

Linear Programming

Lecture notes

Dr Jeremy Budd,
School of Mathematics
University of Birmingham
Birmingham B15 2TT, UK
Email: j.m.budd@bham.ac.uk

Abstract

These are the lecture notes for the module “Linear Programming” covering LP modelling, geometric method, duality theory, the simplex method and some extensions of the simplex method.

Linear programming grew out of attempts to solve systems of linear inequalities, allowing one to optimise linear functions subject to constraints expressed as inequalities. The theory was developed independently at the time of World War II by the Soviet mathematician Kantorovich, for production planning, and by Dantzig, to solve complex military planning problems. Koopmans applied it to shipping problems and the technique enjoyed rapid development in the postwar industrial boom. The first complete algorithm to solve linear programming problems, called the simplex method, was published by Dantzig in 1947 and in the same year von Neumann established the theory of duality. In 1975, Kantorovich and Koopmans shared the Nobel Prize in Economics for their work and Dantzig’s simplex method has been voted the second most important algorithm of the 20th century after the Monte Carlo method. Linear programming is a modern and immensely powerful technique that has numerous applications, not only in business and economics, but also in engineering, transportation, telecommunications, and planning.

By the end of the module students should be able to:

- construct linear programming models of a variety of managerial problems and interpret the results obtained by applying the linear programming techniques to these problems,
- explain why and when the simplex method fails to provide a solution and how to resolve such a situation
- present, prove and use the results of duality theory and interpret them,
- explain the computational complexity of SIMPLEX and LP

Acknowledgments

These lecture notes are based on lecture notes written by Dr Yun-Bin Zhao, Dr Hon Duong, and Prof. Marta Mazzocco. I would like to thank them for sharing their notes.

Contents

| | |
|---|-----------|
| Abstract | 1 |
| 1 Introduction to Linear Programming | 5 |
| 1.1 What is optimisation? | 6 |
| 1.2 What is Linear Programming? | 6 |
| 1.3 Definition of LP: | 8 |
| 1.4 Terminology | 9 |
| 1.5 LP Modeling | 10 |
| 1.6 Standard Forms of LPs: | 17 |
| 1.7 Canonical Form of LPs | 18 |
| 1.8 A brief history of LPs | 19 |
| 2 LPs with 2 Variables: Geometric Method | 20 |
| 3 Introduction to Convex Analysis | 24 |
| 3.1 Basic concepts in convex analysis | 24 |
| 3.2 Extreme Points, Directions, Representation of Polyhedra | 29 |
| 4 Duality theory | 34 |
| 4.1 Motivation and derivation of dual LP problems | 35 |
| 4.2 Primal and dual problems of LPs | 37 |
| 4.3 Duality theorem | 41 |
| 4.4 Network flow problems | 43 |
| 4.5 Optimality Conditions & Complementary slackness | 46 |

| | | |
|----------|---|-----------|
| 4.6 | More Applications of Duality Theory | 51 |
| 5 | The Simplex Method (I): Theory | 53 |
| 5.1 | Basic Feasible Solution | 54 |
| 5.1.1 | Review of Linear Algebra | 54 |
| 5.1.2 | Existence of the Optimal Extreme Point | 55 |
| 5.1.3 | Basic Feasible Solution | 56 |
| 5.1.4 | Correspondence between BFSs & extreme points | 59 |
| 5.2 | The simplex method | 62 |
| 5.2.1 | Reformulation of LP by the current BFS | 63 |
| 5.2.2 | Checking the Optimality | 65 |
| 5.2.3 | Principles and justification of simplex method: | 66 |
| 6 | The simplex method (II): Algorithm | 70 |
| 6.1 | The algorithm for minimisation problems | 70 |
| 6.2 | The simplex method in Tableau Format | 72 |
| 6.2.1 | Pivoting: | 74 |
| 6.2.2 | Algorithm in tableau Format | 76 |
| 6.3 | Degeneracy and cycling in the simplex method | 79 |
| 6.4 | Efficiency of the simplex method | 82 |
| 6.5 | Further Examples for Simplex Methods & Degeneracy | 84 |
| 6.6 | Extracting Information from the Optimal Simplex Tableau | 88 |
| 6.6.1 | Find the inverse of optimal basis | 89 |
| 6.6.2 | Recovering original LP problems from the optimal tableau | 90 |
| 7 | The Dual Simplex Method (non-examinable) | 92 |
| 7.1 | Dual information from the (primal) optimal simplex tableau | 93 |
| 7.2 | Recovering missing data in simplex tableau | 94 |
| 7.3 | Complementary slack condition | 95 |
| 7.4 | Uniqueness of optimal solutions to a linear programming problem | 96 |

| | | |
|----------|---|------------|
| 7.5 | The Dual Simplex Method | 97 |
| 8 | Sensitivity Analysis (non-examinable) | 101 |
| 8.1 | Change in the cost vector | 102 |
| 8.2 | Change in the right-hand-side | 104 |
| 8.3 | Adding a new activity | 106 |
| 8.4 | Adding a new constraint | 108 |
| 9 | The Two-Phase method and the Big-M Method (non-examinable) | 111 |
| 9.1 | The Two-phase Method | 112 |
| 9.2 | The Big-M method | 123 |
| | Bibliography | 128 |

Chapter 1

Introduction to Linear Programming

This chapter provides key concepts of LP that include its definition and terminologies, standard and canonical forms. Some typically important realistic problems that can be modelled and solved using LP are also discussed. The following topics will be covered:

- definition of LP,
- terminologies of LP,
- standard forms of LPs,
- canonical forms of LPs,
- LP modelling.

At the end of this section, students will be able to

- understand the definition of a LP and relevant terminologies.
- write LPs in standard and canonical forms,
- recognise and formulate some typical problems in real applications into LPs.

[BT97, Chapter 1] and [DT97, Chapter 1] discuss similar topics as in this chapter that you may find useful.

1.1 What is optimisation?

Mathematical optimisation (mathematical programming) is a mathematical method for determining a way to achieve the best outcome (e.g., maximum profit or lowest cost) in a given list of requirements (constraints).

We do this by representing the ways to achieve an outcome by a variable, e.g. x , and the quality of the outcome by a function, $f(x)$, called the *objective function*.

We then seek to find

$$\begin{array}{ll} \max \text{ (or min)} & f(x) \\ \text{subject to (s.t.)} & \text{constraints} \end{array}$$

This value is the *optimal value* of the objective function. Sometimes, we will care more about the *optimiser(s)*, the value(s) of the variable which produce this optimal value. This is written

$$\begin{array}{ll} \operatorname{argmax} \text{ (or argmin)} & f(x) \\ \text{subject to (s.t.)} & \text{constraints} \end{array}$$

Optimisation is an extremely important area of modern applied mathematics, with applications in just about every industry.

1.2 What is Linear Programming?

- Linear programming (LP) is a specific case of mathematical optimisation. It is a technique for minimizing (or maximizing) a linear objective function, subject to linear equality and/or linear inequality constraints.

A good understanding of the theory and algorithms for linear programming is essential for understanding

- nonlinear programming (nonlinear optimisation),
- integer programming (integer optimisation),

- combinatorial optimisation problems,
- game theory,
- big-data sparse representation.

So, the theory of LP serves as a guide and motivating force for developing results for other mathematical optimisation problems with either continuous or discrete variables.

Example 1.1. *A small machine shop manufactures two models: standard and deluxe.*

- *Each standard model requires 2 hours of grinding and 4 hours of polishing.*
- *Each deluxe module requires 5 hours of grinding and 2 hours of polishing.*
- *The manufacturer has three grinders and two polishers. Therefore in 40 hours week there are 120 hours of grinding capacity and 80 hours of polishing capacity.*
- *There is a net profit of £3 on each standard model and £4 on each deluxe model.*

To maximise the profit, the manager must decide on the allocation of the available production capacity to standard and deluxe models.

Let x_1 and x_2 be the numbers of standard and deluxe models manufactured, respectively.

For this small machine shop, there are two constraints (restrictions):

- *Grinding restriction:*

$$2x_1 + 5x_2 \leq 120.$$

- *Polishing restriction:*

$$4x_1 + 2x_2 \leq 80.$$

The variables x_1, x_2 are nonnegative:

$$x_1 \geq 0, \quad x_2 \geq 0. \quad (\text{Why?})$$

The net profit is

$$3x_1 + 4x_2.$$

Therefore, we obtain the following model:

$$\begin{aligned}
 &\text{maximise} && 3x_1 + 4x_2 \\
 &\text{subject to} && 2x_1 + 5x_2 \leq 120, \\
 &&& 4x_1 + 2x_2 \leq 80, \\
 &&& x_1, x_2 \geq 0.
 \end{aligned}$$

This is a linear programming problem.

1.3 Definition of LP:

Definition 1.2. A linear programming problem is the problem of minimizing or maximizing a **linear function** subject to **linear constraints**. The constraints may be equalities or inequalities.

$$\begin{aligned}
 &\text{minimise (maximise)} && c_1x_1 + c_2x_2 + \dots + c_nx_n \\
 &\text{Subject to} && a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \{ \leq, =, \geq \} b_1, \\
 &&& a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \{ \leq, =, \geq \} b_2, \\
 &&& \vdots \\
 &&& a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \{ \leq, =, \geq \} b_m, \\
 &&& x_j \geq 0, \quad j = 1, \dots, n.
 \end{aligned}$$

LP in matrix notation:

The LP can be written as

$$\begin{aligned}
 &\text{minimise (maximise)} && c^T x \\
 &\text{subject to} && Ax \{ \geq, =, \leq \} b, \\
 &&& x \geq 0,
 \end{aligned}$$

where $\{ \geq, =, \leq \}$ denotes a diagonal matrix in which the ii -th component is the symbol

$\{\leq, =, \geq\}$ in the i -th equation,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \quad c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

All vectors are understood as column vectors, unless otherwise stated.

The problem can be also represented via the columns of A . Denoting

$$A = [a^1, a^2, \dots, a^n]$$

where a^j is the j th column of A , the LP can be written as

$$\begin{aligned} & \text{minimise (maximise)} \quad \sum_{j=1}^n c_j x_j \\ & \text{subject to} \quad \sum_{j=1}^n a^j x_j \quad (\geq, =, \leq) \quad b, \\ & \quad \quad \quad x \geq 0. \end{aligned}$$

1.4 Terminology

- The function

$$c^T x = c_1 x_1 + \dots + c_n x_n$$

is called the **objective function** (criterion function, goal, or target).

- $c = (c_1, c_2, \dots, c_n)^T$ is called the **cost coefficient**.
- x_1, x_2, \dots, x_n are called the **decision variables**.
- The inequality ' $Ax \leq b$ ' is called the **constraint**, and $\sum_{j=1}^n a_{ij} x_j \leq b_i$ denotes the i th constraint.
- A is called **constraint matrix** or **coefficient matrix**.
- b is called the **right-hand-side vector**.
- If $x = (x_1, x_2, \dots, x_n)^T$ satisfies all the constraints, it is called a **feasible point**, or feasible solution.

- The set of all such points is called the ***feasible region*** (or feasible set).

Example 1.3.

$$\begin{aligned}
 &\text{minimise} && 2x_1 + x_2 \\
 &\text{subject to} && x_1 + x_2 \leq 8, \\
 &&& -3x_1 + 5x_2 \leq 10, \\
 &&& x_1 \geq 0, x_2 \geq 0.
 \end{aligned}$$

- What is the matrix form for this LP?
- What is the feasible region for this LP?

For this example, the problem has two decision variables x_1 and x_2 . The objective function to be minimised is $2x_1 + x_2$. The LP problem is to find a point in the feasible region with the smallest objective value.

1.5 LP Modeling

Linear programming is an extremely powerful tool for addressing a wide range of practical problems. It is used extensively in business, economics, and engineering problems. LP models have been widely used in transportation, energy, telecommunications, manufacturing, production planning, routing, flight crew scheduling, resource allocation, assignment, and design, to name but a few.

Basic steps in the development of an LP model:

- Determine the decision variables.
- Determine the objective function.
- Determine the explicit constraints.
- Determine the implicit constraints.

Example 1.4 (Production Planning).

- A manufacturing firm has to decide on the mix of Ipads and iPhones to be produced.
- A market research indicated that at most 1000 units of Ipads and 4000 units of iPhones can be sold per month.
- The maximum number of work-hours available is 50,000 per month.
- To produce one unit, an Ipad requires 20 work-hours and an iPhone requires 15 work-hours.
- The unit profits of the Ipads and iPhones are £300 and £400 respectively.

It is desired to find the number of unites of Ipads and iPhones that the firm must produce in order to maximise its profit.

Determine the decision variables. Suppose that the number of the units for Ipads is x_1 and the number of the units for iPhones is x_2 .

Determine the explicit constraints. There are two kinds of constraints: The market constraints and work-hours restriction. For the former, we have

$$x_1 \leq 1000,$$

$$x_2 \leq 4000.$$

For the latter, we have

$$20x_1 + 15x_2 \leq 50,000.$$

Determine the objective function. The total profit is given by

$$300x_1 + 400x_2.$$

Determine the implicit constraints. The number of Ipads and iPhones cannot be negative, so $x_1, x_2 \geq 0$.

Therefore, the problem is formulated as the following linear programming problem:

$$\begin{aligned}
 &\text{maximise} && 300x_1 + 400x_2 \\
 &\text{subject to} && 20x_1 + 15x_2 \leq 50,000, \\
 &&& x_1 \leq 1000, \\
 &&& x_2 \leq 4000, \\
 &&& x_1, x_2 \geq 0.
 \end{aligned}$$

Example 1.5 (The Workforce Planning Problem). *Consider a restaurant that is open seven days a week. Based on past experience, the number of workers needed on a particular day is given as follows:*

| Day | Mon | Tue | Wed | Thu | Fri | Sat | Sun |
|--------|-----|-----|-----|-----|-----|-----|-----|
| Number | 14 | 13 | 15 | 16 | 19 | 18 | 11 |

Every worker works five consecutive days, and then takes two days off, repeating this pattern indefinitely. How can we minimise the number of workers that staff the restaurant?

Define the variables: Let the days be numbers 1 through 7 and let x_i be the number of workers who begin their five-day shift on day i .

Let us consider Monday as an example. People work on Monday must start to work on Thur, Fri, Sat, Sun and Monday. Thus the total number of workers on Monday is

$$x_1 + x_4 + x_5 + x_6 + x_7$$

which is at least 14. So we have the following constraints

$$x_1 + x_4 + x_5 + x_6 + x_7 \geq 14.$$

Similarly, we may consider other days. Thus, the linear programming model for this problem is given as follow.

$$\begin{aligned}
& \min \sum_{i=1}^7 x_i \\
& \text{s.t.} \quad x_1 + x_4 + x_5 + x_6 + x_7 \geq 14 \text{ (Mon),} \\
& \quad \quad x_1 + x_2 + x_5 + x_6 + x_7 \geq 13 \text{ (Tue),} \\
& \quad \quad x_1 + x_2 + x_3 + x_6 + x_7 \geq 15 \text{ (Wed),} \\
& \quad \quad x_1 + x_2 + x_3 + x_4 + x_7 \geq 16 \text{ (Thur),} \\
& \quad \quad x_1 + x_2 + x_3 + x_4 + x_5 \geq 19 \text{ (Fri),} \\
& \quad \quad x_2 + x_3 + x_4 + x_5 + x_6 \geq 18 \text{ (Sat),} \\
& \quad \quad x_3 + x_4 + x_5 + x_6 + x_7 \geq 11 \text{ (Sun),} \\
& \quad \quad x_i \geq 0, \quad i = 1, \dots, 7.
\end{aligned}$$

Example 1.6 (The Diet Problem).

- There are m different types of food, F_1, \dots, F_m , that supply varying quantities of the n nutrients, N_1, \dots, N_n , that are essential to good health.
- Let c_j be the minimum daily requirement of nutrient N_j .
- Let b_i be the price per unit of food F_i .
- Let a_{ij} be the amount of nutrient N_j contained in one unit of food F_i .

The problem is to supply the required nutrients at minimum cost.

Let y_i be the number of units of food F_i to be purchased per day.

The cost per day of such a diet is

$$b_1 y_1 + b_2 y_2 + \dots + b_m y_m. \tag{1.1}$$

The amount of nutrient N_j contained in this diet is

$$a_{1j} y_1 + a_{2j} y_2 + \dots + a_{mj} y_m$$

for $j = 1, \dots, n$.

We need to ensure that all the minimum daily requirements are met, that is,

$$a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \geq c_j, \quad j = 1, \dots, n. \quad (1.2)$$

Of course, we cannot purchase a negative amount of food. So we have

$$y_1 \geq 0, \quad y_2 \geq 0, \quad \dots, \quad y_m \geq 0. \quad (1.3)$$

The problem is to minimise (1.1) subject to (1.2) and (1.3). This is exactly the following LP problem:

$$\begin{aligned} &\text{minimise} \quad b_1y_1 + b_2y_2 + \cdots + b_my_m \\ &\text{s.t} \quad a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \geq c_j, \quad j = 1, \dots, n, \\ &\quad y_1 \geq 0, \quad y_2 \geq 0, \quad \dots, \quad y_m \geq 0. \end{aligned}$$

Example 1.7 (The Transportation Problem).

- There are I production plants, P_1, \dots, P_I , that supply a certain commodity, and there are J markets, M_1, \dots, M_J , to which this commodity must be shipped.
- Plant P_i possesses an amount s_i of the commodity ($i = 1, 2, \dots, I$), and market M_j must receive the amount r_j of the commodity ($j = 1, \dots, J$).
- Let c_{ij} be the cost of transporting one unit of the commodity from plant P_i to market M_j .

The problem is to meet the market requirements at minimum transportation cost.

1) Let x_{ij} be the quantity of the commodity shipped from plant P_i to market M_j .

2) The total transportation cost is

$$\sum_{i=1}^I \sum_{j=1}^J x_{ij}c_{ij}. \quad (1.4)$$

The amount sent from plant P_i is

$$\sum_{j=1}^J x_{ij}$$

and since the amount available at plant P_i is s_i , we must have

$$\sum_{j=1}^J x_{ij} \leq s_i, \quad \text{for } i = 1, \dots, I. \quad (1.5)$$

3) The amount sent to market M_j is

$$\sum_{i=1}^I x_{ij}$$

and since the amount required there is r_j , we must have

$$\sum_{i=1}^I x_{ij} \geq r_j, \quad j = 1, \dots, J. \quad (1.6)$$

4) It is assumed that we cannot send a negative amount from plant P_i to M_j . So, we have

$$x_{ij} \geq 0, \quad i = 1, \dots, I, \quad j = 1, \dots, J. \quad (1.7)$$

The problem is to minimise the objective (1.4) subject to the constraints (1.5), (1.6) and (1.7), i.e.,

$$\begin{aligned} & \text{minimise} \quad \sum_{i=1}^I \sum_{j=1}^J x_{ij} c_{ij} \\ & \text{s.t} \quad \sum_{j=1}^J x_{ij} \leq s_i, \quad i = 1, \dots, I, \\ & \quad \sum_{i=1}^I x_{ij} \geq r_j, \quad j = 1, \dots, J, \\ & \quad x_{ij} \geq 0, \quad i = 1, \dots, I, \quad j = 1, \dots, J. \end{aligned}$$

Example 1.8 (The Assignment Problem).

- There are 100 persons available for 10 jobs.
- The value of person i working 1 day at job j is a_{ij} ($i = 1, \dots, 100, j = 1, \dots, 10$).

- Let us suppose that only one person is allowed on a job at a time (two people cannot work on the same job at the same time).

The problem is to choose an assignment of persons to jobs to maximise the total value.

An assignment is a choice of numbers, x_{ij} , for $i = 1, \dots, 100$, and $j = 1, \dots, 10$, where x_{ij} represents the proportion of person i 's time that is to be spent on job j . Thus,

$$\sum_{j=1}^{10} x_{ij} \leq 1, \quad i = 1, \dots, 100, \quad (1.8)$$

$$\sum_{i=1}^{100} x_{ij} \leq 1, \quad j = 1, \dots, 10, \quad (1.9)$$

and

$$x_{ij} \geq 0, \quad i = 1, \dots, 100, \quad j = 1, \dots, 10. \quad (1.10)$$

- Inequality (1.8) reflects the fact that a person cannot spend more than 100 percentage of their time working.
- Inequality (1.9) means that the total time spent on a job is not allowed to exceed a day.
- (1.10) says that no one can work a negative amount of time on any job.

Subject to constraints (1.8), (1.9) and (1.10), we wish to maximise the total value,

$$\sum_{i=1}^{100} \sum_{j=1}^{10} a_{ij} x_{ij}. \quad (1.11)$$

This is the following LP problem:

$$\begin{aligned} \max \quad & \sum_{i=1}^{100} \sum_{j=1}^{10} a_{ij} x_{ij} \\ \text{s.t.} \quad & \sum_{j=1}^{10} x_{ij} \leq 1, \quad i = 1, \dots, 100, \\ & \sum_{i=1}^{100} x_{ij} \leq 1, \quad j = 1, \dots, 10, \\ & x_{ij} \geq 0, \quad i = 1, \dots, 100, \quad j = 1, \dots, 10. \end{aligned}$$

1.6 Standard Forms of LPs:

$$\begin{aligned} &\text{minimise} && c^T x \\ &\text{subject to} && Ax = b, \\ &&& x \geq 0. \end{aligned}$$

Throughout this course, we call the above minimization problem as a standard form. In fact, any LP problem can be converted to this form.

Reduction to the standard form:

- An **inequality** can be easily transformed into an equality by adding a **slack variable**

$$d^T x \leq \alpha \iff d^T x + t = \alpha, \quad t \geq 0,$$

$$d^T x \geq \alpha \iff d^T x - t = \alpha, \quad t \geq 0.$$

In matrix form:

$$Ax \leq b \iff Ax + s = b, \quad s \geq 0.$$

- A **‘free’ variable** can be replaced by two nonnegative variables: If x_j is unrestricted in sign, then

$$x_j = x'_j - x''_j, \quad \text{where } x'_j, \quad x''_j \geq 0.$$

- Converting a **maximization (minimization) problem** into a minimization (maximization) problem: Note that over any feasible region, we have

$$\text{maximum } c^T x = -\text{minimum } (-c^T x).$$

So, by multiplying the coefficients of the objective function by -1 , the maximization problem can be converted into a minimization problem.

- Eliminating **upper and lower bounds**:

$$L_j \leq x_j \leq U_j \iff x_j \leq U_j \text{ and } x_j \geq L_j$$

which are reduced to the first case.

Example 1.9. Write down the standard form of the following linear programming:

$$\begin{aligned} \max \quad & 4x_1 + x_2 - x_3 \\ \text{s.t.} \quad & x_1 + 3x_3 \leq 6, \\ & 3x_1 + x_2 + 3x_3 \geq 9, \\ & x_1, x_2 \geq 0, \ x_3 \text{ is unrestricted in sign.} \end{aligned}$$

Example 1.10. Write down the standard form of the following linear programming:

$$\begin{aligned} \min \quad & 3x_1 - 4x_2 + 2x_3 \\ \text{s.t.} \quad & 5x_1 - 3x_2 + x_3 \geq 7, \\ & -4x_1 + 9x_3 \leq 10, \\ & x_1 + x_2 - 2x_3 = 5, \\ & x_1 \geq 0, \ x_3 \geq 0. \end{aligned}$$

1.7 Canonical Form of LPs

The LP problem

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax \geq b \quad (\text{or in maximum form}) \\ & x \geq 0 \end{array} \quad \begin{array}{ll} \max & c^T x \\ \text{s.t.} & Ax \leq b \\ & x \geq 0 \end{array}$$

is called the ***canonical form of the LP***.

Later, we will see that the canonical form is very useful in the duality representation of LPs. Introducing a slack vector, the canonical form can be easily converted to the standard form

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax - s = b \\ & x \geq 0, \ s \geq 0. \end{aligned}$$

Conversely, a standard form can be easily transformed to a canonical form. (Do this exercise)

1.8 A brief history of LPs

Linear programming was developed during the second world war to plan expenditures and returns in order to reduce costs to the army and increase losses to the enemy.

It was kept secret until 1947. Postwar, many industries found its use in their daily planning.

- **Leonid Kantorovich**, a Russian mathematician, developed linear programming problems in 1939.
- **George B. Dantzig** published the simplex method in 1947.
- **John von Neumann** developed the theory of the duality in the same year.
- The linear programming problem was first shown to be solvable in polynomial time by **Leonid Khachiyan** in 1979.
- A larger theoretical and practical breakthrough in the field came in 1984 when **Narendra Karmarkar** introduced a new interior point method for solving linear programming problems.

Chapter 2

LPs with 2 Variables: Geometric Method

In this chapter, we will study LPs with 2 variables and how to solve them using the geometric method. Topics that will be covered:

- notion of steepest descent,
- graphical method for solving LPs with two variables.

At the end of this chapter students should be able to use geometric method to solve a two variable LP problem determining whether it is infeasible, unbounded or has a finite solution.

[DT97, Chapter 2] presents similar topics to this chapter that you may find helpful.

The following result contains the underlying idea of the geometric method for solving LPs with two variables that we will study in this chapter.

Proposition 2.1. *Let $f(x)$ be a continuously differentiable function where $x \in \mathbb{R}^n$. At any point x , moving in the direction of gradient $\nabla f(x) \neq 0$ for any small stepsize, the function value increases, and moving in the direction $-\nabla f(x)$ for any small stepsize, the function value decreases.*

(Why?) This can be seen from the directional derivative:

$$\lim_{\lambda \rightarrow 0} \frac{f(x + \lambda d) - f(x)}{\lambda} = \nabla f(x)^T d.$$

Corollary 2.2. *Moving in the direction c , the value of the function $c^T x$ increases; moving in the direction $-c$, the function value decreases.*

Geometric solution:

Minimizing $c^T x$ corresponds to moving the line $c^T x = \text{constant}$ in the direction of $-c$ as far as possible.

