Module MA3484: Methods of Mathematical Economics
Hilary Term 2015

D. R. Wilkins

Copyright © David R. Wilkins 2015

Contents

1 Introduction to Linear Programming

1.1 A Furniture Retailing Problem

A retail business is planning to devote a number of retail outlets to the sale of armchairs and sofas.

The retail prices of armchairs and sofas are determined by fierce competition in the furniture retailing business. Armchairs sell for €700 and sofas sell for €1000.

However

- the amount of floor space (and warehouse space) available for stocking the sofas and armchairs is limited;
- the amount of capital available for purchasing the initial stock of sofas and armchairs is limited;
- market research shows that the ratio of armchairs to sofas in stores should neither be too low nor too high.

Specifically:

- there are 1000 square metres of floor space available for stocking the initial purchase of sofas and armchairs;
- each armchair takes up 1 square metre;
- each sofa takes up 2 square metres;
- the amount of capital available for purchasing the initial stock of armchairs and sofas is €351,000;
- the wholesale price of an armchair is €400;
- the wholesale price of a sofa is €600;
- market research shows that between 4 and 9 armchairs should be in stock for each 3 sofas in stock.

We suppose that the retail outlets are stocked with x armchairs and y sofas.

The armchairs (taking up 1 sq. metre each) and the sofas (taking up 2 sq. metres each) cannot altogether take up more than 1000 sq. metres of floor space. Therefore

$$x + 2y \leq 1000 \quad \text{(Floor space constraint).}$$

The cost of stocking the retail outlets with armchairs (costing €400 each) and sofas (costing €600 each) cannot exceed the available capital of €351000. Therefore

$$4x + 6y \leq 3510 \quad \text{(Capital constraint)}.$$

Consumer research indicates that x and y should satisfy

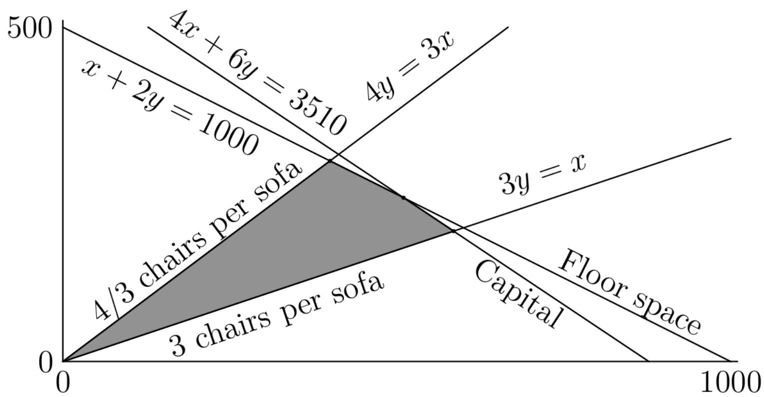
$$4y \leq 3x \leq 9y \quad \text{(Armchair/Sofa ratio)}.$$

An ordered pair (x, y) of real numbers is said to specify a *feasible solution* to the linear programming problem if this pair of values meets all the relevant constraints.

An ordered pair (x, y) constitutes a feasible solution to the the Furniture Retailing problem if and only if all the following constraints are satisfied:

$$\begin{aligned} x - 3y &\leq 0; \\ 4y - 3x &\leq 0; \\ x + 2y &\leq 1000; \\ 4x + 6y &\leq 3510; \\ x &\geq 0; \\ y &\geq 0; \end{aligned}$$

The feasible region for the Furniture Retailing problem is depicted below:



We identify the *vertices* (or *corners*) of the feasible region for the Furniture Retailing problem. There are four of these:

- there is a vertex at $(0, 0)$;
- there is a vertex at $(400, 300)$ where the line $4y = 3x$ intersects the line $x + 2y = 1000$;

- there is a vertex at $(510, 245)$ where the line $x + 2y = 1000$ intersects the line $4x + 6y = 3510$;
- there is a vertex at $(585, 195)$ where the line $3y = x$ intersects the line $4x + 6y = 3510$.

These vertices are identified by inspection of the graph that depicts the constraints that determine the feasible region.

The furniture retail business obviously wants to confirm that the business will make a profit, and will wish to determine how many armchairs and sofas to purchase from the wholesaler to maximize expected profit.

There are fixed costs for wages, rental etc., and we assume that these are independent of the number of armchairs and sofas sold.

The *gross margin* on the sale of an armchair or sofa is the difference between the wholesale and retail prices of that item of furniture.

Armchairs cost €400 wholesale and sell for €700, and thus provide a gross margin of €300.

Sofas cost €600 wholesale and sell for €1000, and thus provide a gross margin of €400.

In a typical linear programming problem, one wishes to determine not merely *feasible* solutions to the problem. One wishes to determine an *optimal* solution that maximizes some *objective function* amongst all feasible solutions to the problem.

The objective function for the Furniture Retailing problem is the gross profit that would accrue from selling the furniture in stock. This gross profit is the difference between the cost of purchasing the furniture from the wholesaler and the return from selling that furniture.

This objective function is thus $f(x, y)$, where

$$f(x, y) = 300x + 400y.$$

We should determine the maximum value of this function on the feasible region.

Because the objective function $f(x, y) = 300x + 400y$ is linear in x and y, its maximum value on the feasable region must be achieved at one of the vertices of the region.

Clearly this function is not maximized at the origin $(0, 0)$!

Now the remaining vertices of the feasible region are located at $(400, 300)$, $(510, 245)$ and $(585, 195)$, and

$$\begin{aligned} f(400, 300) &= 240,000, \\ f(510, 245) &= 251,000, \end{aligned}$$

$$f(585,195) \quad = \quad 253,500.$$

It follows that the objective function is maximized at $(585, 195)$.

The furniture retail business should therefore use up the available capital, stocking 3 armchairs for every sofa, despite the fact that this will not utilize the full amount of floor space available.

1.2 Linear Programming Problems in Von Neumann Maximizing Form

A linear programming problem may be presented in *Von Neumann maximizing form* as follows:

> *given real numbers* c_i, $A_{i,j}$ *and* b_j *for* $i = 1, 2, \ldots, m$ *and* $j = 1, 2, \ldots, n$,
>
> *find real numbers* $x_1, x_2, \ldots, x_n$ *so as to*
>
> *maximize* $c_1x_1 + c_2x_2 + \cdots + c_nx_n$
>
> *subject to constraints*
>
> $x_j \geq 0$ *for* $j = 1, 2, \ldots, n$, *and*
>
> $A_{i,1}x_1 + A_{i,2}x_2 + \cdots + A_{i,n}x_n \leq b_i$ *for* $i = 1, 2, \ldots, m$.

The furniture retailing problem may be presented in von Neumann maximizing form with $n = 2$, $m = 4$,

$$(c_1, c_2) = (300, 400),$$

$$A = \begin{pmatrix} 1 & -3 \\ -3 & 4 \\ 1 & 2 \\ 4 & 6 \end{pmatrix}, \quad \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1000 \\ 3510 \end{pmatrix}.$$

Here A represents the $m \times n$ whose coefficient in the ith row and jth column is $A_{i,j}$.

Linear programming problems may be presented in matrix form. We adopt the following notational conventions with regard to transposes, row and column vectors and vector inequalities:—

- vectors in $\mathbb{R}^m$ and $\mathbb{R}^n$ are represented as column vectors;

- we denote by M^{T} the $n \times m$ matrix that is the transpose of an $m \times n$ matrix M;
- in particular, given $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$, where b and c are represented as column vectors, we denote by b^{T} and c^{T} the corresponding row vectors obtained on transposing the column vectors representing b and c;
- given vectors u and v in $\mathbb{R}^n$ for some positive integer n, we write $u \leq v$ (and $v \geq u$) if and only if $u_j \leq v_j$ for $j = 1, 2, \ldots, n$.

A linear programming problem in von Neumann maximizing form may be presented in matrix notation as follows:—

Given an $m \times n$ matrix A with real coefficients,

and given column vectors $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$,

find $x \in \mathbb{R}^n$ so as to

maximize $c^{\mathrm{T}} x$

subject to constraints $Ax \leq b$ and $x \geq 0$.

2 Bases of Finite-Dimensional Vector Spaces

2.1 Real Vector Spaces

Definition A *real vector space* consists of a set V on which there is defined an operation of vector addition, yielding an element $\mathbf{v}+\mathbf{w}$ of V for each pair $\mathbf{v}, \mathbf{w}$ of elements of V, and an operation of multiplication-by-scalars that yields an element $\lambda\mathbf{v}$ of V for each $\mathbf{v} \in V$ and for each real number λ. The operation of vector addition is required to be commutative and associative. There must exist a zero element $\mathbf{0}_V$ of V that satisfies $\mathbf{v} + \mathbf{0}_V = \mathbf{v}$ for all $\mathbf{v} \in V$, and, for each $\mathbf{v} \in V$ there must exist an element $-\mathbf{v}$ of V for which $\mathbf{v}+(-\mathbf{v}) = \mathbf{0}_V$. The following identities must also be satisfied for all $\mathbf{v}, \mathbf{w} \in V$ and for all real numbers λ and μ:

$$(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}, \quad \lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w},$$

$$\lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}, \quad 1\mathbf{v} = \mathbf{v}.$$

Let n be a positive integer. The set $\mathbb{R}^n$ consisting of all n-tuples of real numbers is then a real vector space, with addition and multiplication-by-scalars defined such that

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots x_n + y_n)$$

and

$$\lambda(x_1, x_2, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$$

for all $(x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \in \mathbb{R}$ and for all real numbers λ.

The set $M_{m,n}(\mathbb{R})$ of all $m \times n$ matrices is a real vector space with respect to the usual operations of matrix addition and multiplication of matrices by real numbers.

2.2 Linear Dependence and Bases

Elements $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ of a real vector space V are said to be *linearly dependent* if there exist real numbers $\lambda_1, \lambda_2, \ldots, \lambda_m$, not all zero, such that

$$\lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \cdots + \lambda_m\mathbf{u}_m = \mathbf{0}_V.$$

If elements $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ of real vector space V are not linearly dependent, then they are said to be *linearly independent*.

Elements $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ of a real vector space V are said to *span* V if, given any element $\mathbf{v}$ of V, there exist real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\mathbf{v} = \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \cdots + \lambda_n\mathbf{u}_n$.

A vector space is said to be *finite-dimensional* if there exists a finite subset of V whose members span V.

Elements $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ of a finite-dimensional real vector space V are said to constitute a *basis* of V if they are linearly independent and span V.

Lemma 2.1 *Elements $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ of a real vector space V constitute a basis of V if and only if, given any element $\mathbf{v}$ of V, there exist uniquely-determined real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that*

$$\mathbf{v} = \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \cdots + \lambda_n \mathbf{u}_n.$$

Proof Suppose that $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ is a basis of V. Let $\mathbf{v}$ be an element V. The requirement that $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ span V ensures that there exist real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that

$$v = \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \cdots + \lambda_n \mathbf{u}_n.$$

If $\mu_1, \mu_2, \ldots, \mu_n$ are real numbers for which

$$v = \mu_1 \mathbf{u}_1 + \mu_2 \mathbf{u}_2 + \cdots + \mu_n \mathbf{u}_n,$$

then

$$(\mu_1 - \lambda_1)\mathbf{u}_1 + (\mu_2 - \lambda_2)\mathbf{u}_2 + \cdots + (\mu_n - \lambda_n)\mathbf{u}_n = \mathbf{0}_V.$$

It then follows from the linear independence of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ that $\mu_i - \lambda_i = 0$ for $i = 1, 2, \ldots, n$, and thus $\mu_i = \lambda_i$ for $i = 1, 2, \ldots, n$. This proves that the coefficients $\lambda_1, \lambda_2, \ldots, \lambda_n$ are uniquely-determined.

Conversely suppose that $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ is a list of elements of V with the property that, given any element $\mathbf{v}$ of V, there exist uniquely-determined real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that

$$v = \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \cdots + \lambda_n \mathbf{u}_n.$$

Then $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ span V. Moreover we can apply this criterion when $\mathbf{v} = 0$. The uniqueness of the coefficients $\lambda_1, \lambda_2, \ldots, \lambda_n$ then ensures that if

$$\lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \cdots + \lambda_n \mathbf{u}_n = \mathbf{0}_V$$

then $\lambda_i = 0$ for $i = 1, 2, \ldots, n$. Thus $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ are linearly independent. This proves that $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ is a basis of V, as required. ∎

Proposition 2.2 *Let V be a finite-dimensional real vector space, let*

$$\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$$

be elements of V that span V, and let K be a subset of $\{1, 2, \ldots, n\}$. Suppose either that $K = \emptyset$ or else that those elements $\mathbf{u}_i$ for which $i \in K$ are linearly independent. Then there exists a basis of V whose members belong to the list $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ which includes all the vectors $\mathbf{u}_i$ for which $i \in K$.

Proof We prove the result by induction on the number of elements in the list $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ of vectors that span V. The result is clearly true when $n = 1$. Thus suppose, as the induction hypothesis, that the result is true for all lists of elements of V that span V and that have fewer than n members.

If the elements $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ are linearly independent, then they constitute the required basis. If not, then there exist real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$, not all zero, such that

$$\lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \cdots + \lambda_n \mathbf{u}_n = \mathbf{0}_V.$$

Now there cannot exist real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$, not all zero, such that both $\sum_{i=1}^{n} \lambda_i \mathbf{u}_i = \mathbf{0}_V$ and also $\lambda_i = 0$ whenever $i \neq K$. Indeed, in the case where $K = \emptyset$, this conclusion follows from the requirement that the real numbers λ_i cannot all be zero, and, in the case where $K \neq \emptyset$, the conclusion follows from the linear independence of those $\mathbf{u}_i$ for which $i \in K$. Therefore there must exist some integer i satisfying $1 \leq i \leq n$ for which $\lambda_i \neq 0$ and $i \notin K$. Without loss of generality, we may suppose that $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ are ordered so that $n \notin K$ and $\lambda_n \neq 0$. Then

$$\mathbf{u}_n = -\sum_{i=1}^{n-1} \frac{\lambda_i}{\lambda_n} \mathbf{u}_i.$$

Let $\mathbf{v}$ be an element of V. Then there exist real numbers $\mu_1, \mu_2, \ldots, \mu_n$ such that $\mathbf{v} = \sum_{i=1}^{n} \mu_i \mathbf{u}_i$, because $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ span V. But then

$$\mathbf{v} = \sum_{i=1}^{n-1} \left(\mu_i - \frac{\mu_n \lambda_i}{\lambda_n} \right) \mathbf{u}_i.$$

We conclude that $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{n-1}$ span the vector space V. The induction hypothesis then ensures that there exists a basis of V consisting of members of this list that includes the linearly independent elements $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$, as required. ∎

Corollary 2.3 *Let V be a finite-dimensional real vector space, and let*

$$\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$$

be elements of V that span the vector space V. Then there exists a basis of V whose elements are members of the list $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$.

Proof This result is a restatement of Proposition 2.2 in the special case where the set K in the statement of that proposition is the empty set. ∎

2.3 Dual Spaces

Definition Let V be a real vector space. A *linear functional* $\varphi\colon V \to \mathbb{R}$ on V is a linear transformation from the vector space V to the field $\mathbb{R}$ of real numbers.

Given linear functionals $\varphi\colon V \to \mathbb{R}$ and $\psi\colon V \to \mathbb{R}$ on a real vector space V, and given any real number λ, we define $\varphi + \psi$ and $\lambda\varphi$ to be the linear functionals on V defined such that $(\varphi + \psi)(\mathbf{v}) = \varphi(\mathbf{v}) + \psi(\mathbf{v})$ and $(\lambda\varphi)(\mathbf{v}) = \lambda\,\varphi(\mathbf{v})$ for all $\mathbf{v} \in V$.

The set V^* of linear functionals on a real vector space V is itself a real vector space with respect to the algebraic operations of addition and multiplication-by-scalars defined above.

Definition Let V be a real vector space. The *dual space* V^* of V is the vector space whose elements are the linear functionals on the vector space V.

Now suppose that the real vector space V is finite-dimensional. Let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ be a basis of V, where $n = \dim V$. Given any $\mathbf{v} \in V$ there exist uniquely-determined real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\mathbf{v} = \sum_{j=1}^{n} \lambda_j \mathbf{u}_j$. It follows that there are well-defined functions $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ from V to the field $\mathbb{R}$ defined such that

$$\varepsilon_i \left(\sum_{j=1}^{n} \lambda_j \mathbf{u}_j \right) = \lambda_i$$

for $i = 1, 2, \ldots, n$ and for all real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$. These functions are linear transformations, and are thus linear functionals on V.

Lemma 2.4 *Let V be a finite-dimensional real vector space, let*

$$\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$$

be a basis of V, and let $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ be the linear functionals on V defined such that

$$\varepsilon_i \left(\sum_{j=1}^{n} \lambda_j \mathbf{u}_j \right) = \lambda_i$$

for $i = 1, 2, \ldots, n$ and for all real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ constitute a basis of the dual space V^* of V. Moreover

$$\varphi = \sum_{i=1}^{n} \varphi(\mathbf{u_i}) \varepsilon_i$$

for all $\varphi \in V^*$.

Proof Let $\mu_1, \mu_2, \ldots, \mu_n$ be real numbers with the property that $\sum_{i=1}^{n} \mu_i \varepsilon_i = \mathbf{0}_{V^*}$. Then

$$0 = \left(\sum_{i=1}^{n} \mu_i \varepsilon_i \right) (\mathbf{u}_j) = \sum_{i=1}^{n} \mu_i \varepsilon_i(\mathbf{u}_j) = \mu_j$$

for $j = 1, 2, \ldots, n$. Thus the linear functionals $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ on V are linearly independent elements of the dual space V^*.

Now let $\varphi: V \to \mathbb{R}$ be a linear functional on V, and let $\mu_i = \varphi(\mathbf{u}_i)$ for $i = 1, 2, \ldots, n$. Now

$$\varepsilon_i(\mathbf{u}_j) = \begin{cases} 1 & \text{if } i = j; \\ 0 & \text{if } i \neq j. \end{cases}$$

It follows that

$$\begin{aligned} \left(\sum_{i=1}^{n} \mu_i \varepsilon_i \right) \left(\sum_{j=1}^{n} \lambda_j \mathbf{u}_j \right) &= \sum_{i=1}^{n} \sum_{j=1}^{n} \mu_i \lambda_j \varepsilon_i(\mathbf{u}_j) = \sum_{j=1}^{n} \mu_j \lambda_j \\ &= \sum_{j=1}^{n} \lambda_j \varphi(\mathbf{u}_j) = \varphi \left(\sum_{j=1}^{n} \lambda_j \mathbf{u}_j \right) \end{aligned}$$

for all real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$.

It follows that

$$\varphi = \sum_{i=1}^{n} \mu_i \varepsilon_i = \sum_{i=1}^{n} \varphi(\mathbf{u}_i) \varepsilon_i.$$

We conclude from this that every linear functional on V can be expressed as a linear combination of $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$. Thus these linear functionals span V^*. We have previously shown that they are linearly independent. It follows that they constitute a basis of V^*. Moreover we have verified that $\varphi = \sum_{i=1}^{n} \varphi(\mathbf{u}_i) \varepsilon_i$ for all $\varphi \in V^*$, as required. ∎

Definition Let V be a finite-dimensional real vector space, let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ be a basis of V. The corresponding *dual basis* of the dual space V^* of V consists of the linear functionals $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ on V, where

$$\varepsilon_i \left(\sum_{j=1}^{n} \lambda_j \mathbf{u}_j \right) = \lambda_i$$

for $i = 1, 2, \ldots, n$ and for all real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Corollary 2.5 *Let V be a finite-dimensional real vector space, and let V^* be the dual space of V. Then* $\dim V^* = \dim V$.

Proof We have shown that any basis of V gives rise to a dual basis of V^*, where the dual basis of V has the same number of elements as the basis of V to which it corresponds. The result follows immediately from the fact that the dimension of a finite-dimensional real vector space is the number of elements in any basis of that vector space. $\blacksquare$

Let V be a real-vector space, and let V^* be the dual space of V. Then V^* is itself a real vector space, and therefore has a dual space V^{**}. Now each element $\mathbf{v}$ of V determines a corresponding linear functional $E_{\mathbf{v}}: V^* \to \mathbb{R}$ on V^*, where $E_{\mathbf{v}}(\varphi) = \varphi(\mathbf{v})$ for all $\varphi \in V^*$. It follows that there exists a function $\iota: V \to V^{**}$ defined so that $\iota(\mathbf{v}) = E_{\mathbf{v}}$ for all $\mathbf{v} \in V$. Then $\iota(\mathbf{v})(\varphi) = \varphi(\mathbf{v})$ for all $\mathbf{v} \in V$ and $\varphi \in V^*$.

Now

$$\iota(\mathbf{v} + \mathbf{w})(\varphi) = \varphi(\mathbf{v} + \mathbf{w}) = \varphi(\mathbf{v}) + \varphi(\mathbf{w}) = (\iota(\mathbf{v}) + \iota(\mathbf{w}))(\varphi)$$

and

$$\iota(\lambda\mathbf{v})(\varphi) = \varphi(\lambda\mathbf{v}) = \lambda\varphi(\mathbf{v}) = (\lambda\iota(\mathbf{v}))(\varphi)$$

for all $\mathbf{v}, \mathbf{w} \in V$ and $\varphi \in V^*$ and for all real numbers λ. It follows that $\iota(\mathbf{v} + \mathbf{w}) = \iota(\mathbf{v}) + \iota(\mathbf{w})$ and $\iota(\lambda\mathbf{v}) = \lambda\iota(\mathbf{v})$ for all $\mathbf{v}, \mathbf{w} \in V$ and for all real numbers λ. Thus $\iota: V \to V^{**}$ is a linear transformation.

Proposition 2.6 *Let V be a finite-dimensional real vector space, and let $\iota: V \to V^{**}$ be the linear transformation defined such that $\iota(\mathbf{v})(\varphi) = \varphi(\mathbf{v})$ for all $\mathbf{v} \in V$ and $\varphi \in V^*$. Then $\iota: V \to V^{**}$ is an isomorphism of real vector spaces.*

Proof Let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ be a basis of V, let $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ be the dual basis of V^*, where

$$\varepsilon_i(\mathbf{u}_j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

and let $\mathbf{v} \in V$. Then there exist real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\mathbf{v} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i$.

Suppose that $\iota(\mathbf{v}) = \mathbf{0}_{V^{**}}$. Then $\varphi(\mathbf{v}) = E_{\mathbf{v}}(\varphi) = 0$ for all $\varphi \in V^*$. In particular $\lambda_i = \varepsilon_i(\mathbf{v}) = 0$ for $i = 1, 2, \ldots, n$, and therefore $\mathbf{v} = \mathbf{0}_V$. We conclude that $\iota: V \to V^{**}$ is injective.

Now let $F: V^* \to \mathbb{R}$ be a linear functional on V^*, let $\lambda_i = F(\varepsilon_i)$ for $i = 1, 2, \ldots, n$, let $\mathbf{v} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i$, and let $\varphi \in V^*$. Then $\varphi = \sum_{i=1}^{n} \varphi(\mathbf{u}_i)\varepsilon_i$ (see

Lemma 2.4), and therefore

$$\begin{aligned}\iota(\mathbf{v})(\varphi) &= \varphi(\mathbf{v}) = \sum_{i=1}^{n} \lambda_i \varphi(\mathbf{u}_i) = \sum_{i=1}^{n} F(\varepsilon_i)\varphi(\mathbf{u}_i) \\ &= F\left(\sum_{i=1}^{n} \varphi(\mathbf{u}_i)\varepsilon_i\right) = F(\varphi).\end{aligned}$$

Thus $\iota(\mathbf{v}) = F$. We conclude that the linear transformation $\iota\colon V \to V^{**}$ is surjective. We have previously shown that this linear transformation is injective. There $\iota\colon V \to V^{**}$ is an isomorphism between the real vector spaces V and V^{**} as required. ∎

The following corollary is an immediate consequence of Proposition 2.6.

Corollary 2.7 *Let V be a finite-dimensional real vector space, and let V^* be the dual space of V. Then, given any linear functional $F\colon V^* \to \mathbb{R}$, there exists some $\mathbf{v} \in V$ such that $F(\varphi) = \varphi(\mathbf{v})$ for all $\varphi \in V^*$.*

3 The Transportation Problem

3.1 Transportation in the Dairy Industry

We discuss an example of the Transportation Problem of Linear Programming, as it might be applied to optimize transportation costs in the dairy industry.

A food business has milk-processing plants located in various towns in a small country. We shall refer to these plants as *dairies*. Raw milk is supplied by numerous farmers with farms located throughout that country, and is transported by milk tanker from the farms to the dairies. The problem is to determine the catchment areas of the dairies so as to minimize transport costs.

We suppose that there are m farms, labelled by integers from 1 to m that supply milk to n dairies, labelled by integers from 1 to n. Suppose that, in a given year, the ith farm has the capacity to produce and supply a s_i litres of milk for $i = 1, 2, \ldots, m$, and that the jth dairy needs to receive at least d_j litres of milk for $j = 1, 2, \ldots, n$ to satisfy the business obligations.

The quantity $\sum_{i=1}^{m} s_i$ then represents that *total supply* of milk, and the quantity $\sum_{j=1}^{n} d_j$ represents the *total demand* for milk.

We suppose that $x_{i,j}$ litres of milk are to be transported from the ith farm to the jth dairy, and that $c_{i,j}$ represents the cost per litre of transporting this milk.

Then the total cost of transporting milk from the farms to the dairies is

$$\sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j} x_{i,j}.$$

The quantities $x_{i,j}$ of milk to be transported from the farms to the dairies should then be determined for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$ so as to minimize the total cost of transporting milk.

However the ith farm can supply no more than s_i litres of milk in a given year, and that jth dairy requires at least d_j litres of milk in that year. It follows that the quantities $x_{i,j}$ of milk to be transported between farms and dairy are constrained by the requirements that

$$\sum_{j=1}^{n} x_{i,j} \leq s_i \quad \text{for } i = 1, 2, \ldots, m$$

and

$$\sum_{i=1}^{m} x_{i,j} \geq d_j \quad \text{for } j = 1, 2, \ldots, m.$$

3.2 The General Transportation Problem

The Transportation Problem can be expressed generally in the following form. Some commodity is supplied by m suppliers and is transported from those suppliers to n recipients. The ith supplier can supply at most to s_i units of the commodity, and the jth recipient requires at least d_j units of the commodity. The cost of transporting a unit of the commodity from the ith supplier to the jth recipient is $c_{i,j}$.

The total transport cost is then

$$\sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j} x_{i,j}.$$

where $x_{i,j}$ denote the number of units of the commodity transported from the ith supplier to the jth recipient.

The Transportation Problem can then be presented as follows:

> *determine* $x_{i,j}$ *for* $i = 1, 2, \ldots, m$ *and* $j = 1, 2, \ldots, n$ *so as minimize* $\sum_{i,j} c_{i,j} x_{i,j}$ *subject to the constraints* $x_{i,j} \geq 0$ *for all* i *and* j, $\sum_{j=1}^{n} x_{i,j} \leq s_i$ *and* $\sum_{i=1}^{m} x_{i,j} \geq d_j$, *where* $s_i \geq 0$ *for all* i, $d_j \geq 0$ *for all* i, *and* $\sum_{i=1}^{m} s_i \geq \sum_{j=1}^{n} d_j$.

3.3 Transportation Problems in which Total Supply equals Total Demand

Consider an instance of the Transportation Problem with m suppliers and n recipients. The following proposition shows that a solution to the Transportation Problem can only exist if total supply of the relevant commodity exceeds total demand for that commodity.

Proposition 3.1 *Let* $s_1, s_2, \ldots, s_m$ *and* $d_1, d_2, \ldots, d_n$ *be non-negative real numbers. Suppose that there exist non-negative real numbers* $x_{i,j}$ *be for* $i = 1, 2, \ldots, m$ *and* $j = 1, 2, \ldots, n$ *that satisfy the inequalities*

$$\sum_{j=1}^{n} x_{i,j} \leq s_i \quad \textit{and} \quad \sum_{i=1}^{m} x_{i,j} \geq d_j.$$

Then

$$\sum_{j=1}^{n} d_j \leq \sum_{i=1}^{m} s_i.$$

Moreover if it is the case that

$$\sum_{j=1}^{n} d_j = \sum_{i=1}^{m} s_i.$$

then

$$\sum_{j=1}^{n} x_{i,j} = s_i \quad \textit{for } i = 1, 2, \ldots, m$$

and

$$\sum_{i=1}^{m} x_{i,j} = d_j \quad \textit{for } j = 1, 2, \ldots, n.$$

Proof The inequalities satisfied by the non-negative real numbers $x_{i,j}$ ensure that

$$\sum_{j=1}^{n} d_j \leq \sum_{i=1}^{m} \sum_{j=1}^{n} x_{i,j} \leq \sum_{i=1}^{m} s_i.$$

Thus the total supply must equal or exceed the total demand.

If it is the case that $\sum_{j=1}^{n} x_{i,j} < s_i$ for at least one value of i then $\sum_{i=1}^{m} \sum_{j=1}^{n} x_{i,j} < \sum_{i=1}^{m} s_i$. Similarly if it is the case that $\sum_{i=1}^{m} x_{i,j} > d_j$ for at least one value of j then $\sum_{i=1}^{m} \sum_{j=1}^{n} x_{i,j} > \sum_{j=1}^{n} d_j$.

It follows that if total supply equals total demand, so that

$$\sum_{i=1}^{m} s_i = \sum_{j=1}^{n} d_j,$$

then

$$\sum_{j=1}^{n} x_{i,j} = s_i \quad \text{for } i = 1, 2, \ldots, m$$

and

$$\sum_{i=1}^{m} x_{i,j} = d_j \quad \text{for } j = 1, 2, \ldots, n,$$

as required. ∎

We analyse the Transportation Problem in the case where total supply equals total demand. The optimization problem in this case can then be stated as follows:—

> *determine* $x_{i,j}$ *for* $i = 1, 2, \ldots, m$ *and* $j = 1, 2, \ldots, n$ *so as minimize* $\sum_{i,j} c_{i,j} x_{i,j}$ *subject to the constraints* $x_{i,j} \geq 0$ *for all* i *and* j, $\sum_{j=1}^{n} x_{i,j} = s_i$ *and* $\sum_{i=1}^{m} x_{i,j} = d_j$, *where* $s_i \geq 0$ *and* $d_j \geq 0$ *for all* i *and* j, *and* $\sum_{i=1}^{m} s_i = \sum_{j=1}^{n} d_j$.

Definition A *feasible* solution to the Transportation Problem (with equality of total supply and total demand) takes the form of real numbers $x_{i,j}$, where

- $x_{i,j} \geq 0$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$;
- $\sum_{j=1}^{n} x_{i,j} = s_i$;
- $\sum_{i=1}^{m} x_{i,j} = d_j$.

Definition A feasible solution $(x_{i,j})$ of the Transportation Problem is said to be *optimal* if it minimizes cost amongst all feasible solutions of the Transportation Problem.

3.4 Row Sums and Column Sums of Matrices

We commence the analysis of the Transportation Problem by studying the interrelationships between the various real vector spaces and linear transformations that arise naturally from the statement of the Transportation Problem.

The quantities $x_{i,j}$ to be determined are coefficients of an $m \times n$ matrix X. This matrix X is represented as an element of the real vector space $M_{m,n}(\mathbb{R})$ that consists of all $m \times n$ matrices with real coefficients.

The non-negative quantities $s_1, s_2, \ldots, s_m$ that specify the sums of the coefficients in the rows of the unknown matrix X are the components of a *supply vector* $\mathbf{s}$ belonging to the m-dimensional real vector space $\mathbb{R}^m$.

Similarly the non-negative quantities $d_1, d_2, \ldots, d_n$ that specify the sums of the coefficients in the columns of the unknown matrix X are the components of a *demand vector* $\mathbf{d}$ belonging to the n-dimensional space $\mathbb{R}^n$.

The requirement that total supply equals total demand translates into a requirement that the sum $\sum_{i=1}^{m} (\mathbf{s})_i$ of the components of the supply vector $\mathbf{s}$ must equal the sum $\sum_{j=1}^{n} (\mathbf{d})_j$ of the components of the demand vector $\mathbf{d}$.

Accordingly we introduce a real vector space W consisting of all ordered pairs $(\mathbf{y}, \mathbf{z})$ for which $\mathbf{y} \in \mathbb{R}^m$, $\mathbf{z} \in \mathbb{R}^n$ and $\sum_{i=1}^{m} (\mathbf{y})_i = \sum_{j=1}^{n} (\mathbf{z})_j$.

Lemma 3.2 *Let m and n be positive integers, and let*

$$W = \left\{ (\mathbf{y}, \mathbf{z}) \in \mathbb{R}^m \times \mathbb{R}^n : \sum_{i=1}^{m} (\mathbf{y})_i = \sum_{j=1}^{n} (\mathbf{z})_j \right\}.$$

Then the dimension of the real vector space W is $m + n - 1$. ▮

Proof It is easy to see that the vector space W is isomorphic to $\mathbb{R}^m$ when $n = 1$. The result then follows directly in the case when $n = 1$. Thus suppose that $n > 1$.

Given real numbers $y_1, y_2, \ldots, y_n$ and $z_1, z_2, \ldots, z_{n-1}$, there exists exactly one element $(\mathbf{y}, \mathbf{z})$ of W that satisfies $(\mathbf{y})_i = y_i$ for $i = 1, 2, \ldots, m$ and $(\mathbf{z})_j = z_j$ for $j = 1, 2, \ldots, n - 1$. The remaining component $(\mathbf{z})_n$ of the n-dimensional vector $\mathbf{z}$ is then determined by the equation

$$(\mathbf{z})_n = \sum_{i=1}^{m} y_i - \sum_{j=1}^{m-1} z_j.$$

It follows from this that $\dim W = m + n - 1$, as required. ▮

The supply and demand constraints on the sums of the rows and columns of the unknown matrix X can then be specified by means of linear transformations

$$\rho \colon M_{m,n}(\mathbb{R}) \to \mathbb{R}^m$$

and

$$\sigma \colon M_{m,n}(\mathbb{R}) \to \mathbb{R}^n,$$

where, for each $X \in M_{m,n}(\mathbb{R})$, the components of the m-dimensional vector $\rho(X)$ are the sums of the coefficients along each row of X, and the components of the n-dimensional vector $\sigma(X)$ are the sums of the coefficients along each column of X.

Let $X \in M_{m,n}(\mathbb{R})$. Then the ith component $\rho(X)_i$ of the vector $\rho(X)$ is determined by the equation

$$\rho(X)_i = \sum_{j=1}^{n} (X)_{i,j} \quad \text{for } i = 1, 2, \ldots, m,$$

for $i = 1, 2, \ldots, m$, and the jth component $\sigma(X)_j$ of $\sigma(X)$ is determined by the equation

$$\sigma(X)_j = \sum_{i=1}^{m} (X)_{i,j} \quad \text{for } j = 1, 2, \ldots, n.$$

for $j = 1, 2, \ldots, n$.

The costs $c_{i,j}$ are the components of an $m \times n$ matrix C, the *cost matrix*, that in turn determines a linear functional

$$f \colon M_{m,n}(\mathbb{R}) \to \mathbb{R}$$

on the vector space $M_{m,n}(\mathbb{R})$ defined such that

$$f(X) = \operatorname{trace}(C^T X) = \sum_{i=1}^{m} \sum_{j=1}^{n} (C)_{i,j} X_{i,j}$$

for all $X \in M_{m,n}(\mathbb{R})$.

An instance of the problem is specified by specifying a supply vector $\mathbf{s}$, demand vector $\mathbf{d}$ and cost matrix C. The components of $\mathbf{s}$ and $\mathbf{d}$ are required to be non-negative real numbers. Moreover $(\mathbf{s}, \mathbf{d}) \in W$, where W is the real vector space consisting of all ordered pairs $(\mathbf{s}, \mathbf{d})$ with $\mathbf{s} \in \mathbb{R}^m$ and $\mathbf{d} \in \mathbb{R}^n$ for which the sum of the components of the vector $\mathbf{s}$ equals the sum of the components of the vector $\mathbf{d}$.

A feasible solution of the Transportation Problem with given supply vector $\mathbf{s}$, demand vector $\mathbf{d}$ and cost matrix C is represented by an $m \times n$ matrix X satisfying the following three conditions:—

- The coefficients of X are all non-negative;
- $\rho(X) = \mathbf{s}$;
- $\sigma(X) = \mathbf{d}$.

The cost functional $f \colon M_{m,n}(\mathbb{R}) \to \mathbb{R}$ is defined so that

$$f(X) = \operatorname{trace}(C^T X)$$

for all $X \in M_{m,n}(\mathbb{R})$.

A feasible solution X of the Transportation problem is optimal if and only if $f(X) \leq f(\overline{X})$ for all feasible solutions $\overline{X}$ of that problem.

Lemma 3.3 *Let $M_{m,n}(\mathbb{R})$ be the real vector space consisting of all $m \times n$ matrices with real coefficients, let $\rho\colon M_{m,n}(\mathbb{R}) \to \mathbb{R}^m$ and $\sigma\colon M_{m,n}(\mathbb{R}) \to \mathbb{R}^n$ be the linear transformations defined so that the ith component $\rho(X)_i$ of $\rho(X)$ satisfies*

$$\rho(X)_i = \sum_{j=1}^{n} (X)_{i,j} \quad \text{for } i = 1, 2, \ldots, m,$$

for $i = 1, 2, \ldots, m$, and the jth component $\sigma(X)_j$ of $\sigma(X)$ satisfies

$$\sigma(X)_j = \sum_{i=1}^{m} (X)_{i,j} \quad \text{for } j = 1, 2, \ldots, n.$$

for $j = 1, 2, \ldots, n$. Then $(\rho(X), \sigma(X)) \in W$ for all $X \in M_{m,n}(\mathbb{R})$, where

$$W = \left\{ (\mathbf{y}, \mathbf{z}) \in \mathbb{R}^m \times \mathbb{R}^n : \sum_{i=1}^{m} (\mathbf{y})_i = \sum_{j=1}^{n} (\mathbf{z})_j \right\}.$$

Proof Let $X \in M_{m,n}(\mathbb{R})$. Then

$$\sum_{i=1}^{m} \rho(X)_i = \sum_{i=1}^{m} \sum_{j=1}^{n} (X)_{i,j} = \sum_{j=1}^{n} \sigma(X)_j.$$

It follows that $(\rho(X), \sigma(X)) \in W$ for all $X \in M_{m,n}(\mathbb{R})$, as required. ∎

3.5 Bases for the Transportation Problem

The real vector space $M_{m,n}(\mathbb{R})$ consisting of all $m \times n$ matrices with real coefficients has a natural basis consisting of the matrices $E^{(i,j)}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, where, for each i and j, the coefficient of the matrix $E^{(i,j)}$ in the ith row and jth column has the value 1, and all other coefficients are zero. Indeed

$$X = \sum_{i=1}^{m} \sum_{j=1}^{n} (X)_{i,j} E^{(i,j)}$$

for all $X \in M_{m,n}(\mathbb{R})$.

Let $\rho\colon M_{m,n}(\mathbb{R}) \to \mathbb{R}^m$ and $\sigma\colon M_{m,n}(\mathbb{R}) \to \mathbb{R}^n$ be the linear transformations defined such that $(\rho(X))_i = \sum_{j=1}^{n} (X)_{i,j}$ for $i = 1, 2, \ldots, m$ and $(\sigma(X))_j = \sum_{i=1}^{m} (X)_{i,j}$ for $j = 1, 2, \ldots, n$. Then $\rho(E^{(i,j)}) = \overline{\mathbf{b}}^{(i)}$ for $i = 1, 2, \ldots, m$, where $\overline{\mathbf{b}}^{(i)}$ denotes the ith vector in the standard basis of $\mathbb{R}^m$, defined such that

$$(\overline{\mathbf{b}}^{(i)})_k = \begin{cases} 1 & \text{if } i = k; \\ 0 & \text{if } i \neq k. \end{cases}$$

Similarly $\sigma(E^{(i,j)}) = \mathbf{b}^{(j)}$ for $j = 1, 2, \ldots, n$, where $\mathbf{b}^{(j)}$ denotes the jth vector in the standard basis of $\mathbb{R}^n$, defined such that

$$(\mathbf{b}^{(j)})_l = \begin{cases} 1 & \text{if } j = l; \\ 0 & \text{if } j \neq l. \end{cases}$$

Now $(\rho(X), \sigma(X)) \in W$ for all $X \in M_{m,n}(\mathbb{R})$, where

$$W = \left\{ (\mathbf{y}, \mathbf{z}) \in \mathbb{R}^m \times \mathbb{R}^n : \sum_{i=1}^{m} (\mathbf{y})_i = \sum_{j=1}^{n} (\mathbf{z})_j \right\}.$$

Let $I = \{1, 2, \ldots, m\}$ and $J = \{1, 2, \ldots, n\}$, and let

$$\beta^{(i,j)} = (\overline{\mathbf{b}}^{(i)}, \mathbf{b}^{(j)})$$

for all $(i, j) \in I \times J$. Then the elements $\beta^{(i,j)}$ span the vector space W. It follows from basic linear algebra that there exist subsets B of $I \times J$ such that the elements $\beta^{(i,j)}$ of W for which $(i, j) \in B$ constitute a basis of the real vector space W (see Corollary 2.3).

