MAT 2009: Operations Research and Optimization

2010/2011

John F. Rayman

Department of Mathematics
University of Surrey

Introduction

The assessment for the this module is based on a class test counting for 10% and an assessed coursework for 15% of total module marks and an examination which counts for the remaining 75%.

Contacting the lecturer

My office is 16AA04, my internal telephone extension 2637. If I am not in my office please feel free to e-mail me on j.rayman@surrey.ac.uk to make an appointment. My website is http://personal.maths.surrey.ac.uk/st/J.Rayman.

Notes

These notes were originally produced by Dr. David Fisher. They are issued with numerous gaps which will be completed during lectures. Supplementary material will also be distributed from time to time.

Lecture attendance is therefore essential to gain a full understanding of the material.

Exercises

There are exercises at the end of each chapter which will be dealt with in tutorial sessions and solutions will be distributed progressively.

Background material

While these notes contain all the material you will need to cover during the module there are numerous excellent textbooks in the library. Although they contain far more material than will be covered during the module they provide interesting and useful background and you are encouraged to look at the early chapters of some of them. Here are some suggestions:

- *Introduction to Operations Research*, by F Hillier and G Lieberman (McGraw-Hill)
- *Operations Research: An Introduction*, by H A Taha (Prentice Hall)
- *Operations Research: Applications and Algorithms*, by Wayne L Winston (Thomson)

- *Operations Research: Principles and Practice*, by Ravindran, Phillips and Solberg (Wiley)
- *Schaum's Outline of Operations Research*, by R Bronson (McGraw-Hill)
- *Linear and Non-linear Programming* by S Nash and A Sofer (McGraw Hill)

Chapter 1

Linear Programming

1.1 Introduction

Mathematical models in business and economics often involve **optimization**, i.e. finding the optimal (best) value that a function can take. This **objective function** might represent profit (to be maximized) or expenditure (to be minimized). There may also be **constraints** such as limits on the amount of money, land and labour that are available.

Operations Research (OR), otherwise known as Operational Research or Management Science, can be traced back to the early twentieth century. The most notable developments in this field took place during World War II when scientists were recruited by the military to help allocate scarce resources. After the war there was a great deal of interest in applying OR methods in business and industry, and the subject expanded greatly during this period.

Many of the techniques that we shall study in the first part of the course were developed at that time. The Simplex algorithm for Linear Programming, which is the principal technique for solving many OR problems, was formulated by George Dantzig in 1947. This method has been applied to problems in a wide variety of disciplines including finance, production planning, timetabling and aircraft scheduling. Nowadays, computers can solve large scale OR problems of enormous complexity.

Later in the course we consider optimization of non-linear functions. The methods here are based on calculus. Non-linear problems with equality constraints will be solved using the method of Lagrange multipliers, named after Joseph Louis Lagrange (1736 - 1812). A similar method for inequality constraints was developed by Karush in 1939 and by Kuhn and Tucker in 1951.

1.2 The graphical method

In a **Linear Programming** (LP) problem, we seek to optimize an **objective function** which is a linear combination of some **decision variables** $x_1, \ldots, x_n$. These variables are restricted by a set of constraints, expressed as linear equations or inequalities.

When there are just two decision variables x_1 and x_2, the constraints can be illustrated graphically in a plane. We represent x_1 on the horizontal axis and x_2 on the vertical axis.

To find the region defined by $ax_1 + bx_2 \leq c$, first draw the straight line with equation $ax_1 + bx_2 = c$. Assuming $a \neq 0$ and $b \neq 0$, this line crosses the axes at $x_1 = \dfrac{c}{a}$ and $x_2 = \dfrac{c}{b}$.

If $b > 0$, then $ax_1 + bx_2 \leq c$ defines the region *on and below* the line and $ax_1 + bx_2 \geq c$ defines the region *on and above* the line. If $b < 0$ then this situation is reversed. It is usual to shade along the side of the line *away from* the region that is being defined.

A strict inequality ($<$ or $>$ rather than $\leq$ or $\geq$) defines the same region, but the line itself is not included. This can be shown by a dotted line.

$x_1 < c$ to the left of the vertical line $x_1 = c$, and $x_1 > c$ to the right of this line.

$x_2 < c$ below the horizontal line $x_2 = c$, and $x_2 > c$ above this line.

A simple way of deciding which side of a line satisfies a given inequality is to consider the origin $(0, 0)$. For example, $2x_1 - 3x_2 \leq 5$ defines that side of the line $2x_1 - 3x_2 = 5$ which contains the origin, since $2(0) - 3(0) \leq 5$.

If the origin lies *on* the line, consider a convenient point which is not on the line, e.g. $(1, 1)$.

For a general LP problem, a set of values of the decision variables can be represented as a vector $(x_1, \ldots, x_n) \in \mathbb{R}^n$. We make the following definitions:

- A vector which satisfies *all* the constraints is called a **feasible point**.
- The set of all feasible points is called the **feasible region** or **solution space** for the problem.
- Points where at least one constraint fails to hold are called **infeasible points**. These points lie outside the feasible region.
- The **optimum** or **optimal value** of the objective function is the maximum or minimum value, whichever the problem requires.
- A feasible point at which the optimal value occurs is called an **optimal point**, and gives an **optimal solution**.

Example 1: Containers problem

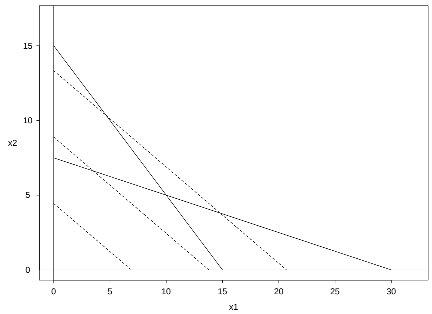
To produce two types of container, A and B, a company uses two machines, M_1 and M_2. Producing one of container A uses M_1 for 2 minutes and M_2 for 4 minutes. Producing one of container B uses M_1 for 8 minutes and M_2 for 4 minutes. The profit made on each container is £30 for type A and £45 for type B. Determine the production plan that maximizes the total profit per hour.

- Identify the decision variables.

- Formulate the objective function and state whether it is to be maximized or minimized.

- List the constraints. The allowable values of the decision variables are restricted by:

The following diagram shows the feasible region for the problem.



There are infinitely many feasible solutions. An optimal one can be found by considering the slope, and direction of increase, of the objective function $z = 30x_1 + 45x_2$.

Different values of the objective function correspond to straight lines with equations of the form $30x_1 + 45x_2 = c$ for varying values of c. These lines are all parallel, with gradient $-\dfrac{2}{3}$.

Draw such a line, e.g. $2x_1 + 3x_2 = 6$ which crosses the axes at $(3, 0)$ and $(0, 2)$. To find the maximum profit we translate this 'profit line' in the direction of increasing z, keeping its slope the same, until moving it any further would take it completely outside the feasible region.

We see that the optimum occurs at a **corner point** (or **extreme point**) of the feasible region, at the intersection of the lines that correspond to the two machine constraints. Hence the optimal values of x_1 and x_2 are the solutions of the simultaneous equations

$$2x_1 + 8x_2 = 60 \text{ and } 4x_1 + 4x_2 = 60,$$

i.e. $(x_1, x_2) = (10, 5)$. The optimal value of the objective function is then $z = 525$.

Thus the optimal production plan is to produce 10 type A containers and 5 type B containers per hour. This plan yields the maximum profit, which is £525 per hour.

We shall show later that *if* a LP problem has an optimal solution then there is an extreme point of the feasible region where this solution occurs.

There can sometimes be optimal solutions which do not occur at extreme points. If two vertices P and Q of the feasible region give equal optimal values for the objective function then the same optimal value occurs at all points on the line segment PQ.

Suppose the profit on container A is changed to £p. If $\frac{45}{4} < p < 45$ then the optimal solution still occurs at $(10, 5)$. If $p = \frac{45}{4}$ or $p = 45$ there are multiple optimal points, all yielding the same profit. If $p > 45$ then $(15, 0)$ is optimal. If $p < \frac{45}{4}$ then $(0, 7.5)$ is optimal, but 7.5 containers cannot be made. The best integer solution can be found by inspection.

Example 2 : Rose-growing problem

A market gardener grows red and white rose bushes. The red and white bushes require an area of 5 dm^2 and 4 dm^2 per bush respectively. Each red bush costs £8 per year to grow and each white bush costs £2 per year. The labour needed per year for a red bush is 1 person-hour, whereas for a white bush it is 5 person-hours. The reds each yield a profit of £2, and the whites £3, per bush per year. The total land available is at most 6100 dm^2, and the total available finance is £8000. The labour available is at most 5000 person-hours per year. How many bushes of each type should be planted to maximize the profit?

Step 1 Summarise the information

	Red	White	Max. resource available
Area (dm^2)			
Finance (£)			
Labour (person-hours)			
Profit			

Step 2 Define the decision variables.

Step 3 Specify the objective function.

Step 4 Identify the constraints and the non-negativity conditions.

Step 5 Sketch the feasible region.

Step 6 **Method 1** Sketch a line of the form $z =$ constant and translate this line, in the direction in which z increases, until it no longer intersects the feasible region.

Method 2 Evaluate z at each of the extreme points and see where it is greatest.

Final answer:

Note that the original problem was stated in words: it did not mention the variables x_1 and x_2 or the objective function z. Your final answer should not use these symbols, as they have been introduced to formulate the mathematical model.

Example 3 : Cattle feed problem

During the winter, farmers feed cattle a combination of oats and hay. The following table shows the nutritional value of oats and hay, and gives the daily nutritional requirements of one cow:

	per unit of hay	per unit of oats	daily requirement per cow
Units of protein	13.2	34.0	65.0
Units of fat	4.3	5.9	14.0
Units of calcium	0.02	0.09	0.12
Units of phosphorus	0.04	0.09	0.15
Cost per unit (£)	0.66	2.08	

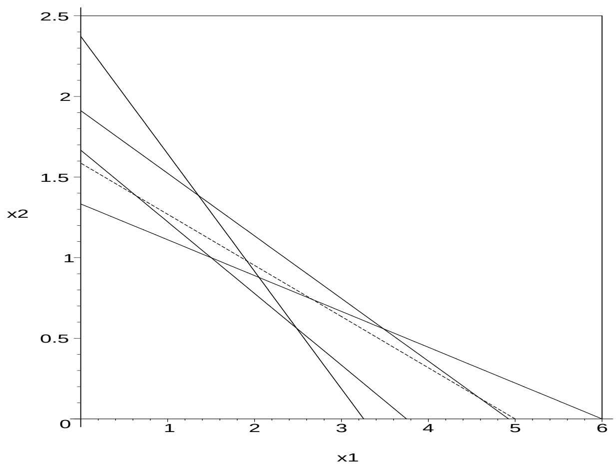
Find the optimal feeding plan and identify any redundant constraint(s).

Decision variables

Objective function

Constraints and non-negativity conditions

Draw the feasible region



Extreme points (calculated to two decimal places where appropriate)

Final Answer

1.3 Convexity and extreme points

Definitions

The feasible region for a LP problem is a subset of the Euclidean vector space $\mathbb{R}^n$, whose elements we call **points**. The **norm** of $\mathbf{x} = (x_1, \ldots, x_n)$ is $|\mathbf{x}| = \sqrt{{x_1}^2 + \cdots + {x_n}^2}$.

An equation of the form $a_1x_1 + \cdots + a_nx_n = b$, or $\mathbf{a}^t\mathbf{x} = b$, defines a **hyperplane** in $\mathbb{R}^n$. A hyperplane in $\mathbb{R}^2$ is a straight line. A hyperplane in $\mathbb{R}^3$ is a plane.

Any hyperplane determines the **half-spaces** given by $\mathbf{a}^t\mathbf{x} \leq b$ and $\mathbf{a}^t\mathbf{x} \geq b$.

A **convex linear combination** of $\mathbf{x}$ and $\mathbf{y}$ is an expression of the form $(1-r)\mathbf{x}+r\mathbf{y}$ where $0 \leq r \leq 1$. In $\mathbb{R}^2$ and $\mathbb{R}^3$, such a *weighted average* of $\mathbf{x}$ and $\mathbf{y}$ represents a point on the straight line segment between $\mathbf{x}$ and $\mathbf{y}$.

More generally, a **convex linear combination** of the points $\mathbf{x}_1, \ldots, \mathbf{x}_m$ is an expression of the form $r_1\mathbf{x}_1 + \cdots + r_m\mathbf{x}_m$, where $r_1 + \cdots + r_m = 1$ and $r_i \geq 0$ for $i = 1, \ldots, m$.

A **convex set** in $\mathbb{R}^n$ is a set S such that if $\mathbf{x}, \mathbf{y} \in S$ then $(1-r)\mathbf{x}+r\mathbf{y} \in S$ for all $r \in [0, 1]$. It follows that a convex linear combination of any number of elements of S is also in S.

A convex set can be interpreted in $\mathbb{R}^2$ and $\mathbb{R}^3$ as a region S such that if A and B are any two points in S, every point on the straight line segment AB lies in S.

Examples of convex sets include: $\mathbb{R}^n$ itself; any vector subspace of $\mathbb{R}^n$; any interval of the real line; any hyperplane; any half-space; the interior of a circle or sphere.

We can often use a graph to decide whether a subset of $\mathbb{R}^2$ is convex, e.g. $\{(x, y) : y \geq x^2\}$ is convex whereas $\{(x, y) : y \leq x^3\}$ is not convex.

If $S_1, \ldots, S_n$ are convex sets, their intersection $S_1 \cap \cdots \cap S_n$ is also convex.

An **extreme point**, **corner point** or **vertex** of a convex set S is a point $\mathbf{z} \in S$ such that there are no two distinct points $\mathbf{x}, \mathbf{y} \in S$ with $\mathbf{z} = (1-r)\mathbf{x} + r\mathbf{y}$ for some $r \in (0, 1)$.

A **neighbourhood** of a point $\mathbf{x} \in \mathbb{R}^n$ is a set of the form $\{\mathbf{y} \in \mathbb{R}^n : |\mathbf{y} - \mathbf{x}| < \varepsilon\}$ for some $\varepsilon > 0$. This can also be called an ε-neighbourhood or an **open ball** of radius ε.

Let S be a subset of $\mathbb{R}^n$. A point $\mathbf{x}$ in S is an **interior point** of S if some neighbourhood of $\mathbf{x}$ is contained in S. S is an **open set** if every point of S is an interior point.

A point $\mathbf{y}$, not necessarily in S, is a **boundary point** of S if *every* neighbourhood of $\mathbf{y}$ contains a point in S and a point not in S. The set of all boundary points of S is called the **boundary** of S. S is a **closed set** if every boundary point of S is in S.

Every point in S is either an interior point or a boundary point of S.

Some sets are neither open nor closed. Others, such as $\mathbb{R}^n$ itself, are both open and closed.

A set $S \subset \mathbb{R}^n$ is **bounded** if there exists a real number M such that $|\mathbf{x}| < M$ for all $\mathbf{x} \in S$.

S is a **compact set** if it is both closed and bounded.

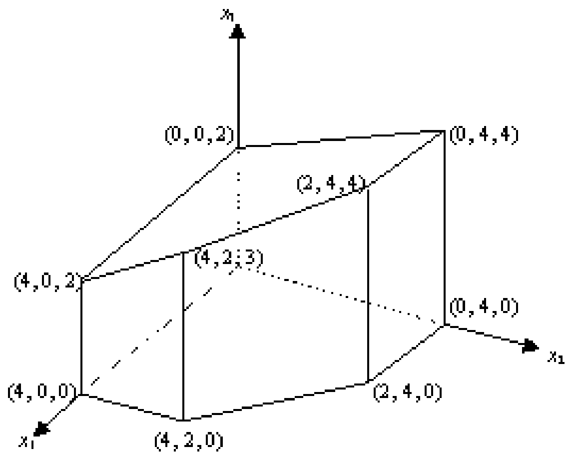
If S is not bounded it has a **direction of unboundedness**, i.e. a vector $\mathbf{u}$ such that for all $\mathbf{x} \in S$ and $k \geq 0$, $\mathbf{x} + k\mathbf{u} \in S$. (In general there are infinitely many such directions.)

The feasible region for a LP problem with n decision variables is the intersection of a finite number of half-spaces in $\mathbb{R}^n$. Hence it is convex. The region is closed, since it includes the hyperplanes which form its boundary, and it has a finite number of extreme points.

Proposition 1.1 (Weierstrass's Theorem, or the Extreme Value Theorem) *Let* f *be a continuous real-valued function defined on a non-empty compact subset S of* $\mathbb{R}^n$. *Then* $\mathrm{f}(\mathbf{x})$ *attains minimum and maximum values on S, i.e. there exist* $\mathbf{x}_m, \mathbf{x}_M \in S$ *such that* $-\infty < \mathrm{f}(\mathbf{x}_m) \leq \mathrm{f}(\mathbf{x}) \leq \mathrm{f}(\mathbf{x}_M) < \infty$ *for all* $\mathbf{x} \in S$.

Example Consider the problem: Maximize $z = 4x_1 + x_2 + 3x_3$, subject to

$x_1 + x_2 \leq 6, \quad -x_2 + 2x_3 \leq 4, \quad 0 \leq x_1 \leq 4, \quad 0 \leq x_2 \leq 4, \quad x_3 \geq 0.$



The feasible set is a convex polyhedron or *simplex* with seven plane faces, as shown. As it is closed and bounded, and any linear function is continuous, Weierstrass's theorem tells us that z attains a greatest and least value over this region. Clearly the minimum occurs at $(0, 0, 0)$. It turns out that z is maximum at $(4, 2, 3)$.

Considering other objective functions with the same feasible region, $x_1+x_2+x_3$ is maximum at $(2, 4, 4)$, while $x_1 - 3x_2 - x_3$ is maximum at $(4, 0, 0)$.

Proposition 1.2 *Let S be the intersection of a finite number of half-planes in* $\mathbb{R}^n$. *Let* $\mathbf{v}_1, \ldots, \mathbf{v}_m$ *be the extreme points of S.*

Then $\mathbf{x} \in S$ *if and only if* $\mathbf{x} = r_1\mathbf{v}_1 + \cdots + r_m\mathbf{v}_m + \mathbf{u}$ *where* $r_1 + \cdots + r_m = 1$, *each* $r_i \geq 0$, $\mathbf{u} = \mathbf{0}$ *if S is bounded and* $\mathbf{u}$ *is a direction of unboundedness otherwise.*

The proof of the 'if' part of the above result is in the Exercises. The 'only if' part is more difficult to prove, but by assuming it we can show the following:

Proposition 1.3 (The Extreme Point Theorem) *Let S be the feasible region for a linear programming problem.*

1. *If S is non-empty and bounded then an optimal solution of the problem exists, and there is an extreme point of S at which it occurs.*

2. *If S is non-empty and not bounded then* if *an optimal solution to the problem exists, there is an extreme point of S at which it occurs.*

Proof

Let the objective function be $z = c_1x_1 + \cdots + c_nx_n = \mathbf{c}^t\mathbf{x}$.
S is closed so in case 1 it is compact, hence by Proposition 1.1 z attains its maximum and minimum values at some points in S. In case 2 we are assuming that the required optimum *is* attained.

Assume z is to be maximized, and takes its maximum value over S at $\mathbf{x}^* \in S$. (If z is to be minimized, apply the following reasoning to $-z$.)

Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be the extreme points of S.

By Proposition 1.2, $\mathbf{x}^* = r_1\mathbf{v}_1 + \cdots + r_m\mathbf{v}_m + \mathbf{u}$ where $r_1 + \cdots + r_m = 1$, each $r_i \geq 0$, and $\mathbf{u}$ is $\mathbf{0}$ in case 1 or a direction of unboundedness in case 2.

Suppose, for a contradiction, that there is no extreme point of S at which z is maximum, so $\mathbf{c}^t\mathbf{v}_i < \mathbf{c}^t\mathbf{x}^*$ for $i = 1, \ldots, m$.

Then
$$\begin{aligned}
\mathbf{c}^t\mathbf{x}^* &= \mathbf{c}^t(r_1\mathbf{v}_1 + \cdots + r_m\mathbf{v}_m + \mathbf{u}) \\
&= r_1\mathbf{c}^t\mathbf{v}_1 + \cdots + r_m\mathbf{c}^t\mathbf{v}_m + \mathbf{c}^t\mathbf{u} \\
&< r_1\mathbf{c}^t\mathbf{x}^* + \cdots + r_m\mathbf{c}^t\mathbf{x}^* + \mathbf{c}^t\mathbf{u} && (\text{since } \mathbf{c}^t\mathbf{v}_i < \mathbf{c}^t\mathbf{x}^* \text{ and } r_i \geq 0 \text{ for } i = 1, \ldots, m) \\
&= (r_1 + \cdots + r_m)\mathbf{c}^t\mathbf{x}^* + \mathbf{c}^t\mathbf{u} \\
&= \mathbf{c}^t\mathbf{x}^* + \mathbf{c}^t\mathbf{u} && (\text{since } r_1 + \cdots + r_m = 1) \\
&= \mathbf{c}^t(\mathbf{x}^* + \mathbf{u}).
\end{aligned}$$

But $\mathbf{x}^* + \mathbf{u} \in S$ and $\mathbf{c}^t\mathbf{x}$ is maximized over S at $\mathbf{x}^*$, so $\mathbf{c}^t\mathbf{x}^* \geq \mathbf{c}^t(\mathbf{x}^* + \mathbf{u})$.

We have a contradiction, so z must take its maximum value at some extreme point of S. $\square$

Exercises 1

1. Solve the following Linear Programming problems graphically.

   (a) Maximize $z = 2x_1 + 3x_2$ subject to
   $2x_1 + 5x_2 \leq 10$, $x_1 - 4x_2 \geq -1$, $x_1 \geq 0$, $x_2 \geq 0$.

   (b) Minimize $z = 4x_2 - 5x_1$ subject to
   $x_1 + x_2 \leq 10$, $-2x_1 + 3x_2 \geq -6$, $6x_1 - 4x_2 \leq 13$.

2. The objective function $z = px + qy$, where $p > 0, q > 0$, is to be maximized subject to the constraints $3x + 2y \leq 6$, $x \geq 0$, $y \geq 0$.

   Find the maximum value of z in terms of p and q. (There are different cases, depending on the relative sizes of p and q.)

3. A company makes two products, A and B, using two components X and Y.

   To produce 1 unit of A requires 5 units of X and 2 units of Y.
   To produce 1 unit of B requires 6 units of X and 3 units of Y.

   At most 85 units of X and 40 units of Y are available per day.

   The company makes a profit of £12 on each unit of A and £15 on each unit of B.

   Assuming that all the units produced can be sold, find the number of units of each product that should be made per day to optimze the profit.

   If the profit on B is fixed, how low or high would the profit on A have to become before the optimal production schedule changed?

4. A brick company manufactures three types of brick in each of its two kilns. Kiln A can produce 2000 standard, 1000 oversize and 500 glazed bricks per day, whereas kiln B can produce 1000 standard, 1500 oversize and 2500 glazed bricks per day. The daily operating cost for kiln A is £400 and for kiln B is £320.

   The brickyard receives an order for 10000 standard, 9000 oversize and 7500 glazed bricks. Determine the production schedule (i.e. the number of days for which each kiln should be operated) which will meet the demand at minimum cost (assuming both kilns can be operated immediately) in each of the following separate cases. (Note: the kilns may be used for fractions of a day.)

   (a) there is no time limit,

   (b) kiln A must be used for at least 5 days,

   (c) there are at most 2 days available on kiln A and 9 days on kiln B.

5. A factory can assemble mobile phones and laptop computers. The maximum amount of the workforce's time that can be spent on this work is 10 hours per day. Before the phones and laptops can be assembled, the component parts must be purchased. The maximum value of the stock that can be held for a day's assembly work is £2200.

   In the factory, a mobile phone takes 10 minutes to assemble using £10 worth of components whereas a laptop takes 1 hour 40 minutes to assemble using £500 worth of components.

   The profit made on a mobile phone is £5 and the profit on a laptop is £100.

   (a) Summarise the above information in a table.

(b) Assuming that the factory can sell all the phones and laptops that it assembles, formulate the above information into a Linear Programming problem.

(c) Determine the number of mobile phones and the number of laptops that should be made in a day to maximise profit.

(d) The market for mobile phones is saturating. In response, retailers are dropping prices which means reduced profits. How low can the profit on a mobile phone go before the factory should switch to assembling only laptops?

6. In the Containers problem (Example 1), suppose we write the machine constraints as $2x_1 + 8x_2 + x_3 = 60$ and $4x_1 + 4x_2 + x_4 = 60$, where x_3 and x_4 are the number of minutes in an hour for which M_1 and M_2 are *not* used, so $x_3 \geq 0$ and $x_4 \geq 0$.

   Show that the objective function can be written as

$$z = 30\left(10 + \frac{1}{6}x_3 - \frac{1}{3}x_4\right) + 45\left(5 - \frac{1}{6}x_3 + \frac{1}{12}x_4\right).$$

   By simplifying this, deduce that that the maximum value of z occurs when $x_3 = x_4 = 0$ and state this maximum value.