- When the feasible region is unbounded, we might move the line indefinitely while always intersecting the feasible region, and hence the problem is ***unbounded***.
- For a bounded and some unbounded feasible regions, the objective line moves and must stop at a certain point, otherwise it goes beyond the feasible region. In such cases, the problem has a ***finite optimal solution***.
- The feasible set might be empty, and hence the problem is ***infeasible***.

Example 2.3 (Finite solution case). *Solve the LP geometrically (or graphically)*

$$\begin{aligned} \text{minimise} \quad & x_1 + x_2 \\ \text{subject to} \quad & 2x_1 + x_2 \leq 6, \\ & -x_1 + x_2 \leq 4, \\ & x_1, x_2 \geq 0. \end{aligned}$$

Three steps to solve the problem:

- 1 . Sketch the feasible set
- 2 . Choose one point in the feasible set and draw the objective hyperplane (in the case of in the direction of 2 variables this is just a line)
- 3 . For a minimisation problem, move the objective hyperplane in the direction of $-c$ to reach the optimal solution. For a maximisation problem, move the objective hyperplane in the direction of c to obtain the optimal solution.

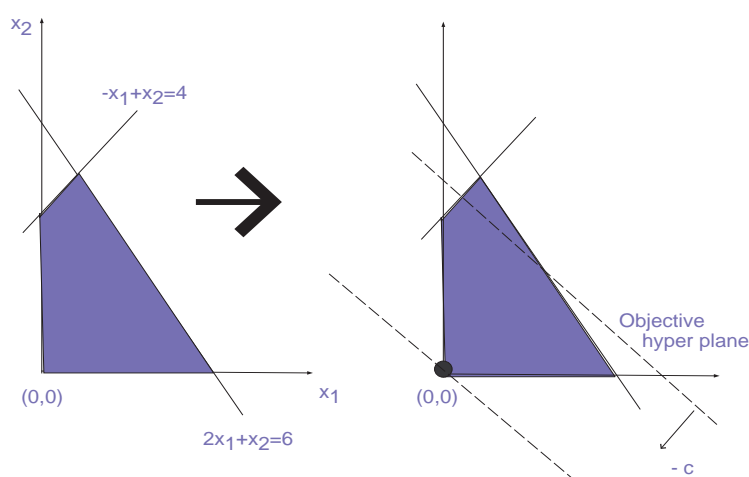


Figure 2.1: Geometric solution

Example 2.4 (Unbounded case). *Solve the LP geometrically*

$$\begin{aligned}
 &\text{maximise} && x_1 + x_2 \\
 &\text{subject to} && -x_1 + x_2 \leq 4, \\
 &&& x_1, x_2 \geq 0.
 \end{aligned}$$

Still, we follow the above three steps and it is easy to see from Figure 2 that this problem has unbounded solution (no finite optimal solution).

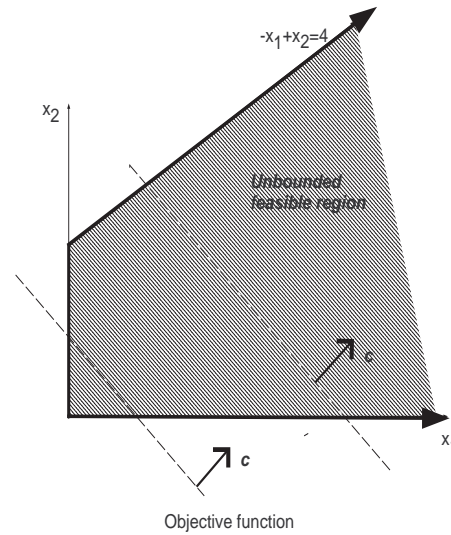


Figure 2.2: Geometric solution

Example 2.5 (Infeasible case). *Solve the LP geometrically*

$$\begin{aligned}
 &\text{minimise} && x_1 + x_2 \\
 &\text{subject to} && -x_1 + x_2 \geq 4, \\
 &&& x_1 - 2x_2 \geq 6, \\
 &&& x_1, x_2 \geq 0.
 \end{aligned}$$

In summary, given an LP, there are only three cases:

- 1 . The problem has a finite optimal solution (unique optimal solution or infinitely many optimal solutions).
- 2 . The solution of LP is unbounded.
- 3 . The problem is infeasible, i.e., the feasible region is empty.

Chapter 3

Introduction to Convex Analysis

In this chapter, we study basic knowledge in convex analysis that will be used in subsequent chapters. The following topics will be covered

- basic concepts in convex analysis (convex set, convex functions, etc)
- extreme points and directions,
- representation of polyhedra.

At the end of this section, students should be able to

- understand basics concepts in convex analysis,
- compute extreme points and extreme directions of simple convex sets.

[BT97, Chapter 2] contains similar material to this chapter that you may find useful.

3.1 Basic concepts in convex analysis

Hyperplane and half-space:

- A *hyperplane* H in \mathbb{R}^n is a set of the form

$$\{x : p^T x = \alpha\}$$

where p is a nonzero vector called the normal or the gradient to the hyperplane.

- A hyperplane divides \mathbb{R}^n into two regions, called *half-spaces*. A half-space is the collection of points of the form

$$\{x : p^T x \geq \alpha\} \text{ or } \{x : p^T x \leq \alpha\}.$$

We often use the following symbols:

$$H = \{x : p^T x = \alpha\},$$

$$H^+ = \{x : p^T x \leq \alpha\},$$

$$H^- = \{x : p^T x \geq \alpha\}.$$

Definition 3.1. • A **polyhedron** $P \subset \mathbb{R}^n$ is the set of points that satisfy a finite number of linear inequalities, i.e.,

$$P = \{x : Ax \leq b\},$$

where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$.

- A bounded polyhedron is called a **polytope**.

Convex set

A set T is a convex set if $x^1, x^2 \in T$ implies that $\lambda x^1 + (1 - \lambda)x^2 \in T$ for all $0 \leq \lambda \leq 1$.

Proposition 3.2. (i) Any hyperplane $H = \{x : p^T x = \beta\}$ is convex.

(ii) The half-spaces $\{x : p^T x \geq \beta\}$ and $\{x : p^T x \leq \beta\}$ are convex.

Proof. Suppose \hat{x} and \bar{x} are elements in H . Then any point of the form

$$x(\alpha) = \alpha \hat{x} + (1 - \alpha) \bar{x}$$

where α can be any number in $(0, 1)$, satisfies

$$p^T x(\alpha) = \alpha p^T \hat{x} + (1 - \alpha) p^T \bar{x} = \alpha \beta + (1 - \alpha) \beta = \beta.$$

The proof for the half spaces is the same by changing the $=$ signs into \geq or \leq . □

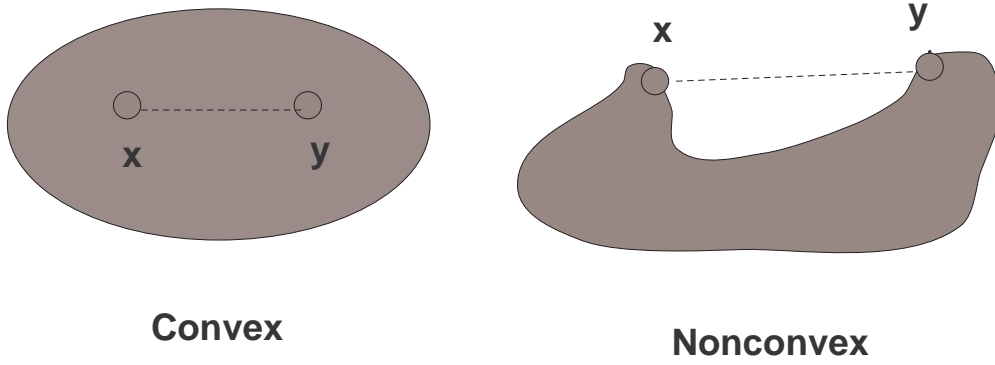


Figure 3.1: Convex and nonconvex sets

Proposition 3.3. (i) *The intersection of convex sets is convex.*

(ii) *A polyhedron is convex.*

Proof of (i). Let S_1, \dots, S_n be convex sets in \mathbb{R}^n . Let

$$S = \bigcap_{i=1}^n S_i$$

be the intersection of S_i 's. We now prove that S is a convex set. Indeed, let $x, y \in S$. We now prove that for any $\alpha \in (0, 1)$, we have $x(\alpha) = \alpha x + (1 - \alpha)y \in S$. In fact, since $x, y \in S$ which is the intersection of S_i 's, the vectors x and y are in S_i for all $i = 1, \dots, n$. By the convexity of S_i , we must have

$$x(\alpha) = \alpha x + (1 - \alpha)y \in S_i, \quad \forall i = 1, \dots, n.$$

Thus $x(\alpha)$ is in the intersection of these sets, and hence $x \in S$.

[Proof of (ii)] Since each half-space is convex, and a polyhedron is the intersection of a finite number of half-spaces, any polyhedron is convex. \square

Proposition 3.4. *If a linear programming problem has a finite optimal solution, then it has either a unique solution or infinitely many solutions.*

Proof. Consider the linear programming problem

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0. \end{aligned}$$

Suppose that the optimal solution of the LP is not unique. Then it has at least two different optimal solutions, denoted by \hat{x} and \bar{x} .

Since \hat{x} and \bar{x} are optimal solutions, they must be feasible to the problem, and the objective values at these two solutions are equal, i.e.

$$A\hat{x} = b, \hat{x} \geq 0, \quad A\bar{x} = b, \bar{x} \geq 0$$

and

$$c^T \hat{x} = c^T \bar{x} = z^*,$$

where z^* denotes the optimal objective value.

Therefore \hat{x} and \bar{x} belong to the same hyperplane $c^T x = z^*$ and to the same polyhedron $\{x | Ax = b\}$ which are convex sets. This proves that any point in the segment $\alpha\hat{x} + (1-\alpha)\bar{x}$, for $\alpha \in (0, 1)$ is a optimal solution. \square

Cone and Convex Cone

A set C is called a **cone** if $x \in C$ implies that $\lambda x \in C$ for all $\lambda \geq 0$. A convex cone is a convex set and cone.

Proposition 3.5. *A polyhedron is a cone if and only if it can be represented as*

$$P = \{x : Ax \leq 0\}$$

for some matrix A .

Proof. If the set can be represented as $P = \{x : Ax \leq 0\}$ for some matrix A , then for any $x \in P$ we have $\alpha x \in P$ for any $\alpha \geq 0$. Thus, P is a cone.

Conversely, if P is a polyhedron and a cone, then there exists a matrix A and a vector b such that P can be represented as

$$P = \{x : Ax \leq b\}.$$

Moreover, by the definition of cone, for any $x \in P$, it must hold that $\alpha x \in P$ for any $\alpha \geq 0$. In particular, by taking $\alpha = 0$, we see that $0 \in P$, thus $b \geq 0$. Let $\alpha > 0$ be any

sufficiently large number, it follows from $A(\alpha x) \leq b$ that $Ax \leq 0$. Thus $P \subseteq \{x : Ax \leq 0\}$. Since $b \geq 0$, it further implies that

$$P \subseteq \{x : Ax \leq 0\} \subseteq P,$$

so $P = \{x : Ax \leq 0\}$.

□

Convex function

A function f of the vector (x_1, x_2, \dots, x_n) is said to be convex if the following inequality holds for any two vectors $x, y \in \mathbb{R}^n$:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad \forall \lambda \in [0, 1].$$

In 1-dimensional case, the foregoing inequality can be interpreted as follows: $\lambda f(x) + (1 - \lambda)f(y)$ represents the height of the segment joining $(x, f(x))$ and $(y, f(y))$ at the point $\lambda x + (1 - \lambda)y$. The convexity means the height of the segment is at least as large as the height of the function itself.

Proposition 3.6. *The objective function $c^T x$ is convex.*

Proof. This can be verified directly. Let $x, y \in \mathbb{R}^d$ and $\lambda \in [0, 1]$, then

$$c^T(\lambda x + (1 - \lambda)y) = \lambda c^T x + (1 - \lambda)c^T y.$$

Example 3.7. *The function $f(x) = x^2$, $x \in \mathbb{R}$, is convex. Indeed, let $x, y \in \mathbb{R}$ and $\lambda \in [0, 1]$. We need to verify that*

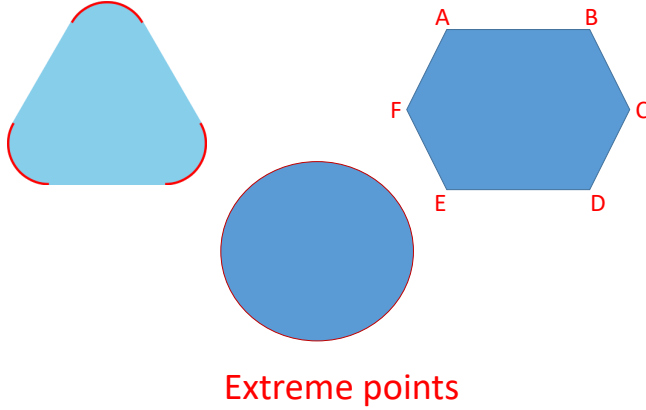
$$f(\lambda x + (1 - \lambda)y) = (\lambda x + (1 - \lambda)y)^2 \leq \lambda f(x) + (1 - \lambda)f(y) = \lambda x^2 + (1 - \lambda)y^2.$$

By subtracting the left-hand side from the right-hand side, the above equality is equivalent to

$$\lambda(1 - \lambda)(x - y)^2 \geq 0,$$

which is true for all $x, y \in \mathbb{R}$ and $\lambda \in [0, 1]$.

□



In summary,

- the feasible set of linear programming is a polyhedron,
- the feasible region of LPs is a convex set, and
- the objective of the LP is a convex function.

3.2 Extreme Points, Directions, Representation of Polyhedra

Definition 3.8 (Extreme point). *A point x in a convex set X is called an extreme point of X , if x cannot be represented as a strict convex combination of two distinct points in X . In other words, if $x = \lambda x^1 + (1 - \lambda)x^2$ with $0 < \lambda < 1$ and $x^1, x^2 \in X$ then it must have $x = x^1 = x^2$.*

Definition 3.9 (Rays). *Let $d \neq 0$ be a vector and x_0 be a point. The collection of points of the form $\{x_0 + \lambda d : \lambda \geq 0\}$ is called a ray.*

x_0 is called the vertex of the ray, and d is the direction of the ray.

Definition 3.10 (Direction of a convex set). *Given a convex set, a vector $d \neq 0$ is called direction (or recession direction) of the set, if for each x_0 in the set, the ray $\{x_0 + \lambda d : \lambda \geq 0\}$ also belongs to the set.*

Therefore, starting at any point x_0 in the set, one can recede along d for any step length $\lambda \geq 0$ and remain within the set.

Example 3.11.

1. What is the direction of bounded convex set?
2. What is the direction of polyhedron $X = \{x : Ax \leq b, \ x \geq 0\}$?
3. What is the direction of polyhedron $X = \{x : Ax = b, \ x \geq 0\}$?

Example 3.12. Consider the set

$$X = \{(x_1, x_2)^T : x_1 - 2x_2 \geq -6, \ x_1 - x_2 \geq -2, \ x_1 \geq 0, \ x_2 \geq 1\}.$$

What is the direction of X ?

Let $\mathbf{x}_0 = (x_1, x_2)^T$ be an arbitrary fixed feasible point. Then $d = (d_1, d_2)^T$ is a direction of X if and only if

$$(d_1, d_2)^T \neq 0$$

and

$$\mathbf{x}_0 + \lambda d = (x_1 + \lambda d_1, x_2 + \lambda d_2)^T \in X, \quad \forall \lambda \geq 0.$$

Therefore,

$$x_1 - 2x_2 + \lambda(d_1 - 2d_2) \geq -6,$$

$$x_1 - x_2 + \lambda(d_1 - d_2) \geq -2,$$

$$x_1 + \lambda d_1 \geq 0,$$

$$x_2 + \lambda d_2 \geq 1,$$

for all $\lambda \geq 0$. Since the last two inequalities must hold for fixed x_1 and x_2 and for all $\lambda \geq 0$, we conclude that $d_1, d_2 \geq 0$. Similarly, from the first two inequalities above, we conclude that

$$d_1 - 2d_2 \geq 0, \quad d_1 - d_2 \geq 0.$$

Since $d_1, d_2 \geq 0$, from $d_1 \geq 2d_2$ it implies that $d_1 \geq d_2$. Therefore, $(d_1, d_2)^T$ is a direction of X if and only if it satisfies the following conditions:

$$(d_1, d_2)^T \neq 0,$$

$$d_1 \geq 0, \quad d_2 \geq 0, \quad d_1 \geq 2d_2.$$

Definition 3.13 (Extreme direction of a convex set). *An extreme direction of a convex set is a direction of the set that cannot be represented as a positive combination of two distinct directions of the set.*

Two directions d and \bar{d} are said to be distinct if d cannot be represented as a positive multiple of \bar{d} , i.e., $d \neq \alpha \bar{d}$ for any $\alpha \geq 0$.

Example 3.14. 1. *What is the extreme direction of the set X in Example 3.12?*

2. *What is the extreme direction of the first orthant $\{x : x \geq 0\}$?*

Theorem 3.15. *A polyhedron*

$$P = \{x \in \mathbb{R}^n : Ax \leq b\} \neq \emptyset$$

has a finite number of extreme points and extreme rays.

Proof. This theorem can be proved directly using convex analysis. However, that approach is beyond the scope of this course. Instead, as will be seen in Theorem 5.5, the set of extreme points of a polyhedron is the same as the set of basic feasible solution of a linear programming problem defined on the polyhedron. Thus the number of extreme points is the same as the number of basic feasible solution, which is finite (cf. Section 1). □

Theorem 3.16. *Let P be a polyhedron. If the LP problem*

$$\min\{c^T x : x \in P\}$$

has a finite optimal solution, then there is an optimal solution that is an extreme point.

Proof. Similarly as Theorem 3.15, this theorem can also be proved directly using convex analysis. However, it can be seen as a direct consequences of Theorem 5.5 and Theorem 5.6 in Chapter 5. □

How to find an extreme direction?

Finding an extreme direction of a polyhedron can be converted into finding an extreme point of another polyhedron called *recession cone*.

1). Consider the set

$$X = \{x : Ax \leq b, x \geq 0\}.$$

The direction of the set is characterized by the following conditions:

$$d \geq 0, \quad d \neq 0, \quad \text{and} \quad Ad \leq 0.$$

This defines a polyhedral cone, called recession cone.

To eliminate duplication, these directions may be normalized, e.g. $e^T d = 1$. Then the set of recession directions of X can be given by

$$D = \{d : Ad \leq 0, e^T d = 1, d \geq 0\}.$$

Therefore, the extreme directions of X are precisely the extreme points of D .

2). Similarly, we may prove that the extreme directions of the nonempty polyhedral set

$$X = \{x : Ax = b, x \geq 0\}$$

correspond to extreme points of the following set:

$$D = \{d : Ad = 0, e^T d = 1, d \geq 0\}.$$

Representation of a Polyhedron

Convex combination

When $0 \leq \alpha \leq 1$, the point $\alpha x + (1 - \alpha)y$ is called the *convex combination* of x and y . It is the line segment between x and y .

In general, we have the following definition.

Definition 3.17. A vector b is said to be convex combination of the vectors u^1, \dots, u^k if

$$b = \sum_{j=1}^k \lambda_j u^j, \quad \text{where} \quad \sum_{j=1}^k \lambda_j = 1, \quad \lambda_j \geq 0, \quad j = 1, \dots, k.$$

Example 3.18. Write $b = \begin{pmatrix} 2/3 \\ 1/2 \end{pmatrix}$ as a convex combination of three vectors

$$u^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad u^3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

We need to find $\alpha, \beta, \gamma \geq 0$ such that

$$b = \alpha u^1 + \beta u^2 + \gamma u^3 \quad \text{and} \quad \alpha + \beta + \gamma = 1.$$

These are equivalent to

$$\begin{cases} \alpha + \gamma = 2/3, \\ \beta + \gamma = 1/2, \\ \alpha + \beta + \gamma = 1, \\ \alpha, \beta, \gamma \geq 0. \end{cases}$$

This system can be easily solved directly to obtain

$$\alpha = 1/2, \quad \beta = 1/3, \gamma = 1/6.$$

Thus b is written as a convex combination of $\{u^1, u^2, u^3\}$ as

$$b = \frac{1}{2}u^1 + \frac{1}{3}u^2 + \frac{1}{6}u^3.$$

The following theorem provides an important representation of a polyhedron in terms of extremes points and extreme rays. The proof is beyond the scope of the lecture notes and hence is omitted.

Theorem 3.19 (Minkowski's Theorem). *If $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ is nonempty and $\text{rank}(A) = n$, then*

$$P = \left\{ x : x = \sum_{j=1}^k \lambda_j x^j + \sum_{i=1}^l \mu_i d^i, \sum_{j=1}^k \lambda_j = 1, \right. \\ \left. \lambda_j \geq 0, j = 1, \dots, k, \mu_i \geq 0, i = 1, \dots, l \right\}$$

where $\{x^1, x^2, \dots, x^k\}$ is the set of extreme points of P and $\{d^1, d^2, \dots, d^l\}$ is the set of extreme rays of P .

Chapter 4

Duality theory

In this chapter we will study an important topic of LP, namely duality theory. Duality in linear programming is essentially a unifying theory that develops the relationships between a given LP problem (the primal problem) and another related LP problem (the dual problem). We also study the *optimality condition*, i.e., the so-called Karush-Kuhn-Tucker (KKT) optimality condition, for LP problems. The duality theory and optimality conditions form the backbone of linear and nonlinear optimisation problems.

The following topics will be covered

- definition of a dual problem,
- weak duality theorem,
- strong duality theorem,
- optimality conditions,
- complementary slackness condition,
- some further applications of duality theory.

At the end of this chapter, students should be able to

- formulate a dual problem of a LP,
- understand weak and strong duality theorems,
- understand the derivation of optimality conditions from duality theorems,

- write down optimality conditions for a LP problem.

[BT97, Chapter 5] and [DT97, Chapter 5] contain relevant topics to this chapter that you may find helpful.

4.1 Motivation and derivation of dual LP problems

Let us consider a common situation in economics: minimising cost vs. maximising profit.

Example 4.1 (minimising cost vs. maximising profit). *There are two products, A and B, that are made up from two elements, e and f. Product A contains 5 units of element e and 2 units of element f and can be sold with price 16 Yuan, while product B contains 3 units of element e and 4 units of element f and can be sold with price 10 Yuan.*

Consumption/customer's perspective: A customer wants to buy some product A and B. To fulfill his/her task, the customer needs at least 100 units of element e and 60 units of element f. How many (units of) product A and B should the customer buy in order to minimise the cost?

Let x and y be the units of product A and B that the customer buys. The above problem can be formulated into a minimising LP problem as follows

$$\text{Minimise } 16x + 10y \quad (\text{the total cost})$$

subject to

$$5x + 3y \geq 100 \quad (\text{the total units of element e}),$$

$$2x + 4y \geq 60 \quad (\text{the total units of element f}),$$

$$x, y \geq 0 \quad (\text{the customer needs both products}).$$

This LP can be written in a matrix form as follows

$$\text{Minimise } c^T \bar{x},$$

subject to

$$\begin{cases} A\bar{x} \geq b, \\ \bar{x} \geq 0, \end{cases}$$

where

$$\bar{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad A = \begin{pmatrix} 5 & 3 \\ 2 & 4 \end{pmatrix}, \quad b = \begin{pmatrix} 100 \\ 60 \end{pmatrix}, \quad c = \begin{pmatrix} 16 \\ 10 \end{pmatrix}.$$

Production's perspective. Now the supplier want to assign price for each unit of element e and f so that the customer's need can be met and the profit is maximal. Let u and v be the price that the supplier should assign to each unit of element e and f . His/her task can be formulated into the following maximising LP problem

$$\text{maximise } 100u + 60v \quad (\text{the total profit})$$

subject to

$$5u + 2v \leq 16 \quad (\text{the price for product A should not exceed 16 Yuan}),$$

$$3u + 4v \leq 10 \quad (\text{the price for product B should not exceed 10}),$$

$$u, v \geq 0 \quad (\text{the prices for each unit of each element}).$$

We also can write this LP in a matrix form using the notations above as

$$\text{maximise } b^T \bar{y},$$

subject to

$$\begin{cases} A^T \bar{y} \leq c, \\ \bar{y} \geq 0, \end{cases}$$

where $\bar{y} = (u, v)^T$. These are two dual problems.

Suppose now in Example above we only know the minimisation problem. How do we attempt to solve this? Since we are looking for a minimum value of the objective function, we first need to find a lower bound for it. To do so, we combine the constraints to obtain a lower bound of the type

$$u(5x + 3y) + v(2x + 4y) = (5u + 2v)x + (3u + 4v)y \geq 100u + 60v, \quad (4.1)$$

by multiplying the first and second constraints by $u \geq 0$ and $v \geq 0$ respectively then adding them up. To ensure that the newly obtained estimate provides a lower bound for the objective function, we require that

$$5u + 2v \leq 16 \quad \text{and} \quad 3u + 4v \leq 10.$$

Finally to achieve the best optimal lower bound for the objective function, we should maximise the lower bound in (4.1)

$$\text{maximise } 100u + 60v.$$

Bringing all steps together, we end up with the following problem

$$\text{maximise } 100u + 60v.$$

subject to

$$\begin{cases} 5u + 2v \leq 16, \\ 3u + 4v \leq 10, \\ u, v \geq 0 \end{cases}$$

This is exactly the problem from the production's perspective in Example 4.1.