Definition Let $I = \{1, 2, \ldots, m\}$ and $J = \{1, 2, \ldots, n\}$, where m and n are positive integers, let

$$W = \left\{ (\mathbf{y}, \mathbf{z}) \in \mathbb{R}^m \times \mathbb{R}^n : \sum_{i=1}^{m} (\mathbf{y})_i = \sum_{j=1}^{n} (\mathbf{z})_j \right\},$$

and, for each $(i, j) \in I \times J$, let $\beta^{(i,j)} = (\overline{\mathbf{b}}^{(i)}, \mathbf{b}^{(j)})$, where $\overline{\mathbf{b}}^{(i)} \in \mathbb{R}^m$ and $\mathbf{b}^{(j)} \in \mathbb{R}^n$ are defined so that the ith component of $\overline{\mathbf{b}}^{(i)}$ and that jth component of $\mathbf{b}^{(j)}$ are equal to 1 and the other components of these vectors are zero. A subset B of $I \times J$ is said to be a *basis* for the Transportation Problem with m suppliers and n recipients if and only if the elements $\beta^{(i,j)}$ for which $(i, j) \in B$ constitute a basis of the real vector space W.

The real vector space W is of dimension $m+n-1$, where m is the number of suppliers and n is the number of recipients. It follows that any basis for the Transportation Problem with m suppliers and n recipients has $m+n-1$ members.

Proposition 3.4 *Let $I = \{1, 2, \ldots, m\}$ and $J = \{1, 2, \ldots, n\}$, where m and n are positive integers. Then a subset B of $I \times J$ is a basis for the transportation problem if and only if, given any vectors $\mathbf{y} \in \mathbb{R}^m$ and $\mathbf{z} \in \mathbb{R}^n$ satisfying $\sum_{i=1}^{m} (\mathbf{y})_m = \sum_{j=1}^{n} (\mathbf{z})_n$, there exists a unique $m \times n$ matrix X with real coefficients satisfying the following properties:—*

(i) $\sum_{j=1}^{n} (X)_{i,j} = (\mathbf{y})_i$ *for* $i = 1, 2, \ldots, m$*;*

(ii) $\sum_{i=1}^{m} (X)_{i,j} = (\mathbf{z})_j$ *for* $j = 1, 2, \ldots, n$*;*

(ii) $(X)_{i,j} = 0$ *unless* $(i, j) \in B$*.*

Proof For each $(i, j) \in I \times J$, let $E^{(i,j)}$ denote the matrix whose coefficient in the ith row and jth column are equal to 1 and whose other coefficients are zero, and let $\rho(X) \in \mathbb{R}^m$ and $\sigma(X) \in \mathbb{R}^n$ be defined for all $m \times n$ matrices X with real coefficients so that $(\rho(X))_i = \sum_{j=1}^{n} (X)_{i,j}$ and $(\sigma(X))_j = \sum_{i=1}^{m} (X)_{i,j}$. Then $\rho(E^{(i,j)}) = \overline{\mathbf{b}}^{(i)}$ for $i = 1, 2, \ldots, m$, where $\overline{\mathbf{b}}^{(i)}$ denotes the vector in $\mathbb{R}^m$ whose ith component is equal to 1 and whose other components are zero. Similarly $\sigma(E^{(i,j)}) = \mathbf{b}^{(j)}$ for $j = 1, 2, \ldots, n$, where $\mathbf{b}^{(j)}$ denotes the vector in $\mathbb{R}^m$ whose jth component is equal to 1 and whose other components are zero.

Let

$$W = \left\{ (\mathbf{y}, \mathbf{z}) \in \mathbb{R}^m \times \mathbb{R}^n : \sum_{i=1}^{m} (\mathbf{y})_i = \sum_{j=1}^{n} (\mathbf{z})_j \right\}.$$

Then $(\rho(X), \sigma(X)) \in W$ for all $X \in M_{m,n}(\mathbb{R})$, and

$$(\rho(E^{(i,j)}), \sigma(E^{(i,j)}) = \beta^{(i,j)}$$

for all $(i, j) \in I \times J$ where

$$\beta^{(i,j)} = (\overline{\mathbf{b}}^{(i)}, \mathbf{b}^{(j)}).$$

Let B be a subset of $I \times J$, let $\mathbf{y}$ and $\mathbf{z}$ be elements of $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively that satisfy $(\mathbf{y}, \mathbf{z}) \in W$, and let X be an $m \times n$ matrix with real coefficients with the property that $(X)_{i,j} = 0$ unless $(i, j) \in B$. Then

$$\rho(X) = \sum_{(i,j) \in B} (X)_{(i,j)} \rho(E^{(i,j)}) = \sum_{(i,j) \in B} (X)_{(i,j)} \overline{\mathbf{b}}^{(i)}$$

and

$$\sigma(X) = \sum_{(i,j) \in B} (X)_{(i,j)} \sigma(E^{(i,j)}) = \sum_{(i,j) \in B} (X)_{(i,j)} \mathbf{b}^{(j)},$$

and therefore

$$(\rho(X), \sigma(X)) = \sum_{(i,j) \in B} (X)_{(i,j)} \beta^{(i,j)}.$$

Suppose that, given any $\mathbf{y} \in \mathbb{R}^m$ and $\mathbf{z} \in \mathbb{R}^n$ for which $(\mathbf{y}, \mathbf{z}) \in W$, there exists a unique $m \times n$ matrix X such that $\mathbf{y} = \rho(X)$, $\mathbf{z} = \sigma(X)$ and $(X)_{i,j} = 0$ for all $(i,j) \in B$. Then the elements $\beta^{(i,j)}$ of W for which $(i,j) \in B$ must span W and must also be linearly independent. These elements must therefore constitute a basis for the vector space B. It then follows that the subset B of $I \times J$ must be a basis for the Transportation Problem.

Conversely if B is a basis for the Transportation Problem then, given any $(\mathbf{y}, \mathbf{z}) \in W$, there must exist a unique $m \times n$ matrix X with real coefficients such that $(\mathbf{y}, \mathbf{z}) = \sum_{(i,j) \in B} (X)_{i,j} \beta^{(i,j)}$ and $(X)_{i,j} = 0$ unless $(i,j) \in B$. The result follows. ∎

Lemma 3.5 *Let m and n be positive integers, let $I = \{1, 2, \ldots, m\}$ and $J = \{1, 2, \ldots, n\}$, and let K be a subset of $I \times J$. Suppose that there is no basis B of the Transportation Problem for which $K \subset B$. Then there exists a non-zero $m \times n$ matrix Y with real coefficients which satisfies the following conditions:*

- $\sum_{j=1}^{n} (Y)_{i,j} = 0$ *for* $i = 1, 2, \ldots, m$;
- $\sum_{i=1}^{m} (Y)_{i,j} = 0$ *for* $j = 1, 2, \ldots, n$;
- $(Y)_{i,j} = 0$ *when* $(i,j) \notin K$.

Proof Let

$$W = \left\{ (\mathbf{y}, \mathbf{z}) \in \mathbb{R}^m \times \mathbb{R}^n : \sum_{i=1}^{m} (\mathbf{y})_i = \sum_{j=1}^{n} (\mathbf{z})_j \right\},$$

let $\overline{\mathbf{b}}^{(1)}, \overline{\mathbf{b}}^{(2)}, \ldots, \overline{\mathbf{b}}^{(m)}$ be the standard basis of $\mathbb{R}^m$ and let $\mathbf{b}^{(1)}, \mathbf{b}^{(2)}, \ldots, \mathbf{b}^{(n)}$ be the standard basis of $\mathbb{R}^n$, where the ith component of $\overline{\mathbf{b}}^{(i)}$ and the jth component of $\mathbf{b}^{(j)}$ are equal to 1 and the other components of these vectors are zero, and let $\beta^{(i,j)} = (\overline{\mathbf{b}}^{(i)}, \mathbf{b}^{(j)})$ for all $(i,j) \in I \times J$.

Now follows from Proposition 2.2 that if the elements $\beta^{(i,j)}$ for which $(i,j) \in K$ were linearly independent then there would exist a subset B of $I \times J$ satisfying $K \subset B$ such that the elements $\beta^{(i,j)}$ for which $(i,j) \in B$ would constitute a basis of W. This subset B of $I \times J$ would then be a basis for the Transportation Problem. But the subset K is not contained in any basis for the Transportation Problem. It follows that the elements $\beta^{(i,j)}$ for which $(i,j) \in K$ must be linearly dependent. Therefore there exists a non-zero $m \times n$ matrix Y with real coefficients such that $(Y)_{i,j} = 0$ when $(i, j) \notin K$ and

$$\sum_{i=1}^{m} \sum_{j=1}^{n} (Y)_{i,j} \beta^{(i,j)} = \mathbf{0}_W.$$

Now $\beta^{(i,j)} = (\overline{\mathbf{b}}^{(i)}, \mathbf{b}^{(j)})$ for all $i \in I$ and $j \in J$. It follows that

$$\sum_{i=1}^{m} \sum_{j=1}^{n} (Y)_{i,j} \overline{\mathbf{b}}^{(i)} = \mathbf{0}$$

and

$$\sum_{i=1}^{m} \sum_{j=1}^{n} (Y)_{i,j} \mathbf{b}^{(j)} = \mathbf{0},$$

and therefore

$$\sum_{j=1}^{n} (Y)_{i,j} = 0 \quad (i = 1, 2, \ldots, m)$$

and

$$\sum_{i=1}^{m} (Y)_{i,j} = 0 \quad (j = 1, 2, \ldots, n),$$

as required. ∎

3.6 Basic Feasible Solutions of Transportation Problems

Consider the Transportation Problem with m suppliers and n recipients, where the ith supplier can provide at most s_i units of some given commodity, where $s_i \geq 0$, and the jth recipient requires at least d_j units of that commodity, where $d_j \geq 0$. We suppose also that total supply equals total demand, so that

$$\sum_{i=1}^{m} s_i = \sum_{j=1}^{n} d_j,$$

The cost of transporting the commodity from the ith supplier to the jth recipient is $c_{i,j}$.

The concept of a *basis* for the Transportation Problem was introduced in Subsection 3.5. We recall some results established in that subsection.

Let $I = \{1, 2, \ldots, m\}$ and $J = \{1, 2, \ldots, n\}$. A subset B of $I \times J$ is a basis for the Transportation Problem if and only if, given any vectors $\mathbf{y} \in \mathbb{R}^m$ and

$\mathbf{z} \in \mathbb{R}^n$ satisfying $\sum_{i=1}^{m} (\mathbf{y})_i = \sum_{j=1}^{n} (\mathbf{z})_j$, there exists a unique matrix X with real coefficients such that $\sum_{j=1}^{n} (X)_{i,j} = (\mathbf{y})_i$ for $i = 1, 2, \ldots, m$, $\sum_{i=1}^{m} (X)_{i,j} = (\mathbf{z})_j$ for $j = 1, 2, \ldots, n$ and $(X)_{i,j} = 0$ unless $(i, j) \in B$ (see Proposition 3.4). A basis for the transportation problem has $m + n - 1$ elements.

Also if K is a subset of $I \times J$ that is not contained in any basis for the Transportation Problem then there exists a non-zero $m \times n$ matrix Y such that $\sum_{j=1}^{n} (Y)_{i,j} = 0$ for $i = 1, 2, \ldots, m$, $\sum_{i=1}^{m} (X)_{i,j} = 0$ for $j = 1, 2, \ldots, n$ and $(Y)_{i,j} = 0$ unless $(i, j) \in K$ (see Lemma 3.5).

Definition A feasible solution $(x_{i,j})$ of a Transportation Problem is said to be *basic* if there exists a basis B for that Transportation Problem such that $x_{i,j} = 0$ whenever $(i, j) \notin B$.

Example Consider the instance of the Transportation Problem where $m = n = 2$, $s_1 = 8$, $s_2 = 3$, $d_1 = 2$, $d_2 = 9$, $c_{1,1} = 2$, $c_{1,2} = 3$, $c_{2,1} = 4$ and $c_{2,2} = 1$.

A feasible solution takes the form of a 2×2 matrix

$$\begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{pmatrix}$$

with non-negative components which satisfies the two matrix equations

$$\begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 8 \\ 3 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{pmatrix} = \begin{pmatrix} 2 & 9 \end{pmatrix}.$$

A basic feasible solution will have at least one component equal to zero. There are four matrices with at least one zero component which satisfy the required equations. They are the following:—

$$\begin{pmatrix} 0 & 8 \\ 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 8 & 0 \\ -6 & 9 \end{pmatrix}, \quad \begin{pmatrix} 2 & 6 \\ 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} -1 & 9 \\ 3 & 0 \end{pmatrix}.$$

The first and third of these matrices have non-negative components. These two matrices represent basic feasible solutions to the problem, and moreover they are the only basic feasible solutions.

The costs associated with the components of the matrices are $c_{1,1} = 2$, $c_{1,2} = 3$, $c_{2,1} = 4$ and $c_{2,2} = 1$.

The cost of the basic feasible solution $\begin{pmatrix} 0 & 8 \\ 2 & 1 \end{pmatrix}$ is

$$8c_{1,2} + 2c_{2,1} + c_{2,2} = 24 + 8 + 1 = 33.$$

The cost of the basic feasible solution $\begin{pmatrix} 2 & 6 \\ 0 & 3 \end{pmatrix}$ is

$$2c_{1,1} + 6c_{1,2} + 3c_{2,2} = 4 + 18 + 3 = 25.$$

Now any 2×2 matrix $\begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{pmatrix}$ satisfying the two matrix equations

$$\begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 8 \\ 3 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{pmatrix} = \begin{pmatrix} 2 & 9 \end{pmatrix}$$

must be of the form

$$\begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{pmatrix} = \begin{pmatrix} \lambda & 8 - \lambda \\ 2 - \lambda & 1 + \lambda \end{pmatrix}$$

for some real number λ.

But the matrix $\begin{pmatrix} \lambda & 8 - \lambda \\ 2 - \lambda & 1 + \lambda \end{pmatrix}$ has non-negative components if and only if $0 \leq \lambda \leq 2$. It follows that the set of feasible solutions of this instance of the transportation problem is

$$\left\{ \begin{pmatrix} \lambda & 8 - \lambda \\ 2 - \lambda & 1 + \lambda \end{pmatrix} : \lambda \in \mathbb{R} \text{ and } 0 \leq \lambda \leq 2 \right\}.$$

The costs associated with the components of the matrices are $c_{1,1} = 2$, $c_{1,2} = 3$, $c_{2,1} = 4$ and $c_{2,2} = 1$. Therefore, for each real number λ satisfying $0 \leq \lambda \leq 2$, the cost $f(\lambda)$ of the feasible solution $\begin{pmatrix} \lambda & 8 - \lambda \\ 2 - \lambda & 1 + \lambda \end{pmatrix}$ is given by

$$f(\lambda) = 2\lambda + 3(8 - \lambda) + 4(2 - \lambda) + (1 + \lambda) = 33 - 4\lambda.$$

Cost is minimized when $\lambda = 2$, and thus $\begin{pmatrix} 2 & 6 \\ 0 & 3 \end{pmatrix}$ is the optimal solution of this instance of the Transportation Problem. The cost of this optimal solution is 25.

Proposition 3.6 *Given any feasible solution of the Transportation Problem, there exists a basic feasible solution with whose cost does not exceed that of the given solution.*

Proof Let m and n be positive integers, and let $\mathbf{s}$ and $\mathbf{d}$ be elements of $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively that satisfy $(\mathbf{s})_i \geq 0$ for $i = 1, 2, \ldots, m$, $(\mathbf{d})_i \geq 0$ for $j = 1, 2, \ldots, n$ and $\sum_{i=1}^{m} (\mathbf{s})_i = \sum_{j=1}^{n} (\mathbf{d})_j$, let C be an $m \times n$ matrix whose components are non-negative real numbers, and let X be a feasible solution of the resulting instance of the Transportation Problem with cost matrix C.

Let $s_i = (\mathbf{s})_i$, $d_j = (\mathbf{d})_j$, $x_{i,j} = (X)_{i,j}$ and $c_{i,j} = (C)_{i,j}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. Then $x_{i,j} \geq 0$ for all i and j, $\sum_{j=1}^{n} x_{i,j} = s_i$ for $i = 1, 2, \ldots, m$ and $\sum_{i=1}^{m} x_{i,j} = d_j$ for $j = 1, 2, \ldots, n$. The cost of the feasible solution X is then $\sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j} x_{i,j}$.

If the feasible solution X is itself basic then there is nothing to prove. Suppose therefore that X is not a basic solution. We show that there then exists a feasible solution $\overline{X}$ with fewer non-zero components than the given feasible solution.

Let $I = \{1, 2, \ldots, m\}$ and $J = \{1, 2, \ldots, n\}$, and let

$$K = \{(i, j) \in I \times J : x_{i,j} > 0\}.$$

Because X is not a basic solution to the Transportation Problem, there does not exist any basis B for the Transportation Problem satisfying $K \subset B$. It therefore follows from Lemma 3.5 that there exists a non-zero $m \times n$ matrix Y which satisfies the following conditions:—

- $\sum_{j=1}^{n} (Y)_{i,j} = 0$ for $i = 1, 2, \ldots, m$;
- $\sum_{i=1}^{m} (Y)_{i,j} = 0$ for $j = 1, 2, \ldots, n$;
- $(Y)_{i,j} = 0$ when $(i, j) \notin K$.

We can assume without loss of generality that $\sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j} (Y)_{i,j} \geq 0$, because otherwise we can replace Y with $-Y$.

Let $Z_\lambda = X - \lambda Y$ for all real numbers λ. Then $(Z_\lambda)_{i,j} = x_{i,j} - \lambda y_{i,j}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, where $x_{i,j} = (X)_{i,j}$ and $y_{i,j} = (Y)_{i,j}$.

Moreover the matrix Z_λ has the following properties:—

- $\sum_{j=1}^{n} (Z_\lambda)_{i,j} = s_i$;
- $\sum_{i=1}^{m} (Z_\lambda)_{i,j} = d_j$;
- $(Z_\lambda)_{i,j} = 0$ whenever $(i,j) \notin K$;
- $\sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j}(Z_\lambda)_{i,j} \leq \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j}(X)_{i,j}$ whenever $\lambda \geq 0$.

Now the matrix Y is a non-zero matrix whose rows and columns all sum to zero. It follows that at least one of its coefficients must be strictly positive. Thus there exists at least one ordered pair (i,j) belonging to the set K for which $y_{i,j} > 0$. Let

$$\lambda_0 = \text{minimum} \left\{ \frac{x_{i,j}}{y_{i,j}} : (i,j) \in K \text{ and } y_{i,j} > 0 \right\}.$$

Then $\lambda_0 > 0$. Moreover if $0 \leq \lambda < \lambda_0$ then $x_{i,j} - \lambda y_{i,j} > 0$ for all $(i,j) \in K$, and if $\lambda > \lambda_0$ then there exists at least one element (i_0, j_0) of K for which $x_{i_0,j_0} - \lambda y_{i_0,j_0} < 0$. It follows that $x_{i,j} - \lambda_0 y_{i,j} \geq 0$ for all $(i,j) \in K$, and $x_{i_0,j_0} - \lambda_0 y_{i_0,j_0} = 0$.

Thus Z_{λ_0} is a feasible solution of the given Transportation Problem whose cost does not exceed that of the given feasible solution X. Moreover Z_{λ_0} has fewer non-zero components than the given feasible solution X.

If Z_{λ_0} is itself a basic feasible solution, then we have found the required basic feasible solution whose cost does not exceed that of the given feasible solution. Otherwise we can iterate the process until we arrive at the required basic feasible solution whose cost does not exceed that of the given feasible solution. ∎

A given instance of the Transportation Problem has only finitely many basic feasible solutions. Indeed there are only finitely many bases for the problem, and any basis is associated with at most one basic feasible solution. Therefore there exists a basic feasible solution whose cost does not exceed the cost of any other basic feasible solution. It then follows from Proposition 3.6 that the cost of this basic feasible solution cannot exceed the cost of any other feasible solution of the given instance of the Transportation Problem. This basic feasible solution is thus a basic optimal solution of the Transportation Problem.

The Transportation Problem determined by the supply vector, demand vector and cost matrix has only finitely many basic feasible solutions, because there are only finitely many bases for the problem, and each basis can determine at most one basic feasible solution. Nevertheless the number of basic feasible solutions may be quite large.

But it can be shown that the Transportation Problem always has a basic optimal solution. It can be found using an algorithm that implements the Simplex Method devised by George B. Dantzig in the 1940s. This algorithm involves passing from one basis to another, lowering the cost at each stage, until one eventually finds a basis that can be shown to determine a basic optimal solution of the Transportation Problem.

3.7 An Example illustrating the Procedure for finding an Initial Basic Feasible Solution to a Transportation Problem using the Minimum Cost Method

We discuss the method for finding a basic optimal solution of the Transportation Problem by working through a particular example. First we find an initial basic feasible solution using a method known as the *Minimum Cost Method*. Then we test whether or not this initial basic feasible solution is optimal. It turns out that, in this example, the initial basic solutions is not optimal. We then commence an iterative process for finding a basic optimal solution.

Let $c_{i,j}$ be the coefficient in the ith row and jth column of the cost matrix C, where

$$C = \begin{pmatrix} 8 & 4 & 16 \\ 3 & 7 & 2 \\ 13 & 8 & 6 \\ 5 & 7 & 8 \end{pmatrix}.$$

and let

$$s_1 = 13, \quad s_2 = 8, \quad s_3 = 11, \quad s_4 = 13,$$

$$d_1 = 19, \quad d_2 = 12, \quad d_3 = 14.$$

We seek to non-negative real numbers $x_{i,j}$ for $i = 1, 2, 3, 4$ and $j = 1, 2, 3$ that minimize $\sum_{i=1}^{4} \sum_{j=1}^{3} c_{i,j} x_{i,j}$ subject to the following constraints:

$$\sum_{j=1}^{3} x_{i,j} = s_i \quad \text{for} \quad i = 1, 2, 3, 4,$$

$$\sum_{i=1}^{4} x_{i,j} = d_j \quad \text{for} \quad j = 1, 2, 3,$$

and $x_{i,j} \geq 0$ for all i and j.

For this problem the supply vector is $(13, 8, 11, 13)$ and the demand vector is $(19, 12, 14)$. The components of both the supply vector and the demand vector add up to 45.

In order to start the process of finding an initial basic solution for this problems, we set up a tableau that records the row sums (or supplies), the column sums (or demands) and the costs $c_{i,j}$ for the given problem, whilst leaving cells to be filled in with the values of the non-negative real numbers $x_{i,j}$ that will specify the initial basic feasible solution. The resultant tableau is structured as follows:—

$c_{i,j} \searrow x_{i,j}$	1	2	3	s_i
1	8 ?	4 ?	16 ?	13
2	3 ?	7 ?	2 ?	8
3	13 ?	8 ?	6 ?	11
4	5 ?	7 ?	8 ?	13
d_j	19	12	14	45

We apply the minimum cost method to find an initial basic solution.

The cell with lowest cost is the cell $(2, 3)$. We assign to this cell the maximum value possible, which is the minimum of s_2, which is 8, and d_3, which is 14. Thus we set $x_{2,3} = 8$. This forces $x_{2,1} = 0$ and $x_{2,2} = 0$. The pair $(2, 3)$ is added to the current basis.

The next undetermined cell of lowest cost is $(1, 2)$. We assign to this cell the minimum of s_1, which is 13, and $d_2 - x_{2,2}$, which is 12. Thus we set $x_{1,2} = 12$. This forces $x_{3,2} = 0$ and $x_{4,2} = 0$. The pair $(1, 2)$ is added to the current basis.

The next undetermined cell of lowest cost is $(4, 1)$. We assign to this cell the minimum of $s_4 - x_{4,2}$, which is 13, and $d_1 - x_{2,1}$, which is 19. Thus we set $x_{4,1} = 13$. This forces $x_{4,3} = 0$. The pair $(4, 1)$ is added to the current basis.

The next undetermined cell of lowest cost is $(3, 3)$. We assign to this cell the minimum of $s_3 - x_{3,2}$, which is 11, and $d_3 - x_{2,3} - x_{4,3}$, which is 6 $(= 14 - 8 - 0)$. Thus we set $x_{3,3} = 6$. This forces $x_{1,3} = 0$. The pair $(3, 3)$ is added to the current basis.

The next undetermined cell of lowest cost is $(1, 1)$. We assign to this cell the minimum of $s_1 - x_{1,2} - x_{1,3}$, which is 1, and $d_1 - x_{2,1} - x_{4,1}$, which is 6. Thus we set $x_{1,1} = 1$. The pair $(1, 1)$ is added to the current basis.

The final undetermined cell is $(3, 1)$. We assign to this cell the common value of $s_3 - x_{3,2} - x_{3,3}$ and $d_1 - x_{1,1} - x_{2,1} - x_{4,1}$, which is 5. Thus we set $x_{3,1} = 5$. The pair $(3, 1)$ is added to the current basis.

The values of the elements $x_{i,j}$ of the initial basic feasible solution are tabulated (with basis elements marked by the • symbol) as follows:—

$c_{i,j} \searrow x_{i,j}$	1		2		3		s_i
1	8	•	4	•	16		
		1		12		0	13
2	3		7		2	•	
		0		0		8	8
3	13	•	8		6	•	
		5		0		6	11
4	5	•	7		8		
		13		0		0	13
d_j		19		12		14	45

Thus the initial basis is B where

$$B = \{(1, 1),\ (1, 2),\ (2, 3),\ (3, 1),\ (3, 3),\ (4, 1)\}.$$

The basic feasible solution is represented by the 6×5 matrix X, where

$$X = \begin{pmatrix} 1 & 12 & 0 \\ 0 & 0 & 8 \\ 5 & 0 & 6 \\ 13 & 0 & 0 \end{pmatrix}.$$

The cost of this initial feasible basic solution is

$$\begin{aligned} & 8 \times 1 + 4 \times 12 + 2 \times 8 + 13 \times 5 + 6 \times 6 \\ & \qquad + 5 \times 13 \\ = {} & 8 + 48 + 16 + 65 + 36 + 65 \\ = {} & 238. \end{aligned}$$

3.8 An Example illustrating the Procedure for finding a Basic Optimal Solution to a Transportation Problem

We continue with the study of the optimization problem discussed in the previous section.

We seek to determine non-negative real numbers $x_{i,j}$ for $i = 1, 2, 3, 4$ and $j = 1, 2, 3$ that minimize $\sum_{i=1}^{4} \sum_{j=1}^{3} c_{i,j} x_{i,j}$, where $c_{i,j}$ is the coefficient in the ith row and jth column of the cost matrix C, where

$$C = \begin{pmatrix} 8 & 4 & 16 \\ 3 & 7 & 2 \\ 13 & 8 & 6 \\ 5 & 7 & 8 \end{pmatrix}.$$

subject to the constraints

$$\sum_{j=1}^{3} x_{i,j} = s_i \quad (i = 1, 2, 3, 4)$$

and

$$\sum_{i=1}^{4} x_{i,j} = d_j \quad (j = 1, 2, 3),$$

where

$$s_1 = 13, \quad s_2 = 8, \quad s_3 = 11, \quad s_4 = 13,$$

$$d_1 = 19, \quad d_2 = 12, \quad d_3 = 14.$$

We have found an initial basic feasible solution by the Minimum Cost Method. This solution satisfies $x_{i,j} = (X)_{i,j}$ for all i and j, where

$$X = \begin{pmatrix} 1 & 12 & 0 \\ 0 & 0 & 8 \\ 5 & 0 & 6 \\ 13 & 0 & 0 \end{pmatrix}.$$

We next determine whether this initial basic feasible solution is an optimal solution, and, if not, how to adjust the basis to obtain a solution of lower cost.

We determine u_1, u_2, u_3, u_4 and v_1, v_2, v_3 such that $c_{i,j} = v_j - u_i$ for all $(i, j) \in B$, where B is the initial basis.

We seek a solution with $u_1 = 0$. We then determine $q_{i,j}$ so that $c_{i,j} = v_j - u_i + q_{i,j}$ for all i and j.

We therefore complete the following tableau:—

$c_{i,j} \searrow q_{i,j}$	1	2	3	u_i
1	8 $\bullet$ 0	4 $\bullet$ 0	16 ?	0
2	3 ?	7 ?	2 $\bullet$ 0	?
3	13 $\bullet$ 0	8 ?	6 $\bullet$ 0	?
4	5 $\bullet$ 0	7 ?	8 ?	?
v_j	?	?	?	

Now $u_1 = 0$, $(1,1) \in B$ and $(1,2) \in B$ force $v_1 = 8$ and $v_2 = 4$.
Then $v_1 = 8$, $(3,1) \in B$ and $(4,1) \in B$ force $u_3 = -5$ and $u_4 = 3$.
Then $u_3 = -5$ and $(3,3) \in B$ force $v_3 = 1$.
Then $v_3 = 1$ and $(2,3) \in B$ force $u_2 = -1$.
After entering the numbers u_i and v_j, the tableau is as follows:—

$c_{i,j} \searrow q_{i,j}$	1	2	3	u_i
1	8 $\bullet$ 0	4 $\bullet$ 0	16 ?	0
2	3 ?	7 ?	2 $\bullet$ 0	-1
3	13 $\bullet$ 0	8 ?	6 $\bullet$ 0	-5
4	5 $\bullet$ 0	7 ?	8 ?	3
v_j	8	4	1	

Computing the numbers $q_{i,j}$ such that $c_{i,j} + u_i = v_j + q_{i,j}$, we find that $q_{1,3} = 15$, $q_{2,1} = -6$, $q_{2,2} = 2$, $q_{3,2} = -1$, $q_{4,2} = 6$ and $q_{4,3} = 10$.

The completed tableau is as follows:—

$c_{i,j} \searrow q_{i,j}$	1		2		3		u_i
1	8	•	4	•	16		0
		0		0		15	
2	3		7		2	•	-1
		-6		2		0	
3	13	•	8		6	•	-5
		0		-1		0	
4	5	•	7		8		3
		0		6		10	
v_j	8		4		1		

The initial basic feasible solution is not optimal because some of the quantities $q_{i,j}$ are negative. To see this, suppose that the numbers $\overline{x}_{i,j}$ for $i = 1, 2, 3, 4$ and $j = 1, 2, 3$ constitute a feasible solution to the given problem. Then $\sum_{j=1}^{3} \overline{x}_{i,j} = s_i$ for $i = 1, 2, 3$ and $\sum_{i=1}^{4} \overline{x}_{i,j} = d_j$ for $j = 1, 2, 3, 4$. It follows that

$$\begin{aligned}\sum_{i=1}^{4}\sum_{j=1}^{3} c_{i,j}\overline{x}_{i,j} &= \sum_{i=1}^{4}\sum_{j=1}^{3}(v_j - u_i + q_{i,j})\overline{x}_{i,j} \\ &= \sum_{j=1}^{3} v_j d_j - \sum_{i=1}^{4} u_i s_i + \sum_{i=1}^{4}\sum_{j=1}^{3} q_{i,j}\overline{x}_{i,j}.\end{aligned}$$

Applying this identity to the initial basic feasible solution, we find that $\sum_{j=1}^{3} v_j d_j - \sum_{i=1}^{4} u_i s_i = 238$, given that 238 is the cost of the initial basic feasible solution. Thus the cost $\overline{C}$ of any feasible solution $(\overline{x}_{i,j})$ satisfies

$$\overline{C} = 238 + 15\overline{x}_{1,3} - 6\overline{x}_{2,1} + 2\overline{x}_{2,2} - \overline{x}_{3,2} + 6\overline{x}_{4,2} + 10\overline{x}_{4,3}.$$

One could construct feasible solutions with $\overline{x}_{2,1} < 0$ and $\overline{x}_{i,j} = 0$ for $(i, j) \notin B \cup \{(2, 1)\}$, and the cost of such feasible solutions would be lower than that of the initial basic solution. We therefore seek to bring $(2, 1)$ into the basis, removing some other element of the basis to ensure that the new basis corresponds to a feasible basic solution.

The procedure for achieving this requires us to determine a 4×3 matrix Y satisfying the following conditions:—

- $y_{2,1} = 1$;
- $y_{i,j} = 0$ when $(i, j) \notin B \cup \{(2, 1)\}$;

- all rows and columns of the matrix Y sum to zero.

Accordingly we fill in the following tableau with those coefficients $y_{i,j}$ of the matrix Y that correspond to cells in the current basis (marked with the • symbol), so that all rows sum to zero and all columns sum to zero:—

$y_{i,j}$	1	2	3	
1	? •	? •		0
2	1 ◦		? •	0
3	? •		? •	0
4	? •			0
	0	0	0	0

The constraints that $y_{2,1} = 1$, $y_{i,j} = 0$ when $(i,j) \notin B$ and the constraints requiring the rows and columns to sum to zero determine the values of $y_{i,j}$ for all $y_{i,j} \in B$. These values are recorded in the following tableau:—

$y_{i,j}$	1	2	3	
1	0 •	0 •		0
2	1 ◦		−1 •	0
3	−1 •		1 •	0
4	0 •			0
	0	0	0	0

We now determine those values of λ for which $X+\lambda Y$ is a feasible solution, where

$$X + \lambda Y = \begin{pmatrix} 1 & 12 & 0 \\ \lambda & 0 & 8-\lambda \\ 5-\lambda & 0 & 6+\lambda \\ 13 & 0 & 0 \end{pmatrix}.$$

In order to drive down the cost as far as possible, we should make λ as large as possible, subject to the requirement that all the coefficients of the above matrix should be non-negative numbers. Accordingly we take $\lambda = 5$. Our new basic feasible solution X is then as follows:—

$$X = \begin{pmatrix} 1 & 12 & 0 \\ 5 & 0 & 3 \\ 0 & 0 & 11 \\ 13 & 0 & 0 \end{pmatrix}.$$

We regard X of as the current feasible basic solution.

The cost of the current feasible basic solution X is

$$\begin{aligned} & 8 \times 1 + 4 \times 12 + 3 \times 5 + 2 \times 3 + 6 \times 11 \\ & \quad + 5 \times 13 \\ = & \; 8 + 48 + 15 + 6 + 66 + 65 \\ = & \; 208. \end{aligned}$$

The cost has gone down by 30, as one would expect (the reduction in the cost being $-\lambda q_{2,1}$ where $\lambda = 5$ and $q_{2,1} = -6$).

The current basic feasible solution X is associated with the basis B where

$$B = \{(1,1),\ (1,2),\ (2,1),\ (2,3),\ (3,3),\ (4,1)\}.$$

We now compute, for the current feasible basic solution We determine, for the current basis B values u_1, u_2, u_3, u_4 and v_1, v_2, v_3 such that $c_{i,j} = v_j - u_i$ for all $(i,j) \in B$. the initial basis.

We seek a solution with $u_1 = 0$. We then determine $q_{i,j}$ so that $c_{i,j} = v_j - u_i + q_{i,j}$ for all i and j.

We therefore complete the following tableau:—

$c_{i,j} \searrow q_{i,j}$	1		2		3		u_i
1	8	•	4	•	16		0
		0		0		?	
2	3	•	7		2	•	?
		0		?		0	
3	13		8		6	•	?
		?		?		0	
4	5	•	7		8		?
		0		?		?	
v_j	?		?		?		

Now $u_1 = 0$, $(1,1) \in B$ and $(1,2) \in B$ force $v_1 = 8$ and $v_2 = 4$.

Then $v_1 = 8$, $(2,1) \in B$ and $(4,1) \in B$ force $u_2 = 5$ and $u_4 = 3$.

Then $u_2 = 5$ and $(3,3) \in B$ force $v_3 = 7$.

Then $v_3 = 7$ and $(3,3) \in B$ force $u_3 = 1$.

Computing the numbers $q_{i,j}$ such that $c_{i,j} + u_i = v_j + q_{i,j}$, we find that $q_{1,3} = 9$, $q_{2,2} = 8$, $q_{3,1} = 6$, $q_{3,2} = 5$, $q_{4,2} = 6$ and $q_{4,3} = 4$.

The completed tableau is as follows:—

$c_{i,j} \searrow q_{i,j}$	1		2		3		u_i
1	8	$\bullet$	4	$\bullet$	16		0
		0		0		9	
2	3	$\bullet$	7		2	$\bullet$	5
		0		8		0	
3	13		8		6	$\bullet$	1
		6		5		0	
4	5	$\bullet$	7		8		3
		0		6		4	
v_j	8		4		7		

All numbers $q_{i,j}$ are non-negative for the current feasible basic solution. This solution is therefore optimal. Indeed, arguing as before we find that the cost $\overline{C}$ of any feasible solution $(\overline{x}_{i,j})$ satisfies

$$\overline{C} = 208 + 9\overline{x}_{1,3} + 8\overline{x}_{2,2} + 6\overline{x}_{3,1} + 5\overline{x}_{3,2} + 6\overline{x}_{4,2} + 4\overline{x}_{4,3}.$$

We conclude that X is an basic optimal solution, where

$$X = \begin{pmatrix} 1 & 12 & 0 \\ 5 & 0 & 3 \\ 0 & 0 & 11 \\ 13 & 0 & 0 \end{pmatrix}.$$

3.9 A Result concerning the Construction of Bases for the Transportation Problem

The following general proposition ensures that certain standard methods for determining an initial basic solution of the Transportation Problem, including the *Northwest Corner Method* and the *Minimum Cost Method* will succeed in determining a basic feasible solution to the Transportation Problem.

Proposition 3.7 *Let $I = \{1, 2, \ldots, m\}$ and $J = \{1, 2, \ldots, n\}$, where m and n are positive integers, let $i_1, i_2, \ldots, i_{m+n-1}$ be elements of I and let $j_1, j_2, \ldots, j_{m+n-1}$ be elements of J, and let*

$$B = \{(i_k, j_k) : k = 1, 2, \ldots, m + n - 1\}.$$

Suppose that there exist subsets $I_0, I_1, \ldots, I_{m+n-1}$ of I and $J_0, J_1, \ldots, J_{m+n+1}$ of J such that $I_0 = I$, $J_0 = J$, and such that, for each integer k between 1 and $m + n - 1$, exactly one of the following two conditions is satisfied:—

(i) $i_k \notin I_k$, $j_k \in J_k$, $I_{k-1} = I_k \cup \{i_k\}$ and $J_{k-1} = J_k$;

(ii) $i_k \in I_k$, $j_k \notin J_k$, $I_{k-1} = I_k$ and $J_{k-1} = J_k \cup \{j_k\}$;

Then, given any real numbers $a_1, a_1, \ldots, a_m$ and $b_1, b_2, \ldots, b_n$ satisfying

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j,$$

there exist uniquely-determined real numbers $x_{i,j}$ for all $i \in I$ and $j \in J$ such that $\sum_{j \in J} x_{i,j} = a_i$ for all $i \in I$, $\sum_{i \in I} x_{i,j} = b_j$ for all $j \in J$, and $x_{i,j} = 0$ whenever $(i, j) \notin B$.

Proof We prove the result by induction on $m + n$. The result is easily seen to be true when $m = n = 1$. Thus suppose as our inductive hypothesis that the corresponding results are true when I and J are replaced by I_1 and J_1, so that, given any real numbers a'_i for $i \in I_1$ and b'_j for $j \in J_1$ satisfying $\sum_{i \in I_1} a'_i = \sum_{j \in J_1} b'_j$, there exist uniquely-determined real numbers $x_{i,j}$ for $i \in I_1$ and $j \in J_1$ such that $\sum_{j \in J_1} x_{i,j} = a_i$ for all $i \in I_1$ and $\sum_{i \in I_1} x_{i,j} = b_j$ for all $j \in I_1$. We prove that the corresponding results are true for the given sets I and J.

Now the conditions in the statement of the Proposition ensure that either $i_1 \notin I_1$ or else $j_1 \notin J_1$.

Suppose that $i_1 \notin I_1$. Then $I = I_1 \cup \{i_1\}$ and $J_1 = J$. Now I_k and J_k are subsets of I_1 and J_1 for $k = 1, 2, \ldots, m+n-1$. Moreover $(i_k, j_k) \in I_{k-1} \times J_{k-1}$

for all integers k satisfying $1 \leq k \leq m+n+1$. It follows that $i_k \in I_1$ and therefore $i_k \neq i_1$ whenever $2 \leq k \leq m$. It follows that the conclusions of the proposition are true if and only if there exist uniquely-determined real numbers $x_{i,j}$ for $i \in I$ and $j \in I$ such that

$$\begin{aligned}
x_{i_1,j_1} &= a_{i_1}, \\
x_{i_1,j} &= 0 \text{ whenever } j \neq j_1, \\
\sum_{j \in J} x_{i,j} &= a_i \text{ whenever } i \neq i_1, \\
\sum_{i \in I_1} x_{i,j_1} &= b_{j_1} - a_{i_1}, \\
\sum_{i \in I_1} x_{i,j} &= b_j \text{ whenever } j \neq j_1, \\
x_{i,j} &= 0 \text{ whenever } (i,j) \notin B
\end{aligned}$$

The induction hypothesis ensures the existence and uniqueness of the real numbers $x_{i,j}$ for $i \in I_1$ and $j \in J$ determined so as to satisfy the above conditions. Thus the induction hypothesis ensures that the required result is true in the case where $i_1 \notin I_1$.

An analogous argument shows that the required result is true in the case where $j_1 \notin J_1$. The result follows. ∎

Proposition 3.7 ensures that if $I = \{1, 2, \ldots, m\}$ and $J = \{1, 2, \ldots, n\}$ and if a subset B of $I \times J$ is determined so as to satisfy the requirements of Proposition 3.7, then that subset B of $I \times J$ is a basis for the Transportation Problem with m suppliers and n recipients.

The algorithms underlying the *Minimal Cost Method* and the *Northwest Corner Method* give rise to subsets I_k and J_k of I and J respectively for $k = 0, 1, 2, \ldots, m+n-1$ that satisfy the conditions of Proposition 3.7. This proposition therefore ensures that *Minimal Cost Method* and the *Northwest Corner Method* do indeed determine basic feasible solutions to the Transportation Problem.