7. By sketching graphs and using the fact that the straight line joining any two points of a convex set lies in the set, decide which of the following subsets of $\mathbb{R}^2$ are convex.

   (a) $\{(x, y) \,:\, xy \leq 1, x \geq 0, y \geq 0\}$,
   (b) $\{(x, y) \,:\, xy \leq -1,\, x \geq 0\}$,
   (c) $\{(x, y) \,:\, y - x^2 \leq 1\}$,
   (d) $\{(x, y) \,:\, 2x^2 + 3y^2 < 6\}$,
   (e) $\{(x, y) \,:\, x^2 - y^2 = 1\}$,
   (f) $\{(x, y) \,:\, y \leq \ln x,\, x > 0\}$.

8. Let $\mathbf{a} = (a_1, \ldots, a_n)$ be a fixed element of $\mathbb{R}^n$ and let b be a real constant. Let $\mathbf{x}_1$ and $\mathbf{x}_2$ lie in the half-space $\mathbf{a}^t\mathbf{x} \leq b$, so that $\mathbf{a}^t\mathbf{x}_1 \leq b$ and $\mathbf{a}^t\mathbf{x}_2 \leq b$. Show that for all $r \in [0, 1]$, $\mathbf{a}^t((1 - r)\mathbf{x}_1 + r\mathbf{x}_2) \leq b$. Deduce that the half-space is a convex set.

9. Let S and T be convex sets. Show that their intersection $S \cap T$ is a convex set.

   (Hint : let $\mathbf{x}, \mathbf{y} \in S \cap T$, so $\mathbf{x}, \mathbf{y} \in S$ and $\mathbf{x}, \mathbf{y} \in T$. Why must $(1 - r)\mathbf{x} + r\mathbf{y}$ be in $S \cap T$ for $0 \leq r \leq 1$?)

   Generalise this to show that if $S_1, \ldots, S_n$ are convex sets then $S = S_1 \cap \cdots \cap S_n$ is convex. Deduce that the feasible region for a linear programming problem is convex.

10. Let S be the feasible region for a Linear Programming problem. If an optimal solution to the problem does *not* exist, what can be deduced about S from Proposition 1.3?

11. Prove that every extreme point of a convex set S is a boundary point of S. (Method: suppose $\mathbf{x}$ is in S and is not a boundary point. Then some neighbourhood of $\mathbf{x}$ must be contained in S (why?) Deduce that $\mathbf{x}$ is a convex linear combination of two points in this neighbourhood and so $\mathbf{x}$ is not an extreme point.)

12. Prove the 'if' part of Proposition 1.2 as follows. Let the half-planes defining S be ${\mathbf{a}_1}^t\mathbf{x} \leq b_1, \ldots, {\mathbf{a}_k}^t\mathbf{x} \leq b_k$. As $\mathbf{v}_i \in S$, all these inequalities hold when $\mathbf{x} = \mathbf{v}_i$ for $i = 1, \ldots, m$.

    Show that $\mathbf{x} = r_1\mathbf{v}_1 + \cdots + r_m\mathbf{v}_m$ also satisfies all the inequalities, where $r_1 + \cdots + r_m = 1$ and each $r_i \geq 0$.

    Deduce further that if S is unbounded and $\mathbf{u}$ is a direction of unboundedness then $r_1\mathbf{v}_1 + \cdots + r_m\mathbf{v}_m + \mathbf{u} \in S$.

Chapter 2

The Simplex Method

2.1 Matrix formulation of Linear Programming problems

If a linear programming problem has more than two decision variables then a graphical solution is not possible. We therefore develop an algebraic, rather than geometric, approach.

Definitions

$\mathbf{x}$ is a **non-negative vector**, written $\mathbf{x} \geq \mathbf{0}$, if every entry of $\mathbf{x}$ is positive or zero.

The set of non-negative vectors in $\mathbb{R}^n$ is denoted by $\mathbb{R}^n_+$.

$\mathbf{u} \geq \mathbf{v}$ means that every entry of the vector $\mathbf{u}$ is greater than or equal to the corresponding entry of $\mathbf{v}$. Thus $\mathbf{u} \geq \mathbf{v}$, or $\mathbf{v} \leq \mathbf{u}$, is equivalent to $\mathbf{u} - \mathbf{v} \geq \mathbf{0}$.

Let $\mathbf{x}$ and $\mathbf{y}$ be non-negative vectors. Then it is easy to show that:

(i) $\mathbf{x} + \mathbf{y} = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{y} = \mathbf{0}$, (ii) $\mathbf{x}^t\mathbf{y} \geq 0$, (iii) $\mathbf{u} \geq \mathbf{v} \Rightarrow \mathbf{x}^t\mathbf{u} \geq \mathbf{x}^t\mathbf{v}$.

Recall that for matrices A and B, $(\mathrm{AB})^t = \mathrm{B}^t\mathrm{A}^t$.

For vectors $\mathbf{x}$ and $\mathbf{y}$, we have $\mathbf{x}^t\mathbf{y} = \mathbf{y}^t\mathbf{x}$ and $\mathbf{x}^t\mathrm{A}\mathbf{y} = \mathbf{y}^t\mathrm{A}^t\mathbf{x}$ (these are all scalars).

We shall sometimes need to work with **partitioned matrices**. Recall that if two matrices can be split into blocks which are conformable for matrix multiplication, then

$$\left(\begin{array}{c|c} \mathrm{A}_{11} & \mathrm{A}_{12} \\ \hline \mathrm{A}_{21} & \mathrm{A}_{22} \end{array}\right)\left(\begin{array}{c|c} \mathrm{B}_{11} & \mathrm{B}_{12} \\ \hline \mathrm{B}_{21} & \mathrm{B}_{22} \end{array}\right) = \left(\begin{array}{c|c} \mathrm{A}_{11}\mathrm{B}_{11} + \mathrm{A}_{12}\mathrm{B}_{21} & \mathrm{A}_{11}\mathrm{B}_{12} + \mathrm{A}_{12}\mathrm{B}_{22} \\ \hline \mathrm{A}_{21}\mathrm{B}_{11} + \mathrm{A}_{22}\mathrm{B}_{21} & \mathrm{A}_{21}\mathrm{B}_{12} + \mathrm{A}_{22}\mathrm{B}_{22} \end{array}\right).$$

Now consider a typical Linear Programming problem. We seek values of the decision variables $x_1, x_2, \ldots, x_n$ which optimize the *objective function*

$$z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

subject to the *constraints*

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & \leq & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & \leq & b_2 \\ & \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & \leq & b_m \end{array}$$

and the *non-negativity restrictions* $x_1 \geq 0, \quad x_2 \geq 0, \quad \ldots, \quad x_n \geq 0.$

A maximization LP problem can be written in matrix-vector form as

$$\begin{array}{ll} \textbf{Maximize} & z = \mathbf{c}^t\mathbf{x} \\ \textbf{subject to} & \mathrm{A}\mathbf{x} \le \mathbf{b} \\ \textbf{and} & \mathbf{x} \ge \mathbf{0} \end{array} \qquad \text{(LP1)}$$

where A is the $m \times n$ matrix with (i, j) entry a_{ij} and

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

The problem (LP1) is said to be **feasible** if the constraints are consistent, i.e. if there exists $\mathbf{x} \in \mathbb{R}^n_+$ such that $\mathrm{A}\mathbf{x} \le \mathbf{b}$. Any such vector $\mathbf{x}$ is a **feasible solution** of (LP1). If a feasible point $\mathbf{x}$ maximizes z subject to the constraints, $\mathbf{x}$ is an **optimal solution**.

The problem is **unbounded** if there is no finite maximum over the feasible region, i.e. there is a sequence of vectors $\{\mathbf{x}_k\}$ satisfying the constraints, such that $\mathbf{c}^t\mathbf{x}_k \to \infty$ as $k \to \infty$.

The **standard form** of a LP problem is defined as follows:

- The objective function is to be *maximized*.
- All constraints are *equations* with *non-negative* right-hand sides.
- All the variables are *non-negative*.

Any LP problem can be converted into standard form by the following methods:

Minimizing f(x) is equivalent to maximizing $-$f(x). Thus the problem

$$\textbf{Minimize} \quad c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

is equivalent to

$$\textbf{Maximize} \quad -c_1x_1 - c_2x_2 - \cdots - c_nx_n$$

subject to the *same* constraints, and the optimum occurs at the same values of $x_1, \ldots, x_n$.

Any *equation* with a negative right hand side can be multiplied through by -1 so that the right-hand side becomes positive. Remember that if an *inequality* is multiplied through by a *negative* number, the inequality sign must be reversed.

A constraint of the form $\le$ can be converted to an equation by adding a non-negative **slack variable** on the left-hand side of the constraint.

A constraint of the form $\ge$ can be converted to an equation by subtracting a non-negative **surplus variable** on the left-hand side of the constraint.

2.1.1 Example

Suppose we start with the problem

$$\textbf{Minimize} \qquad z = -3x_1 + 4x_2 - 5x_3$$

$$\textbf{subject to} \quad \left\{ \begin{array}{rcrcrcr} 3x_1 & + & 2x_2 & - & 3x_3 & \le & 4 \\ 2x_1 & - & 3x_2 & + & x_3 & \le & -5 \end{array} \right.$$

$$\textbf{and} \qquad x_1 \ge 0, x_2 \ge 0, x_3 \ge 0.$$

A variable is **unrestricted** if it is allowed to take both positive and negative values. An unrestricted variable x_j can be expressed in terms of two non-negative variables by substituting $x_j = x'_j - x''_j$ where both x'_j and x''_j are non-negative. The substitution must be used throughout, i.e. in all the constraints and in the objective function.

2.1.2 Example

Suppose we have to maximize $z = x_1 - 2x_2$ subject to the constraints

$x_1 + x_2 \leq 4,\ 2x_1 + 3x_2 \geq 5,\ x_1 \geq 0$ and x_2 unrestricted in sign.

If a LP problem involves n main (original) variables and m inequality constraints then we need m additional (slack or surplus) variables, so the total number of variables becomes $n + m$. Then the problem (LP1) can be written in the form

$$\begin{array}{ll} \textbf{Maximize} & z = \widetilde{\mathbf{c}}^t \widetilde{\mathbf{x}} \\ \textbf{subject to} & \widetilde{\mathrm{A}}\widetilde{\mathbf{x}} = \mathbf{b} \\ \textbf{and} & \widetilde{\mathbf{x}} \geq \mathbf{0} \end{array} \qquad \text{(LP2)}$$

where $\widetilde{\mathrm{A}}$ is an $m \times (n+m)$ matrix, $\widetilde{\mathbf{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_{n+m} \end{pmatrix}$, $\widetilde{\mathbf{c}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

The constraints are now expressed as a linearly independent set of m equations in $n+m$ unknowns, representing m hyperplanes in $\mathbb{R}^{n+m}$. A particular solution can be found by setting any n variables to zero and solving for the remaining m.

A **feasible solution** of (LP2) is a vector $\mathbf{v} \in \mathbb{R}_+^{n+m}$ such that $\widetilde{\mathrm{A}}\mathbf{v} = \mathbf{b}$.

Provided $\mathbf{b} \geq \mathbf{0}$, $\widetilde{\mathrm{A}} = (\mathrm{A} \mid \mathrm{I}_m)$ and then (LP2) has an obvious feasible solution $\widetilde{\mathbf{x}} = \left(\frac{\mathbf{0}}{\mathbf{b}}\right)$.

2.1.3 Example

Consider the Linear Programming problem: Maximize $z = x_1 + 2x_2 + 3x_3$
subject to $2x_1 + 4x_2 + 3x_3 \leq 10$, $3x_1 + 6x_2 + 5x_3 \leq 15$ and $x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$.

Introducing non-negative slack variables x_4 and x_5, the problem can be written as

$$\text{Maximize } z = \begin{pmatrix} 1 & 2 & 3 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \text{ subject to } \begin{pmatrix} 2 & 4 & 3 & 1 & 0 \\ 3 & 6 & 5 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 10 \\ 15 \end{pmatrix},$$

where $x_1, \ldots, x_5 \geq 0$.

Setting any 3 variables to zero, if the resulting equations are consistent we can solve for the other two; e.g. $(1, 2, 0, 0, 0)$ and $(0, 0, 0, 10, 15)$ are feasible solutions for $(x_1, x_2, x_3, x_4, x_5)$.

Note that setting $x_3 = x_4 = x_5 = 0$ gives $2x_1 + 4x_2 = 10$, $3x_1 + 6x_2 = 15$ which are the same equation and thus do *not* have a unique solution for x_1 and x_2.

A *unique* solution of (LP2) obtained by setting n variables to zero is called a **basic solution**.

If it is also feasible, i.e. non-negative, it is called a **basic feasible solution (bfs)**.

The n variables set to zero are called **non-basic variables**. The remaining m (some of which *may* be zero) are **basic variables**. The set of basic variables is called a **basis**.

In the above example, $(0, 0, 10/3, 0, -5/3)$ is a basic infeasible solution. $(1, 2, 0, 0, 0)$ is a feasible solution but not a bfs. $(0, 0, 3, 1, 0)$ is a basic feasible solution, with basis $\{x_3, x_4\}$.

It can be proved that $\widetilde{\mathbf{x}}$ is a basic feasible solution of (LP2) if and only if it is an extreme point of the feasible region for (LP2) in $\mathbb{R}^{n+m}$. The vector $\mathbf{x}$ consisting of the first n components of $\widetilde{\mathbf{x}}$ is then an extreme point of the original feasible region for (LP1) in $\mathbb{R}^n$.

In any LP problem that has an optimal solution, we know by Proposition 1.3 that an optimal solution exists at an extreme point. Hence we are looking for the *basic feasible solution* which optimizes the objective function.

Suppose we have a problem with 25 original variables and 15 constraints, giving rise to 15 slack or surplus variables. In each basic feasible solution, 25 variables are set equal to zero. There are $\binom{40}{25}$ possible sets of 25 variables which could be equated to zero, which is more than 4×10^{10} combinations. Clearly an efficient method for choosing the sets of variables to set to zero is required! Suppose that

- we are able to find an initial basic feasible solution, say $\widetilde{\mathbf{x}}_1$,
- we have a way of checking whether a given bfs is optimal, and

- we have a way of moving from a non-optimal bfs $\widetilde{\mathbf{x}}_i$ to another, $\widetilde{\mathbf{x}}_{i+1}$, that gives a better value of the objective function.

Combining these three steps will yield an algorithm for solving the LP problem.

If $\widetilde{\mathbf{x}}_1$ is optimal, then we are done. If it is not, then we move from $\widetilde{\mathbf{x}}_1$ to $\widetilde{\mathbf{x}}_2$, which by definition is better than $\widetilde{\mathbf{x}}_1$. If $\widetilde{\mathbf{x}}_2$ is not optimal then we move to $\widetilde{\mathbf{x}}_3$ and so on. Since the number of extreme points is finite and we always move towards a better one, we must ultimately find the optimal one. The Simplex method is based on this principle.

2.2 The Simplex algorithm

Consider the Linear Programming problem:

Maximize $z = 12x_1 + 15x_2$ subject to $5x_1 + 6x_2 \leq 85$, $2x_1 + 3x_2 \leq 40$ and $x_1 \geq 0, x_2 \geq 0$.

To use the Simplex method we must first write the problem in standard form:

$$\textbf{Maximize} \qquad z = 12x_1 + 15x_2$$

$$\textbf{subject to} \quad \begin{cases} 5x_1 + 6x_2 + x_3 = 85 & (1) \\ 2x_1 + 3x_2 + x_4 = 40 & (2) \end{cases}$$

$$\textbf{and} \qquad x_1 \geq 0,\ x_2 \geq 0.$$

The two equations in four unknowns have infinitely many solutions. Setting any two of x_1, x_2, x_3, x_4 to zero gives a unique solution for the other two, yielding a basic solution of the problem.

If we take $x_1 = 0, x_2 = 0, x_3 = 85, x_4 = 40$ this is certainly a basic feasible solution; since it gives $z = 0$ it is clearly not optimal. z can be made larger by increasing x_1 or x_2.

Equation (2) gives $x_2 = \dfrac{1}{3}(40 - 2x_1 - x_4)$.

Now express z in terms of x_1 and x_4 only:

$$z = 12x_1 + 15x_2 = 12x_1 + 5(40 - 2x_1 - x_4) = 200 + 2x_1 - 5x_4.$$

Taking $x_1 = x_4 = 0$ gives $z = 200$. This is a great improvement on 0, and corresponds to increasing x_2 to $\dfrac{40}{3}$ so that the second constraint holds as an equation. We have moved from the origin to another vertex of the feasible region. Now $x_3 = 5$, so there are still 5 units of slack in the first constraint. All the x_j are non-negative so we still have a bfs.

z can be improved further by increasing x_1.

Eliminate x_2 between the constraint equations: $(1) - 2 \times (2)$ gives $x_1 = 5 - x_3 + 2x_4$.

We then have $z = 200 + 2x_1 - 5x_4 = 200 + 2(5 - x_3 + 2x_4) - 5x_4 = 210 - 2x_3 - x_4$.

As all the variables are non-negative, increasing x_3 or x_4 above zero will make z smaller than 210. Hence the maximum value of z is 210 and this occurs when $x_3 = x_4 = 0$, i.e. when there is no slack in either constraint. Then $x_1 = 5, x_2 = 10$. We have moved round the feasible region to the vertex where the two constraint lines intersect.

The above working can be set out in an abbreviated way. The objective function is written as $z - 12x_1 - 15x_2 = 0$. This and the constraints are three equations in five unknowns x_1, x_2, x_3, x_4 and z. This system of linear equations holds at every feasible point.

We obtain an equivalent system by forming linear combinations of these equations, so long as the resulting set of equations remains linearly independent. This can be carried out most easily by writing the equations in the form of an augmented matrix and carrying out row operations, as in Gaussian elimination. This matrix is written in a way which helps us to identify the basic variables at each stage, called a **simplex tableau**.

Eventually we should get an expression for the objective function in the form $z = k - \sum \alpha_j x_j$ where each $\alpha_j \geq 0$, such as $z = 210 - 2x_3 - x_4$ in the example above. Then increasing any of the x_j would decrease z, so we have arrived at a maximum value of z.

A **simplex tableau** consists of a grid with headings for each of the variables and a row for each equation, including the objective function. Under the heading 'Basic' are the variables which yield a basic feasible solution when they take the values in the 'Solution' column and the others are all zero. The '=' sign comes immediately before the 'Solution' column.

You will find various forms of the tableau in different books. Some omit the z column and/or the 'Basic' column. The objective function is often placed at the top. Some writers define the standard form to be a minimization rather than a maximization problem.

For the problem on the previous page, the initial simplex tableau is:

Basic	z	x_1	x_2	x_3	x_4	Solution
x_3	0	5	6	1	0	85
x_4	0	2	3	0	1	40
z	1	-12	-15	0	0	0

This tableau represents the initial basic feasible solution $x_1 = x_2 = 0, x_3 = 85, x_4 = 40$ giving $z = 0$. The basic variables at this stage are x_3 and x_4.

The negative values in the z row show that this solution is not optimal: $z - 12x_1 - 15x_2 = 0$ so z can be made larger by increasing x_1 or x_2 from 0.

Increasing x_2 seems likely to give the best improvement in z, as the largest coefficient in z is that of x_2. We carry out row operations on the tableau so as to make x_2 basic. One of the entries in the x_2 column must become 1 and the others must become 0. The right-hand sides must all remain non-negative. This is achieved by choosing the entry '3' as the 'pivot'.

Divide Row 2 by 3 to make the pivot 1. Then combine multiples of this row (only) with each of the other rows so that all other entries in the pivot column become zero:

Basic	z	x_1	x_2	x_3	x_4	Solution	
x_3	0	1	0	1	-2	5	$R_1 := R_1 - 2R_2$
x_2	0	2/3	1	0	1/3	40/3	$R_2 := R_2/3$
z	1	-2	0	0	5	200	$R_3 := R_3 + 5R_2$

This represents the bfs $z = 200$ when $x_1 = 0, x_2 = \dfrac{40}{3}, x_3 = 5, x_4 = 0$. x_4 has left the basis (it is the 'departing variable') and x_2 has entered the basis (it is the 'entering variable').

Now the bottom row says $z - 2x_1 + 5x_4 = 200$ so we can still increase z by increasing x_1 from 0. Thus x_1 must enter the basis. The entry '1' in the top left of the tableau is the new pivot, and all other entries in its column must become 0. x_3 leaves the basis.

Basic	z	x_1	x_2	x_3	x_4	Solution	
x_1	0	1	0	1	-2	5	
x_2	0	0	1	$-2/3$	$5/3$	10	$R_2 := R_2 - \frac{2}{3}R_1$
z	1	0	0	2	1	210	$R_3 := R_3 + 2R_1$

Now there are no negative numbers in the z row, which says $z + 2x_3 + x_4 = 210$.

Increasing either of the non-basic variables x_3, x_4 from 0 cannot increase z, so we have the optimal value $z = 210$ when $x_3 = x_4 = 0$.

The tableau shows that this constrained maximum occurs when $x_1 = 5$ and $x_2 = 10$.

The rose-growing problem revisited

Consider the rose-growing problem from Chapter 1.

Step 1 *Write the problem in standard form.*

Maximise $$z = 2x_1 + 3x_2$$

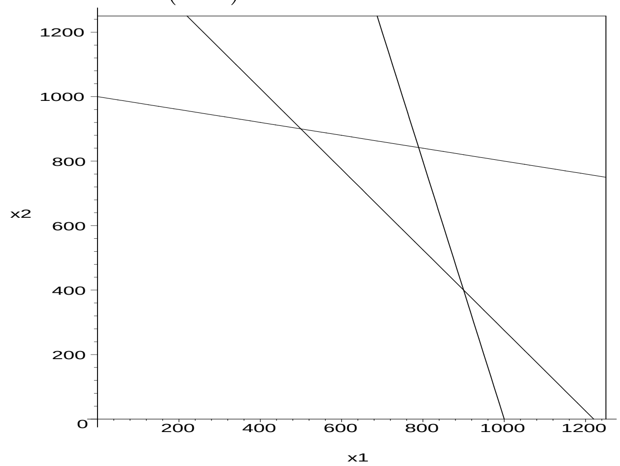
subject to
$$\begin{cases} 5x_1 + 4x_2 + x_3 = 6100 \\ 8x_1 + 2x_2 + x_4 = 8000 \\ x_1 + 5x_2 + x_5 = 5000 \end{cases}$$

and $$x_j \geq 0 \text{ for } j = 1, \ldots, 5.$$

The slack variables x_3, x_4, x_5 represent the amount of spare area, finance and labour that are available. The extreme points of the feasible region are:

Extreme point	x_1	x_2	x_3	x_4	x_5	Objective function z
O	0	0	6100	8000	5000	0
A	0	1000	2100	6000	0	3000
B	500	900	0	2200	0	3700
C	900	400	0	0	2100	3000
D	1000	0	1100	0	4000	2000

The boundary lines each have *one* of $x_1, \ldots, x_5$ equal to zero. Hence the vertices of the feasible region, which occur where two boundary lines intersect, correspond to solutions in which *two* of the five variables in (LP2) are zero.



The simplex method systematically moves around the boundary of this feasible region, starting at the origin and improving the objective function at every stage.

Step 2 *Form the initial tableau.*

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution
x_3	0	5	4	1	0	0	6100
x_4	0	8	2	0	1	0	8000
x_5	0	1	5	0	0	1	5000
z	1	-2	-3	0	0	0	0

Note that the bottom row comes from writing the objective function as $z - 2x_1 - 3x_2 = 0$.

The tableau represents a set of equations which hold simultaneously at every feasible solution. The '=' sign in the equations occurs immediately before the *solution* column. There are $5 - 3 = 2$ basic variables. Each column headed by a basic variable has an entry 1 in exactly one row and 0 in all the other rows. In the above tableau, x_1 and x_2 are non-basic. Setting these to zero, an initial basic feasible solution can be read directly from the tableau: $x_3 = 6100$, $x_4 = 8000$ and $x_5 = 5000$, giving $z = 0$.

Step 3 *Test for optimality: are all the coefficients in the z-row non-negative?*

If so then stop, as the optimal solution has been reached. Otherwise go to **Step 4**.

The bottom row says $z - 2x_1 - 3x_2 = 0$, so z can be increased by increasing x_1 or x_2.

Step 4 *Choose the variable to enter the basis: the* **entering variable**.
This will be increased from its current value of 0.

$z = 2x_1 + 3x_2$, so we should be able to increase z the most if we increase x_2 from 0. (The rate of change of z is greater with respect to x_2; this is called a *steepest ascent* method.) Thus we choose the column with the *most negative* entry in the z-row. Suppose this is in the x_j column. Then x_j is the **entering variable** and its column is the **pivot column**.

Here the entering variable is x_2, as this has the most negative entry (-3) in the z-row.

Step 5 *Choose the variable to leave the basis: the* **departing variable**.
This will be decreased to 0 from its current value.