4.2 Primal and dual problems of LPs

Duality deals with pairs of LPs and the relationship between their solutions. One problem is called primal and the other the dual. It does not matter which problem is called the primal since the dual of the dual is the primal (see Proposition 6.5 below).

Let A be an $m \times n$ matrix, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. We define the standard LP as the primal problem.

Primal Problem:

$$\begin{aligned} (LP) \quad & \text{minimise } c^T x \\ & \text{s.t. } Ax = b, \quad x \geq 0. \end{aligned}$$

The key idea of duality begins by observing that every feasible x produces an *upper bound*, $c^T x$, for the minimum of (LP). Solving (LP) corresponds to minimising upper bounds. But what if we instead maximise lower bounds?

That is, we want to solve a problem that looks like

$$\begin{aligned} & \text{maximise } \xi^T y \\ & \text{s.t. } ???, \end{aligned}$$

so that for every feasible y , $\xi^T y \leq c^T x$ for all feasible x . How is this possible, how can ξ know about x for every feasible x ? Well, one of the only things we know about x is that $Ax = b$, so let's try $\xi = b$ and see what happens. Then,

$$b^T y = (Ax)^T y = x^T A^T y = (A^T y)^T x$$

and so, since $x \geq 0$ for all feasible x , all we need to get our desired $b^T y \leq c^T x$, for all feasible x , is to require that $A^T y \leq c$.

Therefore, (LP)'s dual problem is defined by

Dual problem:

$$\begin{aligned} (DP) \quad & \text{maximise} \quad b^T y \\ & \text{s.t.} \quad A^T y \leq c, \end{aligned}$$

where $y \in \mathbb{R}^m$.

By introducing a slack vector, the dual problem can be equivalently written as

$$\begin{aligned} (DP) \quad & \text{maximise} \quad b^T y \\ & \text{s.t.} \quad A^T y + s = c, \quad s \geq 0 \end{aligned}$$

where $s \in \mathbb{R}^n$ is called the *dual slack variable*.

Example 4.2. Find the dual problem to the LP

$$\begin{aligned} & \text{minimise} \quad 6x_1 + 8x_2 \\ & \text{s.t.} \quad 3x_1 + x_2 - x_3 = 4, \\ & \quad \quad 5x_1 + 2x_2 - x_4 = 7, \\ & \quad \quad x_1, x_2, x_3, x_4 \geq 0. \end{aligned}$$

By the definition, the dual problem of the above LP is

$$\begin{aligned} & \text{maximise} \quad 4w_1 + 7w_2 \\ & \text{s.t.} \quad 3w_1 + 5w_2 \leq 6 \\ & \quad \quad w_1 + 2w_2 \leq 8 \\ & \quad \quad -w_1 \leq 0 \\ & \quad \quad -w_2 \leq 0. \end{aligned}$$

When an LP is not given in standard form, we may first convert them to the standard form, and then write down their dual problems.

Example 4.3. Find the dual problem to the LP

$$\begin{aligned}
 &\text{minimise} && 6x_1 + 8x_2 \\
 &\text{s.t.} && 3x_1 + x_2 - x_3 \geq 4, \\
 &&& 5x_1 + 2x_2 - x_4 \leq 7, \\
 &&& x_1, x_2, x_3 \geq 0, x_4 \text{ is free.}
 \end{aligned}$$

Proposition 4.4. If the primal problem is given by (canonical form)

$$\begin{aligned}
 (LP) \quad &\text{minimise} && c^T x \\
 &\text{s.t.} && Ax \geq b, \quad x \geq 0,
 \end{aligned}$$

then its dual problem is

$$\begin{aligned}
 (DP) \quad &\text{maximise} && b^T y \\
 &\text{s.t.} && A^T y \leq c, \quad y \geq 0.
 \end{aligned}$$

Example 4.5. Find the dual problem to the LP

$$\begin{aligned}
 &\text{minimise} && 6x_1 + 8x_2 \\
 &\text{s.t.} && 3x_1 + x_2 - x_3 \geq 4, \\
 &&& 5x_1 + 2x_2 - x_4 \leq 7, \\
 &&& x_1, x_2, x_3, x_4 \geq 0.
 \end{aligned}$$

Proposition 4.6. The dual of the dual is the primal.

Proof. Suppose that the following standard LP is the primal problem:

$$\begin{aligned}
 (LP) \quad &\text{minimise} && c^T x \\
 &\text{s.t.} && Ax = b, \quad x \geq 0.
 \end{aligned}$$

According to the definition, the dual problem of the LP is given by

$$\begin{aligned}
 (DP) \quad &\text{maximise} && b^T y \\
 &\text{s.t.} && A^T y + s = c, \quad s \geq 0.
 \end{aligned}$$

We now consider the dual of (DP). First we transform this problem as the standard form.

Thus we have

$$(DP) \quad \begin{aligned} &\text{minimise} \quad -b^T y' + b^T y'' \\ &\text{s.t.} \quad A^T y' - A^T y'' + s = c, \quad s \geq 0, \quad y' \geq 0, \quad y'' \geq 0. \end{aligned}$$

This can be written as

$$(DP) \quad \begin{aligned} &\text{minimise} \quad (-b^T, b^T, 0^T) \begin{pmatrix} y' \\ y'' \\ s \end{pmatrix} \\ &\text{s.t.} \quad (A^T, -A^T, I) \begin{pmatrix} y' \\ y'' \\ s \end{pmatrix} = c, \quad \begin{pmatrix} y' \\ y'' \\ s \end{pmatrix} \geq 0. \end{aligned}$$

The dual of this standard form is given by

$$(DDP) \quad \begin{aligned} &\text{maximise} \quad c^T z \\ &\text{s.t.} \quad (A^T, -A^T, I)^T z \leq \begin{pmatrix} -b \\ b \\ 0 \end{pmatrix}, \end{aligned}$$

which can be further written as

$$(DDP) \quad \begin{aligned} &\text{maximise} \quad c^T z \\ &\text{s.t.} \quad Az \leq -b, -Az \leq b, z \leq 0. \end{aligned}$$

It is equivalent to

$$(DDP) \quad \begin{aligned} &\text{minimise} \quad -c^T z \\ &\text{s.t.} \quad -Az = b, \quad z \leq 0. \end{aligned}$$

Setting $w = -z$, the above problem becomes

$$(DDP) \quad \begin{aligned} &\text{minimise} \quad c^T w \\ &\text{s.t.} \quad Aw = b, w \geq 0. \end{aligned}$$

This is the primal problem. □

4.3 Duality theorem

Denote by F_p and F_d the feasible regions of primal and dual problems, respectively, i.e.

$$F_p = \{x : Ax = b, x \geq 0\},$$

$$F_d = \{(y, s) : A^T y + s = c, s \geq 0\}.$$

The following result shows that the objective value at any primal feasible solution is at least as large as the objective value at any feasible dual solution.

Theorem 4.7 (Weak duality Theorem). *Let x be any feasible point of the primal problem, i.e. $x \in F_p$ and y be any feasible point of the dual problem, i.e., $(y, s) \in F_d$. Then*

$$c^T x \geq b^T y.$$

Proof. For any feasible point $x \in F_p$ and $(y, s) \in F_d$, we have $Ax = b$ and $x \geq 0$, and $s = c - A^T y \geq 0$, and thus

$$c^T x - b^T y = c^T x - (Ax)^T y = x^T (c - A^T y) = x^T s \geq 0.$$

The last inequality above follows from the fact that $(x, s) \geq 0$.

□

The quantity $c^T x - b^T y$ is often called the **duality gap**. From the weak duality theorem we have the following immediate consequences:

Corollary 4.8. *If the primal LP is unbounded (i.e., $c^T x \rightarrow -\infty$), then the dual LP must be infeasible.*

Corollary 4.9. *If the dual LP is unbounded, then the primal LP must be infeasible.*

Thus, if either problem has an unbounded objective value, then the other problem possesses no feasible solution.

Corollary 4.10. *If x is feasible to the primal LP, y is feasible to the dual LP, and $c^T x = b^T y$, then x must be optimal to the primal LP and y must be optimal to the dual LP.*

Unboundedness in one problem implies infeasibility in the other problem.

- Is this property symmetric?
- Does infeasibility in one problem imply unboundedness in the other?

The answer is “not necessarily”.

Example 4.11. *Consider the primal problem*

$$\begin{aligned} & \text{minimise} && -x_1 - x_2 \\ & \text{s.t.} && x_1 - x_2 \geq 1, \\ & && -x_1 + x_2 \geq 1, \\ & && x_1, x_2 \geq 0. \end{aligned}$$

Its dual problem is given by

$$\begin{aligned} & \text{maximise} && w_1 + w_2 \\ & \text{s.t.} && w_1 - w_2 \leq -1, \\ & && -w_1 + w_2 \leq -1, \\ & && w_1, w_2 \geq 0. \end{aligned}$$

Both problems are infeasible.

The last corollary above identifies a sufficient condition for the optimality of a primal-dual pair of feasible solutions, namely that their objective values coincide. One natural question to ask is whether this is a necessary condition. The answer is yes, as shown by the next result.

Theorem 4.12 (Strong Duality Theorem). *If both the primal LP and the dual LP have feasible solutions, then they both have optimal solutions, and for any primal optimal solution x and dual optimal solution y we have that $c^T x = b^T y$.*

Proof. As will be seen in Chapter 7 one can solve both the primal and the dual problem simultaneously using the simplex method. □

Therefore, for LPs, exactly one of the following statements is true

- Both problems have optimal solutions x^* and (y^*, s^*) with $c^T x^* = b^T y^*$.
- One problem has unbounded objective value, in which case the other problem must be infeasible.
- Both problems are infeasible.

In Summary,

P optimal \Leftrightarrow D optimal

P unbounded \Rightarrow D infeasible

D unbounded \Rightarrow P infeasible

P infeasible \Rightarrow D unbounded or infeasible

D infeasible \Rightarrow P unbounded or infeasible

4.4 Network flow problems

A flow network (also known as a transportation network) is a directed graph where each edge has a capacity and each edge receives a flow. The amount of flow on an edge cannot exceed the capacity of the edge. Often in operations research, a directed graph is called a network, the vertices are called nodes and the edges are called arcs. A flow must satisfy the restriction that the amount of flow into a node equals the amount of flow out of it, unless it is a source, which has only outgoing flow, or sink, which has only incoming flow. A network can be used to model traffic in a computer network, circulation with demands, fluids in pipes, currents in an electrical circuit, or anything similar in which something travels through a network of nodes.

Definition 4.13. A network $N = (V, E)$ is a directed graph where V and E are respectively the set of vertices (nodes) and edges (arcs). Given two vertices $i, j \in V$ we denote the edge between them by (i, j) .

Definition 4.14. *The capacity of an edge is the maximum amount of flow that can pass through an edge in both directions. Formally it is a map $c : E \rightarrow \mathbb{R}^+$, called the capacity map.*

In most networks there are one or more nodes which are distinguished. A *source* is a node such that all edges through it are oriented away from it. Similarly a *sink* is a node such that all edges through it are oriented towards it.

Given a function g on the edges $g : E \rightarrow \mathbb{R}$, denote by g_{ij} the value of g on $(i, j) \in E$.

Definition 4.15. *A flow is a map $f : E \rightarrow \mathbb{R}^+$ that satisfies the following constraints:*

- *capacity constraint: The flow of an edge cannot exceed its capacity, in other words: $f_{ij} \leq c_{ij}$, for each $(i, j) \in E$;*
- *conservation of flows: The sum of the flows entering a node must equal the sum of the flows exiting that node, except for the sources and the sinks: $\sum_{i:(i,j) \in E} f_{ij} = \sum_{k:(j,k) \in E} f_{jk}$, for each $v \in V \setminus \{S, T\}$.*

In brief: flows generate at sources and terminate at sinks.

Example 4.16 (The max-flow problem). *Let $N = (V, E)$ be a directed network with only one source $S \in V$, only one sink $T \in V$ and a set of intermediate nodes. Assume the flow is never negative, namely if along an edge (i, j) the flow f_{ij} is strictly positive, then the flow f_{ji} in the opposite direction is 0. The maximal flow problem is to maximise the flow from source to sink subject to the network edge capacities and conservation of flow:*

$$\begin{aligned}
 &\text{maximise} && \sum_{i:(S,j) \in E} f_{Sj} \\
 &\text{subject to} && f_{ij} \leq c_{ij}, \text{ for each } (i, j) \in E, \\
 &&& \sum_{i:(i,j) \in E} f_{ij} = \sum_{k:(j,k) \in E} f_{jk}, \text{ for each } j \in V \setminus \{S, T\} \\
 &&& f_{ij} \geq 0 \text{ for each } (i, j) \in E.
 \end{aligned}$$

In the context of flow analysis, there is only an interest in considering how units are transferred between source and sink. Namely, we can re-write the above linear problem by considering the whole paths from S to T . Let P be the set of all possible paths from S to T . We associate with each path $p \in P$ a quantity x_p specifying how much of the flow from

S to T is being transferred along the path. With this view, the conservation constraint above is unnecessary and we obtain the following simpler linear programming problem

$$\begin{aligned} & \text{maximise} && \sum_{p \in P} x_p \\ & \text{subject to} && \sum_{p \in P: (i,j) \in p} x_p \leq c_{ij}, \text{ for each } (i,j) \in E, \\ & && x_p \geq 0, \text{ for all } p \in P. \end{aligned}$$

Example 4.17 (The min-cut problem). *In graph theory, a cut is a partition of the vertices of a graph into two disjoint subsets. Any cut determines a cut-set, the set of edges that have one endpoint in each subset of the partition. These edges are said to cross the cut. In a connected graph, each cut-set determines a unique cut, and in some cases cuts are identified with their cut-sets rather than with their vertex partitions.*

In a flow network, an s - t cut is a cut that requires the source and the sink to be in different subsets, and its cut-set only consists of edges going from the source's side to the sink's side. The capacity of an s - t cut is defined as the sum of the capacity of each edge in the cut-set

In an unweighted undirected graph, the size or weight of a cut is the number of edges crossing the cut. In a weighted graph, the value or weight is defined by the sum of the weights of the edges crossing the cut.

Let $N = (V, E)$ be a directed network with only one source $S \in V$, only one sink $T \in V$ and a set of intermediate nodes. Let y_{ij} denote the weight to assign to an edge (i, j) . The problem is to separate S and T with minimum total weighted capacity. The condition that source and sink are separated can be stated as follows: for every path $p \in P$ from the source S to the sink T , the weight of the path is at least 1.

$$\begin{aligned} & \text{Minimise} && \sum_{(i,j) \in E} c_{ij} y_{ij} \\ & \text{subject to} && \sum_{(i,j) \in p} y_{ij} \geq 1 \quad \forall p \in P \\ & && y_{ij} \geq 0 \quad (i, j) \in E. \end{aligned}$$

This is the dual problem to the max-flow problem.

The max-flow min-cut theorem is a special case of the duality theorem. It states that in a flow network, the maximum amount of flow passing from the source to the sink is

equal to the total weight of the edges in the minimum cut. The max-flow min-cut theorem has many practical applications in networks, transportation and scheduling.

4.5 Optimality Conditions & Complementary slackness

The optimality condition can be obtained by the strong duality theorem from the last lecture.

The strong duality theorem provides the condition to identify optimal solutions:

$x \in \mathbb{R}^n$ is an optimal solution of LP if and only if it satisfies the following three conditions:

- x is feasible to the primal, i.e., $Ax = b, x \geq 0$.
- There exists a $y \in \mathbb{R}^m$ such that y is feasible to the dual, i.e., $A^T y \leq c$.
- There is no duality gap, i.e., $c^T x = b^T y$.

Thus we have the following theorem.

Theorem 4.18 (Optimality Condition). *Consider the standard LP problem:*

$$\begin{aligned} & \text{minimise} \quad c^T x \\ & \text{s.t.} \quad Ax = b, \quad x \geq 0, \end{aligned}$$

where $b \in \mathbb{R}^m, c \in \mathbb{R}^n$, and A an $m \times n$ matrix. Then \bar{x} is an optimal solution to the LP if and only if there exists \bar{y} such that the following three conditions hold

$$A^T \bar{y} \leq c, \tag{4.2}$$

$$A\bar{x} = b, \quad \bar{x} \geq 0, \tag{4.3}$$

$$b^T \bar{y} = c^T \bar{x}. \tag{4.4}$$

Proof. (\Rightarrow) **Suppose that \bar{x} is optimal to the LP:**

- Then \bar{x} is feasible, so (4.3) holds.

- Since LP is feasible and bounded, the dual problem DP is feasible.
- Hence DP has an optimal solution \bar{y} with $c^T \bar{x} = b^T \bar{y}$, by the Strong Duality theorem. Therefore (4.2) and (4.4) hold.

(\Leftarrow) **Suppose there exists \bar{y} such that (4.2), (4.3), and (4.4) hold:**

- By (4.2) and (4.3), \bar{x} is feasible for LP and \bar{y} is feasible for DP.
- Therefore, by the Strong Duality theorem, there exists x^* optimal for LP and y^* optimal for DP, with $c^T x^* = b^T y^*$.
- Therefore $b^T \bar{y} \leq b^T y^* = c^T x^* \leq c^T \bar{x}$.
- Combining this with (4.4), it follows that $c^T x^* = c^T \bar{x}$ and so \bar{x} is optimal.

□

Remark. Similarly, the optimality condition for the duality problem is the same as above.

Therefore, by solving an optimality conditions, we can obtain the optimal solutions for primal and dual problems at the same time.

The optimality conditions (4.2) - (4.4), can be restated by adding the slack variable vector \bar{s} to inequality (4.2), and using the proof of the weak duality problem

$$c^T \bar{x} - b^T \bar{y} = \bar{x}^T \bar{s}.$$

Therefore, we note that at the optimal solution x^* and (\bar{y}, \bar{s}) , we have $\bar{x}^T \bar{s} = 0$, which is equivalent to $\bar{x}_j \bar{s}_j = 0$ for every $j = 1, \dots, n$ since the nonnegativity of \bar{x} and \bar{s} . Thus, at the optimal solution, one of the terms \bar{x}_j and \bar{s}_j must be zero. This is called **complementary slackness condition**.

As a result, we may restate the optimality conditions as follows.

Theorem 4.19 (Optimality Condition). *Consider the standard LP problem*

$$\begin{aligned} & \text{minimise} \quad c^T x \\ & \text{s.t.} \quad Ax = b, \quad x \geq 0 \end{aligned}$$

where $b \in \mathbb{R}^m, c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}$. Then \bar{x} is an optimal solution to the LP if and only if there exists \bar{y} such that the following three conditions hold

$$A^T \bar{y} + \bar{s} = c, \quad \bar{s} \geq 0. \quad (\text{Dual feasibility})$$

$$A\bar{x} = b, \quad \bar{x} \geq 0. \quad (\text{Primal feasibility})$$

$$\bar{x}^T \bar{s} = 0. \quad (\text{Complementary slackness condition})$$

Example 4.20. What is the optimality condition of the following LP?

$$\begin{aligned} &\text{minimise} && -2x_1 - x_2 \\ &\text{s.t.} && x_1 + 2x_2 \geq 15, \\ &&& -x_1 - x_2 \geq -20, \\ &&& x_1, x_2 \geq 0. \end{aligned}$$

Verify that whether $(15, 0)$ and $(20, 0)$ are optimal solutions to the problem.

Outline of the solution:

- Introduce slack variables x_3 and x_4 and transform the problem to the standard form

$$\begin{aligned} &\text{minimise} && -2x_1 - x_2 \\ &\text{s.t.} && x_1 + 2x_2 - x_3 = 15, \\ &&& -x_1 - x_2 - x_4 = -20, \\ &&& x_1, x_2, x_3, x_4 \geq 0. \end{aligned}$$

- The optimality conditions for this LP are given as follows:

$$\begin{aligned} &x_1 + 2x_2 - x_3 = 15, \\ &-x_1 - x_2 - x_4 = -20, \\ &x_1, x_2, x_3, x_4 \geq 0, \\ &y_1 - y_2 \leq -2, \\ &2y_1 - y_2 \leq -1, \\ &-y_1 \leq 0, -y_2 \leq 0, \\ &-2x_1 - x_2 = 15y_1 - 20y_2. \end{aligned}$$

- Consider the point $(x_1, x_2) = (15, 0)$. From the first two equalities above, we see that $(x_3, x_4) = (0, 5)$ satisfies the first three conditions. Thus at $x = (15, 0, 0, 5)^T$ the above optimality conditions are reduced to

$$\begin{aligned} y_1 - y_2 &\leq -2 \\ 2y_1 - y_2 &\leq -1 \\ -y_1 &\leq 0, -y_2 \leq 0 \\ -30 &= 15y_1 - 20y_2. \end{aligned}$$

It is easy to verify that this system is inconsistent (there exists no y satisfies these conditions). Thus, $(x_1, x_2) = (15, 0)$ is not an optimal solution to the LP.

- Consider the point $(x_1, x_2) = (20, 0)$. Clearly, this point, together with $(x_3, x_4) = (5, 0)$, satisfies the first three conditions of the optimality. We now check if there exists a y satisfying the remaining conditions, i.e.,

$$\begin{aligned} y_1 - y_2 &\leq -2 \\ 2y_1 - y_2 &\leq -1 \\ -y_1 &\leq 0, -y_2 \leq 0 \\ -40 &= 15y_1 - 20y_2. \end{aligned}$$

Clearly, $(y_1, y_2) = (0, 2)$ satisfies these conditions, thus the point $(x_1, x_2) = (20, 0)$ is an optimal solution to the LP problem.

Example 4.21. Consider the linear program

$$\begin{aligned} \min \quad & 5x_1 + 12x_2 + 4x_3 \\ \text{s.t.} \quad & x_1 + 2x_2 + x_3 = 10, \\ & 2x_1 - x_2 + 3x_3 = 8, \\ & x_1, x_2, x_3 \geq 0. \end{aligned}$$

You are given the information that x_2 and x_3 are positive in the optimal solution. Use the complementary slackness condition to solve the dual problem.

Solution: The dual problem is

$$\begin{aligned} \max \quad & 10y_1 + 8y_2 \\ \text{s.t.} \quad & y_1 + 2y_2 \leq 5, \\ & 2y_1 - y_2 \leq 12, \\ & y_1 + 3y_2 \leq 4. \end{aligned}$$

Since x_2 and x_3 are positive in the optimal solutions, the corresponding constraints in the dual problem become equalities. Thus $y = (y_1, y_2)$ is an optimal solution to the dual problem if and only if it satisfies the first constraint and the following system

$$2y_1 - y_2 = 12 \quad \text{and} \quad y_1 + 3y_2 = 4.$$

Solving this system gives $y_1 = \frac{40}{7}, y_2 = -\frac{4}{7}$, which satisfies the first constraint since

$$y_1 + 2y_2 = \frac{32}{7} < 5.$$

Thus it is an optimal solution to the dual problem.

Note that we can also use this information to find the optimal solution to the primal problem. In fact, since the first constraint in the dual problem is an inequality, by the complementary slackness conditions, the first variable in the primal problem is zero, that is $x_1 = 0$. Substituting this to the primal constraints, we obtain $x_2 = \frac{22}{7}, x_3 = \frac{26}{7}$, which are nonnegative. Also we can verify that the objective values coincide:

$$10y_1 + 8y_2 = 5x_1 + 12x_2 + 4x_3 = \frac{368}{7}.$$

Thus $y = (40/7, -4/7)$ and $x = (0, 22/7, 26/7)$ are respectively optimal solutions to the primal and the dual problems.

If the solution satisfies that $(x^*)^T s^* = 0$ and $x^* + s^* > 0$. Such solution is called *strictly complementary solution*. The following result shows that an LP always has such a solution,

Theorem 4.22 (Strict Complementarity). *If primal and dual problems both have feasible solutions, then both problems have a pair of strictly complementary solutions $x^* \geq 0$ and $s^* \geq 0$ with*

$$(x^*)^T s^* = 0, \quad x^* + s^* > 0.$$

Moreover, the supports

$$I^* = \{j : x_j^* > 0\} \text{ and } J^* = \{j : s_j^* > 0\}$$

are invariant for all pairs of strictly complementary solutions.

The proof of this theorem can be found in [Gre94] where it is proved that a primal-dual pair of optimal solutions is unique if and only if it is a strictly complementary pair of basic solutions, see also Remark 7.4 in Chapter 8 for more discussions.

4.6 More Applications of Duality Theory

Example 4.23. Solve the following LP

$$\begin{aligned} &\text{minimise} && x_1 + x_2 + 6x_3 \\ &\text{s.t.} && -x_1 + x_2 + x_3 \geq 2, \\ &&& x_1 - 2x_2 + 2x_3 \geq 6, \\ &&& x_1, x_2, x_3 \geq 0. \end{aligned}$$

Example 4.24. Using duality theory to prove the Farkas's Lemma: Only one of the following systems has a solution (feasible)

- (System 1) $Ax = b, \quad x \geq 0.$
- (System 2) $A^T y \leq 0, \quad b^T y > 0.$

Proof. If System 1 has a solution, then for any y such that $A^T y \leq 0$ we have

$$b^T y = (Ax)^T y = x^T (A^T y) \leq 0.$$

The inequality follows from the fact that $A^T y \leq 0$ and $x \geq 0$. So, the System 2 is impossible to have a solution.

