\,, \qquad x_2 = 0\,, \qquad \ldots \qquad x_n = 0\,,$$

and set the slack variables equal to the right-hand sides (RHS's) of the constraints,

$$x_{n+1} = b_1\,, \qquad x_{n+2} = b_2\,, \qquad \ldots \qquad x_{n+m} = b_m\,.$$

Then write the problem in the form:

$$\begin{array}{rcrcrcrcrcrcr}
Z & - & c_1x_1 & - & c_2x_2 & - & \ldots & - & c_nx_n & & & = & 0 \\
 & & a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & + & x_{n+1} & = & b_1 \\
 & & \vdots & & \vdots & & & & \vdots & & & & \vdots \\
 & & a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & + & x_{n+m} & = & b_m
\end{array}$$

or in *tableau* form:

Z	x_1	x_2	$\ldots$	x_n	x_{n+1}	x_{n+2}	$\ldots$	x_{n+m}	RHS
1	$-c_1$	$-c_2$	$\ldots$	$-c_n$	0	0	$\ldots$	0	0
0	a_{11}	a_{12}	$\ldots$	a_{1n}	1	0	$\ldots$	0	b_1
0	a_{21}	a_{22}	$\ldots$	a_{2n}	0	1	$\ldots$	0	b_2
$\vdots$				$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
0	a_{m1}	a_{m2}	$\ldots$	a_{mn}	0	0	$\ldots$	1	b_m

with the decision variables $x_j \geq 0$, $j = 1, \ldots, n$, and the slack variables $x_{n+i} \geq 0$, $i = 1, \ldots, m$. For the drug problem,

$$\begin{array}{rcrcrcrcrcr} Z & - & 3x_1 & - & 5x_2 & & & & & = & 0 \\ & & x_1 & & & + & x_3 & & & = & 4 \\ & & 3x_1 & + & 2x_2 & & & + & x_4 & = & 18 \\ & & & & 2x_2 & & & & + \; x_5 & = & 12 \end{array}$$

and the tableau is

Z	x_1	x_2	x_3	x_4	x_5	RHS
1	-3	-5	0	0	0	0
0	1	0	1	0	0	4
0	3	2	0	1	0	18
0	0	2	0	0	1	12

The initial basic variables are the slack variables,

$$x_3 = 4\,, \qquad x_4 = 18\,, \qquad x_5 = 12\,.$$

The initial non-basic variables are the decision variables,

$$x_1 = 0\,, \qquad x_2 = 0\,.$$

II. Iteration Steps: *Iteration 1*: We have to choose a new basic variable; graphically, we can choose to go to vertex (0,6) or vertex (4,0). We choose to bring x_2 into the basis, since its cost coefficient c_2 $(= 5)$ is larger than c_1, giving a potentially larger increase in Z. We will call this *the largest-coefficient rule*. The actual increase may be larger for x_1 than for x_2, but we have no way of knowing this without performing both the steps. Choosing x_1 would not actually prevent the method from finding the optimal solution, it would only require more iterations.

Which basic variable must leave the basis: x_3, x_4 or x_5 ? To decide this we must determine which of the basic variables (currently the slack variables) x_3, x_4 or x_5 tends to zero *first*, as x_2 increases from zero, while the other non-basic variable x_1 is held fixed and equal to zero. If x_2 were increased further, this variable would become negative, violating feasibility. Thus consider the x_2 column (the column of the entering variable), the basic variables and the right-hand sides of the constraint equations. Since $x_1 = 0$ is fixed, the constraints equations reduce to

$$\begin{array}{rcrcrcr} & x_3 & & & & = & 4 \\ 2x_2 & & + & x_4 & & = & 18 \\ 2x_2 & & & & + \; x_5 & = & 12 \end{array}$$

which imply

$$\begin{aligned} x_3 &= 4 \\ x_4 &= 18 - 2x_2 \\ x_5 &= 12 - 2x_2\,. \end{aligned}$$

Thus x_3 is unaffected as x_2 increases, $x_4 \to 0$ as $x_2 \to 9$ and $x_5 \to 0$ as $x_2 \to 6$. Thus x_5 must leave the basis, as it goes to zero first. This corresponds to the minimum of the ratios b_i/a_{i2}, $i = 1, \ldots, m$, for which $a_{i2} > 0$. In tableau form

Basis	Z	x_1	x_2	x_3	x_4	x_5	RHS	Ratio b_i/a_{i2}	
Z	1	-3	-5	0	0	0	0	–	
x_3	0	1	0	1	0	0	4	–	
x_4	0	3	2	0	1	0	18	9	
x_5	0	0	2	0	0	1	12	6	$\leftarrow$ min

Row 1 contains the objective function expressed in terms of the current non-basic variables (x_1, x_2). The current basic variables (x_3, x_4, x_5), together with Z, are listed in the "basis" column.

We usually omit the basis column. In order to find the new basic feasible solution, perform Gauss-Jordan elimination, pivoting on the element in the column of the entering variable and in the row of the leaving variable. This reduces the pivot element to 1 and the elements above and below it in the same column to zero. The values of the basic variables, including the new one, can then be easily read-off since the non-basic variables are zero. We obtain the following tableau, apart from the last column,

Basis	Z	x_1	x_2	x_3	x_4	x_5	RHS	Ratio b_i/a_{i1}	
Z	1	-3	0	0	0	5/2	30	–	
x_3	0	1	0	1	0	0	4	4	
x_4	0	3	0	0	1	-1	6	2	$\leftarrow$ min
x_2	0	0	1	0	0	1/2	6	–	

Row 1 now contains the objective function expressed in terms of the new non-basic variables (x_1, x_5).

Iteration 2: The entering basic variable is x_1, since it has the largest positive modified cost coefficient $\overline{c}_i$, i.e. the most negative coefficient of the non-basic variables in the Z row. The minimum of the ratios b_i/a_{i1}, $i = 1, \ldots, m$, is 2 for $i = 4$, so x_4 leaves the basis. Pivoting on the element in the x_1 column and the x_4 row,

Basis	Z	x_1	x_2	x_3	x_4	x_5	RHS	Ratio
Z	1	0	0	0	1	3/2	36	
x_3	0	0	0	1	$-1/3$	1/3	2	
x_1	0	1	0	0	1/3	$-1/3$	2	
x_2	0	0	1	0	0	1/2	6	

Row 1 now contains the objective function expressed in terms of the new non-basic variables (x_4, x_5), i.e.

$$Z = -x_4 - \tfrac{3}{2}x_5 + 36\,.$$

To bring a non-basic variable into the basis, we must increase it from 0. Whichever of the two non-basic variables, x_4 or x_5, we bring into the basis, Z will *decrease*, i.e. we cannot move to an adjacent feasible corner point without *decreasing* the profit. Thus the current solution is optimal. This gives the

III. Stopping Rule: When all the coefficients in row 1 are non-negative, i.e. $-\overline{c}_j \geq 0$, the tableau is optimal and the optimal solution is given by the RHS column.

3.3 The Complete Simplex Solution of the Drug Problem

We combine all tableaus for the model drug problem:

	Basis	Z	x_1	x_2	x_3	x_4	x_5	RHS	Ratio
	Z	1	-3	-5	0	0	0	0	–
F_1	x_3	0	1	0	1	0	0	4	–
	x_4	0	3	2	0	1	0	18	9
	x_5	0	0	2	0	0	1	12	6
	Z	1	-3	0	0	0	5/2	30	–
F_5	x_3	0	1	0	1	0	0	4	4
	x_4	0	3	0	0	1	-1	6	2
	x_2	0	0	1	0	0	1/2	6	–
	Z	1	0	0	0	1	3/2	36	Optimal
F_4^*	x_3^*	0	0	0	1	$-1/3$	1/3	2	
	x_1^*	0	1	0	0	1/3	$-1/3$	2	
	x_2^*	0	0	1	0	0	1/2	6	

Since the coefficients of non-basic variables in the Z row are ≥ 0, the tableau is optimal and the algorithm stops. The optimal solution is

$$Z^* = 36\,, \qquad x_1^* = 2\,, \qquad x_2^* = 6\,,$$

since the non-basic variables are zero,

$$x_4^* = x_5^* = 0$$

and the basic variables,

$$x_3^* = 2\,, x_1^* = 2\,, x_2^* = 6\,,$$

can be read off from rows 2, 3, 4, respectively, of the RHS column.

3.4 Further Observations on the Simplex Algorithm

Tie-Breaking Rules

1. If at some iteration of the simplex algorithm, row 1 of the tableau looked like, say

Basis	Z	x_1	x_2	x_3	x_4	x_5	RHS
Z	1	0	-3	-3	0	0	20

   there would be an ambiguity in the choice of variable entering the basis. In practice, it makes no difference which is chosen, though one choice, which cannot be predicted in advance, may require fewer iterations.

2. It is also possible to have a tie for the variable leaving the basis. For example, consider the drug problem, with the second constraint modified to $3x_1 + 2x_2 \leq 12$.

$$\begin{array}{rrcrcrl} \text{Maximise:} & Z & = & 3x_1 & + & 5x_2 & \\ \text{subject to:} & & & x_1 & & & \leq \quad 4 \\ & & & 3x_1 & + & 2x_2 & \leq \quad 12 \\ & & & & & 2x_2 & \leq \quad 12 \\ \text{with:} & & & x_1,\ x_2 \geq 0\,. & & & \end{array}$$

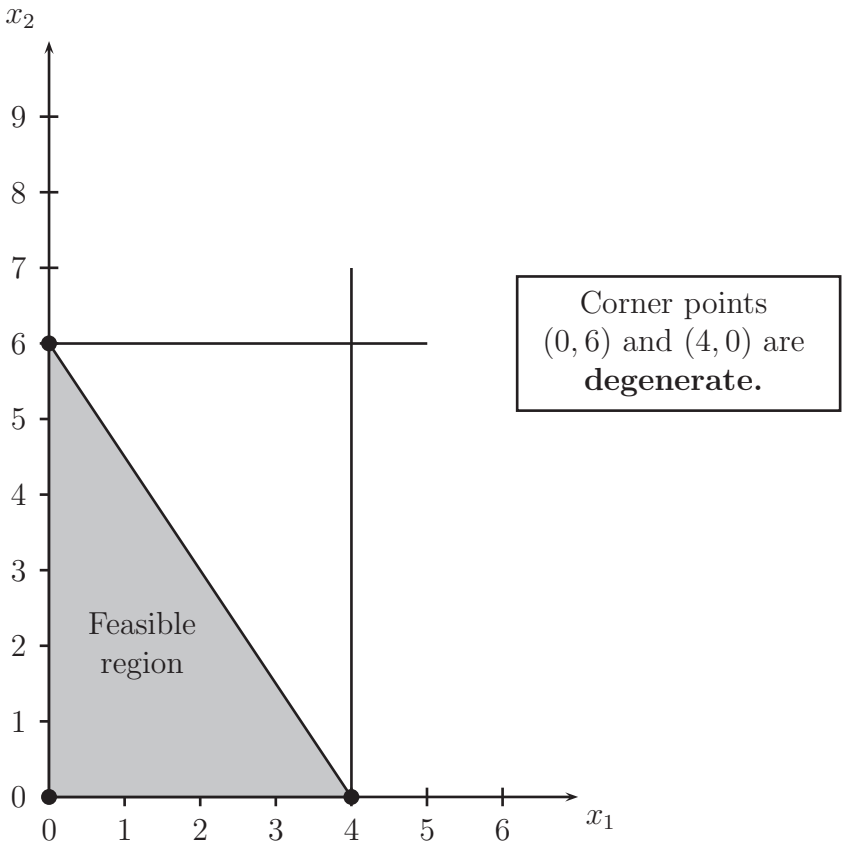


Figure 2.7: Geometric interpretation of a tie for the leaving variable.

The feasible region is illustrated in Figure 2.7.

The simplex tableaus will be:

Basis	Z	x_1	x_2	x_3	x_4	x_5	RHS	Ratio	
Z	1	-3	-5	0	0	0	0	–	
x_3	0	1	0	1	0	0	4	–	
x_4	0	3	2	0	1	0	12	6	← choose
x_5	0	0	2	0	0	1	12	6	← either
Z^*	1	9/2	0	0	5/2	0	30	optimal	
x_3^*	0	1	0	1	0	0	4		
x_2^*	0	3/2	1	0	1/2	0	6		
x_5^*	0	-3	0	0	-1	1	0		← degeneracy

Note : The basic variable $x_5^* = 0$. A *zero basic variable* indicates degeneracy.

It is *theoretically possible* that your choice may lead to an endless loop or *cycling*, but this is so unlikely to occur in practice that the tie can be broken arbitrarily. One rule to prevent

cycling is *the smallest index rule*: to break a tie always choose the variable with the smallest index.

No Variable Leaving the Basis

Consider the following problem:

$$\begin{array}{rrcrcrcl} \text{Maximise:} & Z & = & 3x_1 & + & 5x_2 & & \\ \text{subject to:} & & & -x_1 & + & x_2 & \leq & 4 \\ & & & x_1 & - & x_2 & \leq & 2 \\ \text{with:} & & & x_1,\ x_2 \geq 0. & & & & \end{array}$$

The constraints here correspond to the unbounded feasible region discussed in §2.2.

Introducing slack variables x_3 and x_4 the simplex tableau becomes:

Basis	Z	x_1	x_2	x_3	x_4	RHS	Ratio	
Z	1	-3	-5	0	0	0	–	
x_3	0	-1	1	1	0	4	4	
x_4	0	1	-1	0	1	2	–	
Z	1	-8	0	5	0	20	–	No variable
x_2	0	-1	1	1	0	4	–	to leave
x_4	0	0	0	1	1	6	–	the basis

The conclusion from this condition is that the optimal solution is unbounded, since the *entering variable* and hence Z can be increased indefinitely.

Multiple Optimal Solutions

We observed earlier that if the objective function contour is parallel to a boundary line of the feasible region, there may be an infinite number of optimal solutions, occurring at all points of that boundary line. To see how the simplex algorithm detects this condition, consider:

$$\begin{array}{rrcrcrcl} \text{Maximise:} & Z & = & 6x_1 & + & 4x_2 & & \\ \text{subject to:} & & & x_1 & & & \leq & 4 \\ & & & 3x_1 & + & 2x_2 & \leq & 18 \\ & & & & & 2x_2 & \leq & 12 \\ \text{with:} & & & x_1,\ x_2 \geq 0\,. & & & & \end{array}$$

Note that the level contours of $Z = 6x_1 + 4x_2$ are parallel to $3x_1 + 2x_2 = 18$. The simplex tableaux are given below:

Basis	Z	x_1	x_2	x_3	x_4	x_5	RHS	Ratio
Z	1	-6	-4	0	0	0	0	–
x_3	0	1	0	1	0	0	4	4
x_4	0	3	2	0	1	0	18	6
x_5	0	0	2	0	0	1	12	–
Z	1	0	-4	6	0	0	24	–
x_1	0	1	0	1	0	0	4	–
x_4	0	0	2	-3	1	0	6	3
x_5	0	0	2	0	0	0	12	6
Z^*	1	0	0	0	2	0	36	optimal
x_1	0	1	0	1	0	0	4	
x_2	0	0	1	$-3/2$	$1/2$	0	3	
x_5	0	0	0	3	-1	1	6	

The zero coefficient of the non-basic variable x_3 indicates multiple optimal solutions. The optimal solution is $Z^* = 36$. Row 1 of the final tableau indicates that

$$Z = 36 - 2x_4$$

is *independent* of the other non-basic variable x_3. Hence $Z = 36$ along *all* points of the edge $x_4 = 0$ and there is an infinite number of such points, i.e. changing the value of x_3 does not affect Z. To express the solutions explicitly write down the constraint equationsof the optimal tableau, which contain the non-basic variables associated with modified cost coefficients equal to zero. To these non-basic variables assign arbitrary values and set the remaining non-basic variables to zero. Thus, in the example, $x_3^* = t$, $x_4^* = 0$ and the constraints reduce to

$$x_1^* + t = 4\,, \qquad x_2^* - \tfrac{3}{2}t = 3\,, \qquad 3t + x_5^* = 6\,.$$

Solving for the basic variables gives

$$x_1^* = 4 - t\,, \qquad x_2^* = 3 + \tfrac{3}{2}t\,, \qquad x_5^* = 6 - 3t\,,$$

where $0 \le t \le 2$, since $x_1^* \ge 0$, $x_2^* \ge 0$, $x_3^* \ge 0$, $x_5^* \ge 0$.

4 Summary of the Simplex Algorithm for Standard LP Problems

1. For each less-than-or-equal-to constraint, introduce a slack variable.

2. As the initial feasible solution, choose all decision variables $x_j = 0$, $j = 1, \ldots, n$, and choose each slack variable equal to the corresponding RHS, i.e. $x_{n+i} = b_i$, $i = 1, \ldots, m$. Write the problem in tableau form.

3. At each iteration:

   (a) The entering basic variable x_j has the most negative coefficient in the Z row (i.e. most negative $-\overline{c}_j$, $j = 1, \ldots, n$.

(b) The leaving basic variable x_i corresponds to the row i_0 such that

$$\min_i \frac{b_i}{a_{ij}} = \frac{b_{i_0}}{a_{i_0 j}}$$

for $a_{ij} > 0$, where j is the index corresponding to the entering basic variable.

(c) Use Gauss-Jordan elimination to reduce $a_{i_0 j} = 1$, $a_{ij} = 0$, $i \neq i_0$.

4. Repeat step 3 above until all modified cost coefficients, $-\overline{c}_j$, in the Z row are ≥ 0. Then stop, and read off the optimal solution.

Empirical evidence indicates that the simplex algorithm usually takes about $1.5m - 2m$ iterations, where m is the number of constraints; i.e. the number of iterations depends only weakly (perhaps logarithmically) on the number n of variables. Various attempts have been made to explain this "practical convergence", which is the basis for the great success of the simplex algorithm. The difficulty in explaining the convergence can be shown by the Klee-Minty LP problem:

Maximise:	Z =	$\sum_{j=1}^{n} 10^{n-j} x_j$		
subject to:		$2\sum_{j=1}^{i-1} 10^{i-j} x_j + x_i$	$\leq$ 100^{i-1}	$(i = 1, 2, \ldots, n)$
with:		$x_j \geq 0$		$(j = 1, 2, \ldots, n)$

The simplex algorithm takes a staggering $2^n - 1$ iterations to solve this problem if the largest-coefficient rule is used.

The number of iterations is not the only consideration when considering the efficiency of the simplex algorithm: the time per iteration must also be considered. Practical implementations of the algorithm use the *Revised Simplex Algorithm*, do not use the largest-coefficient rule (it's too slow) and must allow for roundoff error — see Chvátal (1983). For some large scale real examples see Winston (1995), pp.2–3, and the references therein.

5 Adapting the Simplex Algorithm to Non-Standard LP Problems

We consider in this section how to reformulate non-standard LP problems into the standard format of §2.1, where possible, and modifications, where this is not possible.

Minimizing the Objective Function

To minimize $Z = \sum_i c_i x_i$, define a new objective function $\hat{Z} = -Z$. Then

$$\min Z \equiv -\max \hat{Z}\,.$$

Negative Resource Elements

In the standard LP problem we required all resource elements b_j $(i \leq j \leq m)$ to be non-negative. Suppose $b_j = -b < 0$, i.e.

$$\mathbf{a}^T\mathbf{x} = a_1x_1 + a_2x_2 + \cdots + a_nx_n \leq -b\,.$$

This is equivalent to

$$-\mathbf{a}^T\mathbf{x} = -a_1x_1 - a_2x_2 - \cdots - a_nx_n \geq b\,.$$

Thus we can assume that resource elements are always non-negative, if we can modify the simplex algorithm to include greater-than-or-equal-to results.

Greater-Than-Or-Equal-To Constraints

If the LP problem has a constraint of the form

$$\mathbf{a}^T\mathbf{x} = a_1x_1 + a_2x_2 + \cdots + a_nx_n \geq b \geq 0\,,$$

introduce a *surplus variable* x_{n+1} such that

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n - x_{n+1} = b\,, \qquad x_{n+1} \geq 0\,.$$

Negative Decision Variable

If $x_k \leq 0$ introduce a new variable $\hat{x}_k = -x_k$. Then $x_k \leq 0$ is equivalent to $\hat{x}_k \geq 0$.

Unrestricted Decision Variable

If x_k is unrestricted in sign, introduce *two* new variables, $\hat{x}_k \geq 0$ and $\hat{\hat{x}}_k \geq 0$, and let $x_k = \hat{x}_k - \hat{\hat{x}}_k$.

5.1 Finding an Initial FCP Solution

In the standard LP problem:

Maximise:	Z	$=$	$\mathbf{c}^T\mathbf{x}$
subject to:	$\mathbf{A}\mathbf{x}$	$\leq$	$\mathbf{b}$
with:	$\mathbf{x}$	$\geq$	$\mathbf{0}\,.$

where $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{b} = (b_1, \ldots, b_m) \geq \mathbf{0}$, we introduced slack variables $(x_{n+1}, \ldots, x_{n+m})$, one into each constraint. We also noted that an initial FCP could easily be constructed by choosing the decision variables as the non-basic variables and the slack variables as the basic variables:

$$x_j = 0\,, \quad 1 \leq i \leq n\,; \qquad x_{n+i} = b_i\,, \quad 1 \leq j \leq m\,.$$

This is an FCP only because all $b_i \geq 0$ by assumption.

There are two non-standard cases when this initialization procedure will not work:

1. *Equality constraints*: $\mathbf{a}^T\mathbf{x} = b$
2. *Greater-than-or-equal-to constraints*: $\mathbf{a}^T\mathbf{x} \geq b$

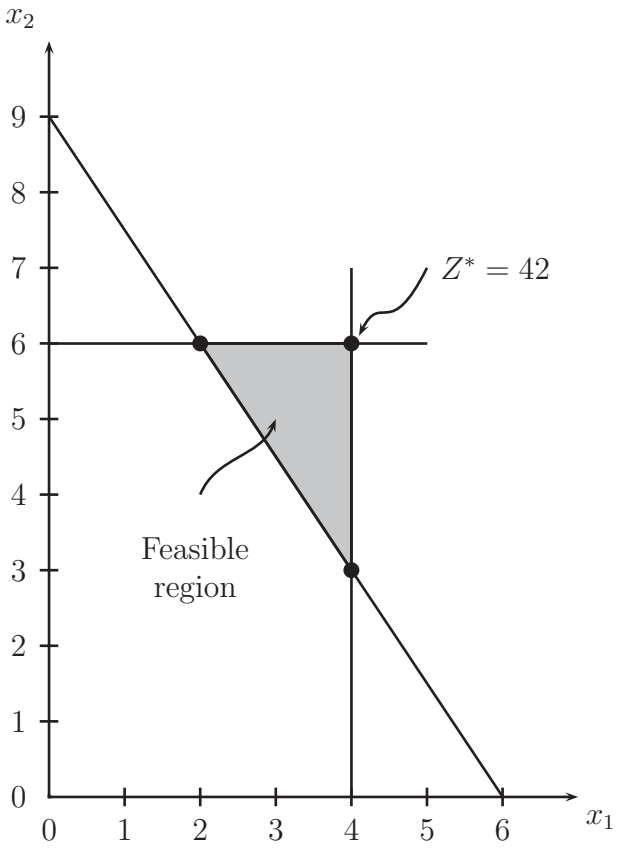


Figure 2.8: The feasible region of a non-standard LP problem.

Case (2) reduces to the case (1), when a surplus variable is introduced.

As an example, consider

$$\begin{array}{rrcrcrcr} \text{Maximise:} & Z & = & 3x_1 & + & 5x_2 & & \\ \text{subject to:} & & & x_1 & & & \leq & 4 \\ & & & 3x_1 & + & 2x_2 & \geq & 18 \\ & & & & & 2x_2 & \leq & 12 \\ \text{with:} & & & x_1,\ x_2 \geq 0. & & & & \end{array}$$

Introduce slack variables x_3, x_5 and a surplus variable x_4. The constraints become

$$\begin{array}{rcrcrcrcr} x_1 & & & + & x_3 & & & = & 4 \\ 3x_1 & + & 2x_2 & & & - & x_4 & = & 18 \\ & & 2x_2 & & & & + \ x_5 & = & 12 \end{array}$$

It is by no means an easy matter to find an initial FCP in non-standard cases. The initial CP solution $(0, 0, 4, -18, 12)$ is clearly *not* feasible since $x_4 = -18 < 0$. To get an initial CP, which is feasible, we introduce a new *artificial* variable into each constraint without a slack variable, i.e.

into each $=$ or $\geq$ contraint. Thus, in the example, introduce the artificial variable $\overline{x}_6 \geq 0$. The constraint equations become

$$\begin{array}{rcrcrcrcrcr} x_1 & & & + & x_3 & & & & & = & 4 \\ 3x_1 & + & 2x_2 & & & - & x_4 & & + \ \overline{x}_6 & = & 18 \\ & & 2x_2 & & & & + \ x_5 & & & = & 12 \end{array}$$

An initial FCP is

$$x_1 = 0\,, \quad x_2 = 0\,, \quad x_3 = 4\,, \quad x_4 = 0\,, \quad x_5 = 12\,, \quad \overline{x}_6 = 18\,;$$

i.e. the decision and surplus variables are non-basic and are set equal to zero, the slack and artificial variables are basic and are set equal to the RHS's of the constraints.

Unfortunately, a feasible solution of the problem with artificial variables is a feasible solution of the original problem *if and only if the artificial variables are zero.* In the example, the second constraint gives

$$3x_1 + 2x_2 - x_4 = 18 - \overline{x}_6 < 18\,,$$

if $\overline{x}_6 > 0$, not

$$3x_1 + 2x_2 - x_4 = 18\,.$$

Thus we must force the artificial variables to zero. We shall discuss two methods of doing so: the *two-phase simplex method*, which is usually preferred, and the alternative *big-M method*.

The Two-Phase Simplex Algorithm

We have seen that we can find an initial basic feasible solution to the non-standard LP problem:

$$\begin{array}{rrcrcrcr} \text{Maximise:} & Z & = & 3x_1 & + & 5x_2 & & \\ \text{subject to:} & & & x_1 & & & \leq & 4 \\ & & & 3x_1 & + & 2x_2 & \geq & 18 \\ & & & & & 2x_2 & \leq & 12 \\ \text{with:} & & & x_1,\ x_2 \geq 0 & & & & \end{array}$$

by casting it into the following form:

$$\begin{array}{rrcrcrcrcrcrcr} \text{Maximise:} & Z & = & 3x_1 & + & 5x_2 & & & & & & \\ \text{subject to:} & & & x_1 & & & + & x_3 & & & & = & 4 \\ & & & 3x_1 & + & 2x_2 & & - \ x_4 & & + \ \overline{x}_6 & & = & 18 \\ & & & & & 2x_2 & & & + \ x_5 & & & = & 12 \\ \text{with:} & & & & & & x_i,\ \overline{x}_6 \geq 0. & & & & & & \end{array}$$

To find a basic feasible solution to the original problem maximize

$$W = -\sum_{\text{artificial}} \overline{x}_k \leq 0\,.$$

If $\max W = 0$, then all artificial variables are zero, $\overline{x}_k = 0$, and we can drop W and $\overline{x}_k$. *If* $\max W < 0$*, there is no basic feasible solution to the original problem.*

We thus apply the simplex algorithm to maximize W, *updating Z as we go.*

The first step in phase I of the two-phase method is to express W in terms of *non-basic* variables. This is done by adding together all rows in which artificial variables appear, and then substituting this sum into W. In the initial tableau replace row-W by row-$W - \sum_k$ row-k, the sum being over all rows containing an artificial variable.

Example of the Two-Phase Simplex Method

Basis	W	Z	x_1	x_2	x_3	x_4	x_5	$\overline{x}_6$	RHS	Ratio
W	1	0	-3	-2	0	1	0	0	-18	–
Z	0	1	-3	-5	0	0	0	0	0	–
x_3	0	0	1	0	1	0	0	0	4	4
$\overline{x}_6$	0	0	3	2	0	-1	0	1	18	6
x_5	0	0	0	2	0	0	1	0	12	–
W	1	0	0	-2	3	1	0	0	-6	–
Z	0	1	0	-5	3	0	0	0	12	–
x_1	0	0	1	0	1	0	0	0	4	–
$\overline{x}_6$	0	0	0	2	-3	-1	0	1	6	3
x_5	0	0	0	2	0	0	1	0	12	6
W	1	0	0	0	0	0	0	1	0	–
Z	0	1	0	0	$-9/2$	$-5/2$	0	$5/2$	27	–
x_1	0	0	1	0	1	0	0	0	4	4
x_2	0	0	0	1	$-3/2$	$-1/2$	0	$1/2$	3	–
x_5	0	0	0	0	3	1	1	-1	6	2

This is the end of phase 1 as there is no variable to enter the basis.
Since $\max W = 0$, drop W row and artificial variable column.

Z	1	0	0	0	-1	$3/2$		36	–
x_1	0	1	0	0	$-1/3$	$-1/3$		2	–
x_2	0	0	1	0	0	$1/2$		6	–
x_3	0	0	0	1	$1/3$	$-1/3$		2	6
Z	1	0	0	3	0	$5/2$		42	Optimal
x_1	0	1	0	1	0	0		4	
x_2	0	0	1	0	0	$1/2$		6	
x_4	0	0	0	3	1	1		6	

Since all coefficients in the Z row are ≥ 0, this tableau is optimal and the optimal solution is

$$Z^* = 42\,, \qquad x_1^* = 4\,, \qquad x_2^* = 6\,.$$

($x_3^* = 0$, $x_4^* = 6$, $x_5^* = 0$, $\overline{x}_6^* = 0$.)

5.2 The Big-M Method

If $Z = \sum_j c_j x_j$ is the original objective function, then in the big-M method, introduce a new objective function

$$\overline{Z} = \sum_j c_j x_j - M \sum_{\text{artificial}} \overline{x}_k \, .$$

Here M is some (very big) number. We now try to max $\overline{Z}$. It is clear that, as M is so big, this must entail setting $\overline{x}_k = 0$, as required. Ensure that the $\overline{Z}$ row of the initial tableau is expressed in terms of non-basic variables by subtracting from it all rows containing artificial variable entries.

The big-M method requires only one pass through the simplex algorithm, while the two-phase method requires two passes, although this does not necessarily imply fewer iterations. The choice of M, particularly in computer implementations, is not always straightforward. Consider:

Example of the Big-M Method

$$\begin{array}{rrcrcrcr}
\text{Maximise:} & Z & = & 3x_1 & + & 5x_2 & & \\
\text{subject to:} & & & x_1 & & & \leq & 4 \\
& & & 3x_1 & + & 2x_2 & \geq & 18 \\
& & & & & 2x_2 & \leq & 12 \\
\text{with:} & & & x_1, \; x_2 \geq 0. & & & &
\end{array}$$

Introduce slack variables x_3 and x_5, a surplus variable x_4 and an artificial variable $\overline{x}_6$.

Basis	$\overline{Z}$	x_1	x_2	x_3	x_4	x_5	$\overline{x}_6$	RHS	Ratio
$\overline{Z}$	1	-3	-5	0	0	0	M	0	–
x_3	0	1	0	1	0	0	0	4	4
$\overline{x}_6$	0	3	2	0	-1	0	1	18	6
x_5	0	0	2	0	0	1	0	12	–
$\overline{Z}$	1	$-3-3M$	$-5-2M$	0	M	0	0	$-18M$	–
x_3	0	1	0	1	0	0	0	4	4
$\overline{x}_6$	0	3	2	0	-1	0	1	18	6
x_5	0	0	2	0	0	1	0	12	–
$\overline{Z}$	1	0	$-5-2M$	$3+3M$	M	0	0	$12-6M$	–
x_1	0	1	0	1	0	0	0	4	–
$\overline{x}_6$	0	0	2	-3	-1	0	1	6	3
x_5	0	0	2	0	0	1	0	12	6
$\overline{Z}$	1	0	0	$-9/2$	$-5/2$	0	$5/2+M$	27	–
x_1	0	1	0	1	0	0	0	4	4
x_2	0	0	1	$-3/2$	$-1/2$	0	$1/2$	3	–
x_5	0	0	0	3	1	1	-1	6	2
$\overline{Z}$	1	0	0	0	-1	$3/2$	$1+M$	36	–
x_1	0	1	0	0	$-1/3$	$-1/3$	$1/3$	2	–
x_2	0	0	1	0	0	$1/2$	0	6	–
x_5	0	0	0	1	$1/3$	$1/3$	$-1/3$	2	6
$\overline{Z}$	1	0	0	3	0	$5/2$	M	42	Optimal
x_1	0	1	0	1	0	0	0	4	
x_2	0	0	1	0	0	$1/2$	0	6	
x_5	0	0	0	3	1	1	-1	6	

The optimal solution is

$$Z^* = \overline{Z}^* = 42\,, \qquad x_1^* = 4\,, \qquad x_2^* = 6\,.$$

($x_3^* = 0$, $x_4^* = 6$, $x_5^* = 0$, $\overline{x}_6^* = 0$.)