Remark One can prove a converse result to Proposition 3.7 which establishes that, given any basis B for an instance of the Transportation Problem with m suppliers and n recipients, there exist subsets I_k of I and J_k of J, for $i = 1, 2, \ldots, m+n-1$, where $I = \{1, 2, \ldots, m\}$ and $J = \{1, 2, \ldots, n\}$, so that these subsets I_k and J_k of I and J are related to one another and to the basis B in the manner described in the statement of Proposition 3.7.

3.10 The Minimum Cost Method

We describe the *Minimum Cost Method* for finding an initial basic feasible solution to the Transportation Problem.

Consider an instance of the Transportation Problem specified by positive integers m and n and non-negative real numbers $s_1, s_2, \ldots, s_m$ and $d_1, d_2, \ldots, d_n$, where $\sum_{i=1}^{m} s_i = \sum_{j=1}^{n} d_j$. Let $I = \{1, 2, \ldots, m\}$ and let $J = \{1, 2, \ldots, n\}$. A feasible solution consists of an array of non-negative real numbers $x_{i,j}$ for $i \in I$ and $j \in J$ with the property that $\sum_{j \in J} x_{i,j} = s_i$ for all $i \in I$ and $\sum_{i \in I} x_{i,j} = d_j$ for all $j \in J$. The objective of the problem is to find a feasible solution that minimizes cost, where the cost of a feasible solution $(x_{i,j} : i \in I \text{ and} j \in J)$ is $\sum_{i \in I} \sum_{j \in J} c_{i,j} x_{i,j}$.

In applying the Minimal Cost Method to find an initial basic solution to the Transportation we apply an algorithm that corresponds to the determination of elements $(i_1, j_1), (i_2, j_2), \ldots, (i_{m+n-1}, j_{m+n-1})$ of $I \times J$ and of subsets $I_0, I_1, \ldots, I_{m+n-1}$ of I and $J_0, J_1, \ldots, J_{m+n-1}$ of J such that the conditions of Proposition 3.7 are satisfied.

Indeed let $I_0 = I$, $J_0 = J$ and $B_0 = \{0\}$. The Minimal Cost Method algorithm is accomplished in $m + n - 1$ stages.

Let k be an integer satisfying $1 \leq k \leq m + n - 1$ and that subsets I_{k-1} of I, J_{k-1} of J and B_{k-1} of $I \times J$ have been determined in accordance with the rules that apply at previous stages of the Minimal Cost algorithm. Suppose also that non-negative real numbers $x_{i,j}$ have been determined for all ordered pairs (i, j) in $I \times J$ that satisfy either $i \notin I_{k-1}$ or $j \notin J_{k-1}$ so as to satisfy the following conditions:—

- $\sum_{j \in J \setminus J_{k-1}} x_{i,j} \leq s_i$ whenever $i \in I_{k-1}$;
- $\sum_{j \in J} x_{i,j} = s_i$ whenever $i \notin I_{k-1}$;
- $\sum_{i \in I \setminus I_{k-1}} x_{i,j} \leq d_j$ whenever $j \in J_{k-1}$;
- $\sum_{i \in I} x_{i,j} = d_j$ whenever $j \notin J_{k-1}$.

The Minimal Cost Method specifies that one should choose $(i_k, j_k) \in I_{k-1} \times J_{k-1}$ so that

$$c_{i_k, j_k} \leq c_{i,j} \quad \text{for all } (i, j) \in I_{k-1} \times J_{k-1},$$

and set $B_k = B_{k-1} \cup \{(i_k, j_k)\}$. Having chosen (i_k, j_k), the non-negative real number x_{i_k,j_k} is then determined so that

$$x_{i_k,j_k} = \min \left(s_{i_k} - \sum_{j \in J \setminus J_{k-1}} x_{i_k,j}, \;\; d_{j_k} - \sum_{i \in I \setminus I_{k-1}} x_{i,j_k} \right).$$

The subsets I_k and J_k of I and J respectively are then determined, along with appropriate values of $x_{i,j}$, according to the following rules:—

(i) if

$$s_{i_k} - \sum_{j \in J \setminus J_{k-1}} x_{i_k,j} < d_{j_k} - \sum_{i \in I \setminus I_{k-1}} x_{i,j_k}$$

then we set $I_k = I_{k-1} \setminus \{i_k\}$ and $J_k = J_{k-1}$, and we also let $x_{i_k,j} = 0$ for all $j \in J_{k-1} \setminus \{j_k\}$;

(ii) if

$$s_{i_k} - \sum_{j \in J \setminus J_{k-1}} x_{i_k,j} > d_{j_k} - \sum_{i \in I \setminus I_{k-1}} x_{i,j_k}$$

then we set $J_k = J_{k-1} \setminus \{j_k\}$ and $I_k = I_{k-1}$, and we also let $x_{i,j_k} = 0$ for all $i \in I_{k-1} \setminus \{i_k\}$;

(iii) if

$$s_{i_k} - \sum_{j \in J \setminus J_{k-1}} x_{i_k,j} = d_{j_k} - \sum_{i \in I \setminus I_{k-1}} x_{i,j_k}$$

then we determine I_k and J_k and the corresponding values of $x_{i,j}$ either in accordance with the specification in rule (i) above or else in accordance with the specification in rule (ii) above.

These rules ensure that the real numbers $x_{i,j}$ determined at this stage are all non-negative, and that the following conditions are satisfied at the conclusion of the kth stage of the Minimal Cost Method algorithm:—

- $\sum_{j \in J \setminus J_k} x_{i,j} \leq s_i$ whenever $i \in I_k$;
- $\sum_{j \in J} x_{i,j} = s_i$ whenever $i \notin I_k$;
- $\sum_{i \in I \setminus I_k} x_{i,j} \leq d_j$ whenever $j \in J_k$;
- $\sum_{i \in I} x_{i,j} = d_j$ whenever $j \notin J_k$.

At the completion of the final stage (for which $k = m + n - 1$) we have determined a subset B of $I \times J$, where $B = B_{m+n-1}$, together with non-negative real numbers $x_{i,j}$ for $i \in I$ and $j \in I$ that constitute a feasible solution to the given instance of the Transportation Problem. Moreover Proposition 3.7 ensures that this feasible solution is a basic feasible solution of the problem with associated basis B.

3.11 The Northwest Corner Method

The *Northwest Corner Method* for finding a basic feasible solution proceeds according to the stages of the *Minimum Cost Method* above, differing only from that method in the choice of the ordered pair (i_k, j_k) at the kth stage of the method. In the Minimum Cost Method, the ordered pair (i_k, j_k) is chosen such that $(i_k, j_k) \in I_{k-1} \times J_{k-1}$ and

$$c_{i_k,j_k} \leq c_{i,j} \quad \text{for all } (i,j) \in I_{k-1} \times J_{k-1}$$

(where the sets I_{k-1}, J_{k-1} are determined as in the specification of the Minimum Cost Method). In applying the Northwest Corner Method, costs associated with ordered pairs (i,j) in $I \times J$ are not taken into account. Instead (i_k, j_k) is chosen so that i_k is the minimum of the integers in I_{k-1} and j_k is the minimum of the integers in J_{k-1}. Otherwise the specification of the Northwest Corner Method corresponds to that of the Minimum Cost Method, and results in a basic feasible solution of the given instance of the Transportation Problem.

3.12 The Iterative Procedure for Solving the Transportation Problem, given an Initial Basic Feasible Solution

We now describe in general terms the method for solving the Transportation Problem, in the case where total supply equals total demand.

We suppose that an initial basic feasible solution has been obtained. We apply an iterative method (based on the general Simplex Method for the solution of linear programming problems) that will test a basic feasible solution for optimality and, in the event that the feasible solution is shown not to be optimal, establishes information that (with the exception of certain 'degenerate' cases of the Transportation Problem) enables one to find a basic feasible solution with lower cost. Iterating this procedure a finite number of times, one should arrive at a basic feasible solution that is optimal for the given instance of the the Transportation Problem.

We suppose that the given instance of the Transportation Problem involves m suppliers and n recipients. The required supplies are specified by non-negative real numbers $s_1, s_2, \ldots, s_m$, and the required demands are specified by non-negative real numbers $d_1, d_2, \ldots, d_n$. We further suppose that $\sum_{i=1}^{m} s_i = \sum_{j=1}^{n} d_j$. A *feasible solution* is represented by non-negative real numbers $x_{i,j}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, where $\sum_{j=1}^{n} x_{i,j} = s_i$ for $i = 1, 2, \ldots, m$ and $\sum_{i=1}^{m} x_{i,j} = d_j$ for $j = 1, 2, \ldots, n$.

Let $I = \{1, 2, \ldots, m\}$ and $J = \{1, 2, \ldots, n\}$. A subset B of $I \times J$ is a *basis* for the Transportation Problem if and only if, given any real numbers $y_1, y_2, \ldots, y_m$ and $z_1, z_2, \ldots, z_n$, there exist uniquely determined real numbers $\overline{x}_{i,j}$ for $i \in I$ and $j \in J$ such that $\sum_{j=1}^{n} \overline{x}_{i,j} = y_i$ for $i \in I$, $\sum_{i=1}^{m} \overline{x}_{i,j} = z_j$ for $j \in J$, where $\overline{x}_{i,j} = 0$ whenever $(i, j) \notin B$ (see Proposition 3.4).

A feasible solution $(x_{i,j})$ is said to be a basic feasible solution associated with the basis B if and only if $x_{i,j} = 0$ for all $i \in I$ and $j \in J$ for which $(i, j) \notin B$.

Let $x_{i,j}$ be a non-negative real number for each $i \in I$ and $j \in J$. Suppose that $(x_{i,j})$ is a basic feasible solution to the Transportation Problem associated with basis B, where $B \subset I \times J$.

The cost associated with a feasible solution $(x_{i,j}$ is given by $\sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j} x_{i,j}$, where the constants $c_{i,j}$ are real numbers for all $i \in I$ and $j \in J$. A feasible solution for the given instance of the Transportation Problems is an optimal solution if and only if it minimizes cost amongst all feasible solutions to the problem.

In order to test for optimality of a basic feasible solution $(x_{i,j})$ associated with a basis B, we determine real numbers $u_1, u_2, \ldots, u_m$ and $v_1, v_2, \ldots, v_n$ with the property that $c_{i,j} = v_j - u_i$ for all $(i, j) \in B$. (Proposition 3.10 below guarantees that, given any basis B, it is always possible to find the required quantities u_i and v_j.) Having calculated these quantities u_i and v_j we determine the values of $q_{i,j}$, where $q_{i,j} = c_{i,j} - v_j + u_i$ for all $i \in I$ and $j \in J$. Then $q_{i,j} = 0$ whenever $(i, j) \in B$.

We claim that a basic feasible solution $(x_{i,j})$ associated with the basis B is optimal if and only if $q_{i,j} \geq 0$ for all $i \in I$ and $j \in J$. This is a consequence of the identity established in the following proposition.

Proposition 3.8 *Let $x_{i,j}$, $c_{i,j}$ and $q_{i,j}$ be real numbers defined for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, and let $u_1, u_2, \ldots, u_m$ and $v_1, v_2, \ldots, v_n$ be real numbers.*

Suppose that

$$c_{i,j} = v_j - u_i + q_{i,j}$$

for $i = 1, 2, \ldots, m$ *and* $j = 1, 2, \ldots, n$. *Then*

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{i,j}x_{i,j} = \sum_{i=1}^{m} v_j d_j - \sum_{j=1}^{n} u_i s_i + \sum_{i=1}^{m}\sum_{j=1}^{n} q_{i,j}x_{i,j},$$

where $s_i = \sum_{j=1}^{n} x_{i,j}$ *for* $i = 1, 2, \ldots, m$ *and* $d_j = \sum_{i=1}^{m} x_{i,j}$ *for* $j = 1, 2, \ldots, n$.

Proof The definitions of the relevant quantities ensure that

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n} c_{i,j}x_{i,j} &= \sum_{i=1}^{m}\sum_{j=1}^{n}(v_j - u_i + q_{i,j})x_{i,j} \\ &= \sum_{j=1}^{n}\left(v_j \sum_{i=1}^{m} x_{i,j}\right) - \sum_{i=1}^{m}\left(u_i \sum_{j=1}^{n} x_{i,j}\right) \\ &\qquad + \sum_{i=1}^{m}\sum_{j=1}^{n} q_{i,j}x_{i,j} \\ &= \sum_{i=1}^{m} v_j d_j - \sum_{j=1}^{n} u_i s_i + \sum_{i=1}^{m}\sum_{j=1}^{n} q_{i,j}x_{i,j},\end{aligned}$$

as required. ∎

Corollary 3.9 *Let* m *and* n *be integers, and let* $I = \{1, 2, \ldots, m\}$ *and* $J = \{1, 2, \ldots, n\}$. *Let* $x_{i,j}$ *and* $c_{i,j}$ *be real numbers defined for all* $i \in I$ *and* $j \in I$, *and let* $u_1, u_2, \ldots, u_m$ *and* $v_1, v_2, \ldots, v_n$ *be real numbers. Suppose that* $c_{i,j} = v_j - u_i$ *for all* $(i, j) \in I \times J$ *for which* $x_{i,j} \neq 0$. *Then*

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{i,j}x_{i,j} = \sum_{i=1}^{m} d_j v_j - \sum_{j=1}^{n} s_i u_i,$$

where $s_i = \sum_{j=1}^{n} x_{i,j}$ *for* $i = 1, 2, \ldots, m$ *and* $d_j = \sum_{i=1}^{m} x_{i,j}$ *for* $j = 1, 2, \ldots, n$.

Proof Let $q_{i,j} = c_{i,j} + u_i - v_j$ for all $i \in I$ and $j \in J$. Then $q_{i,j} = 0$ whenever $x_{i,j} \neq 0$. It follows from this that

$$\sum_{i=1}^{m}\sum_{j=1}^{n} q_{i,j}x_{i,j} = 0.$$

It then follows from Proposition 3.8 that

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{i,j}x_{i,j} = \sum_{i=1}^{m}\sum_{j=1}^{n}(v_j - u_i + q_{i,j})x_{i,j} = \sum_{i=1}^{m} d_j v_j - \sum_{j=1}^{n} s_i u_i,$$

as required. ∎

Let m and n be positive integers, let $I = \{1, 2, \ldots, m\}$ and $J = \{1, 2, \ldots, n\}$, and let the subset B of $I \times J$ be a basis for an instance of the Transportation Problem with m suppliers and n recipients. Let the cost of a feasible solution $(\overline{x}_{i,j})$ be $\sum_{i=1}^{m}\sum_{j=1}^{n} c_{i,j}\overline{x}_{i,j}$. Now $\sum_{j=1}^{n} \overline{x}_{i,j} = s_i$ and $\sum_{i=1}^{m} \overline{x}_{i,j} = d_j$, where the quantities s_i and d_j are determined by the specification of the problem and are the same for all feasible solutions of the problem. Let quantities u_i for $i \in I$ and v_j for $j \in J$ be determined such that $c_{i,j} = v_j - u_i$ for all $(i, j) \in B$, and let $q_{i,j} = c_{i,j} + u_i - v_j$ for all $i \in I$ and $j \in J$. Then $q_{i,j} = 0$ for all $(i, j) \in B$.

It follows from Proposition 3.8 that

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{i,j}\overline{x}_{i,j} = \sum_{i=1}^{m} v_j d_j - \sum_{j=1}^{n} u_i s_i + \sum_{i=1}^{m}\sum_{j=1}^{n} q_{i,j}\overline{x}_{i,j}.$$

Now if the quantities $x_{i,j}$ for $i \in I$ and $j \in J$ constitute a basic feasible solution associated with the basis B then $x_{i,j} = 0$ whenever $(i, j) \notin B$. It follows that $\sum_{i=1}^{m}\sum_{j=1}^{n} q_{i,j}\overline{x}_{i,j} = 0$, and therefore

$$\sum_{i=1}^{m} v_j d_j - \sum_{j=1}^{n} u_i s_i = C,$$

where

$$C = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{i,j}\overline{x}_{i,j}.$$

The cost $\overline{C}$ of the feasible solution $(\overline{x}_{i,j})$ then satisfies the equation

$$\overline{C} = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{i,j}\overline{x}_{i,j} = C + \sum_{i=1}^{m}\sum_{j=1}^{n} q_{i,j}\overline{x}_{i,j}.$$

If $q_{i,j} \geq 0$ for all $i \in I$ and $j \in J$, then the cost $\overline{C}$ of any feasible solution $(\overline{x}_{i,j})$ is bounded below by the cost of the basic feasible solution $(x_{i,j})$. It follows that, in this case, the basic feasible solution $(x_{i,j})$ is optimal.

Suppose that (i_0, j_0) is an element of $I \times J$ for which $q_{i_0,j_0} < 0$. Then $(i_0, j_0) \notin B$. There is no basis for the Transportation Problem that includes the set $B \cup \{(i_0, j_0)\}$. A straightforward application of Lemma 3.5 establishes the existence of quantities $y_{i,j}$ for $i \in I$ and $j \in J$ such that $y_{i_0,j_0} = 1$ and $y_{i,j} = 0$ for all $i \in I$ and $j \in J$ for which $(i, j) \notin B \cup \{(i_0, j_0)\}$.

Let the $m \times n$ matrices X and Y be defined so that $(X)_{i,j} = x_{i,j}$ and $(Y)_{i,j} = y_{i,j}$ for all $i \in I$ and $j \in J$. Suppose that $x_{i,j} > 0$ for all $(i, j) \in B$. Then the components of X in the basis positions are strictly positive. It follows that, if λ is positive but sufficiently small, then the components of the matrix $X + \lambda Y$ in the basis positions are also strictly positive, and therefore the components of the matrix $X + \lambda Y$ are non-negative for all sufficiently small non-negative values of λ. There will then exist a maximum value λ_0 that is an upper bound on the values of λ for which all components of the matrix $X + \lambda Y$ are non-negative. It is then a straightforward exercise in linear algebra to verify that $X + \lambda_0 Y$ is another basic feasible solution associated with a basis that includes (i_0, j_0) together with all but one of the elements of the basis B. Moreover the cost of this new basic feasible solution is $C + \lambda_0 q_{i_0,j_0}$, where C is the cost of the basic feasible solution represented by the matrix X. Thus if $q_{i_0,j_0} < 0$ then the cost of the new basic feasible solution is lower than that of the basic feasible solution X from which it was derived.

Suppose that, for all basic feasible solutions of the given Transportation problem, the coefficients of the matrix specifying the basic feasible solution are strictly positive at the basis positions. Then a finite number of iterations of the procedure discussed above with result in an basic optimal solution of the given instance of the Transportation Problem. Such problems are said to be *non-degenerate.*

However if it turns out that a basic feasible solution $(x_{i,j})$ associated with a basis B satisfies $x_{i,j} = 0$ for some $(i, j) \in B$, then we are in a *degenerate* case of the Transportation Problem. The theory of degenerate cases of linear programming problems is discussed in detail in textbooks that discuss the details of linear programming algorithms.

We now establish the proposition that guarantees that, given any basis B, there exist quantities $u_1, u_2, \ldots, u_m$ and $v_1, v_2, \ldots, v_n$ such that the costs $c_{i,j}$ associated with the given instance of the Transportation Problem satisfy $c_{i,j} = v_j - u_i$ for all $(i, j) \in B$. This result is an essential component of the method described here for testing basic feasible solutions to determine whether or not they are optimal.

Proposition 3.10 *Let m and n be integers, let $I = \{1, 2, \ldots, m\}$ and $J = \{1, 2, \ldots, n\}$, and let B be a subset of $I \times J$ that is a basis for the transporta-*

tion problem with m suppliers and n recipients. For each $(i,j) \in B$ let $c_{i,j}$ be a corresponding real number. Then there exist real numbers u_i for $i \in I$ and v_j for $j \in J$ such that $c_{i,j} = v_j - u_i$ for all $(i,j) \in B$. Moreover if $\overline{u}_i$ and $\overline{v}_j$ are real numbers for $i \in I$ and $j \in J$ that satisfy the equations $c_{i,j} = \overline{v}_j - \overline{u}_i$ for all $(i,j) \in B$, then there exists some real number k such that $\overline{u}_i = u_i + k$ for all $i \in I$ and $\overline{v}_j = v_j + k$ for all $j \in J$.

Proof Let

$$W = \left\{ (\mathbf{y}, \mathbf{z}) \in \mathbb{R}^m \times \mathbb{R}^n : \sum_{i=1}^{m} (\mathbf{y})_i = \sum_{j=1}^{n} (\mathbf{z})_j \right\},$$

let $\rho\colon M_{m,n}(\mathbb{R}) \to \mathbb{R}^m$ and $\sigma\colon M_{m,n}(\mathbb{R}) \to \mathbb{R}^n$ be the linear transformations defined such that $\rho(X)_i = \sum_{j=1}^{n} (X)_{i,j}$ for $i = 1, 2, \ldots, m$ and $\sigma(X)_j = \sum_{i=1}^{m} (X)_{i,j}$ for $j = 1, 2, \ldots, n$, let

$$M_B = \{X \in M_{m,n}(\mathbb{R}) : (X)_{i,j} = 0 \text{ whenever } (i,j) \notin B\},$$

and let C be the $m \times n$ matrix defined such that $(C)_{i,j} = c_{i,j}$ for all $i \in I$ and $j \in J$.

Now, given any element $(\mathbf{y}, \mathbf{z})$ of W, there exists a uniquely-determined $m \times n$ matrix X such that $\sum_{j=1}^{n} (X)_{i,j} = (\mathbf{y})_i$ for $i = 1, 2, \ldots, m$, $\sum_{i=1}^{m} (X)_{i,j} = (\mathbf{z})_j$ for $j = 1, 2, \ldots, n$ and $(X)_{i,j} = 0$ unless $(i,j) \in B$ (see Proposition 3.4). Then X is the unique matrix belonging to M_B that satisfies $\rho(X) = \mathbf{y}$ and $\sigma(X) = \mathbf{z}$. We define

$$g(\mathbf{y}, \mathbf{z}) = \operatorname{trace}(C^T X) = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j} (X)_{i,j}.$$

We obtain in this way a well-defined function $g\colon W \to \mathbb{R}$ characterized by the property that

$$g(\rho(X), \sigma(X)) = \operatorname{trace}(C^T X)$$

for all $X \in M_B$. Now $\rho(\lambda X) = \lambda \mathbf{y}$ and $\sigma(\lambda X) = \lambda \mathbf{z}$ for all real numbers λ. It follows that

$$g(\lambda(\mathbf{y}, \mathbf{z})) = \lambda g(\mathbf{y}, \mathbf{z})$$

for all real numbers λ. Also, given elements $(\mathbf{y}', \mathbf{z}')$ and $(\mathbf{y}'', \mathbf{z}'')$ of W, there exist unique matrices X' and X'' belonging to M_B such that $\rho(X') = \mathbf{y}'$,

$\rho(X'') = \mathbf{y}''$, $\sigma(X') = \mathbf{z}'$ and $\sigma(X'') = \mathbf{z}''$. Then $\rho(X' + X'') = \mathbf{y}' + \mathbf{y}''$ and $\sigma(X' + X'') = \mathbf{z}' + \mathbf{z}''$, and therefore

$$\begin{aligned} g((\mathbf{y}', \mathbf{z}') + (\mathbf{y}'', \mathbf{z}'')) &= \mathrm{trace}(C^T (X' + X'')) \\ &= \mathrm{trace}(C^T X') + \mathrm{trace}(C^T X'') \\ &= g(\mathbf{y}', \mathbf{z}') + g(\mathbf{y}'', \mathbf{z}''). \end{aligned}$$

It follows that the function $g: W \to \mathbb{R}$ is a linear transformation. It is thus a linear functional on the real vector space W.

For each integer i between 1 and m, let $\overline{\mathbf{b}}^{(i)}$ denote the vector in $\mathbb{R}^m$ whose ith component is equal to 1 and whose other components are zero, and, for each integer j between 1 and n, let $\mathbf{b}^{(j)}$ denote the vector in $\mathbb{R}^n$ whose jth component is equal to 1 and whose other components are zero. Then $(\overline{\mathbf{b}}^{(1)} - \overline{\mathbf{b}}^{(i)}, \mathbf{0}) \in W$ for $i = 1, 2, \ldots, m$ and $(\overline{\mathbf{b}}^{(1)}, \mathbf{b}^{(j)}) \in W$ for $j = 1, 2, \ldots, n$. We define $u_i = g(\overline{\mathbf{b}}^{(1)} - \overline{\mathbf{b}}^{(i)}, \mathbf{0})$ for $i = 1, 2, \ldots, m$ and $v_j = g(\overline{\mathbf{b}}^{(1)}, \mathbf{b}^{(j)})$ for $j = 1, 2, \ldots, n$. Then $u_1 = 0$ and

$$\begin{aligned} v_j - u_i &= g(\overline{\mathbf{b}}^{(1)}, \mathbf{b}^{(j)}) - g(\overline{\mathbf{b}}^{(1)} - \overline{\mathbf{b}}^{(i)}, \mathbf{0}) \\ &= g\left((\overline{\mathbf{b}}^{(1)}, \mathbf{b}^{(j)}) - (\overline{\mathbf{b}}^{(1)} - \overline{\mathbf{b}}^{(i)}, \mathbf{0})\right) \\ &= g(\overline{\mathbf{b}}^{(i)}, \mathbf{b}^{(j)}) \end{aligned}$$

for all $i \in I$ and $j \in J$.

If $(i, j) \in B$ then $\overline{\mathbf{b}}^{(i)} = \rho(E^{(i,j)})$ and $\mathbf{b}^{(j)} = \sigma(E^{(i,j)})$, where $E^{(i,j)}$ is the $m \times n$ matrix whose coefficient in the ith row and jth column is equal to 1 and whose other coefficients are zero. Moreover $E^{(i,j)} \in M_B$ for all $(i, j) \in B$. It follows from the definition of the linear functional g that

$$v_j - u_i = g(\overline{\mathbf{b}}^{(i)}, \mathbf{b}^{(j)}) = \mathrm{trace}(C^T E^{(i,j)}) = c_{i,j}$$

for all $(i, j) \in B$.

Now let $\overline{u}_i$ and $\overline{v}_j$ be real numbers for $i \in I$ and $j \in J$ that satisfy the equations $c_{i,j} = \overline{v}_j - \overline{u}_i$ for all $(i, j) \in B$. Let

$$\overline{g}(\mathbf{y}, \mathbf{z}) = \sum_{j=1}^{n} v_j (\mathbf{z})_j - \sum_{i=1}^{m} u_i (\mathbf{y})_i$$

for all $\mathbf{y} \in \mathbb{R}^m$ and $\mathbf{z} \in \mathbb{R}^n$. Then

$$\overline{g}(\rho(E^{(i,j)}), \sigma(E^{(i,j)})) = \overline{g}(\overline{\mathbf{b}}^{(i)}, \mathbf{b}^{(j)}) = \overline{v}_j - \overline{u}_i$$

for all $(i, j) \in I \times J$. It follows that

$$\overline{g}(\rho(E^{(i,j)}), \sigma(E^{(i,j)})) = \overline{v}_j - \overline{u}_i = c_{i,j} = g(\rho(E^{(i,j)}), \sigma(E^{(i,j)})) \quad \text{for all } (i, j) \in B.$$

Now the matrices $E^{(i,j)}$ for all $(i,j) \in B$ constitute a basis of the vector space M_K. It follows that

$$\overline{g}(\rho(X), \sigma(X)) = g(\rho(X), \sigma(X))$$

for all $X \in M_B$. But every element of the vector space W is of the form $(\rho(X), \sigma(X))$ for some $X \in M_B$. (This follows Proposition 3.4, as discussed earlier in the proof.) Thus

$$\overline{g}(\mathbf{y}, \mathbf{z}) = g(\mathbf{y}, \mathbf{z})$$

for all $(\mathbf{y}, \mathbf{z}) \in W$. In particular

$$\overline{u}_i - \overline{u}_1 = \overline{g}(\overline{\mathbf{b}}^{(1)} - \overline{\mathbf{b}}^{(i)}, \mathbf{0}) = g(\overline{\mathbf{b}}^{(1)} - \overline{\mathbf{b}}^{(i)}, \mathbf{0}) = u_i - u_1$$

for all $i \in I$, and

$$\overline{v}_j - \overline{u}_1 = \overline{g}(\overline{\mathbf{b}}^{(1)}, \mathbf{b}^{(j)}) = g(\overline{\mathbf{b}}^{(1)}, \mathbf{b}^{(j)}) = v_j - u_1$$

for all $j \in J$. Let $k = \overline{u}_1 - u_1$. Then $\overline{u}_i = u_i + k$ for all $i \in I$ and $\overline{v}_j = v_j + k$ for all $j \in J$, as required. ∎

4 The Simplex Method

4.1 Vector Inequalities and Notational Conventions

Let $\mathbf{v}$ be an element of the real vector space $\mathbb{R}^n$. We denote by $(\mathbf{v})_j$ the jth component of the vector $\mathbf{v}$. The vector $\mathbf{v}$ can be represented in the usual fashion as an n-tuple $(v_1, v_2, \ldots, v_n)$, where $v_j = (\mathbf{v})_j$ for $j = 1, 2, \ldots, n$. However where an n-dimensional vector appears in matrix equations it will usually be considered to be an $n \times 1$ column vector. The row vector corresponding to an element $\mathbf{v}$ of $\mathbb{R}^n$ will be denoted by $\mathbf{v}^T$ because, considered as a matrix, it is the transpose of the column vector representing $\mathbf{v}$. We denote the zero vector (in the appropriate dimension) by $\mathbf{0}$.

Let $\mathbf{x}$ and $\mathbf{y}$ be vectors belonging to the real vector space $\mathbb{R}^n$ for some positive integer n. We write $\mathbf{x} \leq \mathbf{y}$ (and $\mathbf{y} \geq \mathbf{x}$) when $(\mathbf{x})_j \leq (\mathbf{y})_j$ for $j = 1, 2, \ldots, n$. Also we write $\mathbf{x} \ll \mathbf{y}$ (and $\mathbf{y} \gg \mathbf{x}$) when $(\mathbf{x})_j < (\mathbf{y})_j$ for $j = 1, 2, \ldots, n$.

These notational conventions ensure that $\mathbf{x} \geq \mathbf{0}$ if and only if $(\mathbf{x})_j \geq 0$ for $j = 1, 2, \ldots, n$.

The scalar product of two n-dimensional vectors $\mathbf{u}$ and $\mathbf{v}$ can be represented as the matrix product $\mathbf{u}^T\mathbf{v}$. Thus

$$\mathbf{u}^T\mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, where $u_j = (\mathbf{u})_j$ and $v_j = (\mathbf{v})_j$ for $j = 1, 2, \ldots, n$.

Given an $m \times n$ matrix A, where m and n are positive integers, we denote by $(A)_{i,j}$ the coefficient in the ith row and jth column of the matrix A.

4.2 Feasible and Optimal Solutions of General Linear Programming Problems

A general linear programming problem is one that seeks values of real variables $x_1, x_2, \ldots, x_n$ that maximize or minimize some *objective function*

$$c_1 x_1 + c_2 x_2 + \cdots c_n x_n$$

that is a linear functional of $x_1, x_2, \ldots, x_n$ determined by real constants $c_1, c_2, \ldots, c_n$, where the variables $x_1, x_2, \ldots, x_n$ are subject to a finite number of *constraints* that each place bounds on the value of some linear functional of the variables. These constraints can then be numbered from 1 to m, for an appropriate value of m such that, for each value of i between 1 and m, the ith constraint takes the form of an equation or inequality that can be expressed in one of the following three forms:—

$$a_{i,1} x_1 + a_{i,2} x_2 + \cdots + a_{i,n} x_n = b_i,$$

$$a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n \geq b_i,$$

$$a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n \leq b_i$$

for appropriate values of the real constants $a_{i,1}, a_{i,2}, \ldots, a_{i,n}$ and b_i. In addition some, but not necessarily all, of the variables $x_1, x_2, \ldots, x_n$ may be required to be non-negative. (Of course a constraint requiring a variable to be non-negative can be expressed by an inequality that conforms to one of the three forms described above. Nevertheless constraints that simply require some of the variables to be non-negative are usually listed separately from the other constraints.)

Definition Consider a general linear programming problem with n real variables $x_1, x_2, \ldots, x_n$ whose objective is to maximize or minimize some objective function subject to appropriate constraints. A *feasible solution* of this linear programming problem is specified by an n-dimensional vector $\mathbf{x}$ whose components satisfy the constraints but do not necessarily maximize or minimize the objective function.

Definition Consider a general linear programming problem with n real variables $x_1, x_2, \ldots, x_n$ whose objective is to maximize or minimize some objective function subject to appropriate constraints. A *optimal solution* of this linear programming problem is specified by an n-dimensional vector $\mathbf{x}$ that is a feasible solution that optimizes the value of the objective function amongst all feasible solutions to the linear programming problem.

4.3 Linear Programming Problems in Dantzig Standard Form

Let A be an $m \times n$ matrix of rank m with real coefficients, where $m \leq n$, and let $\mathbf{b} \in \mathbb{R}^m$ and $\mathbf{c} \in \mathbb{R}^n$ be vectors of dimensions m and n respectively. We consider the following linear programming problem:—

> *Determine an n-dimensional vector $\mathbf{x}$ so as to minimize $\mathbf{c}^T\mathbf{x}$ subject to the constraints $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$.*

We refer to linear programming problems presented in this form as being in *Dantzig standard form.* We refer to the $m \times n$ matrix A, the m-dimensional vector $\mathbf{b}$ and the n-dimensional vector $\mathbf{c}$ as the *constraint matrix*, *target vector* and *cost vector* for the linear programming problem.

Remark Nomenclature in Linear Programming textbooks varies. Problems presented in the above form are those to which the basic algorithms of George

B. Dantzig's *Simplex Method* are applicable. In the series of textbooks by George B. Dantzig and Mukund N. Thapa entitled *Linear Programming*, such problems are said to be in *standard form*. In the textbook *Introduction to Linear Programming* by Richard B. Darst, such problems are said to be *standard-form LP*. On the other hand, in the textbook *Methods of Mathematical Economics* by Joel N. Franklin, such problems are said to be in *canonical form*, and the term *standard form* is used for problems which match the form above, except that the vector equality $A\mathbf{x} = \mathbf{b}$ is replaced by a vector inequality $A\mathbf{x} \geq \mathbf{b}$. Accordingly the term *Danztig standard form* is used in these notes both to indicate that such problems are in *standard form* at that term is used by textbooks of which Dantzig is the author, and also to emphasize the connection with the contribution of Dantzig in creating and popularizing the *Simplex Method* for the solution of linear programming problems.

A linear programming problem in Dantzig standard form specified by an $m \times n$ constraint matrix A of rank m, an m-dimensional target vector $\mathbf{b}$ and an n-dimensional cost vector $\mathbf{c}$ has the objective of finding values of real variables $x_1, x_2, \ldots, x_n$ that minimize the value of the *cost*

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

subject to constraints

$$\begin{aligned} A_{1,1}x_1 + A_{1,2}x_2 + \cdots + A_{1,n}x_n &= b_1, \\ A_{2,1}x_1 + A_{2,2}x_2 + \cdots + A_{2,n}x_n &= b_2, \\ &\vdots \\ A_{m,1}x_1 + A_{m,2}x_2 + \cdots + A_{m,n}x_n &= b_m \end{aligned}$$

and

$$x_1 \geq 0, \quad x_2 \geq 0, \ldots, \quad x_n \geq 0.$$

In this above programming problem, the function sending the n-dimensional vector $\mathbf{x}$ to the corresponding *cost* $\mathbf{c}^T\mathbf{x}$ is the objective function for the problem. A feasible solution to the problem consists of an n-dimensional vector $(x_1, x_2, \ldots, x_n)$ whose components satisfy the above constraints but do not necessarily minimize cost. An optimal solution is a feasible solution whose cost does not exceed that of any other feasible solution.

4.4 Basic Feasible Solutions to Linear Programming Problems in Dantzig Standard Form

We define the notion of a *basis* for a linear programming problem in Dantzig standard form.

Definition Let A be an $m \times n$ matrix of rank m with real coefficients, where $m \leq n$, let $\mathbf{b} \in \mathbb{R}^m$ be an m-dimensional column vector, let $\mathbf{c} \in \mathbb{R}^n$ be an n-dimensional column vector. Consider the following programming problem in Dantzig standard form:

> *find* $\mathbf{x} \in \mathbb{R}^n$ *so as to minimize* $\mathbf{c}^T\mathbf{x}$ *subject to constraints* $A\mathbf{x} = \mathbf{b}$ *and* $\mathbf{x} \geq \mathbf{0}$.

For each integer j between 1 and n, let $\mathbf{a}^{(j)}$ denote the m-dimensional vector determined by the jth column of the matrix A, so that $(\mathbf{a}^{(j)})_i = (A)_{i,j}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. A *basis* for this linear programming problem is a set consisting of m distinct integers $j_1, j_2, \ldots, j_m$ between 1 and n for which the corresponding vectors

$$\mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}, \ldots, \mathbf{a}^{(j_m)}$$

constitute a basis of the vector space $\mathbb{R}^m$.

We next define what is meant by saying that a feasible solution of a programming problem Dantzig standard form is a *basic feasible solution* for the programming problem.

Definition Let A be an $m \times n$ matrix of rank m with real coefficients, where $m \leq n$, let $\mathbf{b} \in \mathbb{R}^m$ be an m-dimensional column vector, let $\mathbf{c} \in \mathbb{R}^n$ be an n-dimensional column vector. Consider the following programming problem in Dantzig standard form:—

> *find* $\mathbf{x} \in \mathbb{R}^n$ *so as to minimize* $\mathbf{c}^T\mathbf{x}$ *subject to constraints* $A\mathbf{x} = \mathbf{b}$ *and* $\mathbf{x} \geq \mathbf{0}$.

A feasible solution $\mathbf{x}$ for this programming problem is said to be *basic* if there exists a basis B for the linear programming problem such that $(\mathbf{x})_j = 0$ when $j \notin B$.

Lemma 4.1 *Let A be an $m \times n$ matrix of rank m with real coefficients, where $m \leq n$, let $\mathbf{b} \in \mathbb{R}^m$ be an m-dimensional column vector, let $\mathbf{c} \in \mathbb{R}^n$ be an n-dimensional column vector. Consider the following programming problem in Dantzig standard form:*

find $\mathbf{x} \in \mathbb{R}^n$ *so as to minimize* $\mathbf{c}^T\mathbf{x}$ *subject to constraints* $A\mathbf{x} = \mathbf{b}$ *and* $\mathbf{x} \geq \mathbf{0}$.

Let $\mathbf{a}^{(j)}$ *denote the vector specified by the* j*th column of the matrix* A *for* $j = 1, 2, \ldots, n$. *Let* $\mathbf{x}$ *be a feasible solution of the linear programming problem. Suppose that the* m*-dimensional vectors* $\mathbf{a}^{(j)}$ *for which* $(\mathbf{x})_j > 0$ *are linearly independent. Then* $\mathbf{x}$ *is a basic feasible solution of the linear programming problem.*

Proof Let $\mathbf{x}$ be a feasible solution to the programming problem, let $x_j = (\mathbf{x})_j$ for all $j \in J$, where $J = \{1, 2, \ldots, n\}$, and let $K = \{j \in J : x_j > 0\}$. If the vectors $\mathbf{a}^{(j)}$ for which $j \in K$ are linearly independent then basic linear algebra ensures that further vectors $\mathbf{a}^{(j)}$ can be added to the linearly independent set $\{\mathbf{a}^{(j)} : j \in K\}$ so as to obtain a finite subset of $\mathbb{R}^m$ whose elements constitute a basis of that vector space (see Proposition 2.2). Thus exists a subset B of J satisfying $K \subset B \subset J$ such that the m-dimensional vectors $\mathbf{a}^{(j)}$ for which $j \in B$ constitute a basis of the real vector space $\mathbb{R}^m$. Moreover $(\mathbf{x})_j = 0$ for all $j \in J \setminus B$. It follows that $\mathbf{x}$ is a basic feasible solution to the linear programming problem, as required. ∎

Theorem 4.2 *Let* A *be an* $m \times n$ *matrix of rank* m *with real coefficients, where* $m \leq n$, *let* $\mathbf{b} \in \mathbb{R}^m$ *be an* m*-dimensional column vector, let* $\mathbf{c} \in \mathbb{R}^n$ *be an* n*-dimensional column vector. Consider the following programming problem in Dantzig standard form:*

find $\mathbf{x} \in \mathbb{R}^n$ *so as to minimize* $\mathbf{c}^T\mathbf{x}$ *subject to constraints* $A\mathbf{x} = \mathbf{b}$ *and* $\mathbf{x} \geq \mathbf{0}$.

If there exists a feasible solution to this programming problem then there exists a basic feasible solution to the problem. Moreover if there exists an optimal solution to the programming problem then there exists a basic optimal solution to the problem.

Proof Let $J = \{1, 2, \ldots, n\}$, and let $\mathbf{a}^{(j)}$ denote the vector specified by the jth column of the matrix A for all $j \in J$.

Let $\mathbf{x}$ be a feasible solution to the programming problem, let $x_j = (\mathbf{x})_j$ for all $j \in J$, and let $K = \{j \in J : x_j > 0\}$. Suppose that $\mathbf{x}$ is not basic. Then the vectors $\mathbf{a}^{(j)}$ for which $j \in K$ must be linearly dependent. We show that there then exists a feasible solution with fewer non-zero components than the given feasible solution $\mathbf{x}$.