The x_j column now needs to become 'basic', so we must divide through some row i by a_{ij} to make the entry in the pivot column 1. To keep the right-hand side non-negative we need $a_{ij} > 0$. If there are no positive entries in the pivot column then stop – the problem is unbounded and has no solution. Otherwise a multiple of row i must be added to every other row k to make the entries in the pivot column zero. The operations on rows i and k must be:

$$\left(\begin{array}{ccc|c} \cdots & a_{ij} & \cdots & b_i \\ & & & \\ \cdots & a_{kj} & \cdots & b_k \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} \cdots & 1 & \cdots & \frac{b_i}{a_{ij}} \\ & & & \\ \cdots & 0 & \cdots & b_k - \frac{a_{kj}}{a_{ij}} b_i \end{array}\right) \quad \begin{array}{l} R_i := \frac{1}{a_{ij}} R_i \\ \\ R_k := R_k - \frac{a_{kj}}{a_{ij}} R_i \text{ for all } k \neq i \end{array}$$

To keep all the right-hand sides non-negative requires $b_k - \dfrac{a_{kj}}{a_{ij}} b_i \geq 0$ for all k. Since $a_{ij} > 0$, this certainly holds if $a_{kj} \leq 0$.

If $a_{kj} > 0$ then the condition gives $\dfrac{b_i}{a_{ij}} \leq \dfrac{b_k}{a_{kj}}$ for all k, so row i has to be the row in which

$a_{ij} > 0$ and the **row quotient** $\theta_i = \dfrac{b_i}{a_{ij}}$ is minimum. This row is the **pivot row**.

The element a_{ij} in the selected row and column is called the **pivot element**. The basic variable corresponding to this row is the **departing variable**. It becomes 0 at this iteration.

Step 6 *Form a new tableau.*

As described above, divide the entire pivot row by the pivot element to obtain a 1 in the pivot position. Make every other entry in the pivot column zero, including the entry in the z-row, by carrying out row operations in which multiples of the *pivot row only* are added to or subtracted from the other rows. Then go to **Step 3**.

Applying this procedure to the rose-growing problem, we have:

Initial tableau

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution	θ_i	
x_3	0	5	4	1	0	0	6100	$6100 \div 4 = 1525$	
x_4	0	8	2	0	1	0	8000	$8000 \div 2 = 4000$	
x_5	0	1	5	0	0	1	5000	$5000 \div 5 = 1000$	(smallest)
z	1	-2	-3	0	0	0	0		

The entering variable is x_2, as this has the most negative coefficient in the z-row. Calculating the associated row quotients θ_i shows that x_5 is the departing variable, i.e. the x_5 row is the pivot row for the row operations. The pivot element is 5.

Fill in the next two tableaux:

Second tableau

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution	θ_i
	0							
	0							
	0							
z	1							

Final tableau

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution
	0						
	0						
	0						
z	1						

The coefficients of the non-basic variables in the z-row are both positive. This shows that we have reached the optimum – make sure you can explain why! The algorithm now stops and the optimal values can be read off from the final tableau. $z_{max} = 3700$ when $(x_1, x_2, x_3, x_4, x_5) = (500, 900, 0, 2200, 0)$.

Summary

- We move from one basic feasible solution to the next by taking one variable out of the basis and bringing one variable into the basis.
- The entering variable is chosen by looking at the numbers in the z-row of the current tableau. If they are all non-negative then we already have the optimum value. Otherwise we choose the non-basic column with the *most negative number* in the z-row of the current tableau.

- The departing variable is determined (usually uniquely) by finding the basic variable which will be the first to reach zero as the entering variable is increased. This is identified by finding a row with a positive entry in the pivot column which gives the *smallest non-negative value* of the row quotient $\theta_i = \dfrac{b_i}{a_{ij}}$.

The version of the algorithm described here can be used only on problems which are in standard form. It relies on having an initial basic feasible solution. This is easy to find when all the constraints are of the '$\le$' type. However, in some cases such as the cattle feed problem this is not the case and a modification of the method is needed. We shall return to this in a later section when we consider the dual simplex algorithm.

2.2.1 Further examples

1. Minimize $-2x_1 - 4x_2 + 5x_3$, subject to

$x_1 + 2x_2 + x_3 \le 5, \quad 2x_1 + x_2 - 4x_3 \le 6, \quad 3x_1 - 2x_2 \ge -3, \quad x_j \ge 0$ for $j = 1, 2, 3$.

In standard form, the problem is: Maximize $z = 2x_1 + 4x_2 - 5x_3$, subject to

$x_1 + 2x_2 + x_3 + x_4 = 5, \quad 2x_1 + x_2 - 4x_3 + x_5 = 6, \quad -3x_1 + 2x_2 + x_6 = 3$, all $x_j \ge 0$.

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	Solution
	0							
	0							
	0							
z	1							

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	Solution
	0							
	0							
	0							
z	1							

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	Solution
	0							
	0							
	0							
z	1							

From the last tableau $z = 10 - 7x_3 - 2x_4 - 0x_6$, so any increase in a non-basic variable (x_3, x_4 or x_6) would decrease z. Hence this tableau is optimal. z has a maximum value of 10, so the minimum of the given function is -10 when $x_1 = 0.5, x_2 = 2.25, x_3 = 0$.

$x_5 = 2.75$ shows that strict inequality holds in the second constraint: $2x_1 + x_2 - 4x_3$ falls short of 6 by 2.75. The other constraints are active (binding) at the optimum.

2. Maximize $z = 4x_1 + x_2 + 3x_3$, subject to

$x_1 + x_2 \le 6, \quad -x_2 + 2x_3 \le 4, \quad x_1 \le 4, \quad x_2 \le 4, \quad x_j \ge 0$ for $j = 1, 2, 3$.

In standard form, the constraints become:

$x_1 + x_2 + x_4 = 6, \quad -x_2 + 2x_3 + x_5 = 4, \quad x_1 + x_6 = 4, \quad x_2 + x_7 = 4.$

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	x_7	Solution
	0								
	0								
	0								
	0								
z	1								

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	x_7	Solution
	0								
	0								
	0								
	0								
z	1								

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	x_7	Solution
	0								
	0								
	0								
	0								
z	1								

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	x_7	Solution
	0								
	0								
	0								
	0								
z	1								

The entries in the objective row are all positive; this row reads $z + \frac{5}{2}x_4 + \frac{3}{2}x_5 + \frac{3}{2}x_6 = 27$. (Note that z is always expressed in terms of the *non-basic* variables in each tableau.) Thus z has a maximum value of 27 when $x_4 = x_5 = x_6 = 0$ and $x_1 = 4, x_2 = 2, x_3 = 3, x_7 = 2$. Only one slack variable is basic in the optimal solution, so the at this optimal point the first three constraints are active (they hold as equalities) while the fourth inequality $x_2 \leq 4$ is inactive. Indeed, x_2 is less than 4 by precisely 2, which is the amount of slack in this constraint.

From the final tableau, each basic variable can be expressed in terms of the non-basic variables, e.g. the x_3 row says $x_3 + \frac{1}{2}x_4 + \frac{1}{2}x_5 - \frac{1}{2}x_6 = 3$, so $x_3 = 3 - \frac{1}{2}x_4 - \frac{1}{2}x_5 + \frac{1}{2}x_6$.

2.3 Degeneracy

When one (or more) of the *basic* variables in a basic feasible solution is zero, both the problem and that bfs are said to be **degenerate**.

In a degenerate problem, the same bfs may correspond to more than one basis. This will occur when two rows have the same minimum quotient θ_i. Selecting either of the associated variables to become non-basic results in the one not chosen, which therefore remains basic, becoming equal to zero at the next iteration.

Degeneracy reveals that the LP problem has at least one redundant constraint. The problem of degeneracy is easily dealt with in practice: just keep going! If two or more of the row quotients θ_i are equal, pick any one of them and proceed to the next tableau. There will then be a 0 in the 'Solution' column. Follow the usual rules: the minimum θ_i may be 0, so long as $a_{ij} > 0$. Even if the value of z does not increase, a new basis has been found and we should eventually get to the optimum.

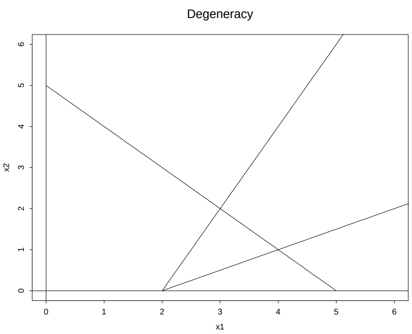
Example: Consider the following problem:

$$\textbf{Maximize} \qquad z = 3x_1 - x_2$$

$$\textbf{subject to} \quad \begin{cases} 2x_1 - x_2 \leq 4 \\ x_1 - 2x_2 \leq 2 \\ x_1 + x_2 \leq 5 \end{cases}$$

$$\textbf{and} \qquad x_j \geq 0, j = 1, 2.$$

Graphically, we see that the optimum occurs when $x_1 = 3$, $x_2 = 2$. The vertex $(2, 0)$ has three lines passing through it. Since, in 2-dimensions, only two lines are needed to define an extreme point, this point is **overdetermined** and one of the constraints is **redundant**.



In standard form, the problem is:

$$\textbf{Maximize} \qquad z = 3x_1 - x_2$$

$$\textbf{subject to} \quad \begin{cases} 2x_1 - x_2 + x_3 = 4 \\ x_1 - 2x_2 + x_4 = 2 \\ x_1 + x_2 + x_5 = 5 \end{cases}$$

$$\textbf{and} \qquad x_j \geq 0 \text{ for } j = 1, \ldots, 5.$$

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution	θ_i
x_3	0	2	-1	1	0	0	4	2
x_4	0	1	-2	0	1	0	2	2
x_5	0	1	1	0	0	1	5	5
z	1	-3	1	0	0	0	0	

The entering variable must be x_1. The departing variable could be either of x_3 or x_4.

We will arbitrarily choose x_4 to depart. The next tableau is:

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution	θ_i
x_3	0	0	3	1	-2	0	0	0
x_1	0	1	-2	0	1	0	2	
x_5	0	0	3	0	-1	1	3	1
z	1	0	-5	0	3	0	6	

Now x_3 is basic but takes the value $x_3 = 0$, so we are at a degenerate bfs. The algorithm has not finished as there is a negative entry in the z-row. The next iteration gives:

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution	θ_i
x_2	0	0	1	1/3	−2/3	0	0	
x_1	0	1	0	2/3	−1/3	0	2	
x_5	0	0	0	−1	1	1	3	3
z	1	0	0	5/3	−1/3	0	6	

This solution is degenerate again, and the objective function has not increased. There is still a negative entry in the z-row. Pivoting on the x_5 row gives:

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution
x_2	0	0	1	−1/3	0	2/3	2
x_1	0	1	0	1/3	0	1/3	3
x_4	0	0	0	−1	1	1	3
z	1	0	0	4/3	0	1/3	7

The algorithm has now terminated. $z_{\max} = 7$ when $x_1 = 3, x_2 = 2$.

Both the second and third tableaux represent the point $(2, 0, 0, 0, 3)$. The only difference is that the decision variables are classified differently as basic and nonbasic at the two stages.

In a degenerate problem, it is possible (though unlikely) that the Simplex algorithm could return at some iteration to a previous tableau. Once caught in this **cycle**, it will go round and round without improving z.

Several methods have been proposed for preventing cycling. One of the simplest is the 'smallest subscript rule' or 'Bland's rule', which states that the simplex method will not cycle provided that whenever there is more than one candidate for the entering or leaving variable, the variable with the smallest subscript is chosen (e.g. x_3 in preference to x_4).

2.4 Theory of the Simplex method

We now investigate the workings of the Simplex algorithm in more detail.

The initial and final tableaux for the rose growing problem were:

Initial tableau

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution
x_3	0	5	4	1	0	0	6100
x_4	0	8	2	0	1	0	8000
x_5	0	1	5	0	0	1	5000
z	1	−2	−3	0	0	0	0

Final tableau

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution
x_1	0	1	0	5/21	0	−4/21	500
x_4	0	0	0	−38/21	1	22/21	2200
x_2	0	0	1	−1/21	0	5/21	900
z	1	0	0	1/3	0	1/3	3700

In the initial tableau, the columns of the 4×4 identity matrix occur in the columns headed x_3, x_4, x_5, z respectively.

The corresponding columns in the final tableau, in the same order, form the matrix

$$\mathrm{P} = \begin{pmatrix} 5/21 & 0 & -4/21 & 0 \\ -38/21 & 1 & 22/21 & 0 \\ -1/21 & 0 & 5/21 & 0 \\ 1/3 & 0 & 1/3 & 1 \end{pmatrix}.$$

The *entire initial tableau*, regarded as a matrix, has been pre-multiplied by P to give the entire final tableau.

P tells us the row operations that would convert the initial tableau directly to the final tableau, e.g. $R_2 := -\frac{38}{21}R_1 + 1R_2 + \frac{22}{21}R_3 + 0R_4$.

In the final tableau, the columns of the 4×4 identity matrix occur under x_1, x_4, x_2 and z respectively. The corresponding columns of the *initial* tableau form the matrix

$\mathrm{P}^{-1} = \begin{pmatrix} 5 & 0 & 4 & 0 \\ 8 & 1 & 2 & 0 \\ 1 & 0 & 5 & 0 \\ -2 & 1 & -3 & 1 \end{pmatrix}$. This would pre-multiply the entire final tableau to give the initial tableau.

To generalize this, suppose we are maximizing $z = \mathbf{c}^t\mathbf{x}$ subject to $\mathrm{A}\mathbf{x} \leq \mathbf{b}$ (where $\mathbf{b} \geq \mathbf{0}$) and $\mathbf{x} \geq \mathbf{0}$.

Let T_1 be the initial simplex tableau, regarded as a matrix (i.e. omit the row and column headings) and let T_f be the final tableau.

T_1 can be written in partitioned form as $\left(\begin{array}{c|c|c||c} \mathbf{0} & \mathrm{A} & \mathrm{I} & \mathbf{b} \\ \hline 1 & -\mathbf{c}^t & \mathbf{0}^t & 0 \end{array}\right)$.

The final tableau is obtained from the initial one by a succesion of row operations, each of which could be achieved by pre-multiplying T_1 by some matrix. Hence $\mathrm{T}_f = \mathrm{PT}_1$ for some matrix P which can be partitioned in the form $\left(\begin{array}{c|c} \mathrm{M} & \mathbf{u} \\ \hline \mathbf{y}^t & \alpha \end{array}\right)$, say. As the z-column is unchanged, $\mathbf{u} = \mathbf{0}$ and $\alpha = 1$.

The process of transforming the initial tableau to the optimal one can be expressed as:

$$\left(\begin{array}{c|c} \mathrm{M} & \mathbf{0} \\ \hline \mathbf{y}^t & 1 \end{array}\right)\left(\begin{array}{c|c|c||c} \mathbf{0} & \mathrm{A} & \mathrm{I} & \mathbf{b} \\ \hline 1 & -\mathbf{c}^t & \mathbf{0}^t & 0 \end{array}\right) = \left(\begin{array}{c|c|c||c} \mathbf{0} & \mathrm{MA} & \mathrm{M} & \mathrm{M}\mathbf{b} \\ \hline 1 & \mathbf{y}^t\mathrm{A} - \mathbf{c}^t & \mathbf{y}^t & \mathbf{y}^t\mathbf{b} \end{array}\right).$$

Thus M and $\mathbf{y}^t$ are found in the columns headed by the slack variables in the final tableau.

As T_f is optimal we must have $\mathbf{y} \geq \mathbf{0}$ and $\mathbf{y}^t\mathrm{A} - \mathbf{c}^t \geq \mathbf{0}$, i.e. $\mathrm{A}^t\mathbf{y} \geq \mathbf{c}$.

Then $z_{\max} = \mathbf{y}^t\mathbf{b}$, or equivalently $z_{\max} = \mathbf{b}^t\mathbf{y}$.

We shall meet these conditions again when we study the *dual* of a LP problem.

Now suppose that in the optimal tableau the basic variables, reading *down* the list in the

'Basic' column, are $x_{B1}, \ldots, x_{Bm}$ and the non-basic variables are $x_{N1}, \ldots, x_{Nn}$. The *initial* tableau could be rearranged as follows:

		Optimal basic variables	Optimal non-basic variables	
Basic	z	$\overbrace{x_{B1} \cdots x_{Bm}}$	$\overbrace{x_{N1} \cdots x_{Nn}}$	Solution
x_{n+1}	0			
$\vdots$	$\vdots$	B	N	$\mathbf{b}$
x_{n+m}	0			
z	1	$-c_{B1} \cdots -c_{Bm}$	$-c_{N1} \cdots -c_{Nn}$	0

Thus the objective function can be written as

$z = c_{B1}x_{B1} + \cdots + c_{Bm}x_{Bm} + c_{N1}x_{N1} + \cdots + c_{Nn}x_{Nn}$.

Let $\mathbf{c}_B = (c_{B1} \ \cdots \ c_{Bm})^t$, $\mathbf{c}_N = (c_{N1} \ \cdots \ c_{Nn})^t$, $\mathbf{x}_B = (x_{B1} \ \cdots \ x_{Bm})^t$, $\mathbf{x}_N = (x_{N1} \ \cdots \ x_{Nn})^t$.

The process of transforming T_1 (with the columns reordered as described) to T_f is now:

$$\left(\begin{array}{c|c} \mathrm{M} & \mathbf{0} \\ \hline \mathbf{y}^t & 1 \end{array}\right)\left(\begin{array}{c|c|c||c} \mathbf{0} & \mathrm{B} & \mathrm{N} & \mathbf{b} \\ \hline 1 & -{\mathbf{c}_B}^t & -{\mathbf{c}_N}^t & 0 \end{array}\right) = \left(\begin{array}{c|c|c||c} \mathbf{0} & \mathrm{MB} & \mathrm{MN} & \mathrm{M}\mathbf{b} \\ \hline 1 & \mathbf{y}^t\mathrm{B} - {\mathbf{c}_B}^t & \mathbf{y}^t\mathrm{N} - {\mathbf{c}_N}^t & \mathbf{y}^t\mathbf{b} \end{array}\right).$$

The basic columns of the final tableau contain an identity matrix and a row of zeros so $\mathrm{MB} = \mathrm{I}$, i.e. $\mathrm{M} = \mathrm{B}^{-1}$. Thus $\mathbf{x}_B = \mathrm{B}^{-1}\mathbf{b}$.

Also $\mathbf{y}^t\mathrm{B} - {\mathbf{c}_B}^t = \mathbf{0}$ which can be written as $\mathbf{y}^t = {\mathbf{c}_B}^t\mathrm{B}^{-1}$.

Every entry in the z-row is non-negative at the optimum, so $\mathbf{y}^t\mathrm{N} - {\mathbf{c}_N}^t \geq \mathbf{0}$,

i.e. ${\mathbf{c}_B}^t\mathrm{B}^{-1}\mathrm{N} - {\mathbf{c}_N}^t \geq \mathbf{0}$. Then $z_{\max} = \mathbf{y}^t\mathbf{b} = {\mathbf{c}_B}^t\mathrm{B}^{-1}\mathbf{b}$. Thus we have:

Proposition 2.1 *Let $x_{B1}, \ldots, x_{Bm}$ be the basic variables in the optimal tableau for problem (LP1) with $\mathbf{b} \geq \mathbf{0}$. Let B be the matrix formed by the columns headed $x_{B1}, \ldots, x_{Bm}$ in the initial tableau (excluding the z-row). Then $(x_{B1} \ \cdots \ x_{Bm})^t = \mathrm{B}^{-1}\mathbf{b}$ and $z_{\max} = {\mathbf{c}_B}^t\mathrm{B}^{-1}\mathbf{b}$.*

2.4.1 Example

Consider the LP problem: Maximize $z = 2x_1 + 4x_2 - 5x_3$, subject to

$x_1 + 2x_2 + x_3 \leq 5, \quad 2x_1 + x_2 - 4x_3 \leq 6, \quad -3x_1 + 2x_2 \leq 3.$

The initial and final tableaux are respectively:

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	Solution
x_4	0	1	2	1	1	0	0	5
x_5	0	2	1	−4	0	1	0	6
x_6	0	−3	2	0	0	0	1	3
z	1	−2	−4	5	0	0	0	0

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	Solution
x_1	0	1	0	1/4	1/4	0	−1/4	1/2
x_5	0	0	0	−39/8	−7/8	1	3/8	11/4
x_2	0	0	1	3/8	3/8	0	1/8	9/4
z	1	0	0	7	2	0	0	10

We now illustrate the theory set out in the last section.

We have been assuming that only extreme points of the feasible region need to be considered as possible solutions. The proof of this is now given; you are not expected to learn it.

Proposition 2.2 $\mathbf{v}$ *is a basic feasible solution of (LP2) if and only if* $\mathbf{v}$ *is an extreme point of the set of feasible solutions* $S = \{\widetilde{\mathbf{x}} \ : \ \widetilde{\mathrm{A}}\widetilde{\mathbf{x}} = \mathbf{b},\, \widetilde{\mathbf{x}} \geq \mathbf{0}\}$ *in* $\mathbb{R}^{n+m}$.

Proof ($\Rightarrow$) Suppose $\mathbf{v}$ is a bfs of (LP2) and is *not* an extreme point of S, so there are distinct feasible solutions $\mathbf{u}$, $\mathbf{w} \in S$ such that, for some $r \in (0,1), \mathbf{v} = (1-r)\mathbf{u} + r\mathbf{w}$.

We can permute the columns of $\widetilde{\mathrm{A}}$, as described earlier, to get $(\mathrm{B} \mid \mathrm{N})$ and permute the entries of $\mathbf{v}$, $\mathbf{u}$, $\mathbf{w}$ correspondingly so that $\left(\dfrac{\mathbf{v}_B}{\mathbf{0}}\right) = (1-r)\left(\dfrac{\mathbf{u}_B}{\mathbf{u}_N}\right) + r\left(\dfrac{\mathbf{w}_B}{\mathbf{w}_N}\right)$

As $0 < r < 1$ and $\mathbf{u}, \mathbf{w}$ are non-negative, it follows that $\mathbf{u}_N = \mathbf{w}_N = \mathbf{0}$.

As $\mathbf{v}$, $\mathbf{u}$, $\mathbf{w}$ are feasible, $\widetilde{\mathrm{A}}\mathbf{v} = \widetilde{\mathrm{A}}\mathbf{u} = \widetilde{\mathrm{A}}\mathbf{w} = \mathbf{b}$,

so $(\mathrm{B} \mid \mathrm{N})\left(\dfrac{\mathbf{v}_B}{\mathbf{0}}\right) = (\mathrm{B} \mid \mathrm{N})\left(\dfrac{\mathbf{u}_B}{\mathbf{0}}\right) = (\mathrm{B} \mid \mathrm{N})\left(\dfrac{\mathbf{w}_B}{\mathbf{0}}\right) = \mathbf{b}$.

Thus $\mathrm{B}\mathbf{v}_B = \mathrm{B}\mathbf{u}_B = \mathrm{B}\mathbf{w}_B = \mathbf{b}$. As B is non-singular, $\mathbf{v}_B = \mathbf{u}_B = \mathbf{w}_B = \mathrm{B}^{-1}\mathbf{b}$.

Hence $\mathbf{v} = \mathbf{u} = \mathbf{w}$, contradicting the assumption that these points are distinct, so $\mathbf{v}$ is an extreme point.

($\Leftarrow$) Suppose $\mathbf{v}$ is in S, so is feasible, but is not a bfs of (LP2).

$\widetilde{\mathrm{A}}\mathbf{v} = \mathbf{b}$, so we can write $(\mathrm{B}' \mid \mathrm{N}')\left(\dfrac{\mathbf{v}'}{\mathbf{0}}\right) = \mathbf{b}$, where $\mathbf{v}'$ contains all the non-zero entries of $\mathbf{v}$. If $\mathbf{v}' \neq \mathbf{v}$ then B' is not unique, but the following reasoning applies to any choice of B'.

If B' were non-singular then $\mathbf{v}' = \mathrm{B}^{-1}\mathbf{b}$, in which case $\mathbf{v}$ would be a bfs.

Hence B' is singular, so there is a vector $\mathbf{p}$ such that $\mathrm{B}'\mathbf{p} = \mathbf{0}$.

As the entries of $\mathbf{v}'$ are strictly positive, we can find $\varepsilon > 0$ so that $\mathbf{v}' - \varepsilon\mathbf{p} \geq \mathbf{0}$ and $\mathbf{v}' + \varepsilon\mathbf{p} \geq \mathbf{0}$. Then $\mathrm{B}'(\mathbf{v}' \pm \varepsilon\mathbf{p}) = \mathrm{B}'\mathbf{v}' \pm \varepsilon\mathrm{A}'\mathbf{p} = \mathrm{B}'\mathbf{v}' \pm \mathbf{0} = \mathbf{b}$, so $(\mathrm{B}' \mid \mathrm{N}')\left(\dfrac{\mathbf{v}' \pm \epsilon\mathbf{p}}{\mathbf{0}}\right) = \mathbf{b}$.

