Chapter 4

DUALITY AND COMPLEMENTARITY

Associated with every linear program, and intimately related to it, is a corresponding dual linear program. Both programs are constructed from the same underlying cost and constraint coefficients but in such a way that if one of these problems is one of minimization the other is one of maximization, and the optimal values of the corresponding objective functions, if finite, are equal. The variables of the dual problem can be interpreted as prices associated with the constraints of the original (primal) problem, and through this association it is possible to give an economically meaningful characterization to the dual whenever there is such a characterization for the primal.

The variables of the dual problem are also intimately related to the calculation of the relative cost coefficients in the simplex method. Thus, a study of duality sharpens our understanding of the simplex procedure and motivates certain alternative solution methods. Indeed, the simultaneous consideration of a problem from both the primal and dual viewpoints often provides significant computational advantage as well as economic insight.

4.1 DUAL LINEAR PROGRAMS

In this section we define the dual program that is associated with a given linear program. Initially, we depart from our usual strategy of considering programs in standard form, since the duality relationship is most symmetric for programs expressed solely in terms of inequalities. Specifically then, we define duality through the pair of programs displayed below.

$$
\begin{array}{llcll}
 & \text{Primal} & & & \text{Dual} \\
\text{minimize} & \mathbf{c}^T\mathbf{x} & & \text{maximize} & \mathbf{y}^T\mathbf{b} \\
\text{subject to} & \mathbf{A}\mathbf{x} \geqslant \mathbf{b} & & \text{subject to} & \mathbf{y}^T\mathbf{A} \leqslant \mathbf{c}^T \\
 & \mathbf{x} \geqslant \mathbf{0} & & & \mathbf{y} \geqslant \mathbf{0}
\end{array}
\qquad (1)
$$

If $\mathbf{A}$ is an $m \times n$ matrix, then $\mathbf{x}$ is an m-dimensional column vector, $\mathbf{b}$ is an n-dimensional column vector, $\mathbf{c}^T$ is an n-dimensional row vector, and $\mathbf{y}^T$ is an m-dimensional row vector. The vector $\mathbf{x}$ is the variable of the primal program, and $\mathbf{y}$ is the variable of the dual program.

The pair of programs (1) is called the *symmetric form* of duality and, as explained below, can be used to define the dual of any linear program. It is important to note that the role of primal and dual can be reversed. Thus, studying in detail the process by which the dual is obtained from the primal: interchange of cost and constraint vectors, transposition of coefficient matrix, reversal of constraint inequalities, and change of minimization to maximization; we see that this same process applied to the dual yields the primal. Put another way, if the dual is transformed, by multiplying the objective and the constraints by minus unity, so that it has the structure of the primal (but is still expressed in terms of $\mathbf{y}$), its corresponding dual will be equivalent to the original primal.

The dual of any linear program can be found by converting the program to the form of the primal shown above. For example, given a linear program in standard form

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geqslant \mathbf{0}, \end{array}$$

we write it in the equivalent form

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} \geqslant \mathbf{b} \\ & -\mathbf{A}\mathbf{x} \geqslant -\mathbf{b} \\ & \mathbf{x} \geqslant \mathbf{0}, \end{array}$$

which is in the form of the primal of (1) but with coefficient matrix $\begin{bmatrix} \mathbf{A} \\ -\mathbf{A} \end{bmatrix}$. Using a dual vector partitioned as $(\mathbf{u}, \mathbf{v})$, the corresponding dual is

$$\begin{array}{ll} \text{maximize} & \mathbf{u}^T\mathbf{b} - \mathbf{v}^T\mathbf{b} \\ \text{subject to} & \mathbf{u}^T\mathbf{A} - \mathbf{v}^T\mathbf{A} \leqslant \mathbf{c}^T \\ & \mathbf{u} \geqslant \mathbf{0},\ \mathbf{v} \geqslant \mathbf{0}. \end{array}$$

Letting $\mathbf{y} = \mathbf{u} - \mathbf{v}$ we may simplify the representation of the dual program so that we obtain the pair of problems displayed below:

$$\begin{array}{llll} \multicolumn{2}{c}{\text{Primal}} & \multicolumn{2}{c}{\text{Dual}} \\ \text{minimize} & \mathbf{c}^T\mathbf{x} & \text{maximize} & \mathbf{y}^T\mathbf{b} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geqslant \mathbf{0} & \text{subject to} & \mathbf{y}^T\mathbf{A} \leqslant \mathbf{c}^T. \end{array} \tag{2}$$

This is the *asymmetric form* of the duality relation. In this form the dual vector $\mathbf{y}$ (which is really a composite of $\mathbf{u}$ and $\mathbf{v}$) is not restricted to be nonnegative.

Similar transformations can be worked out for any linear program to first get the primal in the form (1), calculate the dual, and then simplify the dual to account for special structure.

In general, if some of the linear inequalities in the primal (1) are changed to equality, the corresponding components of $\mathbf{y}$ in the dual become free variables. If some of the components of $\mathbf{x}$ in the primal are free variables, then the corresponding inequalities in $\mathbf{y}^T\mathbf{A} \leqslant \mathbf{c}^T$ are changed to equality in the dual. We mention again that these are not arbitrary rules but are direct consequences of the original definition and the equivalence of various forms of linear programs.

