C&O 350 [1]

1 LPs in Equality Form

A maximization LP in which all the variables are required to be non-negative, and all the constraints, except the non-negativity restrictions, are equalities is said to be in **standard equality form** or simply in **equality form**[2]. Thus, an LP in equality form is one of the following type:

$$\left.\begin{array}{llcll} \text{Maximize} & \sum_{j=1}^{n} c_j x_j & & & \\ \text{subject to} & \sum_{j=1}^{n} a_{ij} x_j & = & b_i & (i = 1, 2, \cdots, m) \\ & x_j & \geq & 0 & (j = 1, 2, \cdots, n) \end{array}\right\} \qquad (1)$$

The above LP may also be stated in matrix notation as follows:

$$\begin{array}{rrcl} \text{maximize} & \mathbf{cx} & & \\ \text{subject to} & \mathbf{Ax} & = & \mathbf{b} \\ & \mathbf{x} & \geq & \mathbf{0} \end{array}$$

where $\mathbf{A}$ is an m by n matrix, the vector $\mathbf{c}$ is a 1 by n row vector[3], $\mathbf{b}$ is an m by 1 column vector, and $\mathbf{x}$ is an n by 1 column vector, as shown below:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{pmatrix}$$

$$\begin{array}{lcl} \mathbf{c} & = & (c_1, \quad c_2, \quad \ldots \quad c_n) \\ \mathbf{x}^t & = & (x_1, \quad x_2, \quad \ldots \quad x_n) \end{array}$$

Note that $\mathbf{x}$ is a column vector; the vector $\mathbf{x}^t$ displayed above is its transpose. We refer to A as the **matrix of the LP**, $\mathbf{c}$ as the **cost vector**, and $\mathbf{b}$ as the **constant vector**.

[1] If you find any mistakes or misprints, please bring them to the attention of one of: whcunnin@math, usrmurty@math, hwolkowicz@orion.

[2] This concept is not used in Chvátal.

[3] For compatibility of matrix multiplication, we should make sure that two matrices being multiplied have compatible orders. Thus if we want to multiply a vector $\mathbf{c}$ in $\mathbf{R}^n$ by another vector $\mathbf{x}$ in $\mathbf{R}^n$, in that order, then $\mathbf{c}$ must be a row vector and $\mathbf{x}$ a column vector. Some authors regard all vectors as column vectors. In that case, we must transpose appropriately to ensure compatibility. For example, if $\mathbf{c}$ and $\mathbf{x}$ are both taken as column vectors, then we must write $\mathbf{c}^t\mathbf{x}$, where $\mathbf{c}^t$ is the transpose of $\mathbf{c}$.

The following is an example of an LP in equality form.

$$\begin{array}{lrcrcrcrcl}
\max & 4x_1 & - & 4x_2 & + & 3x_3 & + & 3x_4 & & \\
\text{s.t.} & x_1 & - & 3x_2 & + & x_3 & + & 9x_4 & = & 4 \\
& 4x_1 & + & x_2 & + & 2x_3 & + & 3x_4 & = & 3 \\
& 3x_1 & + & 4x_2 & + & 2x_3 & + & 6x_4 & = & 1 \\
& x_1 & , & x_2 & , & x_3 & , & x_4 & \geq & 0
\end{array}$$

An LP in standard (inequality) form[4] may be converted to one in equality form by adding slack variables. In fact, for solving an LP in inequality form by the simplex method, we first convert it to an LP in equality form by adding slack variables, and apply the simplex method to the LP in equality form. Consider for example, the following LP in standard (inequality) form

$$\left.\begin{array}{lrcrcrcl}
\max\ z = & 6x_1 & + & 4x_2 & + & 2x_3 & & \\
\text{s.t.} & 2x_1 & + & x_2 & + & x_3 & \leq & 4 \\
& 4x_1 & + & 2x_2 & + & 2x_3 & \leq & 12 \\
& 3x_1 & + & \frac{5}{2}x_2 & + & 2x_3 & \leq & 8 \\
& x_1 & , & x_2 & , & x_3 & \geq & 0
\end{array}\right\} \qquad (2)$$

We may convert this LP to the following equality LP by introducing the slack variables x_4, x_5, and x_6:

$$\left.\begin{array}{lrcrcrcrcrcrcl}
\max\ z = & 6x_1 & + & 4x_2 & + & 2x_3 & & & & & & & & \\
\text{s.t.} & 2x_1 & + & x_2 & + & x_3 & + & x_4 & & & & & = & 4 \\
& 4x_1 & + & 2x_2 & + & 2x_3 & & & + & x_5 & & & = & 12 \\
& 3x_1 & + & \frac{5}{2}x_2 & + & 2x_3 & & & & & + & x_6 & = & 8 \\
& x_1 & , & x_2 & , & x_3 & , & x_4 & , & x_5 & , & x_6 & \geq & 0
\end{array}\right\} \qquad (3)$$

For this LP:

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 4 & 2 & 2 & 0 & 1 & 0 \\ 3 & \frac{5}{2} & 2 & 0 & 0 & 1 \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} 4 \\ 12 \\ 8 \end{pmatrix}$$

$$\begin{array}{rcl}
\mathbf{c} & = & (6,\ 4,\ 2,\ 0,\ 0,\ 0\ ) \\
\mathbf{x}^t & = & (x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6)
\end{array}$$

The purpose of this handout is to provide definitions of the concepts such as bases, basic solutions, dictionaries, etc., used in the development of the simplex method. These definitions are given for equality LPs.

A knowledge of the fundamental notions of linear algebra is necessary for understanding the theory of linear programming, in particular, the development of the simplex method. Students should be familiar with the concepts of linear independence of vectors, bases and dimension of a vector space, ranks of matrices, and the relationship of these notions with solution sets of systems of linear equations.