Thus both of $\left(\dfrac{\mathbf{v}' \pm \varepsilon\mathbf{p}}{\mathbf{0}}\right)$ are permuted feasible solution vectors. But $\left(\dfrac{\mathbf{v}'}{\mathbf{0}}\right)$ is their mean, so $\mathbf{v}$ is a convex linear combination of two points in S, hence $\mathbf{v}$ is not an extreme point of S. It follows that any extreme point of S *is* a bfs of (LP2). $\square$

Exercises 2

1. Use the Simplex method to solve

$$\textbf{Maximize} \qquad z = 4x_1 + 8x_2$$

$$\textbf{subject to} \quad \begin{cases} 5x_1 + x_2 \leq 8 \\ 3x_1 + 2x_2 \leq 4 \end{cases}$$

$$\textbf{and} \qquad x_1 \geq 0, x_2 \geq 0.$$

   Express the objective function in terms of the non-basic variables in the final tableau, and hence explain how you know that the solution is optimal.

2. Use the Simplex method to solve

$$\textbf{Maximize} \qquad z = 5x_1 + 4x_2 + 3x_3$$

$$\textbf{subject to} \quad \begin{cases} 2x_1 + 3x_2 + x_3 \leq 5 \\ 4x_1 + x_2 + 2x_3 \leq 11 \\ 3x_1 + 4x_2 + 2x_3 \leq 8 \end{cases}$$

$$\textbf{and} \qquad x_j \geq 0 \text{ for } j = 1, 2, 3.$$

   Write down a matrix P which would pre-multiply the whole initial tableau to give the final tableau.

   After you have studied Section 2.4 in the notes: write down, from the optimal tableau of this problem, the matrices B, B^{-1} and N and the vectors $\mathbf{y}$, $\mathbf{c}_B$, $\mathbf{c}_N$ as defined on pages 24 - 25. Verify that the optimal solution is $\mathrm{B}^{-1}\mathbf{b}$ and that $z_{\max} = \mathbf{c}_B{}^t\mathrm{B}^{-1}\mathbf{b}$.

3. Express the following LP problem in standard form. (Do not solve it.)

$$\textbf{Minimize} \qquad z = 2x_1 + 3x_2$$

$$\textbf{subject to} \quad \begin{cases} x_1 + x_2 \geq -1 \\ 2x_1 + 3x_2 \leq 5 \\ 3x_1 - 2x_2 \leq 3 \end{cases}$$

$$\textbf{and} \quad x_1 \text{ unrestricted in sign, } x_2 \geq 0.$$

4. The following tableau arose in the solution of a LP problem:

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	x_7	Solution
	0	0	4/3	2/3	0	1	0	−1/3	4
	0	0	1/3	2/3	1	0	1	−1/3	10
	0	1	−1/3	1/6	1/2	0	0	1/6	4
z	1	0	−5/3	−4/3	−1	0	0	5/3	12

   (a) Which variables are basic? What basic feasible solution does the tableau represent? Is it optimal? Explain your answer by writing the z-row as an equation.

   (b) Starting from the given tableau, proceed to find an optimal solution.

   (c) From the optimal tableau, express the objective function and each of the basic variables in terms of the non-basic variables.

5. Suppose one of the constraints in a LP problem is an equation rather than an inequality, as in

$$\textbf{Minimize} \quad 3x_1 + 2x_2 + 4x_3$$

$$\textbf{subject to} \quad \begin{cases} 3x_1 + 4x_2 + 2x_3 \le 6 \\ 3x_1 + 5x_2 + 7x_3 \le 10 \\ x_1 + x_2 + x_3 = 1 \end{cases}$$

$$\textbf{and} \quad x_1 \ge 0, x_2 \ge 0.$$

Use the third constraint to express x_3 in terms of x_1 and x_2. Hence eliminate x_3 from the objective function and the other two constraints. Introduce two slack variables x_4 and x_5, and solve the problem by the Simplex algorithm. Check that your solution satisfies all the constraints. How must this approach be modified if also $x_3 \ge 0$?

6. A furniture manufacturer makes chairs and settees, producing up to 80 chairs and 48 settees per week. The items are sold in suites: Mini - two chairs, Family - three chairs and one settee, Grand - three chairs and two settees. The profits are £20, £30 and £70 per suite, respectively. The total profit is to be maximized. Use the Simplex algorithm to find the maximum profit that can be made in one week. State the profit and the number of each type of suite that should be made.

7. Solve the following problem using the simplex tableau method, **making your decision variable for I_1 basic at the first iteration**. Identify any degenerate basic feasible solutions your calculations produce. Sketch the solution space and explain why degeneracy occurs.

   A foundry produces two kinds of iron, I_1 and I_2, by using three raw materials R_1, R_2 and R_3. Maximise the daily profit.

Raw material	Amount required per tonne of I_1	of I_2	Daily raw material availability (tonnes)
R_1	2	1	16
R_2	1	1	8
R_3	0	1	3.5
Profit per tonne	£150	£300	

8. Solve the following problem by the simplex algorithm, showing that every possible basic feasible solution occurs in the iterations.

$$\textbf{Maximize} \quad z = 100x_1 + 10x_2 + x_3$$

$$\textbf{subject to} \quad \begin{cases} x_1 \le 1 \\ 20x_1 + x_2 \le 100 \\ 200x_1 + 20x_2 + x_3 \le 10000 \end{cases}$$

$$\textbf{and} \quad x_1 \ge 0, x_2 \ge 0, x_3 \ge 0.$$

9. Consider the Linear Programming problem:

$$\textbf{Maximize} \qquad z = x_1 + x_2 + x_3$$

$$\textbf{subject to} \quad \begin{cases} x_1 + 3x_2 - x_3 \leq 10 \\ x_1 - x_2 + 3x_3 \leq 14 \end{cases}$$

$$\textbf{and} \qquad x_1 \geq 0, x_2 \geq 0, x_3 \geq 0.$$

Find a 3×3 matrix which would pre-multiply the initial tableau to give a tableau in which the basic variables are x_1 and x_3. Verify that this tableau is optimal, and hence state the solution of the problem.

If you had not been told which variables to take as basic, how many basic feasible solutions might you have to consider in order to solve the problem this way?

10. Prove that $\mathbf{v}$ is an extreme point of the feasible region R for (LP1) if and only if $\left(\frac{\mathbf{v}}{\mathbf{b} - A\mathbf{v}}\right)$ is an extreme point of the feasible region S for (LP2).

Chapter 3

Further Simplex Methodology

3.1 Sensitivity analysis

Suppose some feature of a LP problem changes after we have found the optimal solution. Do we have to solve the problem again from scratch or can the optimum of the original problem be used as an aid to solving the new problem? These questions are addressed by **sensitivity analysis**, also called **post-optimal analysis**.

If a resource is used up completely then the slack variable in the constraint representing this resource is zero, i.e. the constraint is active at the optimal solution. This type of resource is called **scarce**. In contrast, if the slack variable is non-zero (i.e. the constraint is not active) then this resource is not used up totally in the optimal solution. Resources of this kind are called **abundant**. Increasing the availability of an abundant resource will not in itself yield an improvement in the optimal solution. However, increasing a scarce resource *will* improve the optimal solution.

With the notation of Chapter 2, $z_{\max} = \mathbf{y}^t\mathbf{b}$. Suppose b_i is increased by a small amount δb_i. We say that the ith constraint has been **relaxed** by δb_i. Then as long as the same variables remain basic at the optimal solution, $z_{\max} = y_1b_1 + \cdots + y_i(b_i + \delta b_i) + \cdots + y_mb_m$, i.e. $z_{\max}$ has increased by $y_i\delta b_i$.

y_i is thus the approximate increase in the optimum value of z that results from allowing an extra 1 unit of the ith resource. y_i is called the **shadow price** of the ith resource. It does *not* tell us by how much the resource can be increased while maintaining the same rate of improvement. If the set of constraints which are active at the optimal solution changes, then the shadow price is no longer applicable.

The shadow prices y_i can be read directly from the optimal tableau: they are the numbers in the z-row at the bottom of the slack variable columns.

3.1.1 Example: the rose-growing problem modified

To illustrate these concepts we again consider the rose-growing problem, with the initial and final tableaux as obtained in Chapter 2.

The optimal solution is $(500, 900, 0, 2200, 0)$. The slack variables x_3, x_4 and x_5 were introduced in the land, finance and labour resource constraints respectively. Thus land and labour are scarce resources, and finance is an abundant resource.

We now consider various modifications to the problem and its solution. Most of the methods use the multiplying matrix P which was defined in Section 2.4.

1. *Shadow prices*

2. *Changing the resources*
   The rose grower has the option to buy some more land. What is the maximum area that should be purchased if the other constraints remain the same?

3. *Changing a coefficient in the objective function*
   The profit margin alters on one of the roses. Is the same solution still optimal?

4. *Adding an extra variable.*

Suppose the grower has the option of growing an extra variety, pink roses, with the following requirements per bush: 7 dm^2 area, £6 finance, and 3.5 hours labour. Let the profit per pink bush be £t, and suppose that x_6 pink rose bushes are grown.

5. *Adding an extra constraint.*

Suppose the maximum number of roses that can be sold is 1340, so the constraint $x_1 + x_2 \le 1340$ is added to the system. The current optimum is not feasible as $x_1 + x_2 = 500 + 900 > 1340$.

Adding a slack variable x_6 gives $x_1 + x_2 + x_6 = 1340$.

The above procedure, keeping the z-row coefficients positive and carrying out row operations until the basic variables are non-negative, is called the **dual simplex method** . It can be used, on its own or in combination with the normal simplex method, on problems which are not quite in standard form because some of the right-hand sides are negative.

The dual simplex method provides one way of solving problems for which no initial basic feasible solution is apparent. There are other variations on the simplex method which can be used for such purposes, such as the Two-Phase method (see a textbook).

3.1.2 Example: the garments problem

A company makes three types of garment. The constraints are given by:

Type	A	B	C	
No. of units	x_1	x_2	x_3	Amount available
Labour (hours) per unit	1	2	3	55 hours
Material (m^2) per unit	3	1	4	80 m^2
Profit (£) per unit	7	6	9	

The company wants to make the largest possible profit subject to the constraints on labour and materials, so they must .... $z =$

subject to

and $x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$.

Adding slack variables, the constraints become:

Solve by the Simplex algorithm:

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution
z							

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution
z							

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution
z							

Basic	z	x_1	x_2	x_3	x_4	x_5	Solution
z							

so the maximum profit is £
when of type A, of type B and of type C are made.

From the initial to the final tableau, the middle two rows are multiplied by

$\mathrm{B}^{-1} = \Bigg(\qquad \Bigg)$. This is the inverse of $\mathrm{B} = \Bigg(\qquad \Bigg)$

(using the columns for , respectively in the initial tableau.)

Taking $\mathbf{c}_B = \Bigg(\quad \Bigg)$, $\mathbf{b} = \Bigg(\quad \Bigg)$,

$\mathbf{c}_B{}^t\mathrm{B}^{-1}\mathbf{b} = (\qquad) \Bigg(\quad \Bigg) = \qquad = z_{\max}$.

We now consider the effect of changing some features of the problem.

1. *Shadow prices*

   The shadow prices of labour and material are and respectively.

   Both resources are ....

   The maximum profit increases at the rate of £ per extra hour of labour and £ per extra m^2 of material, within certain limits.

2. *Changing the resources*

   Suppose p hours of labour and q m^2 of material are available.

   The 'solution' column in the initial tableau is $\Bigg(\quad \Bigg)$ so in the final tableau it is

   $$\Bigg(\qquad \Bigg) \Bigg(\quad \Bigg) = \Bigg(\qquad \Bigg).$$

   For the solution to remain optimal when $x_1 > 0, x_2 > 0, x_3 = 0$ we need

   ≥ 0 and ≥ 0, so

   $q \leq$ and $p \leq$. Thus $\leq q \leq$

   If p is kept at 55, this gives $\leq q \leq$

   Provided the amount of material is within this range, $z_{\max} =$

   when $x_1 =$, $x_2 =$, $x_3 =$

   When $q =$ then $z_{\max} =$

3. *Changing a coefficient in the objective function*

   Suppose the profit on type A changes by £t per unit. This does not alter the feasible region, but it may affect the optimal solution.

   The new objective function is $z' = (\qquad)x_1 + 6x_2 + 9x_3$.

   Thus the bottom row of the initial tableau is (1 | .... -6 -9 0 0 || 0), so in the final tableau it is (1 | .... 0 4 11/5 8/5 || 249).

   For x_1 to remain basic we must have 0 at the bottom of the x_1 column, so add .... times Row .... to Row .... to get

(1 | .... || ....).

The original solution $(21, 17, 0, 0, 0)$ is still optimal if no z-row entry is negative.

Therefore the original solution is optimal, with $z'_{\max} =$, provided

$\geq 0,$ $\geq 0,$ ≥ 0, i.e. $\leq t \leq$

The profit on type A can vary from £ to £ without affecting the optimal quantities to produce.

4. *Adding an extra variable.*

Suppose the company can produce a fourth garment, type D, requiring 4 hours of labour and 2 m^2 of material and yielding £k profit per unit.

Let x_0 units of type D be produced.

The new problem is to maximize $z =$

subject to

$= 55,$ $= 80,$

where $x_j \geq 0$ for $j = 0, \ldots, 5$.

The initial tableau is as before with the addition of a column

x_0

$$\begin{pmatrix} 3/5 & -1/5 & 0 \\ -1/5 & 2/5 & 0 \\ 11/5 & 8/5 & 1 \end{pmatrix} \begin{pmatrix} \\ \\ \\ \end{pmatrix} = \begin{pmatrix} \\ \\ \\ \end{pmatrix}$$

, so we have

x_0

in the final tableau.

Thus if $k \leq$, the z-row remains non-negative and the current solution is optimal; D should not be made.

Suppose we take $k = 13$. Then the previously optimal tableau becomes

Basic	z	x_0	x_1	x_2	x_3	x_4	x_5	Solution
x_2	0		0	1	1	3/5	−1/5	17
x_1	0		1	0	1	−1/5	2/5	21
z	1		0	0	4	11/5	8/5	249

After one more iteration this becomes optimal:

Basic	z	x_0	x_1	x_2	x_3	x_4	x_5	Solution
	0							
	0							
z	1							

so units of A and units of D should be made.

(If half-units are not possible, Integer Programming is needed.)

5. *Adding an extra constraint.*

Suppose that in the original 3-product problem there is a further restriction:
$3x_1 + x_2 + 2x_3 \leq 50$, i.e. $3x_1 + x_2 + 2x_3 + x_6 = 50$ where $x_6 \geq 0$ is a slack variable.

$3(21) + 17 + 2(0) > 50$, so the previous optimal point $(21, 17, 0)$ is not feasible.

Adding the new constraint as a row in the original optimal tableau gives:

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	Solution
	0	0	1	1	3/5	−1/5	0	17
	0	1	0	1	−1/5	2/5	0	21
x_6	0							
z	1	0	0	4	11/5	8/5	0	249

This does not represent a valid Simplex iteration because there is only one basic column. Subtracting Row 1 and 3 times Row 2 from Row 3 gives:

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	Solution
x_2	0	0	1	1	3/5	−1/5	0	17
x_1	0	1	0	1	−1/5	2/5	0	21
x_6	0							
z	1	0	0	4	11/5	8/5	0	249

This represents a basic but *infeasible* solution. It would be optimal if the negative entry in the solution column were not there. In this situation we use the **dual simplex method**, as described on page 40.

$x_6 =$, so x_6 must leave leave the basis.

To keep the z-row entries non-negative, choose the entering variable as follows:

Where there is a negative number in the pivot row, divide the z-row entry by this and choose the variable for which the modulus of this quotient is smallest.

| ÷ | < | ÷ | so enters the basis.

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	Solution
x_2	0							
x_1	0							
	0							
z	1							

Making of type A and of type B is now optimal, giving profit £

3.2 The dual simplex method

The method that we used above when negative numbers occur in the solution column can be used on any such problems, not just in the context of sensitivity analysis.

The differences between the original simplex method and the dual simplex method can be summarised as follows:

Original ('primal') simplex

- Starts from a basic feasible solution.
- All $b_j \geq 0$.
- At least one z-row entry is negative.
- Seek to make all z-row entries ≥ 0

Dual simplex

- Starts from a basic infeasible solution.
- At least one $b_j < 0$.
- All z-row entries ≥ 0.
- Seek to make all $b_j \geq 0$ keeping the z-row $\geq$ 0.

The algorithm is:

1. Express the problem in maximization form with slack variables (*not* in standard form).
2. Find the most negative number in the 'solution' column. Suppose this is in row i.
3. For each *negative* a_{ij} in row i, find the smallest absolute value $\left|\dfrac{c_j}{a_{ij}}\right|$.
4. Use a_{ij} as the pivot in the usual way to obtain a new tableau.
5. Return to step 2. Continue until there are no negative entries in the solution column or the z-row.

3.2.1 Example

Use the dual simplex algorithm to solve the problem:

$$\textbf{Minimize} \qquad z = 5x_1 + 4x_2$$

$$\textbf{subject to} \quad \begin{cases} 3x_1 + 2x_2 \geq 6 \\ x_1 + 2x_2 \geq 4 \end{cases}$$

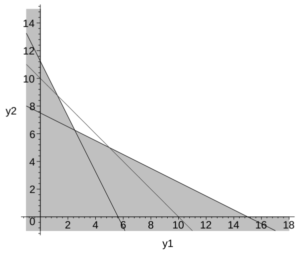
$$\textbf{and} \qquad x_1 \geq 0, x_2 \geq 0.$$

3.3 The dual of a Linear Programming problem

Recall the Containers problem from Chapter 1. Suppose now that the company delegates its production to a contractor who pays them $\pounds y_1$ and $\pounds y_2$ per minute for the use of machines M_1 and M_2 respectively. Let £w be the total hourly charge for using the two machines. The contractor wants to make this hourly charge as small as possible, but must ensure that the company is paid at least as much as it originally made in profit for each container produced: £30 per Type A and £45 per Type B.

Thus the contractor's problem is to minimize $w = 60y_1 + 60y_2$ subject to the constraints $2y_1 + 4y_2 \geq 30, 8y_1 + 4y_2 \geq 45$, where $y_1 \geq 0, y_2 \geq 0$.

The feasible region lies in the first quadrant *above* the boundary lines, as illustrated:



w is minimum where the lines cross, at $(2.5, 6.25)$. Here $w = 60 \times 2.5 + 60 \times 6.25 = 525$.

Thus the contractor should pay £2.50 per minute for M_1 and £6.25 per minute for M_2, so that the company gets £525 per hour – the same as the profit when it made the containers itself! We have solved the **dual** of the original problem.

Every linear programming problem has an associated problem called its **dual**. For now we will restrict attention to pairs of problems of the following form:

Primal: maximize $z = \mathbf{c}^t\mathbf{x}$ subject to $\mathrm{A}\mathbf{x} \leq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$ (P)

Dual: minimize $w = \mathbf{b}^t\mathbf{y}$ subject to $\mathrm{A}^t\mathbf{y} \geq \mathbf{c}$ and $\mathbf{y} \geq \mathbf{0}$ (D)

To obtain the dual problem from the primal problem we swap **c** and **b**, replace A by its transpose A^t, replace '$\leq$' with '$\geq$' in the constraints, and replace 'maximize' with 'minimize'. The non-negativity restrictions remain.

3.3.1 Example

The following is a primal-dual pair of LP problems:

Primal		**Dual**	
Maximize	$z = 6x_1 + 4x_2$	**Minimize**	$w = 5y_1 + 4y_2$
subject to	$\begin{cases} 3x_1 + x_2 \leq 5 \\ 2x_1 + 2x_2 \leq 4 \end{cases}$	**subject to**	$\begin{cases} 3y_1 + 2y_2 \geq 6 \\ y_1 + 2y_2 \geq 4 \end{cases}$
and	$x_1, x_2 \geq 0.$	**and**	$y_1, y_2 \geq 0.$

The primal problem can be solved easily using the standard simplex algorithm. The optimal solution is $z_{\max} = 11$ when $(x_1, x_2, x_3, x_4) = (3/2, 1/2, 0, 0)$.

We solved the dual problem by the dual simplex algorithm in Section 3.2. The optimal solution is $w_{\min} = 11$ when $(y_1, y_2, y_3, y_4) = (1, 3/2, 0, 0)$.

Notice that the objective functions of the primal and dual problems have the same optimum value. Furthermore, in the optimal dual tableau, the objective row coefficients of the slack variables are equal to the optimal primal decision variables.

Consider the cattle-feed problem in Chapter 1. Suppose a chemical company offers the farmer synthetic nutrients at a cost of £y_1 per unit of protein, £y_2 per unit of fat, £y_3 per unit of calcium and £y_4 per unit of phosphorus.

The cost per unit of hay substitute is thus £$(13.2y_1+4.3y_2+0.02y_3+0.04y_4)$. To be economic to the farmer, this must not be more than £0.66. Similarly, considering the oats substitute, $34.0y_1 + 5.9y_2 + 0.09y_3 + 0.09y_4 \le 2.08$.

For feeding one cow, the company will receive £$(65.0y_1 + 14.0y_2 + 0.12y_3 + 0.15y_4)$, which it will wish to maximize.

Thus the company's linear programming problem is :

$$\textbf{Maximize} \qquad w = 65.0y_1 + 14.0y_2 + 0.12y_3 + 0.15y_4$$

$$\textbf{subject to} \quad \begin{cases} 13.2y_1 + 4.3y_2 + 0.02y_3 + 0.04y_4 & \le & 0.66 \\ 34.0y_1 + 5.9y_2 + 0.09y_3 + 0.09y_4 & \le & 2.08 \end{cases}$$

$$\textbf{and} \qquad y_j \ge 0, j = 1, ..., 4.$$

We see that the dual of this is the farmer's original problem. As we shall show next, in fact the two problems are the duals of each other.

Proposition 3.1 *The dual of the dual problem (D) is the primal problem (P).*

Proof The dual problem (D) can be written as follows:

$$\text{maximize } (-\mathbf{b})^t\mathbf{y} \text{ subject to } (-\mathrm{A})^t\mathbf{y} \le -\mathbf{c} \text{ and } \mathbf{y} \ge \mathbf{0}.$$

The dual of this is:

$$\text{minimize } (-\mathbf{c})^t\mathbf{x} \text{ subject to } (-\mathrm{A}^t)^t\mathbf{x} \ge -\mathbf{b} \text{ and } \mathbf{x} \ge \mathbf{0}.$$

This is equivalent to:

$$\text{maximize } \mathbf{c}^t\mathbf{x} \text{ subject to } \mathrm{A}\mathbf{x} \le \mathbf{b} \text{ and } \mathbf{x} \ge \mathbf{0},$$

which is the same as the primal problem (P). □

Thus the dual of either problem may be constructed according to the following rules:

Primal	**Dual**
Maximize $\mathbf{c}^t\mathbf{x}$	Minimize $\mathbf{b}^t\mathbf{y}$
Minimize $\mathbf{c}^t\mathbf{x}$	Maximize $\mathbf{b}^t\mathbf{y}$
Constraints $\mathrm{A}\mathbf{x} \ge \mathbf{b}$	Constraints $\mathrm{A}^t\mathbf{y} \le \mathbf{c}$
Constraints $\mathrm{A}\mathbf{x} \le \mathbf{b}$	Constraints $\mathrm{A}^t\mathbf{y} \ge \mathbf{c}$
$\mathbf{x} \ge \mathbf{0}$	$\mathbf{y} \ge \mathbf{0}$.

- For every primal constraint there is a dual variable.

- For every primal variable there is a dual constraint.
- the constraint coefficients of a primal variable form the left-side coefficients of the corresponding dual constraint; the objective coefficient of the same variable becomes the right-hand side of the dual constraint.

We now investigate how the solutions of the primal and dual problems are related, so that by solving one we automatically solve the other.

Proposition 3.2 (The weak duality theorem) *Let* $\mathbf{x}$ *be any feasible solution to the primal problem (P) and let* $\mathbf{y}$ *be any feasible solution to the dual problem (D).*

(i) $\mathbf{c}^t\mathbf{x} \leq \mathbf{b}^t\mathbf{y}$.

(ii) If $\mathbf{c}^t\mathbf{x} = \mathbf{b}^t\mathbf{y}$ *then* $\mathbf{x}$ *and* $\mathbf{y}$ *are optimal solutions to the primal and dual problems.*

Proof

(i) As $\mathbf{x}$ and $\mathbf{y}$ are feasible, we have $\mathrm{A}\mathbf{x} \leq \mathbf{b}$, $\mathrm{A}^t\mathbf{y} \geq \mathbf{c}$, $\mathbf{x} \geq \mathbf{0}$, $\mathbf{y} \geq \mathbf{0}$.

Thus $\mathbf{c}^t\mathbf{x} \leq (\mathrm{A}^t\mathbf{y})^t\mathbf{x} = \mathbf{y}^t\mathrm{A}\mathbf{x} \leq \mathbf{y}^t\mathbf{b} = \mathbf{b}^t\mathbf{y}$.

(ii) Suppose $\mathbf{c}^t\mathbf{x} = \mathbf{b}^t\mathbf{y}$. If $\mathbf{c}^t\mathbf{x}_0 > \mathbf{c}^t\mathbf{x}$ for any primal feasible $\mathbf{x}_0$, then $\mathbf{c}^t\mathbf{x}_0 > \mathbf{b}^t\mathbf{y}$ which contradicts (i). Hence $\mathbf{c}^t\mathbf{x}_0 \leq \mathbf{c}^t\mathbf{x}$ for all primal feasible $\mathbf{x}_0$, so $\mathbf{x}$ is optimal for (P). Similarly, $\mathbf{y}$ is optimal for (D). □

Proposition 3.3 (The strong duality theorem) *If either the primal or dual problem has a finite optimal solution, then so does the other, and the optimum values of the primal and dual objective functions are equal, i.e.* $z_{\max} = w_{\min}$.