Example 1. (Dual of the diet problem). The diet problem, Example 1, Section 2.2, was the problem faced by a dietitian trying to select a combination of foods to meet certain nutritional requirements at minimum cost. This problem has the form

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} \geqslant \mathbf{b},\ \mathbf{x} \geqslant \mathbf{0} \end{array}$$

and hence can be regarded as the primal program of the symmetric pair above. We describe an interpretation of the dual problem.

Imagine a pharmaceutical company that produces in pill form each of the nutrients considered important by the dietitian. The pharmaceutical company tries to convince the dietitian to buy pills, and thereby supply the nutrients directly rather than through purchase of various foods. The problem faced by the drug company is that of determining positive unit prices $\lambda_1, \lambda_2, \ldots, \lambda_m$ for the nutrients so as to maximize revenue while at the same time being competitive with real food. To be competitive with real food, the cost of a unit of food i made synthetically from pure nutrients bought from the druggist must be no greater than c_i, the market price of the food. Thus, denoting by $\mathbf{a}_i$ the ith food, the company must satisfy $\mathbf{y}^T\mathbf{a}_i \leqslant \mathbf{c}_i$ for each i. In matrix form this is equivalent to $\mathbf{y}^T\mathrm{A} \leqslant \mathbf{c}^T$. Since b_j units of the jth nutrient will be purchased, the problem of the druggist is

$$\begin{array}{ll} \text{maximize} & \mathbf{y}^T\mathbf{b} \\ \text{subject to} & \mathbf{y}^T\mathbf{A} \leqslant \mathbf{c}^T,\ \mathbf{y} \geqslant \mathbf{0}, \end{array}$$

which is the dual problem.

Example 2. (Dual of the transportation problem). The transportation problem, Example 3, Section 2.2, is the problem, faced by a manufacturer, of selecting the pattern of product shipments between several fixed origins and destinations so as to minimize transportation cost while satisfying demand. Referring to (6) and (7) of Chapter 2, the problem is in standard form, and hence the asymmetric version of the duality relation applies. There is a dual variable for each constraint. In this case

we denote the variables $u_i, i = 1, 2, \ldots, m$ for (6) and $v_j,\ j = 1, 2, \ldots, n$ for (7). Accordingly, the dual is

$$\begin{array}{lll} \text{maximize} & \sum_{i=1}^{m} a_i u_i + \sum_{j=1}^{n} b_j v_j & \\ \text{subject to} & u_i + v_j \leqslant c_{ij}, & i = 1, 2, \ldots, m, \\ & & j = 1, 2, \ldots, n. \end{array}$$

To interpret the dual problem, we imagine an entrepreneur who, feeling that he can ship more efficiently, comes to the manufacturer with the offer to buy his product at the plant sites (origins) and sell it at the warehouses (destinations). The product price that is to be used in these transactions varies from point to point, and is determined by the entrepreneur in advance. He must choose these prices, of course, so that his offer will be attractive to the manufacturer.

The entrepreneur, then, must select prices $-u_1, -u_2, \ldots, -u_m$ for the m origins and $v_1, v_2, \ldots, v_n$ for the n destinations. To be competitive with usual transportation modes, his prices must satisfy $u_i + v_j \leqslant c_{ij}$ for all i, j, since $u_i + v_j$ represents the net amount the manufacturer must pay to sell a unit of product at origin i and buy it back again at destination j. Subject to this constraint, the entrepreneur will adjust his prices to maximize his revenue. Thus, his problem is as given above.

4.2 THE DUALITY THEOREM

To this point the relation between the primal and dual programs has been simply a formal one based on what might appear as an arbitrary definition. In this section, however, the deeper connection between a program and its dual, as expressed by the Duality Theorem, is derived.

The proof of the Duality Theorem given in this section relies on the Separating Hyperplane Theorem (Appendix B) and is therefore somewhat more advanced than previous arguments. It is given here so that the most general form of the Duality Theorem is established directly. An alternative approach is to use the theory of the simplex method to derive the duality result. A simplified version of this alternative approach is given in the next section.

Throughout this section we consider the primal program in standard form

$$\begin{aligned} &\text{minimize} \quad \mathbf{c}^T\mathbf{x} \\ &\text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geqslant \mathbf{0} \end{aligned} \tag{3}$$

and its corresponding dual

$$\begin{aligned} &\text{maximize} \quad \mathbf{y}^T\mathbf{b} \\ &\text{subject to} \quad \mathbf{y}^T\mathbf{A} \leqslant \mathbf{c}^T. \end{aligned} \tag{4}$$

In this section it is *not* assumed that $\mathbf{A}$ is necessarily of full rank. The following lemma is easily established and gives us an important relation between the two problems.

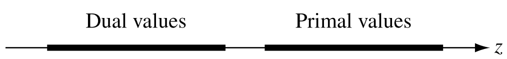

Fig. 4.1 Relation of primal and dual values

Lemma 1. (Weak Duality Lemma). *If* $\mathbf{x}$ *and* $\mathbf{y}$ *are feasible for (3) and (4),respectively, then* $\mathbf{c}^T\boldsymbol{x} \geqslant \mathbf{y}^T\boldsymbol{b}$.