[4]In the early chapters of the textbook, only LPs in inequality form are considered.

2 Equivalence

In the simplex method, we start with a system of linear equations. In each iteration, this system is transformed to a new system by means of elementary row operations. Although, in each iteration, we deal with a different looking system of equations, they all "carry the same information". To make this precise, we need the following important notion.

1 *Definition: Two systems of linear equations are said to be* **equivalent**[5] *if they have the same set of solutions.*

3 Notation

Consider a system
$$\mathbf{Ax} = \mathbf{b}$$
of linear equations where $\mathbf{A}$ is an $m \times n$ matrix. This system may also be rewritten as
$$x_1 A_1 + x_2 A_2 + \ldots + x_n A_n = b$$
where A_j is the jth column of the matrix A. (This shows the correspondence between the columns of $\mathbf{A}$ and the n variables.)

In studying linear equations, it is often convenient to group terms involving subsets of variables. To express such grouping, the following matrix notation is convenient.

For any subset J of the column index set $\{1, 2, ..., n\}$, we shall let $\mathbf{A}_J$ denote the submatrix of $\mathbf{A}$ consisting of the columns with their indices in J. (If $J = (j_1, j_2, ..., j_k)$ is an ordered subset, then $\mathbf{A}_J$ denotes the matrix $(A_{j_1}, A_{j_2}, ..., A_{j_k})$. If there is no specified ordering of the elements of J, then then the columns of $\mathbf{A}_J$ may be listed in any order.)

2 *Example: Consider the LP*
$$\max\ z = \mathbf{cx},\ \text{s.t. } \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0},$$
where
$$\mathbf{A} = \begin{pmatrix} 3 & -6 & 3 & 6 \\ 4 & 1 & 5 & 10 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 12 \\ 25 \end{pmatrix}$$
and
$$\mathbf{c} = (-7, 5, -10, -19).$$
If $B = (1, 2)$ and $N = (3, 4)$, then
$$\mathbf{A}_B = \begin{pmatrix} 3 & -6 \\ 4 & 1 \end{pmatrix}, \quad \mathbf{x}_B = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{A}_N = \begin{pmatrix} 3 & 6 \\ 5 & 10 \end{pmatrix}, \quad \mathbf{x}_N = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix}$$

[5] The dictionaries encountered during the course of an application of the simplex method are all equivalent.

and

$$\mathbf{c}_B = (-7, 5) \quad \text{and} \quad \mathbf{c}_N = (-10, -19).$$

Now the system $\mathbf{Ax} = \mathbf{b}$may be rewritten as

$$\mathbf{A}_B \mathbf{x}_B + \mathbf{A}_N \mathbf{x}_N = \mathbf{b}.$$

This amounts to grouping, in each equation, the terms involving the variables x_1 and x_2 together, and the terms involving the variables x_3 and x_4 together. Similarly,

$$z = \mathbf{cx} = \mathbf{c}_B \mathbf{x}_B + \mathbf{c}_N \mathbf{x}_N.$$

4 Bases

Clearly, a system $\mathbf{Ax} = \mathbf{b}$ is consistent if, and only if, $\mathbf{b}$ is in the column space of the matrix $\mathbf{A}$. In fact, every solution $\mathbf{s}$ corresponds to an expression ($\mathbf{b} = \sum_{i=1}^{n} s_j A_j$) of $\mathbf{b}$ as a linear combination of the columns of $\mathbf{A}$. In particular, $\mathbf{Ax} = \mathbf{b}$ has a solution if, and only if, $\mathbf{b}$ can be expressed as a linear combination of any basis of the column space of $\mathbf{A}$. This observation leads to the central notion of this subject, namely, the notion of a basis.

Consider a system $\mathbf{Ax} = \mathbf{b}$ of linear equations with rank($\mathbf{A}$)$= m$. (See footnote[6])

3 Definition: *An m-element subset B of the column index set $\{1, 2, ..., n\}$ is called a* **basis** *if the m columns of the submatrix $\mathbf{A}_B$ are linearly independent (equivalently, if $\mathbf{A}_B$ is invertible). Note that when B is a basis, the m columns of $\mathbf{A}_B$ constitute a basis for the column space ($\mathbf{R}^m$) of $\mathbf{A}$; this is the reason for calling B a basis.*

4 Definition: *An* **ordered basis** *is a basis with a specified ordering of its elements. If B is an ordered basis, the matrix $\mathbf{A}_B$ is called the* **basis matrix** *corresponding to B.*

5 Definition: *If B is a basis, the variables with subscripts in the set are* **basic variables** *with respect B, and the remaining variables, i.e. those with subscripts not in B are the* **non-basic variables**.

For convenience, we shall denote the complement of B, i.e. the set of the subscripts of non-basic variables, by N.

The vector of variables restricted to B is $\mathbf{x}_B$, and vector of variables restricted to N is $\mathbf{x}_N$. Analogously, $\mathbf{A}_B$ and $\mathbf{c}_B$ are the restrictions of $\mathbf{A}$ and $\mathbf{c}$, respectively, to columns with

[6] Although the systems of equations associated with LPs may be either inconsistent or redundant, the systems that we use in the computational procedures are both consistent and not redundant. (For example, when we add slack variables to an inequality LP, or, when we add artificial variables in Phase 1, we obtain such a system.) For this reason, in studying properties of linear equations relevant to the development of the simplex method, we shall assume that we have a system $\mathbf{Ax} = \mathbf{b}$ which is both consistent and irredundant. This amounts to assuming that the matrix $\mathbf{A}$ is of full rank (i.e. rank($\mathbf{A}$)$= m =$ the number of rows of $\mathbf{A}$).

subscripts in B, and $\mathbf{A}_N$ and $\mathbf{c}_N$ are the restrictions of $\mathbf{A}$ and $\mathbf{c}$, respectively, to the columns with subscripts in N.