Proof

Let $\mathbf{x}$ be a finite optimal solution to the primal, so that $z_{\max} = \mathbf{c}^t\mathbf{x} = z^*$ say.

We have seen that the initial tableau is pre-multiplied as follows to give the final tableau:

$$\left(\begin{array}{c|c} \mathrm{B}^{-1} & \mathbf{0} \\ \hline \mathbf{y}^t & 1 \end{array}\right) \left(\begin{array}{c|c||c} \mathrm{A} & \mathrm{I} & \mathbf{b} \\ \hline -\mathbf{c}^t & \mathbf{0}^t & 0 \end{array}\right) = \left(\begin{array}{c|c||c} \mathrm{B}^{-1}\mathrm{A} & \mathrm{B}^{-1} & \mathrm{B}^{-1}\mathbf{b} \\ \hline \mathbf{y}^t\mathrm{A} - \mathbf{c}^t & \mathbf{y}^t & \mathbf{y}^t\mathbf{b} \end{array}\right).$$

As this is optimal, $\mathbf{y}^t\mathrm{A} - \mathbf{c}^t \geq \mathbf{0}$, so $\mathrm{A}^t\mathbf{y} \geq \mathbf{c}$,

and $\mathbf{y} \geq \mathbf{0}$. Hence $\mathbf{y}$ is feasible for the dual problem.

Now $z^* = \mathbf{y}^t\mathbf{b} = \mathbf{b}^t\mathbf{y}$.

But also $z^* = \mathbf{c}^t\mathbf{x}$ so $\mathbf{x}, \mathbf{y}$ are feasible solutions which give equal values of the primal and dual objective functions respectively.

Thus by Proposition 3.2 (ii), $\mathbf{x}, \mathbf{y}$ are optimal solutions and z^* is the optimal value of both z and w.

As each problem is the dual of the other, the same reasoning applies if we start with a finite optimal solution $\mathbf{y}$ to the dual. □

The above shows that the entries of $\mathbf{y}$ are in fact the optimal values of the main variables in the dual problem. (They are also the shadow prices for the primal constraints).

Furthermore the entries of $\mathbf{y}^t\mathrm{A} - \mathbf{c}^t$ are the values of the dual surplus variables at the optimum, so we shall denote them by $y_{m+1}, \ldots, y_{m+n}$.

Thus the optimal primal tableau contains the following information:

Basic	z	Primal main: $x_1 \cdot \cdot \cdot x_n$	Primal slack: $x_{n+1} \cdot \cdot \cdot x_{n+m}$	Solution
Primal basic variables	0 · 0			Values of primal basic variables
z	1	$y_{m+1} \cdot \cdot \cdot y_{m+n}$ Values of dual surplus variables	$y_1 \cdot \cdot \cdot y_m$ Values of dual main variables	Optimum of objective functions (primal and dual)

The optimal dual solutions may therefore be read from the optimal primal tableau without further calculations. Furthermore, since the dual of the dual is the primal, it does not matter which problem we solve – the optimal solution of one will give us the optimal solution of the other. This is important, as if we are presented with a 'difficult' primal problem, it may be easier to solve it by tackling its dual:

- if the primal constraints are all of the '$\geq$' form then (P) cannot be solved by the normal simplex algorithm, but (D) can;
- if the primal problem has many more constraints than variables then the dual has many fewer constraints than variables, and will in general be quicker to solve.

Proposition 3.4 *If either the primal (P) or the dual (D) has an unbounded optimal solution then the other has no feasible solution.*

Proof: Suppose the dual has a feasible solution $\mathbf{y}$. Then for any primal feasible solution $\mathbf{x}$, $\mathbf{c}^t\mathbf{x} \leq \mathbf{b}^t\mathbf{y}$, so $\mathbf{b}^t\mathbf{y}$ is an upper bound on solutions of the primal. Similarly, if the primal has a feasible solution this places a lower bound on solutions of the dual. It follows that if either problem is unbounded then the other does not have a feasible solution. □

Proposition 3.4 identifies some cases where the duality results do *not* hold, i.e. we cannot say that the primal and dual LP problems have the same optimal values of their objective functions:

1. Primal problem unbounded and dual problem infeasible.
2. Primal problem infeasible and dual problem unbounded.
3. Primal and dual problems **both** infeasible.

3.4 Complementary slackness

Complementary slackness is a very important and useful consequence of the relationship between the primal and dual optima. We continue to work with the primal-dual pair (P) and (D).

In the final tableau for the primal problem, if x_i is non-basic then $x_i = 0$ at the optimum.

Let the primal problem have n main variables and m constraints. Suppose the main variable x_i is basic in the optimal tableau. Then there is a zero in the z row at the bottom of the x_i column, so the dual surplus variable y_{m+i} is zero at the optimum.

We see that for $i = 1, \ldots, n$, either $x_i = 0$ or $y_{m+1} = 0$.

Also, if x_{n+j} is a slack variable in the primal problem then the corresponding dual variable is the main variable y_j. For $j = 1, \ldots, m$, either $y_j = 0$ or $x_{n+j} = 0$.

Thus in every case $x_i y_{m+i} = 0$ and $x_{n+j} y_j = 0$.

So at the optimal solution,

$$\begin{aligned} i\text{th primal main variable} \times i\text{th dual surplus variable} &= 0, \\ j\text{th primal slack variable} \times j\text{th dual main variable} &= 0. \end{aligned}$$

These relationships are called the **complementary slackness** equations.

Thus $(x_1 \ \cdots \ x_n \mid x_{n+1} \ \cdots \ x_{n+m})(y_{m+1} \ \cdots \ y_{m+n} \mid y_1 \ \cdots \ y_m)^t = 0$, since the scalar product of two non-negative vectors is 0 iff the product of each corresponding pair of entries is 0.

Now the entries of $\mathbf{b} - \mathrm{A}\mathbf{x}$ are the primal slack variables $x_{n+1}, \ldots, x_{n+m}$ and the entries of $\mathrm{A}^t\mathbf{y} - \mathbf{c}$ are the dual surplus variables $y_{m+1}, \ldots, y_{m+n}$, so complementary slackness asserts that at the optimal solution,

$$\mathbf{y}^t(\mathbf{b} - \mathrm{A}\mathbf{x}) = 0 \text{ and } \mathbf{x}^t(\mathrm{A}^t\mathbf{y} - \mathbf{c}) = 0.$$

An interpretation of complementary slackness is that if the shadow price of a resource is non-zero then the associated constraint is active at the optimum, i.e. the resource is scarce, but if the constraint is not active (the resource is abundant) then its shadow price is zero.

Proposition 3.5 (The complementary slackness theorem) *A necessary and sufficient condition for* $\mathbf{x}$ *and* $\mathbf{y}$ *to be optimal for the primal and dual problems (P) and (D) is that* $\mathbf{x}$ *is primal feasible,* $\mathbf{y}$ *is dual feasible, and* $\mathbf{x}$ *and* $\mathbf{y}$ *satisfy the complementary slackness conditions* $\mathbf{y}^t(\mathbf{b} - \mathrm{A}\mathbf{x}) = 0$ *and* $\mathbf{x}^t(\mathrm{A}^t\mathbf{y} - \mathbf{c}) = 0$.

Proof:

By the duality theorems, $\mathbf{x}$ and $\mathbf{y}$ are optimal iff they are feasible and $\mathbf{c}^t\mathbf{x} = \mathbf{b}^t\mathbf{y}$.

$$\begin{aligned} \text{Now } \mathbf{c}^t\mathbf{x} = \mathbf{b}^t\mathbf{y} &\Leftrightarrow \mathbf{y}^t\mathbf{b} - \mathbf{x}^t\mathbf{c} = 0 \\ &\Leftrightarrow \mathbf{y}^t\mathbf{b} - \mathbf{y}^t\mathrm{A}\mathbf{x} + \mathbf{x}^t\mathrm{A}^t\mathbf{y} - \mathbf{x}^t\mathbf{c} = 0 \quad (\text{since } \mathbf{y}^t\mathrm{A}\mathbf{x} = \mathbf{x}^t\mathrm{A}^t\mathbf{y}) \\ &\Leftrightarrow \underbrace{\mathbf{y}^t}_{\geq 0}\underbrace{(\mathbf{b} - \mathrm{A}\mathbf{x})}_{\geq 0} + \underbrace{\mathbf{x}^t}_{\geq 0}\underbrace{(\mathrm{A}^t\mathbf{y} - \mathbf{c})}_{\geq 0} = 0 \\ &\Leftrightarrow \mathbf{y}^t(\mathbf{b} - \mathrm{A}\mathbf{x}) = 0 \text{ and } \mathbf{x}^t(\mathrm{A}^t\mathbf{y} - \mathbf{c}) = 0. \end{aligned}$$

□

3.4.1 Examples

1. Consider the rose-growing problem in Chapter 1. The solution of this was

   $(x_1, x_2, x_3, x_4, x_5) = (500, 900, 0, 2200, 0)$.

   The dual problem is : Minimize $w = 6100y_1 + 8000y_2 + 5000y_3$

   subject to $5y_1 + 8y_2 + y_3 \geq 2, 4y_1 + 2y_2 + 5y_3 \geq 3, y_i \geq 0$ for $i = 1, 2, 3$.

   By the strong duality theorem we know that the minimum value of w is the same as the maximum of z in the primal problem, namely 3700. We can read off from the final tableau that $y_1 = y_3 = 1/3, y_2 = y_4 = y_5 = 0$, where y_4, y_5 are the surplus variables in the two constraints of the dual.

If we solved the primal problem graphically, we would only know $x_1 = 500, x_2 = 900$.

Complementary slackness then tells us that $x_1y_4 = x_2y_5 = x_3y_1 = x_4y_2 = x_5y_3 = 0$ so $500y_4 = 900y_5 = 0y_1 = 2200y_2 = 0y_3 = 0$, and since all the y_j are ≥ 0 it follows that $y_2 = y_4 = y_5 = 0$. Thus both dual constraints are active at the optimum, so $5y_1 + y_3 = 2, 4y_1 + 5y_3 = 3$. Solving these gives $y_1 = y_3 = 1/3$.

2. Suppose we wish to verify that $(x_1, x_2) = (10, 5)$ maximizes $z = 30x_1 + 45x_2$ subject to the constraints in the Containers Problem of Chapter 1. (10, 5) is certainly feasible for the primal, i.e. it satisfies the constraints. Then $z = 525$.

   The dual is: Minimize $w = 60y_1 + 60y_2$ subject to $2y_1 + 4y_2 \geq 30, 8y_1 + 4y_2 \geq 45$, $y_1 \geq 0, y_2 \geq 0$. When equality holds in both constraints, $y_1 = 5/2$ and $y_2 = 25/4$.

   Thus $(x_1, x_2, x_3, x_4) = (10, 5, 0, 0)$ is primal feasible, $(y_1, y_2, y_3, y_4) = (5/2, 25/4, 0, 0)$ is dual feasible, and $(10, 5, 0, 0).(0, 0, 5/2, 25/4) = 0$, i.e. the complementary slackness conditions hold. By Theorem 3.5, these solutions are optimal for the primal and dual problems.

   We can further check that when $y_1 = 5/2$ and $y_2 = 25/4$, $w = 525$.

3. Consider again the Cattle Feed problem and its dual from Chapter 1. From the original solution, $z_{\min} = 3.46$ when $x_1 = 3.48, x_2 = 0.56$. If x_3, x_4, x_5, x_6 are the surplus variables in the four constraints, then $x_3 = x_5 = 0$ at the optimum as the first and third constraint are active, but x_4, x_6 are non-zero.

   Let y_5, y_6 be the slack variables in the two dual constraints. By complementary slackness, $(x_1, x_2, \mid x_3, x_4, x_5, x_6).(y_5, y_6, \mid y_1, y_2, y_3, y_4) = 0$. Thus $y_2 = y_4 = y_5 = y_6 = 0$ at the dual optimum.

   Hence $13.2y_1 + 0.02y_3 = 0.66, 34.0y_1 + 0.09y_3 = 2.08$, and solving these gives

   $y_1 = 0.035, y_3 = 9.87$.

   We conclude that the company should charge £0.035 per unit of synthetic protein, £9.87 per unit of synthetic calcium, give away synthetic fat and phosphorus free, and thus charge £3.46 for feeding one cow. No price structure can bring them a higher income without costing the farmer more than before. In accordance with the Strong Duality Theorem, if the farmer and the company both behave rationally (i.e. optimally) then the costs of the normal and synthetic feeding plans are the same. Of course, in practice other considerations might influence the farmer's decision.

   If we solved the dual problem by the simplex algorithm, the solutions for $x_1, \ldots, x_6$ in the primal problem could be read off from the bottom row of the optimal tableau.

   The complementary slackness conditions have the following interpretation here:

   (a) If $(\mathrm{A}\mathbf{x})_i < b_i$ then $y_i = 0$. This means the farmer should buy zero of any nutrient that is overpriced compared to its synthetic equivalent.

   (b) If $(\mathrm{A}^t\mathbf{y})_j > c_j$ then $x_j = 0$. Thus the company should charge zero for any nutrient that is over-supplied in the normal feeding plan.

3.5 Asymmetric duality

The dual problems in Equations (P) and (D) are said to represent **symmetric duality**. We now examine the situation where the variables in the primal problem are unrestricted in sign. Thus the primal problem is:

$$\textbf{Maximize } z = \mathbf{c}^t\mathbf{x} \text{ subject to } \mathrm{A}\mathbf{x} \le \mathbf{b},\ \mathbf{x} \text{ unrestricted.} \qquad \text{(AP)}$$

We can convert this into a LP problem in standard form by letting $\mathbf{x} = \mathbf{x}' - \mathbf{x}''$, where $\mathbf{x}' \ge \mathbf{0}$ and $\mathbf{x}'' \ge \mathbf{0}$.

The constraints then become $\mathrm{A}(\mathbf{x}' - \mathbf{x}'') \le \mathbf{b}$.

This can be written as $(\mathrm{A} \mid -\mathrm{A})\overline{\mathbf{x}} \le \mathbf{b}$, where $\overline{\mathbf{x}} = (x_1' \ \cdots \ x_n' \ \ x_1'' \ \cdots \ x_n'')^t$.

The objective function is $z = \mathbf{c}^t\mathbf{x}' - \mathbf{c}^t\mathbf{x}'' = (\mathbf{c}^t \mid -\mathbf{c}^t)\overline{\mathbf{x}}$, and clearly $\overline{\mathbf{x}} \ge \mathbf{0}$.

Thus the problem becomes

$$\textbf{Maximize } z = (\mathbf{c}^t \mid -\mathbf{c}^t)\overline{\mathbf{x}} \text{ subject to } (\mathrm{A} \mid -\mathrm{A})\overline{\mathbf{x}} \le \mathbf{b},\ \overline{\mathbf{x}} \ge \mathbf{0}.$$

The dual of this problem can be written as:

$$\textbf{Minimize } w = \mathbf{b}^t\mathbf{y} \text{ subject to } \left(\frac{\mathrm{A}^t}{-\mathrm{A}^t}\right)\mathbf{y} \ge \left(\frac{\mathbf{c}}{-\mathbf{c}}\right), \quad \mathbf{y} \ge \mathbf{0}.$$

Now if $-\mathrm{A}^t\mathbf{y} \ge -\mathbf{c}$ then $\mathrm{A}^t\mathbf{y} \le \mathbf{c}$. If this is true simultaneously with $\mathrm{A}^t\mathbf{y} \ge \mathbf{c}$ then we must have $\mathrm{A}^t\mathbf{y} = \mathbf{c}$. Thus the dual of the unrestricted problem (AP) is

$$\textbf{Minimize } w = \mathbf{b}^t\mathbf{y} \text{ subject to } \mathrm{A}^t\mathbf{y} = \mathbf{c},\ \mathbf{y} \ge \mathbf{0}. \qquad \text{(AD)}$$

Conversely, the dual of (AD), a problem with equality constraints, is (AP), a problem in which the variables are unrestricted. (Of course, every LP problem can be expressed as one with equality constraints by including slack and surplus variables.)

Proposition 3.6 *The following two problems are duals of each other:*

Maximize $z = \mathbf{c}^t\mathbf{x}$ *subject to* $\mathrm{A}\mathbf{x} \le \mathbf{b}$, $\mathbf{x}$ *unrestricted,*

Minimize $w = \mathbf{b}^t\mathbf{y}$ *subject to* $\mathrm{A}^t\mathbf{y} = \mathbf{c}$, $\mathbf{y} \ge \mathbf{0}$.

3.5.1 Example

Consider the following asymmetric pair of primal-dual problems:

	Primal		**Dual**
Maximize	$z = x_1 - 2x_2$	**Minimize**	$w = 4y_1 - 5y_2$
subject to	$\begin{cases} x_1 + x_2 \le 4 \\ -2x_1 - 3x_2 \le -5 \end{cases}$	**subject to**	$\begin{cases} y_1 - 2y_2 = 1 \\ y_1 - 3y_2 = -2 \end{cases}$
and	x_1, x_2 unrestricted.	**and**	$y_1, y_2 \ge 0.$

Clearly a solution of the dual can occur only where the two equations hold, i.e. when $y_1 = 7, y_2 = 3$. Thus w has a minimum value of 13, and this must also be the maximum value of z in the primal problem. By complementary slackness we find that $x_1 = 7, x_2 = -3$.

Exercises 3

1. A company which manufactures three products A, B and C, needs to solve the following LP problem in order to maximize their profit.

$$\textbf{Maximize} \qquad z = 3x_1 + x_2 + 5x_3$$

$$\textbf{subject to} \quad \begin{cases} 6x_1 + 3x_2 + 5x_3 \leq 45 \\ 3x_1 + 4x_2 + 5x_3 \leq 30 \end{cases}$$

$$\textbf{and} \qquad x_j \geq 0 \text{ for } j = 1, 2, 3.$$

x_1, x_2 and x_3 are the amounts of A, B and C to be produced. The first constraint is a labour constraint, and the second is a material constraint. The company solves the problem and obtains an optimal solution in which x_1 and x_3 are basic.

   (a) Find the company's optimal solution.

   (b) How much can c_2, the unit profit for B, be increased above 1 without affecting the original optimal solution?

   (c) Find the range of values of c_1, the unit profit for A, for which x_1 and x_3 are still basic at the optimal solution. When this is the case, express the maximum value of z in terms of c_1.

   (d) Find an optimal solution when b_2, the amount of material available, is 60 units.

   (e) The constraint $3x_1 + 2x_2 + 3x_3 \leq 25$ is added to the original problem. How does this affect the original optimal solution?

   (f) A new product D has a unit profit of 5, and its labour and material requirements are 3 units and 4 units respectively. Is it profitable to produce D?

   (g) An additional 15 units of material are available for £10. What should be done?

2. A firm can manufacture four products at its factory. Production is limited by the machine hours available and the number of special components available. The data are given in the table below. Note that production of fractions of a unit is possible.

	Product				Availability
	1	2	3	4	
Machine hours per unit	1	3	8	4	Up to 90 machine hours per day
Components per unit	2	2	1	3	Up to 80 components per day
Production costs (£ per unit)	20	25	40	55	
Sales income (£ per unit)	30	45	80	85	

   (a) Formulate this as a linear programming problem, where x_j is the daily production of product j and the objective is to maximize the daily profit (income minus production costs). Find the optimal solution using the simplex tableau method, and state the optimal profit.

   (b) Write down the shadow prices of machine hours and components, briefly explaining their significance.

   (c) The firm can increase the available machine hours by up to 10 hours per day by hiring extra machinery. The cost of this would be £40 per day. Use sensitivity analysis to decide whether they should hire it, and if so, find the new production schedule.

(d) The production costs of products 1 and 4 are changed by £t per unit. Within what range of values can t lie if the original production schedule is to remain optimal? Find the corresponding range of values of the maximum profit.

(e) Due to a problem at the distributors, the total daily amount produced has to be limited to 25 units. Implement the dual simplex algorithm to find a new production schedule which meets this restriction.

(f) After production has returned to normal (i.e. the original solution is optimal again) the firm considers manufacturing a new product that would require 3 machine hours and 4 components per unit. The production costs would be £45 and sales income £75 per unit. Use sensitivity analysis to decide whether they should go ahead, and if so what the optimum production schedule would be.

3. In each case formulate the dual problem and verify that the given solution is optimal by showing that primal feasibility, dual feasibility and complementary slackness all hold.

   (a) Maximize $19x_1 + 16x_2$ subject to the constraints $x_1 + 4x_2 \leq 20$, $3x_1 + 2x_2 \leq 15$, $x_1 \geq 0$, $x_2 \geq 0$. Solution $(x_1, x_2) = (2, 9/2)$

   (b) Minimize $8x_1 + 11x_2$ subject to the constraints $2x_1 - 2x_2 \geq 2$, $x_1 + 4x_2 \geq -5$, $x_1 \geq 0$, $x_2 \geq 0$. Solution $(x_1, x_2) = (1, 0)$

4. The optimal solution of the problem

$$\textbf{Maximize} \qquad 6x_1 + 4x_2 + 10x_3$$

$$\textbf{subject to} \quad \begin{cases} x_1 + 2x_2 + x_3 \leq 20 \\ 3x_1 + 2x_3 \leq 24 \\ 2x_1 + 2x_2 \leq 22 \end{cases}$$

$$\textbf{and} \qquad x_1, x_2, x_3 \geq 0.$$

   occurs at $(x_1, x_2, x_3) = (0, 4, 12)$. Deduce the solution of the dual problem.

5. Formulate the dual of each of the two problems in Section 2.2.1 and solve them from the optimal primal tableaux using the theory of duality and complementary slackness.

6. By finding and solving the dual problem (without using the simplex algorithm), find the maximum value of

$$z = 5x_1 + 7x_2 + 8x_3 + 4x_4$$

   subject to $x_1 + x_3 \leq 6, x_1 + 2x_4 \leq 5, x_2 + x_3 \leq 9, x_2 + x_4 \leq 3$, where x_1, x_2, x_3, x_4 are unrestricted in sign.

7. Use asymmetric duality to find the solutions (if any) of the following LP problems:

   (a)

$$\textbf{Maximize} \qquad z = x_1 + 2x_2$$

$$\textbf{subject to} \quad \begin{cases} 2x_1 - 3x_2 \leq 1 \\ x_1 + 4x_2 \leq 5 \end{cases}$$

$$\textbf{and} \quad x_1, x_2 \text{ unrestricted.}$$

(b)

$$\textbf{Maximize} \qquad z = 3x_1 + 7x_2 + 5x_3$$

$$\textbf{subject to} \quad \begin{cases} 2x_1 + 5x_2 + 4x_3 \le 9 \\ 2x_1 + 4x_2 + 2x_3 \ge 7 \end{cases}$$

$$\textbf{and} \qquad x_1, x_2, x_3 \text{ unrestricted.}$$

8. When the problem

$$\textbf{Maximize} \qquad 12x_1 + 6x_2 + 4x_3$$

$$\textbf{subject to} \quad \begin{cases} 4x_1 + 2x_2 + x_3 \le 60 \\ 2x_1 + 3x_2 + 3x_3 \le 50 \\ x_1 + 3x_2 + x_3 \le 45 \end{cases}$$

$$\textbf{and} \qquad x_1, x_2, x_3 \ge 0.$$

is solved by the simplex method, using slack variables x_4, x_5, x_6 respectively in the three constraints, the final tableau is

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	Solution
x_1	0	1	3/10	0	3/10	−1/10	0	13
x_3	0	0	4/5	1	−1/5	2/5	0	8
x_6	0	0	19/10	0	−1/10	−3/10	1	24
z	1	0	4/5	0	14/5	2/5	0	188

(a) State the optimal solution and the values of $x_1, \ldots, x_6$ at the optimum.

(b) Write down the dual problem, using y_1, y_2, y_3 for the dual main variables and y_4, y_5, y_6 for the dual surplus variables.

(c) *Using the above tableau*, write down the optimal solution of the dual problem and give the values of $y_1, \ldots, y_6$ at the optimum.

(d) Show how complementary slackness occurs in these solutions.

(e) Convert the dual problem to a maximization problem and solve it by the Dual Simplex method.

(f) Comment on the relationships between the two optimal tableaux.

9. By solving the *dual* problem graphically, solve the LP problem:

$$\textbf{Minimize} \qquad 4x_1 + 3x_2 + x_3$$

$$\textbf{subject to} \quad \begin{cases} 3x_1 + 8x_2 + 2x_3 \ge 3 \\ 2x_1 + 5x_2 + 3x_3 \ge 5 \end{cases}$$

$$\textbf{and} \qquad x_1, x_2, x_3 \ge 0.$$

Give the values of all the dual and primal variables (main, slack and surplus) at the optimum.