Proof. We have

$$\mathbf{y}^T\mathbf{b} = \mathbf{y}^T\mathbf{A}\mathbf{x} \leqslant \mathbf{c}^T\mathbf{x},$$

the last inequality being valid since $\mathbf{x} \geqslant \mathbf{0}$ and $\mathbf{y}^T\mathbf{A} \leqslant \mathbf{c}^T$. ∎

This lemma shows that a feasible vector to either problem yields a bound on the value of the other problem. The values associated with the primal are all larger than the values associated with the dual as illustrated in Fig. 4.1. Since the primal seeks a minimum and the dual seeks a maximum, each seeks to reach the other. From this we have an important corollary.

Corollary. *If* $\mathbf{x}_0$ *and* $\mathbf{y}_0$ *are feasible for (3) and (4), respectively, and if* $\mathbf{c}^T\mathbf{x}_0 = \mathbf{y}_0^T\mathbf{b}$, *then* $\mathbf{x}_0$ *and* $\mathbf{y}_0$ *are optimal for their respective problems.*

The above corollary shows that if a pair of feasible vectors can be found to the primal and dual programs with equal objective values, then these are both optimal. The Duality Theorem of linear programming states that the converse is also true, and that, in fact, the two regions in Fig. 4.1 actually have a common point; there is no "gap."

Duality Theorem of Linear Programming. *If either of the problems* (3) *or* (4) *has a finite optimal solution, so does the other, and the corresponding values of the objective functions are equal. If either problem has an unbounded objective, the other problem has no feasible solution*.

Proof. We note first that the second statement is an immediate consequence of Lemma 1. For if the primal is unbounded and $\mathbf{y}$ is feasible for the dual, we must have $\mathbf{y}^T\mathbf{b} \leqslant -M$ for arbitrarily large M, which is clearly impossible.

Second we note that although the primal and dual are not stated in symmetric form it is sufficient, in proving the first statement, to assume that the primal has a finite optimal solution and then show that the dual has a solution with the same value. This follows because either problem can be converted to standard form and because the roles of primal and dual are reversible.

Suppose (3) has a finite optimal solution with value z_0. In the space E^{m+1} define the convex set

$$C = \{(r,\ \mathbf{w}) : r = tz_0 - \mathbf{c}^T\mathbf{x},\ \mathbf{w} = t\mathbf{b} -\ \mathbf{A}\mathbf{x},\ \mathbf{x} \geqslant \mathbf{0},\ t \geqslant 0\}.$$

It is easily verified that C is in fact a closed convex cone. We show that the point $(1,\ \mathbf{0})$ is not in C. If $\mathbf{w} = t_0\mathbf{b} - \mathbf{A}\mathbf{x}_0 = \mathbf{0}$ with $t_0 > 0,\ \mathbf{x}_0 \geqslant \mathbf{0}$, then $\mathbf{x} = \mathbf{x}_0/t_0$ is feasible for (3) and hence $r/t_0 = z_0 - \mathbf{c}^T\mathbf{x} \leqslant \mathbf{0}$; which means $r \leqslant \mathbf{0}$. If $\mathbf{w} = -\mathbf{A}\mathbf{x}_0 = \mathbf{0}$ with $\mathbf{x}_0 \geqslant \mathbf{0}$ and $\mathbf{c}^T\mathbf{x}_0 = -1$, and if $\mathbf{x}$ is any feasible solution to (3), then $\mathbf{x} + \alpha\mathbf{x}_0$ is feasible for any $\alpha \geqslant 0$ and gives arbitrarily small objective values as α is increased. This contradicts our assumption on the existence of a finite optimum and thus we conclude that no such $\mathbf{x}_0$ exists. Hence $(1, \mathbf{0}) \notin C$.

Now since C is a closed convex set, there is by Theorem 4.4, Section B.3, a hyperplane separating $(1, \mathbf{0})$ and C. Thus there is a nonzero vector $[s, \mathbf{y}] \in E^{m+1}$ and a constant c such that

$$s < c = \inf\{sr + \mathbf{y}^T \mathbf{w} : (r,\ \mathbf{w}) \in C\}.$$

Now since C is a cone, it follows that $c \geqslant 0$. For if there were $(r,\ \mathbf{w}) \in C$ such that $sr + \mathbf{y}^T \mathbf{w} < 0$, then $\alpha(r,\ \mathbf{w})$ for large α would violate the hyperplane inequality. On the other hand, since $(0,\ \mathbf{0}) \in C$ we must have $c \leqslant 0$. Thus $c = 0$. As a consequence $s < 0$, and without loss of generality we may assume $s = -1$.