5.3 Summary of Non-Standard LP Problems

Each of the problems mentioned earlier can be overcome by the introduction of an artificial variable $\bar{x}_{n+k}$, $k = 1, 2, \cdots$etc., where $\bar{x}_{n+k} \geq 0$.

1. *Equality Constraint*:

$$\sum_{j=1}^{n} a_{ij}x_j + \bar{x}_{n+1} = b_i$$

   $\bar{x}_{n+1} = b_i$ is feasible (with $x_j = 0$, $j = 1, \cdots, n$).

2. *Greater-Than-Or-Equal-To Constraint*:

$$\sum_{j=1}^{n} a_{ij}x_j - x_{n+1} + \bar{x}_{n+2} = b_i$$

   $\bar{x}_{n+2} = b_i$ is feasible, with $x_j = 0$, $j = 1, \cdots, n+1$.

3. *Negative RHS With Less-Than-Or-Equal-To Constraint*: multiply through by -1 to get case (2), but now with $-b_i$ positive:

$$-\sum_{j=1}^{n} a_{ij}x_j - x_{n+1} + \bar{x}_{n+2} = -b_i$$

$x_{n+2} = -b_i$ is feasible with $x_j = 0$, $j = 1, \cdots, n+1$.

The introduction of the new variables $\bar{x}_k$ has *changed the problem*. The old problem is only regained if all $\bar{x}_k$ eventually equal zero. This can be done, e.g., by introducing the *artificial objective function* $W = -\sum \bar{x}_k$ and maximizing W. If the maximum value of W is *zero*, then all $\bar{x}_k = 0$, and if use the simplex algorithm to do this, all x_j will still be feasible. Thus we have a feasible solution to the original problem.

6 The Dual Problem

Reconsider the model problem, i.e.

$$\begin{array}{rrcrcrcr} \text{Maximise:} & Z & = & 3x_1 & + & 5x_2 & & \\ \text{subject to:} & & & x_1 & & & \leq & 4 \\ & & & & & 2x_2 & \leq & 12 \\ & & & 3x_1 & + & 2x_2 & \leq & 18 \\ \text{with} & & & x_1, & x_2 \geq 0\,. & & & \end{array}$$

One method to obtain information about the objective function, $Z = 3x_1 + 5x_2$, is to look at the constraints. For example, $3 \times$ first constraint $+ 2.5 \times$ second constraint gives

$$3x_1 + 5x_2 \leq 3 \times 4 + 2.5 \times 12 \leq 42\,,$$

i.e. we could certainly do no better than a maximum profit of 42.

Similarly, $2.5 \times$ constraint three gives $7.5x_1 + 5x_2 \leq 45$, so

$$Z = 3x_1 + 5x_2 \leq 7.5x_1 + 5x_2 \leq 45\,.$$

The most general information about the objective function that can be gleaned from the constraints is obtained by taking the linear combination,

$$y_1(x_1) + y_2(2x_2) + y_3(3x_1 + 2x_2) \leq 4y_1 + 12y_2 + 18y_3$$

or

$$Z = 3x_1 + 5x_2 \leq (y_1 + 3y_3)x_1 + (2y_2 + 2y_3)x_2 \leq 4y_1 + 12y_2 + 18y_3$$

We have assumed that y_1, y_2, y_3 are non-negative. The right-hand side of this inequality is an upper bound on the original objective function Z, if the coefficient of x_1 is at least 3 and the coefficient of x_2 is at least 5; i.e. if we have,

$$\begin{array}{rcrcr} y_1 & + & 3y_3 & \geq & 3 \\ 2y_2 & + & 2y_3 & \geq & 5\,. \end{array}$$

The smallest possible upper bound on Z is obtained if the right-hand side $4y_1 + 12y_2 + 18y_3$ is as small as possible; i.e. we want to minimize $v = 4y_1 + 12y_2 + 18y_3$.

Thus we have defined a new LP problem, the *dual LP Problem*. In this context, the original problem is called the *Primal LP Problem*.

Consider the tableau for the primal:

	$x_1 \geq 0$	$x_2 \geq 0$	
$y_1 \geq 0$	1	0	$\leq 4 = b_1$
$y_2 \geq 0$	0	2	$\leq 12 = b_2$
$y_3 \geq 0$	3	2	$\leq 18 = b_3$
	$\geq 3 = c_1$	$\geq 5 = c_2$	

We can construct the dual LP problem from the tableau for the primal LP problem, i.e.

$$\max Z = c_1x_1 + c_2x_2 \quad \Longrightarrow \begin{cases} \min v &= b_1y_1 + b_2y_2 + b_3y_3 \\ &= 4y_1 + 12y_2 + 18y_3 \end{cases}$$

the constraints for the dual programme are read by column

$$\begin{array}{lrcl} \text{column } 1 \Rightarrow & y_1 + 0y_2 + 3y_3 & \geq & 3 \\ & & \text{sign change} & \\ \text{column } 2 \Rightarrow & 0y_1 + 2y_2 + 2y_3 & \geq & 5 \end{array}$$

the variables y_1, y_2, y_3 are all ≥ 0, as are x_1, x_2.

In other words, the tableau for the dual programme just has the transposed matrix of that for the primal, i.e. schematically,

	$\mathbf{x} \geq \mathbf{0}$	
$\mathbf{y} \geq \mathbf{0}$	$\mathbf{A}$	$\leq \mathbf{b}$
	$\geq \mathbf{c}$	$\max Z$ $\min v$

Thus

Primal				**Dual**			
Maximise:	Z	$=$	$\mathbf{c}^T\mathbf{x}$	Minimise:	v	$=$	$\mathbf{y}^T\mathbf{b}$
subject to:	$\mathbf{Ax}$	$\leq$	$\mathbf{b}$	subject to:	$\mathbf{y}^T\mathbf{A}$	$\geq$	$\mathbf{c}^T$
with:	$\mathbf{x}$	$\geq$	$\mathbf{0}$	with:	$\mathbf{y}$	$\geq$	$\mathbf{0}$

where $\mathbf{c}$ and $\mathbf{x}$ are $n \times 1$ column vectors, $\mathbf{b}$ and $\mathbf{y}$ are $m \times 1$ column vectors, and $\mathbf{A}$ is $m \times n$. The primal problem is general, since the restriction $\mathbf{b} \geq \mathbf{0}$ is not assumed and any equality constraint $\mathbf{a}^T\mathbf{x} = b$ can be replaced by the two equivalent inequalities $\mathbf{a}^T\mathbf{x} \leq b$, $-\mathbf{a}^T\mathbf{x} \leq -b$.

The *dual problem* is not only theoretically interesting but can also have practical applications. First, under certain conditions which we shall discuss later, the optimal solution for Z is also the optimal solution for v. Therefore if we started with a primal problem with, say five variables but only two constraints, we should normally have to solve it using the simplex method. The corresponding dual problem, by contrast, only has *two* variables (and five constraints) and so can be solved graphically.

Second, it can be easier to find an initial feasible solution for the dual problem. Thus, although the primal problem may require the two-phase simplex algorithm, the dual may not.

Third, the number of iterations required by the simplex method depends on the number of constraints. If the primal has many more constraints than decision variables, the dual will take fewer iterations to solve.

Example 1

$$\begin{array}{rrcrcrcrl} \text{Maximise:} & Z & = & 8x_1 & + & 30x_2 & + & 8x_3 & \\ \text{subject to:} & & & x_1 & + & 5x_2 & + & 3x_3 & \leq 10 \\ & & & 4x_1 & + & 6x_2 & + & x_3 & \leq 15 \\ \text{with:} & & & & & x_1,\ x_2,\ x_3 \geq 0. & & & \end{array}$$

The corresponding dual problem is:

$$\begin{array}{rrcrcrcrl} \text{Minimise:} & v & = & 10y_1 & + & 15y_2 & & & \\ \text{subject to:} & & & y_1 & + & 4y_2 & \geq & 8 \\ & & & 5y_1 & + & 6y_2 & \geq & 30 \\ & & & 3y_1 & + & y_2 & \geq & 8 \\ \text{with:} & & & & & y_1,\ y_2 \geq 0\,, & & & \end{array}$$

which can be solved graphically.

The optimal solution is found at the intersection of

$$5y_1 + 6y_2 = 30\,, \qquad y_1 + 4y_2 = 8$$

or $y_1 = 36/7$, $y_2 = 5/7$. The minimum value of $v = 10y_1 + 15y_2 = (360 + 75)/7$, which is also the maximum value of Z — see later.

Example 2

Use the simplex algorithm to

$$\begin{array}{rrcrcrcrl} \text{Minimise:} & Z & = & 3x_1 & + & 2x_2 & & & \\ \text{subject to:} & & & 8x_1 & + & 3x_2 & \geq & 24 \\ & & & 5x_1 & + & 6x_2 & \geq & 30 \\ & & & 2x_1 & + & 9x_2 & \geq & 18 \\ \text{with:} & & & & & x_1,\ x_2 \geq 0\,. & & & \end{array}$$

We would need artificial variables here, because of the $\geq$ sign. Instead, look at the dual problem:

$$\begin{array}{rrcrcrcrl} \text{Maximise:} & v & = & 24y_1 & + & 30y_2 & + & 18y_3 & \\ \text{subject to:} & & & 8y_1 & + & 5y_2 & + & 2y_3 & \leq 3 \\ & & & 3y_1 & + & 6y_2 & + & 9y_3 & \leq 2 \\ \text{with:} & & & & & y_1,\ y_2,\ y_3 \geq 0. & & & \end{array}$$

Introduce slack variables y_4, y_5. The starting vertex is then $(y_1, y_2, y_3, y_4, y_5) = (0, 0, 0, 3, 2)$,

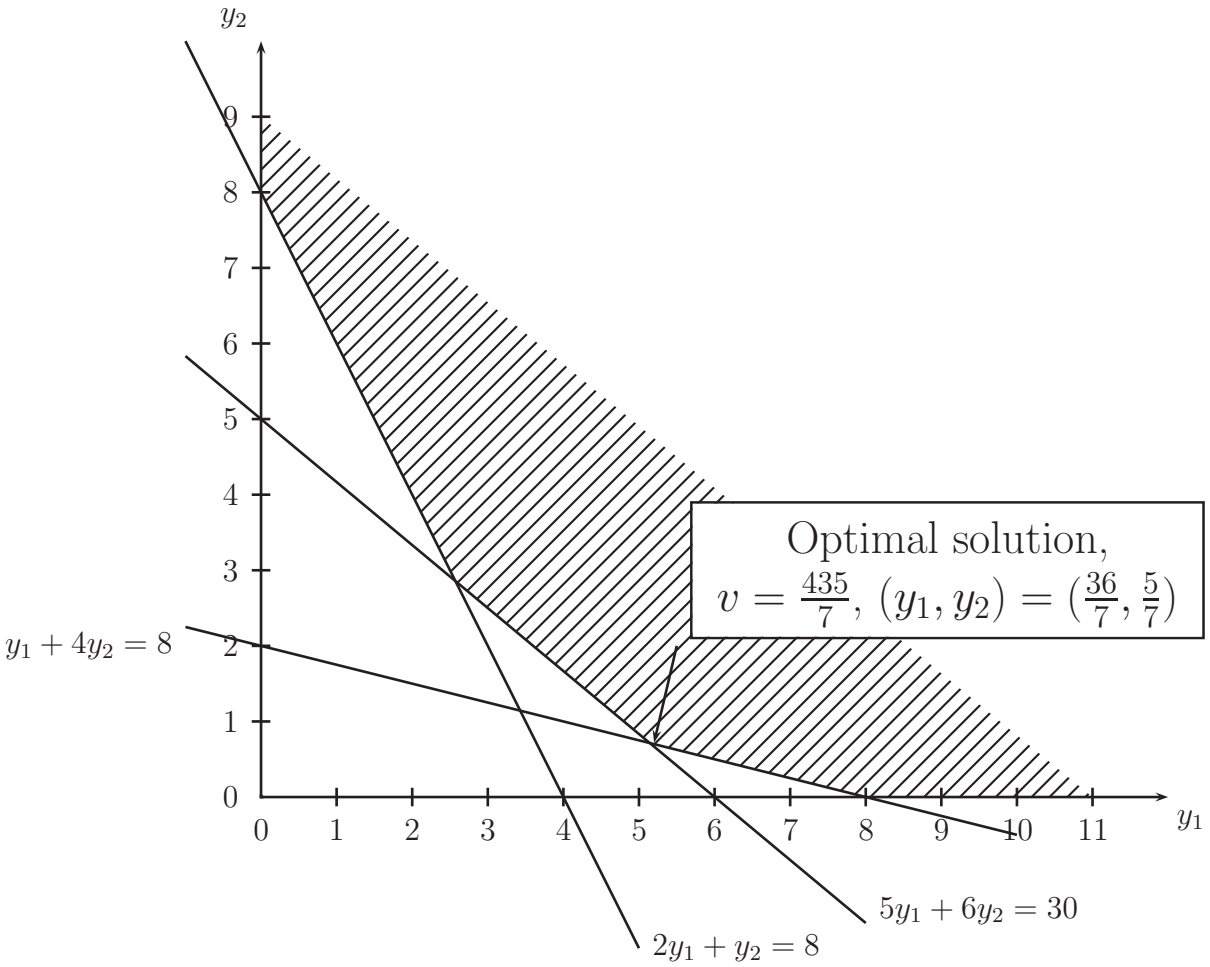


Figure 2.9: Graphical solution of a dual LP problem.

v	y_1	y_2	y_3	y_4	y_5	b_i	Ratio
1	-24	-30	-18	0	0	0	
0	8	5	2	1	0	3	3/5
0	3	6	9	0	1	2	1/3
1	-9	0	27	0	5	10	
0	11/2	0	$-11/2$	1	$-5/6$	4/3	8/33
0	1/2	1	3/2	0	1/6	1/3	2/3
1	0	0	18	18/11	40/11	134/11	
0	1	0	-1	2/11	$-5/33$	8/33	
0	0	1	2	$-1/11$	8/33	7/33	

This is optimal since all coefficients in the objective row are ≥ 0. Then

$$y_1 = \frac{8}{33} \qquad y_2 = \frac{7}{33} \qquad y_3 = 0$$

and the maximum value of v is 134/11.

What are the associated optimal values of x_1, x_2 in the primal problem? They are, in fact, $x_1 = 18/11$, $x_2 = 40/11$, which are simply the coefficients of y_4, y_5 in the objective row.

These numbers represent the multiples of rows 2 and 3, respectively, that have been added to the objective function, i.e. row 1, since the coefficient of y_4 in row 2 was initially 1, as was that for y_5 in row 3,

$$\begin{array}{llrcrcrcrcrcrcr}
 & & v & - & 24y_1 & - & 30y_2 & - & 18y_3 & & & & & = & 0 \\
+ & 18/11\times & & & 8y_1 & + & 5y_2 & + & 2y_3 & + & y_4 & & & = & 3 \\
+ & 40/11\times & & & 3y_1 & + & 6y_2 & + & 9y_3 & & & + & y_5 & = & 2 \\
\hline
\Rightarrow & & v & + & 0y_1 & + & 0y_2 & + & 18y_3 & + & (18/11)y_4 & + & (40/11)y_5 & = & 134/11\,.
\end{array}$$

Thus we have

$$\begin{aligned}
&\text{coefficients of } y_1\text{:} && -24 + \frac{18}{11}\times 8 + \frac{40}{11}\times 3 = 0\,;\\
&\text{coefficients of } y_2\text{:} && -30 + \frac{18}{11}\times 5 + \frac{40}{11}\times 6 = 0\,;\\
&\text{coefficients of } y_3\text{:} && -18 + \frac{18}{11}\times 2 + \frac{40}{11}\times 9 = 18\,.
\end{aligned}$$

Setting $x_1^* = 18/11$ and $x_2^* = 40/11$, these three equations can be written as,

$$\begin{aligned}
8x_1^* + 3x_2^* &= 24\\
5x_1^* + 6x_2^* &= 30\\
2x_1^* + 9x_2^* &= 36 \geq 18\,.
\end{aligned}$$

But this proves (18/11, 40/11) is a *feasible solution* to the primal problem and an optimal solution to the *dual*. Hence, it is optimal for the primal. (This follows from the fundamental duality theorem which we will soon prove.) Note that $Z^* = 3x_1^* + 2x_2^* = 134/11$.

The above procedure can always be used to find the solution to the primal given the optimal tableau for the dual: namely, the optimal values of x_1, x_2, etc. are just the coefficients of the slack variables y_4, y_5, etc. in the objective function.

Similarly the solution to the dual problem can be read off from the tableau of the primal problem, since the dual of the dual is the primal.

6.1 The Fundamental Duality Theorem

Recall

Primal				**Dual**			
Maximise:	Z	$=$	$\mathbf{c}^T\mathbf{x}$	Minimise:	v	$=$	$\mathbf{y}^T\mathbf{b}$
subject to:	$\mathbf{Ax}$	$\leq$	$\mathbf{b}$	subject to:	$\mathbf{y}^T\mathbf{A}$	$\geq$	$\mathbf{c}^T$
	$\mathbf{x}$	$\geq$	$\mathbf{0}$		$\mathbf{y}$	$\geq$	$\mathbf{0}$

Theorem 1 *If* $\mathbf{x}$ *and* $\mathbf{y}$ *are* feasible solutions *of the primal and dual problems, respectively, then*

$$v = \mathbf{y}^T\mathbf{b} \geq \mathbf{y}^T\mathbf{Ax} \geq \mathbf{c}^T\mathbf{x} = Z\,,$$

i.e. every feasible solution of the dual problem is greater than or equal to every feasible solution of the primal problem, including the optimal one.

Proof : $\mathbf{y}^T \geq \mathbf{0}$ and $\mathbf{Ax} \leq \mathbf{b}$ imply

$$\mathbf{y}^T(\mathbf{Ax} - \mathbf{b}) \leq 0\,, \qquad \text{or} \qquad \mathbf{y}^T\mathbf{Ax} \leq \mathbf{y}^T\mathbf{b}\,.$$

Similarly, $\mathbf{x} \geq 0$ and $\mathbf{y}^T\mathbf{A} \geq \mathbf{c}^T$ imply

$$(\mathbf{y}^T\mathbf{A} - \mathbf{c}^T)\mathbf{x} \geq 0\,, \qquad \text{or} \qquad \mathbf{y}^T\mathbf{Ax} \geq \mathbf{c}^T\mathbf{x}\,.$$

Theorem 2 *If* $\mathbf{x}$, $\mathbf{y}$ *are* feasible *primal and dual solutions such that* $\mathbf{c}^T\mathbf{x} = \mathbf{y}^T\mathbf{b}$, *then* $\mathbf{x}$, $\mathbf{y}$ *are dual and primal* optimal *solutions with*

$$v^* = \mathbf{y}^T\mathbf{b} = \mathbf{y}^T\mathbf{Ax} = \mathbf{c}^T\mathbf{x} = Z^*\,.$$

Proof From Theorem 1, for any feasible primal solution,

$$Z \leq \mathbf{y}^T\mathbf{b}\,,$$

i.e. $\mathbf{y}^T\mathbf{b}$ is an upper bound of Z. But, by hypothesis, for the primal solution $\mathbf{x}$,

$$Z = \mathbf{c}^T\mathbf{x} = \mathbf{y}^T\mathbf{b}\,,$$

i.e. Z attains the upper bound, and hence $\mathbf{x}$ is optimal for the primal problem. Similarly one can show $\mathbf{y}$ is optimal for the dual problem.

In the second dual example above, both dual and primal are feasible, and have the same value 134/11 — this is *optimal*, as mentioned there.

Theorem 3 (Fundamental Duality Theorem) *If both an LP and its dual are feasible, then both have optimal solutions and the optimal values of the objective functions of the primal and dual problems are equal.*

We give a demonstration, when the primal problem is in standard form, using the simplex algorithm. This requires the assumption $\mathbf{b} \geq \mathbf{0}$. (The proof fails if cycling occurs, where there is a tie for the leaving basic variable. If we use the smallest-index rule to break the tie cycling does not occur.)

In canonical form, the primal is

Maximise:	$Z = \mathbf{c}^T\mathbf{x}$
subject to:	$\mathbf{A}\mathbf{x} + \mathbf{x}_s = \mathbf{b}$
with:	$\mathbf{x}\,,\ \mathbf{x}_s \geq \mathbf{0}.$

The vector $\mathbf{x}_s$ denotes the slack variables.

Initial Tableau			
Z	$\mathbf{x}^T$	$\mathbf{x}_s^T$	RHS
1	$-\mathbf{c}^T$	$\mathbf{0}$	0
$\mathbf{0}$	$\mathbf{A}$	$\mathbf{I}$	$\mathbf{b}$

Final Tableau			
Z	$\mathbf{x}^T$	$\mathbf{x}_s^T$	RHS
1	$-\mathbf{c}^T + \mathbf{y}^T\mathbf{A}$	$\mathbf{y}^T$	$\mathbf{y}^T\mathbf{b}$
$\mathbf{0}$	$\mathbf{BA}$	$\mathbf{B}$	$\mathbf{Bb}$

The matrix $\mathbf{B}$ keeps track of row operations performed on the constraint rows. As mentioned in an earlier example, $\mathbf{y}^T$ keeps track of the multiples of the original constraint rows that are added to the objective function.

Note that if cycling does not occur, simplex either stops with an optimal solution or with a solution where Z can be increased without bound (the case where all entries in the pivot column are ≤ 0). However, the dual is feasible by assumption, so Z is bounded above (Theorem 1). Hence the stopping solution is optimal.

Since the tableau is optimal, the coefficients in the first row are all ≥ 0. Thus

$$\mathbf{y}^T\mathbf{A} - \mathbf{c}^T \geq \mathbf{0}\,, \qquad \mathbf{y}^T \geq \mathbf{0}\,, \qquad Z^* = \mathbf{y}^T\mathbf{b}\,,$$

i.e.

$$v = \mathbf{y}^T\mathbf{b} = Z^*\,, \qquad \mathbf{y}^T\mathbf{A} \geq \mathbf{c}^T\,, \qquad \mathbf{y} \geq \mathbf{0}\,.$$

Thus $\mathbf{y}$ is a dual feasible solution and the corresponding value of $v = \mathbf{y}^T\mathbf{b} = Z^*$. Thus, by Theorem 2, $\mathbf{y}$ is an *optimal* solution to *the dual*.

The condition $\mathbf{b} \geq \mathbf{0}$ can be omitted: it is only required to find an initial feasible solution. In the theorem the primal problem is feasible by hypothesis, i.e. an initial feasible solution is assumed.

6.2 General LP Problem and its Dual

Each row of the following table matches a primal condition with a dual condition.

Primal		**Dual**
$\max Z = \sum_{j=1}^{n} c_j x_j$		$\min v = \sum_{i=1}^{m} b_i y_i$
$\sum_{j=1}^{n} a_{ij} x_j \le b_i$	$i = 1, \ldots, k$	$y_i \ge 0$
$\sum_{j=1}^{n} a_{ij} x_j \ge b_i$	$i = k+1, \ldots, \ell$	$y_i \le 0$
$\sum_{j=1}^{n} a_{ij} x_j = b_i$	$i = \ell+1, \ldots, m$	y_i unrestricted in sign
$x_j \ge 0$	$j = 1, \ldots, p$	$\sum_{i=1}^{m} y_i a_{ij} \ge c_j$
$x_j \le 0$	$j = p+1, \ldots, q$	$\sum_{i=1}^{m} y_i a_{ij} \le c_j$
x_j unrestricted in sign	$j = q+1, \ldots, n$	$\sum_{i=1}^{m} y_i a_{ij} = c_j$

Tableau

		$1, \ldots, p$ $x_j \ge 0$	$p+1, \ldots, q$ $x_j \le 0$	$q+1, \ldots, n$ x_j	
1 $\vdots$ k	$y_i \ge 0$	A_{11}	A_{12}	A_{13}	$\le b_i$
$k+1$ $\vdots$ ℓ	$y_i \le 0$	A_{21}	A_{22}	A_{23}	$\ge b_i$
$\ell+1$ $\vdots$ m	y_i	A_{31}	A_{32}	A_{33}	$= b_i$
		$\ge c_j$	$\le c_j$	$= c_j$	

Chapter 3

Nonlinear Optimisation without Constraints

1 Introduction

In this chapter we consider the problem of optimising a function of several variables, but without any associated constraints. We look at optimisation problems with constraints in the next chapter. Thus, we wish to find the minimum of an *objective* function $f(x_1, x_2, \ldots, x_n)$ of n variables $x_1, x_2, \ldots, x_n$ defined on a subset Ω of $\mathbb{R}^n$. Often Ω is $\mathbb{R}^n$. Typical examples of objective functions are

$$f = x_1^3 + x_2^2\,, \qquad f = \sin x_1 + 2x_1x_2^3\,, \qquad f = x_2^2 \sin x_1 + x_1^2x_3^3 + x_2^4\,.$$

In vector notation $\mathbf{x} := (x_1, x_2 \ldots, x_n)$ the problem is

$$\min_{\mathbf{x}\in\Omega} f(\mathbf{x})\,.$$

There is no loss of generality in considering only minimisation problems, since a maximisation problem can be converted into a minimisation problem using the identity,

$$\max_{\mathbf{x}\in\Omega} f(\mathbf{x}) = -\min_{\mathbf{x}\in\Omega}\{-f(\mathbf{x})\}\,.$$

We confine ourselves to extrema which occur in the interior of Ω and not on its boundary, since we are considering unconstrained optimisation. Boundary extrema are associated with constrained optimisation problems and hence are considered in the next chapter. We also pursue the easier problem of local extrema of f rather than global extrema on Ω. A point $\mathbf{x}_0$ is a *global minimum* (*global maximum*) of $f(\mathbf{x})$ in $\mathbb{R}^n$, if $f(\mathbf{x}) \geq f(\mathbf{x}_0)$ ($f(\mathbf{x}) \leq f(\mathbf{x}_0)$) for all $\mathbf{x}$ in $\mathbb{R}^n$. A point $\mathbf{x}_0$ is a *local minimum* (*local maximum*) of $f(\mathbf{x})$, if $f(\mathbf{x}) \geq f(\mathbf{x}_0)$ ($f(\mathbf{x}) \leq f(\mathbf{x}_0)$) for all $\mathbf{x}$ in some neighbourhood of $\mathbf{x}_0$.

The idea of finding the zeros of the derivative of a function of one variable to identify the critical points and then differentiating again to find the nature of each critical point (i.e. maximum, minimimum or point of inflexion), as described in Chapter 1, is familiar from elementary calculus. These ideas extend to functions of more than one variable, and provide a direct analytic approach for determining local maxima or minima for unconstrained problems.

1.1 Optimisation Conditions for Functions of One Variable

The necessary and sufficient conditions derived in Chapter 1 for a local maximum or minimum of a function of one variable can be derived using Taylor's Theorem for functions of one variable.

Taylor's Theorem for Functions of One Variable Suppose that $f \in C^p[a,b]$ and the derivative $f^{(p+1)}$ exists on $[a,b]$, and let $x_0 \in [a,b]$. For every $x \in [a,b]$ there exists $\eta(x)$ between x_0 and x such that

$$f(x) = f(x_0) + f^{(1)}(x_0)(x-x_0) + \frac{f^{(2)}(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(p)}(x_0)}{p!}(x-x_0)^p + \frac{f^{(p+1)}(\eta)}{(p+1)!}(x-x_0)^{p+1}\,. \tag{1.1}$$

Proof Use integration by parts:

$$\begin{aligned}
f(x) &= f(x_0) + \int_{x_0}^{x} f^{(1)}(t)\,dt \\
&= f(x_0) - \left[f^{(1)}(t)(x-t)\right]_{x_0}^{x} + \int_{x_0}^{x} f^{(2)}(t)(x-t)\,dt \\
&= f(x_0) - \left[f^{(1)}(t)(x-t) + f^{(2)}(t)\frac{1}{2}(x-t)^2 + \cdots + f^{(p)}(t)\frac{1}{p!}(x-t)^p\right]_{x_0}^{x} \\
&\quad + \frac{1}{p!}\int_{x_0}^{x} f^{(p+1)}(t)(x-t)^p\,dt \\
&= f(x_0) + f^{(1)}(x_0)(x-x_0) + f^{(2)}(x_0)\frac{1}{2}(x-x_0)^2 + \cdots + f^{(p)}(x_0)\frac{1}{p!}(x-x_0)^p \\
&\quad + \frac{1}{p!}\int_{x_0}^{x} f^{(p+1)}(t)(x-t)^p\,dt\,.
\end{aligned}$$

The remainder term is in Young's integral form. Noting that $(x-t)^p$ has one sign on $[x_0,x]$, application of the generalised integral mean value theorem gives

$$\frac{1}{p!}\int_{x_0}^{x} f^{(p+1)}(t)(x-t)^p\,dt = f^{(p+1)}(\eta)\frac{1}{p!}\int_{x_0}^{x}(x-t)^p\,dt = \frac{f^{(p+1)}(\eta)}{(p+1)!}(x-x_0)^{p+1}\,.$$

The following three results may be derived from (1.1).

1. If $f^{(1)}(x_0) = f^{(2)}(x_0) = \cdots = f^{(2m-1)}(x_0) = 0$ for some integer m and $f^{(2m)}(x_0) < 0$, where $f^{(2m)}(x)$ is continuous at x_0, then x_0 is a *local maximum.*

2. If $f^{(1)}(x_0) = f^{(2)}(x_0) = \cdots = f^{(2m-1)}(x_0) = 0$ for some integer m and $f^{(2m)}(x_0) > 0$, where $f^{(2m)}(x)$ is continuous at x_0, then x_0 is a *local minimum.*

3. If $f^{(1)}(x_0) = f^{(2)}(x_0) = \cdots = f^{(2m)}(x_0) = 0$ for some integer n and $f^{(2m+1)}(x_0) \neq 0$, where $f^{(2m+1)}(x)$ is continuous at x_0, then x_0 is a *point of inflection.*

To prove Result 1 choose $p = 2m - 1$ in (1.1), which reduces to

$$f(x) = f(x_0) + \frac{f^{(2m)}(\eta)}{(2m)!}(x - x_0)^{2m},$$

and note that, if x is close enough to x_0, then $f^{(2m)}(\eta) < 0$, since $f^{(2m)}(x_0) < 0$. Thus $f(x) < f(x_0)$, if x lies in some neighbourhood of x_0, excluding x_0, i.e. x_0 is a local maximum of f. Results 2 and 3 can be established by similar arguments.

2 Optimisation Conditions for Functions of Two Variables

The necessary and sufficient conditions for a local maximum or minimum of a function of one variable can be extended to a function $f(x, y)$ of two variables x,y, using Taylor's Theorem for functions of two variables.

Taylor's Theorem for Functions of Two Variables: Suppose that $f(x, y)$ and its partial derivatives of all orders less than or equal to $p + 1$ are continuous on $D = \{(x, y) \mid a \le x \le b\,,\ c \le y \le d\}$ and let $(x_0, y_0) \in D$. For every $(x, y) \in D$, there exists ξ between x and x_0, and η between y and y_0 such that

$$\begin{aligned} f(x,y) = f(x_0,y_0) &+ \left(\frac{\partial f}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial f}{\partial y}\right)_0 (y - y_0) \\ &+ \frac{1}{2!}\left\{ \left(\frac{\partial^2 f}{\partial x^2}\right)_0 (x - x_0)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial y}\right)_0 (x - x_0)(y - y_0) + \left(\frac{\partial^2 f}{\partial y^2}\right)_0 (y - y_0)^2 \right\} \\ &+ \cdots + \frac{1}{p!}\left\{ \sum_{j=0}^{p} \binom{p}{j} \left(\frac{\partial^p f}{\partial x^{p-j} \partial y^j}\right)_0 (x - x_0)^{p-j}(y - y_0)^j \right\} + R_p(x,y) \end{aligned} \quad (2.1)$$

where a subscript zero on the derivatives of f denotes evaluation at (x_0, y_0) and R_p is the remainder,

$$R_p(x,y) = \frac{1}{(p+1)!}\left\{ \sum_{j=0}^{p+1} \binom{p+1}{j} \left(\frac{\partial^{p+1} f}{\partial x^{p+1-j} \partial y^j}\right)_{(\xi,\eta)} (x - x_0)^{p+1-j}(y - y_0)^j \right\}.$$

Proof: Apply the one-dimensional Taylor's theorem about $t = 0$ to the function $F(t)$ of one variable t defined by

$$F(t) := f(x_0 + t(x - x_0), y_0 + t(y - y_0)),$$

noting that

$$\begin{aligned} F(0) &= f(x_0, y_0) \\ F^{(1)}(0) &= f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\ F^{(2)}(0) &= f_{xx}(x_0, y_0)(x - x_0)^2 + 2f_{xy}(x_0, y_0)(x - x_0)(y - y_0) + f_{yy}(x_0, y_0)(y - y_0)^2 \\ &\vdots \end{aligned}$$

where $f_x = \partial f / \partial x$, etc.