Suppose that the System 1 has no solution, then the following LP

$$\begin{aligned} &\text{minimise} && 0^T x \\ &\text{s.t.} && Ax = b, \quad x \geq 0 \end{aligned}$$

is infeasible. According to the duality theorem, its dual problem which is given as

$$\begin{aligned} &\text{maximise} && b^T y \\ &\text{s.t.} && A^T y \leq 0 \end{aligned}$$

is either unbounded or infeasible. However, this dual problem is feasible ($y = 0$ a feasible point), so the dual problem must be unbounded. As a result, there exists a y such that $A^T y \leq 0, b^T y > 0$, and the System 2 has a solution.

□

Chapter 5

The Simplex Method (I): Theory

How to solve a LP problem? George B. Dantzig introduced the simplex method for solving LPs in 1947. The method has been selected by the editors of Computing in Science & Engineering / CiSE (January/February 2000, pp. 22-23, vol. 2) as the second most significant algorithm in the 20-th century. In this chapter, we start to learn this important method. This chapter will introduce essential concepts and develops the theory underlying the simplex method. The implementation and analysis of method will be discussed in subsequent chapters. The following topics will be covered

- review of Linear Algebra,
- basic feasible solutions (BFS),
- reformulation of a LP by the current BFS,
- termination of the iteration and declaration of optimality.

At the end of this chapter, students should be able to

- compute basic feasible solutions,
- formulate a LP by the current BFS,
- understand the principles of the simplex method.

[BT97, Chapter 3] and [DT97, Chapter 3] contain relevant material to this chapter that you may find useful.

5.1 Basic Feasible Solution

5.1.1 Review of Linear Algebra

For matrices, we may perform some elementary row operations. These operations are most helpful in solving a system of linear equations and in finding the inverse of a matrix.

Elementary row operations:

An elementary row operation on a matrix A is one of the following operations:

1. Row i and row j are interchanged.
2. Row i is multiplied by a nonzero scalar α .
3. Row i is replaced by row i plus α times row j .

Elementary row operation on a matrix is equivalent to *pre-multiplying* A by a specific matrix. Elementary column operation on a matrix is equivalent to *post-multiplying* A by a specific matrix.

Solving a system of linear equations:

Theorem 5.1. $Ax = b$ if and only if $A'x = b'$, where (A', b') is obtained from (A, b) by a finite number of elementary row operations. (Any elementary row operations never change the solution of the system of linear equations)

By row operations, A' can be upper triangular, and then we can solve the system by back substitution. The process of reducing A to an upper triangular matrix with ones on the diagonal is called *Gaussian Reduction* of the system, by further reduction, we may reduce A to the identity matrix, the process is called a *Gauss-Jordan* reduction of the system.

Calculation of inverse: By elementary row operations, reducing $(A \ I)$ to $(I \ B)$, then $B = A^{-1}$.

5.1.2 Existence of the Optimal Extreme Point

Still, we consider the standard form of LP:

$$\min\{c^T x : Ax = b, x \geq 0\},$$

for which we have the following result

Theorem 5.2 (Existence of Optimal Extreme Point). *Assume that the feasible region is nonempty. Then*

- *an (finite) optimal solution exists if and only if $c^T d^j \geq 0$ for $j = 1, \dots, l$, where d^1, \dots, d^l are the extreme directions of the feasible region.*
- *Otherwise, if there is an extreme direction such that $c^T d^j < 0$, then the optimal objective value of LP is unbounded.*
- *If an optimal solution exists, then at least one extreme point is optimal.*

Proof. Let x^1, x^2, \dots, x^k be the extreme points of the constraint set. And let d^1, \dots, d^l be the extreme directions. By Minkowski's theorem, any feasible point x can be represented as

$$x = \sum_{j=1}^k \lambda_j x^j + \sum_{j=1}^l \mu_j d^j,$$

where

$$\begin{aligned} \sum_{j=1}^k \lambda_j &= 1, \quad \lambda_j \geq 0, j = 1, \dots, k, \\ \mu_j &\geq 0, \quad j = 1, \dots, l. \end{aligned}$$

Thus, LP can be transformed into the following problem in variables λ and μ :

$$\begin{aligned} \text{Minimize} \quad & \sum_{j=1}^k (c^T x^j) \lambda_j + \sum_{j=1}^l (c^T d^j) \mu_j \\ \text{s.t.} \quad & \sum_{j=1}^k \lambda_j = 1, \quad \lambda_j \geq 0, j = 1, \dots, k, \\ & \mu_j \geq 0, \quad j = 1, \dots, l. \end{aligned}$$

Since the μ_j 's can be made arbitrarily large, the minimum is $-\infty$ if $c^T d^j < 0$ for some j .

Thus, the necessary condition for the existence of the optimal solution is

$$c^T d^j \geq 0, \quad \forall j = 1, \dots, l.$$

If this condition holds, we can show that the optimal solution exists and attains at an extreme point. In fact, if $c^T d^j \geq 0$ for $j \leq l$, in order to minimize the objective function, μ_j should be taken to be zero for all j . In order to minimize the first term of the objective, we simply find the minimum $c^T x^j$, say $c^T x^p$, and let $\lambda_p = 1$, and all other λ_j 's equal to zero. \square

- In summary, the optimal solution of the linear programming is finite if and only if $c^T d^j \geq 0$ for all extreme directions.
- Furthermore, if this is the case, then we find the minimum point by picking the minimum objective value among all extreme points.
- This shows that if an optimal solution exists, we must be able to find an optimal extreme point.

5.1.3 Basic Feasible Solution

To solve a linear programming algebraically, we need to address the following several issues.

- How to identify (recognize) an extreme point?
- If the current extreme point is not optimal, how to move from it to the next better extreme point?
- How to terminate the algorithm and declare optimality?

To address the first issue, we need the concept of **Basic Feasible Solution** (or basic feasible point).

Definition 5.3 (Basic Feasible Solution). *Consider the system*

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

where A is an $m \times n$ and $b \in \mathbb{R}^m$. Suppose A is full rank, i.e., $\text{Rank}(A) = m$. After possibly rearranging the columns of A , let

$$A = [B, N],$$

where B is an $m \times m$ invertible matrix and N is an $m \times (n - m)$ matrix. Denote $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}$. Then the system become

$$[B, N] \begin{bmatrix} x_B \\ x_N \end{bmatrix} = b,$$

i.e.,

$$Bx_B + Nx_N = b.$$

Note that

$$x_B = B^{-1}b, \quad x_N = 0$$

is a solution to the above system. This solution is called a **basic solution**. Thus a basic solution has at least $n - m$ zero variables. In addition, if x is a basis solution then columns of A corresponding to non-zero variables of x are linearly independent.

- B — the basic matrix (or simply basis),
- N — the nonbasic matrix (or simply nonbasis).
- x_B — basic variable
- x_N — nonbasic variable
- If $x_B \geq 0$, $(B^{-1}b, 0)$ is called a **basic feasible solution** of the system.
- If $x_B = B^{-1}b > 0$, then $x = (B^{-1}b, 0)$ is called a **non-degenerate basic feasible solution**, otherwise **degenerate basic feasible solution**.

Remark 1. The possible number of basic feasible solutions is bounded by the number of ways of extracting m columns from A to form the basis, so in general,

$$\text{the number of basic feasible solution} \leq \binom{n}{m} = \frac{n!}{m!(n - m)!}.$$

Example 5.4. Consider the polyhedral set:

$$x_1 + x_2 + x_3 = 6,$$

$$x_2 + x_4 = 3,$$

$$x_1, x_2, x_3, x_4 \geq 0.$$

The constraint matrix

$$A = [a^1, a^2, a^3, a^4] = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

The basic feasible solution corresponds to finding 2×2 invertible submatrix B . The following are possible ways of extracting B out of A .

$$1. \ B = [a^1, a^2] = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \ x_B = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = B^{-1}b = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \ x_N = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

$$2. \ B = [a^1, a^4] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ x_B = \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} = B^{-1}b = \begin{bmatrix} 6 \\ 3 \end{bmatrix}, \ x_N = \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

$$3. \ B = [a^2, a^3] = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \ x_B = \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = B^{-1}b = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \ x_N = \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

$$4. \ B = [a^2, a^4] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \ x_B = \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} = B^{-1}b = \begin{bmatrix} 6 \\ -3 \end{bmatrix}, \ x_N = \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

$$5. \ B = [a^3, a^4] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ x_B = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} = B^{-1}b = \begin{bmatrix} 6 \\ 3 \end{bmatrix}, \ x_N = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Note that the points corresponding to 1, 2, 3, and 5 are basic feasible solutions. The point obtained in 4 is not a basic feasible solution because it violates the nonnegativity restrictions. In other words, we have four basic feasible solutions, namely

$$x_1 = \begin{pmatrix} 3 \\ 3 \\ 0 \\ 0 \end{pmatrix}, \ x_2 = \begin{pmatrix} 6 \\ 0 \\ 0 \\ 3 \end{pmatrix}, \ x_3 = \begin{pmatrix} 0 \\ 3 \\ 3 \\ 0 \end{pmatrix}, \ x_4 = \begin{pmatrix} 0 \\ 0 \\ 6 \\ 3 \end{pmatrix}.$$

These points belong to \mathbb{R}^4 since after introducing the slack variables we have four variables. There basic feasible solutions, projected in \mathbb{R}^2 —that is, in the (x_1, x_2) -plane — give rise to the following four points $(3, 3)$, $(6, 0)$, $(0, 3)$ and $(0, 0)$. These points are precisely the extreme points of the feasible region.

5.1.4 Correspondence between BFSs & extreme points

Theorem 5.5. *A point is a basic feasible solution if and only if it is an extreme point.*

Proof. Let $S = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ be the feasible region. Suppose that x is a basic feasible solution. Then there exists a basis B and a non-basis N such that $x_N = 0$. Suppose that $y, z \in S$ are such that $x = \lambda y + (1 - \lambda)z$ for some $0 \leq \lambda \leq 1$. We have $0 = x_N = \lambda y_N + (1 - \lambda)z_N$. Since $y_N, z_N \geq 0$ we must have $y_N = z_N = 0$. Furthermore since $y, z \in S$, we have

$$Ay = By_B + Ny_N = b = Az = Bz_B + Nx_N$$

which implies $By_B = Bz_B = b$ so $B(y_B - z_B) = 0$. Since B is invertible, we have $y_B = z_B$. Hence $y = z$. This means x must be extreme.

Now suppose that x is not a basic feasible solution, we need to show that it is not extreme. Let $P = \{i : x_i > 0\}$ and $O = \{i : x_i = 0\}$. Since x is feasible $Ax = A_P x_P + A_O x_O = b$, thus $A_P x_P = b$. Since x is not a basic feasible solution, the columns of A_P are linearly dependent. As you have seen in Linear Algebra, there exists a non-zero vector $y_P \in \mathbb{R}^k$, where k is the cardinality of P , such that $A_P y_P = 0$. We take $y_O = 0$ and $y = \begin{pmatrix} y_P \\ y_O \end{pmatrix}$. Consider two points $x + \epsilon y$ and $x - \epsilon y$ for small ϵ . For $\epsilon > 0$ small enough, these two points are both feasible since $A(x \pm \epsilon y) = Ax \pm \epsilon Ay = b$ and $x \pm \epsilon y \geq 0$. Hence

$$x = \frac{1}{2}(x + \epsilon y) + \frac{1}{2}(x - \epsilon y).$$

So x is not extreme, as desired. □

Thus, the set of basic feasible solutions coincides with the set of extreme points (the former is an algebraic description and the latter is a geometric description). Since an LP with a finite optimal value has an optimal solution at an extreme point, an optimal basic feasible solution can always be found for such problems.

Theorem 5.6 (Fundamental theorem of Linear Programming).

i) If there is a feasible solution, i.e., $\{x : Ax = b, x \geq 0\} \neq \emptyset$, then there is a basic feasible solution

ii) If there is an optimal solution for LP, there is a basic feasible solution that is optimal.

Proof. i) Let $P = \{x : Ax = b, x \geq 0\}$ be the feasible region. Suppose that $P \neq \emptyset$ and x is a feasible solution, that is, $Ax = b$ and $x \geq 0$. We will construct a basic feasible solution from x . Suppose that only first p coordinates of x are strictly positive. Thus

$$a^1 x_1 + \dots + a^p x_p = b, \quad (5.1)$$

where a^j is the j -th column of A . We consider two cases.

Case 1: If a^1, \dots, a^p are linearly independent. Then $p \leq m = \text{rank}(A)$.

- If $p = m$, we take $B = [a^1, \dots, a^p]$, then x is a basic feasible solution with respect to this basis.
- If $p < m$, then we can find $(m - p)$ columns from a^{p+1}, \dots, a^n to form a basis B to which x is a basic feasible solution.

Case 2: If $\{a^1, \dots, a^p\}$ are linearly dependent. Then there exist y_1, \dots, y_p , not all zero, such that

$$y_1 a^1 + \dots + y_p a^p = 0. \quad (5.2)$$

Without loss of generality we can assume that $y_i > 0$ for some $i \in \{1, \dots, p\}$ (otherwise, we can multiply the above equation by -1). From (5.1) and (5.2) we deduce that

$$(x_1 - \varepsilon y_1) a^1 + \dots + (x_p - \varepsilon y_p) a^p = b.$$

Hence $A(x - \varepsilon y) = b$ where $y = (y_1, \dots, y_p, 0, \dots, 0)^T$. For sufficiently small ε we have

$$x_1 - \varepsilon y_1 \geq 0, \dots, x_p - \varepsilon y_p \geq 0.$$

We choose ε by

$$\varepsilon = \min \left\{ \frac{x_i}{y_i} \mid i = 1, \dots, p; y_i > 0 \right\}.$$

Then we have

$$(x - \varepsilon y) \geq 0 \quad \text{and} \quad A(x - \varepsilon y) = b.$$

Thus $(x - \varepsilon y)$ is a feasible solution. In addition, due to the choice of ε , the first p coordinates of $(x - \varepsilon y)$ are non-negative and at least one of them is zero. Hence we reduce p by at least 1. By repeating this process, we will either Case 1 or $p = 0$. In both cases, there will be a basic feasible solution. This completes the proof of Part i).

Part ii). Suppose that x is an optimal solution. Without loss of generality we assume that x has minimal number of positive ordinates. We will construct a basic feasible optimal solution from x . If $x = 0$ then x is basic (and the cost is zero). If $x \neq 0$, let $I = \{i_1, \dots, i_p\}$ is the set of index such that $x_{i_j} > 0$. Similarly as in Part i), there are two cases

Case 1: the column vectors $\{a^{i_1}, \dots, a^{i_p}\}$ of A are linearly independent. This case we can proceed exactly the same as in Part i) and obtain a basic feasible optimal solution. Note that the optimality does not change in this case.

Case 2: the column vectors $\{a^{i_1}, \dots, a^{i_p}\}$ of A are linearly dependent. We also proceed similarly as in the second case in Part i. However, we need to show that the optimality is preserved during the process. To show that $x - \varepsilon y$ is optimal, we consider its objective value

$$c^T(x - \varepsilon y) = c^T x - \varepsilon c^T y$$

For sufficiently small $|\varepsilon|$, the vector $x - \varepsilon y$ is a feasible solution (for positive or negative values of ε). It follows that we must have $c^T y = 0$, since otherwise, we could find a small ε of the proper sign such that $c^T(x - \varepsilon y) < c^T x$, which would violate the optimality of x . Thus $x - \varepsilon y$ is still optimal. \square

Example 5.7. Consider the LP problem in standard form with

$$A = \begin{bmatrix} -1 & 3 & 0 & 0 & 0 \\ 0 & 2 & -1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \quad b = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} \quad x = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Let us see how to construct a BFS from the feasible point $x = (3, 2, 1, 1, 1)^T$. The given x has no zero entries, so if we try to replicate the proof of the theorem 5.6, we have $m = 3$ and $p = 5$. Therefore we need to solve the system $Ay = 0$ and pick one solution, for example $y = (0, 0, 1, -1, -1)^T$.

Note that y has only one component y_3 that is strictly positive, so $\varepsilon = \min \left\{ \frac{x_i}{y_i} \mid i = 1, \dots, p; y_i > 0 \right\} = \frac{x_3}{y_3} = 1$ and $x - y = (3, 2, 0, 2, 2)^T$. We take this as our new $\tilde{x} =$

$(3, 2, 0, 2, 2)^T$ this is still a feasible point. It has one coordinate equal to 0, so let us take

$$\tilde{A} = [a^1, a^2, a^4, a^5] = \begin{bmatrix} -1 & 3 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Let us solve the system $\tilde{A}y' = 0$ and pick one solution, for example $y' = (3, 1, 2, 2)^T$. We calculate $\varepsilon = \min \left\{ \frac{\tilde{x}_i}{\tilde{y}_i} \mid i = 1, \dots, 5; y_i > 0 \right\} = \min \left\{ \frac{3}{3}, \frac{2}{1}, \frac{2}{2}, \frac{2}{2} \right\} = 1$. Again we take $\tilde{x} - \tilde{y} = (0, 1, 0, 0, 0)^T$ where $\tilde{y} = (3, 1, 0, 2, 2)^T$. This is a BFS. Infact, if we select the matrix B made of the second, fourth and fifth columns of A , we can see that $(1, 0, 0)^T = B^{-1}b$.

In the next section, we are going to study how to solve LPs by simplex method.

5.2 The simplex method

- The simplex method is to proceed from one BFS to an adjacent or neighboring one in such a way as to improve the value of the objective function (moving from an extreme point to another with a better at least not worse objective).
- It also discovers whether the feasible region is empty, and whether the optimal solution is unbounded. In practice, the method only enumerates a small portion of the extreme points of the feasible region.
- The key to the simplex method lies in recognizing the optimality of a given extreme point solution based on local consideration without having to enumerate all extreme points or basic feasible solutions.
- The basic idea of the simplex method is to confine the search to extreme points of the feasible region in a most intelligent way. The key for the simplex method is to make computers find extreme points.

5.2.1 Reformulation of LP by the current BFS

For simplicity, throughout the remainder of this course we always assume that the problem is in standard form

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b, \quad x \geq 0, \end{aligned}$$

where A is an $m \times n$ matrix with full rank, i.e., $\text{rank}(A) = m$, and b a vector in \mathbb{R}^m , and c is a vector in \mathbb{R}^n .

Initial basis and corresponding objective value

Suppose that we have a basic feasible solution $\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$. The corresponding objective value f_0 is given by

$$f_0 = (c_B^T \quad c_N^T) \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} = c_B^T B^{-1}b.$$

For this given basis B , we would like to check if it is an optimal solution of the problem.

If it is not, we would like to find another basic feasible solution at which the objective value is smaller than the current objective value f_0 .

Reformulation of the problem

Let x be a feasible solution of $Ax = b$, and let x_B and x_N be its basic and nonbasic variables for this given basis. Thus, we have

$$b = Ax = [B, N] \begin{pmatrix} x_B \\ x_N \end{pmatrix} = Bx_B + Nx_N.$$

Multiplying by B^{-1} we have

$$x_B = B^{-1}b - B^{-1}Nx_N.$$

The corresponding objective value is given by

$$\begin{aligned}
f &= c^T x \\
&= c_B^T x_B + c_N^T x_N \\
&= (c_B^T B^{-1} b - c_B^T B^{-1} N x_N) + c_N^T x_N \\
&= c_B^T B^{-1} b - (c_B^T B^{-1} N - c_N^T) x_N \\
&= f_0 - (z^T - c_N^T) x_N,
\end{aligned}$$

where the vector z is defined as

$$z^T = c_B^T B^{-1} N.$$

In components, denote by a^j the j th column vector of A , we observe that

$$B^{-1} N = [B^{-1} a^{m+1}, \dots, B^{-1} a^n] = [y^{m+1}, \dots, y^n]$$

where $y^j = B^{-1} a^j$ for $j = m+1, \dots, n$, so that

$$z_j = c_B^T B^{-1} a^j = c_B^T y^j, \quad j \in R.$$

where R is the set of indices such that a^j is a column of N for $j \in R$.

Therefore, LP can be written as

$$\begin{aligned}
&\text{Minimize} && f = f_0 - \sum_{j \in R} (z_j - c_j) x_j \\
&\text{subject to} && \sum_{j \in R} (y^j) x_j + x_B = \bar{b} \\
&&& x_j \geq 0, j \in R, \text{ and } x_B \geq 0.
\end{aligned} \tag{5.3}$$

For this problem, x_B can be viewed as a slack variable, so the above LP can be further rewritten as

$$\begin{aligned}
&\text{Minimize} && f = f_0 - \sum_{j \in R} (z_j - c_j) x_j \\
&\text{subject to} && \sum_{j \in R} (y^j) x_j \leq \bar{b} \\
&&& x_j \geq 0, j \in R.
\end{aligned} \tag{5.4}$$

Definition: The value $-(z_j - c_j) = c_j - z_j$ is referred to be as *reduced cost coefficient*.

5.2.2 Checking the Optimality

Case 1: $z_j - c_j \leq 0$ for all $j \in R$.

Proposition 5.8. *From (5.4), we see that if $(z_j - c_j) \leq 0$ for all $j \in R$, then the current basic feasible solution is optimal.*

Proof. Since $z_j - c_j \leq 0$ for all $j \in R$, from $f = f_0 - \sum_{j \in R} (z_j - c_j)x_j$ we see that $f \geq f_0$ for any feasible solution. For the current (basic) feasible solution, we know that $f = f_0$ since $x_j = 0$ for all $j \in R$. So, the current feasible solution is optimal. \square

Case 2: $z_j - c_j > 0$ for some $j \in R$.

For such j , there are two possible cases:

- If $z_j - c_j > 0$ and the vector

$$y^j = B^{-1}a^j \leq 0,$$

then LP has an unbounded solution (or simply, we say that the LP is unbounded) since we can send x_j to $+\infty$ in (5.4).

- If $z_j - c_j > 0$ and the vector

$$y^j = B^{-1}a^j = (y_{1j}, \dots, y_{mj})^T$$

has a positive component. In this case, the current objective value can be further decreased (if the current BFS is non-degenerated).

Example 5.9. *Let us consider the following problem in standard form*

$$\begin{aligned} \min \quad & x_1 + 3x_2 \\ \text{s.t.} \quad & Ax = b, \quad x \geq 0, \end{aligned}$$

where the constraint matrix A and the vector b are the same as in example 5.4:

$$A = [a^1, a^2, a^3, a^4] = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad b = \begin{pmatrix} 6 \\ 3 \end{pmatrix}.$$

Consider the basic feasible solution $x^T = (3, 3, 0, 0)$. For this solution $B = [a^1, a^2]$, $N = [a^3, a^4]$, $f_0 = 12$, $c_B^T = (1, 3)$, $c_N^T = (0, 0)$.

Then

$$[y^3, y^4] := B^{-1}N = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix},$$

and

$$z^T := c_B^T B^{-1}N = (1, 2),$$

so that $z_3 - c_3 = 1$ and $z_4 - c_4 = 2$.

So we see that for $j = 3$, $z_j - c_j > 0$. Let us look at the corresponding y^j , namely

$$y^3 = B^{-1}a^3 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

we see that it has a positive component, so we know that the value $f_0 = 12$ is not optimal.

Similarly, for $j = 4$, $z_j - c_j > 0$. Let us look at the corresponding y^j , namely

$$y^4 = B^{-1}a^4 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

again we have a positive component, which tells us that there is scope for improvement.

We shall continue this example at the end of the next subsection, after explaining how to construct a new BFS that improves the objective value.

5.2.3 Principles and justification of simplex method:

The simplex procedure moves from one BFS to an adjacent one, by introducing a non zero component into the zero part of the BFS, and removing a non zero component from it.

The criteria for “entering” and “leaving” are summarised below.

- **Entering:** If $z_k - c_k > 0$, x_k may “enter” the basis.

- **Leaving:** If

$$\frac{\bar{b}_r}{y_{rk}} \equiv \min_{1 \leq i \leq m} \left\{ \frac{\bar{b}_i}{y_{ik}} : y_{ik} > 0 \right\}$$

then x_{B_r} (the r th component of x_B) may be set to 0, namely it “leaves” the basis.

During the process, if

1. $z_j - c_j \leq 0$ for all nonbasic variables x_N , terminate. We have an optimal solution
2. If for some $k \in R$, we have $z_k - c_k > 0$, and $y_k = B^{-1}a_k \leq 0$, then LP is unbounded.
3. Otherwise, according to the above rule to choose entering variable, and leaving variable, to form new basis, and get improved basic feasible solution.

Let us explain why this works. We start from a given BFS, $(\bar{b}, 0, \dots, 0)$. Choose one of the positive number like $z_k - c_k$, and perhaps the most positive of all the $z_j - c_j, j \in R$. Set

$$x_j = 0, \text{ for } j \in R - \{k\}.$$

We want to choose x_k in such a way that $x_N = (0, \dots, 0, x_k, 0, \dots, 0)^T$ leads to a solution of the LP (5.3).

The objective value changes as we can see from (5.3)

$$f = f_0 - (z_k - c_k)x_k \quad (5.5)$$

and

$$x_B = \begin{pmatrix} x_{B_1} \\ x_{B_2} \\ \vdots \\ x_{B_r} \\ \vdots \\ x_{B_m} \end{pmatrix} = \bar{b} - (y^k)x_k = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \vdots \\ \bar{b}_r \\ \vdots \\ \bar{b}_m \end{pmatrix} - \begin{pmatrix} y_{1k} \\ y_{2k} \\ \vdots \\ y_{rk} \\ \vdots \\ y_{mk} \end{pmatrix} x_k, \quad (5.6)$$

where $y_{lk}, l = 1, \dots, m$ are the components of the vector y^k .

Assume the current basic feasible solution is not degenerated, i.e. $\bar{b} > 0$. Now let us analyse what will happen when we increase the value of x_k .

(i) From (5.5), the objective value will decrease.

(ii) For those $y_{ik} \leq 0$, when x_k increases, (5.6) implies that the component x_{B_i} will increase, continuing to be positive.