Now there exist real numbers y_j for $j \in K$, not all zero, such that $\sum_{j \in K} y_j \mathbf{a}^{(j)} = \mathbf{0}$, because the vectors $\mathbf{a}^{(j)}$ for $j \in K$ are linearly dependent.

Let $y_j = 0$ for all $j \in J \setminus K$, and let $\mathbf{y} \in \mathbb{R}^n$ be the n-dimensional vector satisfying $(\mathbf{y})_j = y_j$ for $j = 1, 2, \ldots, n$. Then

$$A\mathbf{y} = \sum_{j \in J} y_j \mathbf{a}^{(j)} = \sum_{j \in K} y_j \mathbf{a}^{(j)} = \mathbf{0}.$$

It follows that $A(\mathbf{x} - \lambda\mathbf{y}) = \mathbf{b}$ for all real numbers λ, and thus $\mathbf{x} - \lambda\mathbf{y}$ is a feasible solution to the programming problem for all real numbers λ for which $\mathbf{x} - \lambda\mathbf{y} \geq \mathbf{0}$.

Now $\mathbf{y}$ is non-zero vector. Replacing $\mathbf{y}$ by $-\mathbf{y}$, if necessary, we can assume, without loss of generality, that at least one component of the vector $\mathbf{y}$ is positive. Let

$$\lambda_0 = \text{minimum}\left(\frac{x_j}{y_j} : j \in K \text{ and } y_j > 0\right),$$

and let j_0 be an element of K for which $\lambda_0 = x_{j_0}/y_{j_0}$. Then $\dfrac{x_j}{y_j} \geq \lambda_0$ for all $j \in J$ for which $y_j > 0$. Multiplying by the positive number y_j, we find that $x_j \geq \lambda_0 y_j$ and thus $x_j - \lambda_0 y_j \geq 0$ when $y_j > 0$. Also $\lambda_0 > 0$ and $x_j \geq 0$, and therefore $x_j - \lambda_0 y_j \geq 0$ when $y_j \leq 0$. Thus $x_j - \lambda_0 y_j \geq 0$ for all $j \in J$. Also $x_{j_0} - \lambda_0 y_{j_0} = 0$, and $x_j - \lambda_0 y_j = 0$ for all $j \in J \setminus K$. Let $\mathbf{x}' = \mathbf{x} - \lambda_0\mathbf{y}$. Then $\mathbf{x}' \geq \mathbf{0}$ and $A\mathbf{x}' = \mathbf{b}$, and thus $\mathbf{x}'$ is a feasible solution to the linear programming problem with fewer non-zero components than the given feasible solution.

Suppose in particular that the feasible solution $\mathbf{x}$ is optimal. Now there exist both positive and negative values of λ for which $\mathbf{x} - \lambda\mathbf{y} \geq \mathbf{0}$. If it were the case that $\mathbf{c}^T\mathbf{y} \neq 0$ then there would exist values of λ for which both $\mathbf{x} - \lambda\mathbf{y} \geq \mathbf{0}$ and $\lambda\mathbf{c}^T\mathbf{y} > 0$. But then $\mathbf{c}^T(x - \lambda\mathbf{y}) < \mathbf{c}^T\mathbf{x}$, contradicting the optimality of $\mathbf{x}$. It follows that $\mathbf{c}^T\mathbf{y} = 0$, and therefore $\mathbf{x} - \lambda\mathbf{y}$ is an optimal solution of the linear programming problem for all values of λ for which $\mathbf{x} - \lambda\mathbf{y} \geq \mathbf{0}$. The previous argument then shows that there exists a real number λ_0 for which $\mathbf{x} - \lambda_0\mathbf{y}$ is an optimal solution with fewer non-zero components than the given optimal solution $\mathbf{x}$.

We have shown that if there exists a feasible solution $\mathbf{x}$ which is not basic then there exists a feasible solution with fewer non-zero components than $\mathbf{x}$. It follows that if a feasible solution $\mathbf{x}$ is chosen such that it has the smallest possible number of non-zero components then it is a basic feasible solution of the linear programming problem.

Similarly we have shown that if there exists an optimal solution $\mathbf{x}$ which is not basic then there exists an optimal solution with fewer non-zero components than $\mathbf{x}$. It follows that if an optimal solution $\mathbf{x}$ is chosen such that it has the smallest possible number of non-zero components then it is a basic optimal solution of the linear programming problem. ∎

4.5 A Simplex Method Example with Five Variables and Two Constraints

We consider the following linear programming problem:—

minimize

$$3x_1 + 4x_2 + 2x_3 + 9x_4 + 5x_5$$

subject to the following constraints:
$5x_1 + 3x_2 + 4x_3 + 7x_4 + 3x_5 = 11$;
$4x_1 + x_2 + 3x_3 + 8x_4 + 4x_5 = 6$;
$x_j \geq 0$ *for* $j = 1, 2, 3, 4, 5$.

The constraints require that x_1, x_2, x_3, x_4, x_5 be non-negative real numbers satisfying the matrix equation

$$\begin{pmatrix} 5 & 3 & 4 & 7 & 3 \\ 4 & 1 & 3 & 8 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 11 \\ 6 \end{pmatrix}.$$

Thus we are required to find a (column) vector $\mathbf{x}$ with components x_1, x_2, x_3, x_4 and x_5 satisfying the equation $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{pmatrix} 5 & 3 & 4 & 7 & 3 \\ 4 & 1 & 3 & 8 & 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 11 \\ 6 \end{pmatrix}.$$

Let

$$\mathbf{a}^{(1)} = \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \quad \mathbf{a}^{(2)} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad \mathbf{a}^{(3)} = \begin{pmatrix} 4 \\ 3 \end{pmatrix},$$

$$\mathbf{a}^{(4)} = \begin{pmatrix} 7 \\ 8 \end{pmatrix} \quad \text{and} \quad \mathbf{a}^{(5)} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

For a feasible solution to the problem we must find non-negative real numbers x_1, x_2, x_3, x_4, x_5 such that

$$x_1\mathbf{a}^{(1)} + x_2\mathbf{a}^{(2)} + x_3\mathbf{a}^{(3)} + x_4\mathbf{a}^{(4)} + x_5\mathbf{a}^{(5)} = \mathbf{b}.$$

An optimal solution to the problem is a feasible solution that minimizes

$$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5$$

amongst all feasible solutions to the problem, where $c_1 = 3$, $c_2 = 4$, $c_3 = 2$, $c_4 = 9$ and $c_5 = 5$.

Let $\mathbf{c}$ denote the column vector whose ith component is c_i respectively. Then

$$\mathbf{c}^T = \begin{pmatrix} 3 & 4 & 2 & 9 & 5 \end{pmatrix},$$

and an optimal solution is a feasible solution that minimizes $\mathbf{c}^T\mathbf{x}$ amongst all feasible solutions to the problem. We refer to the quantity $\mathbf{c}^T\mathbf{x}$ as the *cost* of the feasible solution $\mathbf{x}$.

Let $I = \{1, 2, 3, 4, 5\}$. A basis for this optimization problem is a subset $\{j_1, j_2\}$ of I, where $j_1 \neq j_2$, for which the corresponding vectors $\mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}$ constitute a basis of $\mathbb{R}^2$. By inspection we see that each pair of vectors taken from the list $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \mathbf{a}^{(3)}, \mathbf{a}^{(4)}, \mathbf{a}^{(5)}$ consists of linearly independent vectors, and therefore each pair of vectors from this list constitutes a basis of $\mathbb{R}^2$. It follows that every subset of I with exactly two elements is a basis for the optimization problem.

A feasible solution $(x_1, x_2, x_3, x_4, x_5)$ to this optimization problem is a basic feasible solution if there exists a basis B for the optimization problem such that $x_j = 0$ when $j \neq B$.

In the case of the present problem, all subsets of $\{1, 2, 3, 4, 5\}$ with exactly two elements are bases for the problem. It follows that a feasible solution to the problem is a basic feasible solution if and only if the number of non-zero components of the solution does not exceed 2.

We take as given the following initial basic feasible solution $x_1 = 1$, $x_2 = 2$, $x_3 = x_4 = x_5 = 0$. One can readily verify that $\mathbf{a}^{(1)} + 2\mathbf{a}^{(2)} = \mathbf{b}$. This initial basic feasible solution is associated with the basis $\{1, 2\}$. The cost of this solution is 11.

We apply the procedures of the *simplex method* to test whether or not this basic feasible solution is optimal, and, if not, determine how to improve it.

The basis $\{1, 2\}$ determines a 2×2 minor M_B of A consisting of the first two columns of A. Thus

$$M_B = \begin{pmatrix} 5 & 3 \\ 4 & 1 \end{pmatrix}.$$

We now determine the components of the vector $\mathbf{p} \in \mathbb{R}^2$ whose transpose $\begin{pmatrix} p_1 & p_2 \end{pmatrix}$ satisfies the matrix equation

$$\begin{pmatrix} c_1 & c_2 \end{pmatrix} = \begin{pmatrix} p_1 & p_2 \end{pmatrix} M_B.$$

Now

$$M_B^{-1} = -\frac{1}{7}\begin{pmatrix} 1 & -3 \\ -4 & 5 \end{pmatrix}.$$

It follows that

$$\begin{aligned}
\mathbf{p}^T &= \begin{pmatrix} p_1 & p_2 \end{pmatrix} = \begin{pmatrix} c_1 & c_2 \end{pmatrix} M_B^{-1} \\
&= -\frac{1}{7} \begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ -4 & 5 \end{pmatrix} \\
&= \begin{pmatrix} \frac{13}{7} & -\frac{11}{7} \end{pmatrix}.
\end{aligned}$$

We next compute a vector $\mathbf{q} \in \mathbb{R}^5$, where $\mathbf{q}^T = \mathbf{c}^T - \mathbf{p}^T A$. Solving the equivalent matrix equation for the transpose $\mathbf{q}^T$ of the column vector $\mathbf{q}$, we find that

$$\begin{aligned}
\mathbf{q}^T &= \mathbf{c}^T - \mathbf{p}^T A \\
&= \begin{pmatrix} 3 & 4 & 2 & 9 & 5 \end{pmatrix} - \begin{pmatrix} \frac{13}{7} & -\frac{11}{7} \end{pmatrix} \begin{pmatrix} 5 & 3 & 4 & 7 & 3 \\ 4 & 1 & 3 & 8 & 4 \end{pmatrix} \\
&= \begin{pmatrix} 3 & 4 & 2 & 9 & 5 \end{pmatrix} - \begin{pmatrix} 3 & 4 & \frac{19}{7} & \frac{3}{7} & -\frac{5}{7} \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & -\frac{5}{7} & \frac{60}{7} & \frac{40}{7} \end{pmatrix}.
\end{aligned}$$

We denote the jth component of the vector j by q_j.

Now $q_3 < 0$. We show that this implies that the initial basic feasible solution is not optimal, and that it can be improved by bringing 3 (the index of the third column of A) into the basis.

Suppose that $\overline{\mathbf{x}}$ is a feasible solution of this optimization problem. Then $A\overline{\mathbf{x}} = \mathbf{b}$, and therefore

$$\mathbf{c}^T \overline{\mathbf{x}} = \mathbf{p}^T A \overline{\mathbf{x}} + \mathbf{q}^T \mathbf{x} = \mathbf{p}^T \mathbf{b} + \mathbf{q}^T \overline{\mathbf{x}}.$$

The initial basic feasible solution $\mathbf{x}$ satisfies

$$\mathbf{q}^T \mathbf{x} = \sum_{j=1}^{5} q_j x_j = 0,$$

because $q_1 = q_2 = 0$ and $x_3 = x_4 = x_5 = 0$. This comes about because the manner in which we determined first $\mathbf{p}$ then $\mathbf{q}$ ensures that $q_j = 0$ for all $j \in B$, whereas the components of the basic feasible solution $\mathbf{x}$ associated with the basis B satisfy $x_j = 0$ for $j \notin B$. We find therefore that $\mathbf{p}^T \mathbf{b}$ is the cost of the initial basic feasible solution.

The cost of the initial basic feasible solution is 11, and this is equal to the value of $\mathbf{p}^T \mathbf{b}$. The cost $\mathbf{c}^T \overline{\mathbf{x}}$ of any other basic feasible solution satisfies

$$\mathbf{c}^T \overline{\mathbf{x}} = 11 - \tfrac{5}{7} \overline{x}_3 + \tfrac{60}{7} \overline{x}_4 + \tfrac{40}{7} \overline{x}_5,$$

where $\overline{x}_j$ denotes the jth component of $\overline{\mathbf{x}}$.

We seek to determine a new basic feasible solution $\overline{\mathbf{x}}$ for which $\overline{x}_3 > 0$, $\overline{x}_4 = 0$ and $\overline{x}_5 = 0$. The cost of such a basic feasible solution will then be less than that of our initial basic feasible solution.

In order to find our new basic feasible solution we determine the relationships between the coefficients of a feasible solution $\overline{\mathbf{x}}$ for which $\overline{x}_4 = 0$ and $\overline{x}_5 = 0$. Now such a feasible solution must satisfy

$$\overline{x}_1\mathbf{a}^{(1)} + \overline{x}_2\mathbf{a}^{(2)} + \overline{x}_3\mathbf{a}^{(3)} = \mathbf{b} = x_1\mathbf{a}^{(1)} + x_2\mathbf{a}^{(2)},$$

where x_1 and x_2 are the non-zero coefficients of the initial basic feasible solution. Now the vectors $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$ constitute a basis of the real vector space $\mathbb{R}^2$. It follows that there exist real numbers $t_{1,3}$ and $t_{2,3}$ such that $\mathbf{a}^{(3)} = t_{1,3}\mathbf{a}^{(1)} + t_{2,3}\mathbf{a}^{(2)}$. It follows that

$$(\overline{x}_1 + t_{1,3}\overline{x}_3)\mathbf{a}^{(1)} + (\overline{x}_2 + t_{2,3}\overline{x}_3)\mathbf{a}^{(2)} = x_1\mathbf{a}^{(1)} + x_2\mathbf{a}^{(2)}.$$

The linear independence of $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$ then ensures that $\overline{x}_1 + t_{1,3}\overline{x}_3 = x_1$ and $\overline{x}_2 + t_{2,3}\overline{x}_3 = x_2$. Thus if $\overline{x}_3 = \lambda$, where $\lambda \geq 0$ then

$$\overline{x}_1 = x_1 - \lambda t_{1,3}, \quad \overline{x}_2 = x_2 - \lambda t_{2,3}.$$

Thus, once $t_{1,3}$ and $t_{2,3}$ have been determined, we can determine the range of values of λ that ensure that $\overline{x}_1 \geq 0$ and $\overline{x}_2 \geq 0$.

In order to determine the values of $t_{1,3}$ and $t_{2,3}$ we note that

$$\begin{aligned} \mathbf{a}^{(1)} &= \begin{pmatrix} 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \mathbf{a}^{(2)} &= \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

and therefore

$$\begin{aligned} \mathbf{a}^{(3)} &= t_{3,1}\mathbf{a}^{(1)} + t_{3,2}\mathbf{a}^{(2)} = \begin{pmatrix} 5 & 3 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} t_{3,1} \\ t_{3,2} \end{pmatrix} \\ &= M_B \begin{pmatrix} t_{3,1} \\ t_{3,2} \end{pmatrix}, \end{aligned}$$

where

$$M_B = \begin{pmatrix} 5 & 3 \\ 4 & 1 \end{pmatrix}.$$

It follows that

$$\begin{pmatrix} t_{3,1} \\ t_{3,2} \end{pmatrix} = M_B^{-1}\mathbf{a}^{(3)} = -\frac{1}{7}\begin{pmatrix} 1 & -3 \\ -4 & 5 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \end{pmatrix} = \begin{pmatrix} \frac{5}{7} \\ \frac{1}{7} \end{pmatrix}.$$

Thus $t_{3,1} = \frac{5}{7}$ and $t_{3,2} = \frac{1}{7}$.

We now determine the feasible solutions $\overline{\mathbf{x}}$ of this optimization problem that satisfy $\overline{x}_3 = \lambda$ and $\overline{x}_4 = \overline{x}_5 = 0$. we have already shown that

$$\overline{x}_1 = x_1 - \lambda t_{1,3}, \quad \overline{x}_2 = x_2 - \lambda t_{2,3}.$$

Now $x_1 = 1$, $x_2 = 2$, $t_{1,3} = \frac{5}{7}$ and $t_{2,3} = \frac{1}{7}$. It follows that $\overline{x}_1 = 1 - \frac{5}{7}\lambda$ and $\overline{x}_2 = 2 - \frac{1}{7}\lambda$. Now the components of a feasible solution must satisfy $\overline{x}_1 \geq 0$ and $\overline{x}_2 \geq 0$. it follows that $0 \leq \lambda \leq \frac{7}{5}$. Moreover on setting $\lambda = \frac{7}{5}$ we find that $\overline{x}_1 = 0$ and $\overline{x}_2 = \frac{9}{5}$. We thus obtain a new basic feasible solution $\overline{\mathbf{x}}$ associated to the basis $\{2, 3\}$, where

$$\overline{\mathbf{x}}^T = \begin{pmatrix} 0 & \frac{9}{5} & \frac{7}{5} & 0 & 0 \end{pmatrix}.$$

The cost of this new basic feasible solution is 10.

We now let B' and $\mathbf{x}'$ denote the new basic and new associated basic feasible solution respectively, so that $B' = \{2, 3\}$ and

$$\mathbf{x}'^T = \begin{pmatrix} 0 & \frac{9}{5} & \frac{7}{5} & 0 & 0 \end{pmatrix}.$$

We also let $M_{B'}$ be the 2×2 minor of the matrix A with columns indexed by the new basis B, so that

$$M_{B'} = \begin{pmatrix} 3 & 4 \\ 1 & 3 \end{pmatrix} \quad \text{and} \quad M_{B'}^{-1} = \frac{1}{5}\begin{pmatrix} 3 & -4 \\ -1 & 3 \end{pmatrix}.$$

We now determine the components of the vector $\mathbf{p}' \in \mathbb{R}^2$ whose transpose $\begin{pmatrix} p'_1 & p'_2 \end{pmatrix}$ satisfies the matrix equation

$$\begin{pmatrix} c_2 & c_3 \end{pmatrix} = \begin{pmatrix} p'_1 & p'_2 \end{pmatrix} M_{B'}.$$

We find that

$$\begin{aligned} \begin{pmatrix} p'_1 & p'_2 \end{pmatrix} &= \begin{pmatrix} c_2 & c_3 \end{pmatrix} M_{B'}^{-1} \\ &= \frac{1}{5}\begin{pmatrix} 4 & 2 \end{pmatrix}\begin{pmatrix} 3 & -4 \\ -1 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -2 \end{pmatrix}. \end{aligned}$$

We next compute the components of the vector $\mathbf{q}' \in \mathbb{R}^5$ so as to ensure that

$$\begin{aligned} \mathbf{q}'^T &= \mathbf{c}^T - \mathbf{p}'^T A \\ &= \begin{pmatrix} 3 & 4 & 2 & 9 & 5 \end{pmatrix} - \begin{pmatrix} 2 & -2 \end{pmatrix}\begin{pmatrix} 5 & 3 & 4 & 7 & 3 \\ 4 & 1 & 3 & 8 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 4 & 2 & 9 & 5 \end{pmatrix} - \begin{pmatrix} 2 & 4 & 2 & -2 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 11 & 7 \end{pmatrix}. \end{aligned}$$

The components of the vector $\mathbf{q}'$ determined using the new basis $\{2, 3\}$ are all non-negative. This ensures that the new basic feasible solution is an optimal solution.

Indeed let $\overline{\mathbf{x}}$ be a feasible solution of this optimization problem. Then $A\overline{\mathbf{x}'} = \mathbf{b}$, and therefore

$$\mathbf{c}^T\overline{\mathbf{x}} = \mathbf{p}'^T A\overline{\mathbf{x}} + \mathbf{q}'^T\mathbf{x}' = \mathbf{p}'^T\mathbf{b} + \mathbf{q}'^T\overline{\mathbf{x}}.$$

Moreover $\mathbf{p}'^T\mathbf{b} = 10$. It follows that

$$\mathbf{c}^T\overline{\mathbf{x}} = 10 + \overline{x}_1 + 11\overline{x}_4 + 7\overline{x}_5 \geq 10,$$

and thus the new basic feasible solution $\mathbf{x}'$ is optimal.

We summarize the result we have obtained. The optimization problem was the following:—

minimize

$$3x_1 + 4x_2 + 2x_3 + 9x_4 + 5x_5$$

subject to the following constraints:
$5x_1 + 3x_2 + 4x_3 + 7x_4 + 3x_5 = 11$;
$4x_1 + x_2 + 3x_3 + 8x_4 + 4x_5 = 6$;
$x_j \geq 0$ *for* $j = 1, 2, 3, 4, 5$.

We have found the following basic optimal solution to the problem:

$$x_1 = 0, \quad x_2 = \frac{9}{5}, \quad x_3 = \frac{7}{5}, \quad x_4 = 0, \quad x_5 = 0.$$

We now investigate all bases for this linear programming problem in order to determine which bases are associated with basic feasible solutions.

The problem is to find $\mathbf{x} \in \mathbb{R}^5$ that minimizes $\mathbf{c}^T\mathbf{x}$ subject to the constraints $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq 0$, where

$$A = \begin{pmatrix} 5 & 3 & 4 & 7 & 3 \\ 4 & 1 & 3 & 8 & 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 11 \\ 6 \end{pmatrix}$$

and

$$\mathbf{c}^T = \begin{pmatrix} 3 & 4 & 2 & 9 & 5 \end{pmatrix}.$$

For each two-element subset B of $\{1, 2, 3, 4, 5\}$ we compute M_B, M_B^{-1} and $M_B^{-1}\mathbf{b}$, where M_B is the 2×2 minor of the matrix A whose columns are indexed by the elements of B. We find the following:—

B	M_B	M_B^{-1}	$M_B^{-1}\mathbf{b}$	$\mathbf{c}^T M_B^{-1}\mathbf{b}$
$\{1,2\}$	$\begin{pmatrix} 5 & 3 \\ 4 & 1 \end{pmatrix}$	$-\frac{1}{7}\begin{pmatrix} 1 & -3 \\ -4 & 5 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$	11
$\{1,3\}$	$\begin{pmatrix} 5 & 4 \\ 4 & 3 \end{pmatrix}$	$-\begin{pmatrix} 3 & -4 \\ -4 & 5 \end{pmatrix}$	$\begin{pmatrix} -9 \\ 14 \end{pmatrix}$	1
$\{1,4\}$	$\begin{pmatrix} 5 & 7 \\ 4 & 8 \end{pmatrix}$	$\frac{1}{12}\begin{pmatrix} 8 & -7 \\ -4 & 5 \end{pmatrix}$	$\begin{pmatrix} \frac{23}{6} \\ -\frac{7}{6} \end{pmatrix}$	1
$\{1,5\}$	$\begin{pmatrix} 5 & 3 \\ 4 & 4 \end{pmatrix}$	$\frac{1}{8}\begin{pmatrix} 4 & -3 \\ -4 & 5 \end{pmatrix}$	$\begin{pmatrix} \frac{13}{4} \\ -\frac{7}{4} \end{pmatrix}$	1
$\{2,3\}$	$\begin{pmatrix} 3 & 4 \\ 1 & 3 \end{pmatrix}$	$\frac{1}{5}\begin{pmatrix} 3 & -4 \\ -1 & 3 \end{pmatrix}$	$\begin{pmatrix} \frac{9}{5} \\ \frac{7}{5} \end{pmatrix}$	10
$\{2,4\}$	$\begin{pmatrix} 3 & 7 \\ 1 & 8 \end{pmatrix}$	$\frac{1}{17}\begin{pmatrix} 8 & -7 \\ -1 & 3 \end{pmatrix}$	$\begin{pmatrix} \frac{46}{17} \\ \frac{7}{17} \end{pmatrix}$	$\frac{247}{17}$
$\{2,5\}$	$\begin{pmatrix} 3 & 3 \\ 1 & 4 \end{pmatrix}$	$\frac{1}{9}\begin{pmatrix} 4 & -3 \\ -1 & 3 \end{pmatrix}$	$\begin{pmatrix} \frac{26}{9} \\ \frac{7}{9} \end{pmatrix}$	$\frac{139}{9}$
$\{3,4\}$	$\begin{pmatrix} 4 & 7 \\ 3 & 8 \end{pmatrix}$	$\frac{1}{11}\begin{pmatrix} 8 & -7 \\ -3 & 4 \end{pmatrix}$	$\begin{pmatrix} \frac{46}{11} \\ -\frac{9}{11} \end{pmatrix}$	1
$\{3,5\}$	$\begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}$	$\frac{1}{7}\begin{pmatrix} 4 & -3 \\ -3 & 4 \end{pmatrix}$	$\begin{pmatrix} \frac{26}{7} \\ -\frac{9}{7} \end{pmatrix}$	1
$\{4,5\}$	$\begin{pmatrix} 7 & 3 \\ 8 & 4 \end{pmatrix}$	$\frac{1}{4}\begin{pmatrix} 4 & -3 \\ -8 & 7 \end{pmatrix}$	$\begin{pmatrix} \frac{13}{2} \\ -\frac{23}{2} \end{pmatrix}$	1

From this data, we see that there are four basic feasible solutions to the problem. We tabulate them below:—

B	$\mathbf{x}$	Cost
$\{1,2\}$	$(1,2,0,0,0)$	11
$\{2,3\}$	$(0,\frac{9}{5},\frac{7}{5},0,0)$	10
$\{2,4\}$	$(0,\frac{46}{17},0,\frac{7}{17},0)$	$\frac{247}{17} = 14.529\ldots$
$\{2,5\}$	$(0,\frac{26}{9},0,0,\frac{7}{9})$	$\frac{139}{9} = 15.444\ldots$

4.6 A Linear Tableau Example with Five Variables and Three Constraints

Example Consider the problem of minimizing $\mathbf{c}^T\mathbf{x}$ subject to constraints $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq 0$, where

$$A = \begin{pmatrix} 1 & 2 & 3 & 3 & 5 \\ 2 & 3 & 1 & 2 & 3 \\ 4 & 2 & 5 & 1 & 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 13 \\ 13 \\ 20 \end{pmatrix},$$

$$\mathbf{c}^T = \begin{pmatrix} 2 & 4 & 3 & 1 & 4 \end{pmatrix}.$$

As usual, we denote by $A_{i,j}$ the coefficient of the matrix A in the ith row and jth column, we denote by b_i the ith component of the m-dimensional vector $\mathbf{b}$, and we denote by c_j the jth component of the n-dimensional vector $\mathbf{c}$.

We let $\mathbf{a}^{(j)}$ be the m-dimensional vector specified by the jth column of the matrix A for $j = 1, 2, 3, 4, 5$. Then

$$\mathbf{a}^{(1)} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \quad \mathbf{a}^{(2)} = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}, \quad \mathbf{a}^{(3)} = \begin{pmatrix} 3 \\ 1 \\ 5 \end{pmatrix},$$

$$\mathbf{a}^{(4)} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{a}^{(5)} = \begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix}.$$

A *basis* B for this linear programming problem is a subset of $\{1, 2, 3, 4, 5\}$ consisting of distinct integers j_1, j_2, j_3 for which the corresponding vectors $\mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}, \mathbf{a}^{(j_3)}$ constitute a basis of the real vector space $\mathbb{R}^3$.

Given a basis B for the linear programming programming problem, where $B = \{j_1, j_2, j_3\}$, we denote by M_B the matrix whose columns are specified by the vectors $\mathbf{a}^{(j_1)}$, $\mathbf{a}^{(j_2)}$ and $\mathbf{a}^{(j_3)}$. Thus $(M_B)_{i,k} = A_{i,j_k}$ for $i = 1, 2, 3$ and $k = 1, 2, 3$. We also denote by $\mathbf{c}_B$ the 3-dimensional vector defined such that

$$\mathbf{c}_B^T = \begin{pmatrix} c_{j_1} & c_{j_2} & c_{j_3} \end{pmatrix}.$$

The ordering of the columns of M_B and $\mathbf{c}_B$ is determined by the ordering of the elements j_1, j_2 and j_3 of the basis. However we shall proceed on the basis that some ordering of the elements of a given basis has been chosen, and the matrix M_B and vector $\mathbf{c}_B$ will be determined so as to match the chosen ordering.

Let $j_1 = 1$, $j_2 = 2$ and $j_3 = 3$, and let $B = \{j_1, j_2, j_3\} = \{1, 2, 3\}$. Then B is a basis of the linear programming problem, and the invertible matrix M_B determined by $\mathbf{a}^{(j_k)}$ for $k = 1, 2, 3$ is the following 3×3 matrix:—

$$M_B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 4 & 2 & 5 \end{pmatrix}.$$

This matrix has determinant -23, and

$$M_B^{-1} = \frac{-1}{23} \begin{pmatrix} 13 & -4 & -7 \\ -6 & -7 & 5 \\ -8 & 6 & -1 \end{pmatrix} = \begin{pmatrix} -\frac{13}{23} & \frac{4}{23} & \frac{7}{23} \\ \frac{6}{23} & \frac{7}{23} & -\frac{5}{23} \\ \frac{8}{23} & -\frac{6}{23} & \frac{1}{23} \end{pmatrix}.$$

Then

$$M_B^{-1}\mathbf{a}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad M_B^{-1}\mathbf{a}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad M_B^{-1}\mathbf{a}^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

$$M_B^{-1}\mathbf{a}^{(4)} = \begin{pmatrix} -\frac{24}{23} \\ \frac{27}{23} \\ \frac{13}{23} \end{pmatrix} \quad \text{and} \quad M_B^{-1}\mathbf{a}^{(5)} = \begin{pmatrix} -\frac{25}{23} \\ \frac{31}{23} \\ \frac{26}{23} \end{pmatrix}.$$

Also

$$M_B^{-1}\mathbf{b} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}.$$

It follows that $\mathbf{x}$ is a basic feasible solution of the linear programming problem, where

$$\mathbf{x}^T = \begin{pmatrix} 1 & 3 & 2 & 0 & 0 \end{pmatrix}.$$

The vectors $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \mathbf{a}^{(3)}, \mathbf{a}^{(4)}, \mathbf{a}^{(5)}$, $\mathbf{b}$, $\mathbf{e}^{(1)}$, $\mathbf{e}^{(2)}$ and $\mathbf{e}^{(3)}$ can then be expressed as linear combinations of $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \mathbf{a}^{(3)}$ with coefficients as recorded in the following tableau:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\mathbf{e}^{(3)}$
$\mathbf{a}^{(1)}$	1	0	0	$-\frac{24}{23}$	$-\frac{25}{23}$	1	$-\frac{13}{23}$	$\frac{4}{23}$	$\frac{7}{23}$
$\mathbf{a}^{(2)}$	0	1	0	$\frac{27}{23}$	$\frac{31}{23}$	3	$\frac{6}{23}$	$\frac{7}{23}$	$-\frac{5}{23}$
$\mathbf{a}^{(3)}$	0	0	1	$\frac{13}{23}$	$\frac{26}{23}$	2	$\frac{8}{23}$	$-\frac{6}{23}$	$\frac{1}{23}$
	·	·	·	·	·	·	·	·	·

There is an additional row at the bottom of the tableau. This row is the *criterion row* of the tableau. The values in this row have not yet been calculated, but, when calculated according to the rules described below, the values in the criterion row will establish whether the current basic feasible solution is optimal and, if not, how it can be improved.

Ignoring the criterion row, we can represent the structure of the remainder of the tableau in block form as follows:—

	$\mathbf{a}^{(1)} \quad \cdots \quad \mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)} \quad \cdots \quad \mathbf{e}^{(3)}$
$\mathbf{a}^{(j_1)}$ $\vdots$ $\mathbf{a}^{(j_3)}$	$M_B^{-1}A$	$M_B^{-1}\mathbf{b}$	M_B^{-1}
	$\cdot$	$\cdot$	$\cdot$

We now employ the principles of the Simplex Method in order to determine whether or not the current basic feasible solution is optimal and, if not, how to improve it by changing the basis.

Let $\mathbf{p}$ be the 3-dimensional vector determined so that

$$\mathbf{p}^T = \mathbf{c}_B^T M_B^{-1}.$$

Then $\mathbf{p}^T M_B = \mathbf{c}_B^T$, and therefore $\mathbf{p}^T \mathbf{a}^{(j_k)} = c_{j_k}$ for $k = 1, 2, 3$. It follows that $(\mathbf{p}^T A)_j = c_j$ whenever $j \in B$. Putting in the relevant numerical values, we find that

$$\mathbf{p}^T M_B = \mathbf{c}_B^T = \begin{pmatrix} c_{j_1} & c_{j_2} & c_{j_3} \end{pmatrix} = \begin{pmatrix} c_1 & c_2 & c_3 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 3 \end{pmatrix},$$

and therefore

$$\mathbf{p}^T = \begin{pmatrix} 2 & 4 & 3 \end{pmatrix} M_B^{-1} = \begin{pmatrix} \frac{22}{23} & \frac{18}{23} & \frac{-3}{23} \end{pmatrix}.$$

We enter the values of p_1, p_2 and p_3 into the cells of the criterion row in the columns labelled by $\mathbf{e}^{(1)}$, $\mathbf{e}^{(2)}$ and $\mathbf{e}^{(3)}$ respectively. The tableau with these values entered is then as follows:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\mathbf{e}^{(3)}$
$\mathbf{a}^{(1)}$	1	0	0	$-\frac{24}{23}$	$-\frac{25}{23}$	1	$-\frac{13}{23}$	$\frac{4}{23}$	$\frac{7}{23}$
$\mathbf{a}^{(2)}$	0	1	0	$\frac{27}{23}$	$\frac{31}{23}$	3	$\frac{6}{23}$	$\frac{7}{23}$	$-\frac{5}{23}$
$\mathbf{a}^{(3)}$	0	0	1	$\frac{13}{23}$	$\frac{26}{23}$	2	$\frac{8}{23}$	$-\frac{6}{23}$	$\frac{1}{23}$
	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\frac{22}{23}$	$\frac{18}{23}$	$-\frac{3}{23}$

The values in the criterion row in the columns labelled by $\mathbf{e}^{(1)}$, $\mathbf{e}^{(2)}$ and $\mathbf{e}^{(3)}$ can be calculated from the components of the cost vector $\mathbf{c}$ and the values in these columns of the tableau. Indeed Let $r_{i,k} = (M_B^{-1})_{i,k}$ for $i = 1, 2, 3$ and $k = 1, 2, 3$. Then each $r_{i,k}$ is equal to the value of the tableau element located in the row labelled by $\mathbf{a}^{(j_i)}$ and the column labelled by $\mathbf{e}^{(k)}$. The definition of the vector $\mathbf{p}$ then ensures that

$$p_k = c_{j_1} r_{1,k} + c_{j_2} r_{2,k} + c_{j_3} r_{3,k}$$

for $k = 1, 2, 3$, where, for the current basis, $j_1 = 1$, $j_2 = 2$ and $j_3 = 3$.

The cost C of the current basic feasible solution $\mathbf{x}$ satisfies $C = \mathbf{c}^T\mathbf{x}$. Now $(\mathbf{p}^T A)_j = c_j$ for all $j \in B$, where $B = \{1, 2, 3\}$. Moreover the current basic feasible solution $\mathbf{x}$ satisfies $x_j = 0$ when $j \notin B$, where $x_j = (\mathbf{x})_j$ for $j = 1, 2, 3, 4, 5$. It follows that

$$\begin{aligned} C - \mathbf{p}^T\mathbf{b} &= \mathbf{c}^T\mathbf{x} - \mathbf{p}^T A\mathbf{x} = \sum_{j=1}^{5} (c_j - (\mathbf{p}^T A)_j) x_j \\ &= \sum_{j \in B} (c_j - (\mathbf{p}^T A)_j) x_j = 0, \end{aligned}$$

and thus

$$C = \mathbf{c}^T\mathbf{x} = \mathbf{p}^T\mathbf{b}.$$

Putting in the numerical values, we find that $C = 20$.

We enter the cost C into the criterion row of the tableau in the column labelled by the vector $\mathbf{b}$. The resultant tableau is then as follows:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\mathbf{e}^{(3)}$
$\mathbf{a}^{(1)}$	1	0	0	$-\frac{24}{23}$	$-\frac{25}{23}$	1	$-\frac{13}{23}$	$\frac{4}{23}$	$\frac{7}{23}$
$\mathbf{a}^{(2)}$	0	1	0	$\frac{27}{23}$	$\frac{31}{23}$	3	$\frac{6}{23}$	$\frac{7}{23}$	$-\frac{5}{23}$
$\mathbf{a}^{(3)}$	0	0	1	$\frac{13}{23}$	$\frac{26}{23}$	2	$\frac{8}{23}$	$-\frac{6}{23}$	$\frac{1}{23}$
	·	·	·	·	·	20	$\frac{22}{23}$	$\frac{18}{23}$	$-\frac{3}{23}$

Let s_i denote the value recorded in the tableau in the row labelled by $\mathbf{a}^{(j_i)}$ and the column labelled by $\mathbf{b}$ for $i = 1, 2, 3$. Then the construction of the tableau ensures that

$$\mathbf{b} = s_1 \mathbf{a}^{(j_1)} + s_2 \mathbf{a}^{(j_2)} + s_3 \mathbf{a}^{(j_3)},$$

and thus $s_i = x_{j_i}$ for $i = 1, 2, 3$, where $(x_1, x_2, x_3, x_4, x_5)$ is the current basic feasible solution. It follows that

$$C = c_{j_1} s_1 + c_{j_2} s_2 + c_{j_3} s_3,$$

where, for the current basis, $j_1 = 1$, $j_2 = 2$ and $j_3 = 3$. Thus the cost of the current basic feasible solution can be calculated from the components of the cost vector $\mathbf{c}$ and the values recorded in the rows above the criterion row of the tableau in the column labelled by the vector $\mathbf{b}$.

We next determine a 5-dimensional vector $\mathbf{q}$ such that $\mathbf{c}^T = \mathbf{p}^T A + \mathbf{q}^T$. We find that

$$\begin{aligned} -\mathbf{q}^T &= \mathbf{p}^T A - \mathbf{c}^T \\ &= \begin{pmatrix} \frac{22}{23} & \frac{18}{23} & \frac{-3}{23} \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 3 & 5 \\ 2 & 3 & 1 & 2 & 3 \\ 4 & 2 & 5 & 1 & 4 \end{pmatrix} \\ &\qquad - \begin{pmatrix} 2 & 4 & 3 & 1 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 4 & 3 & \frac{99}{23} & \frac{152}{23} \end{pmatrix} - \begin{pmatrix} 2 & 4 & 3 & 1 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & \frac{76}{23} & \frac{60}{23} \end{pmatrix} \end{aligned}$$

Thus

$$q_1 = 0, \quad q_2 = 0, \quad q_3 = 0, \quad q_4 = -\tfrac{76}{23}, \quad q_5 = -\tfrac{60}{23}.$$

The 4th and 5th components of the vector $\mathbf{q}$ are negative. It follows that the current basic feasible solution is not optimal. Indeed let $\overline{\mathbf{x}}$ be a basic feasible solution to the problem, and let $\overline{x}_j = (\overline{\mathbf{x}})_j$ for $j = 1, 2, 3, 4, 5$. Then the cost $\overline{C}$ of the feasible solution $\overline{\mathbf{x}}$ satisfies

$$\begin{aligned} \overline{C} &= \mathbf{c}^T\overline{\mathbf{x}} = \mathbf{p}^T A\overline{\mathbf{x}} + \mathbf{q}^T\overline{\mathbf{x}} = \mathbf{p}^T\mathbf{b} + \mathbf{q}^T\overline{\mathbf{x}} = C + \mathbf{q}^T\overline{\mathbf{x}} \\ &= C - \frac{76}{23}\overline{x}_4 - \frac{60}{23}\overline{x}_5. \end{aligned}$$

It follows that the basic feasible solution $\overline{\mathbf{x}}$ will have lower cost if either $\overline{x}_4 > 0$ or $\overline{x}_5 > 0$.

We enter the value of $-q_j$ into the criterion row of the tableau in the column labelled by $\mathbf{a}^{(j)}$ for $j = 1, 2, 3, 4, 5$. The completed tableau associated with basis $\{1, 2, 3\}$ is then as follows:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\mathbf{e}^{(3)}$
$\mathbf{a}^{(1)}$	1	0	0	$-\frac{24}{23}$	$-\frac{25}{23}$	1	$-\frac{13}{23}$	$\frac{4}{23}$	$\frac{7}{23}$
$\mathbf{a}^{(2)}$	0	1	0	$\frac{27}{23}$	$\frac{31}{23}$	3	$\frac{6}{23}$	$\frac{7}{23}$	$-\frac{5}{23}$
$\mathbf{a}^{(3)}$	0	0	1	$\frac{13}{23}$	$\frac{26}{23}$	2	$\frac{8}{23}$	$-\frac{6}{23}$	$\frac{1}{23}$
	0	0	0	$\frac{76}{23}$	$\frac{60}{23}$	20	$\frac{22}{23}$	$\frac{18}{23}$	$-\frac{3}{23}$

We refer to this tableau as the *extended simplex tableau* associated with the basis $\{1, 2, 3\}$.