Example: Use Taylor's theorem for functions of two variables to find a quadratic approximation to $\sqrt{1+x-y}$ at $(x,y)=(0,0)$.

Solution: Let $z=\sqrt{1+x-y}$. The derivatives of z up to second order are

$$z_x = \frac{1}{2\sqrt{1+x-y}} = -z_y\,, \qquad z_{xx} = -\frac{1}{4(1+x-y)^{3/2}} = z_{yy} = -z_{xy}\,.$$

Evaluating z and these derivatives at $(0,0)$ gives

$$z=0\,, \qquad z_x = \frac{1}{2} = -z_y\,, \qquad z_{xx} = -\frac{1}{4} = z_{yy} = -z_{xy}\,.$$

Hence

$$z = z(0,0) + z_x(0,0)x + z_y(0,0)y + \frac{1}{2}z_{xx}(0,0)x^2 + z_{xy}(0,0)xy + \tag{2.2}$$

$$\frac{1}{2}z_{yy}(0,0)y^2 + \ldots \tag{2.3}$$

$$= 1 + \frac{1}{2}x - \frac{1}{2}y - \frac{1}{8}x^2 + \frac{1}{4}xy - \frac{1}{8}y^2 + \ldots\,. \tag{2.4}$$

Alternatively, use the binomial expansion to expand $\sqrt{1+z}$ in a Taylor series, $\sqrt{1+z} = 1 + \frac{1}{2}z - \frac{1}{8}z^2 + \ldots$, and then set $z = x - y$.

We now derive necessary and sufficient conditions for the point (x_0, y_0) to be a minimum or a maximum of the function $f(x,y)$ using Taylor's Theorem for functions of two variables (1.1). The conditions use the gradient and Hessian of f. Recall that in two-dimensional cartesian coordinates the *gradient* of f is defined by

$$\nabla f := \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}$$

or, in matrix notation, the gradient of f is the column vector,

$$\nabla f = \begin{pmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{pmatrix}\,.$$

The *Hessian* matrix $\mathbf{H}(x,y)$ of a function $f(x,y)$ in two dimensions is the matrix of second derivatives of f defined by

$$\mathbf{H}(x,y) := \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix}\,.$$

For all objective functions $f(x,y)$, which we consider,

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}\,,$$

so $\mathbf{H}$ is a symmetric matrix, i.e. $H_{12} = H_{21}$.

Let $\mathbf{x} = \mathbf{x}_0 + \mathbf{d}$, where $\mathbf{d}$ is an arbitrary displacement of (x, y) from (x_0, y_0),

$$\mathbf{d} := \mathbf{x} - \mathbf{x}_0 = \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} .$$

Thus (1.1) can be written as

$$f(\mathbf{x}_0 + \mathbf{d}) = f(\mathbf{x}_0) + \{\nabla f(\mathbf{x}_0)\}^T \mathbf{d} + \frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}_0)\mathbf{d} + \cdots , \tag{2.5}$$

Clearly (x_0, y_0) can be an extremum of f only if $\{\nabla f(x_0, y_0)\}^T \mathbf{d} = 0$, since we can replace $\mathbf{d}$ by $-\mathbf{d}$ and choose the point (x, y) sufficiently close to (x_0, y_0) (i.e. $\mathbf{d}$ is small enough), so that $\{\nabla f(x_0, y_0)\}^T \mathbf{d}$ dominates the following terms on the right side of (2.1). But the direction of $\mathbf{d}$ is arbitrary, since (x_0, y_0) is interior to Ω, and hence $\nabla f(x_0, y_0)$ must be zero. Thus a *necessary condition* for (x_0, y_0) to be a maximum or minimum of $f(x, y)$ is

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0 , \qquad \text{at } (x_0, y_0). \tag{2.6}$$

If condition (2.1) holds, then (x_0, y_0) is a *critical point* of f.

If (x_0, y_0) is a critical point of f, then equation (1.2) reduces to

$$f(\mathbf{x}_0 + \mathbf{d}) - f(\mathbf{x}_0) = \frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}_0)\mathbf{d} + \cdots . \tag{2.7}$$

A sufficient condition for a minimum or maximum at the point (x_0, y_0) can be obtained in terms of the Hessian matrix $\mathbf{H}$ at (x_0, y_0). Associated with any symmetric Hessian matrix $\mathbf{H}$ is a quadratic form, $Q := \mathbf{x}^T \mathbf{H}\mathbf{x}$.

Example: If $\mathbf{H}$ is the matrix,

$$H = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

the associated quadratic form Q is

$$Q = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 3x_1x_2 + 3x_2x_1 + x_2^2 = x_1^2 + 6x_1x_2 + x_2^2 .$$

Conversely, associated with any quadratic form, $Q := ax^2 + 2bxy + cy^2$, is a unique symmetric matrix,

$$\mathbf{M} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

such that $Q := \mathbf{x}^T \mathbf{M}\mathbf{x}$.

Example: If Q is quadratic form $Q := ax^2 + bxy + cyx + dy^2$, then the associated symmetric matrix $\mathbf{M}$ is

$$\mathbf{M} = \begin{pmatrix} a & \frac{1}{2}(b+c) \\ \frac{1}{2}(b+c) & d \end{pmatrix} .$$

The quadratic form $Q = \mathbf{x}^T\mathbf{H}\mathbf{x}$, and its associated symmetric matrix $\mathbf{H}$, may be classified as follows:

1. Q and $\mathbf{H}$ are *positive definite*, if $Q > 0$ for $\mathbf{x} \neq \mathbf{0}$.
2. Q and $\mathbf{H}$ are *negative definite*, if $Q < 0$ for $\mathbf{x} \neq \mathbf{0}$.
3. Q and $\mathbf{H}$ are *positive semi-definite*, if $Q \geq 0$ and $Q = 0$ for some $\mathbf{x} \neq \mathbf{0}$.
4. Q and $\mathbf{H}$ are *negative semi-definite*, if $Q \leq 0$ and $Q = 0$ for some $\mathbf{x} \neq \mathbf{0}$.
5. Q and $\mathbf{H}$ are *indefinite*, if there exist $\mathbf{x}_1$, $\mathbf{x}_2$ such that $Q > 0$ at $\mathbf{x} = \mathbf{x}_1$ and $Q < 0$ at $\mathbf{x} = \mathbf{x}_2$.

We now give a method to determine if the quadratic form Q (and its associated symmetric matrix) is positive-definite, negative-definite, positive semi-definite, negative semi-definite or indefinite. and hence whether a critical point is a local minimum or local maximum. The method due to Lagrange is complete the square of Q.

Example: In the previous example,

$$Q = (x_1 + 3x_2)^2 - 9x_2^2 + x_2^2 = (x_1 + 3x_2)^2 - 8x_2^2 = y_1^2 - 8y_2^2 \,,$$

if we introduce new variables $y_1 := x_1 + 3x_2$ and $y_2 := x_2$. If $y_1 \neq 0$ and $y_2 = 0$, which is equivalent to $x_1 \neq 0$ and $x_2 = 0$, then $Q > 0$, whereas, if $y_1 = 0$ and $y_2 \neq 0$, which is equivalent to $x_1 = -3x_2$ with $x_2 \neq 0$, then $Q < 0$. Thus Q is indefinite. This should be clear immediately from the coefficients of the squared terms, y_1^2, y_2^2, which are not of the same sign.

Example: If

$$H = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

then

$$Q = 2x_1^2 + 2x_2^2 + 2x_1x_2 = 2(x_1 + \tfrac{1}{2}x_2)^2 - \tfrac{1}{2}x_2^2 + 2x_2^2 = 2y_1^2 + \tfrac{3}{2}y_2^2 \,,$$

where $y_1 := x_1 + \frac{1}{2}x_2$ and $y_2 := x_2$. Thus Q is positive definite, as the coefficients of the squared terms are all positive and if $Q = 0$ then $y_1 = y_2 = 0$, i.e $x_1 = x_2 = 0$.

Equation (2.2) can be written as $f(\mathbf{x}_0 + \mathbf{d}) - f(\mathbf{x}_0) = \frac{1}{2}Q + \cdots$, where $Q := \mathbf{d}^T\mathbf{H}(\mathbf{x}_0)\mathbf{d}$ is the quadratic form associated with the Hessian $\mathbf{H}(\mathbf{x}_0)$. When $\mathbf{d}$ is small enough Q dominates the higher order terms represented by the dots. We conclude the following *sufficient conditions* for a critical point (x_0, y_0) to be a local minimum or a maximum of f:

1. If $\mathbf{H}$ is positive definite, then $\mathbf{x}_0$ is a local minimum of $f(\mathbf{x})$.
2. If $\mathbf{H}$ is negative definite, then $\mathbf{x}_0$ is a local maximum of $f(\mathbf{x})$.

3. If $\mathbf{H}$ is positive semi-definite, then the test fails.

4. If $\mathbf{H}$ is negative semi-definite, then the test fails.

5. If $\mathbf{H}$ is indefinite, then $\mathbf{x}_0$ is a saddle point of $f(\mathbf{x})$.

In Cases 3 and 4 the test fails because there may be points near $\mathbf{x}_0$, where the neglected terms in (2.7) may be important.

Example: Minimize $f(x,y) = \frac{1}{2}(x-1)^2 + \frac{1}{2}(y-2)^2 + 1$.

Solution: Here it is obvious by inspection that $\min f_{=}1$ when $(x,y) = (1,2)$. Actually, $f = f(x,y)$ is a parabolic bowl with level contours $f =$ constant being the family of concentric circles $(x-1)^2 + (y-2)^2 =$ constant. The minimum point corresponds to the contour of zero radius.

Formally, we get the same solution by writing down the first-order conditions:

$$\frac{\partial f}{\partial x} = x - 1 = 0$$
$$\frac{\partial f}{\partial y} = y - 2 = 0$$

and the Hessian

$$H = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The result follows.

3 Optimisation Conditions for Functions of n Variables

The results of the previous section extend to objective functions $f(x_1, x_2, \ldots, x_n)$ of n variables, $x_1, x_2, \ldots, x_n$.

Taylor's Theorem for Functions of n Variables: Suppose that $f(\mathbf{x})$ and its partial derivatives of all orders less than or equal to $p+1$ are continuous on an open set D and let $\mathbf{x}_0 \in D$ such that the line segment joining $\mathbf{x}$ to $\mathbf{x}_0$ lies in D. For every $\mathbf{x} \in D$, there exists ξ between $\mathbf{x}$ and $\mathbf{x}_0$, such that

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)_0 (x_i - x_{0i}) + \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n} \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_0 (x_i - x_{0i})(x_j - x_{0j})$$
$$+ \cdots + \frac{1}{p!} \sum_{\substack{i_1, i_2, \ldots, i_n = 0 \\ (i_1 + i_2 + \cdots + i_n = p)}}^{p} \frac{p!}{i_1! i_2! \ldots i_n!} \left(\frac{\partial^p f}{\partial x_1^{i_1} \partial x_2^{i_2} \ldots \partial x_n^{i_n}}\right)_0 (x_1 - x_{01})^{i_1} (x_2 - x_{02})^{i_2} \ldots (x_n - x_{0n})^{i_n} + R_p(\mathbf{x}) \tag{3.1}$$

where a subscript zero on the derivatives of f denotes evaluation at (x_0, y_0) and R_p is the remainder,

$$R_p(\mathbf{x}) = \frac{1}{(p+1)!} \sum_{\substack{i_1,i_2,\ldots,i_n=0 \\ (i_1+i_2+\cdots+i_n=p+)}}^{p+1} \frac{(p+1)!}{i_1!i_2!\ldots i_n!} \left(\frac{\partial^{p+1} f}{\partial x_1^{i_1} \partial x_2^{i_2} \ldots \partial x_n^{i_n}} \right)_\xi (x_1 - x_{01})^{i_1} (x_2 - x_{02})^{i_2} \ldots (x_n - x_{0n})^{i_n} .$$

Proof: Apply the one-dimensional Taylor's theorem about $t = 0$ to the function $F(t)$ of one variable t defined by $F(t) := f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))$.

The *gradient* of f in n dimensions is defined by

$$\nabla f := \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} \mathbf{e}_j ,$$

where $\mathbf{e}_j$ is the unit basis vector along the x_j-axis. In matrix form,

$$\nabla f = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix} .$$

The *Hessian* matrix $\mathbf{H}(\mathbf{x})$ of a function $f(\mathbf{x})$ in n-dimensions is the matrix of second derivatives of f defined by

$$\mathbf{H} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{pmatrix} ,$$

For the functions f we consider,

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

so that the Hessian $\mathbf{H}$ is always a symmetric matrix, i.e. $H_{ij} = H_{ji}$.

Thus equation (2.5) still holds for n variables with $\mathbf{x} = \mathbf{x}_0 + \mathbf{d}$ from which we obtain: a *necessary condition* for $\mathbf{x}_0$ to be a maximum or minimum of $f(\mathbf{x})$ is

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \cdots = \frac{\partial f}{\partial x_n} = 0 , \qquad \text{at } \mathbf{x}_0. \tag{3.2}$$

If condition (3.2) holds, then $\mathbf{x}_0$ is a *critical point* of f. These are a system of n simultaneous equations for the n unknowns $x_1, x_2, \ldots, x_n$, which are generally non-linear, but linear if f is quadratic in the x_i.

As in the two variable case, there is associated with the symmetric Hessian matrix $\mathbf{H}$ a quadratic form $Q := \mathbf{x}^T\mathbf{H}\mathbf{x}$. Q and $\mathbf{H}$ may be classified as positive-definite, negative-definite, positive semi-definite, negative semi-definite or indefinite. The method of completing the square in Q, to determine which is the case, works for n variables.

Example: The quadratic form $Q = x_1^2 + 4x_1x_2 + 5x_2^2 + 6x_2x_3 + 2x_3^2 + 2x_1x_3$ has the associated symmetric matrix,

$$\mathbf{M} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 3 \\ 1 & 3 & 2 \end{pmatrix} .$$

To complete the square in Q gather together the terms in x_1 and complete the square in x_1; gather the remaining terms in x_2 and complete the square in x_2; etc. Thus

$$\begin{aligned} Q &= (x_1^2 + 4x_1x_2 + 2x_1x_3) + 5x_2^2 + 6x_2x_3 + 2x_3^2 \\ &= (x_1 + 2x_2 + x_3)^2 - (2x_2 + x_3)^2 + 5x_2^2 + 6x_2x_3 + 2x_3^2 \\ &= (x_1 + 2x_2 + x_3)^2 + x_2^2 + 2x_2x_3 + x_3^2 \\ &= (x_1 + 2x_2 + x_3)^2 + (x_2 + x_3)^2 . \end{aligned}$$

Q is positive, $Q \geq 0$. Further, $Q = 0$ if $x_1 + 2x_2 + x_3 = 0$ and $x_2 + x_3 = 0$, i.e. if $x_1 = s$, $x_2 = -s$, $x_3 = s$, where s is an arbitrary parameter. Since $Q(s, -s, s) = 0$, $s \in \mathbb{R}$, there are points besides $\mathbf{x} = \mathbf{0}$ where Q is zero, so Q is positive semi-definite (and not positive definite).

If $\mathbf{x}_0$ is a critical point of f, then equation (2.7) also still holds for n variables. We conclude the following *sufficient conditions* for a critical point $\mathbf{x}_0$ to be a local minimum or a maximum of f:

1. If $\mathbf{H}$ is positive definite, then $\mathbf{x}_0$ is a local minimum of $f(\mathbf{x})$.
2. If $\mathbf{H}$ is negative definite, then $\mathbf{x}_0$ is a local maximum of $f(\mathbf{x})$.
3. If $\mathbf{H}$ is positive semi-definite, then the test fails.
4. If $\mathbf{H}$ is negative semi-definite, then the test fails.
5. If $\mathbf{H}$ is indefinite, then $\mathbf{x}_0$ is a saddle point of $f(\mathbf{x})$.

4 Determination of Minima and Maxima Using the Hessian

Lagrange's method of completing the square can be used to determine if the associated quadratic form Q of a Hessian $\mathbf{H}$ is positive-definite or negative-definite (or otherwise), and hence whether a critical point is a local minimum or local maximum. We now give a second method to determine if the Hessian $\mathbf{H}$ (and its associated quadratic form Q) is positive-definite or negative-definite (or otherwise), which uses the signs of the eigenvalues of $\mathbf{H}$.

The eigenvalue method is contained in the following theorem, which applies to any (real) symmetric matrix $\mathbf{H}$, not only Hessians. Let $\lambda_1, \lambda_2, \cdots, \lambda_n$ be the eigenvalues of the symmetric matrix $\mathbf{H}$. Thus for each $i = 1, 2, \ldots, n$ there is a non-zero vector $\mathbf{v}_i$, such that $\mathbf{H}\mathbf{v}_i = \lambda_i\mathbf{v}_i$. Since $\mathbf{H}$ is symmetric the eigenvalues are all real. Let us order the eigenvalues so that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.

Eigenvalue Test for a Local Maximum or Minimum:

1. $\mathbf{H}$ is *positive definite* if and only if all its eigenvalues are positive, i.e. $\lambda_1 > 0$.
2. $\mathbf{H}$ is *negative definite* if and only if all its eigenvalues are negative, i.e. $\lambda_n < 0$.
3. $\mathbf{H}$ is *indefinite* if and only if it has positive and negative eigenvalues, i.e. $\lambda_1 < 0$ and $\lambda_n > 0$.
4. $\mathbf{H}$ is *positive semi-definite* if and only if all its eigenvalues are non-negative and at least one is zero, i.e. $\lambda_1 = 0$.
5. $\mathbf{H}$ is *negative semi-definite* if and only if all its eigenvalues are non-positive and at least one is zero, i.e. $\lambda_n = 0$.

The test follows because we can always diagonalize the matrix $\mathbf{H}$. Thus we can write an $n \times n$ matrix $\mathbf{H}$, which possesses a linearly independent set of n eigenvectors, in the form $\mathbf{H} = \mathbf{PDP}^{-1}$, where the columns of the invertible matrix $\mathbf{P}$ are the eigenvectors of $\mathbf{H}$ and the diagonal entries of the diagonal matrix $\mathbf{D}$ are the corresponding eigenvalues. In the special case where $\mathbf{H}$ is symmetric, as it is here, $\mathbf{P}$ can be chosen so that $\mathbf{P}^{-1} = \mathbf{P}^T$, i.e. the columns of $\mathbf{P}$ are orthonormal eigenvectors, i.e. all the eigenvectors are of length one and the scalar product of any two eigenvectors is zero. Thus we have

$$
\begin{aligned}
Q &= \mathbf{x}^T\mathbf{H}\mathbf{x} = \mathbf{x}^T\mathbf{P}\mathbf{D}\mathbf{P}^T\mathbf{x} = \mathbf{y}^T\mathbf{D}\mathbf{y} && (4.1)\\
&= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 && (4.2)\\
&= \begin{pmatrix} y_1 & y_2 & \cdots & y_n \end{pmatrix} \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, && (4.3)
\end{aligned}
$$

where $\mathbf{y} := \mathbf{P}^T\mathbf{y}$ and $\mathbf{x} = \mathbf{P}\mathbf{y}$, so $\mathbf{x}^T = \mathbf{y}^T\mathbf{P}^T$.

The theorem does not use the eigenvectors $\mathbf{v}_i$, so we only need to find the eigenvalues of the Hessian to determine if a critical point is a local maximum, local minimum or saddle, as the theorem above then determines whether the Hessian is positive-definite, negative-definite or indefinite (or otherwise semi-definite). Of course, if the Hessian is semi-definite the test fails.

5 Examples

Example: Find all critical points for the function

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

and determine the nature of the critical points.

First we look for the critical points, i.e. those points at which the first derivatives of f are zero. Take the first derivatives of f with respect to x, y and set these to zero,

$$\frac{\partial f}{\partial x} = 2x\,, \qquad \frac{\partial f}{\partial y} = 2y\,.$$

There is just one critical point at $(x, y) = (0, 0)$.

Then we form the Hessian $\mathbf{H}$, which happens to have constant entries, as the function f is only quadratic in x and y.

$$\mathbf{H} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Now $\mathbf{H}$ happens to be diagonal, so that the eigenvalues can be read off as the diagonal entries, 2,2. Thus $\mathbf{H}$ is positive-definite, so that a local minimum is attained at $(0, 0)$, as would be expected from elementary considerations. (Clearly this is also the global minimum.)

Example: Find all the critical points of the function

$$f(x, y, z) = 3xy - x^3 - y^3 - 3z^2$$

and classify them.

First we find the critical points. Take the first derivatives of f with respect to x, y, z and set these to zero,

$$\begin{aligned} \frac{\partial f}{\partial x} &= 3y - 3x^2 = 0 \\ \frac{\partial f}{\partial y} &= 3x - 3y^2 = 0 \\ \frac{\partial f}{\partial z} &= -6z = 0\,. \end{aligned}$$

The only real solutions to these equations are $x = 0$, $y = 0$, $z = 0$ and $x = 1$, $y = 1$, $z = 0$. Thus there are two critical points, $(0, 0, 0)$ and $(1, 1, 0)$.

The second partial derivatives of f are:

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= -6x \\ \frac{\partial^2 f}{\partial y^2} &= -6y \\ \frac{\partial^2 f}{\partial z^2} &= -6 \\ \frac{\partial^2 f}{\partial x \partial y} &= 3 \\ \frac{\partial^2 f}{\partial x \partial z} &= 0 \\ \frac{\partial^2 f}{\partial y \partial z} &= 0 \end{aligned}$$

So the Hessian is

$$\mathbf{H}(x, y, z) = \begin{pmatrix} -6x & 3 & 0 \\ 3 & -6y & 0 \\ 0 & 0 & -6 \end{pmatrix}\,.$$

At $(0,0,0)$ the Hessian is

$$\mathbf{H}(0,0,0) = \begin{pmatrix} 0 & 3 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & -6 \end{pmatrix} .$$

Since $\mathbf{Hv} = \lambda\mathbf{v}$, where $\mathbf{v}$ is non-zero, the matrix $\mathbf{H} - \lambda\mathbf{I}$ is singular and the eigenvalues are the zeros of the characteristic polynomial $\det(\mathbf{H} - \lambda\mathbf{I})$ of $\mathbf{H}$.

$$\begin{aligned} \det(\mathbf{H} - \lambda\mathbf{I}) = 0 &\Rightarrow -\lambda(-\lambda)(-6-\lambda) - 3(3)(-6-\lambda) = 0 \\ &\Rightarrow -\lambda^3 - 6\lambda^2 + 54 + 9\lambda = 0 \\ &\Rightarrow \lambda^3 + 6\lambda^2 - 9\lambda - 54 = 0 \\ &\Rightarrow (\lambda - 3)(\lambda + 3)(\lambda + 6) = 0 \,. \end{aligned}$$

Thus the eigenvalues of $\mathbf{H}(0,0,0)$ are

$$\lambda = 3,\ -3,\ -6.$$

Therefore $\mathbf{H}(0,0,0)$ is indefinite and $(0,0,0)$ is a saddle point.

At the other critical point $(1,1,0)$,

$$\mathbf{H}(1,1,0) = \begin{pmatrix} -6 & 3 & 0 \\ 3 & -6 & 0 \\ 0 & 0 & -6 \end{pmatrix} .$$

MATLAB or hand calculation as above, shows that $\lambda = -3,\ -6,\ -9$, $\mathbf{H}(1,1,0)$ is negative definite and hence there is a local maximum at $(1,1,0)$.

Chapter 4

Nonlinear Optimisation with Constraints

1 Lagrange Multipliers

We start with a physical example to show how Lagrange multipliers arise naturally in optimisation problems subject to constraints.

Example: *Engine Piston Design.* The problem is to find the relation between the piston radius x and the stroke y, which minimises the rate of heat loss for a cylinder of fixed capacity V (see Figure 4.1). We assume that the rate of heat loss is directly proportional to the surface area of the cylinder. Since the volume of the cylinder is fixed, the problem is equivalent to minimising the rate of heat loss per unit volume for a fixed volume.

The volume of the cylinder is

$$V = \pi x^2 y$$

and the surface area S is

$$S = 2\pi xy + 2\pi x^2 \,.$$

The rate of heat loss per unit volume is proportional to $f(x, y)$, where

$$f(x, y) := \frac{S}{V} = \frac{2\pi xy + 2\pi x^2}{\pi x^2 y} = \frac{2}{x} + \frac{2}{y}.$$

Thus, to minimise heat loss, while keeping the cylinder volume V fixed, we need to solve the problem:

$$\text{Minimise:} \qquad f(x, y) := \frac{2}{x} + \frac{2}{y} \tag{1.1}$$

$$\text{subject to:} \qquad g(x, y) := \pi x^2 y - V = 0 \,. \tag{1.2}$$

Referring to Figure 4.2, we see that the slopes of the objective function $f(x, y)$ and the constraint $g(x, y) = 0$ are equal at the point P. Thus ∇f is parallel to ∇g at P, i.e.

$$\nabla f = -\lambda \nabla g$$

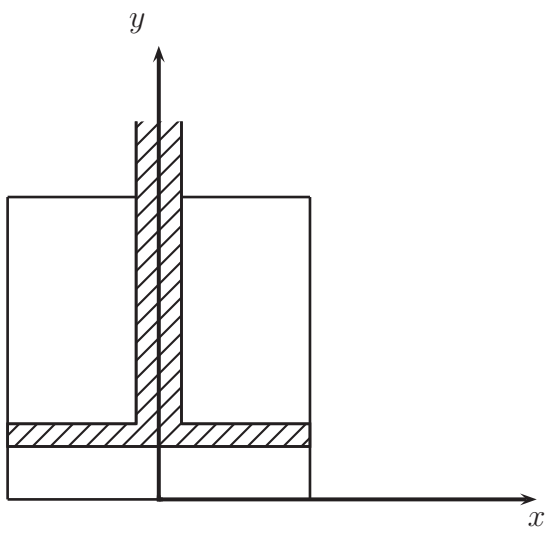


Figure 4.1: Schematic of an engine piston

for some constant λ. The two components of this vector equation are

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0$$
$$\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 0$$

at P. These two conditions together with the constraint, $g(x, y) = 0$, constitute three equations in the three unknowns x, y, λ for the minimum. If we define the *Lagrangian function* or *Lagrangian* L, by

$$L := f + \lambda g \,,$$

then equivalent necessary conditions for a minimum of the problem in terms of L are

$$\frac{\partial L}{\partial x} = 0 \,, \qquad \frac{\partial L}{\partial y} = 0 \,, \qquad \frac{\partial L}{\partial \lambda} = 0 \,.$$

In other words, if we regard the Lagrange multiplier as an additional independent variable, the conditions for a minimum arise by requiring all first derivatives of L to be zero, which is equivalent to seeking a minimum of L. In the particular case of the piston problem, these independent variables are x, y and λ, and we have

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = -\frac{2}{x^2} + 2\pi\lambda xy = 0$$
$$\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = -\frac{2}{y^2} + \pi\lambda x^2 = 0 \,.$$

Thus

$$\lambda = \frac{1}{\pi x^3 y} = \frac{2}{\pi x^2 y^2} \,,$$

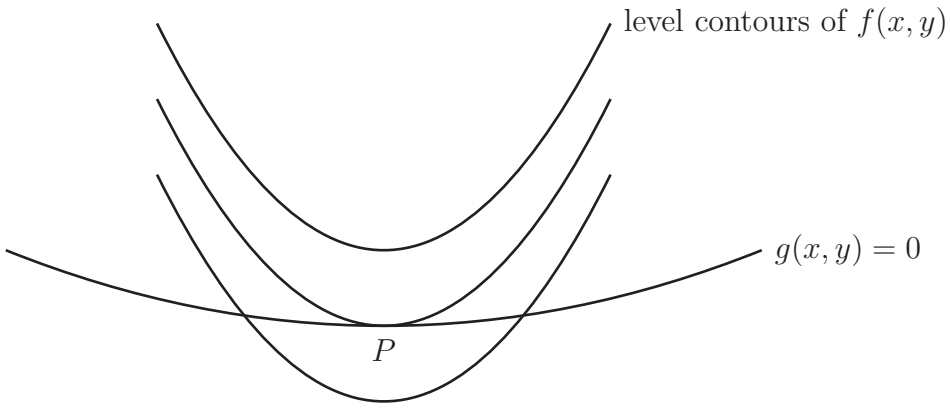


Figure 4.2: Contours of f subject to the constraint $g(x, y) = 0$

which implies $y = 2x$. In this case the constraint equation gives

$$g(x, y) := \pi x^2 y - V = 0 \Rightarrow \frac{\pi y^3}{4} = V\,,$$

so that $y = \sqrt[3]{4V/\pi}$, $x = \frac{1}{2}\sqrt[3]{4V/\pi}$.

2 Equality Constraints

We now consider the constrained optimisation problem in n variables with m equality constraints,

$$\begin{aligned} \text{Minimise:} \quad & f(x_1, x_2, \ldots, x_n) \\ \text{subject to:} \quad & g_1(x_1, x_2, \ldots, x_n) = 0 \\ & g_2(x_1, x_2, \ldots, x_n) = 0 \\ & \qquad\qquad\qquad \vdots \\ & g_m(x_1, x_2, \ldots, x_n) = 0\,, \end{aligned}$$

where $m < n$, i.e. the number of constraints is less than the number of variables. The g_j are the *constraint functions* and f is the objective function. In vector form,

$$\text{Minimise:} \quad f(\mathbf{x}) \tag{2.1}$$

$$\text{subject to:} \quad \mathbf{g}(\mathbf{x}) = \mathbf{0}\,, \tag{2.2}$$

where

$$\mathbf{g} := \begin{pmatrix} g_1(\mathbf{x}) \\ g_2(\mathbf{x}) \\ \vdots \\ g_m(\mathbf{x}) \end{pmatrix}\,.$$

Geometrically, the constraints define a subset Ω of $\mathbb{R}^n$, the *feasible* set, which is of dimension less than $n-m$, in which the minimum lies.

The constrained optimisation problem (1.1) with the single constraint (1.2) led to the introduction of a single Lagrange multiplier. More generally, one Lagrange multiplier is introduced for each constraint. Thus for the general problem the Lagrangian L is defined by

$$L(\mathbf{x}, \boldsymbol{\lambda}) := f + \sum_{i=1}^{m} \lambda_i g_i \,. \tag{2.3}$$

In order to find the minimum of f, subject to all the constraints, we set to zero the first partial derivatives of L with respect to the x_j, $j = 1, \ldots, n$, and the λ_i, $i = 1, \ldots, m$. We obtain the following first-order *necessary conditions* for a minimum:

$$\frac{\partial L}{\partial x_j} = \frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0 \,, \qquad j = 1, \ldots, n \,, \tag{2.4}$$

and

$$\frac{\partial L}{\partial \lambda_i} = g_i = 0 \,, \qquad i = 1, \ldots, m \,. \tag{2.5}$$

There are $n+m$ equations for the n variables x_j and the m Lagrange multipliers λ_i. The necessary conditions are those for the *unconstrained* optimisation of L.

Example: Minimize $f(x, y) = \frac{1}{2}(x-1)^2 + \frac{1}{2}(y-2)^2 + 1$ subject to the single constraint $x + y = 1$.

Solution: The first step is to construct the Lagrangian:

$$L(x, y, \lambda) = \tfrac{1}{2}(x-1)^2 + \tfrac{1}{2}(y-2)^2 + 1 - \lambda(x + y - 1) \,.$$

Then the necessary conditions are:

$$\begin{aligned} \frac{\partial L}{\partial x} &= x - 1 - \lambda = 0 \\ \frac{\partial L}{\partial y} &= y - 2 - \lambda = 0 \\ x + y &= 1 \,. \end{aligned}$$

The solution of this simple 3×3 system of linear equations is

$$\lambda = -1 \,, \qquad x = 0 \,, \qquad y = 1 \,,$$

and the corresponding min $f_=2$. Note this solution could also be obtained by substituting $y = 1 - x$ into L and then minimizing the resulting function of the single variable x.