10. Find the dual of the problem

$$\textbf{Maximize } \mathbf{c}^t\mathbf{x} \text{ subject to } A\mathbf{x} = \mathbf{b},\ \mathbf{x} \ge \mathbf{0}.$$

Chapter 4

The Transportation Problem

4.1 Formulation of the model

The transportation model is concerned with finding the minimum cost of transporting a single commodity from a given number of sources (e.g. factories) to a given number of destinations (e.g. warehouses). Any destination can receive its demand from more than one source. The objective is to find how much should be shipped from each source to each destination so as to minimize the total transportation cost.

Supply		Sources		Destinations		Demand
a_1	$\rightarrow$	S_1	c_{11}	D_1	$\rightarrow$	b_1
a_2	$\rightarrow$	S_2		D_2	$\rightarrow$	b_2
$\vdots$		$\vdots$		$\vdots$		$\vdots$
a_m	$\rightarrow$	S_m	c_{mn}	D_n	$\rightarrow$	b_n

The figure represents a transportation model with m sources and n destinations. Each source or destination is represented by a point. The route between a source and destination is represented by a line joining the two points. The supply available at source i is a_i, and the demand required at destination j is b_j. The cost of transporting one unit between source i and destination j is c_{ij}.

When the total supply is equal to the total demand (i.e. $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$) then the transportation model is said to be **balanced**. In a balanced transportation problem each supply must be entirely used and each demand must be exactly satisfied, so $\sum_{j=1}^{n} x_{ij} = a_i$ for $i = 1, \ldots, m$ and $\sum_{i=1}^{m} x_{ij} = b_j$ for $j = 1, \ldots, n$.

The following is an example of a balanced transportation problem:

	Warehouse 1	Warehouse 2	Warehouse 3	Supply
Factory 1	c_{11}	c_{12}	c_{13}	20
Factory 2	c_{21}	c_{22}	c_{23}	10
Demand	7	10	13	

Total supply $= 20 + 10 = 30 = 7 + 10 + 13 =$ Total demand.

A transportation model in which the total supply and total demand are not equal is called **unbalanced**. It is always possible to balance an unbalanced transportation problem.

Suppose the demand at warehouse 1 above is 9 units. Then the total supply and total demand are unequal, and the problem is unbalanced. In this case it is not possible to satisfy all the demand at each destination simultaneously.

We modify the model as follows: since demand exceeds supply by 2 units we introduce a **dummy source**, Factory 3, which has a capacity of 2. The amount sent from this dummy source to a destination represents the shortfall at that destination.

If supply exceeds demand then a **dummy destination**, Warehouse 4, is introduced to absorb the surplus units. Any units shipped from a source to a dummy destination represent a surplus at that source.

Transportation costs for dummy sources or destinations are allocated as follows:

- If a **penalty cost** is incurred for each unit of unsatisfied demand or unused supply, then the transportation cost is set equal to the penalty cost.
- If there is no penalty cost, the transportation cost is set equal to zero.
- If *no* units may be assigned to a dummy or a particular route, allocate a cost M. This represents a number larger than any other in the problem – think of it as a million!

From now on we shall consider balanced transportation problems only, as any unbalanced problem can be balanced by introducing a dummy.

Let x_{ij} denote the amount transported from source i to destination j. Then the problem is

$$\textbf{Minimize} \qquad z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij},$$

$$\textbf{subject to} \qquad \sum_{j=1}^{n} x_{ij} = a_i \text{ for } i = 1, \ldots, m$$

$$\textbf{and} \qquad \sum_{i=1}^{m} x_{ij} = b_j \text{ for } j = 1, \ldots, n,$$

$$\textbf{where} \qquad x_{ij} \geq 0 \text{ for all } i \text{ and } j.$$

4.2 Solution of the transportation problem

A balanced transportation problem has $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$. Hence one constraint is a linear combination of the others, so there are $n + m - 1$ independent constraint equations.

It is not practicable to use the standard simplex method to solve the transportation problem. However, there is an efficient tableau-based method which makes use of the dual problem.

Starting the algorithm: finding an initial basic feasible solution

Here we examine ways of constructing initial basic feasible solutions, i.e. allocations with $m + n - 1$ basic variables.

Method 1: The North-West Corner Method

Consider the problem represented by the following **transportation tableau**. The number in the bottom right of cell (i, j) is c_{ij}, the cost of transporting 1 unit from source i to destination j.

					Supply
	10	0	20	11	15
	12	7	9	20	25
	0	14	16	18	5
Demand	5	15	15	10	

The **north-west corner method** proceeds as follows:

- Assign as much as possible to the cell in the top-left of the tableau.
- Cross out the row or column whose supply or demand is satisfied. If a row and column are both satisfied then cross out only one of them.
- Adjust the supply and demand for those rows and columns which are not crossed out.
- Repeat the above steps on the remaining tableau until only one row or column remains.

The values of the basic variables x_{ij} are entered in the top left of each cell. There should always be $m+n-1$ of these; in certain (degenerate) cases some of them may be zero. They must always add up to the total supply and demand in each row and column.

Note that some books position the data differently in the cells of the tableau.

Method 2: The Least-Cost Method

This method usually provides a better initial basic feasible solution than the North-West Corner method. Despite its name, it does *not* give the actual minimum cost. It *uses* least available costs to obtain a starting tableau.

- Assign as much as possible to the cell with the smallest unit cost in the entire tableau. If there is a tie then choose arbitrarily. It may be necessary to assign 0.
- Cross out the row or column whose supply or demand is satisfied. If a row and column are both satisfied then cross out only one of them.
- Adjust the supply and demand for those rows and columns which are not crossed out.
- Repeat the above steps on the remaining tableau until only one row or column remains.

For the above example,

					Supply
	10	0	20	11	15
	12	7	9	20	25
	0	14	16	18	5
Demand	5	15	15	10	

Checking for optimality and iterating the algorithm

So far, we have only looked at ways of obtaining an initial basic feasible solution to the balanced transportation problem.

We now develop a method for checking whether the current basic feasible solution is optimal, and a way of moving to a better basic feasible solution if the current solution is not optimal.

Using asymmetric duality, the dual of the transportation problem can be written as

$$\begin{aligned} \textbf{Maximize} \quad & w = \sum_{i=1}^{m} a_i\lambda_i + \sum_{j=1}^{n} b_j\mu_j, \\ \textbf{subject to} \quad & \lambda_i + \mu_j \le c_{ij} \text{ for each } i \text{ and } j \\ \textbf{and} \quad & \lambda_i, \mu_j \text{ unrestricted in sign.} \end{aligned}$$

Introducing slack variables s_{ij}, the constraints can be written as

$$\lambda_i + \mu_j + s_{ij} = c_{ij}, \quad i = 1, \ldots, m, \ j = 1, \ldots, n.$$

By the complementary slackness conditions we must have $x_{ij}s_{ij} = 0$, i.e.

$$x_{ij}\left(c_{ij} - \lambda_i - \mu_j\right) = 0 \text{ for all } i \text{ and } j.$$

As before, primal feasibility, dual feasibility and complementary slackness are necessary and sufficient for optimality. This is the underlying strategy for solving the problem.

For illustrative purposes, we shall start the algorithm for the above example using the bfs that was provided by the North-West Corner method. The Least-Cost method will usually give a better initial allocation.

We assign λ's and μ's which satisfy $\lambda_i + \mu_j = c_{ij}$ to the rows and columns containing the current *basic* variables. (This comes from the complementary slackness condition $x_{ij}s_{ij} = 0$.) As one constraint in the original problem was redundant, we can choose *one* of the λ or μ values arbitrarily. For simplicity, it is conventional to set $\lambda_1 = 0$.

The values of $s_{ij} = c_{ij} - \lambda_i - \mu_j$ are entered in the top right of the cells.

If all the s_{ij} values are non-negative, we have an optimal solution.

Carrying out this procedure, the initial transportation tableau becomes:

	10	0	2	13
0	5 10	10 0	20	11
7	12	5 7	15 9	5 20
5	0	14	16	5 18

We **test for optimality** by checking whether $s_{ij} = c_{ij} - \lambda_i - \mu_j \geq 0$ for all i and j, i.e. in all cells. (This is the dual feasibility condition). If this holds for *every* cell of the tableau then the optimum has been reached.

Otherwise, choose the cell with the **most negative** value of s_{ij}.

This identifies the variable to enter the basis. In this case the entering variable is x_{31}.

Determining the leaving variable

We construct a **closed loop** that starts and ends at the entering variable, and links it to basic variables by a succession of horizontal and vertical segments. It does not matter whether the loop is clockwise or anticlockwise.

Initial tableau

	10	0	2	13
0	5 10	10 0	20	11
7	12	5 7	15 9	5 20
5	0	14	16	5 18

We now see how large the entering variable can be made without violating the feasibility conditions. Suppose x_{31} increases from zero to some level $\varepsilon > 0$. Then x_{11} must change to

$5 - \varepsilon$ to preserve the demand constraint in column 1. This has a knock on effect for x_{12} which therefore changes to $10 + \varepsilon$. This process continues for *all* the corners of the loop.

The departing variable is chosen from among the corners of the loop which *decrease* when the entering variable increases above zero level. It is the one with the smallest current value, as this will be the first to reach zero as the entering variable increases. Any further increase in the entering variable past this value leads to infeasibility.

We may choose any one of x_{11}, x_{22} or x_{34} as the departing variable here. Arbitrarily, we choose x_{34}. The entering variable x_{31} can increase to 5 and feasibility will be preserved.

Second tableau

	10	0	2	13
0	0 10	15 0	20	11
7	12	0 7	15 9	10 20
-10	5 0	14	16	18

Notice that some of the basic variables are zero valued – this solution is degenerate. However, this causes no problem to the general method of solving the problem.

As before, we construct λ's and μ's which satisfy $\lambda_i + \mu_j = c_{ij}$ for the **basic** variables. Then we check for optimality as before. This tableau is not optimal because $s_{ij} \geq 0$ does **not** hold for all the cells. The most negative value of s_{ij} occurs for x_{21}, so this is the entering variable.

Next we construct a loop. Thus ε can only be as large as zero. (This is bound to happen because of the degeneracy of the current solution). We let x_{11} be the departing variable.

Third tableau

	5	0	2	13
0	10	15 0	20	11
7	0 12	0 7	15 9	10 20
-5	5 0	14	16	18

Again, this is a degenerate solution, as some of the basic variables are equal to zero. We construct λ's and μ's as before, and then check for optimality. The tableau is not optimal, and x_{14} is the entering variable. The loop construction shows that ε can be as large as 10, and that x_{24} is the departing variable.

Fourth tableau

	5	0	2	11
0	10	5 0	20	10 11
7	0 12	10 7	15 9	20
-5	5 0	14	16	18

This is now optimal because $\lambda_i + \mu_j \leq c_{ij}$, i.e. $s_{ij} \geq 0$, in every cell. The minimum cost is therefore given by $5 \times 0 + 10 \times 11 + 0 \times 12 + 10 \times 7 + 15 \times 9 + 5 \times 0 = 315$, which occurs when $x_{12} = 5$, $x_{14} = 10$, $x_{22} = 10$, $x_{23} = 15$, $x_{31} = 5$, and all the other decision variables are equal to zero.

4.2.1 Example

This example emphasizes the connection between the transportation algorithm and the primal-dual linear programming problems which underlie the method.

Three factories F_1, F_2, F_3 produce 15000, 25000 and 15000 units respectively of a commodity. Three warehouses W_1, W_2, W_3 require 20000, 19000 and 16000 units respectively.

The cost of transporting from F_i to W_j is £c_{ij} per unit, where $c_{11} = 12$, $c_{12} = 7$, $c_{13} = 10$, $c_{21} = 10$, $c_{22} = 8$, $c_{23} = 6$, $c_{31} = 9$, $c_{32} = 15$, $c_{33} = 8$.

If x_{ij} thousand units are transported from F_i to W_j, the total cost £$1000z$ is given by

$z = 12x_{11} + 7x_{12} + 10x_{13} + 10x_{21} + 8x_{22} + 6x_{23} + 9x_{31} + 15x_{32} + 8x_{33}$

which must be minimized subject to the constraints

$$\begin{array}{ccccccccccccccccccc}
x_{11} & + & x_{12} & + & x_{13} & & & & & & & & & & & & & = & 15 \\
& & & & & & x_{21} & + & x_{22} & + & x_{23} & & & & & & & = & 25 \\
& & & & & & & & & & & & x_{31} & + & x_{32} & + & x_{33} & = & 15 \\
x_{11} & & & & & + & x_{21} & & & & & + & x_{31} & & & & & = & 20 \\
& & x_{12} & & & & & + & x_{22} & & & & & + & x_{32} & & & = & 19 \\
& & & & x_{13} & & & & & + & x_{23} & & & & & + & x_{33} & = & 16
\end{array}$$

where $x_{ij} \geq 0$ for $i, j = 1, 2, 3$.

By asymmetric duality, the dual of this problem is :

Maximize $w = 15\lambda_1 + 25\lambda_2 + 15\lambda_3 + 20\mu_1 + 19\mu_2 + 16\mu_3$

subject to

$$\begin{array}{rcl}
\lambda_1 + \mu_1 \leq 12 & \text{i.e.} & \lambda_1 + \mu_1 + s_{11} = 12 \\
\lambda_1 + \mu_2 \leq 7 & \text{i.e.} & \lambda_1 + \mu_2 + s_{12} = 7 \\
\lambda_1 + \mu_3 \leq 10 & \text{i.e.} & \lambda_1 + \mu_3 + s_{13} = 10 \\
\lambda_2 + \mu_1 \leq 10 & \text{i.e.} & \lambda_2 + \mu_1 + s_{21} = 10 \\
\lambda_2 + \mu_2 \leq 8 & \text{i.e.} & \lambda_2 + \mu_2 + s_{22} = 8 \\
\lambda_2 + \mu_3 \leq 6 & \text{i.e.} & \lambda_2 + \mu_3 + s_{23} = 6 \\
\lambda_3 + \mu_1 \leq 9 & \text{i.e.} & \lambda_3 + \mu_1 + s_{31} = 9 \\
\lambda_3 + \mu_2 \leq 15 & \text{i.e.} & \lambda_3 + \mu_2 + s_{32} = 15 \\
\lambda_3 + \mu_3 \leq 8 & \text{i.e.} & \lambda_3 + \mu_3 + s_{33} = 8
\end{array}$$

where $s_{ij} \geq 0$ for $i, j = 1, 2, 3$ but λ_i, μ_j are unrestricted in sign.

Complementary slackness tells us that $x_{ij}s_{ij} = 0$ for all i and j.

12	7	10
10	8	6
9	15	8

12	7	10
10	8	6
9	15	8

12	7	10
10	8	6
9	15	8

12	7	10
10	8	6
9	15	8

We choose to find an initial bfs by the north-west corner method. After three iterations, all the s_{ij} are non-negative so we have primal feasibility, dual feasibility and complementary slackness. Hence the optimal solution has been found. The minimum cost is £418,000.

The primal solution is $(x_{11}, \ldots, x_{13}, \ldots) = (0, 15, 0, 5, 4, 16, 15, 0, 0,\ \mid 0, 0, 0, 0, 0, 0)$. The last six 0's represent unnecessary slack variables in the primal problem; they are included only to show that complementary slackness does indeed hold when we look at the dual solution $(\lambda_1, \lambda_2, \lambda_3, \mu_1, \mu_2, \mu_3, s_{11}, \ldots, s_{33}) = (0, 1, 0, 9, 7, 5,\ \mid 3, 0, 5, 0, 0, 0, 0, 8, 3)$.

The problem is now modified as follows: The demand at W_2 is increased to 28. There is no link between F_2 and W_2. All the demand at W_3 must be satisfied.

We add a dummy source F_4 with capacity 9.

The costs c_{22} and c_{43} are set equal to a **large number** M. This is the standard method for ensuring that the allocation to a particular cell is always zero. M is to be thought of as larger than any other number in the problem. The tableau becomes:

12	7	10
10	M	6
9	15	8
0	0	M

12	7	10
10	M	6
9	15	8
0	0	M

12	7	10
10	M	6
9	15	8
0	0	M

12	7	10
10	M	6
9	15	8
0	0	M

We can find an initial allocation by the least-cost method, or by adapting the existing optimal tableau.

The minimum cost is £450,000; this *is* uniquely determined, though the allocation which produces it may not be.

Exercises 4

1. For the transportation problem given by the following tableau, find an initial basic feasible solution by the least-cost method and proceed to find an optimal solution.

				Supply
	2	1	3	7
	4	5	6	8
Demand	5	6	4	

2. Formulate the transportation problem in Question 1 in linear programming form. Also state the dual problem. From your final tableau, write down the values of all the primal and dual variables at the optimal solution. Show how this provides a check on your answer.

3. For the transportation problem given by the following tableau, find an initial basic feasible solution by the North-West corner method and then find an optimal solution.

						Supply
	10	15	10	12	20	8
	5	10	8	15	10	7
	15	10	12	12	10	10
Demand	5	9	2	4	5	

   The supply at Source 3 is now reduced from 10 to 6. There is a penalty of 5 for each unit required but not supplied. Find the new optimal solution.

4. Three refineries with maximum daily capacities of 6, 5, and 8 million gallons of oil supply three distribution areas with daily demands of 4, 8 and 7 million gallons. Oil is transported to the three distribution areas through a network of pipes. The transportation cost is 1p per 100 gallons per mile. The mileage table below shows that refinery 1 is not connected to distribution area 3. Formulate the problem as a transportation model and solve it. [Hint: Let the transportation cost for the non-connected route be equal to some large value M say and then proceed as normal.]

		Distribution Area 1	2	3
Refinery	1	120	180	—
	2	300	100	80
	3	200	250	120

5. In Question 4, suppose additionally that the capacity of refinery 3 is reduced to 6 million gallons. Also, distribution area 1 must receive all its demand, and any shortage at areas 2 and 3 will result in a penalty of 5 pence per gallon. Formulate the problem as a transportation model and solve it.

6. In Question 4, suppose the daily demand at area 3 drops to 4 million gallons. Any surplus production at refineries 1 and 2 must be diverted to other distribution areas by tanker. The resulting average transportation costs per 100 gallons are £1.50 from refinery 1 and £2.20 from refinery 2. Refinery 3 can divert its surplus oil to other chemical processes within the plant. Formulate the problem as a transportation model and solve it.

Chapter 5

Non-linear optimization

5.1 Local and global optima

Let f be a real-valued function defined on a domain $S \subset \mathbb{R}^n$, so f is applied to a vector $\mathbf{x} = (x_1, \ldots, x_n) \in S$ to give a real number $f(\mathbf{x})$ as the result.

A vector $\mathbf{x}^* \in S$ is a **global maximizer** of $f(\mathbf{x})$ over S if $f(\mathbf{x}^*) \geq f(\mathbf{x})$ for all $\mathbf{x} \in S$.

$f(\mathbf{x}^*)$ is then the **global maximum value** of $f(\mathbf{x})$ over S.

$\mathbf{x}^*$ is a **strong** global maximizer of $f(\mathbf{x})$ over S if $f(\mathbf{x}^*) > f(\mathbf{x})$ for all $\mathbf{x} \in S$ other than $\mathbf{x}^*$.

Replacing '$\geq$' or '$>$' by '$\leq$' or '$<$' gives the definitions for a (strong) **global minimizer**.

$\mathbf{x}^*$ is a **local maximizer** of $f(\mathbf{x})$ over S if there exists $\varepsilon > 0$ such that $f(\mathbf{x}^*) \geq f(\mathbf{x})$ whenever $\mathbf{x} \in S$ and $|\mathbf{x} - \mathbf{x}^*| < \varepsilon$. That is, $f(\mathbf{x}^*) \geq f(\mathbf{x})$ for all $\mathbf{x} \in S$ sufficiently close to $\mathbf{x}^*$.

Then $f(\mathbf{x}^*)$ is a **local maximum value** of $f(\mathbf{x})$. A local minimum is defined analogously.

An **optimum** or **extremum** is either a maximum or a minimum.

A global optimum is a local optimum, but a local optimum may not be a global optimum.

Recall that if $f(x)$ is a differentiable function of *one* real variable x, local maxima and minima at interior points of the domain occur when $f'(x) = 0$. The second derivative $f''(x)$ is used to determine the nature of the optimum.

For functions of more than one real variable, the corresponding methods involve a vector of first derivatives and a matrix of second derivatives, as follows:

The **gradient vector** of a differentiable function $f(\mathbf{x})$ is $\nabla f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix}$.

A **critical point** of $f(\mathbf{x})$ is a point at which $\nabla f(\mathbf{x}) = \mathbf{0}$.

The **Hessian matrix** of a twice-differentiable function $f(\mathbf{x})$ is the symmetric $n \times n$ matrix

$$H(f(\mathbf{x})) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{pmatrix}.$$

5.1.1 Examples

1. $f(x, y, z) = 4x - 2y + z$

2. $f(x, y) = -2x^2 + 6xy + y^2$

3. $f(x, y, z) = e^{2x}y + \dfrac{\ln 2z}{y}$

Now let f be a twice-differentiable real-valued function defined on a set $S \subset \mathbb{R}^n$.

Let $\mathbf{a} = (a_1, \ldots, a_n)$ be in S and let $\mathbf{u} = (u_1, \ldots, u_n)$ be such that, for some $c > 0$, $\mathbf{a}+r\mathbf{u} \in S$ for all $r \in [0, c]$. Such a vector $\mathbf{u}$ is called a **feasible direction** at $\mathbf{a}$.

Let $\mathbf{x} = \mathbf{a} + r\mathbf{u}$, so $x_i = a_i + ru_i$ for $i = 1, \ldots, n$. For $0 \le r \le c$, define $g(r)$ to be $f(\mathbf{x})$.

By the Chain Rule, $g'(r) = \dfrac{\mathrm{d}g(r)}{\mathrm{d}r} = \dfrac{\mathrm{d}f(\mathbf{x})}{\mathrm{d}r} = \displaystyle\sum_{i=1}^{n} \frac{\partial f(\mathbf{x})}{\partial x_i}\frac{\mathrm{d}x_i}{\mathrm{d}r} = \sum_{i=1}^{n} u_i \frac{\partial f(\mathbf{x})}{\partial x_i} = \mathbf{u}^t\nabla f(\mathbf{x})$.

Hence $g'(0) = \mathbf{u}^t\nabla f(\mathbf{a})$.

When $\mathbf{u}$ is a unit vector, $\mathbf{u}^t\nabla f(\mathbf{a})$ is called the **directional derivative** of f at $\mathbf{a}$ in the direction $\mathbf{u}$. It measures the rate of increase of f as we move from $\mathbf{a}$ in the direction $\mathbf{u}$.

Also $g''(r) = \dfrac{\mathrm{d}}{\mathrm{d}r}\left(\displaystyle\sum_{i=1}^{n} u_i \frac{\partial f(\mathbf{x})}{\partial x_i}\right) = \sum_{i=1}^{n} u_i \sum_{j=1}^{n} \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}\frac{\mathrm{d}x_j}{\mathrm{d}r} = \sum_{i,j=1}^{n} u_i \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} u_j = \mathbf{u}^t H(f(\mathbf{x}))\mathbf{u}$,

so $g''(0) = \mathbf{u}^t H(f(\mathbf{a}))\mathbf{u}$. Thus we have:

Proposition 5.1 *If* $\mathrm{g}(r) = \mathrm{f}(\mathbf{a} + r\mathbf{u})$ *then* $\mathrm{g}'(0) = \mathbf{u}^t \nabla \mathrm{f}(\mathbf{a})$ *and* $\mathrm{g}''(0) = \mathbf{u}^t \mathrm{H}(\mathrm{f}(\mathbf{a}))\mathbf{u}$.

Proposition 5.2 (First order necessary condition for a local optimum.)
Let $\mathrm{f}(\mathbf{x})$ *be a differentiable function on* $S \subset \mathbb{R}^n$. *If* $\mathrm{f}(\mathbf{x})$ *attains a local maximum or minimum value over* S *at an interior point* $\mathbf{x}^* \in S$ *then* $\nabla \mathrm{f}(\mathbf{x}^*) = \mathbf{0}$.

Proof

As $\mathbf{x}^*$ is an interior point of S, we can move from $\mathbf{x}^*$ in any direction $\mathbf{u}$ and still be in S.

If $\mathbf{x}^*$ is a local maximizer, $\mathrm{f}(\mathbf{x})$ must be non-increasing in *every* direction moving away from $\mathbf{x}^*$ so $\mathbf{u}^t \nabla \mathrm{f}(\mathbf{x}^*) \leq 0$ and $(-\mathbf{u})^t \nabla \mathrm{f}(\mathbf{x}^*) \leq 0$. Hence $\mathbf{u}^t \nabla \mathrm{f}(\mathbf{x}^*) = 0$ for *all* $\mathbf{u} \in \mathbb{R}^n$.

Taking $\mathbf{u} = \mathbf{e}_i$ (the vector with ith entry 1 and all other entries 0) shows that the ith component of $\nabla \mathrm{f}(\mathbf{x}^*)$ is 0 for $i = 1, \ldots, n$. Hence $\nabla \mathrm{f}(\mathbf{x}^*) = \mathbf{0}$.