The general structure of the extended simplex tableau is then as follows:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\mathbf{e}^{(3)}$
$\mathbf{a}^{(j_1)}$	$t_{1,1}$	$t_{1,2}$	$t_{1,3}$	$t_{1,4}$	$t_{1,5}$	s_1	$r_{1,1}$	$r_{1,2}$	$r_{1,3}$
$\mathbf{a}^{(j_2)}$	$t_{2,1}$	$t_{2,2}$	$t_{2,3}$	$t_{2,4}$	$t_{2,5}$	s_2	$r_{2,1}$	$r_{2,2}$	$r_{2,3}$
$\mathbf{a}^{(j_3)}$	$t_{3,1}$	$t_{3,2}$	$t_{3,3}$	$t_{3,4}$	$t_{3,5}$	s_3	$r_{3,1}$	$r_{3,2}$	$r_{3,3}$
	$-q_1$	$-q_2$	$-q_3$	$-q_4$	$-q_5$	C	p_1	p_2	p_3

where j_1, j_2 and j_3 are the elements of the current basis, and where the coefficients $t_{i,j}$ s_i and $r_{i,k}$ are determined so that

$$\mathbf{a}^{(j)} = \sum_{i=1}^{3} t_{i,j}\mathbf{a}^{(j_i)}, \quad \mathbf{b} = \sum_{i=1}^{3} s_i\mathbf{a}^{(j_i)}, \quad \mathbf{e}^{(k)} = \sum_{i=1}^{3} r_{i,k}\mathbf{a}^{(j_i)}$$

for $j = 1, 2, 3, 4, 5$ and $k = 1, 2, 3$.

The coefficients of the criterion row can then be calculated according to the following formulae:—

$$p_k = \sum_{i=1}^{3} c_{j_i} r_{i,k}, \quad C = \sum_{i=1}^{3} p_i b_i, \quad -q_j = \sum_{i=1}^{3} p_i A_{i,j} - c_j.$$

The extended simplex tableau can then be represented in block form as follows:—

	$\mathbf{a}^{(1)} \quad \cdots \quad \mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)} \quad \cdots \quad \mathbf{e}^{(3)}$
$\mathbf{a}^{(j_1)}$ $\vdots$ $\mathbf{a}^{(j_3)}$	$M_B^{-1}A$	$M_B^{-1}\mathbf{b}$	M_B^{-1}
	$\mathbf{p}^T A - \mathbf{c}^T$	$\mathbf{p}^T\mathbf{b}$	$\mathbf{p}^T$

The values in the criterion row in any column labelled by some $\mathbf{a}^{(j)}$ can also be calculated from the values in the relevant column in the rows above the criterion row.

To see this we note that the value entered into the tableau in the row labelled by $\mathbf{a}^{(j_i)}$ and the column labelled by $\mathbf{a}^{(j)}$ is equal to $t_{i,j}$, where $t_{i,j}$ is the coefficient in the ith row and jth column of the matrix $M_B^{-1}A$. Also $\mathbf{p}^T = \mathbf{c}_B^T M_B^{-1}$, where $(\mathbf{c}_B)_i = c_{j_i}$ for $i = 1, 2, 3$. It follows that

$$\mathbf{p}^T A = \mathbf{c}_B^T M_B^{-1} A = \sum_{i=1}^{3} c_{j_i} t_{i,j}.$$

Therefore

$$\begin{aligned} -q_j &= (\mathbf{p}^T A)_j - c_j \\ &= c_{j_1} t_{1,j} + c_{j_2} t_{2,j} + c_{j_3} t_{3,j} - c_j \end{aligned}$$

for $j = 1, 2, 3, 4, 5$.

The coefficients of the criterion row can then be calculated according to the formulae

$$p_k = \sum_{i=1}^{3} c_{j_i} r_{i,k}, \quad C = \sum_{i=1}^{3} c_{j_i} s_i, \quad -q_j = \sum_{i=1}^{3} c_{j_i} t_{i,j} - c_j.$$

The extended simplex tableau can therefore also be represented in block form as follows:—

	$\mathbf{a}^{(1)} \quad \cdots \quad \mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)} \quad \cdots \quad \mathbf{e}^{(3)}$
$\mathbf{a}^{(j_1)}$ $\vdots$ $\mathbf{a}^{(j_3)}$	$M_B^{-1}A$	$M_B^{-1}\mathbf{b}$	M_B^{-1}
	$\mathbf{c}_B^T M_B^{-1} A - \mathbf{c}^T$	$\mathbf{c}_B^T M_B^{-1}\mathbf{b}$	$\mathbf{c}_B^T M_B^{-1}$

We now carry through procedures for adjusting the basis and calculating the extended simplex tableau associated with the new basis.

We recall that the extended simplex tableau corresponding to the old basis $\{1, 2, 3\}$ is as follows:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\mathbf{e}^{(3)}$
$\mathbf{a}^{(1)}$	1	0	0	$-\frac{24}{23}$	$-\frac{25}{23}$	1	$-\frac{13}{23}$	$\frac{4}{23}$	$\frac{7}{23}$
$\mathbf{a}^{(2)}$	0	1	0	$\frac{27}{23}$	$\frac{31}{23}$	3	$\frac{6}{23}$	$\frac{7}{23}$	$-\frac{5}{23}$
$\mathbf{a}^{(3)}$	0	0	1	$\frac{13}{23}$	$\frac{26}{23}$	2	$\frac{8}{23}$	$-\frac{6}{23}$	$\frac{1}{23}$
	0	0	0	$\frac{76}{23}$	$\frac{60}{23}$	20	$\frac{22}{23}$	$\frac{18}{23}$	$-\frac{3}{23}$

We now consider which of the indices 4 and 5 to bring into the basis.

Suppose we look for a basis which includes the vector $\mathbf{a}^{(4)}$ together with two of the vectors $\mathbf{a}^{(1)}$, $\mathbf{a}^{(2)}$ and $\mathbf{a}^{(3)}$. A feasible solution $\overline{\mathbf{x}}$ with $\overline{x}_5 = 0$ will satisfy

$$\overline{\mathbf{x}}^T = \begin{pmatrix} 1 + \frac{24}{23}\lambda & 3 - \frac{27}{23}\lambda & 2 - \frac{13}{23}\lambda & \lambda & 0 \end{pmatrix},$$

where $\lambda = \overline{x}_4$. Indeed $A(\overline{\mathbf{x}} - \mathbf{x}) = \mathbf{0}$, where $\mathbf{x}$ is the current basic feasible solution, and therefore

$$(\overline{x}_1 - 1)\mathbf{a}^{(1)} + (\overline{x}_2 - 3)\mathbf{a}^{(2)} + (\overline{x}_3 - 2)\mathbf{a}^{(3)} + \overline{x}_4\mathbf{a}^{(4)} = \mathbf{0}.$$

Now

$$\mathbf{a}^{(4)} = -\tfrac{24}{23}\mathbf{a}^{(1)} + \tfrac{27}{23}\mathbf{a}^{(2)} + \tfrac{13}{23}\mathbf{a}^{(3)},$$

It follows that

$$(\overline{x}_1 - 1 - \tfrac{24}{23}\overline{x}_4)\mathbf{a}^{(1)} + (\overline{x}_2 - 3 + \tfrac{27}{33}\overline{x}_4)\mathbf{a}^{(2)} + (\overline{x}_3 - 2 + \tfrac{13}{23}\overline{x}_4)\mathbf{a}^{(3)} = \mathbf{0}.$$

But the vectors $\mathbf{a}^{(1)}$, $\mathbf{a}^{(2)}$ and $\mathbf{a}^{(3)}$ are linearly independent. Thus if $\overline{x}_4 = \lambda$ and $\overline{x}_5 = 0$ then

$$\overline{x}_1 - 1 - \tfrac{24}{23}\lambda = 0, \quad \overline{x}_2 - 3 + \tfrac{27}{23}\lambda = 0, \quad \overline{x}_3 - 2 + \tfrac{13}{23}\lambda = 0,$$

and thus

$$\overline{x}_1 = 1 + \tfrac{24}{23}\lambda, \quad \overline{x}_2 = 3 - \tfrac{27}{23}\lambda, \quad \overline{x}_3 = 2 - \tfrac{13}{23}\lambda.$$

For the solution $\overline{\mathbf{x}}$ to be feasible the components of $\overline{\mathbf{x}}$ must all be non-negative, and therefore λ must satisfy

$$\lambda \leq \min\left(3 \times \tfrac{23}{27}, \ 2 \times \tfrac{23}{13}\right).$$

Now $3 \times \frac{23}{27} = \frac{69}{27} \approx 2.56$ and $2 \times \frac{23}{13} = \frac{46}{13} \approx 3.54$. It follows that the maximum possible value of λ is $\frac{69}{27}$. The feasible solution corresponding to this value of λ is a basic feasible solution with basis $\{1, 3, 4\}$, and passing from the current basic feasible solution $\mathbf{x}$ to the new feasible basic solution would lower the cost by $-q_4\lambda$, where $-q_4\lambda = \frac{76}{23} \times \frac{69}{27} = \frac{228}{27} \approx 8.44$.

We examine this argument in more generality to see how to calculate the change in the cost that arises if an index j not in the current basis is brought into that basis. Let the current basis be $\{j_1, j_2, j_3\}$. Then

$$\mathbf{b} = s_1\mathbf{a}^{(j_1)} + s_2\mathbf{a}^{(j_2)} + s_3\mathbf{a}^{(j_3)}$$

and

$$\mathbf{a}^{(j)} = t_{1,j}\mathbf{a}^{(j_1)} + t_{2,j}\mathbf{a}^{(j_2)} + t_{3,j}\mathbf{a}^{(j_3)}.$$

Thus if $\overline{\mathbf{x}}$ is a feasible solution, and if $(\overline{\mathbf{x}})_{j'} = 0$ for $j' \notin \{j_1, j_2, j_3, j\}$, then

$$\overline{x}_{j_1}\mathbf{a}^{(j_1)} + \overline{x}_{j_2}\mathbf{a}^{(j_2)} + \overline{x}_{j_3}\mathbf{a}^{(j_3)} + \overline{x}_j\mathbf{a}^{(j)} - \mathbf{b} = \mathbf{0}.$$

Let $\lambda = \overline{x}_j$. Then

$$(\overline{x}_{j_1} + \lambda t_{1,j} - s_1)\mathbf{a}^{(j_1)} + (\overline{x}_{j_2} + \lambda t_{2,j} - s_2)\mathbf{a}^{(j_2)} + (\overline{x}_{j_3} + \lambda t_{3,j} - s_3)\mathbf{a}^{(j_3)} = \mathbf{0}.$$

But the vectors $\mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}, \mathbf{a}^{(j_3)}$ are linearly independent, because $\{j_1, j_2, j_3\}$ is a basis for the linear programming problem. It follows that

$$\overline{x}_{j_i} = s_i - \lambda t_{i,j}$$

for $i = 1, 2, 3$.

For a feasible solution we require $\lambda \geq 0$ and $s_i - \lambda t_{i,j} \geq 0$ for $i = 1, 2, 3$. We therefore require

$$0 \leq \lambda \leq \min \left(\frac{s_i}{t_{i,j}} : t_{i,j} > 0 \right).$$

We could therefore obtain a new basic feasible solution by ejecting from the current basis an index j_i for which the ratio $\dfrac{s_i}{t_{i,j}}$ has its minimum value, where this minimum is taken over those values of i for which $t_{i,j} > 0$. If we set λ equal to this minimum value, then the cost is then reduced by $-q_j\lambda$.

With the current basis we find that $s_2/t_{4,2} = \frac{69}{27}$ and $s_3/t_{4,3} = \frac{46}{13}$. Now $\frac{69}{27} < \frac{46}{13}$. It follows that we could bring the index 4 into the basis, obtaining a new basis $\{1, 3, 4\}$, to obtain a cost reduction equal to $\frac{228}{13}$, given that $\frac{76}{23} \times \frac{69}{27} = \frac{228}{13} \approx 8.44$.

We now calculate the analogous cost reduction that would result from bringing the index 5 into the basis. Now $s_2/t_{5,2} = \frac{69}{31}$ and $s_3/t_{5,3} = \frac{46}{26}$. Moreover $\frac{46}{26} < \frac{69}{31}$. It follows that we could bring the index 5 into the basis, obtaining a new basis $\{1, 2, 5\}$, to obtain a cost reduction equal to $\frac{60}{23} \times \frac{46}{26} = \frac{120}{26} \approx 4.62$.

We thus obtain the better cost reduction by changing basis to $\{1, 3, 4\}$.

We need to calculate the tableau associated with the basis $\{1, 3, 4\}$. We will initially ignore the change to the criterion row, and calculate the updated values in the cells of the other rows. The current tableau with the values in the criterion row deleted is as follows:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\mathbf{e}^{(3)}$
$\mathbf{a}^{(1)}$	1	0	0	$-\frac{24}{23}$	$-\frac{25}{23}$	1	$-\frac{13}{23}$	$\frac{4}{23}$	$\frac{7}{23}$
$\mathbf{a}^{(2)}$	0	1	0	$\frac{27}{23}$	$\frac{31}{23}$	3	$\frac{6}{23}$	$\frac{7}{23}$	$-\frac{5}{23}$
$\mathbf{a}^{(3)}$	0	0	1	$\frac{13}{23}$	$\frac{26}{23}$	2	$\frac{8}{23}$	$-\frac{6}{23}$	$\frac{1}{23}$
	·	·	·	·	·	·	·	·	·

Let $\mathbf{v}$ be a vector in $\mathbb{R}^3$ and suppose that

$$\mathbf{v} = \mu_1\mathbf{a}^{(1)} + \mu_2\mathbf{a}^{(2)} + \mu_3\mathbf{a}^{(3)} = \mu_1'\mathbf{a}^{(1)} + \mu_2'\mathbf{a}^{(4)} + \mu_3'\mathbf{a}^{(3)}.$$

Now

$$\mathbf{a}^{(4)} = -\tfrac{24}{23}\mathbf{a}^{(1)} + \tfrac{27}{23}\mathbf{a}^{(2)} + \tfrac{13}{23}\mathbf{a}^{(3)}.$$

On multiplying this equation by $\frac{23}{27}$, we find that

$$\tfrac{23}{27}\mathbf{a}^{(4)} = -\tfrac{24}{27}\mathbf{a}^{(1)} + \mathbf{a}^{(2)} + \tfrac{13}{27}\mathbf{a}^{(3)},$$

and therefore

$$\mathbf{a}^{(2)} = \tfrac{24}{27}\mathbf{a}^{(1)} + \tfrac{23}{27}\mathbf{a}^{(4)} - \tfrac{13}{27}\mathbf{a}^{(3)}.$$

It follows that

$$\mathbf{v} = (\mu_1 + \tfrac{24}{27}\mu_2)\mathbf{a}^{(1)} + \tfrac{23}{27}\mu_2\mathbf{a}^{(4)} + (\mu_3 - \tfrac{13}{27}\mu_2)\mathbf{a}^{(3)},$$

and thus

$$\mu_1' = \mu_1 + \tfrac{24}{27}\mu_2, \quad \mu_2' = \tfrac{23}{27}\mu_2, \quad \mu_3' = \mu_3 - \tfrac{13}{27}\mu_2.$$

Now each column of the tableau specifies the coefficients of the vector labelling the column of the tableau with respect to the basis specified by the vectors labelling the rows of the tableau.

The *pivot row* of the old tableau is that labelled by the vector $\mathbf{a}^{(2)}$ that is being ejected from the basis. The *pivot column* of the old tableau is that labelled by the vector $\mathbf{a}^{(4)}$ that is being brought into the basis. The *pivot element* of the tableau is the element or value in both the pivot row and the pivot column. In this example the pivot element has the value $\frac{27}{23}$.

We see from the calculations above that the values in the pivot row of the old tableau are transformed by multiplying them by the reciprocal $\frac{23}{27}$ of the pivot element; the entries in the first row of the old tableau are transformed by adding to them the entries below them in the pivot row multiplied by the factor $\frac{24}{27}$; the values in the third row of the old tableau are transformed by subtracting from them the entries above them in the pivot row multiplied by the factor $\frac{13}{27}$.

Indeed the coefficients $t_{i,j}$, s_i, $r_{i,k}$, $t'_{i,j}$, s'_i and $r'_{i,k}$ are defined for $i = 1, 2, 3$, $j = 1, 2, 3, 4, 5$ and $k = 1, 2, 3$ so that

$$\begin{aligned}
\mathbf{a}^{(j)} &= \sum_{i=1}^{3} t_{i,j}\mathbf{a}^{(j_i)} = \sum_{i=1}^{3} t'_{i,j}\mathbf{a}^{(j'_i)}, \\
\mathbf{b} &= \sum_{i=1}^{3} s_i\mathbf{a}^{(j_i)} = \sum_{i=1}^{3} s'_i\mathbf{a}^{(j'_i)}, \\
\mathbf{e}^{(k)} &= \sum_{i=1}^{3} r_{i,k}\mathbf{a}^{(j_i)} = \sum_{i=1}^{3} r'_{i,k}\mathbf{a}^{(j'_i)},
\end{aligned}$$

where $j_1 = j'_1 = 1$, $j_3 = j'_3 = 3$, $j_2 = 2$ and $j'_2 = 4$.

The general rule for transforming the coefficients of a vector when changing from the basis $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \mathbf{a}^{(3)}$ to the basis $\mathbf{a}^{(1)}, \mathbf{a}^{(4)}, \mathbf{a}^{(3)}$ ensure that

$$t'_{2,j} = \frac{1}{t_{2,4}}\, t_{2,j},$$

$$
\begin{aligned}
t'_{i,j} &= t_{i,j} - \frac{t_{i,4}}{t_{2,4}}\, t_{2,j} \quad (i = 1, 3).\\
s'_2 &= \frac{1}{t_{2,4}}\, s_2,\\
s'_i &= s_i - \frac{t_{i,4}}{t_{2,4}}\, s_2 \quad (i = 1, 3).
\end{aligned}
$$

$$
\begin{aligned}
r'_{2,k} &= \frac{1}{t_{2,4}}\, r_{2,j},\\
r'_{i,k} &= r_{i,k} - \frac{t_{i,4}}{t_{2,4}}\, r_{2,k} \quad (i = 1, 3).
\end{aligned}
$$

The quantity $t_{2,4}$ is the value of the pivot element of the old tableau. The quantities $t_{2,j}$, s_2 and $r_{2,k}$ are those that are recorded in the pivot row of that tableau, and the quantities $t_{i,4}$ are those that are recorded in the pivot column of the tableau.

We thus obtain the following tableau:–

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\mathbf{e}^{(3)}$
$\mathbf{a}^{(1)}$	1	$\frac{24}{27}$	0	0	$\frac{3}{27}$	$\frac{99}{27}$	$-\frac{9}{27}$	$\frac{12}{27}$	$\frac{3}{27}$
$\mathbf{a}^{(4)}$	0	$\frac{23}{27}$	0	1	$\frac{31}{27}$	$\frac{69}{27}$	$\frac{6}{27}$	$\frac{7}{27}$	$-\frac{5}{27}$
$\mathbf{a}^{(3)}$	0	$-\frac{13}{27}$	1	0	$\frac{13}{27}$	$\frac{15}{27}$	$\frac{6}{27}$	$-\frac{11}{27}$	$\frac{4}{27}$
	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$

The values in the column of the tableau labelled by the vector $\mathbf{b}$ give us the components of a new basic feasible solution $\mathbf{x}'$. Indeed the column specifies that

$$\mathbf{b} = \tfrac{99}{27}\mathbf{a}^{(1)} + \tfrac{69}{27}\mathbf{a}^{(4)} + \tfrac{15}{27}\mathbf{a}^{(2)},$$

and thus $A\mathbf{x}' = \mathbf{b}$ where

$$\mathbf{x}'^T = \begin{pmatrix} \frac{99}{27} & 0 & \frac{15}{27} & \frac{69}{27} & 0 \end{pmatrix}.$$

We now calculate the new values for the criterion row. The new basis B' is given by $B' = \{j'_1, j'_2, j'_3\}$, where $j'_1 = 1$, $j'_2 = 4$ and $j'_3 = 3$. The values p'_1, p'_2 and p'_3 that are to be recorded in the criterion row of the new tableau in the columns labelled by $\mathbf{e}^{(1)}$, $\mathbf{e}^{(2)}$ and $\mathbf{e}^{(3)}$ respectively are determined by the equation

$$p'_k = c_{j'_1} r'_{1,k} + c_{j'_2} r'_{2,k} + c_{j'_3} r'_{3,k}$$

for $k = 1, 2, 3$, where

$$c_{j_1'} = c_1 = 2, \quad c_{j_2'} = c_4 = 1, \quad c_{j_3'} = c_3 = 3,$$

and where $r'_{i,k}$ denotes the ith component of the vector $\mathbf{e}^{(k)}$ with respect to the basis $\mathbf{a}^{(1)}, \mathbf{a}^{(4)}, \mathbf{a}^{(3)}$ of $\mathbb{R}^3$.

We find that

$$\begin{aligned} p_1' &= c_{j_1'} r'_{1,1} + c_{j_2'} r'_{2,1} + c_{j_3'} r'_{3,1} \\ &= 2 \times \left(-\tfrac{9}{27}\right) + 1 \times \tfrac{6}{27} + 3 \times \tfrac{6}{27} = \tfrac{6}{27}, \\ p_2' &= c_{j_1'} r'_{1,2} + c_{j_2'} r'_{2,2} + c_{j_3'} r'_{3,2} \\ &= 2 \times \tfrac{12}{27} + 1 \times \tfrac{7}{27} + 3 \times \left(-\tfrac{11}{27}\right) = -\tfrac{2}{27}, \\ p_3' &= c_{j_1'} r'_{1,3} + c_{j_2'} r'_{2,3} + c_{j_3'} r'_{3,3} \\ &= 2 \times \tfrac{3}{27} + 1 \times \left(-\tfrac{5}{27}\right) + 3 \times \tfrac{4}{27} = \tfrac{13}{27}. \end{aligned}$$

We next calculate the cost C' of the new basic feasible solution. The quantities s_1', s_2' and s_3' satisfy $s_i' = x'_{j_i}$ for $i = 1, 2, 3$, where $(x_1', x_2', x_3', x_4', x_5')$ is the new basic feasible solution. It follows that

$$C' = c_{j_1} s_1' + c_{j_2} s_2' + c_{j_3} s_3',$$

where s_1, s_2 and s_3 are determined so that

$$\mathbf{b} = s_1' \mathbf{a}^{(j_1')} + s_2' \mathbf{a}^{(j_2')} + s_3' \mathbf{a}^{(j_3')}.$$

The values of s_1', s_2' and s_3' have already been determined, and have been recorded in the column of the new tableau labelled by the vector $\mathbf{b}$.

We can therefore calculate C' as follows:—

$$\begin{aligned} C' &= c_{j_1'} s_1' + c_{j_2} s_2' + c_{j_3} s_3' = c_1 s_1' + c_4 s_2' + c_3 s_3' \\ &= 2 \times \tfrac{99}{27} + \tfrac{69}{27} + 3 \times \tfrac{15}{27} = \tfrac{312}{27}. \end{aligned}$$

Alternatively we can use the identity $C' = \mathbf{p}'^T \mathbf{b}$ to calculate C' as follows:

$$C' = p_1' b_1 + p_2' b_2 + p_3' b_3 = \tfrac{6}{27} \times 13 - \tfrac{2}{27} \times 13 + \tfrac{13}{27} \times 20 = \tfrac{312}{27}.$$

We now enter the values of p_1', p_2', p_3' and C' into the tableau associated with basis $\{1, 4, 3\}$. The tableau then takes the following form:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\mathbf{e}^{(3)}$
$\mathbf{a}^{(1)}$	1	$\frac{24}{27}$	0	0	$\frac{3}{27}$	$\frac{99}{27}$	$-\frac{9}{27}$	$\frac{12}{27}$	$\frac{3}{27}$
$\mathbf{a}^{(4)}$	0	$\frac{23}{27}$	0	1	$\frac{31}{27}$	$\frac{69}{27}$	$\frac{6}{27}$	$\frac{7}{27}$	$-\frac{5}{27}$
$\mathbf{a}^{(3)}$	0	$-\frac{13}{27}$	1	0	$\frac{13}{27}$	$\frac{15}{27}$	$\frac{6}{27}$	$-\frac{11}{27}$	$\frac{4}{27}$
	·	·	·	·	·	$\frac{312}{27}$	$\frac{6}{27}$	$-\frac{2}{23}$	$\frac{13}{23}$

In order to complete the extended tableau, it remains to calculate the values $-q'_j$ for $j = 1, 2, 3, 4, 5$, where q'_j satisfies the equation $-q'_j = \mathbf{p}'^T \mathbf{a}_j - c_j$ for $j = 1, 2, 3, 4, 5$.

Now q'_j is the jth component of the vector $\mathbf{q}'$ that satisfies the matrix equation $-\mathbf{q}'^T = \mathbf{p}'^T A - \mathbf{c}^T$. It follows that

$$
\begin{aligned}
-\mathbf{q}'^T &= \mathbf{p}'^T A - \mathbf{c}^T \\
&= \begin{pmatrix} \frac{6}{27} & \frac{-2}{27} & \frac{13}{27} \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 3 & 5 \\ 2 & 3 & 1 & 2 & 3 \\ 4 & 2 & 5 & 1 & 4 \end{pmatrix} \\
&\qquad - \begin{pmatrix} 2 & 4 & 3 & 1 & 4 \end{pmatrix} \\
&= \begin{pmatrix} 2 & \frac{32}{27} & 3 & 1 & \frac{76}{27} \end{pmatrix} - \begin{pmatrix} 2 & 4 & 3 & 1 & 4 \end{pmatrix} \\
&= \begin{pmatrix} 0 & -\frac{76}{27} & 0 & 0 & -\frac{32}{27} \end{pmatrix}
\end{aligned}
$$

Thus

$$
q'_1 = 0, \quad q'_2 = \tfrac{76}{27}, \quad q'_3 = 0, \quad q'_4 = 0, \quad q'_5 = \tfrac{32}{27}.
$$

The value of each q'_j can also be calculated from the other values recorded in the column of the extended simplex tableau labelled by the vector $\mathbf{a}^{(j)}$. Indeed the vector $\mathbf{p}'$ is determined so as to satisfy the equation $\mathbf{p}'^T \mathbf{a}^{(j')} = c_{j'}$ for all $j' \in B'$. It follows that

$$
\mathbf{p}'^T \mathbf{a}^{(j)} = \sum_{i=1}^{3} t_{i,j} \mathbf{p}'^T \mathbf{a}^{(j'_i)} = \sum_{i=1}^{3} c_{j'_i} t'_{i,j},
$$

and therefore

$$
-q'_j = \sum_{i=1}^{3} c_{j'_i} t'_{i,j} - c_j.
$$

The extended simplex tableau for the basis $\{1, 4, 3\}$ has now been computed, and the completed tableau is as follows:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\mathbf{e}^{(3)}$
$\mathbf{a}^{(1)}$	1	$\frac{24}{27}$	0	0	$\frac{3}{27}$	$\frac{99}{27}$	$-\frac{9}{27}$	$\frac{12}{27}$	$\frac{3}{27}$
$\mathbf{a}^{(4)}$	0	$\frac{23}{27}$	0	1	$\frac{31}{27}$	$\frac{69}{27}$	$\frac{6}{27}$	$\frac{7}{27}$	$-\frac{5}{27}$
$\mathbf{a}^{(3)}$	0	$-\frac{13}{27}$	1	0	$\frac{13}{27}$	$\frac{15}{27}$	$\frac{6}{27}$	$-\frac{11}{27}$	$\frac{4}{27}$
	0	$-\frac{76}{27}$	0	0	$-\frac{32}{27}$	$\frac{312}{27}$	$\frac{6}{27}$	$-\frac{2}{23}$	$\frac{13}{23}$

The fact that $q'_j \geq 0$ for $j = 1, 2, 3, 4, 5$ shows that we have now found our basic optimal solution. Indeed the cost $\overline{C}$ of any feasible solution $\overline{\mathbf{x}}$ satisfies

$$
\overline{C} \;=\; \mathbf{c}^T \overline{\mathbf{x}} = \mathbf{p}'^T A \overline{\mathbf{x}} + \mathbf{q}'^T \overline{\mathbf{x}} = \mathbf{p}'^T \mathbf{b} + \mathbf{q}'^T \overline{\mathbf{x}}
$$

$$
\begin{aligned}
&= C' + \mathbf{q}'^T \overline{\mathbf{x}} \\
&= C' + \frac{76}{27}\overline{x}_2 + \frac{32}{27}\overline{x}_5,
\end{aligned}
$$

where $\overline{x}_2 = (\overline{\mathbf{x}})_2$ and $\overline{x}_5 = (\overline{\mathbf{x}})_5$.

Therefore $\mathbf{x}'$ is a basic optimal solution to the linear programming problem, where

$$\mathbf{x}'^T = \begin{pmatrix} \frac{99}{27} & 0 & \frac{15}{27} & \frac{69}{27} & 0 \end{pmatrix}.$$

It is instructive to compare the pivot row and criterion row of the tableau for the basis $\{1, 2, 3\}$ with the corresponding rows of the tableau for the basis $\{1, 4, 3\}$.

These rows in the old tableau for the basis $\{1, 2, 3\}$ contain the following values:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\mathbf{e}^{(3)}$
$\mathbf{a}^{(2)}$	0	1	0	$\frac{27}{23}$	$\frac{31}{23}$	3	$\frac{6}{23}$	$\frac{7}{23}$	$-\frac{5}{23}$
	0	0	0	$\frac{76}{23}$	$\frac{60}{23}$	20	$\frac{22}{23}$	$\frac{18}{23}$	$-\frac{3}{23}$

The corresponding rows in the new tableau for the basis $\{1, 4, 3\}$ contain the following values:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\mathbf{e}^{(3)}$
$\mathbf{a}^{(4)}$	0	$\frac{23}{27}$	0	1	$\frac{31}{27}$	$\frac{69}{27}$	$\frac{6}{27}$	$\frac{7}{27}$	$-\frac{5}{27}$
	0	$-\frac{76}{27}$	0	0	$-\frac{32}{27}$	$\frac{312}{27}$	$\frac{6}{27}$	$-\frac{2}{23}$	$\frac{13}{23}$

If we examine the values of the criterion row in the new tableau we find that they are obtained from corresponding values in the criterion row of the old tableau by subtracting off the corresponding elements of the pivot row of the old tableau multiplied by the factor $\frac{76}{27}$. As a result, the new tableau has value 0 in the cell of the criterion row in column $\mathbf{a}^{(4)}$. Thus the same rule used to calculate values in other rows of the new tableau would also have yielded the correct elements in the criterion row of the tableau.

We now investigate the reasons why this is so.

First we consider the transformation of the elements of the criterion row in the columns labelled by $\mathbf{a}^{(j)}$ for $j = 1, 2, 3, 4, 5$. Now the coefficients $t_{i,j}$ and $t'_{i,j}$ are defined for $i = 1, 2, 3$ and $j = 1, 2, 3, 4, 5$ so that

$$\mathbf{a}^{(j)} = \sum_{i=1}^{3} t_{i,j} \mathbf{a}^{(j_i)} = \sum_{i=1}^{3} t'_{i,j} \mathbf{a}^{(j'_i)},$$

where $j_1 = j'_1 = 1$, $j_3 = j'_3 = 3$, $j_2 = 2$ and $j'_2 = 4$. Moreover

$$t'_{2,j} = \frac{1}{t_{2,4}}\, t_{2,j}$$

and

$$t'_{i,j} = t_{i,j} - \frac{t_{i,4}}{t_{2,4}}\, t_{2,j} \quad (i = 1, 3).$$

Now

$$\begin{aligned} -q_j &= \sum_{i=1}^{3} c_{j_i} t_{i,j} - c_j \\ &= c_1 t_{1,j} + c_2 t_{2,j} + c_3 t_{3,j} - c_j, \\ -q'_j &= \sum_{i=1}^{3} c_{j'_i} t'_{i,j} - c_j. \\ &= c_1 t'_{1,j} + c_4 t'_{2,j} + c_3 t'_{3,j} - c_j. \end{aligned}$$

Therefore

$$\begin{aligned} q_j - q'_j &= c_1(t'_{1,j} - t_{1,j}) + c_4 t'_{2,j} - c_2 t_{2,j} + c_3(t'_{3,j} - t_{3,j}) \\ &= \frac{1}{t_{2,4}}\, (-c_1 t_{1,4} + c_4 - c_2 t_{2,4} - c_3 t_{3,4})\, t_{2,j} \\ &= \frac{q_4}{t_{2,4}}\, t_{2,j} \end{aligned}$$

and thus

$$-q'_j = -q_j + \frac{q_4}{t_{2,4}}\, t_{2,j}$$

for $j = 1, 2, 3, 4, 5$.

Next we note that

$$\begin{aligned} C &= \sum_{i=1}^{3} c_{j_i} s_i = c_1 s_1 + c_2 s_2 + c_3 s_3, \\ C' &= \sum_{i=1}^{3} c_{j'_i} s'_i = c_1 s'_1 + c_4 s'_2 + c_3 s'_3. \end{aligned}$$

Therefore

$$\begin{aligned} C' - C &= c_1(s'_1 - s_1) + c_4 s'_2 - c_2 s_2 + c_3(s'_3 - s_3) \\ &= \frac{1}{t_{2,4}}\, (-c_1 t_{1,4} + c_4 - c_2 t_{2,4} - c_3 t_{3,4})\, s_2 \\ &= \frac{q_4}{t_{2,4}}\, s_2 \end{aligned}$$

and thus

$$C' = q_k + \frac{q_4}{t_{2,4}}\, s_2$$

for $k = 1, 2, 3$.

To complete the verification that the criterion row of the extended simplex tableau transforms according to the same rule as the other rows we note that

$$\begin{aligned} p_k &= \sum_{i=1}^{3} c_{j_i} r_{i,k} = c_1 r_{1,k} + c_2 r_{2,k} + c_3 r_{3,k}, \\ p'_k &= \sum_{i=1}^{3} c_{j'_i} r'_{i,k} = c_1 r'_{1,k} + c_4 r'_{2,k} + c_3 r'_{3,k}. \end{aligned}$$

Therefore

$$\begin{aligned} p'_k - p_k &= c_1(r'_{1,k} - r_{1,k}) + c_4 r'_{2,k} - c_2 r_{2,k} + c_3(r'_{3,k} - r_{3,k}) \\ &= \frac{1}{t_{2,4}}\, (-c_1 t_{1,4} + c_4 - c_2 t_{2,4} - c_3 t_{3,4})\, r_{2,k} \\ &= \frac{q_4}{t_{2,4}}\, r_{2,k} \end{aligned}$$

and thus

$$p'_k = p_k + \frac{q_4}{t_{2,4}}\, r_{2,k}$$

for $k = 1, 2, 3$.

This completes the discussion of the structure and properties of the extended simplex tableau associated with the optimization problem under discussion.

4.7 Some Results concerning Finite-Dimensional Real Vector Spaces

We consider the representation of vectors belonging to the m-dimensional vector space $\mathbb{R}^m$ as linear combinations of basis vectors belonging to some chosen basis of this m-dimensional real vector space.

Elements of $\mathbb{R}^m$ are normally considered to be *column vectors* represented by $m \times 1$ matrices. Given any $\mathbf{v} \in \mathbb{R}^m$, we denote by $(\mathbf{v})_i$ the ith component of the vector $\mathbf{v}$, and we denote by $\mathbf{v}^T$ the $1 \times m$ row vector that is the transpose of the column vector representing $\mathbf{v} \in \mathbb{R}^m$. Thus

$$\mathbf{v}^T = \left(\begin{array}{cccc} v_1 & v_2 & \cdots & v_m \end{array} \right),$$

where $v_i = (\mathbf{v})_i$ for $i = 1, 2, \ldots, m$.

We define the *standard basis* of the real vector space $\mathbb{R}^m$ to be the basis

$$\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \ldots, \mathbf{e}^{(m)}$$

defined such that

$$(\mathbf{e}^{(k)})_i = \begin{cases} 1 & \text{if } k = i; \\ 0 & \text{if } k \neq i. \end{cases}$$

It follows that $\mathbf{v} = \sum_{i=1}^{m} (\mathbf{v})_i \mathbf{e}^{(i)}$ for all $\mathbf{v} \in \mathbb{R}^m$.

Let $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(m)}$ be a basis of the real vector space $\mathbb{R}^m$, and let $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \ldots, \mathbf{e}^{(m)}$ denote the standard basis of $\mathbb{R}^m$. Then there exists an invertible $m \times m$ matrix M with the property that

$$\mathbf{u}^{(k)} = \sum_{i=1}^{m} (M)_{i,k} \mathbf{e}^{(i)}$$

for $k = 1, 2, \ldots, m$.

The product $M\mathbf{v}$ is defined in the usual fashion for any m-dimensional vector $\mathbf{v}$: the vector $\mathbf{v}$ is expressed as an $m \times 1$ column vector, and the matrix product is then calculated according to the usual rules of matrix multiplication, so that

$$(M\mathbf{v})_i = \sum_{k=1}^{m} (M)_{i,k} (\mathbf{v})_k,$$

and thus

$$M\mathbf{v} = \sum_{i=1}^{m} \sum_{k=1}^{m} M_{i,k} (\mathbf{v})_k \mathbf{e}^{(i)} = \sum_{k=1}^{m} (\mathbf{v})_k \mathbf{u}^{(k)}.$$

Then $M\mathbf{e}^{(i)} = \mathbf{u}^{(i)}$ for $i = 1, 2, \ldots, m$. The inverse matrix M^{-1} of M then satisfies $M^{-1}\mathbf{u}^{(i)} = \mathbf{e}^{(i)}$ for $i = 1, 2, \ldots, m$.

Lemma 4.3 *Let m be a positive integer, let $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(m)}$ be a basis of $\mathbb{R}^m$, let $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \ldots, \mathbf{e}^{(m)}$ denote the standard basis of $\mathbb{R}^m$, and let M be the non-singular matrix that satisfies $M\mathbf{e}^{(i)} = \mathbf{u}^{(i)}$ for $i = 1, 2, \ldots, m$. Let $\mathbf{v}$ be a vector in $\mathbb{R}^m$, and let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be the unique real numbers for which*

$$\mathbf{v} = \sum_{i=1}^{m} \lambda_i \mathbf{u}^{(i)}.$$

Then λ_i is the ith component of the vector $M^{-1}\mathbf{v}$ for $i = 1, 2, \ldots, m$.

Proof The inverse matrix M^{-1} of M satisfies $M^{-1}\mathbf{u}^{(k)} = \mathbf{e}^{(k)}$ for $k = 1, 2, \ldots, m$. It follows that

$$M^{-1}\mathbf{v} = \sum_{k=1}^{m} \lambda_k M^{-1}\mathbf{u}^{(k)} = \sum_{k=1}^{m} \lambda_k \mathbf{e}^{(k)},$$

and thus $\lambda_1, \lambda_2, \ldots, \lambda_m$ are the components of the column vector $M^{-1}\mathbf{v}$, as required. ∎

Lemma 4.4 *Let m be a positive integer, let $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(m)}$ be a basis of $\mathbb{R}^m$, let $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \ldots, \mathbf{e}^{(m)}$ denote the standard basis of $\mathbb{R}^m$, and let M be the non-singular matrix that satisfies $M\mathbf{e}^{(i)} = \mathbf{u}^{(i)}$ for $i = 1, 2, \ldots, m$. Then*

$$\mathbf{e}^{(k)} = \sum_{i=1}^{m} r_{i,k}\mathbf{u}^{(i)},$$

where $r_{i,k}$ is the coefficient $(M^{-1})_{i,k}$ in the ith row and kth column of the inverse M^{-1} of the matrix M.