By grouping variables according to whether or not they are basic, we may rewrite the system $\mathbf{Ax} = \mathbf{b}$ as follows:

$$\mathbf{A}_B\mathbf{x}_B + \mathbf{A}_N\mathbf{x}_N = \mathbf{b}$$

Note that given any basis B, $\mathbf{b}$ can be expressed as a unique linear combination of the columns of $\mathbf{A}_B$. Thus, there is a unique solution to $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{x}_N = \mathbf{0}$, namely, $\mathbf{x}_B = \mathbf{A}_B^{-1}\mathbf{b}$. This solution is called the **basic solution** corresponding to the basis B.

Exercise: Determine, by inspection, all the bases of the following two systems.

$$(i) \qquad \begin{array}{rcrcrcrcr} x_1 & + & x_2 & + & 2x_3 & + & x_3 & = & 1 \\ 2x_1 & + & 2x_2 & + & 4x_3 & + & x_3 & = & 2 \end{array}$$

$$(ii) \quad \begin{array}{rcrcrcrcrcr} x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & = & 7 \\ x_1 & + & 2x_2 & + & 2x_3 & + & 3x_4 & - & 2x_5 & = & 3 \\ 2x_1 & + & 3x_2 & + & 3x_3 & + & 4x_4 & - & x_5 & = & 10 \end{array}$$

5 The system of equations representing an LP

Suppose that we have an equality LP of the form:

$$\max\ z = \mathbf{cx},\ \text{s.t. } \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}$$

where $\mathbf{A}$ is an $m \times n$ matrix. By treating z as a variable, we may associate with this LP, the system

$$\left.\begin{array}{ccccc} & & \mathbf{Ax} & = & \mathbf{b} \\ \hline z & - & \mathbf{cx} & = & 0 \end{array}\right\}$$

of $m + 1$ equations in $n + 1$ variables. (The first m equations of the system represent the equality constraints. The last equation represents the objective function; it is separated from the other equations by a horizontal line because the objective function has a special status.) We may in fact view this system as representing the statement of the LP with the understanding that z is to be maximized, and that the variables x_1, x_2,...,x_n are required to be non-negative.

The augmented coefficient matrix of the system (5) associated with the LP is

$$\left(\begin{array}{c|c|c} \mathbf{0} & \mathbf{A} & \mathbf{b} \\ \hline 1 & -\mathbf{c} & 0 \end{array}\right).$$

If there is a partition of the index set $\{1, 2, ..., n\}$ into two sets B and N, the grouping of terms involving variables with subscripts depending on whether or not they belong to B or N may be indicated by grouping the columns in the augmented coefficient matrix as follows:

$$\left(\begin{array}{c|c|c|c} \mathbf{0} & \mathbf{A}_B & \mathbf{A}_N & \mathbf{b} \\ \hline 1 & -\mathbf{c}_B & -\mathbf{c}_N & 0 \end{array}\right) \tag{4}$$

6 *Example: Consider the following system of equations:*

$$\begin{array}{cccccccl} & x_4 & & + & 2x_1 & - & 4x_3 & = 5 \\ & & x_2 & + & 3x_1 & + & 2x_3 & = 3 \\ \hline z & & & - & 4x_1 & + & 5x_3 & = 6 \end{array}$$

We may represent this system in matrix notation as follows:

$$\begin{array}{cccccc} & \mathbf{I}_2\mathbf{x}_B & + & \overline{\mathbf{A}}_N\mathbf{x}_N & = & \overline{\mathbf{b}} \\ \hline z & & - & \overline{\mathbf{c}}_N\mathbf{x}_N & = & \overline{v} \end{array}$$

where $B = (4, 2)$, and $N = (1, 3)$,

$$\mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{x}_B = \begin{pmatrix} x_4 \\ x_2 \end{pmatrix}, \quad \overline{\mathbf{A}}_N = \begin{pmatrix} 2 & -4 \\ 3 & 2 \end{pmatrix}, \quad \mathbf{x}_N = \begin{pmatrix} x_1 \\ x_3 \end{pmatrix}, \quad \overline{\mathbf{b}} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$$

and

$$\overline{\mathbf{c}}_N = (4, -5), \quad \overline{v} = 6$$

Because $\mathbf{I}_2\mathbf{x}_B = \mathbf{x}_B$, this system may be represented simply, as

$$\begin{array}{cccccc} & \mathbf{x}_B & + & \overline{\mathbf{A}}_N\mathbf{x}_N & = & \overline{\mathbf{b}} \\ \hline z & & - & \overline{\mathbf{c}}_N\mathbf{x}_N & = & \overline{v} \end{array}$$

By rearranging terms, we may also represent the system as

$$\begin{array}{ccccc} \mathbf{x}_B & = & \overline{\mathbf{b}} & - & \overline{\mathbf{A}}_N\mathbf{x}_N \\ \hline z & = & \overline{v} & + & \overline{\mathbf{c}}_N\mathbf{x}_N \end{array}$$

6 Dictionaries

7 *Definition: Let $B = (j_1, j_2, ..., j_m)$ be an ordered index set. Then, a system of equations of the form*

$$D : \left\{ \begin{array}{ccl} x_i & = & \overline{b}_i - \sum_{j \notin B} \overline{a}_{ij} x_j \quad for \quad j_1 \le i \le j_m \\ \hline z & = & \overline{v} + \sum_{j \notin B} \overline{c}_j x_j \end{array} \right.$$

is called a **dictionary** *with respect to B.*

(In matrix notation, D may be represented as

$$\begin{array}{rcccl} \mathbf{x}_B & = & \overline{\mathbf{b}} & - & \overline{\mathbf{A}}_N \mathbf{x}_N \\ \hline z & = & \overline{v} & + & \overline{\mathbf{c}}_N \mathbf{x}_N \end{array}$$

The system

$$\begin{array}{ccccc} & \mathbf{x}_B & + & \overline{\mathbf{A}}_N \mathbf{x}_N & = & \overline{\mathbf{b}} \\ \hline z & & - & \overline{\mathbf{c}}_N \mathbf{x}_N & = & \overline{v} \end{array}$$

obtained by rearranging the terms in D is a **tableau** with respect to B.)