(iii) For those $y_{ik} > 0$, increasing x_k , the corresponding component x_{B_i} will decrease. So we need to increase x_k but not too much. From (5.6), the first basic variable dropping to zero corresponds to the minimum of \bar{b}_i/y_{ik} for positive y_{ik} . More precisely, we can increase x_k until it reaches the following value

$$\frac{\bar{b}_r}{y_{rk}} \equiv \min_{1 \leq i \leq m} \left\{ \frac{\bar{b}_i}{y_{ik}} : y_{ik} > 0 \right\}. \quad (5.7)$$

Let us choose $x_k = \frac{\bar{b}_r}{y_{rk}}$. Under non-degeneracy, $x_k > 0$. From (5.5) and fact that $z_k - c_k > 0$, it follows that objective can be strictly improved, as x_k increases from level 0 to \bar{b}_r/y_{rk} , a new improved feasible solution is obtained. Substituting x_k given by (5.7) into (5.6) yields the following:

$$\begin{cases} x_{B_i} = \bar{b}_i - \frac{y_{ik}}{y_{rk}} \bar{b}_r, & i = 1, 2, \dots, m, \\ x_k = \frac{\bar{b}_r}{y_{rk}}, \\ \text{all other } (x_j)'s \text{ are zero.} \end{cases} \quad (5.8)$$

From (5.8), we have $x_{B_r} = 0$, and hence at most m variables are positive. We want to prove that this is also a BFS.

Proposition 5.10. *(5.8) is the a basic feasible solution corresponding the basis*

$$\tilde{B} = [a^{B_1}, a^{B_2}, \dots, a^{B_{r-1}}, a^k, a^{B_{r+1}}, \dots, a^{B_m}],$$

which is formed by replacing r th column of B with a^k .

Proof. First of all, the matrix \tilde{B} is invertible, or equivalently its columns are linearly independent vectors. This is a consequence of the following lemma (see Linear Algebra):

Lemma 5.11. *Let a_1, a_2, \dots, a_n form a basis of \mathbb{R}^n . Then any vector a can be represented as*

$$a = \sum_{j=1}^n \lambda_j a_j.$$

Let's replace a_j by a , then the vectors $a_1, \dots, a_{j-1}, a, a_{j+1}, \dots, a_n$ are linearly independent if and only if, in the above representation, $\lambda_j \neq 0$.

To apply this lemma to our situation, we notice that by the definition of y_k , i.e., $y_k = B^{-1}a^k$, we have

$$a^k = By_k = a^{B_1}y_{1k} + \dots + a^{B_r}y_{rk} + \dots + a^{B_m}y_{mk}$$

with $y_{rk} \neq 0$. So if we replace the r th column of B by a^k , the resulting vectors are still linearly independent.

Let's verify that

$$[a^{B_1}, a^{B_2}, \dots, a^{B_{r-1}}, a^k, a^{B_{r+1}}, \dots, a^{B_m}] \begin{bmatrix} x_{B_1} \\ x_{B_2} \\ \vdots \\ x_k \\ \vdots \\ x_{B_m} \end{bmatrix} = b$$

where $x_{B_1}, x_{B_2}, \dots, x_k, \dots, x_{B_m}$ is given by (5.8). In fact, the left-hand-side

$$\begin{aligned} & \left(\bar{b}_1 - \frac{y_{1k} \bar{b}_r}{y_{rk}} \right) a^{B_1} + \dots + \frac{\bar{b}_r}{y_{rk}} a^k + \dots + \left(\bar{b}_m - \frac{y_{mk} \bar{b}_r}{y_{rk}} \right) a^{B_m} \\ &= (\bar{b}_1 - y_{1k} x_k) a^{B_1} + \dots + x_k a^k + \dots + (\bar{b}_m - y_{mk} x_k) a^{B_m} \\ &= \sum_{j=1}^m \bar{b}_j a^{B_j} + (-\bar{b}_r a^{B_r} - y_{1k} x_k a^{B_1} - \dots - y_{(r-1)k} x_k a^{B_{r-1}} + x_k a^k - \dots - y_{mk} x_k a^{B_m}) \\ &= \sum_{j=1}^m \bar{b}_j a^{B_j} + x_k \left(-\sum_{j=1}^m y_{jk} a^{B_j} + a^k \right) \\ &= B\bar{b} + x_k(a^k - B(y^k)) \\ &= b + 0 = b, \end{aligned}$$

where the last equality follows from the fact $y^k = B^{-1}a^k$. □

Chapter 6

The simplex method (II): Algorithm

In this chapter, we continue studying the simple method focusing on algorithmic aspect. We will study how to implement the simplex method in practices. The following topics will be covered

- description of the simplex algorithm,
- the simplex method in tableau format,
- further examples of the simplex methods,
- extracting information form optimal simplex tableau,

At the end of the chapter, students should be able to

- understand the main steps of the simplex algorithm,
- apply the algorithm to solve a LP problem using tableau format,
- extract useful information from the optimal simplex tableau.

[BT97, Chapter 3] and [DT97, Chapter 3] have similar material in this chapter that you may find useful.

6.1 The algorithm for minimisation problems

Here we describe the algorithm for minimisation problems in standard form.

Initial Step

Choose a starting basic feasible solution with basis B .

Main Step

Step 1. Solve the system $Bx_B = b$. Set

$$\begin{aligned} x_B &= B^{-1}b = \bar{b}, \\ x_N &= 0, \\ f &= c_B^T x_B = c_B^T B^{-1}b = c_B^T \bar{b}. \end{aligned}$$

Step 2. Set $z = c_B^T B^{-1}N$. Calculate

$$z_j - c_j \quad \text{for all nonbasic variables, i.e. } j \in R.$$

where R is the current set of indices associated with the nonbasic variables. Choose k such that

$$z_k - c_k = \max_{j \in R} \{z_j - c_j\}.$$

If $z_k - c_k \leq 0$, then stop (the current basic feasible solution is an optimal solution). Otherwise, go to Step 3.

Step 3. Calculate $y_k = B^{-1}a_k$. If $y_k \leq 0$, then stop (the optimal value of LP is unbounded). Otherwise, go to step 4.

Step 4. This step deals with removing a variable in the basis and adding one from the non basis. Let x_k enter the basis.

4.1. Determine the index r by the following *minimum ratio test*:

$$\frac{\bar{b}_r}{y_{rk}} \equiv \min_{1 \leq i \leq m} \left\{ \frac{\bar{b}_i}{y_{ik}} : y_{ik} > 0 \right\}$$

where y_{ik} denotes the i th component of y_k . Put $x_k := \frac{\bar{b}_r}{y_{rk}}$.

4.2. Update the basis B where a_k replaces a_{B_r} , update the index set R , and go to step 1.

Remarks: At each iteration, one of the following three actions is executed.

- We may stop with an optimal solution if $z_k - c_k \leq 0$;
- We may stop with an unbounded optimal value if $z_k - c_k > 0$ and $y_k \leq 0$;
- or else we generate a new basic feasible solution if $z_k - c_k > 0$ and $y_k \not\leq 0$, and the objective function strictly decreases.

Therefore, in the absence of degeneracy (i.e. $x_B > 0$), the objective strictly decreases at each iteration, and hence the basic feasible solution generated by the simplex method are distinct. Since there is only a finite number of basic feasible solutions, the method would stop in a finite number of steps with either an optimal solution or an unbounded optimal value.

Theorem 6.1 (Convergence). *In the absence of degeneracy (and assuming feasibility), the simplex method stops in a finite number of iterations, either with an optimal basic feasible optimal solution or with the conclusion that the optimal value is unbounded.*

6.2 The simplex method in Tableau Format

All the operations in simplex method can be performed in a Tableau. Let B be a starting basis. The LP problem can be represented as follows:

$$\begin{aligned} &\text{minimise} && f \\ &\text{subject to} && f - c_B^T x_B - c_N^T x_N = 0 \end{aligned} \tag{6.1}$$

$$Bx_B + Nx_N = b \tag{6.2}$$

$$x_B, x_N \geq 0.$$

This LP has the following constraint matrix

$$\begin{bmatrix} 1 & -c_B^T & -c_N^T \\ 0 & B & N \end{bmatrix}.$$

We now reformulate this LP once again. Observe that (6.2) can be written as

$$x_B + B^{-1}Nx_N = B^{-1}b.$$

Multiplying both sides by c_B^T yields

$$c_B^T x_B + c_B^T B^{-1} N x_N = c_B^T B^{-1} b.$$

Adding this equality to (6.1) leads to

$$f + 0x_B + (c_B^T B^{-1} N - c_N^T) x_N = c_B^T B^{-1} b.$$

So the LP is reformulated as

$$\begin{aligned} & \text{minimise} && f \\ & \text{subject to} && f + 0x_B + (c_B^T B^{-1} N - c_N^T) x_N = c_B^T B^{-1} b, \\ & && x_B + B^{-1} N x_N = B^{-1} b, \\ & && x_B, \quad x_N \geq 0. \end{aligned}$$

Here we think of f as a variable to be minimised. We put all information of the above problem in the following tableau:

| | x_B | x_N | RHS |
|-------|-------|--------------------------|------------------|
| f | 0 | $c_B^T B^{-1} N - c_N^T$ | $c_B^T B^{-1} b$ |
| x_B | I | $B^{-1} N$ | $B^{-1} b$ |

The tableau contains all information of the linear programming that we need to proceed with the simplex method:

- The top element in the RHS column gives us the objective value,
- The bottom element in the RHS column gives basic variable value $B^{-1}b$,
- The top element in the x_N column gives us the values $z_j - c_j$ for nonbasic variables, or in other words the minus *reduced cost*,
- The bottom element in the x_N column gives $B^{-1}N$, namely the vectors $y_j, j \in R$.

We use the following language: the row labelled by f is called *zeroth row*. The rows labelled by x_{B_i} are called *i-th row*.

To minimise f we need to go from one BFS to the next, therefore we need a scheme to do the following:

- (i) Update the basic variable and their values
- (ii) Update the $z_j - c_j$ values of the new nonbasic variables.
- (iii) Update the y_j columns.

In the next subsection, we show that all of the tasks can be simultaneously accomplished by simple *pivoting* operations which are actually equivalent to a series of matrix elementary row operations on the matrix

$$\begin{bmatrix} 1 & c_B^T & c_N^T & 0 \\ 0 & B & N & b \end{bmatrix}.$$

6.2.1 Pivoting:

In general, an element of a matrix, or of an array, which is selected first by an algorithm (e.g. Gaussian elimination, simplex algorithm, etc.), to do certain calculations is called *pivot*. Finding this element is called *pivoting*. Pivoting may be followed by an interchange of rows or columns to bring the pivot to a fixed position and allow the algorithm to proceed successfully, and possibly to reduce round-off error.

If x_k enters the basis and x_{B_r} leaves the basis, then pivoting on y_{rk} includes the following operations:

- Divide row r by y_{rk} .
- For $i = 1, \dots, m$ and $i \neq r$, update the i th row by adding to it $-y_{ik}$ times the new r th row.
- Update row zero by adding to it $-(z_k - c_k)$ times the new r th row.

The following two tables show the situation before and after pivoting operations.

Table 6.1: Before pivoting

| | x_{B_1} | | x_{B_r} | | x_{B_m} | x_j | | x_k | | RHS |
|-----------|-----------|-----|-----------|-----|-----------|-------|-------------|-------|-------------|-----------------|
| f | 0 | ... | 0 | ... | 0 | ... | $z_j - c_j$ | ... | $z_k - c_k$ | $c_B^T \bar{b}$ |
| x_{B_1} | 1 | ... | 0 | ... | 0 | ... | y_{1j} | ... | y_{1k} | \bar{b}_1 |
| \vdots | \vdots | | \vdots | | \vdots | | \vdots | | \vdots | \vdots |
| x_{B_r} | 0 | ... | 1 | ... | 0 | ... | y_{rj} | ... | y_{rk}^* | \bar{b}_r |
| \vdots | \vdots | | \vdots | | \vdots | | \vdots | | \vdots | \vdots |
| x_{B_m} | 0 | ... | 0 | ... | 1 | ... | y_{mj} | ... | y_{mk} | \bar{b}_m |

Table 6.2: After pivoting

| | x_{B_1} | | x_{B_r} | | x_{B_m} | x_j | | x_k | | RHS |
|-----------|-----------|-----|----------------------------|-----|-----------|-------|--|-------|----------|--|
| f | 0 | ... | $\frac{z_k - c_k}{y_{rk}}$ | ... | 0 | ... | $(z_j - c_j) -$ $-\frac{y_{rj}(z_k - c_k)}{y_{rk}}$ | ... | 0 | $c_B^T \bar{b} - (z_k - c_k) \frac{\bar{b}_r}{y_{rk}}$ |
| x_{B_1} | 1 | ... | $\frac{-y_{1k}}{y_{rk}}$ | ... | 0 | ... | $y_{1j} - \frac{y_{rj}}{y_{rk}} y_{1k}$ | ... | 0 | $\bar{b}_1 - \frac{y_{1k}}{y_{rk}} \bar{b}_r$ |
| \vdots | \vdots | | \vdots | | \vdots | | \vdots | | \vdots | \vdots |
| x_k | 0 | ... | $\frac{1}{y_{rk}}$ | ... | 0 | ... | $\frac{y_{rj}}{y_{rk}}$ | ... | 1 | $\frac{\bar{b}_r}{y_{rk}}$ |
| \vdots | \vdots | | \vdots | | \vdots | | \vdots | | \vdots | \vdots |
| x_{B_m} | 0 | ... | $\frac{-y_{mk}}{y_{rk}}$ | ... | 1 | ... | $y_{mj} - \frac{y_{rj}}{y_{rk}} y_{mk}$ | ... | 0 | $\bar{b}_m - \frac{y_{mk}}{y_{rk}} \bar{b}_r$ |

Implication of the pivoting operation:

- x_k entered and x_{B_r} left the basis. This is illustrated on the left-hand-side of the tableau by replacing x_{B_r} with x_k .
- The right-hand-side gives the current values of the basic variables. The nonbasic variables are kept zero.
- Suppose that the original columns for the new basic and nonbasic variables are \hat{B} and \hat{N} , respectively. The pivoting operation which is a sequence of elementary row

operations results in a new tableau that gives the new $\hat{B}^{-1}\hat{N}$ under the nonbasic variables, and updated set of $z_j - c_j$'s for the new nonbasic variables, and the values of the new basic variables and objective.

6.2.2 Algorithm in tableau Format

Initialization step:

Find an initial basic feasible solution with basis B . Form the following initial tableau.

| | x_B | x_N | RHS |
|-------|-------|--------------------------|------------------|
| f | 0 | $c_B^T B^{-1} N - c_N^T$ | $c_B^T B^{-1} b$ |
| x_B | I | $B^{-1} N$ | $B^{-1} b$ |

Main Step:

Step 1. Let $z_k - c_k = \max\{z_j - c_j : j \in R\}$.

- If $z_k - c_k \leq 0$, stop. The current point is optimal.
- Otherwise, x_k will enter, and go to Step 2

Step 2. Examine y_k .

- If $y_k \leq 0$, then stop (The LP is unbounded).
- Otherwise, go to Step 3

Step 3. Determine the index r by *minimum ratio test* to find the variable x_{B_r} to leave.

Step 4.

- Update the tableau by pivoting at y_{rk} .
- Update the basic and nonbasic variables where x_k enters the basis and x_{B_r} leaves the basis, and repeat the main step.

Example 6.2. Solve the following LP problem by the simplex method

$$\begin{aligned}
 &\text{Minimise} && x_1 + x_2 - 4x_3 \\
 &\text{s.t.} && x_1 + x_2 + 2x_3 \leq 9, \\
 &&& x_1 + x_2 - x_3 \leq 2, \\
 &&& -x_1 + x_2 + x_3 \leq 4, \\
 &&& x_1, x_2, x_3 \geq 0.
 \end{aligned}$$

Adding slack variables leads to

$$\begin{aligned}
 &\text{Minimise} && x_1 + x_2 - 4x_3 + 0x_4 + 0x_5 + 0x_6 \\
 &\text{subject to} && x_1 + x_2 + 2x_3 + x_4 = 9, \\
 &&& x_1 + x_2 - x_3 + x_5 = 2, \\
 &&& -x_1 + x_2 + x_3 + x_6 = 4, \\
 &&& x_1, x_2, x_3, x_4, x_5, x_6 \geq 0.
 \end{aligned}$$

Since $b = (9, 2, 4)^T > 0$, initial basis can be chosen as $B = [a_4, a_5, a_6] = I$. so $B^{-1}b = \bar{b} > 0$. Thus, initial tableau is given by

Initial iteration

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|-------|-----------|-------|-------|-------|-------|
| f | -1 | -1 | 4 | 0 | 0 | 0 | 0 |
| x_4 | 1 | 1 | 2 | 1 | 0 | 0 | 9 |
| x_5 | 1 | 1 | -1 | 0 | 1 | 0 | 2 |
| x_6 | -1 | 1 | 1* | 0 | 0 | 1 | 4 |

- From this table, we see that x_3 is a nonbasic variable, and $z_3 - c_3 = 4 > 0$. Thus, the current basic feasible solution $(x_4, x_5, x_6) = (9, 2, 4)$ is not optimal. x_3 will enter the basis.
- Now we are going to determine which vector among x_4, x_5, x_6 will leave the basis. By checking $y_3 = (2, -1, 1)^T$, we note that it has two positive components. By the minimum ratio test: $\min\{\frac{9}{2}, \frac{4}{1}\} = 4$. The vector x_6 leaves the basis, and the marked elements in the above table is the pivoting element.

- We do pivoting operations to get a new basic feasible solution. Thus the algorithm proceeds to the next iteration.

Iteration 1.

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-----------------|-------|-------|-------|-------|-------|-------|
| f | 3 | -5 | 0 | 0 | 0 | -4 | -16 |
| x_4 | \mathcal{Z}^* | -1 | 0 | 1 | 0 | -2 | 1 |
| x_5 | 0 | 2 | 0 | 0 | 1 | 1 | 6 |
| x_3 | -1 | 1 | 1 | 0 | 0 | 1 | 4 |

- x_1 is a nonbasic variable, and $z_1 - c_1 = 3 > 0$. Thus, the current basic feasible solution $(x_4, x_5, x_3) = (1, 6, 4)$ is not optimal. x_1 will enter the basis.
- Now we determine which vector among x_4, x_5, x_3 will leave the basis. By checking $y_1 = (3, 0, -1)^T$ and the minimum ratio test indicates that x_4 should leave the basis.
- We do pivoting operations, and get the following tableau:

Iteration 2.

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | -4 | 0 | -1 | 0 | -2 | -17 |
| x_1 | 1 | -1/3 | 0 | 1/3 | 0 | -2/3 | 1/3 |
| x_5 | 0 | 2 | 0 | 0 | 1 | 1 | 6 |
| x_3 | 0 | 2/3 | 1 | 1/3 | 0 | 1/3 | 13/3 |

Now, $z_j - c_j \leq 0$ for all nonbasic variables, and thus an optimal solution is found and it is given by

$$x_1^* = 1/3, \quad x_2^* = 0, \quad x_3^* = 13/3.$$

The optimal value is

$$z^* = -17.$$

The optimal basis B consists of the column a_1, a_5 and a_3 , i.e., $B = [a_1, a_5, a_3]$.

6.3 Degeneracy and cycling in the simplex method

When applying the simplex method to solve a linear programming problem, it is possible that one starts from some basic feasible solution and ends up at the exact same basic feasible solution after a sequence of pivots. This phenomena is called “cycling” and can occur when a basic feasible solution is degenerate. In the worst scenario, the simplex method can cycle forever. The first cycling example was given by Hoffman in 1952 which had 11 variables and 3 equations. Since then, many examples have been constructed. To demonstrate the cycling phenomena, we consider the following famous example which was given by E. M. L. Beale in 1955 [Bea55]. It is conjectured that this is the simplest example of cycling which is not totally degenerate in the sense that none can be constructed with fewer variables regardless of the number of equations.

Example 6.3 (Beale 1955). *Solve the following LP by using simplex method:*

$$\begin{aligned}
 &\text{Minimise} && -\frac{3}{4}x_4 + 20x_5 - \frac{1}{2}x_6 + 6x_7 \\
 &\text{subject to} && x_1 + \frac{1}{4}x_4 - 8x_5 - x_6 + 9x_7 = 0, \\
 &&& x_2 + \frac{1}{2}x_4 - 12x_5 - \frac{1}{2}x_6 + 3x_7 = 0, \\
 &&& x_3 + x_6 = 1, \\
 &&& x_1, \dots, x_7 \geq 0.
 \end{aligned}$$

In each iteration below, the pivoting entry is marked by an asterisk.

Initial tableau

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | 0 | 0 | 3/4 | -20 | 1/2 | -6 | 0 |
| x_1 | 1 | 0 | 0 | 1/4* | -8 | -1 | 9 | 0 |
| x_2 | 0 | 1 | 0 | 1/2 | -12 | -1/2 | 3 | 0 |
| x_3 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |

Iteration 1. *The maximum value of the minus reduced cost is 3/4, so $k = 4$. The minimum ratio test is satisfied for $r = 1, 2$, for example choose $r = 1$. So we pivot with respect to y_{14}*

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| f | -3 | 0 | 0 | 0 | 4 | 7/2 | -33 | 0 |
| x_4 | 4 | 0 | 0 | 1 | -32 | -4 | 36 | 0 |
| x_2 | -2 | 1 | 0 | 0 | 4* | 3/2 | -15 | 0 |
| x_3 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |

Iteration 2. The maximum value of the minus reduced cost is 4, so $k = 5$. The minimum ratio test is satisfied for $r = 2$, so we pivot with respect to y_{25}

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| f | -1 | -1 | 0 | 0 | 0 | 2 | -18 | 0 |
| x_4 | -12 | 8 | 0 | 1 | 0 | 8* | -84 | 0 |
| x_5 | -1/2 | 1/4 | 0 | 0 | 1 | 3/8 | -15/4 | 0 |
| x_3 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |

Iteration 3. Here we could pivot both at y_{46} and y_{56} , we choose y_{46} :

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| f | 2 | -3 | 0 | -1/4 | 0 | 0 | 3 | 0 |
| x_6 | -3/2 | 1 | 0 | 1/8 | 0 | 1 | -21/2 | 0 |
| x_5 | 1/16 | -1/8 | 0 | -3/64 | 1 | 0 | 3/16* | 0 |
| x_3 | 3/2 | -1 | 1 | -1/8 | 0 | 0 | 21/2 | 1 |

Iteration 4.

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| f | 1 | -1 | 0 | 1/2 | -16 | 0 | 0 | 0 |
| x_6 | 2* | -6 | 0 | -5/2 | 56 | 1 | 0 | 0 |
| x_7 | 1/3 | -2/3 | 0 | -1/4 | 16/3 | 0 | 1 | 0 |
| x_3 | -2 | 6 | 1 | 5/2 | -56 | 0 | 0 | 1 |

Iteration 5.

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|---------|-------|--------|-------|--------|-------|-------|
| f | 0 | 2 | 0 | $7/4$ | -44 | $-1/2$ | 0 | 0 |
| x_1 | 1 | -3 | 0 | $-5/4$ | 28 | $1/2$ | 0 | 0 |
| x_7 | 0 | $1/3^*$ | 0 | $1/6$ | -4 | $-1/6$ | 1 | 0 |
| x_3 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |

Iteration 6.

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|--------|-------|-------|
| f | 0 | 0 | 0 | $3/4$ | -20 | $1/2$ | -6 | 0 |
| x_1 | 1 | 0 | 0 | $1/4$ | -8 | -1 | 9 | 0 |
| x_2 | 0 | 1 | 0 | $1/2$ | -12 | $-1/2$ | 3 | 0 |
| x_3 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |

This tableau is exactly as the initial tableau. Therefore, the simplex method will cycle endlessly. We observe that the basic feasible solution in each iteration is degenerate (but not totally degenerate) and the objective function actually does not change.

It is this degeneracy that causes cycling. We see that throughout the process we were faced several times by a non-unique choice of pivoting. There are several ways to avoid cycling. The most commonly used method is to fix a rule to make choices when faced with more than 1 possible pivoting candidate. The most common rule is Bland's rule which is formally stated as follows:

1. Choose the lowest-numbered (i.e., leftmost) nonbasic column with a negative reduced cost - in the tableau this means "choose the leftmost column with a positive entry y_{il} .
2. If there are more than one row that satisfy the minimum ratio test, choose the row with the lowest-numbered column basic variable in it.

This means that in the previous example we should pivot w.r.t. y_{24} in the first iteration.

Example 6.4. If Bland's rule does not work, we can modify the rule and start again. For example we can change the rule by saying that instead of the lowest-numbered nonbasic column with a negative reduced cost we should choose the highest, or that instead of the highest numbered column basic variable, we should choose the lowest.