Proof It follows from Lemma 4.3 that $\mathbf{e}^{(k)} = \sum_{i=1}^{m} r_{i,k}\mathbf{u}^{(i)}$, where the coefficients $r_{i,k}$ satisfy $r_{i,k} = (M^{-1}\mathbf{e}^{(k)})_i$ for $i = 1, 2, \ldots, m$. But $M^{-1}\mathbf{e}^{(k)}$ is the column vector whose components are those of the kth column of the matrix M^{-1}. The result follows. ∎

Lemma 4.5 *Let m be a positive integer, let $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(m)}$ be a basis of $\mathbb{R}^m$, let $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \ldots, \mathbf{e}^{(m)}$ denote the standard basis of $\mathbb{R}^m$, let M be the non-singular matrix that satisfies $M\mathbf{e}^{(i)} = \mathbf{u}^{(i)}$ for $i = 1, 2, \ldots, m$, and let $r_{i,k} = (M^{-1})_{i,k}$ for $i = 1, 2, \ldots, m$ and $k = 1, 2, \ldots, m$. Let $g_1, g_2, \ldots, g_m$ be real numbers, and let $\mathbf{p} = \sum_{k=1}^{m} p_k\mathbf{e}^{(k)}$, where $p_k = \sum_{i=1}^{m} g_i r_{i,k}$ for $k = 1, 2, \ldots, m$. Then $\mathbf{p}^T\mathbf{u}^{(i)} = g_i$ for $i = 1, 2, \ldots, m$.*

Proof It follows from the definition of the matrix M that $(\mathbf{u}^{(i)})_k = (M)_{k,i}$ for all integers i and k between 1 and m. It follows that the ith component of the row vector $\mathbf{p}^T M$ is equal to $\mathbf{p}^T\mathbf{u}^{(i)}$ for $i = 1, 2, \ldots, m$. But the definition of the vector $\mathbf{p}$ ensures that p_i is the ith component of the row vector $\mathbf{g}^T M^{-1}$, where $\mathbf{g} \in \mathbb{R}$ is defined so that

$$\mathbf{g}^T = \begin{pmatrix} g_1 & g_2 & \cdots & g_m \end{pmatrix}.$$

It follows that $\mathbf{p}^T = \mathbf{g}^T M^{-1}$, and therefore $\mathbf{p}^T M = \mathbf{g}^T$. Taking the ith component of the row vectors on both sides of this equality, we find that $\mathbf{p}^T\mathbf{u}^{(i)} = g_i$, as required. ∎

Lemma 4.6 *Let m be a positive integer, let*

$$\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(m)} \quad \textit{and} \quad \hat{\mathbf{u}}^{(1)}, \hat{\mathbf{u}}^{(2)}, \ldots, \hat{\mathbf{u}}^{(m)}$$

be bases of $\mathbb{R}^m$, and let $\mathbf{v}$ be an element of $\mathbb{R}^m$. Let h be an integer between 1 and m. Suppose that $\hat{\mathbf{u}}^{(h)} = \sum_{i=1}^{m} \mu_i \mathbf{u}^{(i)}$, where $\mu_1, \mu_2, \ldots, \mu_m$ are real numbers, and that $\mathbf{u}^{(i)} = \hat{\mathbf{u}}^{(i)}$ for all integers i between 1 and m satisfying $i \neq h$. Let $\mathbf{v} = \sum_{i=1}^{m} \lambda_i \mathbf{u}^{(i)} = \sum_{i=1}^{m} \hat{\lambda}_i \hat{\mathbf{u}}^{(i)}$. where $\lambda_i \in \mathbb{R}$ and $\hat{\lambda}_i \in \mathbb{R}$ for $i = 1, 2, \ldots, m$. Then

$$\hat{\lambda}^{(i)} = \begin{cases} \dfrac{1}{\mu_h} \lambda_h & \textit{if } i = h; \\ \lambda_i - \dfrac{\mu_i}{\mu_h} \lambda_h & \textit{if } i \neq h. \end{cases}$$

Proof From the representation of $\hat{\mathbf{u}}^{(h)}$ as a linear combination of the basis vectors $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(m)}$ we find that

$$\frac{1}{\mu_h} \hat{\mathbf{u}}^{(h)} = \mathbf{u}^{(h)} + \sum_{\substack{1 \leq i \leq m \\ i \neq h}} \frac{\mu_i}{\mu_h} \mathbf{u}^{(i)}.$$

Moreover $\hat{\mathbf{u}}^{(i)} = \mathbf{u}^{(i)}$ for all integers i between 1 and m satisfying $i \neq h$. It follows that

$$\mathbf{u}^{(h)} = \frac{1}{\mu_h} \hat{\mathbf{u}}^{(h)} - \sum_{\substack{1 \leq i \leq m \\ i \neq h}} \frac{\mu_i}{\mu_h} \hat{\mathbf{u}}^{(i)}.$$

It follows that

$$\begin{aligned} \sum_{i=1}^{n} \hat{\lambda}_i \hat{\mathbf{u}}^{(i)} &= \mathbf{v} = \sum_{i=1}^{n} \lambda_i \mathbf{u}^{(i)} \\ &= \sum_{\substack{1 \leq i \leq m \\ i \neq h}} \lambda_i \hat{\mathbf{u}}^{(i)} + \frac{1}{\mu_h} \lambda_h \hat{\mathbf{u}}^{(h)} - \sum_{\substack{1 \leq i \leq m \\ i \neq h}} \frac{\mu_i}{\mu_h} \lambda_h \mathbf{u}^{(i)}. \\ &= \sum_{\substack{1 \leq i \leq m \\ i \neq h}} \left(\lambda_i - \frac{\mu_i}{\mu_h} \lambda_h \right) \hat{\mathbf{u}}^{(i)} + \frac{1}{\mu_h} \lambda_h \hat{\mathbf{u}}^{(h)} \end{aligned}$$

Therefore, equating coefficients of $\hat{\mathbf{u}}^{(i)}$ for $i = 1, 2, \ldots, n$, we find that

$$\hat{\lambda}^{(i)} = \begin{cases} \dfrac{1}{\mu_h} \lambda_h & \text{if } i = h, \\ \lambda_i - \dfrac{\mu_i}{\mu_h} \lambda_h & \text{if } i \neq h, \end{cases}$$

as required. ∎

4.8 The Extended Simplex Tableau for solving Linear Programming Problems

We now consider the construction of a tableau for a linear programming problem in Dantzig standard form. Such a problem is specified by an $m \times n$ matrix A, an m-dimensional target vector $\mathbf{b} \in \mathbb{R}^m$ and an n-dimensional cost vector $\mathbf{c} \in \mathbb{R}^n$. We suppose moreover that the matrix A is of rank m. We consider procedures for solving the following linear program in Danzig standard form.

> *Determine* $\mathbf{x} \in \mathbb{R}^n$ *so as to minimize* $\mathbf{c}^T\mathbf{x}$ *subject to the constraints* $A\mathbf{x} = \mathbf{b}$ *and* $\mathbf{x} \geq \mathbf{0}$.

We denote by $A_{i,j}$ the component of the matrix A in the ith row and jth column, we denote by b_i the ith component of the target vector $\mathbf{b}$ for $i = 1, 2, \ldots, m$, and we denote by c_j the jth component of the cost vector $\mathbf{c}$ for $j = 1, 2, \ldots, n$.

We recall that a feasible solution to this problem consists of an n-dimensional vector $\mathbf{x}$ that satisfies the constraints $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$ (see Subsection 4.2). A feasible solution of the linear programming problem then consists of non-negative real numbers $x_1, x_2, \ldots, x_n$ for which

$$\sum_{j=1}^{n} x_j \mathbf{a}^{(j)} = \mathbf{b}.$$

A feasible solution determined by $x_1, x_2, \ldots, x_n$ is optimal if it minimizes cost $\sum\limits_{j=1}^{n} c_j x_j$ amongst all feasible solutions to the linear programming problem.

Let $j_1, j_2, \ldots, j_m$ be distinct integers between 1 and n that are the elements of a basis B for the linear programming problem. Then the vectors $\mathbf{a}^{(j)}$ for $j \in B$ constitute a basis of the real vector space $\mathbb{R}^m$. (see Subsection 4.4).

We denote by M_B the invertible $m \times m$ matrix whose component $(M)_{i,k}$ in the ith row and jth column satisfies $(M_B)_{i,k} = (A)_{i,j_k}$ for $i, k = 1, 2, \ldots, m$. Then the kth column of the matrix M_B is specified by the column vector $\mathbf{a}^{(j_k)}$ for $k = 1, 2, \ldots, m$, and thus the columns of the matrix M_B coincide with those columns of the matrix A that are determined by elements of the basis B.

Every vector in $\mathbb{R}^m$ can be expressed as a linear combination of the basis vectors $\mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}, \ldots, \mathbf{a}^{(j_m)}$. It follows that there exist uniquely determined real numbers $t_{i,j}$ and s_i for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$ such that

$$\mathbf{a}^{(j)} = \sum_{i=1}^{m} t_{i,j} \mathbf{a}^{(j_i)} \quad \text{and} \quad \mathbf{b} = \sum_{i=1}^{m} s_i \mathbf{a}^{(j_i)}.$$

It follows from Lemma 4.3 that

$$t_{i,j} = (M_B^{-1}\mathbf{a}^{(j)})_i \quad \text{and} \quad s_i = (M_B^{-1}\mathbf{b})_i$$

for $i = 1, 2, \ldots, m$.

The standard basis $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \ldots, \mathbf{e}^{(m)}$ of $\mathbb{R}^m$ is defined such that

$$(\mathbf{e}^{(k)})_i = \begin{cases} 1 & \text{if } k = i; \\ 0 & \text{if } k \neq i. \end{cases}$$

It follows from Lemma 4.4 that

$$\mathbf{e}^{(k)} = \sum_{i=1}^{m} r_{i,k}\mathbf{u}^{(i)},$$

where $r_{i,k}$ is the coefficient $(M_B^{-1})_{i,k}$ in the ith row and kth column of the inverse M_B^{-1} of the matrix M_B.

We can record the coefficients of the m-dimensional vectors

$$\mathbf{a}^{(1)},\ \mathbf{a}^{(2)}, \ldots,\ \mathbf{a}^{(n)},\ \mathbf{b},\ \mathbf{e}^{(1)},\ \mathbf{e}^{(2)}, \ldots,\ \mathbf{e}^{(m)}$$

with respect to the basis $\mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}, \ldots, \mathbf{a}^{(j_m)}$, of $\mathbb{R}^m$ in a tableau of the following form:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\cdots$	$\mathbf{a}^{(n)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\cdots$	$\mathbf{e}^{(m)}$
$\mathbf{a}^{(j_1)}$	$t_{1,1}$	$t_{1,2}$	$\cdots$	$t_{1,n}$	s_1	$r_{1,1}$	$r_{1,2}$	$\cdots$	$r_{1,m}$
$\mathbf{a}^{(j_2)}$	$t_{2,1}$	$t_{2,2}$	$\cdots$	$t_{2,n}$	s_2	$r_{2,1}$	$r_{2,2}$	$\cdots$	$r_{2,m}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\mathbf{a}^{(j_m)}$	$t_{m,1}$	$t_{m,2}$	$\cdots$	$t_{m,n}$	s_m	$r_{m,1}$	$r_{m,2}$	$\cdots$	$r_{m,m}$
	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$

The definition of the quantities $t_{i,j}$ ensures that

$$t_{i,j_k} = \begin{cases} 1 & \text{if } i = k; \\ 0 & \text{if } i \neq k. \end{cases}$$

Also

$$t_{i,j} = \sum_{k=1}^{m} r_{i,k}A_{i,j}$$

for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, and

$$s_i = \sum_{k=1}^{m} r_{i,k}b_k$$

for $i = 1, 2, \ldots, m$.

If the quantities $s_1, s_2, \ldots, s_m$ are all non-negative then they determine a basic feasible solution $\mathbf{x}$ of the linear programming problem associated with the basis B with components $x_1, x_2, \ldots, x_n$, where $x_{j_i} = s_i$ for $i = 1, 2, \ldots, m$ and $x_j = 0$ for all integers j between 1 and n that do not belong to the basis B. Indeed

$$\sum_{j=1}^{n} x_j \mathbf{a}^{(j)} = \sum_{i=1}^{m} x_{j_i} \mathbf{a}^{(j_i)} = \sum_{i=1}^{m} s_i \mathbf{a}^{(j_i)}.$$

The *cost* C of the basic feasible solution $\mathbf{x}$ is defined to be the value $\overline{c}^T \mathbf{x}$ of the objective function. The definition of the quantities $s_1, s_2, \ldots, s_m$ ensures that

$$C = \sum_{j=1}^{n} c_j x_j = \sum_{i=1}^{m} c_{j_i} s_i.$$

If the quantities $s_1, s_2, \ldots, s_n$ are not all non-negative then there is no basic feasible solution associated with the basis B.

The criterion row at the bottom of the tableau has cells to record quantities $p_1, p_2, \ldots, p_m$ associated with the vectors that constitute the standard basis $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \ldots, \mathbf{e}^{(m)}$ of $\mathbb{R}^m$. These quantities are defined so that

$$p_k = \sum_{i=1}^{m} c_{j_i} r_{i,k}$$

for $k = 1, 2, \ldots, m$, where c_{j_i} is the cost associated with the basis vector $\mathbf{a}^{(j_i)}$ for $i = 1, 2, \ldots, k$,

An application of Lemma 4.5 establishes that

$$\sum_{k=1}^{m} p_k A_{k,j_i} = c_{j_i}$$

for $i = 1, 2, \ldots, k$.

On combining the identities

$$s_i = \sum_{k=1}^{m} r_{i,k} b_k, \quad p_k = \sum_{i=1}^{m} c_{j_i} r_{i,k} \quad \text{and} \quad C = \sum_{i=1}^{m} c_{j_i} s_i$$

derived above, we find that

$$C = \sum_{i=1}^{m} c_{j_i} s_i = \sum_{i=1}^{m} \sum_{k=1}^{m} c_{j_i} r_{i,k} b_k = \sum_{k=1}^{m} p_k b_k.$$

The tableau also has cells in the criterion row to record quantities

$$-q_1, -q_2, \ldots, -q_n,$$

where $q_1, q_2, \ldots, q_n$ are the components of the unique n-dimensional vector $\mathbf{q}$ characterized by the following properties:

- $q_{j_i} = 0$ for $i = 1, 2, \ldots, m$;
- $\mathbf{c}^T\overline{\mathbf{x}} = C + \mathbf{q}^T\overline{\mathbf{x}}$ for all $\overline{\mathbf{x}} \in \mathbb{R}^m$ satisfying the matrix equation $A\overline{\mathbf{x}} = \mathbf{b}$.

First we show that if $\mathbf{q} \in \mathbb{R}^n$ is defined such that $\mathbf{q}^T = \mathbf{c}^T - \mathbf{p}^T A$ then the vector $\mathbf{q}$ has the required properties.

The definition of $p_1, p_2, \ldots, p_k$ ensures (as a consequence of Lemma 4.5, as noted above) that

$$\sum_{k=1}^{m} p_k A_{k,j_i} = c_{j_i}$$

for $i = 1, 2, \ldots, k$. It follows that

$$q_{j_i} = c_{j_i} - (\mathbf{p}^T A)_{j_i} = c_{j_i} - \sum_{k=1}^{m} p_k A_{k,j_i} = 0$$

for $i = 1, 2, \ldots, n$.

Also $\mathbf{p}^T\mathbf{b} = C$. It follows that if $\overline{\mathbf{x}} \in \mathbb{R}^n$ satisfies $A\overline{\mathbf{x}} = \mathbf{b}$ then

$$\mathbf{c}^T\overline{\mathbf{x}} = \mathbf{p}^T A\overline{\mathbf{x}} + \mathbf{q}^T\overline{\mathbf{x}} = \mathbf{p}^T\mathbf{b} + \mathbf{q}^T\overline{\mathbf{x}} = C + \mathbf{q}^T\overline{\mathbf{x}}.$$

Thus if $\mathbf{q}^T = \mathbf{c}^T - \mathbf{p}^T A$ then the vector $\mathbf{q}$ satisfies the properties specified above.

We next show that

$$(\mathbf{p}^T A)_j = \sum_{i=1}^{m} c_{j_i} t_{i,j}$$

for $j = 1, 2, \ldots, n$.

Now

$$t_{i,j} = \sum_{k=1}^{m} r_{i,k} A_{k,j}$$

for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. (This is a consequence of the identities

$$\mathbf{a}^{(j)} = \sum_{k=1}^{m} A_{k,j}\mathbf{e}^{(k)} = \sum_{i=1}^{m}\sum_{k=1}^{m} r_{i,k} A_{k,j}\mathbf{a}^{(j_i)},$$

as noted earlier.)

Also the definition of p_k ensures that

$$p_k = \sum_{i=1}^{m} c_{j_i} r_{i,k}$$

for $k = 1, 2, \ldots, m$. These results ensure that

$$\sum_{i=1}^{m} c_{j_i} t_{i,j} = \sum_{i=1}^{m} \sum_{k=1}^{m} c_{j_i} r_{i,k} A_{k,j} = \sum_{k=1}^{m} p_k A_{k,j} = (\mathbf{p}^T A)_j.$$

It follows that

$$-q_j = \sum_{k=1}^{m} p_k A_{k,j} - c_j = \sum_{i=1}^{m} c_{j_i} t_{i,k} - c_j$$

for $j = 1, 2, \ldots, n$.

The *extended simplex tableau* associated with the basis B is obtained by entering the values of the quantities $-q_j$ (for $j = 1, 2, \ldots, n$), C and p_k (for $k = 1, 2, \ldots, m$) into the bottom row to complete the tableau described previously. The extended simplex tableau has the following structure:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\cdots$	$\mathbf{a}^{(n)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$	$\cdots$	$\mathbf{e}^{(m)}$
$\mathbf{a}^{(j_1)}$	$t_{1,1}$	$t_{1,2}$	$\cdots$	$t_{1,n}$	s_1	$r_{1,1}$	$r_{1,2}$	$\cdots$	$r_{1,m}$
$\mathbf{a}^{(j_2)}$	$t_{2,1}$	$t_{2,2}$	$\cdots$	$t_{2,n}$	s_2	$r_{2,1}$	$r_{2,2}$	$\cdots$	$r_{2,m}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\mathbf{a}^{(j_m)}$	$t_{m,1}$	$t_{m,2}$	$\cdots$	$t_{m,n}$	s_m	$r_{m,1}$	$r_{m,2}$	$\cdots$	$r_{m,m}$
	$-q_1$	$-q_2$	$\cdots$	$-q_n$	C	p_1	p_2	$\cdots$	p_m

The extended simplex tableau can be represented in block form as follows:—

	$\mathbf{a}^{(1)} \;\cdots\; \mathbf{a}^{(n)}$	$\mathbf{b}$	$\mathbf{e}^{(1)} \;\cdots\; \mathbf{e}^{(m)}$
$\mathbf{a}^{(j_1)}$ $\vdots$ $\mathbf{a}^{(j_m)}$	$M_B^{-1} A$	$M_B^{-1}\mathbf{b}$	M_B^{-1}
	$\mathbf{p}^T A - \mathbf{c}^T$	$\mathbf{p}^T\mathbf{b}$	$\mathbf{p}^T$

Let $\mathbf{c}_B$ denote the m-dimensional vector defined so that

$$\mathbf{c}_B^T = \begin{pmatrix} c_{j_1} & c_{j_2} & \cdots & c_{j_m} \end{pmatrix}.$$

The identities we have verified ensure that the extended simplex tableau can therefore also be represented in block form as follows:—

	$\mathbf{a}^{(1)} \quad \cdots \quad \mathbf{a}^{(n)}$	$\mathbf{b}$	$\mathbf{e}^{(1)} \quad \cdots \quad \mathbf{e}^{(m)}$
$\mathbf{a}^{(j_1)}$ $\vdots$ $\mathbf{a}^{(j_m)}$	$M_B^{-1}A$	$M_B^{-1}\mathbf{b}$	M_B^{-1}
	$\mathbf{c}_B^T M_B^{-1}A - \mathbf{c}^T$	$\mathbf{c}_B^T M_B^{-1}\mathbf{b}$	$\mathbf{c}_B^T M_B^{-1}$

Given an $m \times n$ matrix A of rank m, an m-dimensional target vector $\mathbf{b}$, and an n-dimensional cost vector $\mathbf{c}$, there exists an extended simplex tableau associated with any basis B for the linear programming problem, irrespective of whether or not there exists a basic feasible solution associated with the given basis B.

The crucial requirement that enables the construction of the tableau is that the basis B should consist of m distinct integers $j_1, j_2, \ldots, j_m$ between 1 and m for which the corresponding columns of the matrix A constitute a basis of the vector space $\mathbb{R}^m$.

A basis B is associated with a basic feasible solution of the linear programming problem if and only if the values in the column labelled by the target vector $\mathbf{b}$ and the rows labelled by $\mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}, \ldots, \mathbf{a}^{(j_m)}$ should be non-negative. If so, those values will include the non-zero components of the basic feasible solution associated with the basis.

If there exists a basic feasible solution associated with the basis B then that solution is optimal if and only if all the values in the criterion row in the columns labelled by $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ are all non-positive.

Versions of the Simplex Tableau Algorithm for determining a basic optimal solution to the linear programmming problem, given an initial basic feasible solution, rely on the transformation rules that determine how the values in the body of the extended simplex tableau are transformed on passing from an old basis B to an new basis B', where the new basis B' contains all but one of the members of the old basis B.

Let us refer to the rows of the extended simplex tableau labelled by the basis vectors $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ as the *basis rows* of the tableau.

Lemma 4.6 determines how entries in the basis rows of the extended simplex tableau transform which one element of the basis is replaced by an element not belonging to the basis.

Let the old basis B consist of distinct integers $j_1, j_2, \ldots, j_m$ between 1 and n, and let the new basis B' also consist of distinct integers $j'_1, j'_2, \ldots, j'_m$ between 1 and n. We suppose that the new basis B' is obtained from the old basis by replacing an element j_h of the old basis B by some integer j'_h between 1 and n that does not belong to the old basis. We suppose therefore that $j_i = j'_i$ when $i \neq h$, and that j'_h is some integer between 1 and n that does not belong to the basis B.

Let the coefficients $t_{i,j}$, $t'_{i,j}$, s_i, s'_i, $r_{i,k}$ and $r'_{i,k}$ be determined for $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$ and $k = 1, 2, \ldots, m$ so that

$$\mathbf{a}^{(j)} = \sum_{i=1}^{m} t_{i,j} \mathbf{a}^{(j_i)} = \sum_{i=1}^{m} t'_{i,j} \mathbf{a}^{(j'_i)}$$

for $j = 1, 2, \ldots, n$,

$$\mathbf{b} = \sum_{i=1}^{m} s_i \mathbf{a}^{(j_i)} = \sum_{i=1}^{m} s'_i \mathbf{a}^{(j'_i)}$$

and

$$\mathbf{e}^{(k)} = \sum_{i=1}^{m} r_{i,k} \mathbf{a}^{(j_i)} = \sum_{i=1}^{m} r'_{i,k} \mathbf{a}^{(j'_i)}$$

for $k = 1, 2, \ldots, m$.

It then follows from Lemma 4.6 that

$$\begin{aligned}
t'_{h,j} &= \frac{1}{t_{h,j'_h}} \, t_{h,j}, \\
t'_{i,j} &= t_{i,j} - \frac{t_{i,j'_h}}{t_{h,j'_h}} \, t_{h,j} \quad (i \neq h). \\
s'_h &= \frac{1}{t_{h,j'_h}} \, s_h, \\
s'_i &= s_i - \frac{t_{i,j'_h}}{t_{h,j'_h}} \, s_h \quad (i \neq h), \\
r'_{h,k} &= \frac{1}{t_{h,j'_h}} \, r_{h,k}, \\
r'_{i,k} &= r_{i,k} - \frac{t_{i,j'_h}}{t_{h,j'_h}} \, r_{h,k} \quad (i \neq h).
\end{aligned}$$

The *pivot row* of the extended simplex tableau for this change of basis from B to B' is the row labelled by the basis vector $\mathbf{a}^{(j_h)}$ that is to be removed from the current basis. The *pivot column* of the extended simplex tableau for this change of basis is the column labelled by the vector $\mathbf{a}^{(j'_h)}$ that is to be added to the current basis. The *pivot element* for this change of basis is the element t_{h,j'_h} of the tableau located in the pivot row and pivot column of the tableau.

The identities relating the components of $\mathbf{a}^{(j)}$, $\mathbf{b}$ and $\mathbf{e}^{(k)}$ with respect to the old basis to the components of those vectors with respect to the new basis ensure that the rules for transforming the rows of the tableau other than the criterion row can be stated as follows:—

- a value recorded in the pivot row is transformed by dividing it by the pivot element;
- an value recorded in a basis row other than the pivot row is transformed by substracting from it a constant multiple of the value in the same column that is located in the pivot row, where this constant multiple is the ratio of the values in the basis row and pivot row located in the pivot column.

In order to complete the discussion of the rules for transforming the values recorded in the extended simplex tableau under a change of basis that replaces an element of the old basis by an element not in that basis, it remains to analyse the rule that determines how the elements of the criterion row are transformed under this change of basis.

First we consider the transformation of the elements of the criterion row in the columns labelled by $\mathbf{a}^{(j)}$ for $j = 1, 2, \ldots, n$. Now the coefficients $t_{i,j}$ and $t'_{i,j}$ are defined for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$ so that

$$\mathbf{a}^{(j)} = \sum_{i=1}^{m} t_{i,j}\mathbf{a}^{(j_i)} = \sum_{i=1}^{m} t'_{i,j}\mathbf{a}^{(j'_i)},$$

where $j_1 = j'_1 = 1$, $j_3 = j'_3 = 3$, $j_2 = 2$ and $j'_2 = 4$. Moreover

$$t'_{h,j} = \frac{1}{t_{h,j'_h}}\, t_{h,j}$$

and

$$t'_{i,j} = t_{i,j} - \frac{t_{i,j'_h}}{t_{h,j'_h}}\, t_{h,j}$$

for all integers i between 1 and m for which $i \neq h$.

Now

$$-q_j = \sum_{i=1}^{m} c_{j_i} t_{i,j} - c_j$$

and

$$-q'_j = \sum_{i=1}^{m} c_{j'_i} t'_{i,j} - c_j.$$

Therefore

$$q_j - q'_j \quad = \sum_{\substack{1 \le i \le m \\ i \neq h}} c_{j_i}(t'_{i,j} - t_{i,j}) + c_{j'_h} t'_{h,j} - c_{j_h} t_{h,j}$$

$$
\begin{aligned}
&= \frac{1}{t_{h,j'_h}} \left(-\sum_{i=1}^{m} c_{j_i} t_{i,j'_h} + c_{j'_h} \right) t_{h,j} \\
&= \frac{q_{j'_h}}{t_{h,j'_h}} \, t_{h,j}
\end{aligned}
$$

and thus

$$-q'_j = -q_j + \frac{q_{j'_h}}{t_{h,j'_h}} \, t_{h,j}$$

for $j = 1, 2, \ldots, n$.

Next we note that

$$C = \sum_{i=1}^{m} c_{j_i} s_i$$

and

$$C' = \sum_{i=1}^{m} c_{j'_i} s'_i.$$

Therefore

$$
\begin{aligned}
C' - C &= \sum_{\substack{1 \leq i \leq m \\ i \neq h}} c_{j_i}(s'_i - s_i) + c_{j'_h} s'_h - c_{j_h} s_h \\
&= \frac{1}{t_{h,j'_h}} \left(-\sum_{i=1}^{m} c_{j_i} t_{i,j'_h} + c_{j'_h} \right) s_h \\
&= \frac{q_{j'_h}}{t_{h,j'_h}} \, s_h
\end{aligned}
$$

and thus

$$C' = q_k + \frac{q_{j'_h}}{t_{h,j'_h}} \, s_h$$

for $k = 1, 2, \ldots, m$.

To complete the verification that the criterion row of the extended simplex tableau transforms according to the same rule as the other rows we note that

$$p_k = \sum_{i=1}^{m} c_{j_i} r_{i,k}$$

and

$$p'_k = \sum_{i=1}^{m} c_{j'_i} r'_{i,k}.$$

Therefore

$$\begin{aligned} p'_k - p_k &= \sum_{\substack{1\le i\le m\\ i\ne h}} c_{j_i}(r'_{i,k} - r_{i,k}) + c_{j'_h} r'_{h,k} - c_{j_h} r_{h,k} \\ &= \frac{1}{t_{h,j'_h}} \left(-\sum_{i=1}^m c_{j_i} t_{i,j'_h} + c_{j'_h} \right) r_{h,k} \\ &= \frac{q_{j'_h}}{t_{h,j'_h}} r_{h,k} \end{aligned}$$

and thus

$$p'_k = p_k + \frac{q_{j'_h}}{t_{h,j'_h}} r_{h,k}$$

for $k = 1, 2, \ldots, m$.

We conclude that the criterion row of the extended simplex tableau transforms under changes of basis that replace one element of the basis according to a rule analogous to that which applies to the basis rows. Indeed an element of the criterion row is transformed by subtracting from it a constant multiple of the element in the pivot row that belongs to the same column, where the multiplying factor is the ratio of the elements in the criterion row and pivot row of the pivot column.

We have now discussed how the extended simplex tableau associated with a given basis B is constructed from the constraint matrix A, target vector $\mathbf{b}$ and cost vector $\mathbf{c}$ that characterizes the linear programming problem. We have also discussed how the tableau transforms when one element of the given basis is replaced.

It remains how to replace an element of a basis associated with a non-optimal feasible solution so as to obtain a basic feasible solution of lower cost where this is possible.

We use the notation previously established. Let $j_1, j_2, \ldots, j_m$ be the elements of a basis B that is associated with some basic feasible solution of the linear programming problem. Then there are non-negative numbers $s_1, s_2, \ldots, s_m$ such that

$$\mathbf{b} = \sum_{i=1}^m s_i \mathbf{a}^{(j_i)},$$

where $\mathbf{a}^{(j_i)}$ is the m-dimensional vector determined by column j_i of the constraint matrix A.

Let j_0 be an integer between 1 and n that does not belong to the basis B. Then

$$\mathbf{a}^{(j_0)} - \sum_{i=1}^m t_{i,j_0} \mathbf{a}^{(j_i)} = \mathbf{0}.$$

and therefore

$$\lambda \mathbf{a}^{(j_0)} + \sum_{i=1}^{m} (s_i - \lambda t_{i,j_0}) \mathbf{a}^{(j_i)} = \mathbf{b}.$$

This expression representing $\mathbf{b}$ as a linear combination of the basis vectors $\mathbf{a}^{(j_0)}, \mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}, \ldots, \mathbf{a}^{(j_m)}$ determines an n-dimensional vector $\overline{\mathbf{x}}(\lambda)$ satisfying the matrix equation $A\overline{\mathbf{x}}(\lambda) = \mathbf{b}$. Let $\overline{x}_j(\lambda)$ denote the jth component of the vector $\overline{\mathbf{x}}(\lambda)$ for $j = 1, 2, \ldots, n$. Then

- $\overline{x}_{j_0}(\lambda) = \lambda$;
- $\overline{x}_{j_i}(\lambda) = s_i - \lambda t_{i,j_0}$ for $i = 1, 2, \ldots, m$;
- $\overline{x}_j = 0$ when $j \notin \{j_0, j_1, j_2, \ldots, j_m\}$.

The n-dimensional vector $\overline{\mathbf{x}}(\lambda)$ represents a feasible solution of the linear programming problem if and only if all its coefficients are non-negative. The cost is then $C + q_{j_0}\lambda$, where C is the cost of the basic feasible solution determined by the basis B.

Suppose that $q_{j_0} < 0$ and that $t_{i,j_0} \leq 0$ for $i = 1, 2, \ldots, m$. Then $\overline{\mathbf{x}}(\lambda)$ is a feasible solution with cost $C + q_{j_0}\lambda$ for all non-negative real numbers λ. In this situation there is no optimal solution to the linear programming problem, because, given any real number K, it is possible to choose λ so that $C + q_{j_0}\lambda < K$, thereby obtaining a feasible solution whose cost is less than K.

If there does exist an optimal solution to the linear programming problem then there must exist at least one integer i between 1 and m for which $t_{i,j_0} > 0$. We suppose that this is the case. Then $\overline{\mathbf{x}}(\lambda)$ is a feasible solution if and only if λ satisfies $0 \leq \lambda \leq \lambda_0$, where

$$\lambda_0 = \text{minimum} \left(\frac{s_i}{t_{i,j_0}} : t_{i,j_0} > 0 \right).$$

We can then choose some integer h between 1 and n for which

$$\frac{s_h}{t_{h,j_0}} = \lambda_0.$$

Let $j'_i = j_i$ for $i \neq h$, and let $j'_h = j_0$, and let $B' = \{j'_1, j'_2, \ldots, j'_m\}$. Then $\overline{\mathbf{x}}(\lambda_0)$ is a basic feasible solution of the linear programming problem associated with the basis B'. The cost of this basic feasible solution is

$$C + \frac{s_h q_{j_0}}{t_{h,j_0}}.$$

It makes sense to select the replacement column so as to obtain the greatest cost reduction. The procedure for finding this information from the tableau can be described as follows.

We suppose that the simplex tableau for a basic feasible solution has been prepared. Examine the values in the criterion row in the columns labelled by $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$. If all those are non-positive then the basic feasible solution is optimal. If not, then consider in turn those columns $\mathbf{a}^{(j_0)}$ for which the value $-q_{j_0}$ in the criterion row is positive. For each of these columns, examine the coefficients recorded in the column in the basis rows. If these coefficients are all non-positive then there is no optimal solution to the linear programming problem. Otherwise choose h to be the value of i that minimizes the ratio $\dfrac{s_i}{t_{i,j_0}}$ amongst those values of i for which $t_{i,j_0} > 0$. The row labelled by $\mathbf{a}^{(j_h)}$ would then be the pivot row if the column $\mathbf{a}^{(j_0)}$ were used as the pivot column.

Calculate the value of the cost reduction $\dfrac{s_h(-q_{j_0})}{t_{h,j_0}}$ that would result if the column labelled by $\mathbf{a}^{(j_0)}$ were used as the pivot column. Then choose the pivot column to maximize the cost reduction amongst all columns $\mathbf{a}^{(j_0)}$ for which $-q_{j_0} > 0$. Choose the row labelled by $\mathbf{a}^{(j_h)}$, where h is determined as described above. Then apply the procedures for transforming the simplex tableau to that determined by the new basis B', where B' includes j_0 together with j_i for all integers i between 1 and m satisfying $i \neq h$.

4.9 The Simplex Tableau Algorithm

In describing the Simplex Tableau Algorithm, we adopt notation previously introduced. Thus we are concerned with the solution of a linear programming problem in Dantzig standard form, specified by positive integers m and n, an $m \times n$ constraint matrix A of rank m, a target vector $\mathbf{b} \in \mathbb{R}^m$ and a cost vector $\mathbf{c} \in \mathbb{R}^n$. The optimization problem requires us to find a vector $\mathbf{x} \in \mathbb{R}^n$ that minimizes $\mathbf{c}^T\mathbf{x}$ amongst all vectors $\mathbf{x} \in \mathbb{R}^n$ that satisfy the constraints $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$.

We denote by $A_{i,j}$ the coefficient in the ith row and jth column of the matrix A, we denote the ith component of the target vector $\mathbf{b}$ by b_i and we denote the jth component of the cost vector $\mathbf{c}$ by c_j for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

As usual, we define vectors $\mathbf{a}^{(j)} \in \mathbb{R}^m$ for $j = 1, 2, \ldots, n$ such that $(\mathbf{a}^{(j)})_i = A_{i,j}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

Distinct integers $j_1, j_2, \ldots, j_m$ between 1 and n determine a basis B, where

$$B = \{j_1, j_2, \ldots, j_m\},$$

if and only if the corresponding vectors $\mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}, \ldots, \mathbf{a}^{(j_m)}$ constitute a basis of $\mathbb{R}^m$. Given such a basis B we let M_B denote the invertible $m \times m$ matrix defined such that $(M_B)_{i,k} = A_{i,j_k}$ for all integers i and k between 1 and m.

We let $t_{i,j} = (M_B^{-1}A)_{i,j}$ and $s_i = (M_B^{-1}\mathbf{b})_i$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. Then

$$\mathbf{a}^{(j)} = \sum_{i=1}^{m} t_{i,j}\mathbf{a}^{(j_i)}$$

for $j = 1, 2, \ldots, n$, and

$$\mathbf{b} = \sum_{i=1}^{m} s_i \mathbf{a}^{(j_i)}.$$

A basis B determines an associated basic feasible solution if and only if $s_i \geq 0$ for $i = 1, 2, \ldots, m$. We suppose in what follows that the basis B determines a basic feasible solution.

Let

$$C = \sum_{i=1}^{m} c_{j_i} s_i.$$

Then C is the cost of the basic feasible solution associated with the basis B.

Let

$$-q_j = \sum_{i=1}^{m} c_{j_i} t_{i,j} - c_j.$$

Then $q_j = 0$ for all $j \in \{j_1, j_2, \ldots, j_m\}$. Also the cost of any feasible solution $(\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n)$ of the linear programming problem is

$$C + \sum_{j=1}^{n} q_j \overline{x}_j.$$

The *simplex tableau* associated with the basis B is that portion of the extended simplex tableau that omits the columns labelled by $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \ldots, \mathbf{e}^{(m)}$. The simplex table has the following structure:

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\cdots$	$\mathbf{a}^{(n)}$	$\mathbf{b}$
$\mathbf{a}^{(j_1)}$	$t_{1,1}$	$t_{1,2}$	$\cdots$	$t_{1,n}$	s_1
$\mathbf{a}^{(j_2)}$	$t_{2,1}$	$t_{2,2}$	$\cdots$	$t_{2,n}$	s_2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$\mathbf{a}^{(j_m)}$	$t_{m,1}$	$t_{m,2}$	$\cdots$	$t_{m,n}$	s_m
	$-q_1$	$-q_2$	$\cdots$	$-q_n$	C

Let $\mathbf{c}_B$ denote the m-dimensional vector defined such that

$$\mathbf{c}_B^T = \begin{pmatrix} c_{j_1} & c_{j_2} & \cdots & c_{j_m} \end{pmatrix}.$$

Then the simplex tableau can be represented in block form as follows:—

<table>
<tr><td></td><td>$\mathbf{a}^{(1)} \quad \cdots \quad \mathbf{a}^{(n)}$</td><td>$\mathbf{b}$</td></tr>
<tr><td>$\mathbf{a}^{(j_1)}$
$\vdots$
$\mathbf{a}^{(j_m)}$</td><td>$M_B^{-1}A$</td><td>$M_B^{-1}\mathbf{b}$</td></tr>
<tr><td></td><td>$\mathbf{c}_B^T M_B^{-1} A - \mathbf{c}^T$</td><td>$\mathbf{c}_B^T M_B^{-1}\mathbf{b}$</td></tr>
</table>

Example We consider again the following linear programming problem:—

minimize

$$3x_1 + 4x_2 + 2x_3 + 9x_4 + 5x_5$$

subject to the following constraints:
$5x_1 + 3x_2 + 4x_3 + 7x_4 + 3x_5 = 11$;
$4x_1 + x_2 + 3x_3 + 8x_4 + 4x_5 = 6$;
$x_j \geq 0$ *for* $j = 1, 2, 3, 4, 5$.

We are given the following initial basic feasible solution $(1, 2, 0, 0, 0)$. We need to determine whether this initial basic feasible solution is optimal and, if not, how to improve it till we obtain an optimal solution.

The constraints require that x_1, x_2, x_3, x_4, x_5 be non-negative real numbers satisfying the matrix equation

$$\begin{pmatrix} 5 & 3 & 4 & 7 & 3 \\ 4 & 1 & 3 & 8 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 11 \\ 6 \end{pmatrix}.$$

Thus we are required to find a (column) vector $\mathbf{x}$ with components x_1, x_2, x_3, x_4 and x_5 that maximizes $\mathbf{c}^T\mathbf{x}$ subject to the constraints $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$, where

$$A = \begin{pmatrix} 5 & 3 & 4 & 7 & 3 \\ 4 & 1 & 3 & 8 & 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 11 \\ 6 \end{pmatrix},$$

and

$$\mathbf{c}^T = \begin{pmatrix} 3 & 4 & 2 & 9 & 5 \end{pmatrix}.$$

Our initial basis B satisfies $B = \{j_1, j_2\}$, where $j_1 = 1$ and $j_2 = 2$. The first two columns of the matrix A provide the corresponding invertible 2×2 matrix M_B. Thus

$$M_B = \begin{pmatrix} 5 & 3 \\ 4 & 1 \end{pmatrix}.$$

Inverting this matrix, we find that

$$M_B^{-1} = -\frac{1}{7} \begin{pmatrix} 1 & -3 \\ -4 & 5 \end{pmatrix}.$$

For each integer j between 1 and 5, let $\mathbf{a}^{(j)}$ denote the m-dimensional vector whose ith component is $A_{i,j}$ for $i = 1, 2$. Then

$$\mathbf{a}^{(j)} = \sum_{i=1}^{2} t_{i,j} \mathbf{a}^{(j_i)} \quad \text{and} \quad \mathbf{b} = \sum_{i=1}^{2} s_i \mathbf{a}^{(j_i)},$$

where $t_{i,j} = (M_B^{-1}A)_{i,j}$ and $s_i = (M_B^{-1}\mathbf{b})_i$ for $j = 1, 2, 3, 4, 5$ and $i = 1, 2$.

Calculating $M_B^{-1}A$ we find that

$$M_B^{-1}A = \begin{pmatrix} 1 & 0 & \frac{5}{7} & \frac{17}{7} & \frac{9}{7} \\ 0 & 1 & \frac{1}{7} & -\frac{12}{7} & -\frac{8}{7} \end{pmatrix}.$$

Also

$$M_B^{-1}\mathbf{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

The coefficients of these matrices determine the values of $t_{i,j}$ and s_i to be entered into the appropriate cells of the simplex tableau.

The basis rows of the simplex tableau corresponding to the basis $\{1, 2\}$ are thus as follows:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$
$\mathbf{a}^{(1)}$	1	0	$\frac{5}{7}$	$\frac{17}{7}$	$\frac{9}{7}$	1
$\mathbf{a}^{(2)}$	0	1	$\frac{1}{7}$	$-\frac{12}{7}$	$-\frac{8}{7}$	2
	·	·	·	·	·	·

Now the cost C of the current feasible solution satisfies the equation

$$C = \sum_{i=1}^{2} c_{j_i} s_i = c_1 s_1 + c_2 s_2,$$

where $c_1 = 3$, $c_2 = 4$, $s_1 = 1$ and $s_2 = 2$. It follows that $C = 11$.

