C & O 350 **MATRIX FORMULAS**

Consider an LP in standard form:

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

After the introduction of slack variables, the LP becomes:

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

In this form, all constraints of the LP (apart from non-negativity of the variables) are equations, and the LP can be expressed in matrix notation as follows:

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

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

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ . & & & . & & & & \\ . & & & . & & & & \\ . & & & . & & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} & 0 & 0 & \dots & 1 \end{pmatrix} \text{ and } b = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_m \end{pmatrix}$$

$$\begin{array}{lcl} \mathbf{c} & = & (c_1, \quad c_2, \quad \dots \quad c_n, \quad 0 \quad \dots \quad 0) \\ \mathbf{x}^t & = & (x_1, \quad x_2, \quad \dots \quad x_n, \quad x_{n+1} \quad \dots \quad x_{n+m}) \end{array}$$

(Note that $\mathbf{x}$ is a column vector; the vector $\mathbf{x}^t$ displayed above is its transpose.)

There are matrix formulas by means of which we can express any dictionary in terms of the given data ($\mathbf{A}$, $\mathbf{b}$, and $\mathbf{c}$). The following notation and definitions are important for describing those formulas.

NOTATION: Let $\mathbf{A}_j$ denote the jth column of the matrix $\mathbf{A}$. For any subset J of the column index set $\{1, 2, ..., n+m\}$, we shall let $\mathbf{A}_J$ denote the submatrix of consisting of the columns with their indices in J.

Definition: An m-element subset B of the column index set $\{1, 2, ..., n+m\}$ 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.

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

By grouping variables according to whether or not they are basic, we 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 $\mathbf{A}_B$ is invertible, $\mathbf{A}_B^{-1}$ exists. Multiplying both sides of the above equation by $\mathbf{A}_B^{-1}$, we obtain:

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

Now by rearranging terms in the above equation, we can express the basic variables ($\mathbf{x}_B$) in terms of the non-basic variables ($\mathbf{x}_N$):

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

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:

$$\begin{aligned} z &= \mathbf{cx} \\ &= \mathbf{c}_B\mathbf{x}_B + \mathbf{c}_N\mathbf{x}_N \\ &= \mathbf{c}_B(\mathbf{A}_B^{-1}\mathbf{b} - \mathbf{A}_B^{-1}\mathbf{A}_N\mathbf{x}_N) + \mathbf{c}_N\mathbf{x}_N \\ &= \mathbf{c}_B\mathbf{A}_B^{-1}\mathbf{b} + (\mathbf{c}_N - \mathbf{c}_B\mathbf{A}_B^{-1}\mathbf{A}_N)\mathbf{x}_N \end{aligned} \tag{2}$$

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

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

Thus, for any given basis B, the dictionary corresponding to the basis B can be computed directly from the original data using the above formulas. Before giving examples, let us restate the formulas in simpler notation. First of all, if B is a basis, we shall refer to the matrix $\mathbf{A}_B^{-1}$ as the *basis matrix*[1] corresponding to the basis B, and denote it simply by $\mathbf{B}$ (in bold face). Also, since the product $\mathbf{c}_B\mathbf{B}^{-1} = \mathbf{c}_B\mathbf{A}_B^{-1}$ appears repeatedly in the formulas let us simply write $\mathbf{y}$ for the expression $\mathbf{c}_B\mathbf{B}^{-1}$. (The reason for the choice of the letter $\mathbf{y}$ is because of the relavance to the dual. See exercise 4.) With this notation, the above information can be summarized as below: Firstly, the dictionary corresponding to the basis B is as follows:

[1]Note that the basis matrix depends on the basis under consideration, but it is a submatrix of $\mathbf{A}$.

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

(where, to repeat, $\mathbf{B}$ is the basis matrix $\mathbf{A}_B^{-1}$, and $\mathbf{y} = \mathbf{c}_B\mathbf{B}^{-1}$.)

We previously denoted the dictionary corresponding to a basis B as follows:

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

where $\overline{b}_i$ is the value of the basic variable x_i , $\overline{z}$ is the value of the objective function at the basic solution associated with the basis (obtained by setting the values of the non-basic variables equal to zero), $\overline{c}_j$ is the coefficient of x_j (referred to as the 'reduced cost' coefficient of x_j). (In general, the entries of the dictionary corresponding to a basis B are indicated by an overline; those without the overline indicate the original data.) In view of the above derivation of the dictionary corresponding to B, we have the following relationships:

$$\begin{array}{rcl} \overline{b} & = & \mathbf{B}^{-1}\mathbf{b} \\ \overline{\mathbf{A}} & = & \mathbf{B}^{-1}\mathbf{A} \quad (\overline{\mathbf{A}}_j = \mathbf{B}^{-1}\mathbf{A}_j) \\ \overline{z} & = & \mathbf{y}\mathbf{b} \quad \text{and,} \\ \overline{\mathbf{c}} & = & \mathbf{c} - \mathbf{y}\mathbf{A} \quad (\overline{c}_j = c_j - \mathbf{y}\mathbf{A}_j) \end{array}$$