6.4 Efficiency of the simplex method

An interesting property of the simplex method is that it is remarkably efficient in practice. However, in 1972, Klee and Minty [KM72] provided an example, which is now known as the Klee–Minty cube, showing that the worst-case complexity of simplex method as formulated by Dantzig is exponential time. Since then the Klee–Minty cube has been used to analyze the performance of many algorithms both in the worst case and on average. The Klee–Minty example is given by

Example 6.5 (Klee–Minty cube).

$$\begin{aligned}
 &\text{Maximise} && 2^{n-1}x_1 + 2^{n-2}x_2 + \dots + 2x_{n-1} + x_n \\
 &\text{subject to} && x_1 \leq 5, \\
 &&& 4x_1 + x_2 \leq 25, \\
 &&& 8x_1 + 4x_2 + x_3 \leq 125, \\
 &&& \vdots \\
 &&& 2^n x_1 + 2^{n-1}x_2 + \dots + 4x_{n-1} + x_n \leq 5^n, \\
 &&& x_1, \dots, x_n \geq 0.
 \end{aligned}$$

The LP has n variables, n constraints and 2^n extreme points. The simplex method, starting from the origin, will go through each of the extreme points before obtaining the optimal solution at $(0, 0, \dots, 5^n)$. To illustrate this, we consider the Klee–Minty cube for the case $n = 3$, that is

$$\begin{aligned}
 &\text{Maximise} && 4x_1 + 2x_2 + x_3 \\
 &\text{subject to} && x_1 \leq 5, \\
 &&& 4x_1 + x_2 \leq 25, \\
 &&& 8x_1 + 4x_2 + x_3 \leq 125, \\
 &&& x_1, x_2, x_3 \geq 0.
 \end{aligned}$$

It is easy to see that $4x_1 + 2x_2 + x_3 \leq 8x_1 + 4x_2 + x_3 = 125$ and the inequality becomes equality when $x_1 = x_2 = 0, x_3 = 125$. Thus the optimal value is 125 and is achieved at $(x_1, x_2, x_3) = (0, 0, 125)$. Adding s_1, s_2, s_3 as slack variables to obtain the following

problem

$$\begin{aligned}
 &\text{Maximise} && 4x_1 + 2x_2 + x_3 \\
 &\text{subject to} && x_1 + s_1 = 5, \\
 &&& 4x_1 + x_2 + s_2 = 25, \\
 &&& 8x_1 + 4x_2 + x_3 + s_3 = 125, \\
 &&& x_1, x_2, x_3, s_1, s_2, s_3 \geq 0.
 \end{aligned}$$

Below are all the tableaux

| | x_1 | x_2 | x_3 | s_1 | s_2 | s_3 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-----|
| $-f$ | 4 | 2 | 1 | 0 | 0 | 0 | 0 |
| s_1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| s_2 | 4 | 1 | 0 | 0 | 0 | 0 | 0 |
| s_3 | 8 | 4 | 1 | 0 | 0 | 1 | 0 |

| | x_1 | x_2 | x_3 | s_1 | s_2 | s_3 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-----|
| $-f$ | 0 | 2 | 1 | -4 | 0 | 0 | -20 |
| x_1 | 1 | 0 | 0 | 1 | 0 | 0 | 5 |
| s_2 | 0 | 1 | 0 | -4 | 1 | 0 | 5 |
| s_3 | 0 | 4 | 1 | -8 | 0 | 1 | 85 |

| | x_1 | x_2 | x_3 | s_1 | s_2 | s_3 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-----|
| $-f$ | 0 | 0 | 1 | 4 | -2 | 0 | -30 |
| x_1 | 1 | 0 | 0 | 1 | 0 | 0 | 5 |
| x_2 | 0 | 1 | 0 | -4 | 1 | 0 | 5 |
| s_3 | 0 | 0 | 1 | 8 | -4 | 1 | 65 |

| | x_1 | x_2 | x_3 | s_1 | s_2 | s_3 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-----|
| $-f$ | -4 | 0 | 1 | 0 | -2 | 0 | -50 |
| s_1 | 0 | 0 | 0 | 1 | 0 | 0 | 5 |
| x_2 | 4 | 1 | 0 | 0 | 1 | 0 | 25 |
| s_3 | -8 | 0 | 1 | 0 | -4 | 1 | 25 |

| | x_1 | x_2 | x_3 | s_1 | s_2 | s_3 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-----|
| $-f$ | 4 | 0 | 0 | 0 | 2 | -1 | -75 |
| s_1 | 1 | 0 | 0 | 1 | 0 | 0 | 5 |
| x_2 | 4 | 1 | 0 | 0 | 1 | 0 | 25 |
| x_3 | -8 | 0 | 1 | 0 | -4 | 1 | 25 |

| | x_1 | x_2 | x_3 | s_1 | s_2 | s_3 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-----|
| $-f$ | 0 | 0 | 0 | -4 | 2 | -1 | -95 |
| x_1 | 1 | 0 | 0 | 1 | 0 | 0 | 5 |
| x_2 | 0 | 1 | 0 | -4 | 1 | 0 | 5 |
| x_3 | 0 | 0 | 1 | 8 | -4 | 1 | 65 |

| | x_1 | x_2 | x_3 | s_1 | s_2 | s_3 | RHS |
|-------|-------|-------|-------|-------|-------|-------|------|
| $-f$ | 0 | -2 | 0 | 4 | 0 | -1 | -105 |
| x_1 | 1 | 0 | 0 | 1 | 0 | 0 | 5 |
| s_2 | 0 | 1 | 0 | -4 | 1 | 0 | 5 |
| x_3 | 0 | 4 | 1 | -8 | 0 | 1 | 85 |

| | x_1 | x_2 | x_3 | s_1 | s_2 | s_3 | RHS |
|-------|-------|-------|-------|-------|-------|-------|------|
| $-f$ | -4 | -2 | 0 | 0 | 0 | -1 | -125 |
| s_1 | 1 | 0 | 0 | 1 | 0 | 0 | 5 |
| s_2 | 4 | 1 | 0 | 0 | 1 | 0 | 25 |
| x_3 | 8 | 4 | 1 | 0 | 0 | 1 | 125 |

Thus the simplex method applied to the Klee-Minty cube for $n = 3$, starting at the origin, goes through all of the 8 vertices before reaching the optimal value.

Remark 2. *There are other algorithms for solving linear-programming problems. In particular, interior-point methods refer to class of methods for solving linear programming problems that move through the interior of the feasible region such as Khachiyan's ellipsoidal algorithm and Karmarkar's algorithm. This is in contrast to the simplex algorithm, which finds an optimal solution by traversing the edges between vertices of the feasible region. Interior-point methods often have worst-case polynomial complexity; however, in most practical applications, the simplex method is often very efficient.*

Remark 3 (Open problems in Linear Programming). *In 1998 the mathematician Stephen Smale (Fields Medal in 1966) compiled a list of 18 problems in mathematics to be solved in the 21st century, known as Smale's problems. This list was compiled in the spirit of Hilbert's famous list of problems produced in 1900 for the 20th century. The 9-th problem in the list is the following*

The linear programming problem: Find a strongly-polynomial time algorithm which for given matrix $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ decides whether there exists $x \in \mathbb{R}^n$ with $Ax \geq b$.

6.5 Further Examples for Simplex Methods & Degeneracy

Example 6.6. *Solve the following LP by using simplex method:*

$$\begin{aligned} \text{Minimise} \quad & -3x_1 + x_2 \\ \text{s.t.} \quad & x_1 + 2x_2 \leq 4, \\ & -x_1 + x_2 \leq 1, \\ & x_1, x_2, x_3, x_4 \geq 0. \end{aligned}$$

Example 6.7. *Solve the following LP by using simplex method:*

$$\begin{aligned} \text{Minimise} \quad & -x_1 - 3x_2 \\ \text{s.t.} \quad & x_1 - 2x_2 \leq 4, \\ & -x_1 + x_2 \leq 3, \\ & x_1, x_2 \geq 0. \end{aligned}$$

The following tables and figure summarise the result.

| | x_1 | x_2 | x_3 | x_4 | RHS |
|-------|-------|----------------|-------|-------|-------|
| f | 1 | 3 | 0 | 0 | 0 |
| x_3 | 1 | -2 | 1 | 0 | 4 |
| x_4 | -1 | $\mathbf{1}^*$ | 0 | 1 | 3 |

| | x_1 | x_2 | x_3 | x_4 | RHS |
|-------|-------|-------|-------|-------|-------|
| f | 4 | 0 | 0 | -3 | -9 |
| x_3 | -1 | 0 | 1 | 2 | 10 |
| x_2 | -1 | 1 | 0 | 1 | 3 |

Since $z_1 - c_1 = 4 > 0$, x_1 should enter the basis, however, the corresponding $y_1 = (-1, -1)^T \leq 0$. Thus, the LP problem has unbounded optimal value.

Example 6.8. Solve the following LP by using simplex method

$$\begin{aligned}
 &\text{Maximise} && -3x_1 + x_2 \\
 &\text{s.t.} && x_1 - 3x_2 \geq -3, \\
 &&& 2x_1 + x_2 \geq -2, \\
 &&& 2x_1 + x_2 \leq 8, \\
 &&& x_1 \text{ and } x_2 \text{ are unrestricted.}
 \end{aligned}$$

Degeneracy cases

When the right-hand-side vector \bar{b} is not positive (at least one of its components is zero), we can still use the simplex method to solve the problem. The following is such an example.

Example 6.9 (Investment problem). Someone has \$ 20,000 to finance various investments. There are five categories of investments, each with an associated return and risk:

| <i>Investment</i> | <i>Return</i> | <i>Risk</i> |
|------------------------------|---------------|-------------|
| <i>Mortgages</i> | <i>8</i> | <i>2</i> |
| <i>Commercial loans</i> | <i>10</i> | <i>6</i> |
| <i>Government securities</i> | <i>0</i> | <i>6</i> |
| <i>Personal loans</i> | <i>12</i> | <i>8</i> |
| <i>Savings</i> | <i>3</i> | <i>0</i> |

Their goal is to allocate the money to the categories so as to maximise the return, subject to the following constraints:

- (a) An average risk is no more than 6 (all averages taken over the invested money, not over the savings).
- (b) The amount in mortgages and personal loans combined should be no higher than the amount in commercial.

Model the problem:

Since there is no return from government securities, we do not consider such an investment. Denote

- x_1 — the amount invested in mortgages,
- x_2 — the amount allocated to commercial loans,
- x_3 and x_4 are the amounts invested in personal loans and savings, respectively.

Note that the constraint (a) implies that

$$\frac{2x_1 + 6x_2 + 8x_3}{x_1 + x_2 + x_3} \leq 6,$$

which can be written as

$$-4x_1 + 2x_3 \leq 0.$$

We may formulate the problem as follows:

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|-------|-------|-------|-------|-------|---------|
| f | 5 | 7 | 9 | 0 | 0 | 0 | -60,000 |
| x_5 | 1 | -1 | 1* | 0 | 1 | 0 | 0 |
| x_6 | -4 | 0 | 2 | 0 | 0 | 1 | 0 |
| x_4 | 1 | 1 | 1 | 1 | 0 | 0 | 20,000 |

Replacing x_5 by x_3 , the next table will be

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|-------|-------|-------|-------|-------|---------|
| f | -4 | 16 | 0 | 0 | -9 | 0 | -60,000 |
| x_3 | 1 | -1 | 1 | 0 | 1 | 0 | 0 |
| x_6 | -6 | 2* | 0 | 0 | -2 | 1 | 0 |
| x_4 | 0 | 2 | 0 | 1 | -1 | 0 | 20,000 |

Replacing x_6 by x_2 , we have

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|-------|-------|-------|-------|-------|---------|
| f | 44 | 0 | 0 | 0 | 7 | -8 | -60,000 |
| x_3 | -2 | 0 | 1 | 0 | 0 | 1/2 | 0 |
| x_2 | -3 | 1 | 0 | 0 | -1 | 1/2 | 0 |
| x_4 | 6* | 0 | 0 | 1 | 1 | -1 | 20,000 |

Replacing x_4 by x_1 , we have

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|-------|-------|-------|-------|-------|------------|
| f | 0 | 0 | 0 | -22/3 | -1/3 | -2/3 | -620,000/3 |
| x_3 | 0 | 0 | 1 | 1/3 | 1/3 | 1/6 | 20,000/3 |
| x_2 | 0 | 1 | 0 | 1/2 | -1/2 | 0 | 10,000 |
| x_1 | 1 | 0 | 0 | 1/6 | 1/6 | -1/6 | 10,000/3 |

The optimal solution is given by

$$(x_1^*, x_2^*, x_3^*, x_4^*) = \left(\frac{10,000}{3}, 10,000, \frac{20,000}{3}, 0 \right).$$

Thus, he will allocate \$ 10,000/3 to mortgages, \$10,000 to commercial loans, \$20,000/3 to personal loans, and no allocation to savings. Note that the optimal value of problem

(P_1) is $-620,000/3$, and thus the maximum return for the original problem (P_2) is

$$f^* = \frac{620,000}{3}.$$

6.6 Extracting Information from the Optimal Simplex Tableau

Suppose that we get the final tableau (optimal tableau) of simplex method. What information can be obtained from this final tableau? and how this information can be used?

Example 6.10. *The optimal tableau is given as follows:*

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|-------|-------|---------|--------|--------|--------------|
| f | 0 | 0 | 0 | $-22/3$ | $-1/3$ | $-2/3$ | $-620,000/3$ |
| x_3 | 0 | 0 | 1 | $1/3$ | $1/3$ | $1/6$ | $20,000/3$ |
| x_2 | 0 | 1 | 0 | $1/2$ | $-1/2$ | 0 | 10,000 |
| x_1 | 1 | 0 | 0 | $1/6$ | $1/6$ | $-1/6$ | $10,000/3$ |

The optimal solution is given by

$$\begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \\ x_4^* \\ x_5^* \\ x_6^* \end{bmatrix} = \begin{bmatrix} 10,000/3 \\ 10,000 \\ 20,000/3 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

The optimal value of objective function is

$$f^* = -620,000/3.$$

6.6.1 Find the inverse of optimal basis

At each iteration, the current B^{-1} can be obtained immediately from the current tableau, and the inverse of the optimal basis can be found in optimal tableau.

Assume that the original tableau has an identity matrix. The process of reducing the basis matrix B of the original tableau to an identity matrix in the current tableau, is equivalent to premultiplying rows 1 through m of the original tableau by B^{-1} to produce the current tableau. This also converts the identity matrix of the original tableau to B^{-1} . Therefore, B^{-1} can be extracted from the current tableau as the submatrix in rows 1 through m under the original identity columns.

Example 6.11. Consider the problem

$$\begin{aligned}
 &\text{Minimise} && -8x_1 - 10x_2 - 12x_3 - 3x_4 \\
 &\text{subject to} && x_1 - x_2 + x_3 \leq 0, \\
 &&& -4x_1 + 2x_3 \leq 0, \\
 &&& x_1 + x_2 + x_3 + x_4 = 20,000, \\
 &&& x_1, x_2, x_3, x_4 \geq 0.
 \end{aligned}$$

Introducing the slack variables x_5, x_6 leads to

$$\begin{aligned}
 &\text{minimise} && -8x_1 - 10x_2 - 12x_3 - 3x_4 \\
 &\text{subject to} && x_1 - x_2 + x_3 && + x_5 && = 0, \\
 &&& -4x_1 + 2x_3 && && + x_6 && = 0, \\
 &&& x_1 + x_2 + x_3 + x_4 && && && = 20,000, \\
 &&& x_1, x_2, x_3, x_4, x_5, x_6 \geq 0.
 \end{aligned}$$

Initial table is given as

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|-------|-------|-------|-------|-------|--------|
| f | 8 | 10 | 12 | 3 | 0 | 0 | 0 |
| x_5 | 1 | -1 | 1 | 0 | 1 | 0 | 0 |
| x_6 | -4 | 0 | 2 | 0 | 0 | 1 | 0 |
| x_4 | 1 | 1 | 1 | 1 | 0 | 0 | 20,000 |

Final tableau (optimal tableau) is given by

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|-------|-------|---------|--------|--------|--------------|
| f | 0 | 0 | 0 | $-22/3$ | $-1/3$ | $-2/3$ | $-620,000/3$ |
| x_3 | 0 | 0 | 1 | $1/3$ | $1/3$ | $1/6$ | $20,000/3$ |
| x_2 | 0 | 1 | 0 | $1/2$ | $-1/2$ | 0 | 10,000 |
| x_1 | 1 | 0 | 0 | $1/6$ | $1/6$ | $-1/6$ | $10,000/3$ |

What is B and B^{-1} ?

$$B = [a^3, a^2, a^1] = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -4 \\ 1 & 1 & 1 \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} 1/3 & 1/6 & 1/3 \\ -1/2 & 0 & 1/2 \\ 1/6 & -1/6 & 1/6 \end{bmatrix}.$$

Example 6.12. The following simplex tableau shows the optimal solution of a linear programming. It is known that x_4 and x_5 are the slack variables in the first and second constraints of the original problem. The constraints of the original problem are of the ‘ \leq ’ type.

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|--------|-------|--------|-------|-------|
| f | 0 | -2 | 0 | -3 | -2 | -35 |
| x_3 | 0 | $1/4$ | 1 | $1/2$ | 0 | $5/2$ |
| x_1 | 1 | $-1/2$ | 0 | $-1/6$ | $1/3$ | $5/2$ |

The initial identity basis corresponds to x_4 and x_5 . So in the optimal tableau, the inverse of the optimal basis B is given by

$$B^{-1} = \begin{bmatrix} 1/2 & 0 \\ -1/6 & 1/3 \end{bmatrix}.$$

6.6.2 Recovering original LP problems from the optimal tableau

In many situations, from the optimal tableau, we may recover the original LP problem.

Example 6.13. (Same example as Example 6.12) The following simplex tableau shows the optimal solution of a linear programming. It is known that x_4 and x_5 are the slack

variables in the first and second constraints of the original problem. The constraints are of the ' \leq ' type

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | -2 | 0 | -3 | -2 | -35 |
| x_3 | 0 | 1/4 | 1 | 1/2 | 0 | 5/2 |
| x_1 | 1 | -1/2 | 0 | -1/6 | 1/3 | 5/2 |

1. Determine optimal basis B .

The basic variables are x_3 and x_1 , thus, $B = [a_3, a_1]$ and

$$B^{-1} = \begin{bmatrix} 1/2 & 0 \\ -1/6 & 1/3 \end{bmatrix} \Rightarrow B = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}$$

Therefore, the third and first columns of A are obtained.

2. Determine the second column of A , i.e., a^2 .

Since $B^{-1}a^2 = y_2$ in the simplex tableau, where B^{-1} and y_2 are known, i.e.,

$$\begin{bmatrix} 1/2 & 0 \\ -1/6 & 1/3 \end{bmatrix} a^2 = \begin{pmatrix} 1/4 \\ -1/2 \end{pmatrix}.$$

Solving this system yields

$$a^2 = \begin{pmatrix} 1/2 \\ -5/4 \end{pmatrix}$$

Therefore, A is recovered.

3. Find the RHS vector b .

Notice that $\bar{b} = (5/2, 5/2)$ and

$$B^{-1}b = \bar{b}$$

Thus solving the system

$$\begin{bmatrix} 1/2 & 0 \\ -1/6 & 1/3 \end{bmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 5/2 \\ 5/2 \end{pmatrix}$$

leads to

$$b = \begin{pmatrix} 5 \\ 10 \end{pmatrix}$$

4. Determine the cost coefficient c of the objective - this will be done by using the dual simplex method.

Chapter 7

The Dual Simplex Method (non-examinable)

In this chapter, we study another extension/variant of the simplex method, the dual simplex method. This method is useful and necessary in, for instances, the following situations.

- In certain instances, it is difficult to find a starting basic feasible solution (i.e. $\bar{b} \geq 0$) to a linear program without adding artificial variables.
- However, in these same instances, it is often possible to find a starting basis, which is not necessary to be primal feasible, but which is dual feasible (i.e., all $z_j - c_j \leq 0$ for minimization problem).
- The dual simplex method is a variant of the simplex method that would produce a series of simplex tableau that maintain dual feasibility and complementary slackness and strive toward primal feasibility.
- The dual simplex method solves the dual problem directly on the (primal) simplex tableau. At each iteration we move from a basic feasible solution of the dual problem to an improved basic feasible solution until optimality of the dual (and also the primal) is reached, or else until we conclude that the dual is unbounded and that the primal is infeasible.

7.1 Dual information from the (primal) optimal simplex tableau

Consider the following LP in canonical form:

$$\begin{aligned} & \text{Minimise} && c^T x \\ & \text{subject to} && Ax \geq b \\ & && x \geq 0. \end{aligned}$$

As seen in Proposition 4.4, the dual problem is the following

$$\begin{aligned} & \text{Maximise} && b^T w \\ & \text{subject to} && A^T w \leq c \\ & && w \geq 0. \end{aligned}$$

As usual we assume that A is a $m \times n$ matrix of maximal rank m and we introduce m slack variables x_{n+1}, \dots, x_{n+m} in the primal problem to transform it to the standard form.

Assume we run the simplex method for a few iterations and reach an optimal tableau:

| | x_1 | x_2 | \cdots | x_n | x_{n+1} | \cdots | x_{n+m} | RHS |
|-----------|-------------|-------------|----------|-------------|---------------------|----------|---------------------|-------------|
| f | $z_1 - c_1$ | $z_2 - c_2$ | \cdots | $z_n - c_n$ | $z_{n+1} - c_{n+1}$ | \cdots | $z_{n+m} - c_{n+m}$ | $c_B b$ |
| x_{B_1} | y_{11} | y_{12} | \cdots | y_{1n} | $y_{1,n+1}$ | \cdots | $y_{1,n+m}$ | \bar{b}_1 |
| x_{B_2} | y_{21} | y_{22} | \cdots | y_{2n} | $y_{2,n+1}$ | \cdots | $y_{2,n+m}$ | \bar{b}_2 |
| \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots |
| x_{B_m} | y_{m1} | y_{m2} | \cdots | y_{mn} | $y_{m,n+1}$ | \cdots | $y_{m,n+m}$ | \bar{b}_m |

where we assume that for the indices i such that x_i is one of the base variables, say x_{B_i} , we take the column entries y_{ji} to be all 0 except $y_{B_i i}$ that is equal to 1. Similarly for those columns, the values of $z_i - c_i$ are set to be 0. Because the solution is optimal, the slack variables can be chosen to be 0 (they don't contribute to the final cost), so for simplicity assume that the optimal BFS contains the slack variables within the non basic part.

From this optimal tableau, we may also get the optimal solution to the dual problem:

Theorem 7.1. *At optimality of the primal minimisation problem in the canonical form,*

$$\bar{w}^T := c_B^T B^{-1}$$

is an optimal solution to the dual problem and has the following components

$$w_i = c_{n+i} - z_{n+i}, \quad i = 1, \dots, m.$$

Proof. The tableau for the primal is optimal if and only if

$$z_j - c_j \leq 0, \quad \forall j \in R,$$

where R is the set of indices corresponding to the non-basic variables. Since, by definition of z and \bar{w}

$$z_j - c_j = c_B^T B^{-1} a^j - c_j = \bar{w}^T a^j - c_j = [A^T \bar{w}]_j - c_j,$$

we have that $A^T \bar{w} \leq c$ if and only if the tableau is optimal. Moreover, because for the slack variables the columns a^{i+n} are given by the canonical vectors, namely $a^{i+n} = -e_i$, we have

$$z_{n+i} - c_{n+i} = \bar{w}^T a^{n+i} - c_{n+i} = -\bar{w}^T e_i - 0 = -w_i,$$

so that $z_{n+i} - c_{n+i} \leq 0$ if and only if $\bar{w} \geq 0$.

So indeed \bar{w} is feasible for the dual problem if and only if the tableau is optimal. Moreover, \bar{w} is optimal for the dual problem because

$$c^T \bar{x} = c_B^T \bar{x}_B = c_B^T B^{-1} b = \bar{w}^T b = b^T \bar{w}.$$

□

Remark 4. Note that in the previous theorem 7.1, we have assumed that the primal is in canonical form, namely, the constraints are in the form $Ax \geq b$. If the constraints are the other way around, namely $Ax \leq b$, then the components of $w^T := c_B^T B^{-1}$ are given by $w_i = z_{n+i} - c_{n+i}$.

Therefore, when the optimal solution is attained, from the final tableau of the simplex method, we may get both the primal and dual optimal solutions.

7.2 Recovering missing data in simplex tableau

Example 7.2. (Same example as Example 6.13. By using Remark 4, we extract the optimal solution of the dual problem by looking at the 0-row components corresponding to x_4 and x_5 :

$$w^* = (-3, -2)^T.$$

Now, since, thanks to the same theorem, $(-3, -2) = c_B^T B^{-1}$, we calculate

$$c_B^T = (-3, -2) \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix} = (-8, -6).$$

To compute c_N , we use the fact that in the non basic part of the 0-th row, we have the values of $c_B^T B^{-1} N - c_N^T$. In this example, the only component of c_N that we need to find is the second, as the cost of the slack variables is 0. So we have

$$-2 = c_B^T B^{-1} a^2 - c_2 = (-3, -2) \begin{pmatrix} 1/2 \\ -5/4 \end{pmatrix} - c_2$$

from which we obtain $c_2 = 3$.

The data in the simplex tableaux might be lost or corrupted, in which case we need to recover these data.

Example 7.3. The following is the optimal tableau of simplex method for a given minimisation problem. The objective is to minimise $-2x_1 + 3x_2$, and the slack variables are x_3 and x_4 corresponding to the first and second inequalities, respectively. The constraints are of the ' \leq ' type.

| | x_1 | x_2 | x_3 | x_4 | RHS |
|-------|-------|-------|-------|-------|-------|
| f | b | -1 | f | g | -6 |
| x_3 | c | 0 | 1 | $1/5$ | 4 |
| x_1 | d | e | 0 | 2 | a |

Find the values of a, b, c, d, e, f, g in the above tableau.