To complete the simplex tableau, we need to compute $-q_j$ for $j = 1, 2, 3, 4, 5$, where

$$-q_j = \sum_{i=1}^{2} c_{j_i} t_{i,j} - c_j.$$

Let $\mathbf{c}_B$ denote the 2-dimensional vector whose ith component is (c_{j_i}). Then $\mathbf{c}_B = (3, 4)$. Let $\mathbf{q}$ denote the 5-dimensional vector whose jth component is q_j for $j = 1, 2, 3, 4, 5$. Then

$$-\mathbf{q}^T = \mathbf{c}_B^T M_B^{-1} A - \mathbf{c}^T.$$

It follows that

$$\begin{aligned} -\mathbf{q}^T &= \begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 & \frac{5}{7} & \frac{17}{7} & \frac{9}{7} \\ 0 & 1 & \frac{1}{7} & -\frac{12}{7} & -\frac{8}{7} \end{pmatrix} \\ & \qquad - \begin{pmatrix} 3 & 4 & 2 & 9 & 5 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & \frac{5}{7} & -\frac{60}{7} & -\frac{40}{7} \end{pmatrix}. \end{aligned}$$

The simplex tableau corresponding to basis $\{1, 2\}$ is therefore completed as follows:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$
$\mathbf{a}^{(1)}$	1	0	$\frac{5}{7}$	$\frac{17}{7}$	$\frac{9}{7}$	1
$\mathbf{a}^{(2)}$	0	1	$\frac{1}{7}$	$-\frac{12}{7}$	$-\frac{8}{7}$	2
	0	0	$\frac{5}{7}$	$-\frac{60}{7}$	$-\frac{40}{7}$	11

The values of $-q_j$ for $j = 1, 2, 3, 4, 5$ are not all non-positive ensures that the initial basic feasible solution is not optimal. Indeed the cost of a feasible solution $(\overline{x}_1, \overline{x}_2, \overline{x}_3, \overline{x}_4, \overline{x}_5)$ is

$$11 - \tfrac{5}{7}\overline{x}_3 + \tfrac{60}{7}\overline{x}_4 + \tfrac{40}{7}\overline{x}_5.$$

Thus a feasible solution with $\overline{x}_3 > 0$ and $\overline{x}_4 = \overline{x}_5 = 0$ will have lower cost than the initial feasible basic solution. We therefore implement a change of basis whose pivot column is that labelled by the vector $\mathbf{a}^{(3)}$.

We must determine which row to use as the pivot row. We need to determine the value of i that minimizes the ratio $\dfrac{s_i}{t_{i,3}}$, subject to the requirement that $t_{i,3} > 0$. This ratio has the value $\frac{7}{5}$ when $i = 1$ and 14 when $i = 2$. Therefore the pivot row is the row labelled by $\mathbf{a}^{(1)}$. The pivot element $t_{1,3}$ then has the value $\frac{5}{7}$.

The simplex tableau corresponding to basis $\{2, 3\}$ is then obtained by subtracting the pivot row multiplied by $\frac{1}{5}$ from the row labelled by $\mathbf{a}^{(2)}$, subtracting the pivot row from the criterion row, and finally dividing all values in the pivot row by the pivot element $\frac{5}{7}$.

The simplex tableau for the basis $\{2, 3\}$ is thus the following:—

	$\mathbf{a}^{(1)}$	$\mathbf{a}^{(2)}$	$\mathbf{a}^{(3)}$	$\mathbf{a}^{(4)}$	$\mathbf{a}^{(5)}$	$\mathbf{b}$
$\mathbf{a}^{(3)}$	$\frac{7}{5}$	0	1	$\frac{17}{5}$	$\frac{9}{5}$	$\frac{7}{5}$
$\mathbf{a}^{(2)}$	$-\frac{1}{5}$	1	0	$-\frac{11}{5}$	$-\frac{7}{5}$	$\frac{9}{5}$
	-1	0	0	-11	-7	10

All the values in the criterion row to the left of the new cost are non-positive. It follows that we have found a basic optimal solution to the linear programming problem. The values recorded in the column labelled by $\mathbf{b}$ show that this basic optimal solution is

$$(0, \tfrac{9}{5}, \tfrac{7}{5}, 0, 0).$$

4.10 The Revised Simplex Algorithm

The Simplex Tableau Algorithm restricts attention to the columns to the left of the extended simplex tableau. The Revised Simplex Algorithm proceeds by maintaining the columns to the right of the extended simplex tableau, calculating values in the columns to the left of that tableau only as required.

We show how the Revised Simplex Algorithm is implemented by applying it to the example used to demonstrate the implementation of the Simplex Algorithm.

Example We apply the Revised Simplex Algorithm to determine a basic optimal solution to the the following linear programming problem:—

minimize

$$3x_1 + 4x_2 + 2x_3 + 9x_4 + 5x_5$$

subject to the following constraints:
$5x_1 + 3x_2 + 4x_3 + 7x_4 + 3x_5 = 11$*;*
$4x_1 + x_2 + 3x_3 + 8x_4 + 4x_5 = 6$*;*
$x_j \geq 0$ *for* $j = 1, 2, 3, 4, 5$.

We are given the following initial basic feasible solution $(1, 2, 0, 0, 0)$. We need to determine whether this initial basic feasible solution is optimal and, if not, how to improve it till we obtain an optimal solution.

The constraints require that x_1, x_2, x_3, x_4, x_5 be non-negative real numbers satisfying the matrix equation

$$\begin{pmatrix} 5 & 3 & 4 & 7 & 3 \\ 4 & 1 & 3 & 8 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 11 \\ 6 \end{pmatrix}.$$

Thus we are required to find a (column) vector $\mathbf{x}$ with components x_1, x_2, x_3, x_4 and x_5 that maximizes $\mathbf{c}^T\mathbf{x}$ subject to the constraints $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$, where

$$A = \begin{pmatrix} 5 & 3 & 4 & 7 & 3 \\ 4 & 1 & 3 & 8 & 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 11 \\ 6 \end{pmatrix},$$

and

$$\mathbf{c}^T = \begin{pmatrix} 3 & 4 & 2 & 9 & 5 \end{pmatrix}.$$

Our initial basis B satisfies $B = \{j_1, j_2\}$, where $j_1 = 1$ and $j_2 = 2$. The first two columns of the matrix A provide the corresponding invertible 2×2 matrix M_B. Thus

$$M_B = \begin{pmatrix} 5 & 3 \\ 4 & 1 \end{pmatrix}.$$

Inverting this matrix, we find that

$$M_B^{-1} = -\frac{1}{7} \begin{pmatrix} 1 & -3 \\ -4 & 5 \end{pmatrix}.$$

For each integer j between 1 and 5, let $\mathbf{a}^((j)$ denote the m-dimensional vector whose ith component is $A_{i,j}$ for $i = 1, 2$. Then

$$\mathbf{a}^{(j)} = \sum_{i=1}^{2} t_{i,j}\mathbf{a}^{(j_i)} \quad \text{and} \quad \mathbf{b} = \sum_{i=1}^{2} s_i\mathbf{a}^{(j_i)},$$

where $t_{i,j} = (M_B^{-1}A)_{i,j}$ and $s_i = (M_B^{-1}\mathbf{b})_i$ for $j = 1, 2, 3, 4, 5$ and $i = 1, 2$.

Let $r_{i,k} = (M_B^{-1})_{i,k}$ for $i = 1, 2$ and $k = 1, 2$, and let

$$\begin{aligned} C &= c_{j_1}s_1 + c_{j_2}s_2 = c_1s_1 + c_2s_2 = 11 \\ p_1 &= c_{j_1}r_{1,1} + c_{j_2}r_{2,1} = c_1r_{1,1} + c_2r_{2,1} = \tfrac{13}{7} \\ p_2 &= c_{j_1}r_{1,2} + c_{j_2}r_{2,2} = c_1r_{1,2} + c_2r_{2,2} = -\tfrac{11}{7} \end{aligned}$$

The values of s_i, $r_{i,k}$, C and p_k are inserted into the following tableau, which consists of the columns to the right of the extended simplex tableau:—

	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$
$\mathbf{a}^{(1)}$	1	$-\frac{1}{7}$	$\frac{3}{7}$
$\mathbf{a}^{(2)}$	2	$\frac{4}{7}$	$-\frac{5}{7}$
	11	$\frac{13}{7}$	$-\frac{11}{7}$

To proceed with the algorithm, one computes values $-q_j$ for $j \notin B$ using the formula

$$-q_j = p_1 A_{1,j} + p_2 A_{2,j} - c_j,$$

seeking a value of j for which $-q_j > 0$. Were all the values $-q_j$ are non-positive (i.e., if all the q_j are non-negative), then the initial solution would be optimal. Computing $-q_j$ for $j = 5, 4, 3$, we find that

$$\begin{aligned} -q_5 &= \tfrac{13}{7} \times 3 - \tfrac{11}{7} \times 4 - 5 = -\tfrac{40}{7} \\ -q_4 &= = \tfrac{13}{7} \times 7 - \tfrac{11}{7} \times 8 - 9 = -\tfrac{60}{7} \\ -q_3 &= = \tfrac{13}{7} \times 4 - \tfrac{11}{7} \times 3 - 2 = \tfrac{5}{7} \end{aligned}$$

The inequality $q_3 > 0$ shows that the initial basic feasible solution is not optimal, and we should seek to change basis so as to include the vector $\mathbf{a}^{(3)}$. Let

$$\begin{aligned} t_{1,3} &= r_{1,1} A_{1,3} + r_{1,2} A_{2,3} = -\tfrac{1}{7} \times 4 + \tfrac{3}{7} \times 3 = \tfrac{5}{7} \\ t_{2,3} &= r_{2,1} A_{1,3} + r_{2,2} A_{2,3} = \tfrac{4}{7} \times 4 - \tfrac{5}{7} \times 3 = \tfrac{1}{7} \end{aligned}$$

Then

$$\mathbf{a}^{(3)} = t_{1,3}\mathbf{a}^{(j_1)} + t_{2,3}\mathbf{a}^{(j_2)} = \tfrac{5}{7}\mathbf{a}^{(1)} + \tfrac{1}{7}\mathbf{a}^{(2)}.$$

We introduce a column representing the vector $\mathbf{a}^{(3)}$ into the tableau to serve as a pivot column. The resultant tableau is as follows:—

	$\mathbf{a}^{(3)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$
$\mathbf{a}^{(1)}$	$\frac{5}{7}$	1	$-\frac{1}{7}$	$\frac{3}{7}$
$\mathbf{a}^{(2)}$	$\frac{1}{7}$	2	$\frac{4}{7}$	$-\frac{5}{7}$
	$\frac{5}{7}$	11	$\frac{13}{7}$	$-\frac{11}{7}$

To determine a pivot row we must pick the row index i so as to minimize the ratio $\dfrac{s_i}{t_{i,3}}$, subject to the requirement that $t_{i,3} > 0$. In the context of this example, we should pick $i = 1$. Accordingly the row labelled by the vector $\mathbf{a}^{(1)}$ is the pivot row. To implement the change of basis we must subtract from the second row the values above them in the pivot row, multiplied by $\frac{1}{5}$; we must subtract the values in the pivot row from the values below them in the criterion row, and we must divide the values in the pivot row itself by the pivot element $\frac{5}{7}$.

The resultant tableau corresponding to the basis $2, 3$ is then as follows:—

	$\mathbf{a}^{(3)}$	$\mathbf{b}$	$\mathbf{e}^{(1)}$	$\mathbf{e}^{(2)}$
$\mathbf{a}^{(3)}$	1	$\frac{7}{5}$	$-\frac{1}{5}$	$\frac{3}{5}$
$\mathbf{a}^{(2)}$	0	$\frac{9}{5}$	$\frac{3}{5}$	$-\frac{4}{7}$
	0	10	2	-2

A straightforward computation then shows that if

$$\mathbf{p}^T = \begin{pmatrix} 2 & -2 \end{pmatrix}$$

then

$$\mathbf{p}^T A - \mathbf{c}^T = \begin{pmatrix} -1 & 0 & 0 & -11 & -7 \end{pmatrix}.$$

The components of this row vector are all non-positive. It follows that the basis $\{2, 3\}$ determines a basic optimal solution

$$(0, \tfrac{9}{5}, \tfrac{7}{5}, 0, 0).$$

4.11 Finding an Initial Basic Solution to a Linear Programming Problem

Suppose that we are given a linear programming problem in Dantzig standard form, specified by positive integers m and n, an $m \times n$ matrix A of rank m, an m-dimensional target vector $\mathbf{b} \in \mathbb{R}^m$ and an n-dimensional cost vector $\mathbf{c} \in \mathbb{R}^n$. The problem requires us to find an n-dimensional vector $\mathbf{x}$ that minimizes the objective function $\mathbf{c}^T\mathbf{x}$ subject to the constraints $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$.

The Simplex Tableau Algorithm and the Revised Simplex Algorithm provided methods for passing from an initial basic feasible solution to a basic optimal solution, provided that such a basic optimal solution exists. However, we need first to find an initial basic feasible solution for this linear programming problem.

One can find such an initial basic feasible solution by solving an auxiliary linear programming problem. This auxiliary problem requires us to find n-dimensional vectors $\mathbf{x}$ and $\mathbf{z}$ that minimize the objective function $\sum_{j=1}^{n} (\mathbf{z})_j$ subject to the constraints $A\mathbf{x} + \mathbf{z} = \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{z} \geq \mathbf{0}$.

This auxiliary linear programming problem is itself in Dantzig standard form. Moreover it has an initial basic feasible solution specified by the simultaneous equations $\mathbf{x} = \mathbf{0}$ and $\mathbf{z} = \mathbf{b}$. The objective function of a feasible solution is always non-negative. Applications of algorithms based on the Simplex Method should identify a basic optimal solution $(\mathbf{x}, \mathbf{z})$ for this problem. If the cost $\sum_{j=1}^{n} (\mathbf{z})_j$ of this basic optimal solution is equal to zero then $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$. If the cost of the basic optimal solution is positive then the problem does not have any basic feasible solutions.

The process of solving a linear programming problem in Dantzig standard form thus typically consists of two *phases*. The *Phase I* calculation aims to solve the auxiliary linear programming problem of seeking n-dimensional vectors $\mathbf{x}$ and $\mathbf{z}$ that minimize $\sum_{i=1}^{n} (\mathbf{z})_j$ subject to the constraints $A\mathbf{x} + \mathbf{z} = \mathbf{b}$, $\mathbf{x} \geq 0$ and $\mathbf{z} \geq 0$. If the optimal solution $(\mathbf{x}, \mathbf{z})$ of the auxiliary problem satisfies $\mathbf{z} \neq \mathbf{0}$ then there is no initial basic solution of the original linear programming problem. But if $\mathbf{z} = \mathbf{0}$ then $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$, and thus the Phase I calculation has identified an initial basic feasible solution of the original linear programmming problem. The *Phase II* calculation is the process of successively changing bases to lower the cost of the corresponding basic feasible solutions until either a basic optimal solution has been found or else it has been demonstated that no such basic optimal solution exists.

5 General Linear Programming Problems, Duality and Complementary Slackness

5.1 General Linear Programming Problems

Linear programming is concerned with problems seeking to maximize or minimize a linear functional of several real variables subject to a finite collection of constraints, where each constraint either fixes the values of some linear function of the variables or else requires those values to be bounded, above or below, by some fixed quantity.

The objective of such a problem involving n real variables $x_1, x_2, \ldots, x_n$ is to maximize or minimize an *objective function* of those variables that is of the form

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n,$$

subject to appropriate constraints. The coefficients $c_1, c_2, \ldots, c_n$ that determine the objective function are then fixed real numbers.

Now such an optimization problem may be presented as a minimization problem, because simply changing the signs of all the coefficients $c_1, c_2, \ldots, c_n$ converts any maximization problem into a minimization problem. We therefore suppose, without loss of generality, that the objective of the linear programming problem is to find a feasible solution satisfying appropriate constraints which minimizes the value of the objective function amongst all such feasible solutions to the problem.

Some of the constraints may simply require specific variables to be non-negative or non-positive. Now a constraint that requires a particular variable x_j to be non-positive can be reformulated as one requiring a variable to be non-negative by substituting x_j for $-x_j$ in the statement of the problem. Thus, without loss of generality, we may suppose that all constraints that simply specify the sign of a variable x_j will require that variable to be non-negative. Then all such constraints can be specified by specifying a subset J^+ of $\{1, 2, \ldots, n\}$: the constraints then require that $x_j \geq 0$ for all $j \in J^+$.

There may be further constraints in addition to those that simply specify whether one of the variables is required to be non-positive or non-negative. Suppose that there are m such additional constraints, and let them be numbered between 1 and m. Then, for each integer i between 1 and m, there exist real numbers $A_{i,1}, A_{i,2}, \ldots, A_{i,n}$ and b_i that allow the ith constraint to be expressed either as an *inequality constraint* of the form

$$A_{i,1}x_1 + A_{i,2}x_2 + \ldots + A_{i,n}x_n \geq b_i$$

or else as an *equality constraint* of the form

$$A_{i,1}x_1 + A_{i,2}x_2 + \ldots + A_{i,n}x_n = b_i.$$

It follows from the previous discussion that the statement of a general linear programming problem can be transformed, by changing the signs of some of the variables and constants in the statement of the problem, so as to ensure that the statement of the problem conforms to the following restrictions:—

- the objective function is to be minimized;
- some of the variables may be required to be non-negative;
- other constraints are either inequality constraints placing a lower bound on the value of some linear function of the variables or else equality constraints fixing the value of some linear function of the variables.

Let us describe the statement of a linear programming problem as being in *general primal form* if it conforms to the restrictions just described.

A linear programming problem is expressed in general primal form if the specification of the problem conforms to the following restrictions:—

- the objective of the problem is to find an optimal solution minimizing the objective function amongst all feasible solutions to the problem;
- any variables whose sign is prescribed are required to be non-negative, not non-positive;
- all inequality constraints are expressed by prescribing a lower bound on the value on some linear function of the variables.

A linear programming problem in general primal form can be specified by specifying the following data: an $m \times n$ matrix A with real coefficients, an m-dimensional vector $\mathbf{b}$ with real components; an n-dimensional vector $\mathbf{c}$ with real components; a subset I^+ of $\{1, 2, \ldots, m\}$; and a subset J^+ of $\{1, 2, \ldots, n\}$. The linear programming programming problem specified by this data is the following:—

> *seek* $\mathbf{x} \in \mathbb{R}^n$ *that minimizes the objective function* $\mathbf{c}^T\mathbf{x}$ *subject to the following constraints:—*
>
> - $A\mathbf{x} \geq \mathbf{b}$*;*
> - $(A\mathbf{x})_i = (\mathbf{b})_i$ *unless* $i \in I^+$*;*

- $(\mathbf{x})_j \geq 0$ *for all* $j \in J^+$.

We refer to the $m \times n$ matrix A, the m-dimensional vector $\mathbf{b}$ and the n-dimensional vector $\mathbf{c}$ employed in specifying a linear programming problem in general primal form as the *constraint matrix*, *target vector* and *cost vector* respectively for the linear programming problem. Let us refer to the subset I^+ of $\{1, 2, \ldots, m\}$ specifying those constraints that are inequality constraints as the *inequality constraint specifier* for the problem, and let us refer to the subset J^+ of $\{1, 2, \ldots, n\}$ that specifies those variables that are required to be non-negative for a feasible solution as the *variable sign specifier* for the problem.

We denote by $\text{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ the linear programming problem whose specification in general primal form is determined by a constraint matrix A, target vector $\mathbf{b}$, cost vector $\mathbf{c}$, inequality constraint specifier I^+ and variable sign specifier J^+.

A linear programming problem formulated in general primal form can be reformulated as a problem in Dantzig standard form, thus enabling the use of the Simplex Method to find solutions to the problem.

Indeed consider a linear programming problem $\text{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ where the constraint matrix A is an $m \times n$ matrix with real coefficients, the target vector $\mathbf{b}$ and the cost vector $\mathbf{c}$ are vectors of dimension m and n respectively with real coefficients. Then the inequality constraint specifier I^+ is a subset of $\{1, 2, \ldots, m\}$ and the variable sign specifier J^+ is a subset of $\{1, 2, \ldots, n\}$. The problem is already in Dantzig standard form if and only if $I^+ = \emptyset$ and $J^+ = \{1, 2, \ldots, n\}$.

If the problem is not in Dantzig standard form, then each variable x_j for $j \notin J^+$ can be replaced by a pair of variables x_j^+ and x_j^- satisfying the constraints $x_j^+ \geq 0$ and $x_j^- \geq 0$: the difference $x_j^+ - x_j^-$ of these new variables is substituted for x_j in the objective function and the constraints. Also a *slack variable* z_i can be introduced for each $i \in I^+$, where z_i is required to satisfy the sign constraint $z_i \geq 0$, and the inequality constraint

$$A_{i,1}x_1 + A_{i,2}x_2 + \ldots + A_{i,n}x_n \geq b_i$$

is then replaced by the corresponding equality constraint

$$A_{i,1}x_1 + A_{i,2}x_2 + \ldots + A_{i,n}x_n - z_i = b_i.$$

The linear programming problem $\text{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ in general primal form can therefore be reformulated as a linear programming problem in Dantzig standard form as follows:—

determine values of x_j for all $j \in J^+$, x_j^+ and x_j^- for all $j \in J^0$, where $J^0 = \{1, 2, \ldots, n\} \setminus J^+$, and z_i for all $i \in I^+$ so as to minimize the objective function

$$\sum_{j \in J^+} c_j x_j + \sum_{j \in J^0} c_j x_j^+ - \sum_{j \in J^0} c_j x_j^-$$

subject to the following constraints:—

(i) $\sum_{j \in J^+} A_{i,j} x_j + \sum_{j \in J^0} A_{i,j} x_j^+ - \sum_{j \in J^0} A_{i,j} x_j^- = b_i$ *for each* $i \in \{1, 2, \ldots, n\} \setminus I^+$;

1. $\sum_{j \in J^+} A_{i,j} x_j + \sum_{j \in J^0} A_{i,j} x_j^+ - \sum_{j \in J^0} A_{i,j} x_j^- - z_i = b_i$ *for each* $i \in I^+$;

(ii) $x_j \geq 0$ *for all* $j \in J^+$;

(iii) $x_j^+ \geq 0$ *and* $x_j^- \geq 0$ *for all* $j \in J^0$;

(iv) $z_i \geq 0$ *for all* $i \in I^+$.

Once the problem has been reformulated in Dantzig standard form, techiques based on the Simplex Method can be employed in the search for solutions to the problem.

5.2 Complementary Slackness and the Weak Duality Theorem

Every linear programming problem $\text{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ in general primal form determines a corresponding linear programming problem $\text{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ in *general dual form*. The second linear programming problem is referred to as the *dual* of the first, and the first linear programming problem is referred to as the *primal* of its dual.

We shall give the definition of the dual problem associated with a given linear programming problem, and investigate some important relationships between the primal linear programming problem and its dual.

Let $\text{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ be a linear programming problem in general primal form specified in terms of an $m \times n$ constraint matrix A, m-dimensional target vector $\mathbf{b}$, n-dimensional cost vector $\mathbf{c}$, inequality constraint specifier I^+ and variable sign specifier J^+. The corresponding dual problem $\text{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ may be specified in *general dual form* as follows:

seek $\mathbf{p} \in \mathbb{R}^m$ *that maximizes the objective function* $\mathbf{p}^T\mathbf{b}$ *subject to the following constraints:—*

- $\mathbf{p}^T A \leq \mathbf{c}^T$*;*
- $(\mathbf{p})_i \geq 0$ *for all* $i \in I^+$*;*
- $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ *unless* $j \in J^+$*.*

Lemma 5.1 *Let* $\mathrm{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ *be a linear programming problem expressed in general primal form with constraint matrix* A *with* m *rows and* n *columns, target vector* $\mathbf{b}$*, cost vector* $\mathbf{c}$*, inequality constraint specifier* I^+ *and variable sign specifier* J^+*. Then the feasible and optimal solutions of the corresponding dual linear programming problem* $\mathrm{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ *are those of the problem* $\mathrm{Primal}(-A^T, -\mathbf{c}, -\mathbf{b}, J^+, I^+)$*.*

Proof An m-dimensional vector $\mathbf{p}$ satisfies the constraints of the dual linear programming problem $\mathrm{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ if and only if $\mathbf{p}^T A \leq \mathbf{c}^T$, $(\mathbf{p})_i \geq 0$ for all $i \in I^+$ and $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ unless $j \in J^+$. On taking the transposes of the relevant matrix equations and inequalities, we see that these conditions are satisfied if and only if $-A^T\mathbf{p} \geq -\mathbf{c}$, $(\mathbf{p})_i \geq 0$ for all $i \in I^+$ and $(-A^T\mathbf{p})_j = (-\mathbf{c})_j$ unless $j \in J^+$. But these are the requirements that the vector $\mathbf{p}$ must satisfy in order to be a feasible solution of the linear programming problem $\mathrm{Primal}(-A^T, -\mathbf{c}, -\mathbf{b}, J^+, I^+)$. Moreover $\mathbf{p}$ is an optimal solution of $\mathrm{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ if and only if it maximizes the value of $\mathbf{p}^T\mathbf{b}$, and this is the case if and only if it minimizes the value of $-\mathbf{b}^T\mathbf{p}$. The result follows. ∎

A linear programming problem in Dantzig standard form is specified by specifying integers m and n a constraint matrix A which is an $m \times n$ matrix with real coefficients, a target vector $\mathbf{b}$ belonging to the real vector space $\mathbb{R}^m$ and a cost vector $\mathbf{c}$ belonging to the real vector space $\mathbb{R}^m$. The objective of the problem is to find a feasible solution to the problem that minimizes the quantity $\mathbf{c}^T\mathbf{x}$ amongst all n-dimensional vectors $\mathbf{x}$ for which $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$.

The objective of the dual problem is then to find some feasible solution to the problem that maximizes the quantity $\mathbf{p}^T\mathbf{b}$ amongst all m-dimensional vectors $\mathbf{p}$ for which $\mathbf{p}^T A \leq \mathbf{c}$.

Theorem 5.2 (Weak Duality Theorem for Linear Programming Problems in Dantzig Standard Form)
Let m *and* n *be integers, let* A *be an* $m \times n$ *matrix with real coefficients, let* $\mathbf{b} \in \mathbb{R}^m$ *and let* $\mathbf{c} \in \mathbb{R}^n$*. Let* $\mathbf{x} \in \mathbb{R}^n$ *satisfy the constraints* $A\mathbf{x} = \mathbf{b}$ *and* $\mathbf{x} \geq \mathbf{0}$*, and let* $\mathbf{p} \in \mathbb{R}^m$ *satisfy the constraint* $\mathbf{p}^T A \leq \mathbf{c}$*. Then* $\mathbf{p}^T\mathbf{b} \leq \mathbf{c}^T\mathbf{x}$*. Moreover* $\mathbf{p}^T\mathbf{b} = \mathbf{c}^T\mathbf{x}$ *if and only if the following complementary slackness condition is satisfied:*

- $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ *for all integers* j *between* 1 *and* n *for which* $(\mathbf{x})_j > 0$.

Proof The constraints satisfied by the vectors $\mathbf{x}$ and $\mathbf{p}$ ensure that

$$\begin{aligned}\mathbf{c}^T\mathbf{x} - \mathbf{p}^T\mathbf{b} &= (\mathbf{c}^T - \mathbf{p}^T A)\mathbf{x} + \mathbf{p}^T(A\mathbf{x} - \mathbf{b}) \\ &= (\mathbf{c}^T - \mathbf{p}^T A)\mathbf{x},\end{aligned}$$

because $A\mathbf{x} - \mathbf{b} = \mathbf{0}$. But also $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{c}^T - \mathbf{p}^T A) \geq \mathbf{0}$, and therefore $(\mathbf{c}^T - \mathbf{p}^T A)\mathbf{x} \geq 0$. Moreover $(\mathbf{c}^T - \mathbf{p}^T A)\mathbf{x} = 0$ if and only if $(\mathbf{c}^T - \mathbf{p}^T A)_j = 0$ for all integers j between 1 and n for which $(\mathbf{x})_j > 0$. The result follows. ∎

Corollary 5.3 *Let a linear programming problem in Dantzig standard form be specified by an* $m \times n$ *constraint matrix* A, *and* m*-dimensional target vector* $\mathbf{b}$ *and an* n*-dimensional cost vector* $\mathbf{c}$. *Let* $\mathbf{x}^*$ *be a feasible solution of this primal problem, and let* $\mathbf{p}^*$ *be a solution of the dual problem. Then* $\mathbf{p}^{*T}A \leq \mathbf{c}^T$. *Suppose that the complementary slackness conditions for this primal-dual pair are satisfied, so that* $(\mathbf{p}^{*T}A)_j = (\mathbf{c})_j$ *for all integers* j *between* 1 *and* n *for which* $(\mathbf{x}^*)_j > 0$. *Then* $\mathbf{x}^*$ *is an optimal solution of the primal problem, and* $\mathbf{p}^*$ *is an optimal solution of the dual problem.*

Proof Because the complementary slackness conditions for this primal-dual pair are satisfied, it follows from the Weak Duality Theorem that $\mathbf{c}^T\mathbf{x}^* = \mathbf{p}^{*T}\mathbf{b}$ (see Theorem 5.2). But it then also follows from the Weak Duality Theorem that

$$\mathbf{c}^T\mathbf{x} \geq \mathbf{p}^{*T}\mathbf{b} = \mathbf{c}^T\mathbf{x}^*$$

for all feasible solutions $\mathbf{x}$ of the primal problem. It follows that $\mathbf{x}^*$ is an optimal solution of the primal problem. Similarly

$$\mathbf{p}^T\mathbf{b} \leq \mathbf{c}^T\mathbf{x}^* = \mathbf{p}^{*T}\mathbf{b}$$

for all feasible solutions $\mathbf{p}$ of the dual problem. It follows that $\mathbf{p}^*$ is an optimal solution of the dual problem, as required. ∎

Another special case of duality in linear programming is exemplified by a primal-dual pair of problems in *Von Neumann Symmetric Form.* In this case the primal and dual problems are specified in terms of an $m \times n$ constraint matrix A, an m-dimensional target vector $\mathbf{b}$ and an n-dimensional cost vector $\mathbf{c}$. The objective of the problem is minimize $\mathbf{c}^T\mathbf{x}$ amongst n-dimensional vectors $\mathbf{x}$ that satisfy the constraints $A\mathbf{x} \geq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$. The dual problem is to maximize $\mathbf{p}^T\mathbf{b}$ amongst m-dimensional vectors $\mathbf{p}$ that satisfy the constraints $\mathbf{p}^T A \leq \mathbf{c}^T$ and $\mathbf{p} \geq \mathbf{0}$.

Theorem 5.4 (Weak Duality Theorem for Linear Programming Problems in Von Neumann Symmetric Form)

Let m and n be integers, let A be an $m \times n$ matrix with real coefficients, let $\mathbf{b} \in \mathbb{R}^m$ and let $\mathbf{c} \in \mathbb{R}^n$. Let $\mathbf{x} \in \mathbb{R}^n$ satisfy the constraints $A\mathbf{x} \geq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$, and let $\mathbf{p} \in \mathbb{R}^m$ satisfy the constraints $\mathbf{p}^T A \leq \mathbf{c}$ and $\mathbf{p}^T \geq \mathbf{0}$. Then $\mathbf{p}^T\mathbf{b} \leq \mathbf{c}^T\mathbf{x}$. Moreover $\mathbf{p}^T\mathbf{b} = \mathbf{c}^T\mathbf{x}$ if and only if the following complementary slackness conditions are satisfied:

- $(A\mathbf{x})_i = (\mathbf{b})_i$ *for all integers i between 1 and m for which* $(\mathbf{p})_i > 0$;
- $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ *for all integers j between 1 and n for which* $(\mathbf{x})_j > 0$;

Proof The constraints satisfied by the vectors $\mathbf{x}$ and $\mathbf{p}$ ensure that

$$\mathbf{c}^T\mathbf{x} - \mathbf{p}^T\mathbf{b} = (\mathbf{c}^T - \mathbf{p}^T A)\mathbf{x} + \mathbf{p}^T(A\mathbf{x} - \mathbf{b}).$$

But $\mathbf{x} \geq \mathbf{0}$, $\mathbf{p} \geq \mathbf{0}$, $A\mathbf{x} - \mathbf{b} \geq \mathbf{0}$ and $\mathbf{c}^T - \mathbf{p}^T A \geq \mathbf{0}$. It follows that $\mathbf{c}^T\mathbf{x} - \mathbf{p}^T\mathbf{b} \geq 0$. and therefore $\mathbf{c}^T\mathbf{x} \geq \mathbf{p}^T\mathbf{b}$. Moreover $\mathbf{c}^T\mathbf{x} - \mathbf{p}^T\mathbf{b} = 0$ if and only if $(\mathbf{c}^T - \mathbf{p}^T A)_j(\mathbf{x})_j = 0$ for $j = 1, 2, \ldots, n$ and $(\mathbf{p})_i(A\mathbf{x} - \mathbf{b})_i = 0$, and therefore $\mathbf{c}^T\mathbf{x} = \mathbf{p}^T\mathbf{b}$ if and only if the complementary slackness conditions are satisfied. ∎

Theorem 5.5 (Weak Duality Theorem for Linear Programming Problems in General Primal Form)

Let $\mathbf{x} \in \mathbb{R}^\times$ be a feasible solution to a linear programming problem

$$\text{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$$

expressed in general primal form with constraint matrix A with m rows and n columns, target vector $\mathbf{b}$, cost vector $\mathbf{c}$, inequality constraint specifier I^+ and variable sign specifier J^+, and let $\mathbf{p} \in \mathbb{R}^m$ be a feasible solution to the corresponding dual programming problem

$$\text{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+).$$

Then $\mathbf{p}^T\mathbf{b} \leq \mathbf{c}^T\mathbf{x}$. Moreover $\mathbf{p}^T\mathbf{b} = \mathbf{c}^T\mathbf{x}$ if and only if the following complementary slackness conditions are satisfied:—

- $(A\mathbf{x})_i = \mathbf{b}_i$ *whenever* $(\mathbf{p})_i \neq 0$;
- $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ *whenever* $(\mathbf{x})_j \neq 0$.

Proof The feasible solution $\mathbf{x}$ to the primal problem satisfies the following constraints:—

- $A\mathbf{x} \geq \mathbf{b}$;
- $(A\mathbf{x})_i = (\mathbf{b})_i$ unless $i \in I^+$;
- $(\mathbf{x})_j \geq 0$ for all $j \in J^+$.

The feasible solution $\mathbf{p}$ to the dual problem satisfies the following constraints:—

- $\mathbf{p}^T A \leq \mathbf{c}^T$;
- $(\mathbf{p})_i \geq 0$ for all $i \in I^+$;
- $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ unless $j \in J^+$.

Now

$$\begin{aligned} \mathbf{c}^T\mathbf{x} - \mathbf{p}^T\mathbf{b} &= (\mathbf{c}^T - \mathbf{p}^T A)\mathbf{x} + \mathbf{p}^T(A\mathbf{x} - \mathbf{b}) \\ &= \sum_{j=1}^{n} (\mathbf{c}^T - \mathbf{p}^T A)_j (\mathbf{x})_j + \sum_{i=1}^{m} (\mathbf{p})_i (A\mathbf{x} - \mathbf{b})_i. \end{aligned}$$

Let j be an integer between 1 and n. If $j \in J^+$ then $(\mathbf{x})_j \geq 0$ and $(\mathbf{c}^T - \mathbf{p}^T A)_j \geq 0$, and therefore $(\mathbf{c}^T - \mathbf{p}^T A)_j (\mathbf{x})_j \geq 0$. If $j \notin J^+$ then $(\mathbf{p}^T A)_j = (\mathbf{c})_j$, and therefore $(\mathbf{c}^T - \mathbf{p}^T A)_j (\mathbf{x})_j = 0$, irrespective of whether $(\mathbf{x})_j$ is positive, negative or zero. It follows that

$$\sum_{j=1}^{n} (\mathbf{c}^T - \mathbf{p}^T A)_j (\mathbf{x})_j \geq 0.$$

Moreover

$$\sum_{j=1}^{n} (\mathbf{c}^T - \mathbf{p}^T A)_j (\mathbf{x})_j = 0$$

if and only if $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ for all indices j for which $(\mathbf{x})_j \neq 0$.

Next let i be an index between 1 and m. If $i \in I^+$ then $(\mathbf{p})_i \geq 0$ and $(A\mathbf{x} - \mathbf{b})_i \geq 0$, and therefore $(\mathbf{p})_i (A\mathbf{x} - \mathbf{b})_i \geq 0$. If $i \notin I^+$ then $(A\mathbf{x})_i = (\mathbf{b})_i$, and therefore $(\mathbf{p})_i (A\mathbf{x} - \mathbf{b})_i = 0$, irrespective of whether $(\mathbf{p})_i$ is positive, negative or zero. It follows that

$$\sum_{i=1}^{m} (\mathbf{p})_i (A\mathbf{x} - \mathbf{p})_i \geq 0.$$

Moreover

$$\sum_{i=1}^{m} (\mathbf{p})_i (A\mathbf{x} - \mathbf{p})_i = 0.$$

if and only if $(A\mathbf{x})_i = (\mathbf{b})_i$ for all indices i for which $(\mathbf{p})_i \neq 0$. The result follows. ∎

Corollary 5.6 *Let $\mathbf{x}^* \in \mathbb{R}^\ltimes$ be a feasible solution to a linear programming problem* $\text{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ *expressed in general primal form with constraint matrix A with m rows and n columns, target vector $\mathbf{b}$, cost vector $\mathbf{c}$, inequality constraint specifier I^+ and variable sign specifier J^+, and let $\mathbf{p}^* \in \mathbb{R}^m$ be a feasible solution to the corresponding dual programming problem* $\text{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$. *Suppose that the complementary slackness conditions are satisfied for this pair of problems, so that $(A\mathbf{x})_i = \mathbf{b}_i$ whenever $(\mathbf{p})_i \neq 0$, and $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ whenever $(\mathbf{x})_j \neq 0$. Then $\mathbf{x}^*$ is an optimal solution for the primal problem and $\mathbf{p}^*$ is an optimal solution for the dual problem.*

Proof Because the complementary slackness conditions for this primal-dual pair are satisfied, it follows from the Weak Duality Theorem that $\mathbf{c}^T\mathbf{x}^* = \mathbf{p}^{*T}\mathbf{b}$ (see Theorem 5.5). But it then also follows from the Weak Duality Theorem that

$$\mathbf{c}^T\mathbf{x} \geq \mathbf{p}^{*T}\mathbf{b} = \mathbf{c}^T\mathbf{x}^*$$

for all feasible solutions $\mathbf{x}$ of the primal problem. It follows that $\mathbf{x}^*$ is an optimal solution of the primal problem. Similarly

$$\mathbf{p}^T\mathbf{b} \leq \mathbf{c}^T\mathbf{x}^* = \mathbf{p}^{*T}\mathbf{b}$$

for all feasible solutions $\mathbf{p}$ of the dual problem. It follows that $\mathbf{p}^*$ is an optimal solution of the dual problem, as required. ∎

Example Consider the following linear programming problem in general primal form:—

> *find values of x_1, x_2, x_3 and x_4 so as to minimize the objective function*
>
> $$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4$$
>
> *subject to the following constraints:—*
>
> - $a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + a_{1,4}x_4 = b_1$;
> - $a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + a_{2,4}x_4 = b_2$;
> - $a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 + a_{3,4}x_4 \geq b_3$;
> - $x_1 \geq 0$ *and* $x_3 \geq 0$.

Here $a_{i,j}$, b_i and c_j are constants for $i = 1, 2, 3$ and $j = 1, 2, 3, 4$. The dual problem is the following:—

find values of p_1, p_2 and p_3 so as to maximize the objective function

$$p_1 b_1 + p_2 b_2 + p_3 b_3$$

subject to the following constraints:—

- $p_1 a_{1,1} + p_2 a_{2,1} + p_3 a_{3,1} \leq c_1$;
- $p_1 a_{1,2} + p_2 a_{2,2} + p_3 a_{3,2} = c_2$;
- $p_1 a_{1,3} + p_2 a_{2,3} + p_3 a_{3,3} \leq c_3$;
- $p_1 a_{1,4} + p_2 a_{2,4} + p_3 a_{3,4} = c_4$;
- $p_3 \geq 0$.

We refer to the first and second problems as the *primal problem* and the *dual problem* respectively. Let (x_1, x_2, x_3, x_4) be a feasible solution of the primal problem, and let (p_1, p_2, p_3) be a feasible solution of the dual problem. Then

$$\begin{aligned}\sum_{j=1}^{4} c_j x_j - \sum_{i=1}^{3} p_i b_i &= \sum_{j=1}^{4} \left(c_j - \sum_{i=1}^{3} p_i a_{i,j} \right) x_j \\ &\quad + \sum_{i=1}^{3} p_i \left(\sum_{j=1}^{4} a_{i,j} x_j - b_i \right).\end{aligned}$$

Now the quantity $c_j - \sum\limits_{i=1}^{3} p_i a_{i,j} = 0$ for $j = 2$ and $j = 4$, and $\sum\limits_{j=1}^{4} a_{i,j} x_j - b_i = 0$ for $i = 1$ and $i = 2$. It follows that

$$\begin{aligned}\sum_{j=1}^{4} c_j x_j - \sum_{i=1}^{3} p_i b_i &= \left(c_1 - \sum_{i=1}^{3} p_i a_{i,1} \right) x_1 \\ &\quad + \left(c_3 - \sum_{i=1}^{3} p_i a_{i,3} \right) x_3 \\ &\quad + p_3 \left(\sum_{j=1}^{4} a_{3,j} x_j - b_3 \right).\end{aligned}$$

Now $x_1 \geq 0$, $x_3 \geq 0$ and $p_3 \geq 0$. Also

$$c_1 - \sum_{i=1}^{3} p_i a_{i,1} \geq 0, \quad c_3 - \sum_{i=1}^{3} p_i a_{i,3} \geq 0$$

and

$$\sum_{j=1}^{4} a_{3,j} x_j - b_3 \geq 0.$$

It follows that

$$\sum_{j=1}^{4} c_j x_j - \sum_{i=1}^{3} p_i b_i \geq 0.$$

and thus

$$\sum_{j=1}^{4} c_j x_j \geq \sum_{i=1}^{3} p_i b_i.$$

Now suppose that

$$\sum_{j=1}^{4} c_j x_j = \sum_{i=1}^{3} p_i b_i.$$

Then

$$\begin{aligned} \left(c_1 - \sum_{i=1}^{3} p_i a_{i,1}\right) x_1 &= 0, \\ \left(c_3 - \sum_{i=1}^{3} p_i a_{i,3}\right) x_3 &= 0, \\ p_3 \left(\sum_{j=1}^{4} a_{3,j} x_j - b_3\right) &= 0, \end{aligned}$$

because a sum of three non-negative quantities is equal to zero if and only if each of those quantities is equal to zero.