3 Convex Functions and Convex Sets

The first-order necessary conditions (2.4), (2.5) are also sufficient, if the function f is *convex* and the constraints are linear.

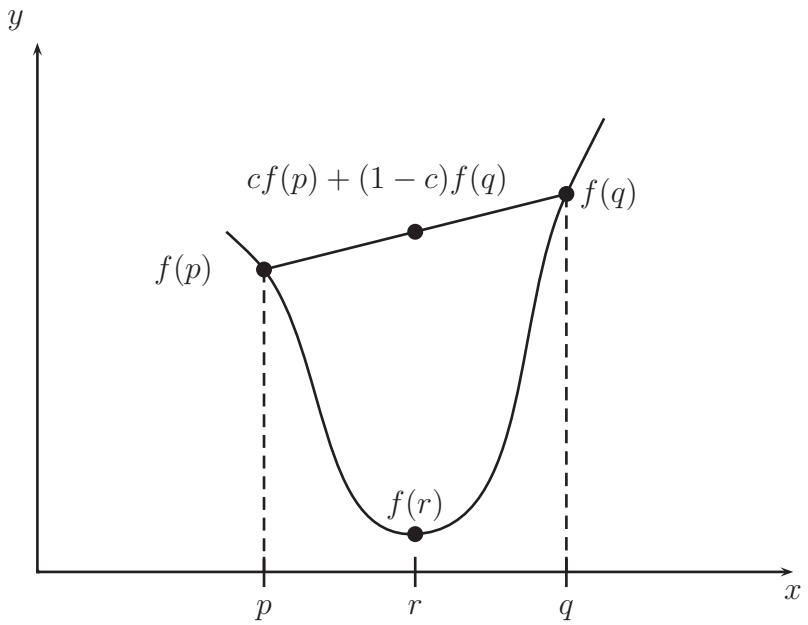


Figure 4.3: A convex function of a single variable $f(x)$

Definition: A set $\Omega \in \mathbb{R}^n$ is *convex*, if for any points P, Q in Ω, the line segment PQ joining P and Q lies in Ω. If $\mathbf{p}$ and $\boldsymbol{q}$ are the position vectors of P and Q, then the point R with position vector $\mathbf{r} = c\mathbf{p} + (1-c)\boldsymbol{q}$ lies in Ω, where $0 \leq c \leq 1$. A function $f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n)$ is *convex* on the convex set Ω, if

$$f(c\mathbf{p} + (1-c)\boldsymbol{q}) \leq cf(\mathbf{p}) + (1-c)f(\boldsymbol{q})\,,$$

where $0 \leq c \leq 1$. This is illustrated in Figure 4.3 for the one-dimensional case. Note in particular, that linear functions are convex.

Example: The problem is to find

$$\begin{aligned} \text{Minimise:} \quad & f(x,y) = x^2 + y^2 - xy - 5y \\ \text{subject to:} \quad & x + 2y = 5\,. \end{aligned}$$

The constraint is linear, so we expect that if we can find suitable λ, then we can find an optimal solution. The Lagrangian is

$$L(x, y, \lambda) = x^2 + y^2 - xy - 5y + \lambda(x + 2y - 5)\,,$$

so that

$$\frac{\partial L}{\partial x} = 2x - y + \lambda = 0 \tag{3.1}$$

$$\frac{\partial L}{\partial y} = 2y - x - 5 + 2\lambda = 0 \tag{3.2}$$

$$\frac{\partial L}{\partial \lambda} = x + 2y - 5 = 0\,. \tag{3.3}$$

Now equation (3.1) implies

$$y = 2x + \lambda$$

and equation (3.2) implies

$$x = 2y - 5 + 2\lambda\,,$$

so that

$$y = \tfrac{10}{3} - \tfrac{5}{3}\lambda\,, \qquad x = \tfrac{5}{3} - \tfrac{4}{3}\lambda\,.$$

If we substitute these two expressions into (3.3) we find that $\lambda = \frac{5}{7}$, so that finally

$$x = \tfrac{5}{7}\,, \qquad y = \tfrac{15}{7}\,.$$

4 Inequality Constraints

We now consider constrained optimisation problems with m inequality constraints,

$$\begin{aligned} \text{Minimise:} \quad & f(x_1, x_2, \ldots, x_n) \\ \text{subject to:} \quad & h_1(x_1, x_2, \ldots, x_n) \le 0 \\ & h_2(x_1, x_2, \ldots, x_n) \le 0 \\ & \qquad\qquad \vdots \\ & h_m(x_1, x_2, \ldots, x_n) \le 0\,. \end{aligned}$$

In vector form the problem is

$$\text{Minimise:} \quad f(\mathbf{x}) \tag{4.1}$$

$$\text{subject to:} \quad \mathbf{h}(\mathbf{x}) \le \mathbf{0},. \tag{4.2}$$

(If $\mathbf{u}$ and $\mathbf{v}$ are two vectors in $\mathbb{R}^m$, the vector inequality $\mathbf{u} \le \mathbf{v}$ is defined by $u_i \le v_i$ for $i = 1, \ldots, m$.) The formulation (4.1), (4.2) in terms of $\le$ constraints is, in fact, quite general. Any $\ge$ constraint can be converted to $\le$ constraint by multiplying it by -1, while equalities such as $h(\mathbf{x}) = 0$ are equivalent to two $\le$ constraints, $h(\mathbf{x}) \le 0$ and $-h(\mathbf{x}) \le 0$. (The latter inequality is equivalent to $h(\mathbf{x}) \ge 0$.)

4.1 The Kuhn-Tucker Conditions

A point $\mathbf{x}$ is a minimum of the optimisation problem, (4.1) and (4.2), only if the following first-order necessary conditions, which are known as the *Kuhn-Tucker conditions*, and (4.2) are satisfied at $\mathbf{x}$: there exist m Lagrange multipliers λ_i, $i = 1, \ldots, m$, such that

$$\frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial h_i}{\partial x_j} = 0\,, \qquad j = 1, \ldots, n\,; \tag{4.3}$$

$$\lambda_i h_i = 0\,, \qquad i = 1, \ldots, m\,; \tag{4.4}$$

$$\lambda_i \geq 0\,, \qquad i = 1, \ldots, m\,. \tag{4.5}$$

Note that in a maximisation problem the sign of f must be changed.

The conditions (4.4) and (4.5), together with the original constraints (4.2), are equivalent to either

$$h_i(\mathbf{x}) = 0\,, \qquad \text{if } \lambda_i > 0,\ i = 1, \ldots, m \tag{4.6}$$

or

$$h_i(\mathbf{x}) < 0\,, \qquad \text{if } \lambda_i = 0,\ i = 1, \ldots, m. \tag{4.7}$$

First we note that these relations reduce to the familiar Lagrange multipliers case, if we have *equality* constraints. Although we appear to have an additional requirement, that the λ_i's should all be greater than or equal to zero, any original equality constraint $h = 0$ must have been replaced by two inequality constraints $h \leq 0$, $-h \leq 0$ for which there must be two Lagrange multipliers, say λ^+, λ^-. Condition (4.4) implies both $\lambda^+ h = 0$ and $-\lambda^- h = 0$, and hence $h = 0$.

These conditions are always necessary. If the objective function f and the constraint functions h_i, $i = 1, \ldots, m$, are convex, then the Kuhn-Tucker conditions are also sufficient. Therefore, in particular, the conditions are sufficient for a positive definite quadratic objective function with linear constraints, an important class of problem known as *quadratic programming*.

That the Kuhn-Tucker conditions are necessary can be demonstrated as follows. First the λ_i must all be greater than or equal to zero. Each of the constraints in (4.2) can be converted to an equality by adding a slack variable s_i^2 to each constraint, where the square is taken to ensure that the contribution from the slack variable is greater than or equal to zero. Then we can write down the Lagrangian,

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{s}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i [h_i(\mathbf{x}) + s_i^2]\,,$$

where $\boldsymbol{s}$ is a vector in $\mathbb{R}^m$ with i-th component s_i. The Lagrangian L depends on the $n + 2m$ *independent* variables $x_1, x_2, \ldots, x_n$, $\lambda_1, \lambda_2, \ldots, \lambda_m$ and $s_1, s_2, \ldots, s_m$. To find the constrained minimum we must differentiate L with respect to each of these variables and set the result to zero in each case, i.e. we are effectively looking for a minimum of L. For a well-defined minimum of L, it is necessary that $\lambda_i \geq 0$ for $i = 1, \ldots, m$, otherwise, if $\lambda_i < 0$, then $L \to -\infty$ as $\lambda_i \to -\infty$ due to the term $\lambda_i s_i^2$ and L has no minimum.

Equating to zero the derivative of L with respect to x_j, i.e. $\partial L/\partial x_j = 0$, gives the equations (4.3). Differentiation of L with respect to λ_i and equating the result to zero, i.e. $\partial L/\partial \lambda_i = 0$, yields

$$h_i + s_i^2 = 0\,, \qquad i = 1, \ldots, m\,. \tag{4.8}$$

Finally, setting to zero the derivative of L with respect to s_i, i.e. $\partial L/\partial s_i = 0$, gives

$$2\lambda_i s_i = 0\,, \qquad i = 1, \ldots, m\,. \tag{4.9}$$

Now either $\lambda_i > 0$ or $\lambda_i = 0$. In the first case $s_i = 0$ from (4.9) and hence (4.8) implies (4.6), while in the second case we obtain (4.7) using (4.2). As these conditions are equivalent to (4.4) and (4.5) together with the original constraints (4.2) we are done.

4.2 Examples

We now consider the following simple examples to demonstrate how these conditions work in practice.

Example 1: Consider the general problem

$$\begin{aligned} \text{Minimise:} \quad & f(x) \\ \text{subject to:} \quad & a \le x \le b\,. \end{aligned}$$

The Kuhn-Tucker conditions follow from the modified formulation

$$\begin{aligned} \text{Minimise:} \quad & f(x) \\ \text{subject to:} \quad & -x \le -a \\ & x \le b\,. \end{aligned}$$

Solution: The Lagrangian L is given by

$$L = f + \lambda_1(-x + a + s_1^2) + \lambda_2(x - b + s_2^2)\,.$$

Then

$$\frac{\partial L}{\partial x} = f'(x) - \lambda_1 + \lambda_2 = 0$$

and the partial derivatives with respect to the λ_i give

$$\begin{aligned} -x + a + s_1^2 &= 0 \\ x - b + s_2^2 &= 0\,. \end{aligned}$$

Finally, on differentiating with respect to the s_i we find

$$2\lambda_1 s_1 = 0\,, \qquad 2\lambda_2 s_2 = 0\,.$$

Considering the equations involving λ_1 and s_1, we have

$$\lambda_1 = 0\,, \qquad -x + a + s_1^2 = 0$$

or

$$\lambda_1 > 0\,, \qquad -x + a = 0 \qquad \text{as } s_1 = 0,$$

with similar equations for λ_2.

The minimum either occurs at an interior turning point, or at one of the two ends of the interval, dependent on the form of $f(x)$. The Kuhn-Tucker conditions give all three possibilities, as follows.

DOWN-variable $\lambda_1 > 0 \quad (\Rightarrow s_1 = 0), \quad \lambda_2 = 0.$

This implies that $x = a$, and $f'(x) = f'(a) = \lambda_1 > 0$, so that the minimum lies at $x = a$, which is a boundary point of the domain. See Figure 4.4.

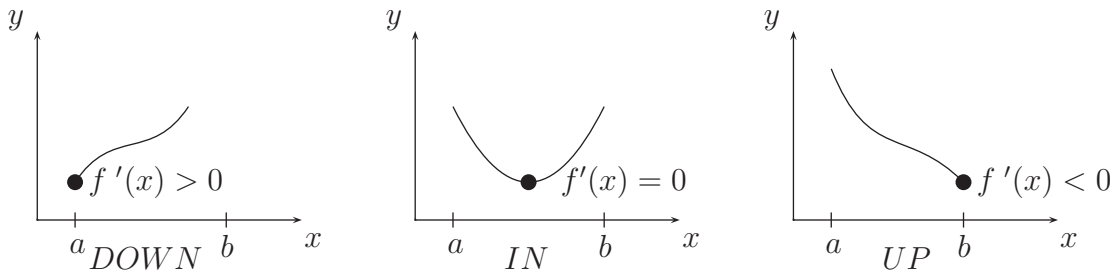


Figure 4.4: The three cases in Example 1.

IN-variable $\lambda_1\,, \quad \lambda_2 = 0.$

This implies $f'(x) = 0$, for some $x \in [a, b]$ determined by the specific $f(x)$, and thus there is a turning point somewhere in the interval $[a, b]$.

UP-variable $\lambda_1 = 0\,, \quad \lambda_2 > 0 \quad (\Rightarrow s_2 = 0).$

This implies $x = b$, so that $f'(x) = f'(b) = -\lambda_2 < 0$ and we have the minimum at $x = b$, which is the other boundary point of the domain.

Thus the Kuhn-Tucker conditions have given us the three possibilities in this simple case.

Example 2:

$$\begin{aligned} \text{Minimise:} \qquad & f(x_1, x_2, x_3) = x_1(x_1 - 10) + x_2(x_2 - 50) - 2x_3 \\ \text{subject to:} \qquad & x_1 + x_2 \leq 10 \\ & x_3 \leq 10\,. \end{aligned}$$

We set

$$L = f + \lambda_1(g_1 - b_1) + \lambda_2(g_2 - b_2)\,,$$

where

$$g_1 = x_1 + x_2 + s_1^2$$

and

$$g_2 = x_3 + s_2^2\,,$$

with s_1, s_2 slack variables, and where the right sides are $b_1 = 10$, $b_2 = 10$. Then

$$\begin{aligned} \frac{\partial L}{\partial x_1} &= \frac{\partial f}{\partial x_1} + \lambda_1 \frac{\partial g_1}{\partial x_1} + \lambda_2 \frac{\partial g_2}{\partial x_1} \\ &= 2x_1 - 10 + \lambda_1 = 0 \qquad (4.10) \\ \frac{\partial L}{\partial x_2} &= 2x_2 - 50 + \lambda_1 = 0 \qquad (4.11) \\ \frac{\partial L}{\partial x_3} &= -2 + \lambda_2 = 0 \qquad (4.12) \end{aligned}$$

and the remaining Kuhn-Tucker conditions are

$$\lambda_1(x_1 + x_2 - 10) = 0 \tag{4.13}$$
$$\lambda_2(x_3 - 10) = 0 \tag{4.14}$$

and

$$\lambda_1\,,\ \lambda_2 \geq 0\,. \tag{4.15}$$

From (4.12) we find that $\lambda_2 = 2$. Thus from (4.14) $x_3 = 10$. Now there are two cases to consider.

Case 1 $\lambda_1 = 0$. From equation (4.10) we find that $x_1 = 5$, and from (4.11) we find $x_2 = 25$. However, this does not satisfy the original constraint $x_1 + x_2 \leq 10$, and so we reject this solution. (It is in fact the unconstrained minimum.)

Case 2 $\lambda_1 > 0$ From equation (4.13) $x_1 + x_2 = 10$. When this is used in (4.11) we find:

$$-2x_1 - 30 + \lambda_1 = 0\,.$$

But (4.10) is

$$2x_1 - 10 + \lambda_1 = 0\,,$$

so that $\lambda_1 = 20$ and $x_1 = -5$, $x_2 = 15$.

Thus finally we find that the feasible minimum occurs at $\mathbf{x} = (-5, 15, 10)$, (with $\lambda_1 = 20$, $\lambda_2 = 2$). The function value $f_{\min} = -450$ at this point.

4.3 Bounded Variable Constraints

We extend Example 1 of the previous section to n variables: we consider here inequality constraints of the form

$$L_i \leq x_i \leq U_i\,,$$

where L_i, U_i are constants describing respectively the lower and upper bounds of the variable x_i. In vector form, $\mathbf{L} \leq \mathbf{x} \leq \mathbf{U}$.

The optimisation problem for this situation can be stated as:

$$\begin{aligned} \text{Minimise:} \quad & f(\mathbf{x}) \\ \text{subject to:} \quad & \mathbf{g}(\mathbf{x}) = \mathbf{b} \\ & \mathbf{L} \leq \mathbf{x} \leq \mathbf{U}\,. \end{aligned}$$

For inequality constraints there are three cases to consider for each x_i as in Figure 4.4.

Thus the feasible minimum point lies either in the region $L_i < x_i < U_i$ or on either boundary $x_i = L_i$; $x_i = U_i$. These three cases are summarized in the table below:

Type	Condition	$\partial F/\partial x_i$	Constraint
DOWN	$x_i = L_i$	> 0	L_i-Active
IN	$L_i < x_i < U_i$	0	Both Inactive
UP	$x_i = U_i$	< 0	U_i-Active

The optimal solution is found by introducing Lagrange parameters: one for each inequality constraint imposed. The Lagrangian now takes the form:

$$L = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i [g_i(\mathbf{x}) - b_i] + \boldsymbol{\ell}^T(\mathbf{L} - \mathbf{x}) + \boldsymbol{u}^T(\mathbf{x} - \mathbf{U})\,,$$

where $\boldsymbol{\ell}$ and $\boldsymbol{u}$ are vectors of Lagrangian parameters for the lower and upper inequality constraints.

The first-order Kuhn-Tucker conditions for the optimal solution are

$$\begin{aligned} \frac{\partial L}{\partial x_j} = \frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} - \ell_j + u_j = 0 \\ g_i(\mathbf{x}) = 0 \\ \ell_j(x_j - L_j) = 0\,, \qquad \ell_j \geq 0 \\ u_j(U_j - x_j) = 0\,, \qquad u_j \geq 0\,. \end{aligned}$$

Remarks:

1. The equation $\ell_j(x_j - L_j) = 0$ implies that either $\ell_j = 0$ and $x_j > L_j$ (inactive); or $\ell_j > 0$ and $x_j = L_j$ (active). A similar interpretation applies to the equation $u_j(x_j - U_j) = 0$.
2. Let $L(\mathbf{x}) = f(\mathbf{x}) + \sum \lambda_i g_i(\mathbf{x})$. If x_j is an IN-variable we have $\ell_j = u_j = 0$, implying also that $\partial L/\partial x_j = 0$.
3. If $x_j = L_j$, the lower constraint is active and $\ell_j > 0$, $u_j = 0$. Hence $\partial L/\partial x_j = \ell_j > 0$.
4. If $x_j = U_j$, the upper constraint is active and $\ell_j = 0$, $u_j > 0$. Hence $\partial L/\partial x_j = -u_j < 0$.
5. If there is no inequality constraint $x_j \geq L_j$ or $x_j \leq U_j$ we may set $\ell_j = 0$ (or $u_j = 0$), equivalent to an inactive constraint.

Example: Minimize $f(x, y) = \frac{1}{2}(x-1)^2 + \frac{1}{2}(y-2)^2 + 1$, subject to $x + y = 1$ and the inequality constraints $x \geq \frac{1}{2}$ and $y \leq \frac{3}{4}$.

Solution: Here we have an optimisation problem with one equality constraint and two inequality constraints. The Lagrangian is given by:

$$L = f(x, y) + \lambda(x + y - 1) + \ell(\tfrac{1}{2} - x) + u(y - \tfrac{3}{4})\,.$$

The KT-conditions are

$$\begin{aligned} \frac{\partial L}{\partial x} = x - 1 + \lambda - \ell = 0 \\ \frac{\partial L}{\partial y} = y - 2 + \lambda + u \\ x + y = 1 \\ \ell(x - \tfrac{1}{2}) = 0\,, \qquad \ell \geq 0 \\ u(\tfrac{3}{4} - y) = 0\,, \qquad u \geq 0\,. \end{aligned}$$

We solve these equations by trial and error.

1. Assume x, y are both IN-variables. Then $\ell = u = 0$ and we get $x = 0$ and $y = 1$. This violates the inequality constraints and we conclude that x, y cannot both be IN.

2. Assume x is IN and $y = 3/4$ is UP. Then $\ell = 0$ and $u > 0$ and we get the solution: $x = 1/4$ which violates $x \geq 1/2$. Hence this is not the correct solution either.

3. Assume $x = 1/2$ is DOWN and y is IN. Then $\ell > 0$ and $u = 0$ and we obtain: $y = 1/2$, $\ell = 1 > 0$. This solution satisfies all the KT-conditions.

   The unique solution is

$$x = \tfrac{1}{2}\,, \qquad y = \tfrac{1}{2}\,, \qquad \lambda = \tfrac{3}{2}\,, \qquad \ell = 1\,, \qquad u = 0\,.$$

The corresponding $f_{\min} = \frac{9}{4}$. Note that the lower constraint $x \geq \frac{1}{2}$ is active, while the upper constraint $y \leq \frac{3}{4}$ is inactive. The value $\ell = 1$ tells us that the partial derivative $\partial L/\partial x = 1$ at the active boundary $x = \frac{1}{2}$.

Chapter 5

Probability Review

This chapter contains a brief review of the elements of mathematical probability required in this course. One of the key concepts needed is the *expectation* of a random variable, denoted by the operator symbol $\mathbb{E}$.

1 Discrete Random Variables

X is a *discrete random variable* (drv) if it takes values from a finite (or countable) set of real numbers $(x_1, x_2, \ldots, x_n)$ with probabilities $(p_1, p_2, \ldots, p_n)$ $(0 \leq p_i \leq 1$ and $\sum_i p_i = 1)$. The set of p_i is called the *probability function* or *probability distribution* of X and we write

$$p_i = \mathbb{P}\{X = x_i\}; \qquad i = 1, 2, \ldots, n\,.$$

Example: Let X be the number of heads occurring when a fair coin is tossed three times. Then we readily identify

$$x_i = \{0, 1, 2, 3\} \qquad \text{and} \qquad p_i = \{\tfrac{1}{8}, \tfrac{3}{8}, \tfrac{3}{8}, \tfrac{1}{8}\}$$

If the coin is not a fair one so that the probability of a head is $p \neq \frac{1}{2}$, we get, assuming each toss is independent of any other,

$$p_i = \{q^3, 3q^2p, 3qp^2, p^3\}$$

where $q = 1 - p$ is the probability of getting a tail.

This distribution is a special case of the *discrete binomial distribution* defined by

$$\mathbb{P}\{X = k\} = \left(\begin{array}{c} n \\ k \end{array} \right) p^k q^{n-k}; \qquad k = 0, 1, 2, \ldots, n$$

We may think of X, in a more general framework, as the number of *successes* in n independent *trials*, given that p is the probability of success in a single trial.

2 The Expectation Operator

Let X be a drv with probability function p_i on the set x_i for $i = 1, 2, \ldots, n$. Then the *expectation* (also called *expected value* and *mean value*) of X is defined to be

$$\boxed{\mathbb{E}\{X\} = \mu = \sum_{i=1}^{n} p_i x_i} \tag{2.1}$$

$\mathbb{E}$ is called the *expectation operator*. We see that $\mathbb{E}\{X\}$ is the weighted average of each allowed x_i, by its probability of occurrence. Note that $\mu = \mathbb{E}\{X\}$ is not a random variable.

Theorem: $\mathbb{E}$ *is a linear operator. That is, if* X *and* Y *are any two drv's and* a, b *are any non-random scalars, then*

$$\boxed{\mathbb{E}\{aX + bY\} = a\mathbb{E}\{X\} + b\mathbb{E}\{Y\}} \tag{2.2}$$

Proof: Let

$$p_{ij} = \mathbb{P}\{X = x_i \text{ and } Y = y_j\}$$

p_{ij} is called the *joint probability function* of X and Y. Define

$$p_i = \sum_j p_{ij} \qquad \text{and} \qquad q_j = \sum_i p_{ij}$$

as the *marginal distributions* of X and Y respectively. Then

$$\begin{aligned} \mathbb{E}\{aX + bY\} &= \sum_i \sum_j p_{ij}(ax_i + by_j) \\ &= a\sum_i p_i x_i + b\sum_j q_j y_j \\ &= a\mathbb{E}\{X\} + b\mathbb{E}\{Y\} \qquad \square \end{aligned}$$

Three special cases of this theorem will help explain its importance:

$$\begin{aligned} \mathbb{E}\{a\} &= a \\ \mathbb{E}\{X + Y\} &= \mathbb{E}\{X\} + \mathbb{E}\{Y\} \\ \mathbb{E}\{aX\} &= a\mathbb{E}\{X\} \end{aligned}$$

The first of these states that the expectation of any scalar (or non-random variable) is equal to that scalar; the second is the superposition principle: the expectation of a sum of random variables is equal to the sum of their expectations; the third is the scaling rule: the expectation of a times a random variable is equal to a times its expectation.

Example A2 For a fair coin tossed three times: $p = 1/2$ and $n = 3$ we get the expected number of heads equal to 1.5 from

$$\mathbb{E}\{X\} = (0 \times 1/8) + (1 \times 3/8) + (2 \times 3/8) + (3 \times 1/8) = 1.5$$

For the general binomial distribution $\mathbb{E}\{X\} = np$.

3 Expectation of $f(X)$

Let X be a drv on x_i with probability function p_i and let $Y = f(X)$ for any single-valued function f. Then Y is also a drv on $y_i = f(x_i)$ with the *same* probability function p_i. That is

$$\mathbb{P}\{X = x_i\} = p_i \qquad \text{implies} \qquad \mathbb{P}\{Y = y_i = f(x_i)\} = p_i$$

We therefore have the important result: $\mathbb{E}\{Y\} = \sum_i p_i y_i$ or equivalently

$$\boxed{\mathbb{E}\{f(X)\} = \sum_i p_i f(x_i).} \tag{3.1}$$

Definition: The *variance* of a drv X with mean $\mu = \mathbb{E}\{X\}$ is defined to be

$$\boxed{\mathbb{V}\{X\} = \sigma^2 = \mathbb{E}\{(X-\mu)^2\}} \tag{3.2}$$

Variance is a measure of the spread of a random variable about its mean. Thus variance is also a measure of uncertainty. In financial terms, we shall identify variance of a random security with its degree of risk.
Note that variance $\sigma^2 \geq 0$ and is equal to zero, if and only if $X = \mu = \mathbb{E}\{X\}$; that is if X is non-random. The parameter σ is called the *standard deviation*.

Theorem: For any drv X

$$\boxed{\mathbb{V}\{X\} = \mathbb{E}\{X^2\} - \mu^2} \tag{3.3}$$

Proof:

$$\begin{aligned} \sigma^2 &= \mathbb{E}\{(X-\mu)^2\} \\ &= \mathbb{E}\{X^2 - 2\mu X + \mu^2\} \\ &= \mathbb{E}\{X^2\} - 2\mu\mathbb{E}\{X\} + \mu^2 \qquad \text{by linearity of } \mathbb{E} \\ &= \mathbb{E}\{X^2\} - \mu^2 \qquad \text{since } \mathbb{E}\{X\} = \mu \qquad \square \end{aligned}$$

Example A3 For the discrete binomial distribution, $\sigma^2 = npq$. Hence, for the 3-toss of the coin problem, $n = 3, p = q = 1/2$ and we obtain $\sigma^2 = 3/4$.

Theorem: If X, Y are any *independent* random variables and a, b are any scalars (*i.e.* non-random constants), then

$$\boxed{\mathbb{V}\{aX + bY\} = a^2\mathbb{V}\{X\} + b^2\mathbb{V}\{Y\}.} \tag{3.4}$$

4 Continuous Random Variables

If the space of allowed values of a random variable X is continuous, we refer to X as a *continuous random variable* or crv for short. We then call its probability function, the *probability density function* or pdf for short. Loosely speaking, we may define the pdf $f(x)$ on some domain $x \in \Omega$ by:

$$\mathbb{P}\{x < X \leq x + dx\} \approx f(x)\, dx$$

where

$$f(x) \geq 0 \qquad \text{and} \qquad \int_\Omega f(x)\,dx = 1.$$

Notes:

1. A pdf need not satisfy the constraint $f(x) \leq 1$.

2. The domain of a crv is often extended to the whole real line $-\infty < x < \infty$ by defining $f(x) = 0$ for all x not in Ω.

A more formal way of defining the pdf of a continuous random variable is through its *cumulative distribution function* or cdf for short. This is the function

$$F(x) = \mathbb{P}\{X \leq x\}; \qquad -\infty < x < \infty.$$

We then have the following relations between the pdf and the cdf:

$$F(x) = \int_\infty^x f(x')\,dx'; \qquad f(x) = \frac{dF}{dx}.$$

Mean,Variance and Expectation
Let X be a continuous random variable on Ω. Then the mean, variance and general expectation of an arbitrary function $g(X)$ are defined by:

1. Mean $\mathbb{E}\{X\} = \mu = \int\limits_\Omega x f(x)\,dx$
2. Variance $\mathbb{V}\{X\} = \sigma_2 = \int\limits_\Omega (x-\mu)^2 f(x)\,dx$
3. Expectation $\mathbb{E}\{g(X)\} = \int\limits_\Omega g(x) f(x)\,dx$
4. Moment Generating Function (MGF) $M(\theta) = \mathbb{E}\{e^\theta X\} = \int\limits_\Omega e^{\theta x} f(x)\,dx.$

Basically all the properties of expectations of drv's carry over to crv's, with discrete probability functions being replaced by pdf's and sums by integrals over the domain Ω.

5 Some Standard Probability Distributions

In this section we shall list a few of the best known probability distributions together with their means and variances.

1. **Discrete Binomial Distribution**
   X = No. successes in $n-$independent trials where p is the probability of success in a single trial.
   Prob. function $p_k = \left(\begin{array}{c} n \\ k \end{array} \right) p^k q^{n-k}$ with $k = 0, 1, 2, \ldots, n$ and $q = 1 - p$.
   Mean $\mu = np$; Variance $\sigma^2 = npq$; MGF $M(\theta) = (pe^\theta + q)^n$.

2. **Discrete Poisson Distribution**
   X = No. of successes in some interval of time (or space), given m = average no. successes in that interval.
   Prob. function $p_r = m^r e^{-m}/r!$ for $r = 0, 1, 2, \ldots, \infty$.
   Mean $\mu = m$; Variance $\sigma^2 = m$; MGF $M(\theta) = e^{(m-1)\theta}$.

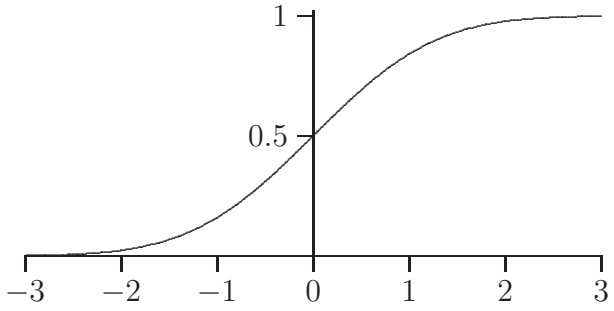
3. **Discrete Geometric Distribution**
   X = No. of independent trials to achieve the first success, given p = probability of success of a single trial.
   Prob. function $p_k = q^{k-1}p$ for $k = 1, 2, \ldots, \infty$.
   Mean $\mu = 1/p$; Variance $\sigma^2 = q/p^2$; MGF $M(\theta) = p/(e^{-\theta} - q)$.

4. **Continuous Uniform Distribution**
   X is uniformly distributed on the interval (a, b) if its pdf is given by:
   Pdf: $f(x) = \begin{cases} \frac{1}{b-a} & \text{if} \quad a < x < b \\ 0 & \text{otherwise} \end{cases}$
   Mean $\mu = \frac{1}{2}(a+b)$; Variance $\sigma^2 = \frac{1}{12}(b-a)^2$; MGF $M(\theta) = \frac{e^{b\theta} - e^{a\theta}}{(b-a)\theta}$.

5. **Continuous Exponential Distribution**
   Pdf: $f(x) = \begin{cases} \frac{1}{\lambda}e^{-x/\lambda} & \text{if} \quad x > 0 \\ 0 & \text{otherwise} \end{cases}$
   Mean $\mu = \lambda$; Variance $\sigma^2 = \lambda^2$; MGF $M(\theta) = (1 - \lambda\theta)^{-1}$.

6. **Continuous Normal (Gaussian) Distribution**
   This is the celebrated *bell curve* with:
   Pdf: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} = \frac{1}{\sigma}\phi(\frac{x-\mu}{\sigma})$ for $-\infty < x < \infty$, where $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ is the pdf of a standard $N(0, 1)$ variate.
   The mean and variance are of course μ and σ^2 respectively.
   The notation $X \sim N(\mu, \sigma^2)$ is used to indicate that the crv X is normally distributed with mean μ and variance σ^2 and the MGF is $M(\theta) = e^{\mu\theta + \frac{1}{2}\sigma^2\theta^2}$.