The variables x_i, $i \in B$ are referred to as the basic variables, and the variables x_j, $j \notin B$ as the non-basic variables. The reason for this nomenclature is the following:

8 *Theorem: If a dictionary D with respect to B is equivalent to the system of linear equations representing an LP*

$$\max \mathbf{cx} \quad \mathbf{Ax} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0},$$

then B must be a basis of $\mathbf{Ax} = \mathbf{b}$.

Proof: When we set $\mathbf{x}_N = \mathbf{0}$ in D, we obtain the unique solution $\mathbf{x}_B = \overline{\mathbf{b}}$, $\mathbf{x}_N = \mathbf{0}$, and $z = \overline{v}$. Thus, if D is equivalent to the system of equations representing the LP, $\mathbf{x}_B = \overline{\mathbf{b}}$, $\mathbf{x}_N = \mathbf{0}$ must be the unique solution to the system $\mathbf{Ax} = \mathbf{b}$ when we set $\mathbf{x}_N = \mathbf{0}$. This, in turn, means that $\mathbf{x}_B = \overline{\mathbf{b}}$ must be the unique solution to the system $\mathbf{A}_B \mathbf{x}_B = \mathbf{b}$. But this is possible only if B is a basis of $\mathbf{Ax} = \mathbf{b}$. ■

9 *Definitions: In view of the above theorem, we refer to the index set B as a basis, and the variables x_i, $i \in B$ as the* **basic variables**, *and the variables x_j, $j \notin B$ as the* **non-basic variables** *in D, and the solution $\mathbf{x}_B = \overline{\mathbf{b}}$, $\mathbf{x}_N = \mathbf{0}$ as the* **basic solution** *associated with D.*

A dictionary D is **feasible**[7] *if $\overline{\mathbf{b}} \geq \mathbf{0}$.*

A dictionary D satisfies the optimality condition if $\overline{\mathbf{c}}_N \leq \mathbf{0}$.

A feasible dictionary D which satisfies the optimality condition is an **optimal** *dictionary. (In this case, the basic feasible solution associated with the dictionary is an optimal solution for the LP.)*

The converse of the above theorem is also true. That is, given any ordered basis B of $\mathbf{A}$, there is a dictionary D with respect to B that is equivalent to the system of equations representing the LP. Such a dictionary can be obtained using algebraic operations. This derivation will be discussed in the next section. We conclude this section with the following important theorem.

10 *Theorem: If D and D' are two equivalent dictionaries with respect to the same ordered basis B, then D and D' are identical.*

(A proof of this theorem appears on page 33 of the textbook.)

[7] In order to apply the simplex method to solve an LP, we must have a feasible dictionary equivalent to the given LP as the input. In fact, all the dictionaries encountered during the course of an application of the simplex method are feasible; the pivot rules are designed to maintain feasibility. (Note that, if a feasible dictionary is available, the question of infeasibility does not arise.)

7 The dictionary corresponding to an ordered basis

11 <u>*Theorem:*</u> *Consider an LP of the form:*

$$\max\ \mathbf{cx},\ \ s.t.\ \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq 0.$$

in which the rank of $\mathbf{A}$ *is* m*, and let* B *be an ordered basis of* $\mathbf{Ax} = \mathbf{b}$*. Then there is a unique dictionary with respect to* B *which is equivalent to the system of equations associated with the LP.*

The required dictionary can be obtained by performing elementary row operations on the system of equations corresponding to the LP, or equivalently, by suitable matrix operations. First, we shall illustrate the former method by considering the LP in example (2). We shall then establish the validity of the theorem in general by describing the algebraic procedure in terms of matrices.

12 <u>*Example:*</u> *Consider the following system of equations associated with the LP in example (2):*

$$\begin{array}{rrrrrrrrrrr}
 & 3x_1 & - & 6x_2 & + & 3x_3 & + & 6x_4 & = & 12 \\
 & 4x_1 & + & x_2 & + & 5x_3 & + & 10x_4 & = & 25 \\
\hline
z+ & 7x_1 & - & 5x_2 & + & 10x_3 & + & 19x_4 & = & 0
\end{array}$$

To obtain the tableau with respect to the basis $(1,2)$, we must transform the equations so that x_1 appears with coefficient 1 in the first equation, and with coefficient 0 in the other equations, and x_2 appears with coefficient 1 in the second equation, and with coefficient 0 in the other equations. To achieve this, first let us divide the first equation by 3, the coefficient of x_1 in the first equation, and then subtract suitable multiples of the first equation from the other equations to eliminate x_1 from them. (We say that we are **pivoting** on the coefficient of x_1 in equation 1.) This gives us the following system:

$$\begin{array}{rrrrrrrrrr}
 & x_1 & - & 2x_2 & + & x_3 & + & 2x_4 & = & 4 \\
 & & & 9x_2 & + & x_3 & + & 2x_4 & = & 9 \\
\hline
z & & & 9x_2 & + & 3x_3 & + & 5x_4 & = & -28
\end{array}$$

Now devide the second equation by 9, the coefficient of x_2 in the second equation, and then subtract suitable multiples of the first equation from the other equations to eliminate x_2 from them. (That is, pivot on the coefficient of x_2 in the second equation.) This gives us the following system:

$$\begin{array}{rrrrrrrrr}
 & x_1 & & + & \frac{11}{9}x_3 & + & \frac{22}{9}x_4 & = & 6 \\
 & & x_2 & + & \frac{1}{9}x_3 & + & \frac{2}{9}x_4 & = & 1 \\
\hline
z & & & + & 2x_3 & + & 3x_4 & = & -37
\end{array}$$