We have to this point established the existence of $\mathbf{y} \in E^m$ such that

$$-r + \mathbf{y}^T \mathbf{w} \geqslant 0$$

for all $(r,\ \mathbf{w}) \in C$. Equivalently, using the definition of C,

$$(\mathrm{c} - \mathbf{y}^T \mathbf{A})\mathbf{x} - tz_0 + t\mathbf{y}^T \mathbf{b} \geqslant 0$$

for all $\mathbf{x} \geqslant \mathbf{0}, t \geqslant 0$. Setting $t = 0$ yields $\mathbf{y}^T \mathbf{A} \leqslant \mathbf{c}^T$, which says $\mathbf{y}$ is feasible for the dual. Setting $\mathbf{x} = \mathbf{0}$ and $t = 1$ yields $\mathbf{y}^T \mathbf{b} \geqslant z_0$, which in view of Lemma 1 and its corollary shows that $\mathbf{y}$ is optimal for the dual. ∎

4.3 RELATIONS TO THE SIMPLEX PROCEDURE

In this section the Duality Theorem is proved by making explicit use of the characteristics of the simplex procedure. As a result of this proof it becomes clear that once the primal is solved by the simplex procedure a solution to the dual is readily obtainable.

Suppose that for the linear program

$$\begin{aligned} &\text{minimize} \quad \mathbf{c}^T \mathbf{x} \\ &\text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geqslant \mathbf{0}, \end{aligned} \tag{5}$$

we have the optimal basic feasible solution $\mathbf{x} = (\mathbf{x}_{\mathrm{B}},\ \mathbf{0})$ with corresponding basis $\mathbf{B}$. We shall determine a solution of the dual program

$$\begin{aligned} &\text{maximize} \quad \mathbf{y}^T \mathbf{b} \\ &\text{subject to} \quad \mathbf{y}^T \mathbf{A} \leqslant \mathbf{c}^T \end{aligned} \tag{6}$$

in terms of $\mathbf{B}$.

We partition $\mathbf{A}$ as $\mathbf{A} = [\mathbf{B},\ \mathbf{D}]$. Since the basic feasible solution $\mathbf{x_B} = \mathbf{B}^{-1}\mathbf{b}$ is optimal, the relative cost vector $\mathbf{r}$ must be nonnegative in each component. From Section 3.6 we have

$$\mathbf{r}_\mathbf{D}^T = \mathbf{c}_\mathbf{D}^T - \mathbf{c}_\mathbf{B}^T \mathbf{B}^{-1} \mathbf{D},$$

and since $\mathbf{r_D}$ is nonnegative in each component we have $\mathbf{c_B}^T\mathbf{B}^{-1}\mathbf{D} \leqslant \mathbf{c_D}^T$.

Now define $\mathbf{y}^T = \mathbf{c_B}^T\mathbf{B}^{-1}$. We show that this choice of $\mathbf{y}$ solves the dual problem. We have

$$\mathbf{y}^T\mathbf{A} = [\mathbf{y}^T\mathbf{B},\ \mathbf{y}^T\mathbf{D}] = [\mathbf{c_B}^T,\ \mathbf{c_B}^T\mathbf{B}^{-1}\mathbf{D}] \leqslant [\mathbf{c_B}^T,\ \mathbf{c_D}^T] = \mathbf{c}^T.$$

Thus since $\mathbf{y}^T\mathbf{A} \leqslant \mathbf{c}^T$, $\mathbf{y}$ is feasible for the dual. On the other hand,

$$\mathbf{y}^T\mathbf{b} = \mathbf{c_B}^T\mathbf{B}^{-1}\mathbf{b} = \mathbf{c_B}^T\mathbf{x_B},$$

and thus the value of the dual objective function for this $\mathbf{y}$ is equal to the value of the primal problem. This, in view of Lemma 1, Section 4.2, establishes the optimality of $\mathbf{y}$ for the dual. The above discussion yields an alternative derivation of the main portion of the Duality Theorem.

Theorem. *Let the linear program (5) have an optimal basic feasible solution corresponding to the basis* $\boldsymbol{B}$. *Then the vector* $\mathbf{y}$ *satisfying* $\mathbf{y}^T = \boldsymbol{c}_{\boldsymbol{B}}^T\boldsymbol{B}^{-1}$ *is an optimal solution to the dual program (6). The optimal values of both problems are equal.*

We turn now to a discussion of how the solution of the dual can be obtained directly from the final simplex tableau of the primal. Suppose that embedded in the original matrix $\mathbf{A}$ is an $m \times m$ identity matrix. This will be the case if, for example, m slack variables are employed to convert inequalities to equalities. Then in the final tableau the matrix $\mathbf{B}^{-1}$ appears where the identity appeared in the beginning. Furthermore, in the last row the components corresponding to this identity matrix will be $\mathbf{c_I}^T - \mathbf{c_B}^T\mathbf{B}^{-1}$, where $\mathbf{c_I}$ is the m-vector representing the cost coefficients of the variables corresponding to the columns of the original identity matrix. Thus by subtracting these cost coefficients from the corresponding elements in the last row, the negative of the solution $\mathbf{y}^T = \mathbf{c_B}^T\mathbf{B}^{-1}$ to the dual is obtained. In particular, if, as is the case with slack variables, $\mathbf{c_I} = \mathbf{0}$, then the elements in the last row under $\mathbf{B}^{-1}$ are equal to the negative of components of the solution to the dual.