Dictionaries are convenient for understanding the way the simplex method works. But to visualize the above relationships, tableaux are more convenient. Note that a tableau simply consists of the constraint equations written without rearranging terms, followed by the equation representing the objective function: (The tableau to the left is the initial tableau, and the tableau to the right is a subsequent tableau. Note that the number in the lower right-hand corner represents the negative of the value of the objective function at the corresponding basic solution.)

$$\left.\begin{array}{rrcl} & \sum a_{ij}x_j & = & b_i \\ \hline -z+ & \sum c_j x_j & = & 0 \end{array}\right\} \quad \Rightarrow \quad \begin{array}{rrcl} & \sum \overline{a}_{ij}x_j & = & \overline{b}_i \\ \hline -z+ & \sum \overline{c}_j x_j & = & -\overline{z} \end{array}$$

The relationships stated above can be displayed in tableau format (just using the coefficients) as follows:

Let us emphasise that there is a one-to-one correpondence between bases and dictionaries. We shall call a basis *feasible*, if the the correponding basic solution is feasible (i.e. nonnegative), and we shall refer to the basis corresponding to an optimal basic feasible solution as an *optimal basis*[2].

Whenever a basis is specified, we can derive the corresponding dictionary using the above formulas[3]. There are situations when we might wish to compute the dictionary corresponding to a specified basis directly from the given data. For example, we might, for verification purposes, wish to reconstruct the final dictionary (corresponding to the optimal solution) without going through all the intermediate dictionaries. Another situation involves sensitivity analysis: after an LP has been solved, we are told that some numbers in the data need to be modified slightly. In order to decide if the basis which is optimal with the old data remains optimal after changes, it is not necessary to solve the revised problem from scratch: we can use the matrix formulas.

EXERCISES:

1. Consider the following LP (obtained from an LP in standard form by adding the slack variables x_4, and x_5):

$$\begin{array}{lcccccccccccc} \max & x_1 & + & 4x_2 & + & 3x_3 & & & & & & & \\ \text{s.t.} & x_1 & + & x_2 & + & 2x_3 & + & x_4 & & & = & 5 \\ & 2x_1 & + & x_2 & + & 4x_3 & & & + & x_5 & = & 7 \\ & x_1 & , & x_2 & , & x_3 & , & x_4 & , & x_5 & \geq & 0 \end{array}$$

   (a) For the basis $\{1, 2\}$, using the matrix formulas, compute the following:

   (i) The values of the basic variables in the corresponding basic solution (i.e $\bar{b}$),

   (ii) $\mathbf{y}$,

   (iii) the value of the objective function at the corresponding basic solution, and

   (iv) the corresponding reduced costs.

   Using the above information, decide whether or not $\{1, 2\}$ is an optimal basis.

   (b) Do the same for the basis $\{2, 3\}$.

2. For the LP on page 98 (also 102) of the text, verify, using matrix formulas that $\{1, 3, 4\}$ is an optimal basis. Then show that the same basis remains optimal if the number of hours of labour available is increased by ten.

3. Given a basis B, we can compute $\mathbf{B}^{-1}$, the inverse of the basis matrix, and $\mathbf{y}$, and then, using the matrix formulas we can construct the dictionary corresponding to the basis B. Conversely, given the dictionary corresponding to the basis B, we can find $\mathbf{B}^{-1}$ and $\mathbf{y}$. ($\mathbf{B}^{-1}$ appears 'under' the slack variables, and y_i is the negative of the reduced cost of the ith slack variable, (more precise explanation will be given in class).)

   Consider the final dictionary for example 1 (page 17 of the text). Recover $\mathbf{B}^{-1}$ and $\mathbf{y}$.

[2] 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).

[3] Important Remark: All the above defintions and formulas are valid for any LP of the form $\max \mathbf{cx}$: s.t. $\mathbf{Ax} = \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$, provided that the rank of the matrix $\mathbf{A}$ is m. It is not necessary for the LP to have been obtained from an LP in standard form by adding slack variables.

4. (a) Write, in matrix notation, the dual of an LP in standard form. (b) Let $\mathbf{y}^*$ denote the $\mathbf{y}$-vector corresponding to an optimal basis. Show that $\mathbf{y}^*$ is an optimal solution to the dual. (This would be proof of the duality theorem, stated in matrix notation.)