Normal Distribution Function $\mathcal{N}(x)$



6 Covariance and Correlation

Let X_1 and X_2 be two drv's with means μ_1, μ_2 and variances σ_1^2, σ_2^2. Then

Definition: The *covariance* of X_1 and X_2 is defined to be

$$\boxed{\text{cov}\{X_1, X_2\} = \sigma_{12} = \mathbb{E}\{(X_1 - \mu_1)(X_2 - \mu_2)\}} \tag{6.1}$$

The covariance of two random variables is a measure of the degree of variation of one variable with respect to the other. The covariance may be positive, negative or zero. It will be positive if an increase in one of the variables tends to be followed by an increase in the other. It will be negative if an increase in one is followed by a decrease in the other. If the covariance is zero we shall say the two random variables are *uncorrelated*. This is slightly different from saying two r.v.'s are *independent*.
Properties :

1. $\text{cov}\{X, X\} = \mathbb{V}\{X\} = \sigma^2$
2. $\text{cov}\{X_1, X_2\} = \mathbb{E}\{X_1 X_2\} - \mu_1\mu_2$
3. $\mathbb{E}\{X_1 X_2\} = \mathbb{E}\{X_1\} \cdot \mathbb{E}\{X_2\}$ if X_1, X_2 are uncorrelated
4. $\mathbb{V}\{a_1X_1 + a_2X_2\} = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + 2a_1a_2\sigma_{12}$
5. $\mathbb{V}\{X_1 + X_2\} = \mathbb{V}\{X_1\} + \mathbb{V}\{X_2\}$ if X_1, X_2 are uncorrelated

Proofs:

1. When $X_1 = X_2 = X$, we get

$$\text{cov}\{X, X\} = \mathbb{E}\{(X - \mu)^2\} = \mathbb{V}\{X\} = \sigma^2$$

2.

$$\begin{aligned} \text{cov}\{X_1, X_2\} &= \mathbb{E}\{(X_1 - \mu_1)(X_2 - \mu_2)\} \\ &= \mathbb{E}\{X_1X_2\} - \mu_1\mathbb{E}\{X_2\} - \mu_2\mathbb{E}\{X_1\} + \mu_1\mu_2 \\ &= \mathbb{E}\{X_1X_2\} - \mu_1\mu_2 \qquad \text{since } \mathbb{E}\{X_i\} = \mu_i \end{aligned}$$

3. An important corollary of the previous result is the following:
If X_1 and X_2 are uncorrelated random variables, then $\text{cov}\{X_1, X_2\} = 0$ and $\mathbb{E}\{X_1X_2\} = \mu_1\mu_2$.

4.

$$\begin{aligned} \mathbb{V}\{a_1X_1 + a_2X_2\} &= \mathbb{E}\{[a_1(X_1 - \mu_1) + a_2(X_2 - \mu_2)]^2\} \\ &= \left[\begin{array}{l} a_1^2\mathbb{E}\{(X_1 - \mu_1)^2\} + a_2^2\mathbb{E}\{(X_2 - \mu_2)^2\} \\ +2a_1a_2\mathbb{E}\{(X_1 - \mu_1)(X_2 - \mu_2)\} \end{array}\right. \\ &= a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + 2a_1a_2\sigma_{12} \end{aligned}$$

5. This last result implies that $\mathbb{V}\{aX\} = a^2\mathbb{V}\{X\}$ (quadratic scaling) and $\mathbb{V}\{X_1 + X_2\} = \mathbb{V}\{X_1\} + \mathbb{V}\{X_2\}$ only if X_1 and X_2 are uncorrelated.

Definition: Let X_1 and X_2 be two drv's with variances σ_1^2, σ_2^2 and covariance σ_{12}. Then we define their *correlation coefficient* to be

$$\boxed{\rho = \text{corr}\{X_1, X_2\} = \frac{\sigma_{12}}{\sigma_1 \sigma_2} = \frac{\text{cov}\{X_1, X_2\}}{\sqrt{\mathbb{V}\{X_1\}\mathbb{V}\{X_2\}}}} \tag{6.2}$$

Theorem: $-1 \leq \rho \leq 1$. The proof follows from an application of the Cauchy-Schwartz Inequality.

The correlation coefficient is a normalized covariance measure, whose properties are illustrated by the following five special cases:

1. $\rho = -1$: Perfect negative correlation; X_1 and X_2 are linearly dependent of the form $X_2 = aX_1 + b$ with $a < 0$. Thus as X_1 increases, X_2 decreases linearly.
2. $-1 < \rho < 0$: X_1 and X_2 are negatively correlated; as X_1 increases, X_2 will tend to decrease.
3. $\rho = 0$: X_1 and X_2 are uncorrelated; as X_1 increases, X_2 shows no tendency to either increase or decrease.
4. $0 < \rho < 1$: X_1 and X_2 are positively correlated; as X_1 increases, X_2 also tends to increase.
5. $\rho = +1$: Perfect positive correlation; X_1 and X_2 are linearly dependent of the form $X_2 = aX_1 + b$ with $a > 0$. Thus as X_1 increases, X_2 increases linearly.

7 Linear Regression

Consider observations x_i, y_i of two random variables X, Y. A linear fit of the two sets of observations is a relation of the form

$$y_i = \alpha + \beta x_i + \epsilon_i$$

where α, β are constants and ϵ_i is called the *residual* of the i^{th} observation.

The line $Y = \alpha + \beta X$ is called the line of 'best' fit or the line of regression of Y on X if α, β are chosen to minimize the sum of squares of the residuals, *i.e.* $\sum \epsilon_i^2$.

This is a standard minimization problem with solution:

$$\boxed{\alpha = \mu_Y - \beta \mu_X \qquad \beta = \sigma_{XY}/\sigma_X^2.} \tag{7.1}$$

8 Covariance and Correlation Matrices

Let X_i for $i = 1, 2, \ldots, n$ be $n-$discrete random variables with means μ_i, variances σ_i^2 and covariances σ_{ij}. Then the matrix S with $ij-$element $\sigma_{ij} = \text{cov}\{X_i, X_j\}$ is called the *covariance matrix*.

$$S = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \ldots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \ldots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \ldots & \sigma_n^2 \end{pmatrix}.$$

The corresponding *correlation matrix* with $ij-$element ρ_{ij} is

$$P = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1n} \\ \rho_{21} & 1 & \dots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \dots & 1 \end{pmatrix}.$$

Note that we have used $\text{corr}\{X, X\} = 1$ for any random variable X.

Properties:

1. S and P are symmetric positive semi-definite matrices with $S = DPD$ where $D^2 = \text{diag}(S)$.

2. If the X_i are mutually independent random variables, then S is a diagonal matrix with the variances σ_i^2 on the diagonal; and P is the order-n identity matrix.

3. Let Y be a drv which is some linear combination of the X_i. Then we may write Y in vector form as

$$Y = \sum_{i=1}^{n} a_i X_i = \underset{\sim}{a}\,' \underset{\sim}{X}$$

where $\underset{\sim}{a}, \underset{\sim}{X}$ are $(n \times 1)$-column vectors and the dash denotes the transpose. If Y_1 and Y_2 are two such linear combinations of the X_i with coefficients $\underset{\sim}{a}_1$ and $\underset{\sim}{a}_2$ then

$$\boxed{\text{cov}\{Y_1, Y_2\} = \underset{\sim}{a}_1\,' S \underset{\sim}{a}_2} \tag{8.1}$$

—oOo—

Chapter 6

Risky Securities and Utility Theory

A financial *security* is a document which records some entitlement or obligation, and which can readily be traded for money. Some common examples are Stocks and Options, Bonds, Forwards and Futures, Treasury Bills, Swaps. The return and the fair price or value of a *riskless* security is known with certainty.
But what if, as is likely in reality, the return is not known with certainty? What if the return is a *random variable*? In this case, we refer to the security as a *risky security*. The question now arises: how do we obtain the value or price of such a risky security? The answer to this question requires the use of concepts in probability theory reviewed briefly in the previous chapter.

Example: Consider a simple security which pays a single amount A. In the world of certainty, the security would be valued as $S_0 = A$. But suppose A is a random variable. In particular, assume A has the value A_1 if event E_1 occurs and the value A_2 if event E_2 occurs. In this case, A is a discrete binomial random variable. Define probabilities

$$\mathbb{P}\{E_1\} = p_1; \qquad \mathbb{P}\{E_2\} = p_2; \qquad p_1 + p_2 = 1.$$

Then a reasonable price of the risky security might be

$$S_0 = p_1 A_1 + p_2 A_2 \, .$$

That is we weight the possible payments by their probabilities of occurrence. For one thing, this approach is consistent with the limiting cases $p_2 = 0$ (A_1 is paid with certainty) and $p_1 = 0$ (A_2 is paid with certainty).

Valuing the security in this way turns out to be a consistent *model*. But as we shall see in this Chapter, it does not tell the whole story, because it ignores the risk preferences of different investors.

1 Principle of Expected Return

Suppose now that the return R of a risky security is regarded as a random variable with known distribution. A consistent model for pricing the security is then given by the Principle of Expected

Return (also called the *Expectations Hypothesis*). If $\mathbb{E}$ denotes the expectation operator, then according to the Expectations Hypothesis the "fair price" of the security is

$$S_0 = \mathbb{E}\{R\}\,. \tag{1.1}$$

This expectation is calculated using the probability distribution of the random variable R.

The term "fair price" gives the impression that this price should be acceptable to buyers and sellers, but this has not been demonstrated in the present context and in fact may not be true (as we shall see shortly).

2 The St. Petersburg Paradox

As the following example illustrates, there is a difficulty with the Principle of Expected Return as a pricing model in some types of problems. Consider the following question:

> What is the fair price of a coin tossing game which pays an amount $C_t = 2^{t-1} \times A$ if the first head appears on toss number t ?

Thus the pay-off of the game is A if the 1st-head occurs on the 1st-toss; $2A$ if it occurs on the 2nd toss; 44 if on the 3rd toss; $8A$ on the 4th toss etc.

This is an example of a single pay-off (cash flow) at an unknown time (toss) in the future, with zero interest rates. The pay-off in this case is however not constant, but depends on when it is paid. Let p_t be the probability that the first head occurs on toss t. Then the first t tosses must be $TTT\ldots TH$ which occurs with probability $p_t = 1/2^t$ (assuming an unbiased coin). The pay-off for this case is given as $C_t = 2^{t-1}A$. Hence, by the Principle of Expected Return with $V_t = 1$, the fair price of the game is given by

$$\begin{aligned} S_0 &= \mathbb{E}\{C_T\} = \sum_{t=1}^{\infty} p_t C_t = \sum_{t=1}^{\infty} 2^{t-1} \cdot \frac{A}{2^t} = \sum_{t=1}^{\infty} \tfrac{A}{2} \\ &= \tfrac{A}{2} + \tfrac{A}{2} + \tfrac{A}{2} + \ldots \\ &= \infty \end{aligned}$$

This seems absurd! In order to resolve this apparent paradox, we need to understand the concepts of the following sections.

3 Utility Theory

It is conjectured that people are more concerned with the *utility* of income they might receive than with the actual amount. Utility is a hypothetical quantity which is supposed to measure the satisfaction or pleasure an individual gets from receiving a given sum of money. Generally, an increment of wealth on a large fortune has less utility than an equal increment on a small fortune. This concept

is sometimes called the *law of diminishing return*.

For example, a \$1,000 increment on savings of only \$1,000 is likely to have a lot more utility than a \$1,000 increment on \$1-million. In the former, the individual's fortune has doubled; in the latter the relative increase is a meagre 0.1 of a percent. Unless our millionaire is exceedingly parsimonious, he/she will experience considerably less satisfaction than persons who double their fortunes.

Let $U(x)$ denote the utility associated with the receipt of an amount x by an individual. Then $U(x)$ is called the *utility function* for that individual. This is postulated to satisfy a simple but important requirement:

(1) $U'(x) > 0$ – more wealth is preferred to less.

This is sometimes called the *law of non-satiety*, and can be interpreted as saying "You can never have too much money". (This reasonableness of this assertion does not really depend on one's political or moral perspective; after all, somebody who aspires to poverty can always give away unwanted money!)

The utility function will usually also satisfy a second condition:

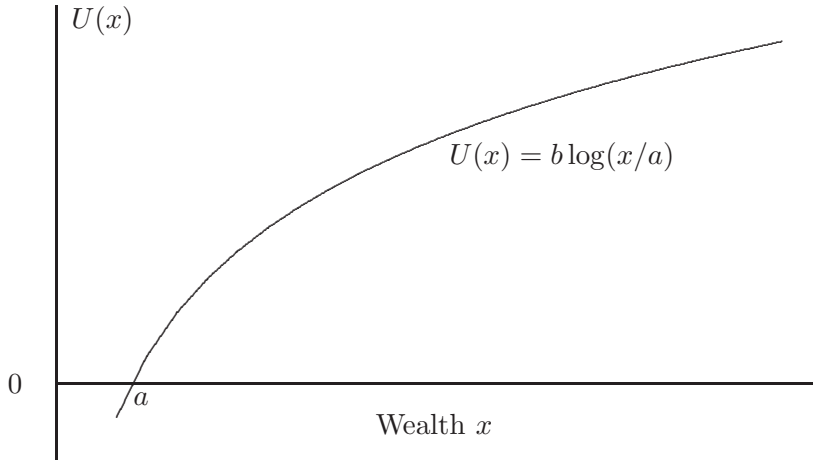
(2) $U''(x) < 0$ – marginal utility decreases as wealth increases.

This is the mathematical statement of the *law of diminishing returns*, and ensures that $U'(x)$ is a decreasing function. Thinking of $U'(x)$ as a measure of the satisfaction associated with a small increment in cash returns, this means that the prospect of an extra dollar or two is less exciting if you are getting a lot of money anyway.

As we will see later, the second assumption implies that an individual will generally try to avoid risky investments. While this holds for most large investors, it isn't always true for individuals risking small amounts; after all, many people enjoy the excitement associated with gambling. (For such people, the hope of winning presumably outweighs the fear of losing.) In general, however, when the stakes are high most people prefer to avoid risks. Consequently, when considering sizeable investments, condition (2) can generally be assumed.

When conditions (1) and (2) both hold, we see that $U(x)$ is a monotonic increasing, concave function of wealth x with the general appearance of the graph below.

Logarithmic Utility Function



Clearly, if condition (2) does not hold, then the graph may not be concave everywhere.

Because $U(x)$ is an increasing function, an infinitesimal increment in the amount x to be received must result in a positive increment dU in the utility. It has been argued that a good model for risk-averse investors is one for which the increment in utility dU is proportional to the increment dx and *inversely* proportional to the resulting wealth $x+W$, where W denotes the individual's current wealth. This ensures that an individual's satisfaction is determined by the *proportion* by which her total wealth increases, rather than the dollar value of the increase. For example, it ensures that a doubling of wealth will produce the same amount of satisfaction regardless of the current level of wealth.

Mathematically, this model implies that there exists a positive constant b such that $dU = b\,dx/(x+W)$. We then have a first order differential equation for $U(x)$ with solution

$$\boxed{U(x) = b\,\log\left(\tfrac{x+W}{W}\right) + a = b\,\log\left(1+\tfrac{x}{W}\right) + a} \tag{3.1}$$

with a a constant of integration. If we suppose that $U = 0$ when $x = 0$ (i.e. that no satisfaction comes from merely maintaining a steady level of wealth) then we must set $a = 0$.

This is the logarithmic utility function first proposed by Bernoulli. Each individual with logarithmic utility may have a different value of the parameters a and b. These respectively determine the zero level and scale of the utility function. In fact we will see that the values of a and b are immaterial (provided b is positive); regardless of what values are chosen, the Principle of Expected Utility will always lead to the same choice of investment.

Observe that the logarithmic utility function does satisfy the two conditions imposed above: $U'(x) = b/x > 0$ and $U''(x) = -b/x^2 < 0$.

So how do we use an individual's utility function to price a risky security ?
The answer is given next.

4 Principle of Expected Utility

The *Principle of Expected Utility* states simply that an individual will seek to *maximise expected utility*, rather than expected return.

Thus, if given the choice between an investment with a net return of X_1 and one with a net return of X_2, with X_1 and X_2 random variables, the investor will choose the investment for which $\mathbb{E}\{U(X_i)\}$ is greater.

We can use the Principle of Expected Utility to determine how much to pay for a given security. Suppose that the current purchase price of the security is S, and that after its purchase the security will produce cash-flows with a combined (present) value X. (We suppose that X is a random variable with known statistical properties.) Then the net return from the security is $X - S$ and so the expected utility resulting from its purchase is

$$\mathbb{E}\{U(X - S)\}. \tag{4.1}$$

In this investment worthwhile? In other words, is it a better investment than doing nothing at all? If the investor does nothing, the return will be zero and so the expected utility will be just

$$\mathbb{E}\{U(0)\} = U(0). \tag{4.2}$$

According to the Principle of Expected Utility, we must compare the expected utilities to decide which alternative is better.

In fact, we see that the purchase of the security will be worthwhile (better than doing nothing) if and only if

$$\mathbb{E}\{U(X - S)\} > U(0). \tag{4.3}$$

On the other hand, it will be better to do nothing if and only if

$$\mathbb{E}\{U(X - S)\} < U(0). \tag{4.4}$$

If neither of these equalities holds, i.e. if $\mathbb{E}\{U(X - S)\} = U(0)$, then neither alternative is preferred to the other; i.e. it will be a matter of indifference whether or not the security is purchased at the price S.

Therefore, the root S_0 of the equation

$$\boxed{\mathbb{E}\{U(X - S_0)\} = U(0)} \tag{4.5}$$

must represent the price at which the purchase of the security is neither advantageous or disadvantageous; and also the maximum price at which the purchase of the security might be considered. (At any lower price the security would be considered a bargain.)

For example; suppose that the security will be worth either $M = \$100$ or nothing, with equal probability. The maximum price that would be paid for the security by an investor with utility function $U(x)$ is then the root S_0 of the equation

$$\frac{1}{2}U(M - S_0) + \frac{1}{2}U(-S_0) = U(0). \tag{4.6}$$

If the investor uses the logarithmic utility function $U(x) = a + b\log(1 + (x/W))$, where W denotes his total wealth, then this reduces to

$$\frac{1}{2}\left[a + b\log\left(1 + \frac{M - S_0}{W}\right)\right] + \frac{1}{2}\left[a + b\log\left(1 + \frac{-S_0}{W}\right)\right] = a \tag{4.7}$$

or (since we assume $b > 0$)

$$\log\left(1 + \frac{M - S_0}{W}\right) + \log\left(1 - \frac{S_0}{W}\right) = 0 \tag{4.8}$$

and hence

$$\left(1 + \frac{M - S_0}{W}\right)\left(1 - \frac{S_0}{W}\right) = 1 \tag{4.9}$$

which has roots

$$S_0 = W + \frac{M}{2} \pm \sqrt{W^2 + \left(\frac{M}{2}\right)^2}. \tag{4.10}$$

We discard the larger root, as this gives a price which exceeds the amount W that the investor can afford to pay. If the investor's current wealth W is much larger than M, so that $M/W << 1$, then we can use the binomial series expansion for $\sqrt{1 + (M/2W)^2}$ to approximate the remaining root as

$$S_0 = W + \frac{M}{2} - W\sqrt{1 + \left(\frac{M}{2W}\right)^2} \simeq \frac{M}{2} - \frac{M^2}{8W} + \frac{M^4}{128W^3}\ldots \tag{4.11}$$

According to the Principle of Expected Utility, this is the maximum price the investor would consider paying. It is slightly less than the fair price $M/2$ derived from the Principle of Expected Return, which does not take into account the investor's aversion to the risk. (For example, with $W = \$10,000$ the Principle of Expected Utility gives $S_0 = \$49.875$, while the Principle of Expected Return gives \$50.) The discrepancy between the two prices is called the *risk premium*, since it is a monetary measure of the extent to which the investor's valuation of the security is diminished due to the uncertainty in the return. In this example (as in most cases), the risk premium is a decreasing function of the investor's wealth W.

Although the expression obtained for S_0 depends on the investor's wealth W, it doesn't depend at all on the values of the constants a and b appearing in the definition of the utility function $U(x)$. As remarked above, the values assigned to these constants are immaterial.

Theorem: *Investment decisions based on the Principle of Expected Utility are unchanged by positive linear transformation of the utility function.*

Proof: First, let us be clear what we mean by a positive linear transformation. We say that $\bar{U}(x) = a + bU(x)$ is a positive linear transformation of the utility function $U(x)$ if a and b are constants with $b > 0$.

We need to show that an investor guided by the Principle of Expected Utility will make the same decisions regardless of whether she uses the original utility function $U(x)$ or the transformed utility function $\bar{U}(x)$.

Suppose that the random variables X_1 and X_2 represent the net returns from two alternative investments, and that

$$\mathbb{E}\{U(X_1)\} > \mathbb{E}\{U(X_2)\}. \tag{4.12}$$

Then, according to the Principle of Expected Utility, an investor with utility function $U(x)$ will choose investment 1 rather than investment 2.

It is easy to verify that the decison will be unchanged if the investor uses the new utility function $\bar{U}(x) \equiv a + bU(x)$ instead of $U(x)$. Indeed, we see that

$$\begin{aligned}
\mathbb{E}\{\bar{U}(X_1)\} - \mathbb{E}\{\bar{U}(X_2)\} &= \mathbb{E}\{a + bU(X_1)\} - \mathbb{E}\{a + bU(X_2)\} \\
&= a + b\mathbb{E}\{U(X_1)\} - a - b\mathbb{E}\{U(X_2)\} \\
&= b[\mathbb{E}\{U(X_1)\} - \mathbb{E}\{U(X_2)\}] \\
&> 0
\end{aligned}$$

since we assumed that $b > 0$ and $\mathbb{E}\{U(X_1)\} > \mathbb{E}\{U(X_2)\}$. Consequently,

$$\mathbb{E}\{\bar{U}(X_1)\} > \mathbb{E}\{\bar{U}(X_2)\} \tag{4.13}$$

and so an investment decision based on the utility function $\bar{U}(x)$ will also favour the first investment.Thus, it doesn't matter whether the investor takes $U(x)$ or $\bar{U}(x) = a + bU(x)$ as her utility function; she will make exactly the same decision.

—oOo—

Utility functions $U(x)$ and $\bar{U}(x)$ which are related by positive linear transformations are said to be *equivalent*, since they lead to identical investment decisions.

Corollary: Prices calculated using the Principle of Expected Utility are unaffected by positive linear transformation of the utility function.

—oOo—

As a special case of this result, observe that the logarithmic utility function $U(x) = a+b\log(1+x/W)$ (with $b > 0$) is a positive linear transformation of $V(x) = \log(1 + x/W)$. Hence, as far as pricing risky securities with the Principle of Expected Utility is concerned, $U(x) = a + b\log(1 + x/W)$ and $V(x) = \log(1 + x/W)$ are equivalent utility functions. This explains why the maximum price S_0 derived in the preceding example was found to be independent of a and b.

We conclude this section by noting that the method outlined above can be used to provide a solution to the St Petersburg Paradox. To determine the maximum price S_0 an investor should consider paying to play the game, we need to solve the equation

$$U(0) = \mathbb{E}\{U(C_T - S_0)\} = \sum_{t=1}^{\infty} p_t U(C_T - S_0). \tag{4.14}$$

If we use the logarithmic utility function $U(x) = \log(1 + x/W)$, this reduces to the equation

$$0 = \sum_{t=1}^{\infty} \frac{1}{2^t} \log\left(1 + \frac{2^{t-1}A - S_0}{W}\right) \tag{4.15}$$

which admits a positive but finite solution S_0. (Unfortunately, the proof of this statement is beyond the scope of this course.) Thus, as one might expect, there is an upper limit to the price the individual might pay for the right to play the game. This conclusion seems much more reasonable than that obtained earlier using the Principle of Expected Return.

5 The Certainty Equivalent

The valuation of a risky security using the Principle of Expected Utility is *subjective*, in the sense that it depends on the investor's attitudes to risk (as well as her assessment of the probabilities). In fact it also depends on the nature of the transaction; for example, a security received as a gift may be valued slightly differently from one which is purchased.

To see this, consider suppose an investor is offered a choice between the following gifts:

- a fixed amount of cash C
- a risky security with net return X (a random variable).

If the investor takes the cash, the (expected) utility will be simply $U(C)$; if she chooses the security, the expected utility will be $\mathbb{E}\{U(X)\}$.

According to the Principle of Expected Utility, she will choose the security if $\mathbb{E}\{U(X)\} > U(C)$, and the cash if $\mathbb{E}\{U(X)\} < U(C)$. However, she will have no clear preference if

$$\mathbb{E}\{U(X)\} = U(C). \tag{5.1}$$

When this equality is satisfied, the cash and the security will be equally attractive to the investor; so this particular value of C may be thought of as the cash value of the security when received by

the investor as a gift. The value of C which solves this equation is referred to as the "certainty equivalent" value of the security to the investor.

It is important to note that this may differ slightly from the previously introduced quantity S_0, which represented the maximum price at which the investor would consider purchasing the security. (In general C is slightly higher than S_0 because the investor will be less concerned about the risks associated with a gift than with a purchase. For example, in the case of the security discussed earlier which returns either M or nothing with equal probability, and an investor with utility function $U(x) = \log(1 + x/W)$, the certainty equivalent value of the security is found to be $C = W\sqrt{1 + (M/W)} - W = M[\frac{1}{2} - (M/8W) + (M^2/16W^2)\ldots]$ which is about $7M^2/128W^3$ greater than the maximum price S_0 at which the investor would consider purchasing the security.)

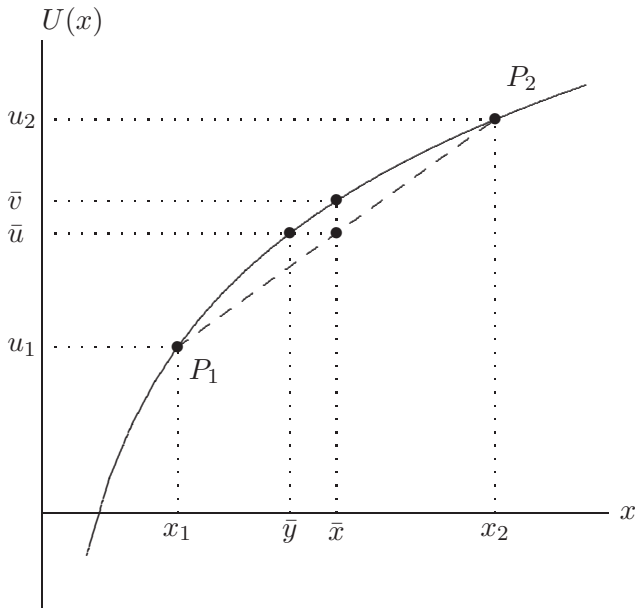
The certainty equivalent C is generally a very good approximation to the maximum price S_0 the investor might consider paying. As it is usually easier to calculate than S_0, it is much more widely used in practice.

6 Utility and Risk Attitudes

Let X be a discrete binomial random variable which takes the values x_1 and x_2 with probabilities p_1 and $p_2 = 1 - p_1$ respectively. For the utility function $U(x)$ let $u_i = U(x_i)$ and define

$$\begin{aligned} \bar{x} &= \mathbb{E}\{X\} = p_1 x_1 + p_2 x_2 \\ \bar{u} &= \mathbb{E}\{U(X)\} = p_1 u_1 + p_2 u_2 = U(\bar{y}) \\ \bar{v} &= U(\bar{x}) \end{aligned}$$

Assume that $U(x)$ is concave, monotonic increasing. Then we have the situation appearing in the graph below:



Since the utility function $U(x)$ is concave, chord P_1P_2 will always lie *below* $U(x)$. Hence $\bar{v} > \bar{u}$ or equivalently

$$U(\mathbb{E}\{X\}) > \mathbb{E}\{U(X)\}.$$

This inequality is generally true for any random variable X and any concave function $U(x)$ and is known as Jensen's Inequality.

In the present context, the inequality has an important economic interpretation. It implies that an investor with utility function $U(x)$ will prefer the certainty of receiving an amount of cash $\bar{x} = \mathbb{E}\{X\}$ to a risky security with return X. This preference is typical of *risk-averse* investors. According to the Principle of Expected Utility, a risk-averse investor will value the risky security as $\bar{y}$.

The difference between $\bar{x}$ and $\bar{y}$ is called the *risk premium*:

$$\boxed{\rho = \bar{x} - \bar{y} = \mathbb{E}\{X\} - U^{-1}\mathbb{E}\{U(X)\}} \tag{6.1}$$

It represents the amount an investor with utility function $U(x)$ must be compensated to accept the risky investment instead of a riskless investment with the same expected net return. It is also the difference between the expected return price and certainty equivalent of the security. For a risk-averse investor, the risk premium ρ is always positive.

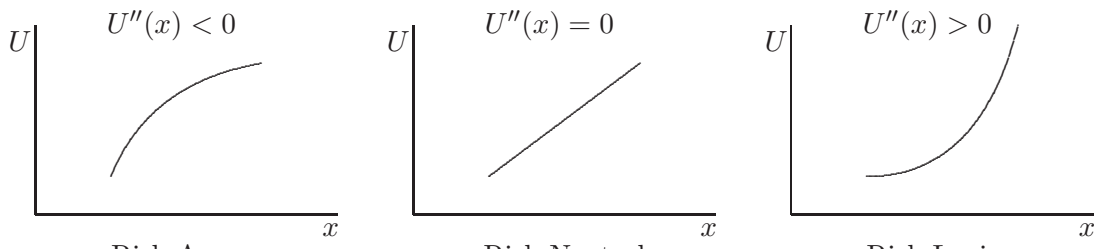
Note that since $U(x)$ is monotonic it has a well-defined inverse function U^{-1}, such that if $U^{-1}(u) = x$ then $U(x) = u$.

The above discussion is readily extended to other types of risk attitudes, that are summarized in the next Table.

Classification of Risk Attitudes

Type	$U''(x)$	ρ	*Risk Attitudes*
Risk-averse	< 0 (concave)	$\rho > 0$	Prefers certainty to risk for the same expected return
Risk-neutral	$= 0$ (linear)	$\rho = 0$	Indifferent between certainty and risk for the same expected return
Risk-loving	> 0 (convex)	$\rho < 0$	Prefers risk to certainty for the same expected return

If these three investors value a particular security respectively as S_-, S_0 and S_+, then $S_- < S_0 < S_+$.



Note:
A risk-neutral investor has linear utility function $U(x) = a + bx$ (equivalent to $U(x) = x$ by the theorem on positive linear transformations of utility functions) and the Principle of Expected Utility reduces to the Principle of Expected Return.

Example: Consider two securities: A has PV(cash-flows) $X = 100$ with probability one (i.e. non-random); B has PV(cash-flows) $Y = 50$ or $Y = 150$ with equal probability $\frac{1}{2}$ (i.e. B is stochastic). Let I_1, I_2, I_3 be three investors with the following utility functions:

$$U_1(x) = 2 + 3\log(x); \quad U_2(x) = -4 + 12x; \quad U_3(x) = 7 + 6x^2.$$

Determine their risk attitudes and obtain each investor's preference between A and B; and each investor's risk premium for the risky security.

Solution:
The theorem on positive linear transformations allows us (as far as pricing and ranking investments by expected utility is concerned) to replace U_i by the equivalent utility functions:

$$V_1(x) = \log(x); \quad V_2(x) = x; \quad V_3(x) = x^2.$$

This will make the analysis a little simpler.
The inverse functions are then:

$$V^{-1}(x) = e^x; \quad V_2^{-1}(x) = x; \quad V_3^{-1}(x) = \sqrt{x}.$$

Note that A and B have the same expected return of 100 units.

The following table is easily computed:

	$A : V(X)$	$B : \mathbb{E}\{V(Y)\}$	$\rho(Y) = \mathbb{E}(Y) - V^{-1}\mathbb{E}\{V(Y)\}$	
I_1	4.61	4.46	$100 - e^{\frac{1}{2}(\log 50 + \log 150)}$	$= 13.4$
I_2	100	100	$100 - 100$	$= 0$
I_3	10,000	12,500	$100 - \sqrt{\frac{1}{2}(50^2 + 150^2)}$	$= -11.8$

We therefore conclude that:
$V_1''(x) = -1/x^2 < 0$ implying risk averse attitude; I_1 prefers A to B (i.e. prefers certainty to risk for the same return); risk-premium $\rho_1 = +13.4 > 0$.

$V_2''(x) = 0$ implying risk neutral attitude; I_2 is indifferent between A and B; risk premium $\rho_2 = 0$.

$V_3''(x) = 2x > 0$ implying risk loving attitude; I_3 prefers B to A (i.e. prefers risk to certainty for the same return); risk premium $\rho_3 = -11.8 < 0$.