To obtain the dictionary, it is a simply a matter of rearranging terms:

$$\begin{array}{rcrcrcr}
x_1 & = & 6 & - & \frac{11}{9}x_3 & - & \frac{22}{9}x_4 \\
x_2 & = & 1 & - & \frac{1}{9}x_3 & - & \frac{2}{9}x_4 \\
\hline
z & = & -37 & - & 2x_3 & - & 3x_4
\end{array}$$

■

This procedure can be described in general, using matrix algebra. (Remember that performing any sequence of elementary row operations on a matrix amounts to premultiplying the matrix by a suitable non-singular matrix.)

By grouping variables according to whether or not they are basic, we may rewrite the system $\mathbf{Ax} = \mathbf{b}$ as follows:

$$\mathbf{A}_B\mathbf{x}_B + \mathbf{A}_N\mathbf{x}_N = \mathbf{b}$$

Now since the basis matrix $\mathbf{A}_B$ is invertible, $\mathbf{A}_B^{-1}$ exists. Multiplying both sides of the above equation by $\mathbf{A}_B^{-1}$, we obtain:

$$\begin{aligned}
\mathbf{A}_B^{-1}\mathbf{A}_B\mathbf{x}_B + \mathbf{A}_B^{-1}\mathbf{A}_N\mathbf{x}_N &= \mathbf{A}_B^{-1}\mathbf{b} \quad \text{or,} \\
\mathbf{x}_B + \mathbf{A}_B^{-1}\mathbf{A}_N\mathbf{x}_N &= \mathbf{A}_B^{-1}\mathbf{b}
\end{aligned}$$

Note that in the above system, which is equivalent to the given system $\mathbf{Ax} = \mathbf{b}$, each basic variable appears precisely in one equation, and where it does, it appears with coefficient 1. Thus, we may solve for each of the m basic variables in terms of the $n-m$ non-basic variables. In matrix notation, the relationship between the basic variables ($\mathbf{x}_B$) and non-basic variables ($\mathbf{x}_N$) is obtained by rearranging the terms in the above matrix equation:

$$\mathbf{x}_B = \mathbf{A}_B^{-1}\mathbf{b} - \mathbf{A}_B^{-1}\mathbf{A}_N\mathbf{x}_N$$

This equation describes, in matrix notation, the set of solutions of the system $\mathbf{Ax} = \mathbf{b}$ in terms of $\mathbf{x}_N$.

The above equation represents, in matrix notation, the first m rows of the dictionary corresponding to the basis B. To obtain the last row, we must express the objective function $z = \mathbf{cx}$ in terms of the non-basic variables. This is simply a matter of substitution:

$$\begin{aligned}
z &= \mathbf{cx} \\
&= \mathbf{c}_B\mathbf{x}_B + \mathbf{c}_N\mathbf{x}_N \\
&= \mathbf{c}_B(\mathbf{A}_B^{-1}\mathbf{b} - \mathbf{A}_B^{-1}\mathbf{A}_N\mathbf{x}_N) + \mathbf{c}_N\mathbf{x}_N \\
&= \mathbf{c}_B\mathbf{A}_B^{-1}\mathbf{b} + (\mathbf{c}_N - \mathbf{c}_B\mathbf{A}_B^{-1}\mathbf{A}_N)\mathbf{x}_N \qquad (5)
\end{aligned}$$

By writing equation (2) underneath equation (1), we obtain the dictionary corresponding to the basis B:

$$\left.\begin{array}{rcrcl}
\mathbf{x}_B & = & \mathbf{A}_B^{-1}\mathbf{b} & - & \mathbf{A}_B^{-1}\mathbf{A}_N\mathbf{x}_N \\
\hline
z & = & \mathbf{c}_B\mathbf{A}_B^{-1}\mathbf{b} & + & (\mathbf{c}_N - \mathbf{c}_B\mathbf{A}_B^{-1}\mathbf{A}_N)\mathbf{x}_N
\end{array}\right\}$$

Thus, for any given basis B, the dictionary corresponding to the basis B can be computed directly from the original data using the above procedure[8].

This material is covered in the seventh chapter of the textbook. We shall return to the matrix formulation of dictionaries later on.

As a result of the above discussion, we have:

13 _Theorem:_ *There is a one-to-one correspondence between ordered bases and dictionaries.*

However, different orderings of the same basis give rise to the same basic solution[9].

8 Basic Solutions

Consider an LP of the form

$$\max z = \mathbf{c}\mathbf{x}, \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}$$

with[10] rank($\mathbf{A}$)=m = number of rows of $\mathbf{A}$.

Let B be any basis of the system $\mathbf{A}\mathbf{x} = \mathbf{b}$. We have seen that a solution to the system can be obtained by arbitrarily assigning values to the non-basic variables $\{x_j : j \notin B\}$. In particular we may assign the value zero to each of the non-basic variables. The values of basic variables are then uniquely determined. The solution thus obtained is called the **basic solution** corresponding to the basis B. (Note that the ordering of elements of B is irrelevent for this definition.)

If we know the dictionary associated with B (with any ordering of B), we may read the basic solution corresponding to B from the dictionary. For example, if the dictionary is D is

$$\begin{array}{rcl} x_i & = & \overline{b}_i - \sum_{j \notin B} \overline{a}_{ij} x_j \quad \text{for} \quad i \in B, \\ \hline z & = & \overline{v} + \sum_{j \notin B} \overline{c}_j x_j. \end{array}$$

Then, $x_i = \overline{b}_i$, for $i \in B$, $x_j = 0$, for $j \notin B$ is the basic solution corresponding to B.