If $\mathbf{x}^*$ is a local minimizer, the same reasoning applies with 'decrease' replaced by 'increase' and '$\leq$' by '$\geq$'. □

This condition is not *sufficient* for a maximum or minimum, as $\nabla \mathrm{f}(\mathbf{x}) = \mathbf{0}$ also holds at a **saddle point**, i.e. a critical point where $\mathrm{f}(\mathbf{x})$ is locally neither minimum nor maximum.

5.2 Quadratic forms

Recall that if A is a square matrix, the function $\mathbf{x}^t \mathrm{A}\mathbf{x}$ is called a **quadratic form**. Any expression consisting entirely of second-order terms can be written in this form. We can always choose A to be symmetric.

For example, ${x_1}^2 - 8x_1x_2 + 5{x_2}^2 = \mathbf{x}^t \mathrm{A}\mathbf{x}$, where $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $\mathrm{A} = \begin{pmatrix} 1 & -4 \\ -4 & 5 \end{pmatrix}$.

Similarly $2x^2 + y^2 - z^2 + 4xy - 6yz = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 2 & 2 & 0 \\ 2 & 1 & -3 \\ 0 & -3 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

A real square matrix A, and equivalently the quadratic form $\mathbf{x}^t \mathrm{A}\mathbf{x}$, is defined to be

positive definite if $\mathbf{x}^t \mathrm{A}\mathbf{x} > 0 \;\forall\; \mathbf{x} \neq \mathbf{0}$, **positive semi-definite** if $\mathbf{x}^t \mathrm{A}\mathbf{x} \geq 0 \;\forall\; \mathbf{x}$,

negative definite if $\mathbf{x}^t \mathrm{A}\mathbf{x} < 0 \;\forall\; \mathbf{x} \neq \mathbf{0}$, **negative semi-definite** if $\mathbf{x}^t \mathrm{A}\mathbf{x} \leq 0 \;\forall\; \mathbf{x}$,

indefinite if none of the above is true. (Note that 'semi-definite' includes 'definite'.)

When $\mathbf{x} \in \mathbb{R}^2$, we can visualise these definitions graphically; e.g. if A is positive definite, the surface $z = \mathbf{x}^t \mathrm{A}\mathbf{x}$ lies wholly above the (x, y) plane and meets it only at $(0, 0, 0)$.

By orthogonally diagonalizing the symmetric matrix, any quadratic form can be expressed as $\lambda_1 {y_1}^2 + \cdots + \lambda_n {y_n}^2$ where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of the matrix. Using this we showed in Level 1 Linear Algebra that a real symmetric matrix is positive definite if and only if all its eigenvalues are positive, and negative definite if and only if all its eigenvalues are negative.

The determinant of a square matrix is the product of the eigenvalues, so *if* a real symmetric matrix is positive definite then its determinant is strictly positive.

The converse of this is *not* true, but we can say that if a real symmetric matrix A is either positive or negative definite then it is non-singular. If A is semi-definite but not definite, $\mathbf{x}^t \mathrm{A}\mathbf{x} = 0$ for some $\mathbf{x} \neq \mathbf{0}$, hence some $\lambda_i = 0$ so A is singular.

The kth order **principal minors** of an $n \times n$ matrix A are the determinants of all the $k \times k$ matrices that can be formed using k rows and the corresponding k columns of A.

The kth **leading principal minor** of A is the determinant of the $k \times k$ matrix formed from the *first* k rows and the *first* k columns of A.

Thus the first leading principal minor is just the top left entry and the nth one is det(A).

For example, $\begin{pmatrix} 4 & 2 & -3 \\ 2 & -1 & 0 \\ -3 & 0 & 3 \end{pmatrix}$ has principal minors $4, -1, 3$ (first order), $-8, 3, -3$ (second order), -15 (third order). The leading ones are $4, -8, -15$.

Proposition 5.3 *A real symmetric matrix* A *is positive definite if and only if all its leading principal minors are strictly positive.*

Proof

($\Leftarrow$) Let M_j be the submatrix of A consisting of the first j rows and the first j columns.

We can write $\mathrm{A} = \left(\begin{array}{c|c} \mathrm{M}_j & \mathrm{Q}_j \\ \hline \mathrm{R}_j & \mathrm{S}_j \end{array}\right)$ where Q_j, R_j, S_j are matrices of appropriate sizes (0×0 when $j = n$).

Suppose all leading principal minors of A are positive, so $\det(\mathrm{M}_1), \ldots, \det(\mathrm{M}_n)$ are all > 0.

For $j = 1, \ldots, n$, let $\mathbf{u}_j$ be the vector in $\mathbb{R}^j$ with jth entry 1 and all other entries (if any) 0 and let $\overline{\mathbf{p}}_j = \begin{pmatrix} p_{1j} \\ \vdots \\ p_{jj} \end{pmatrix}$ satisfy the linear system of equations $\mathrm{M}_j \overline{\mathbf{p}}_j = \mathbf{u}_j$.

By Cramer's rule, $p_{jj} = \dfrac{\det(\mathrm{M}_{j-1})}{\det(\mathrm{M}_j)}$ which is greater than 0.

Let P be the upper triangular matrix with entries p_{ij} for $i \leq j$ and 0 for $i > j$.

The jth column of P is $\mathbf{p}_j = \left(\begin{array}{c} \overline{\mathbf{p}}_j \\ \hline \mathbf{0} \end{array}\right)$.

Thus column j of AP is $\left(\begin{array}{c|c} \mathrm{M}_j & \mathrm{Q}_j \\ \hline \mathrm{R}_j & \mathrm{S}_j \end{array}\right) \left(\begin{array}{c} \overline{\mathbf{p}}_j \\ \hline \mathbf{0} \end{array}\right) = \begin{pmatrix} \mathrm{M}_j \overline{\mathbf{p}}_j \\ \mathrm{R}_j \overline{\mathbf{p}}_j \end{pmatrix} = \begin{pmatrix} \mathbf{u}_j \\ \mathrm{R}_j \overline{\mathbf{p}}_j \end{pmatrix}$.

Hence AP is lower-triangular, with all the entries on its main diagonal equal to 1.

Let $\mathrm{C} = \mathrm{P}^t\mathrm{AP}$. C is a product of two lower-triangular matrices, so is itself lower-triangular. But $\mathrm{C}^t = \mathrm{P}^t\mathrm{A}^t\mathrm{P} = \mathrm{P}^t\mathrm{AP}$ as A is symmetric, so C is symmetric. Hence C is a diagonal matrix, and its diagonal entries are $p_{11}, \ldots, p_{nn}$ which we have seen are all positive.

Let $\mathbf{x} = \mathrm{P}\mathbf{y}$. Then $\mathbf{x}^t\mathrm{A}\mathbf{x} = \mathbf{y}^t\mathrm{C}\mathbf{y} = \sum_{j=1}^{n} p_{jj}{y_j}^2 > 0$ for all $\mathbf{y} \neq \mathbf{0}$.

P is non-singular, so $\mathbf{y} \neq \mathbf{0}$ iff $\mathbf{x} \neq \mathbf{0}$. Hence $\mathbf{x}^t\mathrm{A}\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$, so A is positive definite.

($\Rightarrow$) Suppose A is positive definite.

Let $\mathbf{y}$ be any non-zero vector in $\mathbb{R}^j$ and let $\mathbf{x} = \left(\begin{array}{c} \mathbf{y} \\ \hline \mathbf{0} \end{array}\right) \in \mathbb{R}^n$.

Then $\mathbf{x}^t\mathrm{A}\mathbf{x} = (\mathbf{y}^t \mid \mathbf{0}) \left(\begin{array}{c|c} \mathrm{M}_j & \mathrm{Q}_j \\ \hline \mathrm{R}_j & \mathrm{S}_j \end{array}\right) \left(\begin{array}{c} \mathbf{y} \\ \hline \mathbf{0} \end{array}\right) = (\mathbf{y}^t \mid \mathbf{0}) \left(\begin{array}{c} \mathrm{M}_j\mathbf{y} \\ \hline \mathrm{R}_j\mathbf{y} \end{array}\right) = \mathbf{y}^t\mathrm{M}_j\mathbf{y}$.

Since $\mathbf{x} \neq \mathbf{0}$ and A is positive definite, $\mathbf{x}^t \mathrm{A} \mathbf{x} > 0$, so $\mathbf{y}^t \mathrm{M}_j \mathbf{y} > 0$ for all $\mathbf{y} \neq \mathbf{0}$.

Hence M_j is positive definite, and thus $\det(\mathrm{M}_j) > 0$ for $j = 1, \ldots, n$. □

Note that A is positive definite if and only if $-\mathrm{A}$ is negative definite. The leading principal minors of $-\mathrm{A}$ are then alternately negative and positive.

The results can be further extended to semi-definite matrices. The following table gives necessary and sufficient conditions for a real *symmetric* matrix A, and the associated quadratic form $\mathbf{x}^t \mathrm{A} \mathbf{x}$, to be positive / negative (semi-) definite. *Either* the eigenvalues condition *or* the principal minors condition may be used. The latter is generally easier.

Symmetric matrix	Definition	Eigenvalues	Principal Minors
Positive definite	$\mathbf{x}^t \mathrm{A} \mathbf{x} > 0 \ \forall \mathbf{x} \neq \mathbf{0}$	All > 0	Leading ones all positive.
Positive semi-definite	$\mathbf{x}^t \mathrm{A} \mathbf{x} \geq 0 \ \forall \mathbf{x}$	All ≥ 0	All non-negative.
Negative definite	$\mathbf{x}^t \mathrm{A} \mathbf{x} < 0 \ \forall \mathbf{x} \neq \mathbf{0}$	All < 0	jth l.p.m. has sign of $(-1)^j$
Negative semi-definite	$\mathbf{x}^t \mathrm{A} \mathbf{x} \leq 0 \ \forall \mathbf{x}$	All ≤ 0	Those of $-\mathrm{A}$ non-negative
Indefinite	None of above	Some +, some −	None of the above

5.2.1 Examples

1. $\mathrm{A} = \begin{pmatrix} -3 & 2 \\ 2 & -4 \end{pmatrix}$.

2. $\mathrm{B} = \begin{pmatrix} 3 & -6 \\ -6 & 12 \end{pmatrix}$.

3. $\mathrm{C} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 3 & -1 \\ -2 & -1 & 5 \end{pmatrix}$.

4. $D = \begin{pmatrix} 4 & 2 & -3 \\ 2 & -1 & 0 \\ -3 & 0 & 3 \end{pmatrix}.$

Proposition 5.4 (Second order sufficient conditions for local optima.)
Let $\mathrm{f}(\mathbf{x})$ *be a twice-differentiable function on* $S \subset \mathbb{R}^n$*. Suppose* $\nabla \mathrm{f}(\mathbf{x}^*) = \mathbf{0}$*.*

(i) If $\mathrm{H}(\mathrm{f}(\mathbf{x}^*))$ *is positive definite then* $\mathbf{x}^*$ *is a strong local minimizer of* $\mathrm{f}(\mathbf{x})$*.*

(ii) If $\mathrm{H}(\mathrm{f}(\mathbf{x}^*))$ *is negative definite then* $\mathbf{x}^*$ *is a strong local maximizer of* $\mathrm{f}(\mathbf{x})$*.*

(iii) If $\mathrm{H}(\mathrm{f}(\mathbf{x}^*))$ *is indefinite and non-singular then* $\mathrm{f}(\mathbf{x})$ *has a saddle-point at* $\mathbf{x}^*$*.*

Proof of (i)

Let $\mathbf{u}$ be any feasible direction at $\mathbf{x}^*$. Let $\mathrm{g}(r) = \mathrm{f}(\mathbf{x}^* + r\mathbf{u})$, so $\mathrm{g}(0) = \mathrm{f}(\mathbf{x}^*)$.

By Proposition 5.1, $\mathrm{g}'(0) = \mathbf{u}^t \nabla \mathrm{f}(\mathbf{x}^*)$ and $\mathrm{g}''(0) = \mathbf{u}^t \mathrm{H}(\mathrm{f}(\mathbf{x}^*))\mathbf{u}$.

Thus if $\nabla \mathrm{f}(\mathbf{x}^*) = \mathbf{0}$ then $\mathrm{g}'(0) = 0$.

If also $\mathrm{H}(\mathrm{f}(\mathbf{x}^*))$ is positive definite, i.e. $\mathbf{u}^t \mathrm{H}(\mathrm{f}(\mathbf{x}^*))\mathbf{u} > 0$ for all $\mathbf{u} \neq \mathbf{0}$, then $\mathrm{g}''(0) > 0$

Hence $r = 0$ is a strong local minimizer of $\mathrm{g}(r)$, so $\mathrm{g}(0) < \mathrm{g}(r)$ for all r close to 0.

Equivalently, $\mathrm{f}(\mathbf{x}^*) < \mathrm{f}(\mathbf{x}^* + r\mathbf{u})$ for small enough r and all feasible directions $\mathbf{u}$. Thus $\mathbf{x}^*$ is a strong local minimizer of $\mathrm{f}(\mathbf{x})$ over S. □

5.2.2 Examples

(i) $f(\mathbf{x}) = xyz - x^2 - y^2 - z^2$.

(ii) $f(\mathbf{x}) = 5x + 4y + 2z - 2x^2 - 3y^2 - 5z^2 - 2xy - 6xz - 4yz$.

5.3 Convex and concave functions

A twice-differentiable function f defined on an interval $S \subset \mathbb{R}$ is **convex** on S if $\mathrm{f}''(x) \geq 0$ for all $x \in S$, and **concave** on S if $\mathrm{f}''(x) \leq 0$ for all $x \in S$. (Note: *this* definition of convexity applies only to functions of *one* variable. The general definition is given below.)

Examples of *convex* functions of one variable x are: $x^2, x^4, \mathrm{e}^x, \cosh x$.

Examples of *concave* functions of one variable x are: $\sqrt{x}, \ln x, \mathrm{e}^{-x}, \arctan x$.

Geometrically, if S is the interval $[a, b]$, a convex function f lies *below* (or on) the straight line joining A $(a,\ \mathrm{f}(a))$ and B $(b,\ \mathrm{f}(b))$.

If $c \in S$, then $c = (1-r)a + rb$ for some r with $0 \leq r \leq 1$. Then $\mathrm{f}(c) = \mathrm{f}((1-r)a + rb)$, while the point on the straight line AB at which $x = c$ has y-coordinate $(1-r)\mathrm{f}(a) + r\mathrm{f}(b)$. Thus $\mathrm{f}((1-r)a + rb) \leq (1-r)\mathrm{f}(a) + r\mathrm{f}(b)$ for all $r \in [0, 1]$.

For functions of any number of variables, the definitions are as follows:

Let f be a real-valued function defined on a *convex* subset S of $\mathbb{R}^n$.

f is a **convex function** on S if for any $\mathbf{x}$, $\mathbf{y} \in S$,

$$\mathrm{f}((1-r)\mathbf{x} + r\mathbf{y}) \leq (1-r)\mathrm{f}(\mathbf{x}) + r\mathrm{f}(\mathbf{y}) \text{ for all } r \in [0, 1].$$

f is a **concave function** if for any $\mathbf{x}$, $\mathbf{y} \in S$,

$$\mathrm{f}((1-r)\mathbf{x} + r\mathbf{y}) \geq (1-r)\mathrm{f}(\mathbf{x}) + r\mathrm{f}(\mathbf{y}) \text{ for all } r \in [0, 1].$$

f is a **strictly** $\left\{\begin{matrix}\textbf{convex}\\ \textbf{concave}\end{matrix}\right\}$ function on S if for any $\mathbf{x}$, $\mathbf{y} \in S$ with $\mathbf{x} \neq \mathbf{y}$,

$$\mathrm{f}((1-r)\mathbf{x} + r\mathbf{y}) \quad \left\{\begin{matrix}<\\ >\end{matrix}\right\} \ (1-r)\mathrm{f}(\mathbf{x}) + r\mathrm{f}(\mathbf{y}) \text{ for all } r \in (0, 1).$$

Clearly f is concave if and only if $-\mathrm{f}$ is convex.

If f is a *convex* function of one or two variables, its graph lies *below* (or on) the straight line joining any two points on it, so the region *above* it is a convex set.

If f is a *concave* function of one or two variables, its graph lies *above* (or on) the straight line joining any two points on it, so the region *below* it is a convex set.

- Any *linear* function is both convex and concave, but not strictly so.
- A positive multiple of a convex / concave function is itself convex / concave. For example, x^2 is convex on $\mathbb{R}$, so ax^2 is convex on $\mathbb{R}$ for any $a > 0$.
- A sum of convex / concave functions is itself convex / concave. For example, x^2, y^4 and z^6 are each convex, so $x^2 + y^4 + z^6$ is convex on $\mathbb{R}^3$.

It may not be clear from the definitions whether a function of several variables is convex, concave or neither. For functions of more than one variable it is *not* enough just to consider the second derivatives. We therefore derive criteria which are easier to use.

Proposition 5.5 *A twice-differentiable function* f *defined on a convex set* $S \subset \mathbb{R}^n$ *is*

(i) convex on S *if and only if, at each* $\mathbf{x} \in S$, $\mathrm{H}(\mathrm{f}(\mathbf{x}))$ *is positive semi-definite;*

(ii) concave on S *if and only if, at each* $\mathbf{x} \in S$, $\mathrm{H}(\mathrm{f}(\mathbf{x}))$ *is negative semi-definite;*

(iii) strictly convex on S if, at each $\mathbf{x} \in S$, $\mathrm{H}(\mathrm{f}(\mathbf{x}))$ *is positive definite;*

(iv) strictly concave on S if, at each $\mathbf{x} \in S$, $\mathrm{H}(\mathrm{f}(\mathbf{x}))$ *is negative definite.*

(Note that 'semi-definite' includes 'definite'.)

Proof: We prove the 'if' part of (i).

Let $\mathbf{x}$ and $\mathbf{y}$ be in S and let $\mathbf{u} = \mathbf{y} - \mathbf{x}$.

For $0 \leq r \leq 1$, let $\mathbf{z} = (1-r)\mathbf{x} + r\mathbf{y} = \mathbf{x} + r\mathbf{u}$. As S is a convex set, $\mathbf{z} \in S$.

Let $\mathrm{g}(r) = \mathrm{f}(\mathbf{z})$, so $\mathrm{g}(0) = \mathrm{f}(\mathbf{x})$ and $\mathrm{g}(1) = \mathrm{f}(\mathbf{y})$.

As in the proof of Proposition 5.1, $\mathrm{g}''(r) = \mathbf{u}^t\mathrm{H}(\mathrm{f}(\mathbf{z}))\mathbf{u}$.

If $\mathrm{H}(\mathrm{f}(\mathbf{x}))$ is positive semi-definite for all $\mathbf{x} \in S$ then $\mathbf{u}^t\mathrm{H}(\mathrm{f}(\mathbf{z}))\mathbf{u} \geq 0$.

Thus $\mathrm{g}''(r) \geq 0$ for $0 \leq r \leq 1$ so g is a convex function on $[0,1]$.

$\mathrm{f}((1-r)\mathbf{x} + r\mathbf{y}) = \mathrm{f}(\mathbf{z}) = \mathrm{g}(r) = \mathrm{g}((1-r)(0) + r(1)) \leq (1-r)\mathrm{g}(0) + r\mathrm{g}(1)$ (as g is convex)

$= (1-r)\mathrm{f}(\mathbf{x}) + r\mathrm{f}(\mathbf{y})$. Thus f is convex on S. □

Proposition 5.6 *Let* f *be a* $\left\{\begin{array}{c}\textit{convex}\\\textit{concave}\end{array}\right\}$ *function on a convex set $S \subset \mathbb{R}^n$. If there is a point $\mathbf{x}^* \in S$ for which* $\nabla\mathrm{f}(\mathbf{x}^*) = \mathbf{0}$ *then* $\mathbf{x}^*$ $\left\{\begin{array}{c}\textit{minimizes}\\\textit{maximizes}\end{array}\right\}$ $\mathrm{f}(\mathbf{x})$ *globally over S.*

Proof (Convex case)

Suppose $\nabla\mathrm{f}(\mathbf{x}^*) = \mathbf{0}$. Let $\mathbf{u}$ be any vector such that $\mathbf{y} = \mathbf{x}^* + \mathbf{u} \in S$.

$\mathbf{x}^* + r\mathbf{u} = (1-r)\mathbf{x}^* + r\mathbf{y} \in S$ for all $r \in [0,1]$, as S is convex.

Let $\mathrm{g}(r) = \mathrm{f}(\mathbf{x}^* + r\mathbf{u})$, so $\mathrm{g}(0) = \mathrm{f}(\mathbf{x}^*)$ and $\mathrm{g}(1) = \mathrm{f}(\mathbf{y})$.

Assume f is convex, so $\mathrm{f}((1-r)\mathbf{x}^* + r\mathbf{y}) \leq (1-r)\mathrm{f}(\mathbf{x}^*) + r\mathrm{f}(\mathbf{y})$.

Thus $\mathrm{g}(r) \leq (1-r)\mathrm{g}(0) + r\mathrm{g}(1)$, so if $r > 0$ then $\dfrac{\mathrm{g}(r) - \mathrm{g}(0)}{r} \leq \mathrm{g}(1) - \mathrm{g}(0)$.

As $r \to 0$, the left hand side tends to $\mathrm{g}'(0)$. But $\mathrm{g}'(0) = \mathbf{u}^t\nabla\mathrm{f}(\mathbf{x}^*) = 0$ since $\nabla\mathrm{f}(\mathbf{x}^*) = \mathbf{0}$.

Hence $0 \leq \mathrm{g}(1) - \mathrm{g}(0)$ so $\mathrm{g}(1) \geq \mathrm{g}(0)$, i.e. $\mathrm{f}(\mathbf{y}) \geq \mathrm{f}(\mathbf{x}^*)$, for all $\mathbf{y} \in S$.

Thus $\mathbf{x}^*$ minimizes $\mathrm{f}(\mathbf{x})$ globally over S. □

It can also be shown that every local $\left\{\begin{array}{c}\text{minimizer}\\\text{maximizer}\end{array}\right\}$ of a $\left\{\begin{array}{c}\text{convex}\\\text{concave}\end{array}\right\}$ function is global, even if it does not occur at a critical point.

For example, a concave function might have a local maximizer which is not at a critical point, on the boundary of its domain. This would have to be a global maximizer.

5.3.1 Examples

(i) $\mathrm{f}(x, y) = 3x^2 + 4y^2 - 5xy + 2x - 3y$.

(ii) $\mathrm{f}(x, y, z) = \ln(xy) - \dfrac{y}{z}$ where $x, y, z > 0$. (iii) Let $\mathrm{f}(x, y) = x^4 + y^4$.

Exercises 5

1. Express each of the following in the form $\mathbf{x}^t \mathrm{A}\mathbf{x}$, where A is a *symmetric* matrix.

   (a) $4x^2 + 3y^2$, (b) $-3x^2 + 4xy - 6y^2$, (c) $x^2 + y^2 + z^2$,

   (d) $3x^2 + 2y^2 + z^2 + 4xy + 4yz$, (e) $3x^2 + 5y^2 + 2z^2 + 4xy - 2xz - 4yz$.

   Determine whether each is positive or negative definite or semi-definite or indefinite.

2. Using the Hessian matrices where appropriate, find and classify the local optima of

   (a) $2x^2 - \ln|2x|$, (b) $3x^2 + 4xy - 8y$, (c) $x^3 + 4xy - 6y^2$,

   (d) $x^2 + 3y^2 + 4z^2 - 2xy - 2xz + 3yz - y$, (e) $x^3 + y^3 + z^3 - 9xy - 9xz + 27x$.

3. Determine whether each of the following is convex, concave or neither on S.

   (a) $\mathrm{f}(x) = \frac{1}{x}$, $S = (0, \infty)$, (b) $\mathrm{f}(x) = x^\alpha$, $0 \leq \alpha \leq 1$, $S = (0, \infty)$,

   (c) $\mathrm{f}(x, y) = x^3 + 3xy + y^2, S = \mathbb{R}^2$, (d) $\mathrm{f}(x, y, z) = -x^2 - y^2 - 2z^2 + \frac{1}{2}xy, S = \mathbb{R}^3$.

4. Let $\mathrm{f}(x, y, z) = x^2 - y^2 - xy - x^3$

   Find S, the largest convex subset of $\mathbb{R}^2$ on which f is a concave function.

   Find, with justification, the global maximum value of f over S.

5. Find the values of x, y and z which minimize $\mathrm{f}(x, y, z) = x^2 + 2y^2 + 3z^2 + 2xy + 2xz - 4y$ over $\mathbb{R}^3$. Justify that the point you find is a global minimizer.

6. The function f is *quasi-convex* on $\mathbb{R}^n$ if for every real number a, the set $S = \{\mathbf{x} \in \mathbb{R}^n : \mathrm{f}(\mathbf{x}) \leq a\}$ is a convex set. Prove that any convex function on $\mathbb{R}^n$ is quasi-convex.

Chapter 6

Constrained optimization

6.1 Lagrange multipliers

Consider the constrained optimization problem:

$$\text{maximize or minimize } \mathrm{f}(\mathbf{x}) = \mathrm{f}(x_1, \ldots, x_n)$$
$$\text{subject to } \mathrm{g}_j(\mathbf{x}) = b_j \text{ for } j = 1, \ldots, m,$$
$$\text{where } \mathbf{x} \in S \subset \mathbb{R}^n.$$

Here, f and $\mathrm{g}_1, \ldots, \mathrm{g}_m$ are functions from $\mathbb{R}^n$ to $\mathbb{R}$, and $m < n$.