Example. Consider the primal program

$$\begin{array}{rl}
\text{minimize} & -x_1 - 4x_2 - 3x_3 \\
\text{subject to} & 2x_1 + 2x_2 + x_3 \leqslant 4 \\
& x_1 + 2x_2 + 2x_3 \leqslant 6 \\
& x_1 \geqslant 0,\ x_2 \geqslant 0,\ x_3 \geqslant 0.
\end{array}$$

This can be solved by introducing slack variables and using the simplex procedure. The appropriate sequence of tableaus is given below without explanation.

2	2	1	1	0	4
1	2	2	0	1	6
−1	−4	−3	0	0	0
1	1	1/2	1/2	0	2
−1	0	1	−1	1	2
3	0	−1	2	0	8
3/2	1	0	1	−1/2	1
−1	0	1	−1	1	2
2	0	0	1	1	10

The optimal solution is $x_1 = 0,\ x_2 = 1,\ x_3 = 2$. The corresponding dual program is

$$\begin{array}{ll} \text{maximize} & 4\lambda_1 + 6\lambda_2 \\ \text{subject to} & 2\lambda_1 + \lambda_2 \leqslant -1 \\ & 2\lambda_1 + 2\lambda_2 \leqslant -4 \\ & \lambda_1 + 2\lambda_2 \leqslant -3 \\ & \lambda_1 \leqslant 0,\ \lambda_2 \leqslant 0. \end{array}$$

The optimal solution to the dual is obtained directly from the last row of the simplex tableau under the columns where the identity appeared in the first tableau: $\lambda_1 = -1,\ \lambda_2 = -1$.

Geometric Interpretation

The duality relations can be viewed in terms of the dual interpretations of linear constraints emphasized in Chapter 3. Consider a linear program in standard form. For sake of concreteness we consider the problem

$$\begin{array}{ll} \text{minimize} & 18x_1 + 12x_2 + 2x_3 + 6x_4 \\ \text{subject to} & 3x_1 + x_2 - 2x_3 + x_4 = 2 \\ & x_1 + 3x_2 - x_4 = 2 \\ & x_1 \geqslant 0,\ x_2 \geqslant 0,\ x_3 \geqslant 0,\ x_4 \geqslant 0. \end{array}$$

The columns of the constraints are represented in requirements space in Fig. 4.2. A basic solution represents construction of $\mathbf{b}$ with positive weights on two of the $\mathbf{a}_i$'s. The dual problem is

$$\begin{array}{ll} \text{maximize} & 2\lambda_1 + 2\lambda_2 \\ \text{subject to} & 3\lambda_1 + \lambda_2 \leqslant 18 \\ & \lambda_1 + 3\lambda_2 \leqslant 12 \\ & -2\lambda_1 \leqslant 2 \\ & \lambda_1 - \lambda_2 \leqslant 6. \end{array}$$

The dual problem is shown geometrically in Fig. 4.3. Each column $\mathbf{a}_i$ of the primal defines a constraint of the dual as a half-space whose boundary is orthogonal

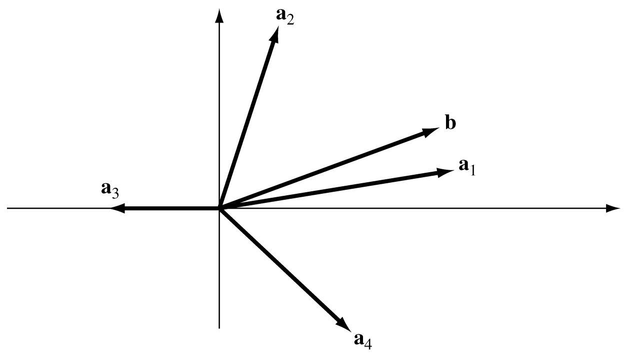

Fig. 4.2 The primal requirements space

to that column vector and is located at a point determined by c_i. The dual objective is maximized at an extreme point of the dual feasible region. At this point exactly two dual constraints are active. These active constraints correspond to an optimal basis of the primal. In fact, the vector defining the dual objective is a positive linear combination of the vectors. In the specific example, **b** is a positive combination of $\mathbf{a}_1$ and $\mathbf{a}_2$. The weights in this combination are the x_i's in the solution of the primal.