A non-negative basic solution is called a **basic feasible solution**. (This is often abbreviated as **bfs**.) We shall call a basis **feasible** if the corresponding basic solution is feasible,

---

[8] There is a slick way of describing the above computations as follows: We can obtain the augmented coefficient matrix of the tableau corresponding to B, by premultiplying (4) by a non-singular matrix as below:

$$\left(\begin{array}{c|c} \mathbf{A}_B^{-1} & \mathbf{0} \\ \hline \mathbf{c}_B\mathbf{A}_B^{-1} & 1 \end{array}\right) \left(\begin{array}{c|c|c|c} \mathbf{0} & \mathbf{A}_B & \mathbf{A}_N & \mathbf{b} \\ \hline 1 & -\mathbf{c}_B & -\mathbf{c}_N & 0 \end{array}\right) = \left(\begin{array}{c|c|c|c} \mathbf{0} & \mathbf{I} & \mathbf{A}_B^{-1}\mathbf{A}_N & \mathbf{A}_B^{-1}\mathbf{b} \\ \hline 1 & \mathbf{0} & -\mathbf{c}_N + \mathbf{c}_B\mathbf{A}_B^{-1}\mathbf{A}_N & \mathbf{c}_B\mathbf{A}_B^{-1}\mathbf{b} \end{array}\right)$$

[9] The correspondence between bases and basic solutions is not one-to-one. Corresponding to each basis, there is a unique vector that is its basic solution, but the same vector may be the basic solution corresponding to different bases. An example is given in the next section

[10] This assumption is mainly for the sake of simplifying the presentation. With minor adjustments, we can make all the definitions in this section work for any equality LP.

and we shall refer to the basis corresponding to an optimal basic feasible solution as an **optimal basis**[11].

Example: Consider the following system of equations

$$\begin{array}{rcrcl} x_1 & & - & 3x_3 & = & -3 \\ & x_2 & + & 4x_3 & = & 5 \end{array}$$

The following list shows various types of solutions:

- $(1, -\frac{1}{3}, \frac{4}{3})$ is a solution but is not feasible,
- $(\frac{1}{4}, \frac{2}{3}, \frac{13}{12})$ is a feasible solution but not a basic solution,
- $(-3, 5, 0)$ is a basic solution but is not a bfs,
- $(0, 1, 1)$ is a bfs,
- $(\frac{3}{4}, 0, \frac{5}{4})$ is a bfs.

The following theorem is a characterization of basic feasible solutions.

14 *Theorem: A feasible solution is a basic feasible solution if and only if the columns of* $\mathbf{A}$ *corresponding to variables with positive value in the solution constitute a linearly independent set.*

Proof. Suppose that $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ is a basic feasible solution corresponding to a basis B. Then the only s_i that can be positive have their subscripts in B. By definition, any subset of columns whose subscripts are in B is linearly independent. This proves the "only if" part.

Conversely, suppose that $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ is a feasible solution which satisfies the given condition. Suppose that I is the index set such that $s_i > 0$ for all $i \in I$ and $s_j = 0$ for all $j \notin I$. By the hypothesis $\{A_i : i \in I\}$ is linearly independent. This linearly independent set is contained in subset of columns of $\mathbf{A}$ which constitutes a basis for the column space of $\mathbf{A}$. Let $\{A_i : i \in B\}$ be such a basis. Since $B \supseteq I$, $s_j = 0$ for all $j \notin B$. It follows that $\mathbf{s}$ is the basic solution corresponding to B. ■

15 *Example: Consider the following system of linear equations*

$$\begin{array}{rrrrrrrrrrrl} x_1 & & & + & x_4 & + & 2x_5 & - & 2x_6 & + & x_7 & = 1 \\ & x_2 & & + & x_4 & + & 2x_5 & - & 2x_6 & + & x_7 & = 1 \\ & & x_3 & & & + & x_5 & - & 2x_6 & + & x_7 & = 0 \end{array}$$

[11] There is a practical meaning of an optimal basis. In a manufacturing problem, we typically have the option to make a number (n) of products. The fact the the optimal solution obtained by the simplex method is a basic solution tells us that, in order to maximize profit, it suffices to produce a selection of m of the products (those corresponding to the basic variables).

(a) $(0,0,0,0,1,3,5)$ is not a bfs to this system because $\mathbf{A}_5, \mathbf{A}_6$ and $\mathbf{A}_7$ (the fifth, sixth and seventh columns of the coefficient matrix) are not linearly independent.

(b) $(1,1,0,0,0,0,0)$ is a bfs because the columns A_1 and A_2 are linearly independent. $\{A_1, A_2\}$ can be extended to a basis. $\{A_1, A_2, A_3\}$ is one of them. In fact there are four bases which contain $\{A_1, A_2\}$.

Two remarks are in order here:

(1) The same vector may arise as a bfs corresponding to different bases.

(2) A basic variable may assume the value zero in a bfs (if a basic variable takes the value zero in a bfs then the solution is a **degenerate** bfs).

Both these are exemplified by (b) above.

The following theorem brings out the importance of basic feasible solutions in the theory of linear programming.

16 <u>*Theorem:*</u> *Consider the linear program*

$$\begin{array}{rl} \textit{Maximize} & z = \mathbf{cx} \\ \textit{subject to} & \mathbf{Ax} = \mathbf{b} \\ \textit{and} & \mathbf{x} \geq 0. \end{array}$$

(1) If the LP has a feasible solution, then it has a basic feasible solution.
(2) If the LP has an optimal solution then it has an optimal basic feasible solution.

These theorems provide a motivation for the simplex method. It is possible to give simple direct proofs of them, but they do follow from the fact that the simplex method works.

Exercises.