7.3 Complementary slack condition

Introducing slack variables in both primal and dual we can state the dual problems as:

| Primal | Dual |
|-------------------------|----------------------------|
| Minimise $c^T x$ | Maximise $b^T w$ |
| subject to $Ax + t = b$ | subject to $A^T w + s = c$ |
| $x \geq 0$ | $w \geq 0.$ |

Thanks to the complementary slackness condition, we have that if (\bar{x}^T, \bar{t}^T) and (\bar{w}^T, \bar{s}^T) are optimal, then

$$\bar{x}^T \bar{s} = \bar{w}^T \bar{t} = 0.$$

This means that

1. *Slack variables of the primal problem correspond to the dual variables*
2. *The variables of the primal problem correspond to the dual slack variables.*

7.4 Uniqueness of optimal solutions to a linear programming problem

It has been proved [Gre94] that a primal-dual pair of optimal solutions is unique if and only if it is a strictly complementary pair of basic solutions (see Theorem 4.22 for the definition of a strictly complementary solution). However, in general, it may happen that both the primal and the dual linear programming problems have multiple optimal solutions. For example, consider the following LP problem

$$\begin{aligned} \text{Minimise} \quad & x_1 + x_2 + 2x_3 \\ \text{subject to} \quad & x_1 + x_3 = 1, \\ & x_2 + x_3 = 1, \\ & x_3 \leq 0, \\ & x_1, x_2 \geq 0. \end{aligned}$$

This problem has multiple solution, for instance $(1, 1, 0)$ and $(0, 1, 1)$. In fact, any feasible point is optimal. The dual problem is

$$\begin{aligned} \text{Maximise} \quad & y_1 + y_2 \\ \text{subject to} \quad & y_1 \leq 1, \\ & y_2 \geq 1, \\ & y_1 + y_2 \leq 2. \end{aligned}$$

The dual also has multiple solutions, for example $(1, 1)$ and $(0, 2)$.

7.5 The Dual Simplex Method

Consider the linear program:

$$\min\{cx : Ax = b, x \geq 0\},$$

and suppose that we have a dual feasible point, i.e., the initial table of simplex method is given as follows:

| | x_1 | x_2 | \cdots | x_n | RHS |
|-----------|-------------|-------------|----------|-------------|-------------|
| f | $z_1 - c_1$ | $z_2 - c_2$ | \cdots | $z_n - c_n$ | $c_B^T b$ |
| x_{B_1} | y_{11} | y_{12} | \cdots | y_{1n} | \bar{b}_1 |
| x_{B_2} | y_{21} | y_{22} | \cdots | y_{2n} | \bar{b}_2 |
| \vdots | \vdots | \vdots | \vdots | \vdots | \vdots |
| x_{B_m} | y_{m1} | y_{m2} | \cdots | y_{mn} | \bar{b}_m |

Where

$$z_j - c_j \leq 0, \quad i = 1, \dots, n.$$

We have several cases:

Case 1. If $\bar{b} \geq 0$, then the current point is already optimal to the linear program.

Case 2. If $\bar{b}_r < 0$ for some r , and the r th row vector is nonnegative, i.e., $y_{rj} \geq 0$, for all $j = 1, \dots, n$, then the dual is unbounded (and hence the primal problem is infeasible). To see why this is true, we need to go back to section 5.2.1. There we saw that the primal problem in standard form can be reformulated as in (5.4) :

$$\begin{aligned} \text{Minimise} \quad & f = f_0 - \sum_{j \in R} (z_j - c_j)x_j \\ \text{subject to} \quad & \sum_{j \in R} (y^j)x_j \leq \bar{b} \\ & x_j \geq 0, j \in R, \end{aligned} \tag{7.1}$$

where R is the set of indices belonging to the non basis. The constraints can be written in matrix form as

$$\sum_{j \in R} (y^j)x_j = B^{-1}Nx_N \leq \bar{b},$$

so that we can write the dual problem as follows

$$\begin{aligned}
 & \text{Maximise} && \bar{b}^T w_N \\
 & \text{subject to} && N^T B^{-T} w_N \geq c - z \\
 & && w \geq 0,
 \end{aligned} \tag{7.2}$$

where we denote by w_N the dual variable to x_N . Note that the entries of the matrix $N^T B^{-T}$ are all strictly positive, so w_N is unbounded.

Case 3. If $\bar{b}_r < 0$ for some r , and there exists at least one element $y_{rj} < 0$ in the r th row. In this case, let x_{B_r} leave, and use the following minimum ratio test to determine the variable to enter:

$$\frac{z_k - c_k}{y_{rk}} = \min \left\{ \frac{z_j - c_j}{y_{rj}} : y_{rj} < 0 \right\}.$$

After pivoting at y_{rk} , it is easy to see that

$$(z_j - c_j)' = (z_j - c_j) - \frac{y_{rj}}{y_{rk}}(z_k - c_k) \leq 0,$$

and hence the dual feasibility is kept. Furthermore, we may show that the dual objective is improved, in fact, the new objective value is

$$c_B^T B^{-1} b - (z_k - c_k) \bar{b}_r / y_{rk} \geq c_B^T B^{-1} b,$$

which follows from the fact

$$z_k - c_k \leq 0, \quad \bar{b}_r < 0, \quad y_{rk} < 0.$$

Thus, after a pivoting, the dual simplex method moves from one dual basic feasible point to another one with improved function value. Under non-degeneracy assumption, the algorithm terminates after finite iterations and finds the primal and dual optimal solutions.

Summary of the dual simplex method (minimisation problem)

Initialization Step: Find a basis B of the primal such that $z_j - c_j \leq 0$ for all j .

Main Step:

Step 1. If $\bar{b} = B^{-1}b \geq 0$, stop, the current solution is optimal. Otherwise, select the pivot row r with $\bar{b}_r < 0$, say, $\bar{b}_r = \min\{\bar{b}_i\}$.

Step 2. If $y_{rj} \geq 0$ for all j , stop, the dual is unbounded (and hence the primal is infeasible). Otherwise, select the pivot column k by the following minimum ratio test:

$$\frac{z_k - c_k}{y_{rk}} = \min \left\{ \frac{z_j - c_j}{y_{rj}} : y_{rj} < 0 \right\}.$$

Step 3. Pivot at y_{rk} and return to Step 1.

Example 7.4. Solve the following LP problem:

$$\begin{aligned} &\text{Minimise} \quad 2x_1 + 3x_2 + 4x_3 \\ &s.t. \quad x_1 + 2x_2 + x_3 \geq 3, \\ &\quad \quad 2x_1 - x_2 + 3x_3 \geq 4, \\ &\quad \quad x_1, x_2, x_3 \geq 0. \end{aligned}$$

Multiplying both sides of the inequalities by -1 and adding slack variables x_4 and x_5 , we may form the first table as follows:

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|--------|-------|-------|-------|-------|-------|
| f | -2 | -3 | -4 | 0 | 0 | 0 |
| x_4 | -1 | -2 | -1 | 1 | 0 | -3 |
| x_5 | -2^* | 1 | -3 | 0 | 1 | -4 |

Since $b_2 = -4 < 0$, x_5 leaves the basis. Which one to enter the basis? By minimum ratio test,

$$\min \left\{ \frac{z_1 - c_1}{y_{21}}, \frac{z_3 - c_3}{y_{23}} \right\} = \min \left\{ \frac{-2}{-2}, \frac{-4}{-3} \right\} = \min \{1, 4/3\} = 1$$

which indicates that x_1 enters the basis. So after pivoting operations, we get the next tableau.

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|------------|-------|-------|-------|-------|
| f | 0 | -4 | -1 | 0 | -1 | 4 |
| x_4 | 0 | $-(5/2)^*$ | 1/2 | 1 | -1/2 | -1 |
| x_1 | 1 | -1/2 | 3/2 | 0 | -1/2 | 2 |

Since $\bar{b}_1 = -1 < 0$, x_4 must leave the basis. By the minimum ratio test, x_2 enters the basis. After pivoting, we get the optimal tableau

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | 0 | -9/5 | -8/5 | -1/5 | 28/5 |
| x_2 | 0 | 1 | -1/5 | -2/5 | 1/5 | 2/5 |
| x_1 | 1 | 0 | 7/5 | -1/5 | -2/5 | 11/5 |

Thus the optimal solution is

$$x^* = (11/5, 2/5, 0),$$

and optimal value

$$f^* = 28/5.$$

Chapter 8

Sensitivity Analysis (non-examinable)

What is sensitivity and why do we have to care about it?

- In most practical applications, some of the problem data are uncertain or cannot be known exactly. In this situation, the problem data are estimated. It is important to be able to find the new optimal solution of the problem as other estimates of some of the data become available, without the expensive task of resolving the problem from scratch.
- Also at early stages of problem formulation some factors may be overlooked. It is important to update the current solution in a way that takes care of these factors.
- Furthermore, in many situations the constraints are not very rigid. For example, a constraint may reflect the availability of some resources. This availability can be increased by extra purchase, and the like. It is desirable to examine the effect of relaxing some of the constraints on the value of the optimal objective without having to resolve the problem.

These and other related topics constitute sensitivity analysis.

[BT97, Chapter 5] and [DT97, Chapter 7] have similar material in this chapter that you may find useful.

We still consider the standard linear programming problem:

$$\min\{cx : Ax = b, x \geq 0\}.$$

Suppose that the simplex method produces an optimal basis B .

We shall describe how to make use of the optimality conditions in order to find a new optimal solution, if some of the problem data change, without resolving the problem from scratch. In particular, the following variations in the problem will be considered.

- Change in the cost vector c .
- Change in the right-hand-side vector b .
- Change in the constraint matrix A .
- Addition of a new activity.
- Addition of a new constraint.

8.1 Change in the cost vector

Changes of the cost vector c will change the cost row of the final tableau, that is, the dual feasibility may be affected (lost). Suppose c_k is changed to c'_k . This includes two cases:

[Case I]: x_k is a nonbasic variable.

[Case II]: x_k is a basic variable.

Case I

In this case c_k is changed to c'_k and x_k is nonbasic.

In this case, c_B is not changed, and hence $z_j = c_B B^{-1} a^j$ is not changed.

- For $j \neq k$, the reduced cost coefficient $z_j - c_j$ is not changed, and it satisfies that $z_j - c_j \leq 0$.

- For $j = k$, we have

$$z_k - c'_k = (z_k - c_k) + (c_k - c'_k).$$

Thus, if c_k is increased, i.e., if $c'_k \geq c_k$, then we still have $z_k - c'_k \leq 0$, the current solution remains to be optimal. More general, if $c'_k \geq z_k$, the solution is no change. Otherwise, $z_k - c'_k > 0$, in which case x_k must be introduced into the basis and the simplex method is continued as usual.

Example 8.1. Consider the problem

$$\begin{aligned} \text{Minimise} \quad & -2x_1 + x_2 - x_3 \\ \text{s.t.} \quad & x_1 + x_2 + x_3 \leq 6, \\ & -x_1 + 2x_2 \leq 4, \\ & x_1, x_2, x_3 \geq 0. \end{aligned}$$

The optimal tableau is given by

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | -3 | -1 | -2 | 0 | -12 |
| x_1 | 1 | 1 | 1 | 1 | 0 | 6 |
| x_5 | 0 | 3 | 1 | 1 | 1 | 10 |

Suppose that $c_2 = 1$ is changed to $c'_2 = -3$. Since x_2 is nonbasic, then

$$z_2 - c'_2 = (z_2 - c_2) + (c_2 - c'_2) = -3 + 4 = 1 > 0,$$

and all other $z_j - c_j$ are unaffected. Hence x_2 enters the basis to replace x_5 in the basis.

So we get the following table

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|------------------|-------|-------|-------|-------|
| f | 0 | 1 | -1 | -2 | 0 | -12 |
| x_1 | 1 | 1 | 1 | 1 | 0 | 6 |
| x_5 | 0 | $\mathfrak{3}^*$ | 1 | 1 | 1 | 10 |

Continue the simplex iteration, and find the new optimal solution as follows.

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|--------|--------|--------|---------|
| f | 0 | 0 | $-4/3$ | $-7/3$ | $-1/3$ | $-46/3$ |
| x_1 | 1 | 0 | $2/3$ | $2/3$ | $-1/3$ | $8/3$ |
| x_2 | 0 | 1 | $1/3$ | $1/3$ | $1/3$ | $10/3$ |

So the solution of the changed LP is $(x_1^*, x_2^*) = (8/3, 10/3)$, and the optimal value is $-46/3$.

Case II

In this case x_k is basic, for example $x_k = x_{B_1}$. The change of c to c' changes the value of z to $z' = c_B'^T B^{-1}N$ and hence

$$z'_j - c_j = c_B'^T B^{-1}a^j - c^j = c_B^T B^{-1}a^j - c_j + (0, \dots, 0, c'_{B_1} - c_{B_1}, 0, \dots, 0)y^j = z_j - c_j + (c'_{B_1} - c_{B_1})y_{B_1j}.$$

This means that we can just update the 0-th row by replacing $z_j - c_j$ with $z_j - c_j + (c'_{B_1} - c_{B_1})y_{B_1j}$ for j non basic, while all entries corresponding to the basic variables remain equal to 0. The new objective function is also changed to

$$c_B'^T B^{-1}b = c_B^T B^{-1}b + (c'_{B_1} - c_{B_1})\bar{b}_{B_1}.$$

Example 8.2. In the LP case as in Exercise 8.1, imagine to change the cost c_1 to $c'_1 = -4$, so that $c'_1 - c_1 = -2$. The the optimal tableau is updated to:

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | -5 | -3 | -4 | 0 | -24 |
| x_1 | 1 | 1 | 1 | 1 | 0 | 6 |
| x_5 | 0 | 3 | 1 | 1 | 1 | 10 |

We see that this is no longer an optimal solution as we have a positive number in the 0-th row. We need to run the simplex method to find an optimal solution.

8.2 Change in the right-hand-side

Suppose that the RHS vector b is changed to b' . Then

- $z_j - c_j = c_B^T B^{-1}a^j - c_j$ has no change,

- and the solution $B^{-1}b$ is changed to $B^{-1}b'$. Thus, the primal feasibility might be affected. If one of the components of $B^{-1}b'$ is negative, the simplex method can be continued on the final tableau by using the dual simplex method.

Otherwise, if $B^{-1}b'$ is nonnegative, the current basis remains to be an optimal basis, and the new optimal value would be $c_B B^{-1}b'$.

Example 8.3. Consider the problem

$$\begin{aligned}
 &\text{Minimise} && -2x_1 + x_2 - x_3 \\
 &\text{s.t.} && x_1 + x_2 + x_3 \leq 6, \\
 &&& -x_1 + 2x_2 \leq 4, \\
 &&& x_1, x_2, x_3 \geq 0.
 \end{aligned}$$

The optimal tableau is given by

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | -3 | -1 | -2 | 0 | -12 |
| x_1 | 1 | 1 | 1 | 1 | 0 | 6 |
| x_5 | 0 | 3 | 1 | 1 | 1 | 10 |

- i) Suppose that the right-hand-side vector $b = (6, 4)^T$ is replaced by $b' = (5, 4)^T$. Check if this change will affect the current optimal solution of the problem.
- ii) What if the RHS vector is replaced by $b' = (-3, 4)^T$?

Solutions: the initial identity basis corresponds to x_4 and x_5 . So in the optimal tableau, the inverse of the optimal basis B is given by

$$B^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Suppose that the RHS vector $b = (6, 4)^T$ is replaced by $b' = (5, 4)^T$. We compute

$$B^{-1}b' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ 9 \end{pmatrix} > 0.$$

Thus the current basis is still optimal. The optimal solution become

$$x_B = \begin{pmatrix} x_1 \\ x_5 \end{pmatrix} = B^{-1}b' = \begin{pmatrix} 5 \\ 9 \end{pmatrix}.$$

The optimal objective value is

$$c_B B^{-1}b' = \begin{pmatrix} -2 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ 9 \end{pmatrix} = -10.$$

Now suppose that the RHS vector is changed to $b' = (-3, 4)^T$. We compute

$$B^{-1}b' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -3 \\ 4 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \end{pmatrix}.$$

Since $B^{-1}b'$ has a negative entry, the current basis becomes infeasible. The dual simplex method implies that the dual problem is unbounded. Hence the changed problem is infeasible. For this specific case, we can deduce that the changed problem is infeasible directly since the first constraint is replaced by

$$x_1 + x_2 + x_3 \leq -3,$$

which is impossible for $x_1, x_2, x_3 \geq 0$.

8.3 Adding a new activity

Suppose that a new activity x_{n+1} with cost c_{n+1} and consumption column a^{n+1} is considered for possible production. Without resolving the problem, we can easily determine whether producing x_{n+1} is worthwhile. First calculate $z_{n+1} - c_{n+1}$.

- If $z_{n+1} - c_{n+1} \leq 0$ (minimization problem), then $x_{n+1}^* = 0$ since it is nonbasic variable, and current solution is optimal.
- If $z_{n+1} - c_{n+1} > 0$, then x_{n+1} is introduced into the basis and the simplex method continues to find the new optimal solution.

Example 8.4. Consider the problem

$$\begin{aligned}
 &\text{Minimise} && -2x_1 + x_2 - x_3 \\
 &\text{s.t.} && x_1 + x_2 + x_3 \leq 6, \\
 &&& -x_1 + 2x_2 \leq 4, \\
 &&& x_1, x_2, x_3 \geq 0.
 \end{aligned}$$

The optimal tableau is given by

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | -3 | -1 | -2 | 0 | -12 |
| x_1 | 1 | 1 | 1 | 1 | 0 | 6 |
| x_5 | 0 | 3 | 1 | 1 | 1 | 10 |

Let's find new optimal solution if a new activity $x_6 \geq 0$ with $c_6 = 1$ and $a_6 = (-1, 2)^T$ is introduced.

From the optimal tableau, we find that the the inverse of the optimal basis B is

$$B^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

We compute

$$z_6 - c_6 = c_B^T B^{-1} a^6 - c_6 = \begin{pmatrix} -2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \end{pmatrix} - 1 = 1 > 0.$$

$$y_6 = B^{-1} a^6 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

We incorporate this information into the last optimal simplex tableau:

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | -3 | -1 | -2 | 0 | 1 | -12 |
| x_1 | 1 | 1 | 1 | 1 | 0 | -1 | 6 |
| x_5 | 0 | 3 | 1 | 1 | 1 | 1* | 10 |

We continue using the simplex method: x_6 enters the basis and x_5 leaves. By pivoting around the pivoting entry, which is marked with a star on the above tableau, we obtain the following tableau

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | -6 | -2 | -3 | -1 | 0 | -22 |
| x_1 | 1 | 4 | 2 | 2 | 1 | 0 | 16 |
| x_6 | 0 | 3 | 1 | 1 | 1 | 1 | 10 |

This is an optimal solution. Thus the optimal solution to the changed problem is $(x_1^*, x_2^*, x_3^*, x_6^*) = (16, 0, 0, 10)$, the optimal value is -22 .

8.4 Adding a new constraint

Suppose that a new constraint is added to the problem.

- If the optimal solution to the original problem satisfies the added constraint, it is then obvious that the point is also an optimal solution to the new problem.
- If the optimal solution to the original problem does not satisfy the new constraint, i.e., if the constraint cut away the optimal point, we can use the dual simplex method to find the new optimal solution.

Example 8.5. Consider the problem

$$\begin{aligned}
 &\text{Minimise} && -2x_1 + x_2 - x_3 \\
 &s.t. && x_1 + x_2 + x_3 \leq 6, \\
 &&& -x_1 + 2x_2 \leq 4, \\
 &&& x_1, x_2, x_3 \geq 0.
 \end{aligned}$$

The optimal tableau is given by

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | -3 | -1 | -2 | 0 | -12 |
| x_1 | 1 | 1 | 1 | 1 | 0 | 6 |
| x_5 | 0 | 3 | 1 | 1 | 1 | 10 |

Consider the problem with the added constraint

$$-x_1 + 2x_3 \geq 2.$$

Clearly the optimal point $(x_1, x_2, x_3) = (6, 0, 0)$ does not satisfy this constraint. The constraint $-x_1 + 2x_2 \geq 2$ is rewritten as

$$x_1 - 2x_3 + x_6 = -2,$$

where x_6 is a nonnegative slack variable. This row is added to the optimal tableau of the above problem to obtain the following tableau.

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | -3 | -1 | -2 | 0 | 0 | -12 |
| x_1 | 1 | 1 | 1 | 1 | 0 | 0 | 6 |
| x_5 | 0 | 3 | 1 | 1 | 1 | 0 | 10 |
| x_6 | 1 | 0 | -2 | 0 | 0 | 1 | -2 |

Multiply row 1 by -1 and add to row 3 in order to restore column x_1 to a unit vector.

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|-------|--------|-------|-------|-------|-------|
| f | 0 | -3 | -1 | -2 | 0 | 0 | -12 |
| x_1 | 1 | 1 | 1 | 1 | 0 | 0 | 6 |
| x_5 | 0 | 3 | 1 | 1 | 1 | 0 | 10 |
| x_6 | 0 | -1 | -3^* | -1 | 0 | 1 | -8 |

The dual simplex method can then be applied to the above tableau. By minimum ratio test of the dual simplex method, we see that x_3 enters the basis to replace x_6 . Thus, we get the following tableau.

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | RHS |
|-------|-------|--------|-------|--------|-------|--------|---------|
| f | 0 | $-8/3$ | 0 | $-5/3$ | 0 | $-1/3$ | $-28/3$ |
| x_1 | 1 | $2/3$ | 0 | $2/3$ | 0 | $1/3$ | $10/3$ |
| x_5 | 0 | $8/3$ | 0 | $2/3$ | 1 | $1/3$ | $22/3$ |
| x_3 | 0 | $1/3$ | 1 | $1/3$ | 0 | $-1/3$ | $8/3$ |

So the optimal solution of the changed problem is

$$(x_1^*, x_2^*, x_3^*) = (10/3, 0, 8/3).$$

In summary, when a constraint is added, restore the simplex tableau first, and then continue simplex iterations (by using dual simplex method) until new solution is found.

The above example introduces the idea of how to handle the case when the optimal solution to the original problem does not satisfy the new constraint. Suppose that B is the optimal basis before the constraint

$$a^{m+1}x \leq b_{m+1}$$

is added. As seen in section 6.2, the corresponding system is given by the following system

$$f + (c_B B^{-1} N - c_N)x_N = c_B B^{-1} b,$$

$$x_B + B^{-1} N x_N = B^{-1} b.$$

Notice that the new constraint can be written as

$$a_B^{m+1} x_B + a_N^{m+1} x_N + x_{n+1} = b_{m+1}.$$

Here a^{m+1} is decomposed into a_B^{m+1} , a_N^{m+1} and x_{n+1} is a nonnegative slack variable. Multiplying the above second equation by a_B^{m+1} and subtracting from the new constraint gives the following system:

$$\begin{cases} f + (c_B B^{-1} N - c_N)x_N = c_B B^{-1} b, \\ x_B + B^{-1} N x_N = B^{-1} b, \\ (a_N^{m+1} - a_B^{m+1} B^{-1} N)x_N + x_{n+1} = b_{m+1} - a_B^{m+1} B^{-1} b. \end{cases}$$

The basic feasible solution to this system is

$$x_B = B^{-1} b$$

$$x_N = 0$$

$$x_{n+1} = b_{m+1} - a_B^{m+1} B^{-1} b.$$

The only possible violation of optimality of the new problem is the sign of $b_{m+1} - a_B^{m+1} B^{-1} b$.

If

$$b_{m+1} - a_B^{m+1} B^{-1} b \geq 0,$$

then the current solution is optimal. Otherwise, if

$$b_{m+1} - a_B^{m+1} B^{-1} b < 0,$$

then the dual simplex method is used to restore feasibility.

Chapter 9

The Two-Phase method and the Big-M Method (non-examinable)

The simplex method assumes that an initial basic feasible solution (BFS) is at hand. In many cases, such a BFS is not readily available. There are some work may be needed to get the simplex method started.

Example 9.1. *Consider the following two groups of constraints*

1.

$$x_1 + 2x_2 \leq 4,$$

$$-x_1 + x_2 \leq 1,$$

$$x_1, x_2 \geq 0.$$

2.

$$x_1 + x_2 + x_3 \leq 6,$$

$$-2x_1 + 3x_2 + 2x_3 \geq 3,$$

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

For the first group of constraints, an initial basis can be found without any difficulty. However, the second one is not. Some extra work is needed to get the initial basis.

In this chapter we will study two procedures that can be used to construct initial basis:

- The two-phase method.
- The big-M method.

Both involve artificial variables to obtain an initial basic feasible solution.

At the end of this chapter, students should be able to apply the two-phase method and the big-M method to solve some LPs.

9.1 The Two-phase Method

Suppose we want to solve the following LP:

$$\begin{aligned} &\text{Minimise} && c^T x \\ &\text{subject to} && Ax \leq b \\ &&& x \geq 0. \end{aligned}$$

with $b \geq 0$. The most convenient way to find a BFS is to introduce some slack variables s so that the problem is re-formulated as

$$\begin{aligned} &\text{Minimise} && c^T x \\ &\text{subject to} && Ax + s = b \\ &&& x, s \geq 0. \end{aligned}$$

Now we pick $x = 0$ and $s = b$ as BFS. This is feasible because $b \geq 0$. Since the basic variables coincide with the slack variables, $z = 0$ and $f = 0$. So our initial tableau is very simple:

| | x | s | RHS |
|-----|--------|-----|-----|
| f | $-c^T$ | 0 | 0 |
| s | A | 1 | b |

and we can start our simplex method.

However, in general, we may not be that lucky. In particular we may have that we don't have enough slack variables to construct a BFS, or that the non all component of b are non negative and/or A does not have an identity sub-matrix.

Example 9.2. Consider the following constraints

$$\begin{aligned}x_1 + x_2 + x_3 &\leq 6, \\ -2x_1 + 3x_2 + 2x_3 &\geq 3, \\ x_2, x_3 &\geq 0.\end{aligned}$$

Note that x_1 is unrestricted, therefore we introduce $x'_1, x''_1 \geq 0$, and put $x_1 = x'_1 - x''_1$. We also introduce slack variables x_4, x_5 :

$$\begin{aligned}x'_1 - x''_1 + x_2 + x_3 + x_4 &= 6, \\ -2x'_1 + 2x''_1 + 3x_2 + 2x_3 - x_5 &= 3, \\ x'_1, x''_1, x_2, x_3, x_4, x_5 &\geq 0.\end{aligned}$$

We see that the new constraint matrix does not contain an identity 2×2 sub-matrix and therefore an obvious feasible basis cannot be extracted.