—oOo—

7 Measures of Risk Aversity

There is no unique measure of risk aversity, but as we have already seen the risk premium and utility function concavity can be used as such a measure. We shall define three measures of risk aversity in this course: they are the *risk premium*, the *absolute risk aversion* and the *relative risk aversion* defined below. The last two depend on the concavity of the utility function.

1. *Risk premium*
   The risk premium measures risk aversity in monetary units – it is the amount that a risk averter must be compensated to take on a risky investment and is equal to the excess of the expected utility price above the risk neutral price of the investment.

   $$\rho = \mathbb{E}(X) - U^{-1}\mathbb{E}\{U(X)\}.$$

   Note that this measure depends not only on the investor's risk attitude (through $U(x)$, but also on the actual investment.

2. *Absolute risk aversion*

   $$\rho_{\text{abs}} = -\frac{U''(x)}{U'(x)}.$$

   This measure depends only on the investor's risk attitude through the utility function $U(x)$.

3. *Relative risk aversion*

   $$\rho_{\text{rel}} = -\frac{x\,U''(x)}{U'(x)}.$$

   This also depends only the investor's risk attitude through $U(x)$.

In all three cases we have the classification:

$$\begin{array}{lcll} \rho & > & 0; & \text{risk aversion} \\ \rho & = & 0; & \text{risk neutrality} \\ \rho & < & 0; & \text{risk loving} \end{array}$$

Chapter 7

Portfolio Theory

A financial *portfolio* is an investment consisting of any collection of securities or assets. In practice the assets will be risky and are therefore modelled by random variables. Portfolios in financial institutions are administered by *fund managers* and their portfolios may contain several hundreds of securities covering many different economic sectors. Examples of such securities include stocks, equity options, futures, foreign currencies, bonds and bond options etc.

Mean-Variance Portfolio Theory (MVPT) stems from the work of Harry Markowitz who received the 1990 Nobel Prize in Economics for his contribution to the field. We shall adopt the Markowitz mean-variance approach exclusively in this Chapter.

The following assumptions underlie MVPT:

1. The market provides perfect information and perfect competition
2. Asset returns refer to a single fixed time period
3. Asset returns are normally distributed with known means and covariance matrix
4. Investors act rationally and are risk averse
5. Investors base their decisions solely on the means and covariances of asset returns

Note: By *asset return* we shall mean the one-period holding rate defined by: $R = (X_1 - X_0)/X_0$ where X_0, X_1 are the asset prices at the start and end of the period. X_0 (today's price) is known, but X_1 (future price) is unknown and therefore a random variable. This makes R a random variable.

Portfolio theory is really all about managing risk. At present there is no universally accepted definition of *risk*. However, a simple proxy for the risk of an asset is the standard deviation of its return. This fits in well with our intuition that an asset return with large variance is very risky, while one with zero variance is riskless. We use this proxy for risk throughout this chapter. One very important feature of a portfolio is that a judicious choice of assets within the portfolio can lead to reduced risk – even lower risk than that of the individual assets within the portfolio. The

degree of risk reduction depends critically on the asset covariances.

Example 1.1
Consider two assets with random returns $R_1,\ R_2$; expected returns r_1 and r_2 and standard deviation of returns σ_1 and σ_2. That is $r_i = \mathbb{E}\{R_i\}$ and $\sigma_i^2 = \mathbb{V}\{R_i\}$. Suppose the asset returns have perfect negative correlation, so that the correlation coefficient (see Chapter 5) is $\rho = -1$. Let us construct a simple portfolio of one share of asset-1 and h shares of asset-2. The parameter h is sometimes called the *hedge ratio*. The random return on this two asset portfolio is $R = R_1 + hR_2$ with expected return $\mu = r_1 + hr_2$. The variance of return on the portfolio will be (see Chapter 5)

$$\begin{aligned} \sigma^2 &= \sigma_1^2 + h^2\sigma_2^2 + 2h\rho\sigma_1\sigma_2 \\ &= \sigma_1^2 + h^2\sigma_2^2 - 2h\sigma_1\sigma_2 \qquad \text{when } \rho = -1 \\ &= (\sigma_1 - h\sigma_2)^2 \end{aligned}$$

With the choice $h = \sigma_1/\sigma_2$ we obtain $\sigma = 0$ with the extra-ordinary conclusion that this portfolio of two risky assets is itself riskless !

The above example is a dramatic illustration of risk reduction. In practice no two assets are perfectly correlated so it will not be possible to reduce the portfolio risk to zero, but depending on the level of correlation, overall risk reduction will generally still occur. And of course this is why portfolio theory is so important in practice.

1 Portfolio Basics

Let $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n$ be $n-$risky securities with returns described by $n-$random variables $R_1, R_2, \ldots, R_n$. It is assumed that the expected returns and their covariances are known. Thus for $i, j = 1, 2, , \ldots, n$ let

$$\begin{array}{rcll} r_i &=& \mathbb{E}[R_i] & \text{Expected return of } \mathcal{S}_i \\ s_{ij} &=& \text{cov}[R_i, R_j] & \text{Covariance of returns of } \mathcal{S}_i,\ \mathcal{S}_j\ (i \neq j) \\ \sigma_i^2 &=& \mathbb{V}[R_i] & \text{Variance of return of } \mathcal{S}_i;\ (= s_{ii}) \end{array}$$

Recall from Chapter 5 that

$$\text{cov}[R_i, R_j] = \mathbb{E}\{(R_i - r_i)(R_j - r_j)\} = \mathbb{E}\{R_i R_j\} - r_i r_j$$

The covariance matrix also contains the correlation matrix with components ρ_{ij} through the relation $s_{ij} = \rho_{ij}\sigma_i\sigma_j$. Hence if we define the diagonal matrix D by the relation $D^2 = \text{diag}(S)$ where S is the covariance matrix, then the correlation matrix $P = \rho_{ij}$ satisfies $S = DPD$ or equivalently $P = D^{-1}SD^{-1}$.

A *portfolio* $\mathcal{P}$ is an investment containing a selection of the securities $\mathcal{S}_i$ in some proportion. Let W_0 denote the total wealth available for investment at the start of the period and let x_i denote the proportion of W_0 invested in $\mathcal{S}_i$. It is further assumed that all capital W_0 must be invested in $\mathcal{P}$. This *budget constraint* defines the decision domain $\mathcal{D}$ for portfolio selection as a hyperplane in an $n-$dimensional vector space.

$$\boxed{\mathcal{D} = \{x_i \mid \sum_{i=1}^{n} x_i = 1\}} \tag{1.1}$$

We shall meet other constraints on the x_i in this Chapter, the most important one being the *no short selling constraint*:

$$\boxed{x_i \geq 0 \quad \text{if no short-selling of asset } \mathcal{S}_i} \tag{1.2}$$

If short selling of a security is allowed, then the corresponding x_i may be negative and may also have absolute value exceeding unity (i.e. $|x_i| > 1$), provided of course that the budget condition (1.1) is not violated. In effect short selling raises extra capital to invest in other assets in the portfolio.

The return on the whole portfolio $\mathcal{P}$ is the random variable [1]

$$\boxed{R = \sum_{i=1}^{n} x_i \, R_i} \tag{1.3}$$

The mean μ and and variance σ^2 of the portfolio return are then given by:

$$\boxed{\begin{array}{rcl} \mu & = & \mathbb{E}[R] = \sum_{i=1}^{n} x_i r_i = \underset{\sim}{x}' \underset{\sim}{r} \\ \sigma^2 & = & \mathbb{V}[R] = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j s_{ij} = \underset{\sim}{x}' S \underset{\sim}{x} \end{array}} \tag{1.4}$$

These equations are often simply called the *Portfolio Equations*. The final expressions above are their matrix representations which we shall find very useful. S is the covariance matrix of returns with off-diagonal components $s_{ij} = \sigma_i \sigma_j \rho_{ij}$ and diagonal components $s_{ii} = \sigma_i^2$ where $\rho_{ij} = \text{corr}[R_i, R_j]$ is the correlation coefficient of the i and $j-$returns.

It is useful to think of Equation (1.4) as a transformation or mapping from the $n-$dimensional allocation $\underset{\sim}{x}-$space, to the 2-dimensional portfolio $\mu\sigma-$plane.

Proof: The portfolio equations are derived as follows:

$$\begin{array}{rcll} \mu & = & \mathbb{E}\{R\} = \mathbb{E}\{\sum_i x_i R_i\} & \\ & = & \sum_i x_i \mathbb{E}\{R_i\} & \text{(by linearity of } \mathbb{E}) \\ & = & \sum_i x_i r_i & \square \end{array}$$

and

$$\sigma^2 \;=\; \mathbb{V}\{R\} = \mathbb{E}\{(R-\mu)^2\}$$

[1] This result is non-trivial and is the subject of one of the Tutorial problems.

$$
\begin{aligned}
&= \mathbb{E}\{(\sum_i x_i R_i - \mu)^2\} \\
&= \mathbb{E}\{\sum_i x_i (R_i - r_i) \cdot \sum_j x_j (R_j - r_j)\} \\
&= \sum_i \sum_j x_i x_j \mathbb{E}\{(R_i - r_i)(R_j - r_j)\} \\
&= \sum_i \sum_j x_i x_j s_{ij} \qquad \square
\end{aligned}
$$

2 The Feasible Set

Definition: *The feasible set $\mathcal{F}$ is the set of $x_i \in \mathcal{D}$ which also satisfy all other imposed constraints.*

If there are no such constraints (other than the budget condition) then we shall term the portfolio *unrestricted*. The class of restricted portfolios we shall consider in these notes will mostly be single asset constraints of the form:

$$\boxed{L_i \leq x_i \leq U_i.} \tag{2.1}$$

Thus L_i and U_i denote lower and upper bounds on the proportion x_i invested in asset $\mathcal{S}_i$.

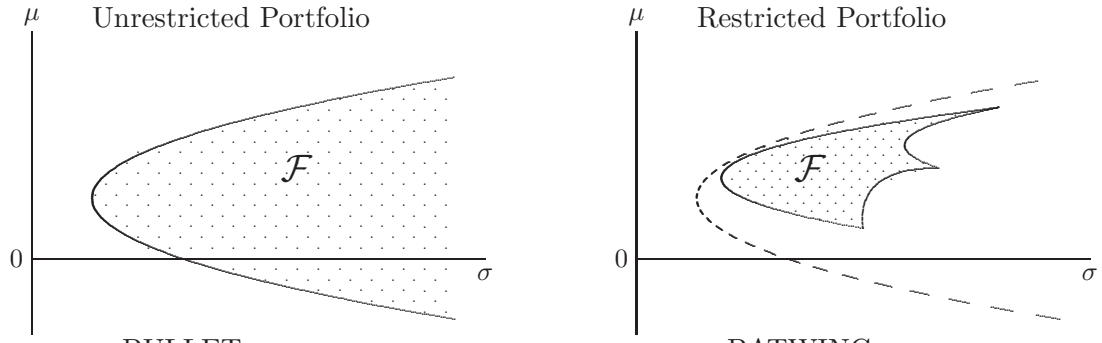
For example, if we forbid short selling of the single asset $\mathcal{S}_i$, then $L_i = 0$ and $U_i = \infty$. If we forbid short selling of all assets, then for all i we have $L_i = 0$ and $U_i = 1$.

Sector Constraints
Another popular and practical constraint is the *sector constraint*. Suppose assets can be classified into different economic sectors such as: real estate, banking, mining, industrial, communications, IT, media etc. Let $\underset{\sim}{b}$ a vector of maximum asset allocations within each sector. Then a sector constraint has the general linear form

$$\boxed{A\underset{\sim}{x} \leq \underset{\sim}{b}; \qquad A_{ij} = \begin{cases} 1 & \text{if asset-}j\text{ belongs to sector-}i \\ 0 & \text{otherwise} \end{cases}} \tag{2.2}$$

In this context the vector $\underset{\sim}{b}$ can be thought of as a resource vector and A is the constraint matrix (whose entries are either zero or one).

The feasible set $\mathcal{F}$ for a given portfolio problem, when mapped to the $\mu\sigma-$plane may turn out to be empty, a single point, a curve, a closed region or an open region.



We shall later derive the boundary of the feasible set for unrestricted portfolios. It turns out to be an hyperbola, whose interior $\mathcal{F}$, is somewhat colourfully called the *bullet*.

The shape of the feasible set for restricted portfolios is considerably more complicated, but will be a subset of the 'bullet'. It too has a descriptive name: the *batwing* (or *umbrella*).

3 Portfolio Selection Rules

In the rest of this Chapter we shall assume that the portfolio variance σ^2 measures its level of risk.

Consider two feasible portfolios (μ_1, σ_1) and (μ_2, σ_2). The following two portfolio selection rules apply to all rational risk-averse investors.

1. If $\mu_1 = \mu_2$ then select the portfolio with smaller σ (i.e. with lower risk).
2. If $\sigma_1 = \sigma_2$ then select the portfolio with larger μ (i.e. with higher return)

These two rules can be combined into a single rule called the:

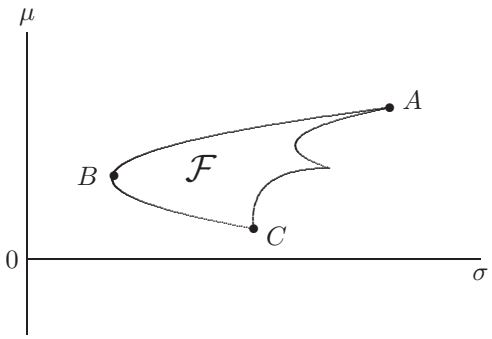
North-West Rule: *Given two feasible portfolios P_1 and P_2 in the $\mu\sigma-$plane, select the one lying in any region NW of the other.*

Clearly portfolios P which are NW of portfolios Q will have both higher return and lower risk. Any rational risk-averse investor would prefer P to Q in such circumstances.

If neither portfolio lies NW of the other, then the NW-rule makes no preferential selection.

4 The Efficient Frontier

Let A, B and C denote points (shown in the next figure) on the boundary of the feasible set $\mathcal{F}$.



We make the following observations:

1. Point A is the maximum return, maximum risk portfolio; point B is the minimum risk portfolio and point C is the minimum return portfolio.

2. Points on the curve ABC represent minimum risk portfolios for any given feasible return. That is a horizontal line corresponding to a given return μ, has smallest σ (risk) on this curve. The curve ABC is therefore called the *minimum variance frontier* or MVF for short.

3. Each point on the upper branch AB of the MVF has the special property that no other point of the feasible set lies NW of it. That is, points on AB are optimal portfolios in the sense that no other can have a higher return *and* a lower risk. The curve AB, the upper branch of the MVF, is called the *efficient frontier* or EF for short.

Proposition: *All rational risk-averse investors will select portfolios on the efficient frontier.*

Exactly which portfolio on the EF is chosen depends on the investor's risk-return preferences. Mr Cowboy will select portfolio A and obtain the highest possible return regardless of the risk. On the other hand, Mr Chicken will select portfolio B to ensure the least possible risk, regardless of the portfolio return. All other risk-averse investors will of course lie somewhere between these two extremes. Only a Mr Stupid would select a portfolio on the segment BC (the *inefficient frontier*).

Markowitz Criterion: *All rational risk-averse investors have risk-return preferences derived from a utility function $U = U(\mu, \sigma)$, which depends only on the portfolio return μ and portfolio variance σ^2.*

This criterion (which is actually a model of investor behaviour) is also called the *mean-variance criterion*. Its consequences are explored in the next section.

5 Indifference Curves

It is clear that an investor can obtain a high portfolio return μ only by increasing the portfolio variance (risk) σ^2. According to the Markowitz criterion, the degree of risk required for a given level of return will depend on the investor's utility function $U(\mu, \sigma)$.

We shall adopt a utility function of the form

$$U(\mu, \sigma) = F(t\mu - \tfrac{1}{2}\sigma^2)$$

where t is a positive parameter and $F(y)$ is a concave monotonic increasing function of its argument y. However, it is important to realise that this is only one choice of any number of models.

Let the negative of U's argument,

$$\boxed{Z(\mu, \sigma) = -t\mu + \tfrac{1}{2}\sigma^2} \tag{5.1}$$

denote the investor's *objective function*. Z can also be expressed as a quadratic function of the asset allocations x_i:

$$\boxed{Z(\underset{\sim}{x}) = -t(\underset{\sim}{r}'\underset{\sim}{x}) + \tfrac{1}{2}\underset{\sim}{x}'S\underset{\sim}{x}} \tag{5.2}$$

Observe that the utility function U is *maximised* with respect to x_i when Z is *minimised* with respect to x_i. Hence the basic portfolio problem is to find the x_i which minimises the investor's objective function $Z(\underset{\sim}{x})$.

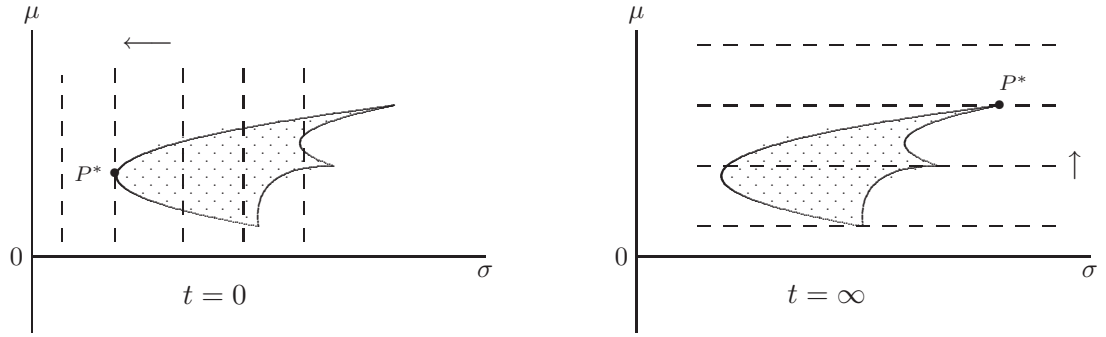
The parameter t for $t \geq 0$ is a *risk aversion parameter*, whose properties are discussed next.

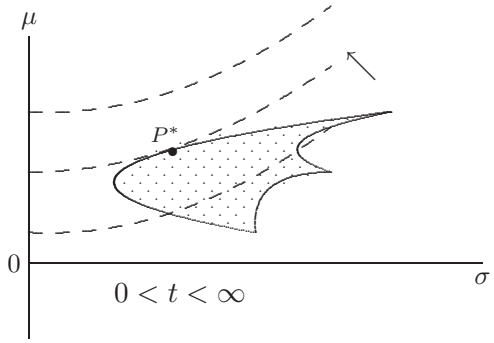
Case $t = 0$ The problem reduces to: min $Z = \frac{1}{2}\sigma^2$. This investor is interested only in minimizing risk without regard to the portfolio return (our Mr Chicken).

Case $t \to \infty$ The problem reduces to: max $Z = \mu$. This investor wishes to maximise portfolio return regardless of the risk (our Mr Cowboy).

Other investors will have a risk aversion parameter somewhere in between: *i.e.* $0 < t < \infty$.

Curves of constant Z in the $\mu\sigma-$plane are called *indifference curves*, because an investor has no preference for any portfolio lying on a given indifference curve. The curves $Z(\mu, \sigma) = Z_0$ are a family of parallel convex parabolas. This family has two degenerate cases: when $t = 0$ the indifferent curves are vertical lines in the $\mu\sigma-$plane; when $t \to \infty$ the indifferent curves become horizontal lines in the $\mu\sigma-$planes.





The optimal portfolio P^* is indicated in each diagram above (the arrow points in the direction of *decreasing* objective function value Z_0). As expected P^* is the minimum risk portfolio when $t = 0$ and is the maximum return portfolio when $t = \infty$. In the intermediate case, when $0 < t < \infty$, the optimum portfolio P^* occurs on the efficient frontier where it is tangential to the indifference curve. Since the efficient frontier is concave and the indifference curve is convex, there will always be a unique optimum portfolio $P^* = (\mu^*, \sigma^*)$.

Example 1.2 (a) An investor I with risk aversion parameter $t = 1.2$ is offered a choice of two portfolios:

$$\begin{aligned} P_1: \quad & \mu_1 = 10\%; \quad & \sigma_1 = 25\% \\ P_2: \quad & \mu_2 = 12\%; \quad & \sigma_2 = 30\% \end{aligned}$$

Which portfolio will I choose?
(b) Which investor is indifferent bewteen the two portfolios above?

Solution: (a) Compute the objective functions:

$$\begin{aligned} Z_1 &= -t\mu_1 + \tfrac{1}{2}\sigma_1^2 = -0.12 + 0.03125 = -0.08875 \\ Z_2 &= -t\mu_2 + \tfrac{1}{2}\sigma_2^2 = -0.144 + 0.045 = -0.099 \end{aligned}$$

Since $Z_2 < Z_1$ investor I will prefer P_2 to P_1.

(b) Invetsor $J(t)$ will be indifferent to the two portfolios if $Z_1(t) = Z_2(t)$. This leads to an equation for the risk aversion parameter t:

$$-0.10t + 0.03125 = -0.12t + 0.045$$

with solution $t = 0.6875$. □

Example 1.3 Suppose the EF is given by the straight line $\mu = a + b\sigma$ in the $\mu\sigma-$plane. Find the optimal portfolio $P^*(\mu^*, \sigma^*)$ for an investor with risk aversion parameter t.

Solution: The optimal portfolio requires that the indifference curve $-t\mu + \frac{1}{2}\sigma^2 = \text{const}$ is tangential to the EF at P^*. Differentiating we get $-t\mu' + \sigma = 0$ giving

$$\mu' = \frac{d\mu}{d\sigma} = \sigma/t = b \qquad \text{at } P^*$$

Hence $\mu^* = a + b^2 t$ and $\sigma^* = bt$ yields the optimal portfolio. □

6 The Two-Asset Portfolio

In the case $n = 2$, we can set $x_1 = x$ and $x_2 = 1 - x$ and the budget constraint $x_1 + x_2 = 1$ is automatically satisfied. The portfolio return and variance are then functions of the single variable x and are given by

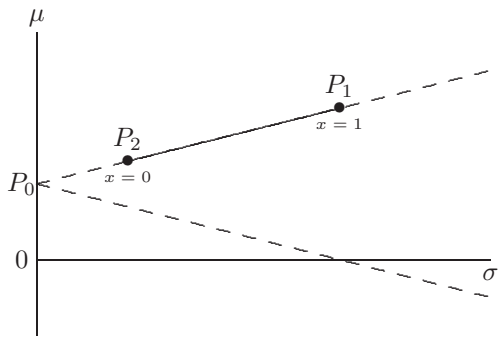
$$\boxed{\begin{array}{lcl} \mu & = & r_1 x + r_2(1-x) \\ \sigma^2 & = & \sigma_1^2 x^2 + \sigma_2^2 (1-x)^2 + 2\rho\sigma_1\sigma_2 x(1-x) \end{array}} \tag{6.1}$$

where $\rho = \text{corr}[R_1, R_2]$. These are the parametric equations of an hyperbola in the $\mu\sigma-$plane. If there are no other constraints, this hyperbola is both the feasible set $\mathcal{F}$ and the minimum variance frontier (MVF) [2]. Before finding its explicit equation, we consider in detail, three degenerate cases.

1. The case $\rho = 1$
When the asset returns R_1 and R_2 are perfectly correlated ($\rho = +1$), the portfolio equations above reduce to the simpler pair:

$$\begin{array}{lcl} \mu & = & r_1 x + r_2(1-x) \\ \sigma & = & |\sigma_1 x + \sigma_2(1-x)| \end{array}$$

These represent a pair of straight lines as sketched below (assuming $r_1 > r_2$, $\sigma_1 > \sigma_2$):



The feasible set consists of all points (portfolios) which lie on this pair of lines. In particular, P_1 is the portfolio ($x = 1$) in which all wealth is invested in asset $\mathcal{S}_1$ and none in asset $\mathcal{S}_2$. Similarly, P_2 corresponds to the case in which all wealth is invested in $\mathcal{S}_2$. The segment joining P_1 and P_2 corresponds to $0 \leq x \leq 1$ i.e. no short selling of either asset. The upper and lower dashed segments correspond respectively to the cases $x > 1$ ($\mathcal{S}_2$ sold short); and $x < 0$ ($\mathcal{S}_1$ sold short).

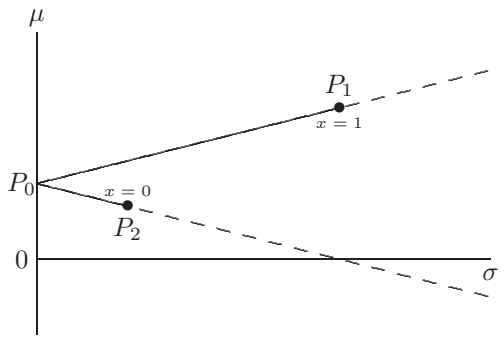
In this example it is possible to obtain a portfolio with zero risk, represented by the point P_0 in the previous graph. The parameters of this portfolio are found to be:

$$x_0 = \frac{-\sigma_2}{\sigma_1 - \sigma_2}; \qquad \mu_0 = \frac{\sigma_1 r_2 - \sigma_2 r_1}{\sigma_1 - \sigma_2}; \qquad \sigma_0 = 0.$$

[2] For each level of portfolio return μ there is only one value of σ and hence is trivially a minimum value.

2. The case $\rho = -1$
In this case the returns R_1 and R_2 are perfect negatively correlated. The feasible set reduces to another pair of straight lines given parametrically by

$$\begin{aligned} \mu &= r_1 x + r_2(1-x) \\ \sigma &= |\sigma_1 x - \sigma_2(1-x)| \end{aligned}$$



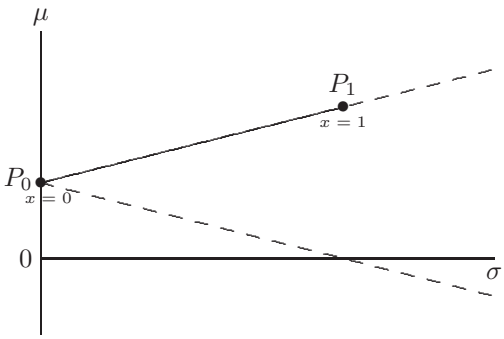
In this case it is again possible to obtain a riskless portfolio P_0, but without short selling this time.

The parameters of P_0 are: [3]

$$x_0 = \frac{\sigma_2}{\sigma_1 + \sigma_2}; \qquad \mu_0 = \frac{\sigma_1 r_2 + \sigma_2 r_1}{\sigma_1 + \sigma_2}; \qquad \sigma_0 = 0.$$

3. The case $\sigma_2 = 0$
The third degenerate case occurs when one of the assets ($\mathcal{S}_2$ say), is a riskless asset. In this case the correlation coefficient $\rho = 0$ since $s_{12} = 0$. The feasible set is once again a pair of straight lines:

$$\begin{aligned} \mu &= r_1 x + r_0(1-x) \qquad (r_2 = r_0) \\ \sigma &= |\sigma_1 x| \end{aligned}$$

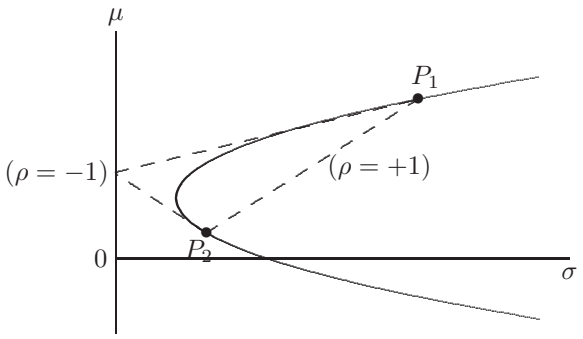


The slope of the line P_0P_1 is $\frac{r_1 - r_0}{\sigma_1}$ which is called the *price of risk* of asset $\mathcal{S}_1$, since it gives the increase in expected return per unit increase in risk σ_1.

[3]The *hedge ratio* is $(1 - x_0)/x_0 = \sigma_1/\sigma_2$ in agreement with the example at the start of this Chapter.

The General Two-Asset Case

When none of the above degenerate cases occurs, the feasible set $\mathcal{F}$ is the hyperbola described parametrically by Equation (6.1).



Example 1.4

Find the feasible set and efficient frontier for the two asset portfolio with parameters: $r_1 = 1; r_2 = \frac{1}{2}; \sigma_1^2 = 1; \sigma_2^2 = \frac{1}{2}$ and $\rho = 0$.

Solution:

The input parameters are:

$$\underset{\sim}{r} = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}; \qquad S = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

The portfolio Equations (1.4) (for the feasible) set reduce to:

$$\begin{aligned} \mu &= \tfrac{1}{2}(1+x) \\ \sigma^2 &= [x^2 + \tfrac{1}{2}(1-x)^2] \end{aligned}$$

Now eliminate x using $x = 2\mu - 1$ and $1 - x = 2(1-\mu)$ to obtain:

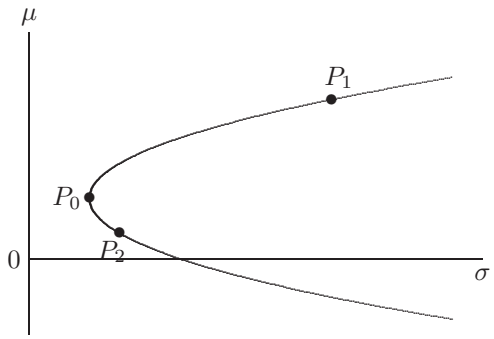
$$\begin{aligned} \sigma^2 &= (2\mu-1)^2 + 2(1-\mu)^2 \\ &= 6\mu^2 - 8\mu + 3 \\ &= 6(\mu - 2/3)^2 + 1/3 \end{aligned}$$

We therefore obtain the feasible set as an hyperbola in the standard form:

$$\sigma^2 - 6(\mu - \tfrac{2}{3})^2 = \tfrac{1}{3}.$$

The vertex P_0 of this hyperbola with $\mu_0 = \frac{2}{3}$ and $\sigma_0 = 1/\sqrt{3}$ corresponds to the minimum risk portfolio. The asymptotes of the hyperbola are the lines:

$$\sigma = \pm\sqrt{6}|\mu - \tfrac{2}{3}|.$$



The feasible set above is the set of all points on the hyperbola; i.e. all combinations of P_1 and P_2 will lie somewhere on this set. The efficient frontier is its upper half. The labelled points (not to scale) are: $P_1(1,1)$, $P_2(\frac{1}{2}, \frac{1}{\sqrt{2}})$ and the minimum risk portfolio $P_0(\frac{2}{3}, \frac{1}{\sqrt{3}})$.

7 Unrestricted $n-$Asset Portfolios

Consider the following optimisation problem:

$$\boxed{\text{Minimise} \quad Z(\underset{\sim}{x}) = -t(\underset{\sim}{r}'\underset{\sim}{x}) + \tfrac{1}{2}\underset{\sim}{x}'S\underset{\sim}{x} \quad \text{s.t.} \quad \underset{\sim}{x}'\underset{\sim}{e} = 1} \tag{7.1}$$

In this problem $\underset{\sim}{e}$ is a column vector of $n-$ones: *i.e.* $(1, 1, \ldots, 1)'$ so that the condition $\underset{\sim}{x}'\underset{\sim}{e} = 1$ is equivalent to the budget constraint (1.1): $\sum_i x_i = 1$.

Mathematically speaking, this is a quadratic programming (QP) problem subject to one equality constraint. (As noted previously, if the only constraint acting is the budget constraint, we term the problem an *unrestricted* portfolio problem). We seek a parametric solution of the form $\underset{\sim}{x} = \underset{\sim}{x}(t)$ for all allowed $t \geq 0$. This solution can then be mapped to the $\mu\sigma-$plane through the portfolio equations

$$\mu = \underset{\sim}{x}'\underset{\sim}{r} = \mu(t) \qquad \text{and} \qquad \sigma^2 = \underset{\sim}{x}'S\underset{\sim}{x} = \sigma^2(t)$$

Our task is therefore to find $\underset{\sim}{x}(t), \mu(t)$ and $\sigma^2(t)$ for $t \geq 0$.

Solution using Lagrange Parameters

The standard way of solving for the minimum of a function $V(\underset{\sim}{x})$ of several variables, subject to constraints of the form $G_1(\underset{\sim}{x}) = 0, G_2\underset{\sim}{x}) = 0, \ldots$ is by the method of Lagrange Parameters.