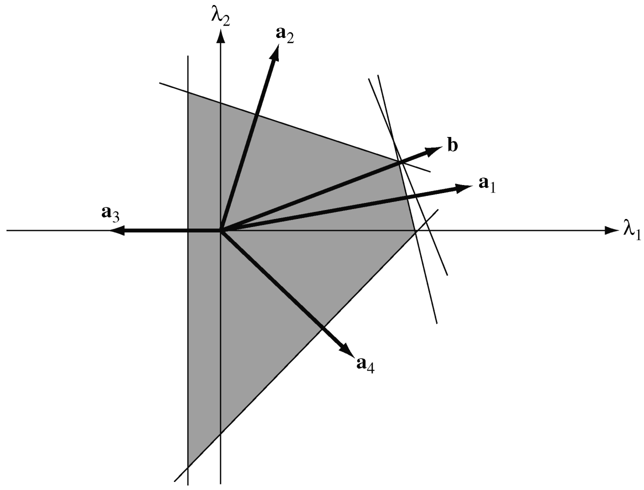

Fig. 4.3 The dual in activity space

Simplex Multipliers

We conclude this section by giving an economic interpretation of the relation between the simplex basis and the vector $\mathbf{y}$. At any point in the simplex procedure we may form the vector $\mathbf{y}$ satisfying $\mathbf{y}^T = \mathbf{c}_\mathbf{B}^T\mathbf{B}^{-1}$. This vector is not a solution to the dual unless $\mathbf{B}$ is an optimal basis for the primal, but nevertheless, it has an economic interpretation. Furthermore, as we have seen in the development of the revised simplex method, this $\mathbf{y}$ vector can be used at every step to calculate the relative cost coefficients. For this reason $\mathbf{y}^T = \mathbf{c}_\mathbf{B}^T\mathbf{B}^{-1}$, corresponding to any basis, is often called the vector of *simplex multipliers*.

Let us pursue the economic interpretation of these simplex multipliers. As usual, denote the columns of $\mathbf{A}$ by $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ and denote by $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m$ the m unit vectors in E^m. The components of the $\mathbf{a}_i$'s and $\mathbf{b}$ tell how to construct these vectors from the $\mathbf{e}_i$'s.

Given any basis $\mathbf{B}$, however, consisting of m columns of $\mathbf{A}$, any other vector can be constructed (synthetically) as a linear combination of these basis vectors. If there is a unit cost c_i associated with each basis vector $\mathbf{a}_i$, then the cost of a (synthetic) vector constructed from the basis can be calculated as the corresponding linear combination of the c_i's associated with the basis. In particular, the cost of the jth unit vector, $\mathbf{e}_j$, when constructed from the basis $\mathbf{B}$, is λ_j, the jth component of $\mathbf{y}^T = \mathbf{c}_\mathbf{B}^T\mathbf{B}^{-1}$. Thus the λ_j's can be interpreted as synthetic prices of the unit vectors.

Now, any vector can be expressed in terms of the basis $\mathbf{B}$ in two steps: (i) express the unit vectors in terms of the basis, and then (ii) express the desired vector as a linear combination of unit vectors. The corresponding synthetic cost of a vector constructed from the basis $\mathbf{B}$ can correspondingly be computed directly by: (i) finding the synthetic price of the unit vectors, and then (ii) using these prices to evaluate the cost of the linear combination of unit vectors. Thus, the simplex multipliers can be used to quickly evaluate the synthetic cost of any vector that is expressed in terms of the unit vectors. The difference between the true cost of this vector and the synthetic cost is the relative cost. The process of calculating the synthetic cost of a vector, with respect to a given basis, by using the simplex multipliers is sometimes referred to as *pricing out* the vector.

Optimality of the primal corresponds to the situation where every vector $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ is cheaper when constructed from the basis than when purchased directly at its own price. Thus we have $\mathbf{y}^T\mathbf{a}_i \leqslant c_i$ for $i = 1, 2, \ldots, n$ or equivalently $\mathbf{y}^T\mathbf{A} \leqslant \mathbf{c}^T$.

4.4 SENSITIVITY AND COMPLEMENTARY SLACKNESS

The optimal values of the dual variables in a linear program can, as we have seen, be interpreted as prices. In this section this interpretation is explored in further detail.

Sensitivity

Suppose in the linear program

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geqslant \mathbf{0}, \end{array} \tag{7}$$

the optimal basis is $\mathbf{B}$ with corresponding solution $(\mathbf{x}_\mathbf{B},\ \mathbf{0})$, where $\mathbf{x}_\mathbf{B} = \mathbf{B}^{-1}\mathbf{b}$. A solution to the corresponding dual is $\mathbf{y}^T = \mathbf{c}_\mathbf{B}^T\mathbf{B}^{-1}$.

Now, assuming nondegeneracy, small changes in the vector $\mathbf{b}$ will not cause the optimal basis to change. Thus for $\mathbf{b} + \boldsymbol{\Delta}\mathbf{b}$ the optimal solution is

$$\mathbf{x} = (\mathbf{x}_\mathbf{B} + \boldsymbol{\Delta}\mathbf{x}_\mathbf{B},\ \mathbf{0}),$$

where $\boldsymbol{\Delta}\mathbf{x}_\mathbf{B} = \mathbf{B}^{-1}\boldsymbol{\Delta}\mathbf{b}$. Thus the corresponding increment in the cost function is

$$\boldsymbol{\Delta}z = \mathbf{c}_\mathbf{B}^T\boldsymbol{\Delta}\mathbf{x}_\mathbf{B} = \mathbf{y}^T\boldsymbol{\Delta}\mathbf{b}. \tag{8}$$

This equation shows that $\mathbf{y}$ gives the sensitivity of the optimal cost with respect to small changes in the vector $\mathbf{b}$. In other words, if a new program were solved with $\mathbf{b}$ changed to $\mathbf{b} + \boldsymbol{\Delta}\mathbf{b}$, the change in the optimal value of the objective function would be $\mathbf{y}^T\boldsymbol{\Delta}\mathbf{b}$.