Let us start with an LP in the following form:

$$\begin{aligned}\text{Minimise} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0,\end{aligned}$$

where $b \geq 0$. Indeed, given any LP in standard form, if some component b_r of b is negative, we can just multiply the r -th row by -1 to map it to this form.

If A has a $m \times m$ identity sub-matrix, then we know a quick way to select out BFS. Otherwise, we may face some difficulty. In this case, we apply the 2-phase method. This consists in adding further *artificial variables* x_a to the constraints in order to force the appearance of an identity sub-matrix. Specifically we modify the constraints as follows,

$$Ax + x_a = b, \quad x, x_a \geq 0,$$

so that the new constraint matrix is $[A, 1]$ and we have an immediate BFS in the form $x = 0, x_a = b$.

Note the difference between slack variables and artificial variables: slack variables are introduced to transform problems in non-standard form to problems in standard form. The artificial variables are introduced after having reduced the problem in standard form and are not legitimate variables.

The two-phase method consist in a Phase I problem which is used to find a basic feasible solution for the original constraints, then we can initiate Phase II wherein the Simplex Method is applied to solve the original linear programming problem.

Phase 1. Solve the following linear programming problem starting with the basic feasible solution $x = 0$ and $x_a = b$:

$$\begin{aligned} &\text{Minimise} && e^T x_a \\ &\text{subject to} && Ax + x_a = b, \\ &&& x, x_a \geq 0, \end{aligned}$$

where e denotes the vector with all components equal to 1.

- If at optimality $x_a \neq 0$, then stop. The original problem has no feasible solution because if it did have a feasible solution x^* then $(x^*, 0)$ would be an optimal solution of the Phase I problem.
- Otherwise, this means we have reached optimality in the form $(\bar{x}, 0)$. By construction, \bar{x} is feasible for the original problem. Let's split it in basic and nonbasic variables be x_B and x_N . This step is simple: put all non zero variables in the basic and then add enough 0 to reach a total of m variables.

Phase II. Having chosen x_B , we now have B and we can solve the following linear programming problem starting with the basic feasible solution $x_B = B^{-1}b$ and $x_N = 0$:

$$\begin{aligned} &\text{Minimise} && c_B x_B + c_N x_N \\ &\text{subject to} && x_B + B^{-1} N x_N = B^{-1} b, \\ &&& x_B, x_N \geq 0. \end{aligned}$$

The purpose of the phase I is to get us to an extreme point of the feasible region, while phase II takes us from this feasible point to an optimal point.

Remark: For the system $Ax = b$ and $x \geq 0$, we change the restrictions by adding an artificial vector x_a leading to the system

$$Ax + x_a = b, \quad x \geq 0, \quad x_a \geq 0.$$

This gives an immediate basic feasible solution of the new system, namely $x_a = b$ and $x = 0$. Even though we now have a starting basic feasible solution and the simplex method can be applied, we have in fact changed the problem. In order to get back to our original problem, we must force these artificial variables to zero, because $Ax = b$ if and only if $Ax + x_a = b$ with $x_a = 0$. In other words, artificial variables are only a tool to get the simplex method started, however, we must guarantee that these variables will eventually drop to zero.

Example 9.3. *Solve the following LP by using two-phase method*

$$\begin{aligned} \text{Minimise} \quad & x_1 - 2x_2 \\ \text{subject to} \quad & x_1 + x_2 \geq 2, \\ & -x_1 + x_2 \geq 1, \\ & x_2 \leq 3, \\ & x_1, x_2 \geq 0. \end{aligned}$$

By introducing the slack variables x_3, x_4 and x_5 , the problem becomes

$$\begin{aligned} \text{Minimise} \quad & x_1 - 2x_2 \\ \text{subject to} \quad & x_1 + x_2 - x_3 = 2, \\ & -x_1 + x_2 - x_4 = 1, \\ & x_2 + x_5 = 3, \\ & x_1, x_2, x_3, x_4, x_5 \geq 0. \end{aligned}$$

Phase I

Adding artificial variables x_6 and x_7 and consider the Phase-I problem:

$$\begin{aligned}
 &\text{Minimise} \quad x_6 + x_7 \\
 &\text{subject to} \quad x_1 + x_2 - x_3 \quad \quad + x_6 = 2 \\
 &\quad \quad \quad -x_1 + x_2 \quad \quad - x_4 \quad \quad + x_7 = 1 \\
 &\quad \quad \quad \quad \quad x_2 \quad \quad + x_5 \quad \quad = 3 \\
 &\quad \quad \quad x_1, x_2, x_3, x_4, x_5 \geq 0.
 \end{aligned}$$

Let $f_0 = x_6 + x_7$. The Phase I minimises the objective f_0 . Since x_5, x_6, x_7 are basic variables, $c_B^T = (0, 1, 1)$ and we have

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-----------|-------|-------|-------|-------|-------|-----|
| f_0 | 0 | 2 | -1 | -1 | 0 | 0 | 0 | 3 |
| x_6 | 1 | 1 | -1 | 0 | 0 | 1 | 0 | 2 |
| x_7 | -1 | 1* | 0 | -1 | 0 | 0 | 1 | 1 |
| x_5 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 3 |

Since $z_2 - c_2 = 2 > 0$, x_7 is replaced by x_2 , after pivoting operations, we have the following table

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-----------|-------|-------|-------|-------|-------|-------|-----|
| f_0 | 2 | 0 | -1 | 1 | 0 | 0 | -2 | 1 |
| x_6 | 2* | 0 | -1 | 1 | 0 | 1 | -1 | 1 |
| x_2 | -1 | 1 | 0 | -1 | 0 | 0 | 1 | 1 |
| x_5 | 1 | 0 | 0 | 1 | 1 | 0 | -1 | 2 |

Replacing x_6 by x_1 , we have

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|-----|
| f_0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | 0 |
| x_1 | 1 | 0 | -1/2 | 1/2 | 0 | 1/2 | -1/2 | 1/2 |
| x_2 | 0 | 1 | -1/2 | -1/2 | 0 | 1/2 | 1/2 | 3/2 |
| x_5 | 0 | 0 | 1/2 | 1/2 | 1 | -1/2 | -1/2 | 3/2 |

which is the optimal tableau of the Phase I. There is no artificial variable in the optimal basis of the first phase. We have a starting basic feasible solution for proceeding to the second phase.

Phase II

Removing the artificial variables from the table, we do not consider it any more. Replace the first row by the coefficients of the equation $f - x_1 + 2x_2 = 0$ since the original objective is $f = x_1 - 2x_2$.

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|--------|--------|-------|-------|
| f | -1 | 2 | 0 | 0 | 0 | 0 |
| x_1 | 1 | 0 | $-1/2$ | $1/2$ | 0 | $1/2$ |
| x_2 | 0 | 1 | $-1/2$ | $-1/2$ | 0 | $3/2$ |
| x_5 | 0 | 0 | $1/2$ | $1/2$ | 1 | $3/2$ |

Restoring the simplex tableau gives rise to

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|--------|-----------|-------|--------|
| f | 0 | 0 | $1/2$ | $3/2$ | 0 | $-5/2$ |
| x_1 | 1 | 0 | $-1/2$ | $(1/2)^*$ | 0 | $1/2$ |
| x_2 | 0 | 1 | $-1/2$ | $-1/2$ | 0 | $3/2$ |
| x_5 | 0 | 0 | $1/2$ | $1/2$ | 1 | $3/2$ |

Since $z_4 - c_4 = 3/2 > 0$, x_4 is the entering variable. By minimum ratio test, x_1 is the leaving variable.

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|-------|-------|-------|-------|
| f | -3 | 0 | 2 | 0 | 0 | -4 |
| x_4 | 2 | 0 | -1 | 1 | 0 | 1 |
| x_2 | 1 | 1 | -1 | 0 | 0 | 2 |
| x_5 | -1 | 0 | 1^* | 0 | 1 | 1 |

It is evident that x_3 is the entering variable and x_5 is the leaving variable.

| | f | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-----|-------|-------|-------|-------|-------|-------|
| f | 1 | -1 | 0 | 0 | 0 | -2 | -6 |
| x_4 | 0 | 1 | 0 | 0 | 1 | 1 | 2 |
| x_2 | 0 | 0 | 1 | 0 | 0 | 1 | 3 |
| x_3 | 0 | -1 | 0 | 1 | 0 | 1 | 1 |

The optimal solution to the original LP problem is given by

$$x^* = (x_1^*, x_2^*) = (0, 3).$$

Note that phase I moves from the infeasible point $(0,0)$ to the infeasible point $(0,1)$, and finally to the feasible point $(1/2, 3/2)$. From this extreme point, phase II moves to the feasible point $(0,2)$ and finally to the optimal point $(0, 3)$.

Analysis of the Two-Phase Method:

At the end of phase I, there are possible cases:

Case A: $x_a \neq 0$. In this case, the original problem is infeasible.

Proof. In fact, if there is an $x \geq 0$ with $Ax = b$, then $(x, 0)$ is a feasible solution of the phase I problem and objective value at this feasible solution is 0, violating the optimality of x_a at which the objective is positive.

Case B: $x_a = 0$. There are two subcases:

- If all artificial variables are out of the basis, remove from the last form the artificial and construct the initial simplex step since there is a basic feasible solution at hand, and hence proceed to Phase II.
- If some artificial variables are in the basis at zero level, we have two ways to continue. We may first eliminate the columns corresponding to the nonbasic artificial variables of Phase I, and then proceed to phase II directly to continue to replace those left artificial variables without increasing the original objective values.

Example 9.4. Solve the following LP problem:

$$\begin{aligned} &\text{Minimise} && 2x_1 + 3x_2 + 4x_3 \\ &s.t. && x_1 + 2x_2 + x_3 \geq 3, \\ &&& 2x_1 - x_2 + 3x_3 \geq 4, \\ &&& x_1, x_2, x_3 \geq 0. \end{aligned}$$

First, we introduce slack variables x_4 and x_5 , to write the problem in the standard form

$$\begin{aligned}
 &\text{Minimise} \quad 2x_1 + 3x_2 + 4x_3 \\
 &\text{s.t.} \quad x_1 + 2x_2 + x_3 - x_4 = 3, \\
 &\quad \quad 2x_1 - x_2 + 3x_3 - x_5 = 4, \\
 &\quad \quad x_1, \dots, x_5 \geq 0.
 \end{aligned}$$

Phase I: now we introduce artificial variables x_6 and x_7 and consider the following phase-I problem

$$\begin{aligned}
 &\text{Minimise} \quad x_6 + x_7 \\
 &\text{s.t.} \quad x_1 + 2x_2 + x_3 - x_4 + x_6 = 3, \\
 &\quad \quad 2x_1 - x_2 + 3x_3 - x_5 + x_7 = 4, \\
 &\quad \quad x_1, \dots, x_7 \geq 0.
 \end{aligned}$$

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| f_0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | 0 |
| x_6 | 1 | 2 | 1 | -1 | 0 | 1 | 0 | 3 |
| x_7 | 2 | 1 | 3 | 0 | -1 | 0 | 1 | 4 |

Restore the initial simplex method.

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| f_0 | 3 | 1 | 4 | -1 | -1 | 0 | 0 | 7 |
| x_6 | 1 | 2 | 1 | -1 | 0 | 1 | 0 | 3 |
| x_7 | 2 | 1 | 3* | 0 | -1 | 0 | 1 | 4 |

x_3 enters, x_7 leaves

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| f_0 | 1/3 | 7/3 | 0 | -1 | 1/3 | 0 | -4/3 | 5/3 |
| x_6 | 1/3 | 7/3 | 0 | -1 | 1/3 | 1 | -1/3 | 5/3 |
| x_3 | 2/3 | -1/3 | 1 | 0 | -1/3 | 0 | 1/3 | 4/3 |

x_2 enters, x_6 leaves.

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|--------|--------|-------|--------|--------|
| f_0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | 0 |
| x_2 | $1/7$ | 1 | 0 | $-3/7$ | $1/7$ | $3/7$ | $-1/7$ | $5/7$ |
| x_3 | $5/7$ | 0 | 1 | $-1/7$ | $-2/7$ | $1/7$ | $3/7$ | $11/7$ |

This is an optimal solution to the phase I problem. The artificial variables x_6 and x_7 are not basic variable at optimality. Thus we remove columns x_6 and x_7 and move to phase II.

Phase II: we replace the original cost function to obtain the tableau:

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|-------|--------|--------|--------|
| f | -2 | -3 | -4 | 0 | 0 | 0 |
| x_2 | $1/7$ | 1 | 0 | $-3/7$ | $1/7$ | $5/7$ |
| x_3 | $5/7$ | 0 | 1 | $-1/7$ | $-2/7$ | $11/7$ |

We restore the simplex tableau:

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|---------|-------|-------|---------|--------|--------|
| f | $9/7$ | 0 | 0 | $-13/7$ | $-5/7$ | $59/7$ |
| x_2 | $1/7$ | 1 | 0 | $-3/7$ | $1/7$ | $5/7$ |
| x_3 | $5/7^*$ | 0 | 1 | $-1/7$ | $-2/7$ | $11/7$ |

x_1 enters, x_3 leaves.

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|--------|--------|--------|--------|
| f | 0 | 0 | $-9/5$ | $-8/5$ | $-1/5$ | $28/5$ |
| x_2 | 0 | 1 | $-1/5$ | $-2/5$ | $1/5$ | $2/5$ |
| x_1 | 1 | 0 | $7/5$ | $-1/5$ | $-2/5$ | $11/5$ |

Thus the optimal solution is

$$x^* = (11/5, 2/5, 0),$$

and optimal value

$$f^* = 28/5.$$

Example 9.5. Solve the following LP problem:

$$\begin{aligned}
 &\text{Minimise} && -x_1 + 2x_2 - 3x_3 \\
 &s.t. && x_1 + x_2 + x_3 = 6, \\
 &&& -x_1 + x_2 + 2x_3 = 4, \\
 &&& 2x_2 + 3x_3 = 10, \\
 &&& x_3 \leq 2, \\
 &&& x_1, x_2, x_3 \geq 0
 \end{aligned}$$

First, we introduce the slack variable x_4 to write the problem in the standard form

$$\begin{aligned}
 &\text{Minimise} && -x_1 + 2x_2 - 3x_3 \\
 &s.t. && x_1 + x_2 + x_3 = 6, \\
 &&& -x_1 + x_2 + 2x_3 = 4, \\
 &&& 2x_2 + 3x_3 = 10, \\
 &&& x_3 + x_4 = 2, \\
 &&& x_1, x_2, x_3, x_4 \geq 0.
 \end{aligned}$$

Phase I:

$$\begin{aligned}
 &\text{Minimise} && x_5 + x_6 + x_7 \\
 &s.t. && x_1 + x_2 + x_3 + x_5 = 6, \\
 &&& -x_1 + x_2 + 2x_3 + x_6 = 4, \\
 &&& 2x_2 + 3x_3 + x_7 = 10, \\
 &&& x_3 + x_4 = 2, \\
 &&& x_1, x_2, x_3, x_4, x_5, x_6, x_7 \geq 0.
 \end{aligned}$$

This corresponds to the tableau (because $c_B^T = (0, 1, 1, 1)$)

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | 4 | 6 | 0 | 0 | 0 | 0 | 20 |
| x_5 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 6 |
| x_6 | -1 | 1 | 2 | 0 | 0 | 1 | 0 | 4 |
| x_7 | 0 | 2 | 3 | 0 | 0 | 0 | 1 | 10 |
| x_4 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 2 |

Pivot at y_{43} :

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | 4 | 0 | -6 | 0 | 0 | 0 | 8 |
| x_5 | 1 | 1 | 0 | -1 | 1 | 0 | 0 | 4 |
| x_6 | -1 | 1 | 0 | -2 | 0 | 1 | 0 | 0 |
| x_7 | 0 | 2 | 0 | -3 | 0 | 0 | 1 | 4 |
| x_3 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 2 |

Pivot at y_{62} :

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| f | 4 | 0 | 0 | 2 | 0 | -4 | 0 | 8 |
| x_5 | 2 | 0 | 0 | 1 | 1 | -1 | 0 | 4 |
| x_2 | -1 | 1 | 0 | -2 | 0 | 1 | 0 | 0 |
| x_7 | 2 | 0 | 0 | 1 | 0 | -2 | 1 | 4 |
| x_3 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 2 |

Pivot at y_{51} :

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | RHS |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| f | 0 | 0 | 0 | 0 | -2 | -2 | 0 | 0 |
| x_1 | 1 | 0 | 0 | 1/2 | 1/2 | -1/2 | 0 | 2 |
| x_2 | 0 | 1 | 0 | -3/2 | 1/2 | 1/2 | 0 | 2 |
| x_7 | 0 | 0 | 0 | 0 | -1 | -1 | 1 | 0 |
| x_3 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 2 |

We have reached optimality. However, we see that we still have an artificial variable in the basis. This is not a problem because its value is 0, therefore we can eliminate the corresponding row all together. The Phase II tableau is then obtained by inserting the original cost function at row 0:

| | x_1 | x_2 | x_3 | x_4 | RHS |
|-------|-------|-------|-------|-------|-------|
| f | 1 | -2 | 3 | 0 | 0 |
| x_1 | 1 | 0 | 0 | 1/2 | 2 |
| x_2 | 0 | 1 | 0 | -3/2 | 2 |
| x_3 | 0 | 0 | 1 | 1 | 2 |

9.2 The Big-M method

The big-M approach is also based on introducing artificial variables, but to get rid of the artificial variables by assigning coefficients for those artificial variables in the original objective function in such way as to make their presence in the basis at a positive level very unattractive from the objective function point of view. Consider the following problem:

$$\begin{aligned} \text{Minimize} \quad & c^T x + M(e^T x_a) \\ \text{subject to} \quad & Ax + x_a = b, \\ & x, x_a \geq 0. \end{aligned}$$

where M is a very large positive number.

The term $Me^T x_a$ can be interpreted as a penalty to be paid by any solution with $x_a \neq 0$.

Even though the starting solution $x = 0, x_a = b$ is feasible to the above problem, it has a very unattractive objective value.

Therefore, the simplex method will try to get the artificial variables out the basis, and then continue to find an optimal solution of the original problem.

Analysis of the Big-M method.

Two cases may arise after solving the Big-M problem by the simplex method:

- Arrive at an optimal solution of Big-M problem.
- The Big-M problem has an unbounded optimal solution.

Relationship between the Big-M problem and the Original LP problem

Case A: The Big-M problem has an finite optimal solution. There are two subcases:

- $(x, x_a) = (x^*, 0)$ is an optimal solution of the Big-M problem. Then x^* is the optimal solution to the original problem.

Proof. In fact, if x is an arbitrary feasible point of the original LP, then $(x, 0)$ is a feasible solution to the Big-M problem. Since $(x^*, 0)$ is the optimal solution to the Big-M problem, we have

$$c^T x^* + 0 \leq c^T x + 0,$$

i.e.,

$$c^T x^* \leq c^T x$$

which implies that x^* is indeed an optimal solution of the original LP problem.

- (x^*, x_a^*) is an optimal solution to the Big-M problem and $x_a^* \neq 0$.

In this case, the original LP has no feasible point, i.e., the problem is infeasible.

Proof. Suppose by contradiction that x is a feasible point of the original LP. Then $(x^T, 0)$ is feasible for the Big-M problem. By optimality of (x^*, x_a^*) , we have

$$c^T x^* + M x_a^* \leq c^T x$$

which is impossible because M is very large and $x_a^* > 0$.

Case B: The Big-M problem has an unbounded optimal value. Assume that $z_k - c_k = \max(z_j - c_j) > 0$ and $y_k \leq 0$, then we have the following two cases:

- all the artificial are equal to zero, then the original LP has an unbounded optimal solution.

Proof. In this case, the optimal solution of the Big-M problem has the form $(x^{*T}, 0)$, therefore x^* is feasible for the original LP, but because $(x^{*T}, 0)$ is unbounded, then x^* is unbounded. This solution is also optimal because if it wasn't there would be another optimal solution \bar{x} such of the LP and $(\bar{x}, 0)$ would be a feasible solution of the Big-M problem with better objective value, which is not allowed.

- not all the artificial are equal to zero, then the original LP is infeasible.

Proof. Omitted.

In summary:

| Solve the Big-M problem | | | |
|--|--|--|--|
| A. Optimal is finite | | B. Optimal is unbounded | |
| A1. $\bar{x}_a = 0$ \bar{x} optimal OLP | A2. $\bar{x}_a \neq 0$ infeasible OLP | B1. $\bar{x}_a = 0$ \bar{x} unbounded OLP | B2. $\bar{x}_a \neq 0$ infeasible OLP |

Example 9.6. *Solve the LP by Big-M method*

$$\begin{aligned}
 &\text{Minimise} && x_1 - 2x_2 \\
 &\text{subject to} && x_1 + x_2 \geq 2, \\
 &&& x_2 \leq 3, \\
 &&& x_1, x_2 \geq 0.
 \end{aligned}$$

Adding slack variables x_3 and x_4 and artificial variables x_5 , the Big-M LP can be written as

$$\begin{aligned}
 &\text{Minimise} && x_1 - 2x_2 + Mx_5 \\
 &\text{subject to} && x_1 + x_2 - x_3 + x_5 = 2, \\
 &&& x_2 + x_4 = 3, \\
 &&& x_1, x_2, x_3, x_4, x_5 \geq 0.
 \end{aligned}$$

The initial basic variables are x_5 and x_4 ,

$$\begin{aligned}
 B &= [a_5 \ a_4] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad N = [a_1 \ a_2 \ a_3] = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \\
 c_B &= (c_5 \ c_4) = (M \ 0), \quad c_N = (c_1 \ c_2 \ c_3) = (1 \ -2 \ 0).
 \end{aligned}$$

Note that in the simplex method, we work with tableau (including the initial one) of the form

| | x_B | x_N | RHS |
|-------|-------|----------------------|----------------|
| f | 0 | $c_B B^{-1} N - c_N$ | $c_B B^{-1} b$ |
| x_B | I | $B^{-1} N$ | $B^{-1} b$ |

where the reduce cost $z_j - c_j$ are zero for all basic variables.

Let us find the initial tableau:

$$c_B B^{-1} N - c_N = \begin{pmatrix} M & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & -2 & 0 \end{pmatrix} = \begin{pmatrix} M-1 & M+2 & -M \end{pmatrix},$$

$$c_B B^{-1} b = \begin{pmatrix} M & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2M, \quad B^{-1} N = N = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad B^{-1} b = b = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Thus we obtain the following initial tableau for the simplex method

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|----------------|-------|-------|-------|-------|
| f | $M-1$ | $M+2$ | $-M$ | 0 | 0 | $2M$ |
| x_5 | 1 | $\mathbf{1}^*$ | -1 | 0 | 1 | 2 |
| x_4 | 0 | 1 | 0 | 1 | 0 | 3 |

Now we apply the simplex method: x_2 enters and x_5 leaves:

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|----------------|-------|----------|-------|
| f | -3 | 0 | 2 | 0 | $-(M+2)$ | -4 |
| x_2 | 1 | 1 | -1 | 0 | 1 | 2 |
| x_4 | -1 | 0 | $\mathbf{1}^*$ | 1 | -1 | 1 |

x_3 enters and x_4 leaves:

| | x_1 | x_2 | x_3 | x_4 | x_5 | RHS |
|-------|-------|-------|-------|-------|-------|-------|
| f | -1 | 0 | 0 | -2 | $-M$ | -6 |
| x_2 | 0 | 1 | 0 | 1 | 0 | 3 |
| x_3 | -1 | 0 | 1 | 1 | -1 | 1 |

So the optimal solution of the original problem is

$$x^* = (x_1^*, x_2^*) = (0, 3)$$

and the optimal value is

$$f^* = -6.$$

Example 9.7. *Solve the LP by using Big-M method*

$$\begin{aligned}
 &\text{Minimise} && -x_1 - 3x_2 + x_3 \\
 &\text{subject to} && x_1 + x_2 + 2x_3 \leq 4, \\
 &&& -x_1 + x_3 \geq 4, \\
 &&& x_3 \geq 3, \\
 &&& x_1, x_2, x_3 \geq 0.
 \end{aligned}$$

Example 9.8. *Solve the LP by using Big-M method*

$$\begin{aligned}
 &\text{Minimise} && -x_1 - x_2 \\
 &\text{subject to} && x_1 - x_2 - x_3 = 1, \\
 &&& -x_1 + x_2 + 2x_3 - x_4 = 1, \\
 &&& x_1, x_2, x_3, x_4 \geq 0.
 \end{aligned}$$

Bibliography

- [Bea55] E. M. L. Beale. Cycling in the dual simplex algorithm. *Naval Research Logistics Quarterly*, pages 269–275, 1955.
- [BT97] D. Bertsimas and J.N. Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific series in optimization and neural computation. Athena Scientific, 1997.
- [DT97] George B. Dantzig and Mukund N. Thapa. *Linear programming. 1*. Springer Series in Operations Research. Springer-Verlag, New York, 1997. Introduction, With 1 CD-ROM for Windows.
- [Gre94] Harvey J. Greenberg. The use of the optimal partition in a linear programming solution for postoptimal analysis. *Operations Research Letters*, 15(4):179 – 185, 1994.
- [KM72] V. Klee and G.J. Minty. How good is the simplex algorithm? *In O. Shisha, editor, Inequalities, III*, pages 159–175, 1972.