It follows that

$$\sum_{j=1}^{4} c_j x_j = \sum_{i=1}^{3} p_i b_i$$

if and only if the following three complementary slackness conditions are satisfied:—

- $\sum_{i=1}^{3} p_i a_{i,1} = c_1$ if $x_1 > 0$;
- $\sum_{i=1}^{3} p_i a_{i,3} = c_3$ if $x_3 > 0$;
- $\sum_{j=1}^{4} a_{3,j} x_j = b_3$ if $p_3 > 0$.

5.3 Open and Closed Sets in Euclidean Spaces

Let m be a positive integer. The *Euclidean norm* $|\mathbf{x}|$ of an element $\mathbf{x}$ of $\mathbb{R}^m$ is defined such that

$$|\mathbf{x}|^2 = \sum_{i=1}^{m} (\mathbf{x})_i^2.$$

The *Euclidean distance function* d on $\mathbb{R}^m$ is defined such that

$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{y} - \mathbf{x}|$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$. The Euclidean distance function satisfies the Triangle Inequality, together with all the other basic properties required of a distance function on a metric space, and therefore $\mathbb{R}^m$ with the Euclidean distance function is a metric space.

A subset U of $\mathbb{R}^m$ is said to be *open* in $\mathbb{R}^m$ if, given any point $\mathbf{b}$ of U, there exists some real number ε satisfying $\varepsilon > 0$ such that

$$\{\mathbf{x} \in \mathbb{R}^m : |\mathbf{x} - \mathbf{b}| < \varepsilon\} \subset U.$$

A subset of $\mathbb{R}^m$ is *closed* in $\mathbb{R}^m$ if and only if its complement is open in $\mathbb{R}^m$.

Every union of open sets in $\mathbb{R}^m$ is open in $\mathbb{R}^m$, and every finite intersection of open sets in $\mathbb{R}^m$ is open in $\mathbb{R}^m$.

Every intersection of closed sets in $\mathbb{R}^m$ is closed in $\mathbb{R}^m$, and every finite union of closed sets in $\mathbb{R}^m$ is closed in $\mathbb{R}^m$.

Lemma 5.7 *Let m be a positive integer, let $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(m)}$ be a basis of $\mathbb{R}^m$, and let*

$$F = \left\{ \sum_{i=1}^{m} s_i \mathbf{u}^{(i)} : s_i \geq 0 \text{ for } i = 1, 2, \ldots, m \right\}.$$

Then F is a closed set in $\mathbb{R}^m$.

Proof Let $T\colon \mathbb{R}^m \to \mathbb{R}^m$ be defined such that

$$T(s_1, s_2, \ldots, s_m) = \sum_{i=1}^{m} s_i \mathbf{u}^{(i)}$$

for all real numbers $s_1, s_2, \ldots, s_m$. Then T is an invertible linear operator on $\mathbb{R}^m$, and $F = T(G)$, where

$$G = \{\mathbf{x} \in \mathbb{R}^m : (\mathbf{x})_i \geq 0 \text{ for } i = 1, 2, \ldots, m\}.$$

Moreover the subset G of $\mathbb{R}^m$ is closed in $\mathbb{R}^m$.

Now it is a standard result of real analysis that every linear operator on a finite-dimensional vector space is continuous. Therefore $T^{-1}: \mathbb{R}^m \to \mathbb{R}^m$ is continuous. Moreover $T(G)$ is the preimage of the closed set G under the continuous map T^{-1}, and the preimage of any closed set under a continuous map is itself closed. It follows that $T(G)$ is closed in $\mathbb{R}^m$. Thus F is closed in $\mathbb{R}^m$, as required. ∎

Lemma 5.8 *Let m be a positive integer, let F be a non-empty closed set in $\mathbb{R}^m$, and let $\mathbf{b}$ be a vector in $\mathbb{R}^m$. Then there exists an element $\mathbf{g}$ of F such that $|\mathbf{x} - \mathbf{b}| \geq |\mathbf{g} - \mathbf{b}|$ for all $\mathbf{x} \in F$.*

Proof Let R be a positive real number chosen large enough to ensure that the set F_0 is non-empty, where

$$F_0 = F \cap \{\mathbf{x} \in \mathbb{R}^m : |\mathbf{x} - \mathbf{b}| \leq R\}.$$

Then F_0 is a closed bounded subset of $\mathbb{R}^m$. Let $f: F_0 \to \mathbb{R}$ be defined such that $f(\mathbf{x}) = |\mathbf{x} - \mathbf{b}|$ for all $\mathbf{x} \in F$. Then $f: F_0 \to \mathbb{R}$ is a continuous function on F_0.

Now it is a standard result of real analysis that any continuous real-valued function on a closed bounded subset of a finite-dimensional Euclidean space attains a minimum value at some point of that set. It follows that there exists an element $\mathbf{g}$ of F_0 such that

$$|\mathbf{x} - \mathbf{b}| \geq |\mathbf{g} - \mathbf{b}|$$

for all $\mathbf{x} \in F_0$. If $\mathbf{x} \in F \setminus F_0$ then

$$|\mathbf{x} - \mathbf{b}| \geq R \geq |\mathbf{g} - \mathbf{b}|.$$

It follows that

$$|\mathbf{x} - \mathbf{b}| \geq |\mathbf{g} - \mathbf{b}|$$

for all $\mathbf{x} \in F$, as required. ∎

5.4 A Separating Hyperplane Theorem

Definition A subset K of $\mathbb{R}^m$ is said to be *convex* if $(1-\mu)\mathbf{x} + \mu\mathbf{x}' \in K$ for all elements $\mathbf{x}$ and $\mathbf{x}'$ of K and for all real numbers μ satisfying $0 \leq \mu \leq 1$.

It follows from the above definition that a subset K of $\mathbb{R}^{>}$ is a convex subset of $\mathbb{R}^m$ if and only if, given any two points of K, the line segment joining those two points is wholly contained in K.

Theorem 5.9 *Let m be a positive integer, let K be a closed convex set in $\mathbb{R}^m$, and let $\mathbf{b}$ be a vector in $\mathbb{R}^m$, where $\mathbf{b} \notin K$. Then there exists a linear functional $\varphi\colon \mathbb{R}^m \to \mathbb{R}$ and a real number c such that $\varphi(\mathbf{x}) > c$ for all $\mathbf{x} \in K$ and $\varphi(\mathbf{b}) < c$.*

Proof It follows from Lemma 5.8 that there exists a point $\mathbf{g}$ of K such that $|\mathbf{x} - \mathbf{b}| \geq |\mathbf{g} - \mathbf{b}|$ for all $\mathbf{x} \in K$. Let $\mathbf{x} \in K$. Then $(1 - \lambda)\mathbf{g} + \lambda\mathbf{x} \in K$ for all real numbers λ satisfying $0 \leq \lambda \leq 1$, because the set K is convex, and therefore

$$|(1 - \lambda)\mathbf{g} + \lambda\mathbf{x} - \mathbf{b}| \geq |\mathbf{g} - \mathbf{b}|$$

for all real numbers λ satisfying $0 \leq \lambda \leq 1$. Now

$$(1 - \lambda)\mathbf{g} + \lambda\mathbf{x} - \mathbf{b} = \mathbf{g} - \mathbf{b} + \lambda(\mathbf{x} - \mathbf{g}).$$

It follows by a straightforward calculation from the definition of the Euclidean norm that

$$\begin{aligned} |\mathbf{g} - \mathbf{b}|^2 &\leq |(1 - \lambda)\mathbf{g} + \lambda\mathbf{x} - \mathbf{b}|^2 \\ &= |\mathbf{g} - \mathbf{b}|^2 + 2\lambda(\mathbf{g} - \mathbf{b})^T(\mathbf{x} - \mathbf{g}) \\ &\qquad + \lambda^2|\mathbf{x} - \mathbf{g}|^2 \end{aligned}$$

for all real numbers λ satisfying $0 \leq \lambda \leq 1$. In particular, this inequality holds for all sufficiently small positive values of λ, and therefore

$$(\mathbf{g} - \mathbf{b})^T(\mathbf{x} - \mathbf{g}) \geq 0$$

for all $\mathbf{x} \in K$.

Let

$$\varphi(\mathbf{x}) = (\mathbf{g} - \mathbf{b})^T\mathbf{x}$$

for all $\mathbf{x} \in \mathbb{R}^m$. Then $\varphi\colon \mathbb{R}^m \to \mathbb{R}$ is a linear functional on $\mathbb{R}^m$, and $\varphi(\mathbf{x}) \geq \varphi(\mathbf{g})$ for all $\mathbf{x} \in K$. Moreover

$$\varphi(\mathbf{g}) - \varphi(\mathbf{b}) = |\mathbf{g} - \mathbf{b}|^2 > 0,$$

and therefore $\varphi(\mathbf{g}) > \varphi(\mathbf{b})$. It follows that $\varphi(\mathbf{x}) > c$ for all $\mathbf{x} \in K$, where $c = \frac{1}{2}\varphi(\mathbf{b}) + \frac{1}{2}\varphi(\mathbf{g})$, and that $\varphi(\mathbf{b}) < c$. The result follows. ∎

5.5 Convex Cones

Definition Let m be a positive integer. A subset C of $\mathbb{R}^m$ is said to be a *convex cone* in $\mathbb{R}^m$ if $\lambda\mathbf{v} + \mu\mathbf{w} \in C$ for all $\mathbf{v}, \mathbf{w} \in C$ and for all real numbers λ and μ satisfying $\lambda \geq 0$ and $\mu \geq 0$.

Lemma 5.10 *Let m be a positive integer. Then every convex cone in $\mathbb{R}^m$ is a convex subset of $\mathbb{R}^m$.*

Proof Let C be a convex cone in $\mathbb{R}^m$ and let $\mathbf{v}, \mathbf{w} \in C$. Then $\lambda\mathbf{v} + \mu\mathbf{w} \in C$ for all non-negative real numbers λ and μ. In particular $(1-\lambda)\mathbf{w} + \lambda\mathbf{v} \in C$. whenever $0 \leq \lambda \leq 1$, and thus the convex cone C is a convex set in $\mathbb{R}^m$, as required. ∎

Lemma 5.11 *Let S be a subset of $\mathbb{R}^m$, and let C be the set of all elements of $\mathbb{R}^m$ that can be expressed as a linear combination of the form*

$$s_1\mathbf{a}^{(1)} + s_2\mathbf{a}^{(2)} + \cdots + s_n\mathbf{a}^{(n)},$$

where $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ are vectors belonging to S and $s_1, s_2, \ldots, s_n$ are non-negative real numbers. Then C is a convex cone in $\mathbb{R}^m$.

Proof Let $\mathbf{v}$ and $\mathbf{w}$ be elements of C. Then there exist finite subsets S_1 and S_2 of S such that $\mathbf{v}$ can be expressed as a linear combination of the elements of S_1 with non-negative coefficients and $\mathbf{w}$ can be expressed as a linear combination of the elements of S_2 with non-negative coefficients. Let

$$S_1 \cup S_2 = \{\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}\}.$$

Then there exist non-negative real numbers $s_1, s_2, \ldots, s_n$ and $t_1, t_2, \ldots, t_n$ such that

$$\mathbf{v} = \sum_{j=1}^{n} s_j\mathbf{a}^{(j)} \quad \text{and} \quad \mathbf{w} = \sum_{j=1}^{n} t_j\mathbf{a}^{(j)}.$$

Let λ and μ be non-negative real numbers. Then

$$\lambda\mathbf{v} + \mu\mathbf{w} = \sum_{j=1}^{n} (\lambda s_j + \mu t_j)\mathbf{a}^{(j)},$$

and $\lambda s_j + \mu t_j \geq 0$ for $j = 1, 2, \ldots, n$. It follows that $\lambda\mathbf{v} + \mu\mathbf{w} \in S$, as required. ∎

Proposition 5.12 *Let m be a positive integer, let $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)} \in \mathbb{R}^m$, and let C be the subset of $\mathbb{R}^m$ defined such that*

$$C = \left\{ \sum_{j=1}^{n} t_j\mathbf{a}^{(j)} : t_j \geq 0 \text{ for } j = 1, 2, \ldots, n \right\}.$$

Then C is a closed convex cone in $\mathbb{R}^m$.

Proof It follows from Lemma 5.11 that C is a convex cone in $\mathbb{R}^m$. We must prove that this convex cone is a closed set.

The vectors $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ span a vector subspace V of $\mathbb{R}^m$ that is isomorphic as a real vector space to $\mathbb{R}^k$ for some integer k satisfying $0 \leq k \leq m$. This vector subspace V of $\mathbb{R}^m$ is a closed subset of $\mathbb{R}^m$, and therefore any subset of V that is closed in V will also be closed in $\mathbb{R}^m$. Replacing $\mathbb{R}^m$ by $\mathbb{R}^k$, if necessary, we may assume, without loss of generality that the vectors $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ span the vector space $\mathbb{R}^m$. Thus if A is the $m \times n$ matrix defined such that $(A)_{i,j} = (\mathbf{a}^{(j)})_i$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$ then the matrix A is of rank m.

Let $\mathcal{B}$ be the collection consisting of all subsets B of $\{1, 2, \ldots, n\}$ for which the members of the set $\{\mathbf{a}^{(j)} : j \in B\}$ constitute a basis of the real vector space $\mathbb{R}^m$ and, for each $B \in \mathcal{B}$, let

$$C_B = \left\{ \sum_{i=1}^{m} s_i \mathbf{a}^{(j_i)} : s_i \geq 0 \text{ for } i = 1, 2, \ldots, m \right\},$$

where $j_1, j_2, \ldots, j_m$ are distinct and are the elements of the set B. It follows from Lemma 5.7 that the set C_B is closed in $\mathbb{R}^m$ for all $B \in \mathbb{B}$.

Let $\mathbf{b} \in C$. The definition of C then ensures that there exists some $\mathbf{x} \in \mathbb{R}^n$ that satisfies $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$. Thus the problem of determining $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$ has a feasible solution. It follows from Theorem 4.2 that there exists a basic feasible solution to this problem, and thus there exist distinct integers $j_1, j_2, \ldots, j_m$ between 1 and n and non-negative real numbers $s_1, s_2, \ldots, s_m$ such that $\mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}, \ldots, \mathbf{a}^{(j_m)}$ are linearly independent and

$$\mathbf{b} = \sum_{i=1}^{m} s_i \mathbf{a}^{(j_i)}.$$

Therefore $\mathbf{b} \in C_B$ where

$$B = \{j_1, j_2, \ldots, j_m\}.$$

We have thus shown that, given any element $\mathbf{b}$ of C, there exists a subset B of $\{1, 2, \ldots, n\}$ belonging to $\mathcal{B}$ for which $\mathbf{b} \in C_B$. It follows from this that the subset C of $\mathbb{R}^m$ is the union of the closed sets C_B taken over all elements B of the finite set $\mathcal{B}$. Thus C is a finite union of closed subsets of $\mathbb{R}^m$, and is thus itself a closed subset of $\mathbb{R}^m$, as required. ∎

5.6 Farkas' Lemma

Proposition 5.13 *Let C be a closed convex cone in $\mathbb{R}^m$ and let $\mathbf{b}$ be a vector in $\mathbb{R}^m$. Suppose that $\mathbf{b} \notin C$. Then there exists a linear functional $\varphi: \mathbb{R}^m \to \mathbb{R}$*

such that $\varphi(\mathbf{v}) \geq 0$ *for all* $\mathbf{v} \in C$ *and* $\varphi(\mathbf{b}) < 0$.

Proof Suppose that $\mathbf{b} \notin C$. The cone C is a closed convex set. It follows from Theorem 5.9 that there exists a linear functional $\varphi: \mathbb{R}^n \to \mathbb{R}$ and a real number c such that $\varphi(\mathbf{v}) > c$ for all $\mathbf{v} \in C$ and $\varphi(\mathbf{b}) < c$.

Now $\mathbf{0} \in C$, and $\varphi(\mathbf{0}) = 0$. It follows that $c < 0$, and therefore $\varphi(\mathbf{b}) \leq c < 0$.

Let $\mathbf{v} \in C$. Then $\lambda \mathbf{v} \in C$ for all real numbers λ satisfying $\lambda > 0$. It follows that $\lambda \varphi(\mathbf{v}) = \varphi(\lambda \mathbf{v}) > c$ and thus $\varphi(\mathbf{v}) > \dfrac{c}{\lambda}$ for all real numbers λ satisfying $\lambda > 0$, and therefore

$$\varphi(\mathbf{v}) \geq \lim_{\lambda \to +\infty} \frac{c}{\lambda} = 0.$$

We conclude that $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in C$.

Thus $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in C$ and $\varphi(\mathbf{b}) < 0$, as required. ∎

Lemma 5.14 (Farkas' Lemma) *Let* A *be a* $m \times n$ *matrix with real coefficients, and let* $\mathbf{b} \in \mathbb{R}^m$ *be an* m*-dimensional real vector. Then exactly one of the following two statements is true:—*

(i) there exists $\mathbf{x} \in \mathbb{R}^n$ *such that* $A\mathbf{x} = \mathbf{b}$ *and* $\mathbf{x} \geq \mathbf{0}$*;*

(ii) there exists $\mathbf{y} \in \mathbb{R}^m$ *such that* $\mathbf{y}^T A \geq \mathbf{0}$ *and* $\mathbf{y}^T \mathbf{b} < 0$.

Proof Let $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ be the vectors in $\mathbb{R}^m$ determined by the columns of the matrix A, so that $(\mathbf{a}^{(j)})_i = (A)_{i,j}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, and let

$$C = \left\{ \sum_{j=1}^{n} x_j \mathbf{a}^{(j)} : x_j \geq 0 \text{ for } j = 1, 2, \ldots, n \right\}.$$

It follows from Proposition 5.12 that C is a closed convex cone in $\mathbb{R}^m$. Moreover

$$C = \{ A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{x} \geq \mathbf{0} \}.$$

Thus $\mathbf{b} \in C$ if and only if there exists $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{b} = A\mathbf{x}$ and $\mathbf{x} \geq \mathbf{0}$. Therefore statement (i) in the statement of Farkas' Lemma is true if and only if $\mathbf{b} \in C$.

If $\mathbf{b} \notin C$ then it follows from Proposition 5.13 that there exists a linear functional $\varphi: \mathbb{R}^m \to \mathbb{R}$ such that $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in C$ and $\varphi(\mathbf{b}) < 0$. Then there exists $\mathbf{y} \in \mathbb{R}^m$ with the property that $\varphi(\mathbf{v}) = \mathbf{y}^T \mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^m$. Now $A\mathbf{x} \in C$ for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $\mathbf{x} \geq \mathbf{0}$. It follows that $\mathbf{y}^T A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $\mathbf{x} \geq \mathbf{0}$. In particular $(\mathbf{y}^T A)_i = \mathbf{y}^T A\mathbf{e}^{(i)} \geq 0$ for

$i = 1, 2, \ldots, m$, where $\mathbf{e}^{(i)}$ is the vector in $\mathbb{R}^m$ whose ith component is equal to 1 and whose other components are zero. Thus if $\mathbf{b} \notin C$ then there exists $\mathbf{y} \in \mathbb{R}^m$ for which $\mathbf{y}^T A \geq \mathbf{0}$ and $\mathbf{y}^T \mathbf{b} < 0$.

Conversely suppose that there exists $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T A \geq 0$ and $\mathbf{y}^T \mathbf{b} < 0$. Then $\mathbf{y}^T A \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $\mathbf{x} \geq \mathbf{0}$, and therefore $\mathbf{y}^T \mathbf{v} \geq 0$ for all $\mathbf{v} \in C$. But $\mathbf{y}^T \mathbf{b} < 0$. It follows that $\mathbf{b} \notin C$. Thus statement (ii) in the statement of Farkas's Lemma is true if and only if $\mathbf{b} \notin C$. The result follows. ∎

Corollary 5.15 *Let A be a $m \times n$ matrix with real coefficients, and let $\mathbf{c} \in \mathbb{R}^n$ be an n-dimensional real vector. Then exactly one of the following two statements is true:—*

(i) there exists $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T A = \mathbf{c}^T$ and $\mathbf{y} \geq \mathbf{0}$;

(ii) there exists $\mathbf{v} \in \mathbb{R}^n$ such that $A\mathbf{v} \geq \mathbf{0}$ and $\mathbf{c}^T \mathbf{v} < 0$.

Proof It follows on applying Farkas's Lemma to the transpose of the matrix A that exactly one of the following statements is true:—

(i) there exists $\mathbf{y} \in \mathbb{R}^m$ such that $A^T \mathbf{y} = \mathbf{c}$ and $\mathbf{y} \geq \mathbf{0}$;

(ii) there exists $\mathbf{v} \in \mathbb{R}^m$ such that $\mathbf{v}^T A^T \geq \mathbf{0}$ and $\mathbf{v}^T \mathbf{c} < 0$.

But $\mathbf{v}^T \mathbf{c} = \mathbf{c}^T \mathbf{v}$. Also $A^T \mathbf{y} = \mathbf{c}$ if and only if $\mathbf{y}^T A = \mathbf{c}^T$, and $\mathbf{v}^T A^T \geq \mathbf{0}$ if and only if $A\mathbf{v} \geq \mathbf{0}$. The result follows. ∎

Corollary 5.16 *Let A be a $m \times n$ matrix with real coefficients, and let $\mathbf{c} \in \mathbb{R}^n$ be an n-dimensional real vector. Suppose that $\mathbf{c}^T \mathbf{v} \geq 0$ for all $\mathbf{v} \in \mathbb{R}^n$ satisfying $A\mathbf{v} \geq 0$. Then there exists some there exists $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T A = \mathbf{c}^T$ and $\mathbf{y} \geq \mathbf{0}$.*

Proof Statement (ii) in the statement of Corollary 5.15 is false, by assumption, and therefore statement (i) in the statement of that corollary must be true. The result follows. ∎

Proposition 5.17 *Let n be a positive integer, let I be a non-empty finite set, let $\varphi: \mathbb{R}^n \to \mathbb{R}$ be a linear functional on $\mathbb{R}^n$, and, for each $i \in I$, let $\eta_i: \mathbb{R}^n \to \mathbb{R}$ be a linear functional on $\mathbb{R}^n$. Suppose that $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in \mathbb{R}^n$ with the property that $\eta_i(\mathbf{v}) \geq 0$ for all $i \in I$. Then there exist non-negative real numbers g_i for all $i \in I$ such that $\varphi = \sum_{i \in I} g_i \eta_i$.*

Proof We may suppose that $I = \{1, 2, \ldots, m\}$ for some positive integer m. For each $i \in I$ there exist real numbers $A_{i,1}, A_{i,2}, \ldots, A_{i,n}$ such that

$$\eta_i(v_1, v_2, \ldots, v_n) = \sum_{j=1}^{n} A_{i,j} v_j$$

for $i = 1, 2, \ldots, m$ and for all real numbers $v_1, v_2, \ldots, v_n$. Let A be the $m \times n$ matrix whose coefficient in the ith row and jth column is the real number $A_{i,j}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. Then an n-dimensional vector $\mathbf{v} \in \mathbb{R}^n$ satisfies $\eta_i(\mathbf{v}) \geq 0$ for all $i \in I$ if and only if $A\mathbf{v} \geq \mathbf{0}$.

There exists an n-dimensional vector $\mathbf{c} \in \mathbb{R}^n$ such that $\varphi(\mathbf{v}) = \mathbf{c}^T\mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^n$. Then $\mathbf{c}^T\mathbf{v} \geq 0$ for all $\mathbf{v} \in \mathbb{R}^n$ satisfying $A\mathbf{v} \geq 0$. It then follows from Corollary 5.16 that there exists $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T A = \mathbf{c}^T$ and $\mathbf{y} \geq \mathbf{0}$. Let $g_i = (\mathbf{y})_i$ for $i = 1, 2, \ldots, m$. Then $g_i \geq 0$ for $i = 1, 2, \ldots, m$ and $\sum\limits_{i \in I} g_i \eta_i = \varphi$, as required. ∎

Remark The result of Proposition 5.17 can also be viewed as a consequence of Proposition 5.13 applied to the convex cone in the dual space $\mathbb{R}^{n*}$ of the real vector space $\mathbb{R}^n$ generated by the linear functionals η_i for $i \in I$. Indeed let C be the subset of $\mathbb{R}^{n*}$ defined such that

$$C = \left\{ \sum_{i \in I} g_i \eta_i : g_i \geq 0 \text{ for all } i \in I \right\}.$$

It follows from Proposition 5.12 that C is a closed convex cone in the dual space $\mathbb{R}^{n*}$ of $\mathbb{R}^n$. If the linear functional φ did not belong to this cone then it would follow from Proposition 5.13 that there would exist a linear functional $V\colon \mathbb{R}^{n*} \to \mathbb{R}$ with the property that $V(\eta_i) \geq 0$ for all $i \in I$ and $V(\varphi) < 0$.

But given any linear functional on the dual space of a given finite-dimensional vector space, there exists some vector belonging to the given vector space such that the linear functional on the dual space evaluates elements of the dual space at that vector (see Corollary 2.7). It follows that there would exist $\mathbf{v} \in \mathbb{R}^n$ such that $V(\psi) = \psi(\mathbf{v})$ for all $\psi \in \mathbb{R}^{n*}$. But then $\eta_i(\mathbf{v}) \geq 0$ for all $i \in I$ and $\varphi(\mathbf{v}) < 0$. This contradicts the requirement that $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in \mathbb{R}^n$ satisfying $\eta_i(\mathbf{v}) \geq 0$ for all $i \in I$. To avoid this contradiction it must be the case that $\varphi \in C$, and therefore there must exist non-negative real numbers g_i for all $i \in I$ such that $\varphi = \sum_{i \in I} g_i \eta_i$.

Corollary 5.18 *Let n be a positive integer, let I be a non-empty finite set, let $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ be a linear functional on $\mathbb{R}^n$, and, for each $i \in I$, let $\eta_i\colon \mathbb{R}^n \to \mathbb{R}$ be a linear functional on $\mathbb{R}^n$. Suppose that there exists a subset I_0 of I such that*

$\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in \mathbb{R}^n$ with the property that $\eta_i(\mathbf{v}) \geq 0$ for all $i \in I_0$. Then there exist non-negative real numbers g_i for all $i \in I$ such that $\varphi = \sum_{i \in I} g_i \eta_i$ and $g_i = 0$ when $i \notin I_0$.

Proof It follows directly from Proposition 5.17 that there exist non-negative real numbers g_i for all $i \in I_0$ such that $\varphi = \sum_{i \in I_0} g_i \eta_i$. Let $g_i = 0$ for all $i \in I \setminus I_0$. Then $\varphi = \sum_{i \in I_0} g_i \eta_i$, as required. ∎

Definition A subset X is said to be a *convex polytope* if there exist linear functionals $\eta_1, \eta_2, \ldots, \eta_m$ on $\mathbb{R}^n$ and real numbers $s_1, s_2, \ldots, s_m$ such that

$$X = \{\mathbf{x} \in \mathbb{R}^n : \eta_i(\mathbf{x}) \geq s_i \text{ for } i = 1, 2, \ldots, m\}.$$

Let $(\eta_i : i \in I)$ be a finite collection of linear functionals on $\mathbb{R}^n$ indexed by a finite set I, let s_i be a real number for all $i \in I$, and let

$$X = \bigcap_{i \in I} \{\mathbf{x} \in \mathbb{R} : \eta_i(\mathbf{x}) \geq s_i\}.$$

Then X is a convex polytope in $\mathbb{R}^n$. A point $\mathbf{x}$ of $\mathbb{R}^n$ belongs to the convex polytope X if and only if $\eta_i(\mathbf{x}) \geq s_i$ for all $i \in I$.

Proposition 5.19 *Let n be a positive integer, let I be a non-empty finite set, and, for each $i \in I$, let $\eta_i \colon \mathbb{R}^n \to \mathbb{R}$ be non-zero linear functional and let s_i be a real number. Let X be the convex polytope defined such that*

$$X = \bigcap_{i \in I} \{\mathbf{x} \in \mathbb{R} : \eta_i(\mathbf{x}) \geq s_i\}.$$

(Thus a point $\mathbf{x}$ of $\mathbb{R}^n$ belongs to the convex polytope X if and only if $\eta_i(\mathbf{x}) \geq s_i$ for all $i \in I$.) Let $\varphi \colon \mathbb{R}^n \to \mathbb{R}$ be a non-zero linear functional on $\mathbb{R}^n$, and let $\mathbf{x}^* \in X$. Then $\varphi(\mathbf{x}^*) \leq \varphi(\mathbf{x})$ for all $\mathbf{x} \in X$ if and only if there exist non-negative real numbers g_i for all $i \in I$ such that $\varphi = \sum_{i \in I} g_i \eta_i$ and $g_i = 0$ whenever $\eta_i(\mathbf{x}^*) > s_i$.

Proof Let $K = \{i \in I : \eta_i(\mathbf{x}^*) > s_i\}$. Suppose that there do not exist non-negative real numbers g_i for all $i \in I$ such that $\varphi = \sum_{i \in I} g_i \eta_i$ and $g_i = 0$ when $i \in K$. Corollary 5.18 then ensures that there must exist some $\mathbf{v} \in \mathbb{R}^n$ such that $\eta_i(\mathbf{v}) \geq 0$ for all $i \in I \setminus K$ and $\varphi(\mathbf{v}) < 0$. Then

$$\eta_i(\mathbf{x}^* + \lambda \mathbf{v}) = \eta_i(\mathbf{x}^*) + \lambda \eta_i(\mathbf{v}) \geq s_i$$

for all $i \in I \setminus K$ and for all $\lambda \geq 0$. If $i \in K$ then $\eta_i(\mathbf{x}^*) > s_i$. The set K is finite. It follows that there must exist some real number λ_0 satisfying $\lambda_0 > 0$ such that $\eta_i(\mathbf{x}^* + \lambda\mathbf{v}) \geq s_i$ for all $i \in K$ and for all real numbers λ satisfying $0 \leq \lambda \leq \lambda_0$.

Combining the results in the cases when $i \in I \setminus K$ and when $i \in K$, we find that $\eta_i(\mathbf{x}^* + \lambda\mathbf{v}) \geq s_i$ for all $i \in I$ and $\lambda \in [0, \lambda_0]$, and therefore $\mathbf{x}^* + \lambda\mathbf{v} \in X$ for all real numbers λ satisfying $0 \leq \lambda \leq \lambda_0$. But

$$\varphi(\mathbf{x}^* + \lambda\mathbf{v}) = \varphi(\mathbf{x}^*) + \lambda\varphi(\mathbf{v}) < \varphi(\mathbf{x}^*)$$

whenever $\lambda > 0$. It follows that the linear functional φ cannot attain a minimum value in X at any point $\mathbf{x}^*$ for which either $K = I$ or for which K is a proper subset of I but there exist non-negative real numbers g_i for all $i \in I \setminus K$ such that $\varphi = \sum_{i \in I \setminus K} g_i \eta_i$. The result follows. ∎

5.7 Strong Duality

Example Consider again the following linear programming problem in general primal form:—

> *find values of x_1, x_2, x_3 and x_4 so as to minimize the objective function*
>
> $$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4$$
>
> *subject to the following constraints:—*
>
> - $a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + a_{1,4}x_4 = b_1$;
> - $a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + a_{2,4}x_4 = b_2$;
> - $a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 + a_{3,4}x_4 \geq b_3$;
> - $x_1 \geq 0$ *and* $x_3 \geq 0$.

Now the constraint

$$a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + a_{1,4}x_4 = b_1$$

can be expressed as a pair of inequality constraints as follows:

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + a_{1,4}x_4 &\geq b_1 \\ -a_{1,1}x_1 - a_{1,2}x_2 - a_{1,3}x_3 - a_{1,4}x_4 &\geq -b_1. \end{aligned}$$

Similarly the equality constraint involving b_2 can be expressed as a pair or inequality constraints.

Therefore the problem can be reformulated as follows:—

find values of x_1, x_2, x_3 and x_4 so as to minimize the objective function

$$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4$$

subject to the following constraints:—

- $a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + a_{1,4}x_4 \geq b_1$;
- $-a_{1,1}x_1 - a_{1,2}x_2 - a_{1,3}x_3 - a_{1,4}x_4 \geq -b_1$;
- $a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + a_{2,4}x_4 \geq b_2$;
- $-a_{2,1}x_1 - a_{2,2}x_2 - a_{2,3}x_3 - a_{2,4}x_4 \geq -b_2$;
- $a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 + a_{3,4}x_4 \geq b_3$;
- $x_1 \geq 0$;
- $x_3 \geq 0$.

Let

$$\varphi(x_1, x_2, x_3, x_4) = c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4,$$

and let

$$\begin{aligned}
\eta_1^+(x_1, x_2, x_3, x_4) &= a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + a_{1,4}x_4, \\
\eta_1^-(x_1, x_2, x_3, x_4) &= -\eta_1(x_1, x_2, x_3, x_4), \\
\eta_2^+(x_1, x_2, x_3, x_4) &= a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + a_{2,4}x_4, \\
\eta_2^-(x_1, x_2, x_3, x_4) &= -\eta_3(x_1, x_2, x_3, x_4), \\
\eta_3(x_1, x_2, x_3, x_4) &= a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 + a_{3,4}x_4, \\
\zeta_1(x_1, x_2, x_3, x_4) &= x_1, \\
\zeta_3(x_1, x_2, x_3, x_4) &= x_3,
\end{aligned}$$

Then (x_1, x_2, x_3, x_4) is a feasible solution to the primal problem if and only if this element of $\mathbb{R}^4$ belongs to the convex polytope X, where X is the subset of $\mathbb{R}^4$ consisting of all points $\mathbf{x}$ of $\mathbb{R}^4$ that satisfy the following constraints:—

- $\eta_1^+(\mathbf{x}) \geq b_1$;
- $\eta_1^-(\mathbf{x}) \geq -b_1$;
- $\eta_2^+(\mathbf{x}) \geq b_2$;
- $\eta_2^-(\mathbf{x}) \geq -b_2$;
- $\eta_3(\mathbf{x}) \geq b_3$;

- $\zeta_1(\mathbf{x}) \geq 0$;
- $\zeta_3(\mathbf{x}) \geq 0$.

An inequality constraint is said to be *binding* for a particular feasible solution $\mathbf{x}$ if equality holds in that constraint at the feasible solution. Thus the constraints on the values of η_1^+, η_1^-, η_2^+ and η_2^- are always binding at points of the convex polytope X, but the constraints determined by η_3, ζ_1 and ζ_3 need not be binding.

Suppose that the linear functional φ attains its minimum value at a point $\mathbf{x}^*$ of X, where $\mathbf{x}^* = (x_1^*, x_2^*, x_3^*, x_4^*)$. It then follows from Proposition 5.19 that there exist non-negative real numbers p_1^+, p_1^-, p_2^+, p_2^-, p_3, q_1 and q_3 such that

$$p_1^+\eta_1^+ + p_1^-\eta_1^- + p_2^+\eta_2^+ + p_2^-\eta_2^- + p_3\eta_3 + q_1\zeta_1 + q_3\zeta_3 = \varphi.$$

Moreover $p_3 = 0$ if $\eta_3(\mathbf{x}^*) > b_3$, $q_1 = 0$ if $\zeta_1(\mathbf{x}^*) > 0$, and $q_3 = 0$ if $\zeta_3(\mathbf{x}^*) > 0$.

Now $\eta_1^- = -\eta_1^+$ and $\eta_2^- = -\eta_2^+$. It follows that

$$p_1\eta_1^+ + p_2\eta_2^+ + p_3\eta_3 + q_1\zeta_1 + q_3\zeta_3 = \varphi,$$

where $p_1 = p_1^+ - p_1^-$ and $p_2 = p_2^+ - p_2^-$. Moreover $p_3 = 0$ if $\sum_{i=1}^{4} a_{3,j}x_j^* > b_3$, $q_1 = 0$ if $x_1^* > 0$, and $q_3 = 0$ if $x_3^* > 0$.

It follows that

$$\begin{aligned}
p_1a_{1,1} + p_2a_{2,1} + p_3a_{3,1} &\leq c_1,\\
p_1a_{1,2} + p_2a_{2,2} + p_3a_{3,2} &= c_2,\\
p_1a_{1,3} + p_2a_{2,3} + p_3a_{3,3} &\leq c_3,\\
p_1a_{1,4} + p_2a_{2,4} + p_3a_{3,4} &= c_4,\\
p_3 &\geq 0.
\end{aligned}$$

Moreover $p_3 = 0$ if $\sum_{i=1}^{4} a_{3,j}x_j^* > b_3$, $\sum_{i=1}^{3} p_i a_{i,1} = c_1$ if $x_1^* > 0$, and $\sum_{i=1}^{3} p_i a_{i,3} = c_3$ if $x_3^* > 0$. It follows that (p_1, p_2, p_3) is a feasible solution of the dual problem to the feasible primal problem.

Moreover the complementary slackness conditions determined by the primal problem are satisfied. It therefore follows from the Weak Duality Theorem (Theorem 5.5) that (p_1, p_2, p_3) is an optimal solution to the dual problem.

Theorem 5.20 (Strong Duality for Linear Programming Problems with Optimal Solutions)
Let $\mathbf{x}^* \in \mathbb{R}^{\kappa}$ be an optimal solution to a linear programming problem

$$\text{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$$

expressed in general primal form with constraint matrix A with m rows and n columns, target vector $\mathbf{b}$, cost vector $\mathbf{c}$, inequality constraint specifier I^+ and variable sign specifier J^+. Then there exists an optimal solution $\mathbf{p}^*$ to the corresponding dual programming problem

$$\mathrm{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+),$$

and moreover $\mathbf{p}^{*T}\mathbf{b} = \mathbf{c}^T\mathbf{x}^*$.

Proof Let $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)$, and let $A_{i,j} = (A)_{i,j}$, $b_i = (\mathbf{b})_i$ and $c_j = (\mathbf{c})_j$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. Then optimal solution $\mathbf{x}^*$ minimizes $\mathbf{c}^T\mathbf{x}^*$ subject to the following constraints:—

- $A\mathbf{x}^* \geq \mathbf{b}$;
- $(A\mathbf{x}^*)_i = b_i$ unless $i \in I^+$;
- $x_j^* \geq 0$ for all $j \in J^+$.

Let $\mathbf{p}$ be a feasible solution to the dual problem, and let $\mathbf{p} = (p_1, p_2, \ldots, p_m)$. Then $\mathbf{p}$ must satisfy the following constraints:—

- $\mathbf{p}^T A \leq \mathbf{c}^T$;
- $p_i \geq 0$ for all $i \in I^+$;
- $(\mathbf{p}^T A)_j = c_j$ unless $j \in J^+$.

Now the constraints of the primal problem can be expressed in inequality form as follows:—

- $(A\mathbf{x}^*)_i \geq b_i$ for all $i \in I^+$;
- $(A\mathbf{x}^*)_i \geq b_i$ for all $i \in I \setminus I^+$; $(-A\mathbf{x}^*)_i \geq -b_i$ for all $i \in I \setminus I^+$;
- $x_j^* \geq 0$ for all $j \in J^+$.

Let

$$\varphi(x_1, x_2, \ldots, x_n) = \sum_{j=1}^{n} c_j x_j,$$

$$\eta_i(x_1, x_2, \ldots, x_n) = \sum_{j=1}^{n} (A)_{i,j} x_j \quad (i = 1, 2, \ldots, m)$$

$$\zeta_j(x_1, x_2, \ldots, x_n) = x_j \quad (j = 1, 2, \ldots, n)$$

It follows from Proposition 5.19 that if there exists an optimal solution to the primal problem then there exist non-negative quantities p_i for all $i \in I^+$, p_i^+ and p_- for all $i \in I \setminus I^+$ and q_j for all $j \in J^+$ such that

$$\varphi = \sum_{i \in I^+} p_i \eta_i + \sum_{i \in I \setminus I^+} (p_i^+ - p_i^-) \eta_i + \sum_{j \in J^+} q_j \zeta_j.$$

Moreover $p_i = 0$ whenever $i \in I^+$ and $\eta_i(x_1^*, x_2^*, \ldots, x_n^*)_i > b_i$ and $q_j = 0$ whenever $x_j^* > 0$. Let $\mathbf{p}^* \in \mathbb{R}^m$ be defined such that $(\mathbf{p}^*)_i = p_i$ for all $i \in I^+$ and $(\mathbf{p}^*)_i = p_i^+ - p_i^-$ for all $i \in I \setminus I^+$. Then $(\mathbf{p}^{*T} A)_j \leq c_j$ for $j = 1, 2, \ldots, n$, $(\mathbf{p}^*)_i \geq 0$ for all $i \in I^+$, and $(\mathbf{p}^{*T} A)_j = c_j$ unless $j \in J^+$. Moreover $(\mathbf{p}^*)_i = 0$ whenever $(A\mathbf{x}^*)_i > b_i$ and $q_i = 0$ whenever $x_j > 0$. It follows that $\mathbf{p}^*$ is a feasible solution of the dual problem. Moreover the relevant complementary slackness conditions are satisfied by $\mathbf{x}^*$ and $\mathbf{p}^*$. It is then a consequence of the Weak Duality Theorem that $\mathbf{c}^T \mathbf{x}^* = \mathbf{p}^{*T} \mathbf{b}$, and that therefore $\mathbf{p}^*$ is an optimal solution of the dual problem (see Corollary 5.6). The result follows. ∎