1. Show that if $\mathbf{A}$ is an $m \times n$ matrix with rank($\mathbf{A}$)=m, then $\mathbf{Ax} = \mathbf{b}$ has at most $\binom{n}{m}$ basic feasible solutions.

2. Consider the following system

$$\begin{array}{rcrcrcrcrcrcl} x_1 & & & & & + & x_4 & + & 2x_5 & - & x_6 & = & 1 \\ & + & x_2 & & & + & 2x_4 & + & 4x_5 & + & 2x_6 & = & 2 \\ 2x_1 & & & + & x_3 & + & 3x_4 & + & 6x_5 & + & x_6 & = & 3. \end{array}$$

In each of the following cases verify whether or not the given vector is a basic feasible solution to the system

(i) $(1,2,1,0,0,0)$ $\qquad$ (ii) $(0,0,0,\frac{1}{3},\frac{1}{3},0)$

3. Show that if $\mathbf{0}$ is a feasible solution to $\mathbf{Ax} = \mathbf{b}$ then it is a basic feasible solution of $\mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}$.

4. Show that a non-negative solution $\mathbf{s}$ to a system $\mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}$ is a bfs if and only if there is a linear function $z = \mathbf{cx}$ such that $\mathbf{s}$ is the unique optimum solution to the LP

$$\begin{aligned} \text{Maximize} \quad & z = \mathbf{cx} \\ \text{subject to} \quad & \mathbf{Ax} = \mathbf{b} \\ \text{and} \quad & \mathbf{x} \geq \mathbf{0}. \end{aligned}$$

9 Geometrical Interpretation

The feasible region of an LP in n variables is a convex polyhedron in $\mathbf{R}^n$, and the basic feasible solutions to the LP are the extreme points (corners) of this polyhedron. In the simplex method, we start at one extreme point. In each iteration, we move along an edge of the polyhedron from the the current extreme point to an extreme point adjacent to it.

A subset K of R^n is called a **convex set** if for any two points $\mathbf{x}$ and $\mathbf{y}$ in K, the whole line segment joining $\mathbf{x}$ and $\mathbf{y}$ is entirely in K; that is if $\mathbf{x}, \mathbf{y} \in K \Rightarrow \lambda\mathbf{x} + (1 - \lambda)\mathbf{y} \in K$ for all λ such that $0 \leq \lambda \leq 1$.

17 *Examples:*

(1) The set of all solutions to one linear equation, $\sum_{j=1}^{n} a_j x_j = b$, with not all $a_j = 0$, is called a **hyperplane** in R^n. (In two-dimensions, hyperplanes are lines, in three-dimensions, hyperplanes are planes.) If $\sum_{j=1}^{n} a_j x_j = b$ is a hyperplane, then the two sets

$$\{(x_1, x_2, \ldots, x_n) | \sum_{j=1}^{n} a_j x_j \geq b\} \text{ and}$$

$$\{(x_1, x_2, \ldots, x_n) | \sum_{j=1}^{n} a_j x_j \leq b\},$$

are called the *half-spaces* generated by the hyperplane.

Any hyperplane or a half-space is a convex set.

(2) The set of points inside a sphere is a convex set.

18 *Theorem: The intersection of a finite number of convex sets is convex.*

19 *Definition: The intersection of a finite number of half-spaces is called a* **convex polyhedron**.

In an LP, each constraint either defines a hyperplane or a half-space[12]. Thus, the set of points satisfying any one constraint is a convex set. The feasible region of the LP is the set of all points which satisfy all the constraints. Thus, it is the intersection of the convex sets defined by the individual constraints. Hence we have:

20 *Theorem: The feasible region of an LP is a convex polyhedron.*

A point $\mathbf{x}$ in a convex set K is called an **extreme point** of K if

$$\mathbf{x} = \frac{1}{2}\mathbf{x}_1 + \frac{1}{2}\mathbf{x}_2 \text{ and } \mathbf{x}_1, \mathbf{x}_2 \in K \Longrightarrow \mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}.$$

In the case of a convex polygon, the extreme points are precisely its corner points. Each point on the surface of a closed ball is its extreme points.

21 *Theorem: A feasible solution to* $\mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}$ *is an extreme point of the feasible region if, and only if, it is a basic feasible solution.*

Sketch of Proof: Let $\mathbf{x}$ be a basic feasible solution of $\mathbf{Ax} = \mathbf{b}$. Assume without loss of generality that $\mathbf{x}$ corresponds to the ordered basis $B = (1, 2, \ldots, m)$. If $\mathbf{x}_1$ and $\mathbf{x}_2$ are feasible solutions such that $\mathbf{x} = \frac{1}{2}\mathbf{x}_1 + \frac{1}{2}\mathbf{x}_2$ then $\mathbf{x}_1$ and $\mathbf{x}_2$ agree with $\mathbf{x}$ in the last $n - m$ co-ordinates. It follows (exercise) that $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}$. Hence $\mathbf{x}$ is an extreme point of the set of feasible solutions to $\mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}$.