The method states that the constrained optimisation of $Z(\underset{\sim}{x})$ is equivalent to the unconstrained optimisation of $\mathcal{L}(\underset{\sim}{x})$ define by

$$\mathcal{L}(\underset{\sim}{x}) = Z(\underset{\sim}{x}) - \lambda_1 G_1(\underset{\sim}{x}) - \lambda_2 G_2(\underset{\sim}{x}) + \ldots$$

where the λ's are Lagrange parameters and $\mathcal{L}(\underset{\sim}{x})$ is called the Lagrangian function. The optimal solution satisfies the functional equations:

$$\frac{\partial \mathcal{L}}{\partial \underset{\sim}{x}} = 0; \qquad G_1(\underset{\sim}{x}) = 0; \quad G_2(\underset{\sim}{x}) = 0; \quad \ldots$$

In the current problem we have:

$$\mathcal{L}(\underset{\sim}{x}) = -t(\underset{\sim}{x}'\underset{\sim}{r}) + \tfrac{1}{2}\underset{\sim}{x}'S\underset{\sim}{x} - \lambda(\underset{\sim}{x}'\underset{\sim}{e} - 1)$$

and the optimal solution satisfies the $(n+1)-$ linear equations [4]:

$$\begin{cases} S\underset{\sim}{x} &= t\underset{\sim}{r} + \lambda\underset{\sim}{e} & (n - \text{equations}) \\ \underset{\sim}{x}'\underset{\sim}{e} &= 1 & (1 - \text{equation}) \end{cases}$$

When the covariance matrix S is non-singular we can write the solution as

$$\underset{\sim}{x} = S^{-1}\,(t\underset{\sim}{r} + \lambda\underset{\sim}{e}).$$

The constraint $\underset{\sim}{x}'\underset{\sim}{e} = \underset{\sim}{e}'\underset{\sim}{x} = 1$ now yields the equation $a\lambda + bt = 1$ for the Lagrange parameter $\lambda = \lambda(t)$, where

$$\boxed{a = \underset{\sim}{e}'S^{-1}\underset{\sim}{e}; \qquad \text{and} \qquad b = \underset{\sim}{e}'S^{-1}\underset{\sim}{r}.} \tag{7.2}$$

We also define the related parameters (which we shall need presently)

$$\boxed{c = \underset{\sim}{r}'S^{-1}\underset{\sim}{r} \qquad \text{and} \qquad d = ac - b^2.} \tag{7.3}$$

Note that S^{-1} is a symmetric matrix (since S is), so b can also written as $b = \underset{\sim}{r}'S^{-1}\underset{\sim}{e}$.

Solving for $\lambda(t)$ we get:

$$\boxed{\lambda(t) = \tfrac{1-bt}{a}.} \tag{7.4}$$

Theorem: *a, c and d are positive.*

Proof: Since S is positive definite, so is S^{-1} (all covariance matrices are positive semi-definite and we are assuming here S is also non-singular). Hence $a > 0$ and $c > 0$, since for any pos. def. matrix A, $\underset{\sim}{x}'A\underset{\sim}{x} > 0$ for all $\underset{\sim}{x}$.

To prove that $d > 0$, consider $(b\underset{\sim}{r} - c\underset{\sim}{e})'S^{-1}(b\underset{\sim}{r} - c\underset{\sim}{e}) > 0$.
Multiplying this out we get $b^2c - 2bc(b) + c^2a = c(ac - b^2) = cd > 0$ and then as $c > 0$ we can infer that $d > 0$. □

8 The Critical Line

Since λ from (7.4) is linear in the parameter t, so is $\underset{\sim}{x}$. Hence

$$\boxed{\underset{\sim}{x}(t) = \underset{\sim}{\alpha} + \underset{\sim}{\beta}\,t} \tag{8.1}$$

where the vectors $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ are readily found to be:

[4] We use: $\frac{\partial}{\partial \underset{\sim}{x}}\ \underset{\sim}{x}'\underset{\sim}{e} = \underset{\sim}{e}$ and $\frac{\partial}{\partial \underset{\sim}{x}}\ \underset{\sim}{x}'S\underset{\sim}{x} = 2S\underset{\sim}{x}$ if S is symmetric.

$$\boxed{\underset{\sim}{\alpha} = \tfrac{1}{a} S^{-1} \underset{\sim}{e}; \qquad \text{and} \qquad \underset{\sim}{\beta} = S^{-1}(\underset{\sim}{r} - \tfrac{b}{a}\underset{\sim}{e})} \tag{8.2}$$

This is the equation of a straight line in $n-$dimensional $\underset{\sim}{x}-$space. Markowitz calls this line the *critical line*. We shall see that the critical line plays a very important rôle in solving portfolio optimisation problems. The budget constraint $\underset{\sim}{e}'\underset{\sim}{x} = 1$ for all t imlies:

$$\underset{\sim}{e}'\underset{\sim}{\alpha} = \sum_i \alpha_i = 1 \qquad \text{and} \qquad \underset{\sim}{e}'\underset{\sim}{\beta} = \sum_i \beta_i = 0.$$

The Efficient Frontier
The image of the critical line when it is mapped to the $\mu\sigma-$plane defines the EF for this portfolio problem. We find:

$$\begin{aligned} \mu &= \underset{\sim}{x}'\underset{\sim}{r} = \underset{\sim}{r}'\underset{\sim}{x} \\ &= \lambda(\underset{\sim}{r}' S^{-1}\underset{\sim}{e}) + t(\underset{\sim}{r}' S^{-1}\underset{\sim}{r}) \\ &= b\lambda + ct \\ &= \tfrac{b}{a} + (c - \tfrac{b^2}{a})t \qquad \text{from (7.4)} \\ &= \frac{b + d\,t}{a} \end{aligned}$$

and

$$\begin{aligned} \sigma^2 &= \underset{\sim}{x}' S \underset{\sim}{x} \\ &= \underset{\sim}{x}'(\lambda\underset{\sim}{e} + t\underset{\sim}{r}) \\ &= \lambda + \mu t \\ &= \left(\tfrac{1}{a} - \tfrac{b}{a}t\right) + \left(\tfrac{b}{a} + \tfrac{d}{a}t\right) t \\ &= \frac{1 + d\,t^2}{a} \end{aligned}$$

Eliminating the risk aversion parameter t from this pair of parametric equations now gives

$$\begin{aligned} t &= \tfrac{a}{d}\left(\mu - \tfrac{b}{a}\right) \\ \sigma^2 &= \tfrac{1}{a} + \tfrac{a}{d}\left(\mu - \tfrac{b}{a}\right)^2 \end{aligned}$$

Hence, the MVF is given by either of the equations:

$$\boxed{\mu = \tfrac{b+dt}{a}; \qquad \sigma^2 = \tfrac{1+dt^2}{a}} \tag{8.3}$$

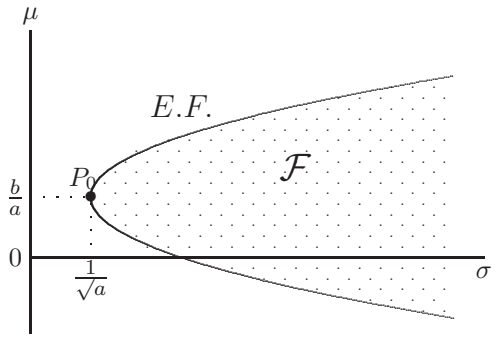
or

$$\boxed{\sigma^2 - \tfrac{a}{d}\left(\mu - \tfrac{b}{a}\right)^2 = \tfrac{1}{a}} \tag{8.4}$$

Theorem: In the $\mu\sigma-$plane:

1. *The minimum variance frontier is an hyperbola and is the image of the Critical Line C.L.*
2. *The efficient frontier is the upper branch of this hyperbola ($\mu > \frac{b}{a}$) and is parameterised by the C.L. with $t \geq 0$.*
3. *The minimum risk portfolio is the vertex of this hyperbola corresponding to $t = 0$, $\mu_0 = \frac{b}{a}$, $\sigma_0^2 = \frac{1}{a}$ and $\underset{\sim}{x}_0 = \underset{\sim}{\alpha}$.*
4. The lower branch of the hyperbola ($\mu < \frac{b}{a}$, the inefficient frontier) corresponds to all values of the C.L. with $t < 0$.
5. *Provide $n \geq 3$, the feasible set is the interior of the hyperbola (the bullet). If $n = 2$, the FS is just the hyperbola itself.*

Proof: The only problem is to show that Equation (8.4) actually corresponds to a hyperbola, rather than an ellipse or parabola. This follows from the inequalities $a > 0$ and $d > 0$. Note that the lower branch for $t < 0$ does not correspond to any rational investor since all risk aversion parameters must necessarily be non-negative. □



Example 1.5 Consider a 3-asset portfolio with parameters:

$$\underset{\sim}{r} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}; \qquad S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that since S is a diagonal matrix, the three assets in this problem have uncorrelated returns.

Obtain the following:
1. the MVF (minimum variance frontier) and EF (efficient frontier);
2. the MRP (minimum risk portfolio);
3. the projection of the critical line in the x_1x_2-plane;
4. the optimal portfolio for an investor with risk-aversion parameter $t = \frac{1}{2}$.

Solution: We shall solve this problem by applying the general formulas just derived.

First compute the vectors $S^{-1}\underset{\sim}{e}, S^{-1}\underset{\sim}{r}$, the parameters a, b, c, d and $\underset{\sim}{\alpha}, \underset{\sim}{\beta}$:

$$\begin{aligned}
S^{-1}\underset{\sim}{e} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix} \\
S^{-1}\underset{\sim}{r} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \\
a &= \underset{\sim}{e}' S^{-1}\underset{\sim}{e} = (1,1,1) \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix} = 5/2 \\
b &= \underset{\sim}{r}' S^{-1}\underset{\sim}{e} = (1,2,0) \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix} = 2 \\
c &= \underset{\sim}{r}' S^{-1}\underset{\sim}{r} = (1,2,0) \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 3 \\
d &= ac - b^2 = 7/2 \\
\underset{\sim}{\alpha} &= \frac{1}{a} S^{-1}\underset{\sim}{e} = \frac{1}{5} \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \\
\underset{\sim}{\beta} &= S^{-1}(\underset{\sim}{r} - \tfrac{b}{a}\underset{\sim}{e}) = \frac{1}{5} \begin{pmatrix} 1 \\ 3 \\ -4 \end{pmatrix}
\end{aligned}$$

The critical line $\underset{\sim}{x} = \underset{\sim}{\alpha} + t\underset{\sim}{\beta}$ becomes (in component form)

$$x_1 = \tfrac{1}{5}(t+2); \qquad x_2 = \tfrac{1}{5}(3t+1); \qquad x_3 = \tfrac{2}{5}(-2t+1).$$

Note (as required): $\sum \alpha_i = 1$ and $\sum \beta_i = 0$.

1. The formula for the MVF: $\sigma^2 - \frac{a}{d}(\mu - \frac{b}{a})^2 = \frac{1}{a}$ translates into the hyperbola $\sigma^2 - \frac{5}{7}(\mu - \frac{4}{5})^2 = \frac{2}{5}$. The efficient frontier is the upper branch of this hyperbola.

2. The MRP is the point (portfolio) $P_0(b/a, 1/\sqrt{a})$ in the $\mu\sigma-$plane, has risk-aversion $t = 0$ and leads to the parameters:

$$\mu = \tfrac{4}{5};\ \sigma^2 = \tfrac{2}{5}; \qquad x_1 = \tfrac{2}{5};\ x_2 = \tfrac{1}{5};\ x_3 = \tfrac{2}{5}.$$

This means that Mr Chicken (who always selects the MRP) would invest 40% of his funds in asset-1, 20% in asset-2 and 40% in asset-3. His expected return would then be 0.8 units at a risk of $\sqrt{0.4} = 0.632$ units.

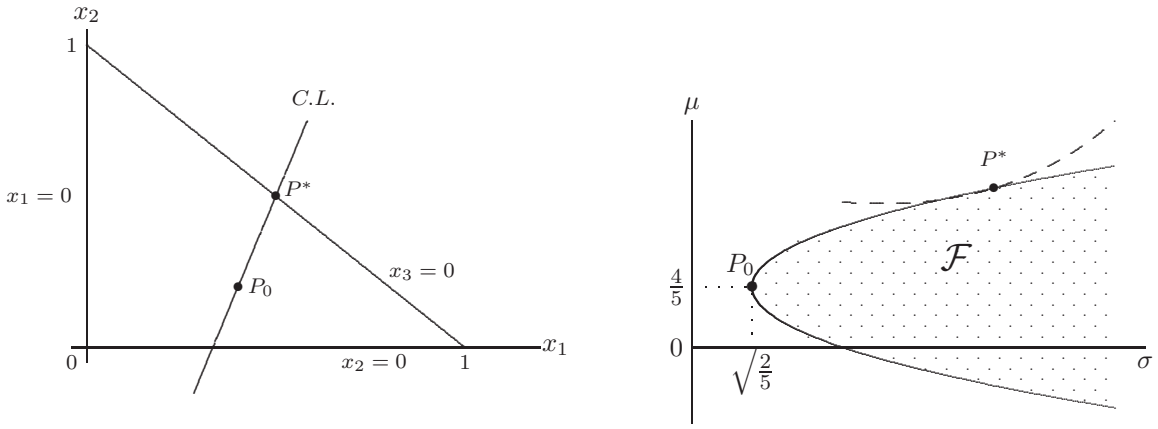
3. The projection of the critical line in the x_1x_2-plane is given parametrically by the pair of equations $x_1 = \frac{1}{5}(t+2)$ and $x_2 = \frac{1}{5}(3t+1)$. Eliminating t, this leads $3x_1 - x_2 = 1$.

4. For the investor with risk aversion parameter $t = \frac{1}{2}$:

$$\begin{aligned} \mu &= \frac{b + dt}{a} = \tfrac{3}{2} \\ \sigma^2 &= \frac{1 + dt^2}{a} = \tfrac{3}{4} \\ x_1 &= \tfrac{1}{2}; \quad x_2 = \tfrac{1}{2}; \quad x_3 = 0. \end{aligned}$$

Thus $t = \frac{1}{2}$ investors obtain their optimal portfolios by investing half their wealth in asset-1, half in asset-2 and nothing in asset-3. Their expected portfolio return is $\mu = 3/2$ units with a portfolio risk given by $\sigma = \sqrt{3/4} = 0.866$ units.

The solution for this problem is sketched in the following diagrams.



Note that the whole of 3D $\underset{\sim}{x}-$space in this example is mapped to the (shaded) feasible set $\mathcal{F}$ in the $\mu\sigma-$plane. No points x_i map to the region outside of $\mathcal{F}$. The critical line C.L. is mapped to the MVF, which is the boundary of $\mathcal{F}$. The half-line from P_0 (the MRP) towards P^* (the optimal portfolio for $t = \frac{1}{2}$) corresponds to the EF ($t \geq 0$); while the half-line from P_0 ($t = 0$) in the opposite direction, corresponds to the inefficient frontier.

8.1 Two-Fund Theorem

Let $\mathcal{P}$ and $\mathcal{Q}$ be any two distinct efficient portfolios or funds. Then the two-fund theorem states:

> All investors seeking efficient portfolios need only invest in combinations of funds $\mathcal{P}$ and $\mathcal{Q}$.

The funds are assumed to have no constraints other than the budget condition.

Proof:
Let $\underset{\sim}{x}^P$ and $\underset{\sim}{x}^Q$ be the allocation vectors for $\mathcal{P}$ and $\mathcal{Q}$ respectively. Then since these are efficient

portfolios there exists vectors $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ and non-negative parameters t^P and t^Q such that (see Eq (8.1)):

$$\underset{\sim}{x}^P = \underset{\sim}{\alpha} + t^P\underset{\sim}{\beta} \qquad \text{and} \qquad \underset{\sim}{x}^Q = \underset{\sim}{\alpha} + t^Q\underset{\sim}{\beta}$$

Now consider a new portfolio containing a proportion $(1-\theta)$ of $\mathcal{P}$ and a proportion θ of $\mathcal{Q}$. This portfolio will have allocation vector

$$\begin{aligned} \underset{\sim}{x} &= (1-\theta)\underset{\sim}{x}^P + \theta\underset{\sim}{x}^Q \\ &= \underset{\sim}{\alpha} + t\underset{\sim}{\beta} \\ \text{if} \quad t &= (1-\theta)t^P + \theta t^Q \end{aligned}$$

Hence the combined portfolio is efficient provided we choose θ such that the effective risk aversion parameter $t \geq 0$ [5]. □

The two-fund theorem has dramatic implications. Provided all the assumptions underlying MVPT are satisfied, two mutual funds could provide a complete investment service for everyone. Investors need not purchase individual stocks; they could just purchase shares in the two mutual funds. Furthermore, once we know two points on the efficient frontier, we can get all other points on the EF for appropriate choices of the parameter θ.

9 Restricted $n-$Asset Portfolios

We now consider the effect of including additional constraints of the form (2.1)

$$L_i \leq x_i \leq U_i; \qquad \text{for each } i = 1, 2, \ldots n$$

Some of the L_i may be set to $-\infty$ (no lower constraint on x_i); and some of the U_i may be set to $+\infty$ (no upper constraint on x_i). The efficient frontier (EF) is then obtained by solving the constrained optimisation problem:

Minimise	$Z = -t(\underset{\sim}{x}'\underset{\sim}{r}) + \frac{1}{2}\underset{\sim}{x}'S\underset{\sim}{x}$
s.t.	$\underset{\sim}{x}'\underset{\sim}{e} = 1$
and	$\underset{\sim}{L} \leq \underset{\sim}{x} \leq \underset{\sim}{U}$

(9.1)

This problem is to be solved for all $t \geq 0$ for which the solution is feasible. For restricted problems of this type, not all values of t in $(0, \infty)$ will necessarily lead to feasible solutions. Hence, part of the problem is to determine the allowed values of the parameter t.

An Indirect Method of Solution

We shall illustrate the approach by extending the 3-asset problem (Example 1.5) considered previously to the case of no short-selling of asset-3. The new problem is therefore:

Example 1.6

$$\underset{\sim}{r} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}; \quad S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad x_3 \geq 0$$

[5] In fact for any value of θ the combined portfolio will lie on the MVF.

We have seen that the critical line L_1 for the unrestricted version of this problem is given by

$$L_1: \quad \underset{\sim}{x} = \tfrac{1}{5}(2+t,\ 1+3t,\ 2-4t)$$

This solution becomes infeasible when $t > \frac{1}{2}$, giving $x_3 < 0$ in violation of no short-selling asset-3. For all $t > \frac{1}{2}$ we must therefore set $x_3 = 0$. This effectively reduces the 3-asset problem to a 2-asset problem for (x_1, x_2). So calculate (a, b, c, d) for these two assets to get:

$$\begin{aligned} a &= (1,1)\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \tfrac{3}{2} \\ b &= (1,2)\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2 \\ c &= (1,2)\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = 3 \\ d &= ac - b^2 = \tfrac{1}{2} \end{aligned}$$

Now get the 2-asset critical line L_2 from

$$\begin{aligned} \underset{\sim}{\alpha} &= \tfrac{1}{a} S^{-1}\underset{\sim}{e} = \tfrac{2}{3}\begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \end{pmatrix} \\ \underset{\sim}{\beta} &= S^{-1}(\underset{\sim}{r} - \tfrac{b}{a}\underset{\sim}{e}) = \begin{pmatrix} \frac{1}{3} \\ -\frac{1}{3} \end{pmatrix} \\ \underset{\sim}{x} &= \underset{\sim}{\alpha} + t\underset{\sim}{\beta}; \qquad t > \tfrac{1}{2} \\ L_2: \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \tfrac{1}{3}\begin{pmatrix} 2-t \\ 1+t \end{pmatrix} \end{aligned}$$

In summary the critical line for the constrained problem consists of two line segments in allocation $(\underset{\sim}{x})$ space: the line segment L_1 for $t \le \frac{1}{2}$ and the line segment L_2 for $t > \frac{1}{2}$. If these Markovitz critical lines are mapped to the portfolio $\mu\sigma-$plane we get two hyperbola segments that are smoothly joined at $t = \frac{1}{2}$. These hyperbola segments are given by:

$$\begin{aligned} \mathcal{H}_0: \quad \sigma^2 - \tfrac{5}{7}\left(\mu - \tfrac{4}{5}\right)^2 &= \tfrac{2}{5} \\ \mathcal{H}_1: \quad \sigma^2 - 3\left(\mu - \tfrac{4}{3}\right)^2 &= \tfrac{2}{3} \end{aligned}$$

Here are the calculations.

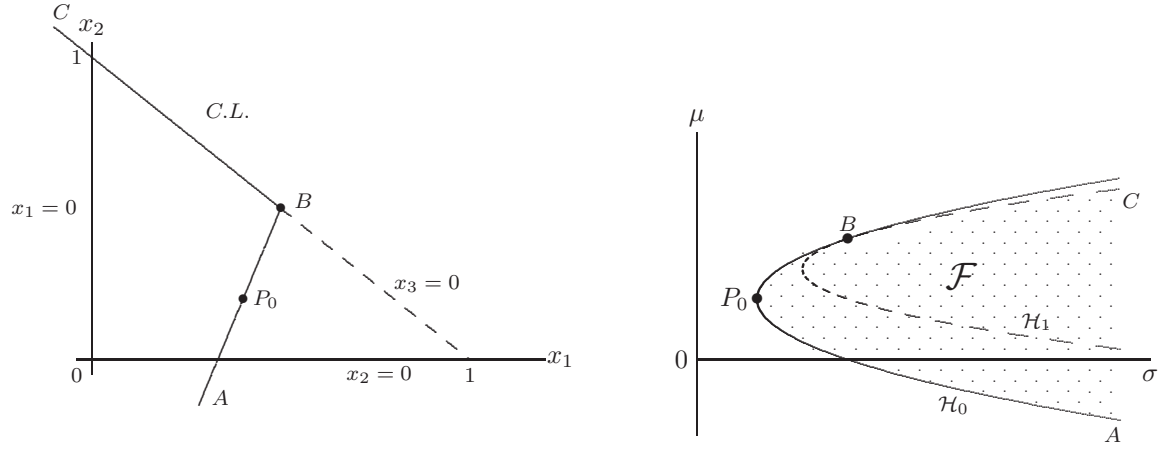
(a) For $t < \frac{1}{2}$, we have part of the original unrestricted MVF-$\mathcal{H}_0$:

$$\begin{aligned} \mu &= \underset{\sim}{x}'\underset{\sim}{r} = x_1 + 2x_2 = \tfrac{1}{5}(4+7t) \\ \sigma^2 &= \underset{\sim}{x}' S \underset{\sim}{x} = x_1^2 + 2x_2^2 + x_3^2 \\ &= \tfrac{1}{25}\left[(t+2)^2 + 2(3t+1)^2 + 4(1-2t)^2\right] \\ &= \tfrac{5}{7}(\mu - \tfrac{4}{5})^2 + \tfrac{2}{5} \qquad \text{for } \mu < \tfrac{3}{2} \end{aligned}$$

(b) For $t > \frac{1}{2}$, we have part of the new hyperbola $\mathcal{H}_1$:

$$\mu = x_1 + 2x_2 = \tfrac{1}{3}(4+t)$$

$$\begin{aligned}\sigma^2 &= x_1^2 + 2x_2^2 + x_3^2 \\ &= \tfrac{1}{9}\left[(2-t)^2 + 2(1+t)^2 + 0^2\right] \\ &= 3(\mu - \tfrac{4}{3})^2 + \tfrac{2}{3} \qquad \text{for } \mu > \tfrac{3}{2}\end{aligned}$$



Description in the x_1x_2-plane
The critical line consists of 2 segments: AB and BC. The segment AB is described by the equation $3x_1 - x_2 = 1; \quad x_3 > 0$. It corresponds to part of the original unconstrained critical line. The segment BC is described by the equation $x_3 = 0; \quad x_2 > \frac{1}{2}$. The critical line therefore satisfies the constraint $x_3 \geq 0$ at all its points. The point B, connecting the two segments, is a *corner portfolio*. The part of the C.L. mapping to the efficient frontier consists of the segments: $P_0B + BC$.

Description in the $\mu\sigma-$plane
It is easy to check that the two hyperbolas $\mathcal{H}_0$ and $\mathcal{H}_1$ are tangential to each other at the point B where $t = \frac{1}{2}, \mu = \frac{3}{2}$. So we can now describe the MVF for the restricted portfolio in the $\mu\sigma-$plane. We start at A ($t = -\infty$) on $\mathcal{H}_0$ (solid curve) and move along this hyperbola to the point B ($\mu = \frac{3}{2}$), passing the minimum risk portfolio (MRP) $P_0(t = 0,\ \mu = \frac{1}{2})$ on the way. From B we then move tangentially onto $\mathcal{H}_1$ (dashed curve), finishing eventually at C ($t = +\infty$). The feasible set $\mathcal{F}$ is the interior of this pair of touching hyperbolas and the EF corresponds to the upper branch: from the MRP at P_0 to B, then from B to C. The point B is a corner portfolio for this set of parameters and constraints.

Solution using KT-Conditions: (Advanced* Level Only)
Since the above problem has one equality constraint and one or more inequality constraints, its solution is obtained by using the method of Lagrange parameters together with the Kuhn-Tucker conditions, discussed in Chapter 4.

The Lagrangian and KT-conditions for (9.1) are as follows:

Lagrangian:

$$\boxed{\mathcal{L} = \begin{cases} \frac{1}{2}\underset{\sim}{x}'S\underset{\sim}{x} - t(\underset{\sim}{x}'\underset{\sim}{r}) - \lambda(\underset{\sim}{x}'\underset{\sim}{e} - 1) \\ -\underset{\sim}{\ell}'(\underset{\sim}{x} - \underset{\sim}{L}) - \underset{\sim}{u}'(\underset{\sim}{U} - \underset{\sim}{x}) \end{cases}} \tag{9.2}$$

KT-Conditions:

$$\begin{array}{lcll} S\underset{\sim}{x} & = & t\underset{\sim}{r} + \lambda\underset{\sim}{e} + \underset{\sim}{\ell} - \underset{\sim}{u} & \\ \underset{\sim}{x}'\underset{\sim}{e} & = & 1 & \\ \ell_i(x_i - L_i) & = & 0; & \ell_i \geq 0 \\ u_i(U_i - x_i) & = & 0; & u_i \geq 0 \end{array} \tag{9.3}$$

Remarks:

1. If there is no inequality constraint on variable x_k (for one or more $1 \leq k \leq n$) of the form $x_k \geq L_k$ (*i.e.* $L_k = -\infty$), then $\ell_k = 0$. Similarly, if there is no constraint of the form $x_k \leq U_k$ (*i.e.* $U_k = +\infty$), then $u_k = 0$.

2. There are only three possible cases for the Lagrange parameters ℓ_i and u_i:

$$\begin{array}{ccccc} \text{Either} & x_i = U_i & \text{UP} & \ell_i = 0 & u_i > 0 \\ \text{or} & x_i = L_i & \text{DOWN} & \ell_i > 0 & u_i = 0 \\ \text{or} & L_i < x_i < U_i & \text{IN} & \ell_i = 0 & u_i = 0 \end{array}$$

   If x_i is an UP- or DOWN-variable, then we say the corresponding constraint is *active*; otherwise if x_i is an IN-variable, the constraints are *inactive*.

3. Let m be the total number of inequality constraints of the form $x_i \geq L_i$ or $x_i \leq U_i$. Then the KT-conditions constitute $(n + m + 1)-$equations for the same number of unknowns: viz $n \sim \underset{\sim}{x}$ variables; $m \sim \underset{\sim}{\ell}, \underset{\sim}{u}$ variables; and $1 \sim \lambda$ variable.

4. If a solution $\underset{\sim}{x}(t)$ exists for some $t \geq 0$, then this $\underset{\sim}{x}$ lies on the Markowitz critical line. It transpires that the critical line, in the presence of inequality constraints of the type we are considering, will consist of a number of straight line segments. These segments are made up from the original unconstrained critical line and (possibly) parts of the boundary lines $\underset{\sim}{x} = \underset{\sim}{L}$ and $\underset{\sim}{x} = \underset{\sim}{U}$. Each straight line segment of the critical line maps to a hyperbola in the $\mu\sigma-$plane.

   The places in $\underset{\sim}{x}-$space where the critical line changes slope (or turns a corner) are called *corner portfolios*. It follows that the efficient frontier in the $\mu\sigma-$plane may also contain one or more corner portfolios. The EF then consists of different hyperbolas connecting the corner portfolios. It transpires that hyperbolas at corner portfolios are always tangential to each other.

   Depending on the nature of the inequality constraints, the critical line may have finite or infinite length. In some cases, there are values of $t \geq 0$, for which there are no feasible solutions.

5. In the special case of no short selling of any asset, we have inequality constraints of the form $x_i \geq 0$ for all i. The feasible region in $\underset{\sim}{x}-$space will then be a finite $n-$dimensional simplex (tetrahedron if $n = 3$). The corresponding image of this region in the $\mu\sigma-$plane (the feasible set $\mathcal{F}$) will then also be a finite region, which we earlier called the *batwing*.

—oOo—

*__Example 1.6 (cont.)__ Consider our previous portfolio with parameters $n = 3$ and :

$$\underset{\sim}{r} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}; \qquad S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \qquad x_3 > 0$$

Solution: The Lagrangian for this problem is now:

$$\mathcal{L} = \tfrac{1}{2}\, x_1^2 + x_2^2 + \tfrac{1}{2} x_3^2 - t(x_1 + 2x_2) - \lambda(x_1 + x_2 + x_3 - 1) - \ell x_3.$$

The single inequality constraint $x_3 \geq 0$ introduces exactly one extra Lagrange parameter ℓ.

The KT-conditions are:

$$\begin{array}{rcrcll} \frac{\partial \mathcal{L}}{\partial x_1} & = & x_1 - t - \lambda & = & 0 & \\ \frac{\partial \mathcal{L}}{\partial x_2} & = & 2x_2 - 2t - \lambda & = & 0 & \\ \frac{\partial \mathcal{L}}{\partial x_3} & = & x_3 - \lambda - \ell & = & 0 & \\ & & x_1 + x_2 + x_3 & = & 1 & \\ & & \ell x_3 & = & 0; & \ell \geq 0 \end{array}$$

These are 5-equations for the 5-unknowns: $(x_1, x_2, x_3, \lambda, \ell)$.

Although there exist systematic methods of solving such systems of equations, we shall employ a trial and error approach.

Begin by assuming $\ell = 0$ and $x_3 > 0$ (i.e. x_3 is an IN-variable). This leads to the same (unrestricted) solution obtained previously:

$$x_1 = \tfrac{1}{5}(t + 2); \quad x_2 = \tfrac{1}{5}(3t + 1); \quad x_3 = \tfrac{2}{5}(1 - 2t).$$

But this solution is feasible ($x_3 > 0$) if and only if $t < \frac{1}{2}$. Thus for $0 < t < \frac{1}{2}$, we have the critical line segment:

$$3x_1 - x_2 = 1; \qquad x_3 > 0.$$

For $t > \frac{1}{2}$, the above solution is no longer valid and we are forced to take $x_3 = 0$ and $\ell > 0$ (i.e. x_3 is now a DOWN-variable). In this case, the budget constraint reduces to: $x_1 + x_2 = 1$ and we obtain the solution

$$x_1 = \tfrac{1}{3}(2 - t); \quad x_2 = \tfrac{1}{3}(1 + t); \quad x_3 = 0; \quad \lambda = -\tfrac{1}{3}(2t - 1); \quad \ell = \tfrac{1}{3}(2t - 1).$$

Thus the critical line for this problem consists of two segments: (i) part of the original unrestricted critical line; and (ii) part of the boundary $x_3 = 0$. The C.L. changes slope at $t = \frac{1}{2}$ where $\underset{\sim}{x} = (\frac{1}{2}, \frac{1}{2}, 0)$ so this point is a corner portfolio for the problem [6].

—oOo—

[6] It was just a coincidence that this corner portfolio was also the optimal unrestricted portfolio for our previous $t = \frac{1}{2}$ investor

Chapter 8

CAPM

1 Single Index Model

As an introduction to this Chapter we consider what has come to be called the *single index model* portfolio. For a portfolio of many assets (*e.g.* $n > 100$), there are many input parameters required. The covariance matrix will require $\frac{1}{2}n(n+1)$ parameters (equal to 5050 when $n = 100$). In order to simplify the input specification, we postulate the existence of a linear model for the random return of each asset:

$$R_i = a_i + \beta_i R_m; \qquad i = 1, 2, \ldots, n$$

where a_i and R_m are random variables and β_i is non-random.

R_m is the random return on the *market portfolio* $\mathcal{P}_m$. This portfolio is a very special one consisting of all assets weighted by their market capitalisation [1].

a_i is the component of R_i that is independent of the market and β_i is a constant that determines how R_i varies relative to changes in the market return R_m.