This interpretation of the dual vector $\mathbf{y}$ is intimately related to its interpretation as a vector of simplex multipliers. Since λ_j is the price of the unit vector e_j when constructed from the basis $\mathbf{B}$, it directly measures the change in cost due to a change in the jth component of the vector $\mathbf{b}$. Thus, λ_j may equivalently be considered as the *marginal price* of the component b_j, since if b_j is changed to $b_j + \Delta b_j$ the value of the optimal solution changes by $\lambda_j \Delta b_j$.

If the linear program is interpreted as a diet problem, for instance, then λ_j is the maximum price per unit that the dietitian would be willing to pay for a small amount of the jth nutrient, because decreasing the amount of nutrient that must be supplied by food will reduce the food bill by λ_j dollars per unit. If, as another example, the linear program is interpreted as the problem faced by a manufacturer who must select levels $x_1,\ x_2,\ \ldots,\ x_n$ of n production activities in order to meet certain required levels of output $b_1,\ b_2,\ \ldots,\ b_m$ while minimizing production costs, the λ_i's are the marginal prices of the outputs. They show directly how much the production cost varies if a small change is made in the output levels.

Complementary Slackness

The optimal solutions to primal and dual programs satisfy an additional relation that has an economic interpretation. This relation can be stated for any pair of dual linear programs, but we state it here only for the asymmetric and the symmetric pairs defined in Section 4.1.

Theorem. (Complementary slackness—asymmetric form). *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be feasible solutions for the primal and dual programs, respectively, in the pair (2). A necessary and sufficient condition that they both be optimal solutions is that*† *for all i*

i) $x_i > 0 \Rightarrow \mathbf{y}^T\boldsymbol{a}_i = c_i$
ii) $x_i = 0 \Leftarrow \mathbf{y}^T\boldsymbol{a}_j < c_j$.

Proof. If the stated conditions hold, then clearly $(\mathbf{y}^T\mathbf{A} - \mathbf{c}^T)\mathbf{x} = 0$. Thus $\mathbf{y}^T\mathbf{b} = \mathbf{c}^T\mathbf{x}$, and by the corollary to Lemma 1, Section 4.2, the two solutions are optimal. Conversely, if the two solutions are optimal, it must hold, by the Duality Theorem, that $\mathbf{y}^T\mathbf{b} = \mathbf{c}^T\mathbf{x}$ and hence that $(\mathbf{y}^T\mathbf{A} - \mathbf{c}^T)\mathbf{x} = 0$. Since each component of $\mathbf{x}$ is nonnegative and each component of $\mathbf{y}^T\mathbf{A} - \mathbf{c}^T$ is nonpositive, the conditions (i) and (ii) must hold. ∎

Theorem. (Complementary slackness—symmetric form). *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be feasible solutions for the primal and dual programs, respectively, in the pair (1). A necessary and sufficient condition that they both be optimal solutions is that for all i and j*

i) $x_i > 0 \Rightarrow \mathbf{y}^T\mathbf{a}_i = c_i$
ii) $x_i = 0 \Leftarrow \mathbf{y}^T\mathbf{a}_i < c_i$
iii) $\lambda_j > 0 \Rightarrow \mathbf{a}^j\mathbf{x} = b_j$
iv) $\lambda_j = 0 \Leftarrow \mathbf{a}^j\mathbf{x} > b_j$,

(where $\mathbf{a}^j$ *is the jth row of* $\mathbf{A}$*).*

Proof. This follows by transforming the previous theorem. ∎

The complementary slackness conditions have a rather obvious economic interpretation. Thinking in terms of the diet problem, for example, which is the primal part of a symmetric pair of dual problems, suppose that the optimal diet supplies more than b_j units of the jth nutrient. This means that the dietitian would be unwilling to pay anything for small quantities of that nutrient, since availability of it would not reduce the cost of the optimal diet. This, in view of our previous interpretation of λ_j as a marginal price, implies $\lambda_j = 0$ which is (iv) of Theorem 4.4. The other conditions have similar interpretations which the reader can work out.

4.5 MAX FLOW–MIN CUT THEOREM

One of the most exemplary pairs of linear primal and dual problems is the max-flow and min-cut theorem, which we describe in this section. The maximal flow problem described in Chapter 2 can be expressed more compactly in terms of the node–arc incidence matrix (see Appendix D). Let $\mathbf{x}$ be the vector of arc flows x_{ij} (ordered in any way). Let $\mathbf{A}$ be the corresponding node-arc incidence matrix. Finally, let $\mathbf{e}$ be a vector with dimension equal to the number of nodes and having a + 1 component on

† The symbol ⇒ means ”implies” and ⇐ means “is implied by.”

node 1, a – 1 on node m, and all other components zero. The maximal flow problem is then

$$\begin{aligned} \text{maximize} \quad & f \\ \text{subject to} \quad & \mathbf{Ax} - f\mathbf{e} = \mathbf{0} \\ & \mathbf{x} \leqslant \mathbf{k}. \end{aligned} \tag{9}$$

The coefficient matrix of this problem is equal to the node–arc incidence matrix with an additional column for the flow variable f. Any basis of this matrix is triangular, and hence as indicated by the theory in the transportation problem in Chapter 3, the simplex method can be effectively employed to solve this problem. However, instead of the simplex method, a simple algorithm based on the tree algorithm (also see Appendix D) can be used.