Let $\mathbf{g}(\mathbf{x}) = (\mathrm{g}_1(\mathbf{x}) \ \cdots \ \mathrm{g}_m(\mathbf{x}))^t$, $\mathbf{b} = (b_1 \ \cdots \ b_m)^t$, so the constraints can be expressed as $\mathbf{g}(\mathbf{x}) = \mathbf{b}$.

Let $\boldsymbol{\lambda} = (\lambda_1 \ \cdots \ \lambda_m)^t$. The **Lagrangean function** for the optimization problem is

$$\mathcal{L}(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) = \mathrm{f}(x_1, \ldots, x_n) + \sum_{j=1}^{m} \lambda_j (b_j - \mathrm{g}_j(x_1, \ldots, x_n))$$

which can be written as $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \mathrm{f}(\mathbf{x}) + \boldsymbol{\lambda}^t(\mathbf{b} - \mathbf{g}(\mathbf{x}))$.

$\lambda_1, \ldots, \lambda_m$ are called **Lagrange multipliers**. There is one for each constraint.

6.1.1 Example

Suppose we have to maximize $\mathrm{f}(x_1, x_2, x_3) = x_1{}^2 + 2x_2x_3$ subject to the constraints $x_1 + x_2 = 5$, $x_1{}^2 - x_2{}^2 = 4$ and $x_1x_2x_3 = 7$.

Proposition 6.1 (Lagrange sufficiency theorem)
(i) If there exist $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ *and* $\mathbf{x}^* \in S$ *such that* $\mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) \geq \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^*)$ *for all* $\mathbf{x} \in S$*, and* $\mathbf{g}(\mathbf{x}^*) = \mathbf{b}$*, then* $\mathbf{x}^*$ *maximizes* $\mathrm{f}(\mathbf{x})$ *over* S *subject to* $\mathbf{g}(\mathbf{x}) = \mathbf{b}$.

(ii) If there exist $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ *and* $\mathbf{x}^* \in S$ *such that* $\mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) \leq \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^*)$ *for all* $\mathbf{x} \in S$*, and* $\mathbf{g}(\mathbf{x}^*) = \mathbf{b}$*, then* $\mathbf{x}^*$ *minimizes* $\mathrm{f}(\mathbf{x})$ *over* S *subject to* $\mathbf{g}(\mathbf{x}) = \mathbf{b}$.

Proof of (i) [the proof of (ii) is similar]

Assume $\mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) \geq \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^*)$ for all $\mathbf{x} \in S$,

so $\mathrm{f}(\mathbf{x}^*) + (\boldsymbol{\lambda}^*)^t(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) \geq \mathrm{f}(\mathbf{x}) + (\boldsymbol{\lambda}^*)^t(\mathbf{b} - \mathbf{g}(\mathbf{x}))$ for all $\mathbf{x} \in S$.

By assumption $\mathbf{g}(\mathbf{x}^*) = \mathbf{b}$, so $\mathbf{b} - \mathbf{g}(\mathbf{x}^*) = \mathbf{0}$.

Hence $\mathrm{f}(\mathbf{x}^*) \geq \mathrm{f}(\mathbf{x})$ for all $\mathbf{x} \in S$ with $\mathbf{g}(\mathbf{x}) = \mathbf{b}$,

so $\mathbf{x}^*$ maximizes $\mathrm{f}(\mathbf{x})$ over S subject to $\mathbf{g}(\mathbf{x}) = \mathbf{b}$. □

Now suppose f and each g_i are differentiable and that a constrained optimum of $\mathrm{f}(\mathbf{x})$ occurs at an interior point $\mathbf{x}^*$ in S.

If f and the g_i are fixed, we can treat $\mathbf{x}^*$ as a function of $\mathbf{b}$: call it $\mathbf{x}^*(\mathbf{b})$.

Let $\mathrm{v}(\mathbf{b}) = \mathrm{f}(\mathbf{x}^*(\mathbf{b}))$ and let $\mathrm{h}(\mathbf{x}) = \mathrm{f}(\mathbf{x}) - \mathrm{v}(\mathbf{g}(\mathbf{x}))$.

Then where the constraints hold, $\mathrm{h}(\mathbf{x}) = \mathrm{f}(\mathbf{x}) - \mathrm{v}(\mathbf{b})$. In particular, for any $\mathbf{b}$, $\mathrm{h}(\mathbf{x}^*(\mathbf{b})) = \mathrm{f}(\mathbf{x}^*(\mathbf{b})) - \mathrm{v}(\mathbf{b}) = 0$.

For each $\mathbf{x}$, $\mathrm{v}(\mathbf{g}(\mathbf{x}))$ is the optimal value of $\mathrm{f}(\mathbf{y})$ subject to $\mathbf{g}(\mathbf{y}) = \mathbf{g}(\mathbf{x})$. As $\mathbf{x}$ is feasible for this problem, $\mathrm{v}(\mathbf{g}(\mathbf{x})) \geq \mathrm{f}(\mathbf{x})$, i.e. $\mathrm{h}(\mathbf{x}) \leq 0$, for all $x \in S$ if the problem is a maximization, and $\mathrm{v}(\mathbf{g}(\mathbf{x})) \leq \mathrm{f}(\mathbf{x})$, i.e. $\mathrm{h}(\mathbf{x}) \geq 0$, for all $x \in S$ if the problem is a minimization.

Thus $\mathbf{x}^*$ is an interior point of S at which $\mathrm{h}(\mathbf{x})$ is either maximized or minimized, so $\nabla\mathrm{h}(\mathbf{x}^*) = 0$.

Now $\nabla\mathrm{h}(\mathbf{x}) = \nabla\mathrm{f}(\mathbf{x}) - \nabla\mathrm{v}(\mathbf{g}(\mathbf{x}))$.

By the chain rule, $\dfrac{\partial \mathrm{h}(\mathbf{x})}{\partial x_i} = \dfrac{\partial \mathrm{f}(\mathbf{x})}{\partial x_i} - \sum_{j=1}^{m} \dfrac{\partial \mathrm{v}(\mathbf{g}(\mathbf{x}))}{\partial \mathrm{g}_j(\mathbf{x})} \dfrac{\partial \mathrm{g}_j(\mathbf{x})}{\partial x_i} = \dfrac{\partial \mathrm{f}(\mathbf{x})}{\partial x_i} - \sum_{j=1}^{m} \lambda_j(\mathbf{x}) \dfrac{\partial \mathrm{g}_j(\mathbf{x})}{\partial x_i}$

where $\lambda_j(\mathbf{x}) = \dfrac{\partial \mathrm{v}(\mathbf{g}(\mathbf{x}))}{\partial \mathrm{g}_j(\mathbf{x})}$.

Let ${\lambda_j}^* = \lambda_j(\mathbf{x}^*)$. As $\nabla \mathrm{h}(\mathbf{x}^*) = 0$, we have $\dfrac{\partial \mathrm{f}}{\partial x_i}(\mathbf{x}^*) = \sum_{j=1}^{m} {\lambda_j}^* \dfrac{\partial \mathrm{g}_j}{\partial x_i}(\mathbf{x}^*)$,

so $\nabla \mathrm{f}(\mathbf{x}^*) = \sum_{j=1}^{m} {\lambda_j}^* \nabla \mathrm{g}_j(\mathbf{x}^*)$. Equivalently, $\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \mathbf{0}$, where $\nabla_{\mathbf{x}}$ denotes the gradient vector with respect to $(x_1, \ldots, x_n)$.

If the m vectors $\nabla \mathrm{g}_j(\mathbf{x}^*)$ form a linearly independent set then $\nabla \mathrm{f}(\mathbf{x}^*)$ can be expressed as a linear combination of them in only one way, so the multipliers ${\lambda_j}^*$ are uniquely determined. Hence we have:

Proposition 6.2 (Lagrange necessity theorem)
Let $\mathbf{x}^*$ be an interior point of a set S which is a local maximizer or minimizer of $\mathrm{f}(\mathbf{x})$ over S subject to $\mathbf{g}(\mathbf{x}) = \mathbf{b}$. Suppose the set $\{\nabla \mathrm{g}_1(\mathbf{x}^*), \ldots, \nabla \mathrm{g}_m(\mathbf{x}^*)\}$ is linearly independent. Let $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \mathrm{f}(\mathbf{x}) + \boldsymbol{\lambda}^t(\mathbf{b} - \mathbf{g}(\mathbf{x}))$.

Then there is a unique vector $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ such that $\mathbf{x}^*, \boldsymbol{\lambda}^*$ satisfy $\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \mathbf{0}$.

6.1.2 Examples

(i) Maximise $\mathrm{f}(x, y) = xy$ over $\mathbb{R}^2$ subject to $\mathrm{g}(x, y) = x + y = 8$.

(ii) Find the optimal value(s) of the function $\mathrm{f}(x, y, z) = 4x^{1/4}y^{1/4}z^{1/4}$ over $\mathbb{R}^3_+$ subject to the constraints $x + y = 3$, $y + z = 3$.

(iii) Minimize $\mathrm{f}(x, y, z) = x^2 + y^2 + z^2$ over $\mathbb{R}^3$, subject to $x + 2y + z = 1$ and $2x - y - 3z = 4$.
Let $\mathcal{L}(x, y, z, \lambda_1, \lambda_2) = x^2 + y^2 + z^2 + \lambda_1(1 - x - 2y - z) + \lambda_2(4 - 2x + y + 3z)$.

Sensitivity analysis

In the proof of the Lagrange necessity theorem, we defined $\lambda_j(\mathbf{x})$ to be $\dfrac{\partial \mathrm{v}(\mathbf{g}(\mathbf{x}))}{\partial \mathrm{g}_j(\mathbf{x})}$.

Thus $\lambda_j{}^* = \dfrac{\partial \mathrm{v}(\mathbf{g}(\mathbf{x}^*))}{\partial \mathrm{g}_j(\mathbf{x}^*)} = \dfrac{\partial \mathrm{v}(\mathbf{b})}{\partial b_j}$,

i.e. the optimal value of the Lagrange multiplier for the jth constraint is equal to the rate of change in the maximal value of the objective function as the jth constraint is relaxed.

If an increase δb_j in the right-hand side of a constraint yields an increase δV in the optimal value of $\mathrm{f}(\mathbf{x})$ then $\delta V \approx \lambda_j{}^* \, \delta b_j$.

If the constraints arise because of limits on some resources, then $\lambda_j{}^*$ is called the **shadow price** of the jth resource.

If *all* the b_j can vary, $\delta V \approx \sum_{j=1}^{m} \dfrac{\partial V}{\partial b_j} \delta b_j = \sum_{j=1}^{m} \lambda_j{}^* \, \delta b_j$. Thus we have:

Proposition 6.3 *Suppose* $\mathbf{x}^*$ *optimizes* $\mathrm{f}(\mathbf{x})$ *over* S *subject to* $\mathrm{g}_j(\mathbf{x}) = b_j$ *for* $j = 1, \ldots, m$. *Let* $\lambda_1{}^*, \ldots, \lambda_m{}^*$ *be the values of the Lagrange multipliers at the optimal point.*

If each b_j *is increased by a small amount* δb_j*, then the increase in the optimal value of* $\mathrm{f}(\mathbf{x})$ *is approximately* $\sum_{j=1}^{m} \lambda_j{}^* \, \delta b_j$.

In Example 6.1.2 (ii), $\mathrm{f}(\mathbf{x}^*) = 4\sqrt{2}, b_1 = b_2 = 3, \lambda_1{}^* = \lambda_2{}^* = \dfrac{1}{\sqrt{2}}$.

If b_1 and b_2 are slightly increased, say to 3.1 and 3.2 respectively, we would expect the constrained maximum to be increased by about $0.3 \times \dfrac{1}{\sqrt{2}}$, to approximately 5.87.

6.2 Constrained optimization of quadratic forms

Let A be a symmetric real matrix. In this section we seek to optimize the quadratic form $\mathrm{q}(\mathbf{x}) = \mathbf{x}^t \mathrm{A} \mathbf{x}$ subject to the non-linear constraint $\mid \mathbf{x} \mid = 1$, which can also be written as $\mathbf{x}^t \mathbf{x} = 1$ or $x_1{}^2 + \cdots + x_n{}^2 = 1$.

First we note that $\nabla(\mathbf{x}^t \mathbf{x}) = \nabla(x_1{}^2 + \cdots + x_n{}^2) = (2x_1, \ldots, 2x_n) = 2\mathbf{x}$

and if A is symmetric, $\nabla(\mathbf{x}^t \mathrm{A} \mathbf{x}) = \nabla \left(\sum_{i,j=1}^{n} a_{ij} x_i x_j \right)$. The partial derivative of this with respect to x_i is $2a_{ii}x_i + 2 \sum_{j \neq i} a_{ij} x_j = 2 \sum_{j=1}^{n} a_{ij} x_j$, so $\nabla(\mathbf{x}^t \mathrm{A} \mathbf{x}) = 2\mathrm{A}\mathbf{x}$ when A is symmetric.

Proposition 6.4 *Let* A *be a symmetric real matrix and let* $q(\mathbf{x}) = \mathbf{x}^t \mathrm{A} \mathbf{x}$.

The minimum value of $q(\mathbf{x})$ *subject to* $\mid \mathbf{x} \mid = 1$ *is equal to the smallest eigenvalue of* A*, and occurs when* $\mathbf{x}$ *is a corresponding unit eigenvector.*

The maximum value of $q(\mathbf{x})$ *subject to* $\mid \mathbf{x} \mid = 1$ *is equal to the largest eigenvalue of* A*, and occurs when* $\mathbf{x}$ *is a corresponding unit eigenvector.*

Proof

The constraint is $\mathbf{x}^t \mathbf{x} = 1$, i.e. $1 - \mathbf{x}^t \mathbf{x} = 0$.

The Lagrangean function is $\mathcal{L}(\mathbf{x}, \lambda) = \mathbf{x}^t \mathrm{A} \mathbf{x} + \lambda(1 - \mathbf{x}^t \mathbf{x})$.

For $\mathrm{q}(\mathbf{x})$ to be optimal subject to the constraint, $\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda) = \mathbf{0}$.

Hence $2\mathrm{A}\mathbf{x} - 2\lambda\mathbf{x} = 0$, giving $\mathrm{A}\mathbf{x} = \lambda\mathbf{x}$. Also $\mathbf{x}^t\mathbf{x} = 1$.

Thus λ is an eigenvalue of A, with eigenvector $\mathbf{x}$ such that $\mathbf{x}^t\mathbf{x} = 1$, i.e. $\mathbf{x}$ is a unit vector.

So the *necessary* condition for an optimum of $\mathrm{q}(\mathbf{x})$ is satisfied when $\mathbf{x}$ is such an eigenvector of A. Then $\mathrm{q}(\mathbf{x}) = \mathbf{x}^t\mathrm{A}\mathbf{x} = \mathbf{x}^t(\lambda\mathbf{x}) = \lambda(\mathbf{x}^t\mathbf{x}) = \lambda$.

The feasible set is compact, so by Weierstrass's Theorem the minimum and maximum exist. Thus they must be among the values we have found, which are in $\mathbb{R}$ since the eigenvalues of a symmetric real matrix are real. The result follows. □

6.2.1 Example

Find the maximum and minimum values of $q = x^2+z^2+4xz-6yz$ subject to $x^2+y^2+z^2 = 1$.

6.2.2 Example

A council plans to repair x hundred miles of roads and improve y hundred acres of parks. Budget restrictions lead to the constraint $4x^2 + 9y^2 = 36$. The benefit obtained from the possible work schedules is $U = xy$. Find the schedule that maximizes U.

Exercises 6

1. A firm has to minimize its cost function $C(x, y) = rx + wy$ subject to the constraint $x^{1/2}y^{1/4} = 8$. Find the minimum cost in terms of the constants r and w.

2. Find (i) the minimum value of $x^2 + y^2$ subject to the constraint $2x - 3y = 4$,

   (ii) the maximum value of x^2y^2 subject to the constraint $2x^2 + y^2 = 3$,

   Interpret your answers graphically.

3. $\mathrm{f}(x, y, z) = x^{1/3}y^{1/3}z^{1/3}$ where $x > 0, y > 0, z > 0$.

   Maximize $\mathrm{f}(x, y, z)$ subject to the constraints $x + y = 3$ and $y + z = 3$.

4. The total profit £z thousand which a company makes from producing and selling x thousand units of one commodity X and y thousand units of another commodity Y is given by $z = 10 + 50x - 5x^2 + 16y - y^2$.

   (a) Find the maximum value of z if there are no constraints on x and y, explaining why the value you find is a maximum.

   (b) Find the maximum value of z if the total cost of production is to be £12,000, given that each unit of X costs £4 to produce and each unit of Y costs £3.20.

   (c) The company now increases the money available for production to £12,500. Use sensitivity analysis to estimate the new maximum profit.

5. A firm's production function, using quantities x, y, z of three inputs, is defined by $P(x, y, z) = x^{\frac{1}{2}} \ln y - z^2$, where $x > 0, y > 0, z > 0$.

   Find the largest region of $\mathbb{R}^3_+$ on which P is a concave function.

   Maximize $P(x, y, z)$ over this region subject to the constraints $x + y - z = \mathrm{e}^4$ and $-x + 2y + z = 2\mathrm{e}^4$.

6. A consumer's satisfaction with three foods is measured by $U(x_1, x_2, x_3) = x_1 + \ln x_2 + 2x_3$, where $x_1 > 0, x_2 > 0, x_3 > 0$ and x_i is the number of units of food i consumed.

   Foods 1, 2 and 3 cost £2, £1, £0 per unit and contain 0, 100 and 200 calories per unit respectively.

   The consumer wants to spend exactly £10 and consume 1000 calories.

   Find the maximum value of $U(x_1, x_2, x_3)$. Justify that the value you have found is a maximum.

   Using sensitivity analysis, estimate the increase in this maximum value if an extra £1 may be spent and an extra 50 calories consumed.

7. Use Proposition 6.4 to find the maximum and minimum values of the following quadratic forms subject to the constraint $\sum {x_i}^2 = 1$. Also find the values of the variables x_i at which the optimal values are attained.

   (a) $q(x_1, x_2) = 5{x_1}^2 + 5{x_2}^2 - 4x_1x_2$, (b) $q(x_1, x_2) = 7{x_1}^2 + 3{x_2}^2 + 3x_1x_2$,

   (c) $q(x_1, x_2, x_3) = -2{x_1}^2 - {x_2}^2 + 4x_1x_2 + 4x_2x_3$.

Sample Class test

1. A predator needs a daily intake of at least 10 units of food A, 12 units of food B and 12 units of food C. These requirements are satisfied by feeding on two species of prey.

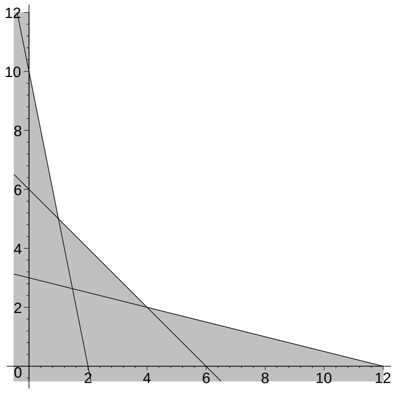
   One animal of species 1 contains 5 units of A, 2 units of B amd 1 unit of C.

   One animal of species 2 contains 1 unit of A, 2 units of B and 4 units of C.

   To catch and digest each animal takes 3 units of energy for species 1 and 2 units of energy for species 2. The predator needs to minimise its daily expenditure of energy.

   (a) If x_1 animals of species 1 and x_2 of species 2 are consumed in a day, formulate the above information as a linear programming problem. [4]

   (b) The feasible region for the problem is shown in the diagram:



   From the graph, find how many of each species the predator should eat to satisfy its daily requirements with the minimum expenditure of energy. State the optimal number of units of energy used. [4]

   (c) Species 2 is becoming tougher, and to catch and digest one animal now requires q units of energy. How large does q have to become before the optimal feeding plan changes? State how many of each species should be consumed per day when q slightly exceeds this value. [4]

2. $z = 3x_1 + 2x_2 + 5x_3$ is to be maximized subject to the constraints

$$\begin{array}{rcrcrcl} x_1 & + & 2x_2 & + & x_3 & \leq & 430 \\ 3x_1 & + & & + & 2x_3 & \leq & 460 \\ x_1 & + & 4x_2 & & & \leq & 420 \end{array}$$

where $x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$.

   (a) Write the constraints in standard form. [2]

   (b) Using the simplex algorithm, show that the maximum value of z is 1350. [10]

   (c) State the values of x_1, x_2 and x_3 at which the maximum occurs. [2]

   (d) Express z in terms of the variables which are non-basic in the final tableau. Hence explain why the solution is optimal. [4]

(e) B^{-1} is the 3×3 matrix which pre-multiplies the middle three rows of the initial tableau to give those rows in the final tableau. Write down B^{-1} and B. [4]

(f) If the first constraint is changed to $x_1 + 2x_2 + x_3 \leq k$, find the range of values of k for which the same variables remain basic at the optimum. If k lies in this range, express the new optimal value of z in terms of k. [6]

3. Give definitions of the following:

(a) a **non-negative vector** in $\mathbb{R}^n$, [2]

(b) a **convex** subset of $\mathbb{R}^n$, [2]

(c) an **extreme point** of a convex set, [2]

(d) a **bounded** subset of $\mathbb{R}^n$, [2]

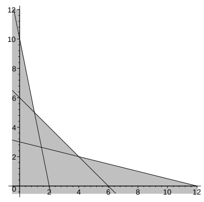
(e) an algorithm of **exponential complexity**. [2]

[TOTAL MARKS: 50]

Solution to Sample Class test

1. (a) **Minimise** $z = 3x_1 + 2x_2$
   subject to $5x_1 + x_2 \geq 10,\ 2x_1 + 2x_2 \geq 12,\ x_1 + 4x_2 \geq 12$
   and $x_1 \geq 0, x_2 \geq 0$.

   (b) Feasible region:



   By considering lines of the form $3x_1 + 2x_2 = c$ we see that z is minimum where $5x_1 + x_2 = 10$ meets $2x_1 + 2x_2 = 12$, i.e. at (1, 5).
   The predator should eat 1 of species 1 and 5 of species 2, using 13 units of energy.

   (c) Increasing the coefficient of x_2 makes the constraint line less steep. When it is parallel to $2x_1 + 2x_2 = 12$, the optimal point changes. This occurs when $q = 3$. Then the optimal point becomes (4, 2), so eat 4 of species 1 and 2 of species 2.

2. (a) In standard form, constraints are
   $3x_1 + 2x_2 + x_3 + x_4 = 430,\ 3x_1 + 2x_3 + x_5 = 460,\ x_1 + 4x_2 + x_6 = 420.$
   where $x_1, \ldots, x_6 \geq 0$.

   (b) Write objective function as $z - 3x_1 - 2x_2 - 5x_3 = 0$.

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	Solution	θ_i
x_4	0	1	2	1	1	0	0	430	430
x_5	0	3	0	2	0	1	0	460	230
x_6	0	1	4	0	0	0	1	420	
z	1	-3	-2	-5	0	0	0	0	

   Minimum $\theta_i = 230$. x_3 enters basis, x_5 leaves.

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	Solution	θ_i
x_4	0	$-1/2$	2	0	1	$-1/2$	0	200	100
x_3	0	$3/2$	0	1	0	$1/2$	0	230	
x_6	0	1	4	0	0	0	1	420	105
z	1	$9/2$	-2	0	0	$5/2$	0	1150	

   Minimum $\theta_i = 100$. x_2 enters basis, x_4 leaves.

Basic	z	x_1	x_2	x_3	x_4	x_5	x_6	Solution
x_2	0	$-1/4$	1	0	$1/2$	$-1/4$	0	100
x_3	0	$3/2$	0	1	0	$1/2$	0	230
x_6	0	2	0	0	-2	1	1	20
z	1	4	0	0	1	2	0	1350

   (c) Optimum occurs when $x_1 = 0, x_2 = 100, x_3 = 230$.

   (d) $z = 1350 - 4x_1 - x_4 - 2x_5$.
   Increasing any of x_1, x_4, x_5 from 0 would decrease z, so 1350 is maximum.

(e) $B^{-1} = \begin{pmatrix} 1/2 & -1/4 & 0 \\ 0 & 1/2 & 0 \\ -2 & 1 & 1 \end{pmatrix}$. (under x_4, x_5, x_6 in optimal tableau)

$B = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 4 & 0 & 1 \end{pmatrix}$. (under x_2, x_3, x_6 in optimal tableau)

(f) If original solution column is $\mathbf{b} = (k \quad 460 \quad 420)^t$, final solution column is

$B^{-1}\mathbf{b} = \left(\frac{k}{2} - 115 \quad 230 \quad 880 - 2k\right)^t$.

For this to be non-negative, $k \geq 230$ and $k \leq 440$.

Then $z_{\max} = (1 \;\; 2 \;\; 0)\mathbf{b} = k + 920$.