Since a_i is random, we can write it in the more convenient form $a_i = \alpha_i + \epsilon_i$ where $\alpha_i = \mathbb{E}\{a_i\}$ and $\mathbb{E}\{\epsilon_i\} = 0$. We further assume that the ϵ_i are uncorrelated with the other ϵ_j and also uncorrelated with R_m. Then our linear model becomes:

$$\boxed{R_i = \alpha_i + \beta_i R_m + \epsilon_i} \tag{1.1}$$

with $\text{cov}\{\epsilon_i, \epsilon_j\} = 0$ when $i \neq j$ and $\text{cov}\{\epsilon_i, R_m\} = 0$. This means that the covariance of R_i and R_j is due only to their common relation to the market return R_m. In fact if we let $r_i = \mathbb{E}\{R_i\}$, $r_m = \mathbb{E}\{R_m\}$, $\sigma_i^2 = \mathbb{V}\{R_i\}$ and $\sigma_{\epsilon i}^2 = \mathbb{V}\{\epsilon_i\}$ then we find

$$\begin{aligned} r_i &= \alpha_i + \beta_i r_m \\ \sigma_i^2 &= \beta_i^2 \sigma_m^2 + \sigma_{\epsilon i}^2 \end{aligned}$$

[1] Thus if the market contains n_i outstanding shares of asset-i with current price π_i, the portfolio weights for $\mathcal{P}_m$ are given by $w_i = \frac{n_i \pi_i}{T}$ where $T = \sum_{i=1}^{n} n_i \pi_i$ denotes the total market capitalisation of all assets.

$$\sigma_{ij} \;=\; \mathrm{cov}\{R_i, R_j\} = \beta_i\beta_j\sigma_m^2 \qquad (i \neq j)$$

$\beta_i\sigma_m$ is called the *systematic risk*, while the parameter $\sigma_{\epsilon i}$ is called the *non-systematic risk*. In a well diversified portfolio of assets, the non-systematic risk can be made negligible compared to the systematic risk. For example, for a portfolio of risky assets with weights x_i we obtain:

$$\mu_P = \alpha_p + \beta_P r_m; \qquad \alpha_P = \sum_i x_i\alpha_i; \quad \beta_P = \sum_i x_i\beta_i$$

and

$$\sigma_P^2 = \beta_P^2\sigma_m^2 + \sum_i x_i^2\sigma_{\epsilon i}^2$$

If the portfolio is well-diversified we have approximately: $x_i \approx \frac{1}{n}$ for each i and the portfolio variance reduces to

$$\sigma_P^2 = \beta_P^2\sigma_m^2 + \frac{1}{n}\bar{\sigma}_\epsilon^2; \qquad \bar{\sigma}_\epsilon^2 = \frac{1}{n}\sum_i \sigma_{\epsilon i}^2$$

For large values of n the second term is much smaller than the first and can be neglected. Thus we have: for a well-diversified portfolio:

$$\boxed{\sigma_P \approx \beta_P\sigma_m} \tag{1.2}$$

The variable R_m in this model plays the role of an index such as the All Ordinaries Index on the Australian market or the Dow Jones Index on the NY Stock Exchange. For this single index model we need only specify $n-$parameters: the $n-$betas.

Estimating the Beta's

In practice the beta's are estimated from historical asset return data, by fitting the data to a straight line in the R_iR_m-plane. The line of best fit, *i.e.* the regression line of R_i onto R_m (see Ch. 5) is the line with equation given by (1.1) with estimated slope and intercept:

$$\hat{\beta}_i = \frac{\mathrm{cov}\{R_i, R_m\}}{\sigma_m^2} \qquad \text{and} \qquad \hat{\alpha}_i = \mathbb{E}\{R_i\} - \hat{\beta}_i\mathbb{E}\{R_m\}. \tag{1.3}$$

2 Adding a Riskless Asset

The Capital Asset Pricing Model or CAPM [2] is a model for pricing risky securities in an equilibrium frictionless market (*i.e.* one in which there are no transaction costs, taxes, restrictions on short selling or other market imperfections). The model is an extension of the Portfolio Theory considered in the previous Chapter. The key feature of CAPM postulates the existence of a riskless asset in the market, which is available for borrowing or lending to all investors. The presence of this riskless asset leads, as we shall prove, to a degenerate form of the EF (efficient frontier). It becomes a straight line, rather than an hyperbola, in the $\mu\sigma$ =plane and is called the *Capital Market Line*.

[2] Pronounced *cap-em*

2.1 The Capital Market Line

Consider a portfolio of $n-$risky assets $\mathcal{S}_1, \mathcal{S}_2, \ldots \mathcal{S}_n$ (as in Chapter 7) and a single riskless asset $\mathcal{S}_0$. Assume that the return on $\mathcal{S}_0$ is r_0. That is, r_0 is the riskless rate of return. Since $\mathcal{S}_0$ is riskless, we may write, in the notation of Chapter 7:

$$\mathbb{E}\{R_0\} = R_0 = r_0 \qquad \text{and} \qquad \mathbb{V}\{R_0\} = 0.$$

Furthermore

$$\text{cov}\{R_0, R_i\} = 0 \quad \text{for } i = 0, 1, 2, \ldots n$$

since $\text{cov}\{R_0, R_i\} = \mathbb{E}\{(R_0 - r_0)(R_i - r_i)\} = 0$ when $R_0 = r_0$.

The covariance matrix for the corresponding $(n+1)-$asset portfolio therefore has the structure

$$\hat{S} = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 0 & s_{11} & \ldots & s_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & s_{n1} & \ldots & s_{nn} \end{pmatrix}.$$

Clearly $\hat{S}$ is singular (*i.e.* the inverse $\hat{S}^{-1}$ does not exist.) Hence the portfolio analysis of Chapter 7 cannot be applied when one of the assets is riskless. It is in this sense that the problem becomes degenerate. However, a solution is still available.

Let $\underset{\sim}{x}, S, \underset{\sim}{e}$ and $\underset{\sim}{r}$ denote the same quantities defined in Chapter 7 for the case of $n-$risky assets. We shall use the notation $\hat{\mu}, \hat{\sigma}$ etc to denote quantities which *include* the riskless asset.

Let x_0 denote the allocation to the riskless asset. Then the budget constraint becomes: $\hat{\underset{\sim}{x}}'\hat{\underset{\sim}{e}} = x_0 + \underset{\sim}{x}'\underset{\sim}{e} = 1$, or equivalently

$$\boxed{x_0 = 1 - \underset{\sim}{x}'\underset{\sim}{e}} \tag{2.1}$$

Next, the return on the portfolio can be written as $\hat{\mu} = \hat{\underset{\sim}{x}}'\hat{\underset{\sim}{r}} = x_0 r_0 + \underset{\sim}{x}'\underset{\sim}{r}$, or equivalently using (2.1) above [3]

$$\boxed{\hat{\mu} = r_0 + \underset{\sim}{x}'\bar{\underset{\sim}{r}}; \qquad \bar{\underset{\sim}{r}} = \underset{\sim}{r} - r_0\underset{\sim}{e}} \tag{2.2}$$

Finally, the presence of the riskless asset cannot affect the portfolio variance and so

$$\boxed{\hat{\sigma}^2 = \hat{\underset{\sim}{x}}'\hat{S}\hat{\underset{\sim}{x}} = \underset{\sim}{x}'S\underset{\sim}{x} = \sigma^2} \tag{2.3}$$

still applies.

The EF for the $(n+1)-$asset portfolio is therefore the solution to the unconstrained optimisation problem:

$$\boxed{\min \quad Z(\underset{\sim}{x}) = -t\,\hat{\mu} + \tfrac{1}{2}\hat{\sigma}^2 = -t(r_0 + \underset{\sim}{x}'\bar{\underset{\sim}{r}}) + \tfrac{1}{2}\underset{\sim}{x}'S\underset{\sim}{x}} \tag{2.4}$$

[3] The vector $\bar{\underset{\sim}{r}}$ is called the vector of *excess* returns, since it gives the returns in excess of the riskless rate r_0.

In this form, the budget constraint is automatically satisfied (since we have removed the dependence of the problem on x_0). Thus, we minimise Z with respect to the *risky* allocations $\underset{\sim}{x}$ only. The matrix S is the corresponding covariance matrix of the risky assets, and is non-singular. As in Chapter 7, the critical line (which maps to the EF in the $\hat{\mu}\hat{\sigma}-$plane) is given parametrically by a vector equation $\underset{\sim}{x} = \underset{\sim}{x}(t)$ for $t \geq 0$.

The first order equations for the minimum are:

$$\frac{\partial Z}{\partial \underset{\sim}{x}} = S\underset{\sim}{x} - t\,\bar{\underset{\sim}{r}} = 0$$

Since S is non-singular

$$\boxed{\underset{\sim}{x} = t(S^{-1}\bar{\underset{\sim}{r}})} \tag{2.5}$$

is the required solution, giving

$$\begin{aligned} \hat{\mu} &= r_0 + t(\bar{\underset{\sim}{r}}'S^{-1}\bar{\underset{\sim}{r}}) = r_0 + \bar{c}\,t \\ \hat{\sigma}^2 &= t^2(\bar{\underset{\sim}{r}}'S^{-1}\bar{\underset{\sim}{r}}) = \bar{c}\,t^2 = (\hat{\mu} - r_0)t \end{aligned}$$

The common term $\bar{c} = \bar{\underset{\sim}{r}}'S^{-1}\bar{\underset{\sim}{r}}$ can be evaluated in terms of known quantities as follows:

$$\begin{aligned} \bar{c} &= \bar{\underset{\sim}{r}}'S^{-1}\bar{\underset{\sim}{r}} \\ &= (\underset{\sim}{r} - r_0\underset{\sim}{e})'S^{-1}(\underset{\sim}{r} - r_0\underset{\sim}{e}) \\ &= ar_0^2 - 2br_0 + c \qquad (a, b, c \quad \text{as in Chapter 7}) \\ &= d\left[\tfrac{a}{d}(r_0 - \tfrac{b}{a})^2 + \tfrac{1}{a}\right] \\ &= d\sigma_0^2 \end{aligned}$$

where σ_0^2 is the portfolio variance on the EF *for the risky assets only*, corresponding to a portfolio return of r_0, *i.e.*

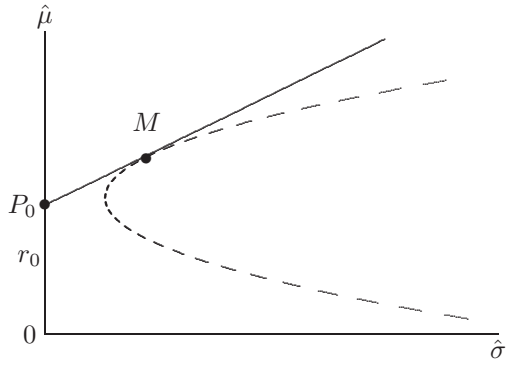
$$\sigma_0^2 = \frac{ar_0^2 - 2br_0 + c}{d} = \left[\tfrac{a}{d}(r_0 - \tfrac{b}{a})^2 + \tfrac{1}{a}\right]. \tag{2.6}$$

The Critical Line
We have $\underset{\sim}{x} = t(S^{-1}\bar{\underset{\sim}{r}})$. This is the vector equation of a straight line in $n-$dimensional $\underset{\sim}{x}-$ space (of risky assets) passing through the origin ($\underset{\sim}{x} = 0$) at $t = 0$. Thus the $t = 0$ investor (Mr Chicken, again) puts all his assets in the riskless security and none in the risky securities.

The EF (Capital Market Line)
Using the above, we find $\hat{\sigma}^2 = d\sigma_0^2\,t^2 = [(\hat{\mu} - r_0)^2/d\sigma_0^2]$ and since $d > 0$ we obtain

$$\boxed{\hat{\sigma} = \tfrac{\hat{\mu} - r_0}{\sigma_0\sqrt{d}}} \tag{2.7}$$

This straight line in the $\hat{\mu}\hat{\sigma}-$plane is the degenerate form of the efficient frontier, when one of the assets is riskless. In this context, it is called the *Capital Market Line* or CML for short.



In the above diagram, the CML is the solid line P_0M and the risky-asset MVF is the dashed hyperbola.
We prove in the next section that the CML is tangential to the risky-asset EF. Technically, there are three cases to consider, but only the first is of practical importance, since the other two are unlikely to occur in real markets.

1. When $r_0 < b/a$ (the practical case), the CML is tangential to the risky-asset EF at the point M. The portfolio associated with M is called the *Tangency Portfolio* and is constructed from proportions of risky assets only (*i.e.* $x_0 = 0$ at M).

2. If $r_0 > b/a$, then the line $\hat{\sigma} = -(\hat{\mu} - r_0)/\sigma_0\sqrt{d}$ with negative slope, is tangential to the lower branch of the risky-asset MVF. This case is unlikely to ever occur in practice.

3. In the remaining case, $r_0 = b/a$, we find that the pair of lines $\hat{\sigma} = \pm(\hat{\mu} - r_0)/\sigma_0\sqrt{d}$ are asymptotes to the risky-asset MVF.

In all cases the new feasible set is the wedge-shaped region contained in the straight lines $\hat{\sigma} = \pm(\hat{\mu} - r_0)/\sigma_0\sqrt{d}$.

Proof that CML is Tangential to Risky-Asset EF
Assume that $r_0 < b/a$. Then we have:

$$\begin{aligned} \text{CML} \quad \hat{\sigma} &= \frac{\hat{\mu} - r_0}{\sigma_0\sqrt{d}} \\ \text{Risky-Asset EF} \quad \hat{\sigma}^2 &= (a\hat{\mu}^2 - 2b\hat{\mu} + c)/d \end{aligned}$$

These intersect where:

$$(\hat{\mu} - r_0)^2 = \sigma_0^2(a\hat{\mu}^2 - 2b\hat{\mu} + c)$$

or

$$(1 - a\sigma_0^2)\hat{\mu}^2 - 2(r_0 - b\sigma_0^2)\hat{\mu} + (r_0^2 - c\sigma_0^2) = 0.$$

The discriminant of this quadratic is

$$\Delta \quad = \quad \text{“}B^2 - 4AC\text{”}$$

$$\begin{aligned} &= 4(r_0 - b\sigma_0^2)^2 - 4(1 - a\sigma_0^2)(r_0^2 - c\sigma_0^2) \\ &= 4(r_0^2 - 2br_0\sigma_0^2 + b^2\sigma_0^4) - 4(r_0^2 - ar_0^2\sigma_0^2 - c\sigma_0^2 + ac\sigma_0^4) \\ &= 4\sigma_0^2(ar_0^2 - 2br_0 + c) + 4\sigma_0^4(b^2 - ac) \\ &= 4(d\sigma_0^4 - d\sigma_0^4) \\ &= 0 \end{aligned}$$

Hence the CML and risky-asset EF intersect at a single point, which must therefore be a point of tangency. This point is the tangency portfolio M. Note that since M lies on the the risky-asset EF it must satisfy $\underset{\sim}{x}'\underset{\sim}{e} = 1$, and so by (2.1) we have $x_0 = 0$ at M. In words: the tangency portfolio contains no riskless asset.

We leave as a tutorial exercise the problem of showing that the tangency portfolio M has parameters:

$$\boxed{\mu_M = \tfrac{c-br_0}{b-ar_0}; \quad \sigma_M = \tfrac{\sigma_0\sqrt{d}}{b-ar_0}; \quad t_M = \tfrac{1}{b-ar_0}} \tag{2.8}$$

Proposition: There is an irresistable temptation to equate the tangency portfolio M with the previously discussed market portfolio. This is the key assumption that underlies the *Capital Asset Pricing Model* considered next.

3 CAPM

Let $\underset{\sim}{x}$ be the allocation vector for any feasible portfolio (but not necessarily efficient) and let $\underset{\sim}{x}_m$ denote the allocation vector for the market portfolio M (containing only risky assets).
Then

$$\begin{aligned} \text{cov}\{R, R_m\} &= \underset{\sim}{x}' S \underset{\sim}{x}_m \qquad \text{(see Appendix A.8)} \\ &= t_m \, \underset{\sim}{x}' \underset{\sim}{\bar{r}} \\ &= t_m(\hat{\mu} - r_0) \\ &= \sigma_m^2 \left(\frac{\hat{\mu} - r_0}{\hat{\mu}_m - r_0} \right) \end{aligned}$$

where $\hat{\mu} = \mathbb{E}\{R\}$ and $\hat{\mu}_m = \mathbb{E}\{R_m\}$ and in the last line we have used $\sigma_m^2 = t_m(\hat{\mu}_m - r_0)$. Hence we obtain $\hat{\mu} - r_0 = \beta(\hat{\mu}_m - r_0)$ or

$$\boxed{\mathbb{E}\{R\} = r_0 + \beta \left(\mathbb{E}\{R_m\} - r_0 \right)} \tag{3.1}$$

where

$$\boxed{\beta = \tfrac{\text{COV}\{R, R_m\}}{\sigma_m^2}} \tag{3.2}$$

This pair of equations is the mathematical statement of the CAPM. It implies that the expected return on any asset (or portfolio of assets) is a linear combination of the expected returns on the market portfolio and the riskless asset. It is important to realise that the asset (or portfolio) under

consideration need not lie on the CML.

The parameter β (or *beta*) is a measure of the risk of the asset (or portfolio) relative to the market portfolio. A high beta ($\beta > 1$) asset has higher risk than the market portfolio and is rewarded by a higher expected return. Similarly, a low beta ($\beta < 1$) asset has lower risk than the market portfolio and is penalised by a lower expected return.

CAPM also infers a linear relation between μ and β, *i.e.* between the expected return and the associated risk of an asset or portfolio. The corresponding line in the $\mu\beta-$plane is called the *Security Market Line* or SML for short.

4 The One-Fund Theorem

The one-fund theorem states:

> There exists a single fund M of risky assets such that any efficient portfolio can be constructed as a combination of the fund M and the risk-free asset P_0.

In equilibrium, all investors will select their portfolios on the CML since this is the new EF in the presence of a riskless asset. However we already know two points on the CML, namely P_0, the riskless asset and M, the market portfolio of risky assets. These two points uniquely determine the CML. Indeed, if r_m denotes the expected return on the market portfolio we have

$$\begin{aligned} \hat{\mu} &= x_0 r_0 + (1 - x_0) r_m \\ \hat{\sigma} &= (1 - x_0)\sigma_m \end{aligned}$$

or equivalently, on eliminating x_0,

$$\boxed{\frac{\hat{\mu} - r_0}{\hat{\sigma}} = \frac{r_m - r_0}{\sigma_m}} \tag{4.1}$$

This last equation states the important result:

> The *price of risk* of any asset or portfolio on the CML is equal to the price of risk of the market portfolio.

All investors will choose portfolios which are equivalent to some proportion x_0 of the riskless asset P_0 and a proportion $(1-x_0)$ of the market portfolio M. Furthermore, the relative proportions of the risky assets contained in M will be the same for all investors, since these quantities are independent of investor preferences.

Efficient portfolios between P_0 and M are called *lending* portfolios, since these can be obtained by purchasing the riskless asset (*i.e.* lending) and the market portfolio.

Efficient portfolios between M and infinity are called *borrowing* portfolios, since they can be obtained by short-selling the riskless asset (*i.e.* borrowing) and investing the proceeds in the market portfolio.

What is the market portfolio M ? Since all investors should purchase the same fund of risky securities, a little reflection indicates that it might be the portfolio of risky assets, with weights proportional to their market capitalisation. That is, the weight of a risky asset in the market portfolio is equal to the proportion of that asset's total capital value to the total market value. Thus, in equilibrium, M should be precisely the same market portfolio considered in Section 1.

5 The Characteristic Line

An interesting observation of the CAPM given by Eqs (3.1) and (3.2) is that the CML is consistent with the regression line (see Eq (1.3) connecting the random variables R and R_m. This regression line is called the *Characteristic Line* in the context of the CAPM and can be written in the form

$$\boxed{R = r_0 + \beta(R_m - r_0) + \epsilon} \tag{5.1}$$

where ϵ is a zero mean random variable or residual.

In this form of the CAPM, we see that the return on an asset or portfolio of assets contains two sources of risk: one through the parameter β and the other through the parameter ϵ. The beta of an asset determines the systematic or non-diversifiable risk. This is the risk related to the covariance of the asset and the market portfolio. The second component ϵ, determines the non-systematic or diversifiable risk, because its effects may be virtually eliminated through portfolio diversification.

6 The Pricing Model

Consider an asset whose price today ($t = 0$) is X_0 (known) and whose price at time T in the future X_T is unknown. We ask how is X_0 related to X_T ? This is none other than the problem of pricing a risky security. There exist a number of models for pricing such a risky asset, including the Principle of Expected Return and the Principle of Expected Utility. The CAPM provides another pricing model. Starting from Eq (3.1)

$$r = r_0 + \beta(r_m - r_0)$$

where r is the expected rate of return on an asset, we can identify

$$r = \mathbb{E}\{R\} = \mathbb{E}\left\{\frac{X_T - X_0}{X_0}\right\} = \frac{\mathbb{E}\{X_T\}}{X_0} - 1$$

Also

$$\begin{aligned}
\beta &= \frac{\text{cov}\{R, R_m\}}{\sigma_m^2} \\
&= \frac{\text{cov}\{X_T/X_0 - 1, R_m\}}{\sigma_m^2} \\
&= \frac{\text{cov}\{X_T, R_m\}}{X_0\sigma_m^2}
\end{aligned}$$

Hence, for any risky asset, the CAPM is equivalent to

$$\frac{\mathbb{E}\{X_T\}}{X_0} - 1 = r_0 + \frac{\mathrm{cov}\{X_T, R_m\}}{X_0 \sigma_m^2}(r_m - r_0)$$

Solving for X_0 we obtain

$$\boxed{X_0 = \frac{1}{1+r_0}\left[\mathbb{E}\{X_T\} - \frac{\mathrm{cov}\{X_T, R_m\}}{\sigma_m^2}(r_m - r_0)\right]} \tag{6.1}$$

This is the pricing formulation of CAPM for a risky asset.

Observe that the first term $(1 + r_0)^{-1}\mathbb{E}\{X_T\}$ is just the discounted expected future pay-off of the asset, corresponding to the price determined by the Principle of Expected Return. The second term is a risk premium dependent on the asset's covariance with the market and the excess return of the market over the risk-free rate.

Another important observation about this formula is that it is linear in X_T. This ensures that it is consistent with arbitrage-free pricing. For example, if two assets have random pay-outs X_T and Y_T at time T in the future, the combination $P_T = aX_T + bY_T$ must have price today equal to $P_0 = aX_0 + bY_0$. If this were not the case, we could lock in an arbitrage profit by purchasing the two asset portfolio if $P_0 < aX_0 + bY_0$, or selling the portfolio if $P_0 > aX_0 + bY_0$.

—oOo—

Chapter 9

Dynamic Programming

Dynamic Programming is a general method that can be applied to a wide class of optimisation problems. We start with a simple example.

1 Preliminary Example: the Shortest Distance Problem

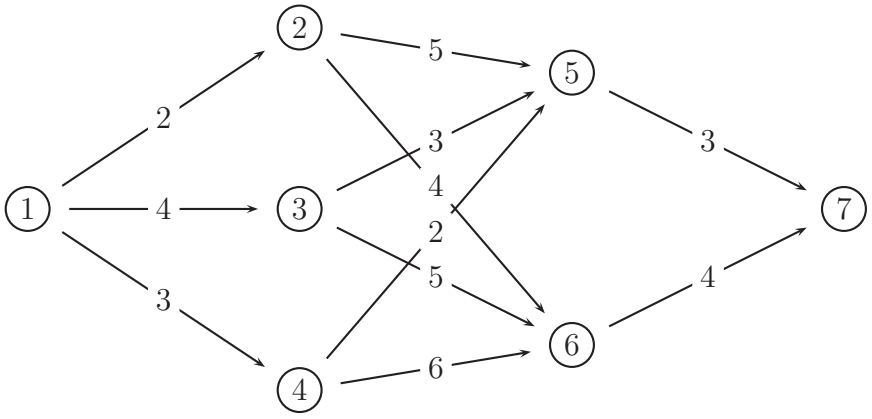


Figure 9.1: Shortest route problem

In Figure 9.1, let the circled nodes be towns and the numbers on the lines distances between them, in km say. What is the shortest route from town 1 to town 7?
The problem could be solved by *enumerating* all possibilities. This is clearly impossible, however, for large problems. We work backwards, dividing the problem into 3 *stages*, each with a certain number of *states*.

Stage 3 There are two towns 5, 6; i.e. there are two states.

Stage 2 Three towns 2, 3, 4; i.e. three states.

Stage 1 A single state, namely town 1.

At each stage a decision must be made, e.g. in stage 2, we must decide, if in town 3, say, whether to go to town 5 or 6.
Work backwards from stage 3 $\longrightarrow$ stage 2 $\longrightarrow$ stage 1.
Decision variables $x_j\,, j = 1, 2, 3$ are used to define which town we move on to, given that we are currently in stage j.
For example, one possible route is

$$1 \longrightarrow 3 \longrightarrow 5 \longrightarrow 7$$

so $x_1 = 3$, $x_2 = 5$, $x_3 = 7$ in this case.
Let s denote the state, or number of the town. Then $f_n(s, x_n)$ is the total distance for the best overall policy for the following stages. From s we go to x_n, and we have to find x_n^* that minimises $f_n(s, x_n) = f_n(s, x_n^*) = f_n^*(s)$.
At stage 3 the choices are trivial. From town 5 or 6 go to 7 $\Longrightarrow x_3^* = 7$.
Stage 3

	$f_3(s, x_3^*) = f_3^*(s)$	$f_3^*(s)$	x_3^*
s $\backslash x_3$	7	3	7
5	3	3	7
6	4	4	7

Now we know the optimal policy for the last stage of the journey, for both possible states.
Move back to stage 2.
The distance from s in stage 2 to the final town 7 is

$$\begin{aligned} & f_2(s, x_2) &=& \; d_{s,x_2} + f_3^*(x_2) \\ \text{then} \quad & f_2(s, x_2^*) &=& \; f_2^*(s) = \min_{x_2} \left[d_{s,x_2} + f_3^*(x_2) \right] \end{aligned}$$

There are now three towns (states) in stage 2 to consider: towns 2, 3 or 4.
There are two possible destinations at the next stage, so $x_2 = 5$ or 6.
Stage 2

	$f_2(s, x_2) = d_{s,x_2} + f_3^*(x_2)$			
s $\backslash x_2$	5	6	$f_2^*(s)$	x_2^*
2	$5 + 3 = 8$	$4 + 4 = 8$	8	5,6
3	$3 + 3 = 6$	$5 + 4 = 9$	6	5
4	$2 + 3 = 5$	$6 + 4 = 10$	5	5

Here $f_2(s, x_2)$ equals the distance from s to x_2 + the shortest distance from x_2 to the end, which can be read off from the previous table for stage 3.
Finally, for stage 1

$$f_1(s, x_1) = d_{s,x_1} + f_2^*(x_1).$$

Town 1, i.e. $s = 1$ is the only state in stage 1.

	$f_1(s, x_1) = d_{s,x_1} + f_2^*(x_1)$				
s $\backslash x_1$	2	3	4	$f_1^*(s)$	x_1^*
1	$2 + 8 = 10$	$4 + 6 = 10$	$3 + 5 = 8$	8	4

We can now reconstruct the solution by working through our tables of optimal strategies for each stage .
Starting at town 1, $s = 1$, the shortest route to town 7 is then $= f_1^*(1) = 8$ km.

Stage 1		Stage 2		Stage 3		Stage 4
1	$\longrightarrow$	$x_1^* = 4$	$\longrightarrow$	$x_2^* = 5$	$\longrightarrow$	$x_3^* = 7$
total distance		total distance		total distance		
$= f_1^*(1) = 8$		remaining $= f_2^*(4) = 5$		remaining $= f_3^*(5) = 3$		

So the optimal route is

$$1 \xrightarrow{3\text{ km}} 4 \xrightarrow{2\text{ km}} 5 \xrightarrow{3\text{ km}} 7$$

This method will find *all* optimal routes if there are more than one.

2 Characteristics of Dynamic Programming Problems

1. The problem can be broken up into *stages*, at each of which a decision must be taken.

2. There are a number of *states* associated with each stage — these describe all the possibilities for the conditions in the system at a particular stage (the number of states may be infinite).

3. The decision we take, given that we are in state s at stage i say, will be determined by the best choice for moving to state r at stage $(i + 1)$.

4. Note that the *optimal policy* for the remaining stages does not depend on the policy adopted in previous stages. This is clearly true for our previous example problem. This is known as the Markovian assumption in probabilistic examples.

5. A *recurrence relation* is required that allows the optimal policy for each state at stage i to be determined given the optimal policy for each state at stage $i + 1$. In the shortest route example, this was

$$f_i^*(s) = \min_{x_i} \left\{ d_{s,x_i} + f_{i+1}^*(x_i) \right\}$$

   In more general terms, this relationship is of the form

$$f_i^*(s) = \max_{x_i} / \min_{x_i} \{ f_i(s, x_i) \}$$

   where $f_i(s, x_i)$ depends on, in general, s, x_i, $f_{i+1}^*(x_i)$.

6. Because f_i^* is defined in terms of f_{i+1}^* we must solve the problem by moving backwards from the last stage to the first, finding the optimal policy at each step. Note that there is no requirement that this optimal policy should be unique.

There are two major classes of problem that can be dealt with, namely deterministic and stochastic problems. In stochastic problems the state at the next stage is not completely defined by knowing the current state and then making a policy decision — there is usually some probability distribution associated with transition to the next stage, and so optimal policy in such a case would, for example, maximise *expected* profit.
Dynamic programming is a fairly diverse field, with a variety of different types of problems that can be considered. For example, *linear programming* problems can be formulated as dynamic programming problems (usually inefficiently) and the fundamental results of the *calculus of variations* can be obtained using a dynamic programming approach.

2.1 Example: Assignment of Workers

The manager of a factory has six workers and wishes to assign them to four distinct production lines. It would be inefficient to assign a worker to more than one line, but it is possible to assign no workers to any one line if they can accomplish more on the other lines.
The following table gives the estimated % increase in productivity for each production line if allocated a given number of workers.

Number of workers	Production line 1	2	3	4
0	0	0	0	0
1	20	25	18	28
2	42	45	39	47
3	60	57	61	65
4	75	65	78	74
5	85	70	90	80
6	90	73	95	85

The problem is to maximise the productivity, subject to the above constraints.
Solution
Stages correspond to lines.
The **state** corresponds to the number of workers left to assign.
The **recurrence relation** is

$$f_n^*(s) = \max_{x_n=0,1,\cdots,s} \left\{ P(x_n) + f_{n+1}^*(s - x_n) \right\}$$

There are four stages; so we start with stage 4.
Stage 4 ($n = 4$)

$s\backslash$	$f_4^*(s) = P(x_4^*)$	x_4^*
0	0	0
1	28	1
2	47	2
3	65	3
4	74	4
5	80	5
6	85	6

so at this stage there are no choices to be made.

Stage 3 ($n = 3$)

	$f_3(s, x_3) = P(x_3) + f_4^*(s - x_3)$ number actually allocated								
$s \backslash x_3$	0	1	2	3	4	5	6	$f_3^*(s)$	x_3^*
0	0							0	0
1	28	18						28	0
2	47	46	39					47	0
3	65	65	67	61				67	2
4	74	83	86	89	78			89	3
5	80	92	104	108	106	90		108	3
6	85	98	113	126	125	118	95	126	3

Stage 2 ($n = 2$)

	$f_2(s, x_2) = P(x_2) + f_3^*(s - x_2)$								
$s \backslash x_2$	0	1	2	3	4	5	6	$f_2^*(s)$	x_2^*
0	0							0	0
1	28	25						28	0
2	47	53	45					53	1
3	67	72	73	57				73	2
4	89	92	92	85	65			92	1, 2
5	108	114	112	104	93	70		114	1
6	126	133	134	124	112	98	73	134	2

Stage 1 ($n = 1$)

	$f_1(s, x_1) = P(x_1) + f_2^*(s - x_1)$								
$s \backslash x_1$	0	1	2	3	4	5	6	$f_1^*(s)$	x_1^*
6	134	134	134	133	128	113	90	134	0, 1, 2

So (all) optimal allocations are:

stage 1		stage 2		stage 3		stage 4
0	$\rightarrow$	2	$\rightarrow$	3	$\rightarrow$	1
1	$\rightarrow$	1	$\rightarrow$	3	$\rightarrow$	1
2	$\rightarrow$	1	$\rightarrow$	2	$\rightarrow$	1
	$\searrow$	2	$\rightarrow$	0	$\rightarrow$	2