Conversely suppose that $\mathbf{s}$ is an extreme point of the set of feasible solutions to $\mathbf{Ax} = \mathbf{b}$. Assume without loss of generality taht the first p co-ordinates of S are positive and the remaining $n - p$ co-ordinates of $\mathbf{s}$ are zero. If $\mathbf{s}$ is not a basic feasible solution, then, by Theorem (14), the first p columns of $\mathbf{A}$ are linearly dependent. Thus, there exist $\lambda_1, \lambda_2, \ldots, \lambda_p$, not all zero, such that

$$\sum_{j=1}^{p} \lambda_j A_j = 0.$$

Let λ denote the n-vector $(\lambda_1, \lambda_2, \ldots, \lambda_p, 0, \ldots, 0)^t$. Consider the two vectors

$$\mathbf{x} = \mathbf{s} + t\lambda$$

and

$$\mathbf{y} = \mathbf{s} - t\lambda.$$

If t is a positive real number such that

$$t \leq \min_j \left\{ \frac{s_j}{|\lambda_j|}, \lambda_j \neq 0 \right\}$$

[12]Note that a hyperplane itself is the intersection of the two half-spaces it generates.

then both $\mathbf{x}$ and $\mathbf{y}$ are feasible solutions (exercise), and

$$\mathbf{s} = \frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y}.$$

This contradicts the assumption that $\mathbf{s}$ is an extreme point of the convex set of feasible solutions. Hence $\mathbf{s}$ is a basic feasible solution. ■

Suppose that we have an LP max $z = \mathbf{cx}$ s.t. $\mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}$, with rank $\mathbf{A} = m =$ the number of rows of $\mathbf{A}$. If B is a feasible basis, the basic feasible solution corresponding to B is the unique solution to the system

$$\begin{aligned} \mathbf{A}_B\mathbf{x}_B &= \mathbf{b}, \\ \mathbf{x}_N &= \mathbf{0} \end{aligned} \tag{6}$$

of n equations in n variables. Suppose that the current bfs is not optimal and suppose that a variable x_s is chosen as the entering variable. To determine the leaving variable, we consider the system of equations

$$\begin{aligned} \mathbf{A}_B\mathbf{x}_B &= \mathbf{b} - x_s\mathbf{A}_s, \\ \mathbf{x}_{N-s} &= \mathbf{0}. \end{aligned} \tag{7}$$

When $x_s = 0$, the solution to this system is the current bfs. Let us denote by L the set of all non-negative solutions to (7). It may be the case that L is just a singleton consisting of the current bfs[13]. For clarity, let us consider the case in which there are points in L at which x_s takes positive values. Then L is an 'edge' of the polyhedron (see exercise 4). Moving along this edge, away from the current bfs, results in an increase in the value of the objective function. If, for all non-negative values of x_s, (7) has feasible solutions, then the LP is unbounded. If not, there is a maximum value of x_s. A basic variable x_r that becomes zero when x_s is given this maximum possible value is chosen as the leaving variable. The new basis is $B' = B - \{r\} \cup \{s\}$ (and $N' = N - \{s\} \cup \{r\}$). We move to the basic feasible solution which is the unique solution to the system

$$\begin{aligned} \mathbf{A}_{B'}\mathbf{x}_{B'} &= \mathbf{b}, \\ \mathbf{x}_{N'} &= \mathbf{0} \end{aligned} \tag{8}$$

In this case L is the edge of the feasible region joining the two extreme points which are the basic feasible solutions corresponding to B and B', respectively[14].

Exercises.

1. Show that the vectors $\mathbf{x}$ and $\mathbf{y}$ defined in the proof of theorem 6 are feasible solutions to $\mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}$ if $0 < t \leq \min_j \left\{\frac{s_j}{|\lambda_j|}, \lambda_j \neq 0\right\}$.

---

[13]This happens in case of degeneracy.

[14]In case of a degenerate pivot, the two systems (6) and (8) have the same solution.

2. Consider the LP:

$$\begin{array}{lrcrcrcr}
\max\ z = & 6x_1 & + & 4x_2 & + & 2x_3 & & \\
\text{s.t.} & 2x_1 & + & x_2 & + & x_3 & \leq & 4 \\
 & 4x_1 & + & 2x_2 & + & 2x_3 & \leq & 12 \\
 & 3x_1 & + & \frac{5}{2}x_2 & + & 2x_3 & \leq & 8 \\
 & x_1 & , & x_2 & , & x_3 & \geq & 0
\end{array}$$

Determine the polyhedron of feasible solutions to this LP. What are the extreme points of this polyhedron?

3. A hyperplane H defined by $\sum c_i x_i = c$ is called a **supporting hyperplane** of a convex set K if (i) $H \cap K \neq \emptyset$ and (ii) K lies in one closed half-space generated by H.

   Consider the LP $\max z = \mathbf{cx}$ s.t. $\mathbf{Ax} = \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$. Let v^* be the optimal value of z. Show that the hyperplane defined by $\mathbf{cx} = v^*$ is a supporting hyperplane to the convex set of feasible solutions to the LP.

4. (a) Show that, in case of non-degeneracy, the set L of feasible solutions to (7) is indeed an edge of the polyhedron of feasible solutions. (Firstly show that L is 1-dimensional. Then show that if $\mathbf{x}$ is on L, and $\mathbf{x}_1$ and $\mathbf{x}_2$ are feasible solutions to the LP such that $\mathbf{x} = \frac{1}{2}\mathbf{x}_1 + \frac{1}{2}\mathbf{x}_2$, then $\mathbf{x}_1$ and $\mathbf{x}_2$ are also on L.

   (b) Describe L parametrically in terms of x_s.

5*. If $\mathbf{x}_1$, $\mathbf{x}_2$,...,$\mathbf{x}_k$, are k vectors in $\mathbf{R}^n$, λ_1, λ_2,...,λ_k are real numbers such that $0 \leq \lambda_i \leq 1$, for $1 \leq i \leq k$, and $\sum_{i=1}^k \lambda_i = 1$, then $\sum_{i=1}^k \lambda_i \mathbf{x}_i$ is called a **convex linear combination** of $\mathbf{x}_1$, $\mathbf{x}_2$,...,$\mathbf{x}_k$. Show that a convex set is closed under the operation of taking convex linear combinations. (For example, show that if K is a convex set, and $\mathbf{x}_1$, $\mathbf{x}_2$,$\mathbf{x}_3$ belong to K, then $\frac{1}{3}\mathbf{x}_1 + \frac{1}{3}\mathbf{x}_2 + \frac{1}{3}\mathbf{x}_3$ also belongs to K.)