Max Flow Augmenting Algorithm

The basic strategy of the algorithm is quite simple. First we recognize that it is possible to send nonzero flow from node 1 to node m only if node m is reachable from node 1. The tree procedure can be used to determine if m is in fact reachable; and if it is reachable, the algorithm will produce a path from 1 to m. By examining the arcs along this path, we can determine the one with minimum capacity. We may then construct a flow equal to this capacity from 1 to m by using this path. This gives us a strictly positive (and integer-valued) initial flow.

Next consider the nature of the network at this point in terms of additional flows that might be assigned. If there is already flow x_{ij} in the arc $(i,\ j)$, then the effective capacity of that arc is reduced by x_{ij}(to $k_{ij} - x_{ij}$), since that is the maximal amount of additional flow that can be assigned to that arc. On the other hand, the effective reverse capacity, on the arc (j, i), is increased by x_{ij}(to $k_{ji} + x_{ij}$), since a small incremental backward flow is actually realized as a reduction in the forward flow through that arc. Once these changes in capacities have been made, the tree procedure can again be used to find a path from node 1 to node m on which to assign additional flow. (Such a path is termed an *augmenting path*.) Finally, if m is not reachable from 1, no additional flow can be assigned, and the procedure is complete.

It is seen that the method outlined above is based on repeated application of the tree procedure, which is implemented by labeling and scanning. By including slightly more information in the labels than in the basic tree algorithm, the minimum arc capacity of the augmenting path can be determined during the initial scanning, instead of by reexamining the arcs after the path is found. A typical label at a node i has the form $(k,\ c_i)$, where k denotes a precursor node and c_i is the maximal flow that can be sent from the source to node i through the path created by the previous labeling and scanning. The complete procedure is this:

Step 0. Set all $x_{ij} = 0$ and $f = 0$.

Step 1. Label node 1 $(-,\ \infty)$. All other nodes are unlabeled.

Step 2. Select any labeled node i for scanning. Say it has label (k, c_i). For all unlabeled nodes j such that (i, j) is an arc with $x_{ij} < k_{ij}$, assign the label (i, c_j), where $c_j = \min\{c_i, k_{ij} - x_{ij}\}$. For all unlabeled nodes j such that (j, i) is an arc with $x_{ji} > 0$, assign the label (i, c_j), where $c_j = \min\{c_i, x_{ji}\}$.

Step 3. Repeat Step 2 until either node m is labeled or until no more labels can be assigned. In this latter case, the current solution is optimal.

Step 4. (Augmentation.) If the node m is labeled (i, c_m), then increase f and the flow on arc (i, m) by c_m. Continue to work backward along the augmenting path determined by the nodes, increasing the flow on each arc of the path by c_m. Return to Step 1.

The validity of the algorithm should be fairly apparent, that is, the finite termination of the algorithm. However, a complete proof is deferred until we consider the max flow-min cut theorem below.

Example. An example of the above procedure is shown in Fig. 4.4. Node 1 is the source, and node 6 is the sink. The original network with capacities indicated on the arcs is shown in Fig. 4.4(a). Also shown in that figure are the initial labels obtained by the procedure. In this case the sink node is labeled, indicating that a flow of 1 unit can be achieved. The augmenting path of this flow is shown in Fig. 4.4(b). Numbers in square boxes indicate the total flow in an arc. The new labels are then found and added to that figure. Note that node 2 cannot be labeled from node 1 because there is no unused capacity in that direction. Node 2 can, however, be labeled from node 4, since the existing flow provides a reverse capacity of 1 unit. Again the sink is labeled, and 1 unit more flow can be constructed. The augmenting path is shown in Fig. 4.4(c). A new labeling is appended to that figure. Again the sink is labeled, and an additional 1 unit of flow can be sent from source to sink. The path of this 1 unit is shown in Fig. 4.4(d). Note that it includes a flow from node 4 to node 2, even though flow was not allowed in this direction in the original network. This flow is allowable now, however, because there is already flow in the opposite direction. The total flow at this point is shown in Fig. 4.4(e). The flow levels are again in square boxes. This flow is maximal, since only the source node can be labeled.

Max Flow–Min Cut Theorem

A great deal of insight and some further results can be obtained through the introduction of the notion of *cuts* in a network. Given a network with source node 1 and sink node m, divide the nodes arbitrarily into two sets S and $\bar{S}$ such that the source node is in S and the sink is in $\bar{S}$. The set of arcs from S to $\bar{S}$ is a *cut* and is denoted $(S, \bar{S})$. The *capacity* of the cut is the sum of the capacities of the arcs in the cut.

An example of a cut is shown in Fig. 4.5. The set S consists of nodes 1 and 2, while $\bar{S}$ consists of 3, 4, 5, 6. The capacity of this cut is 4.

It should be clear that a path from node 1 to node m must include at least one arc in any cut, for the path must have an arc from the set S to the set $\bar{S}$. Furthermore, it

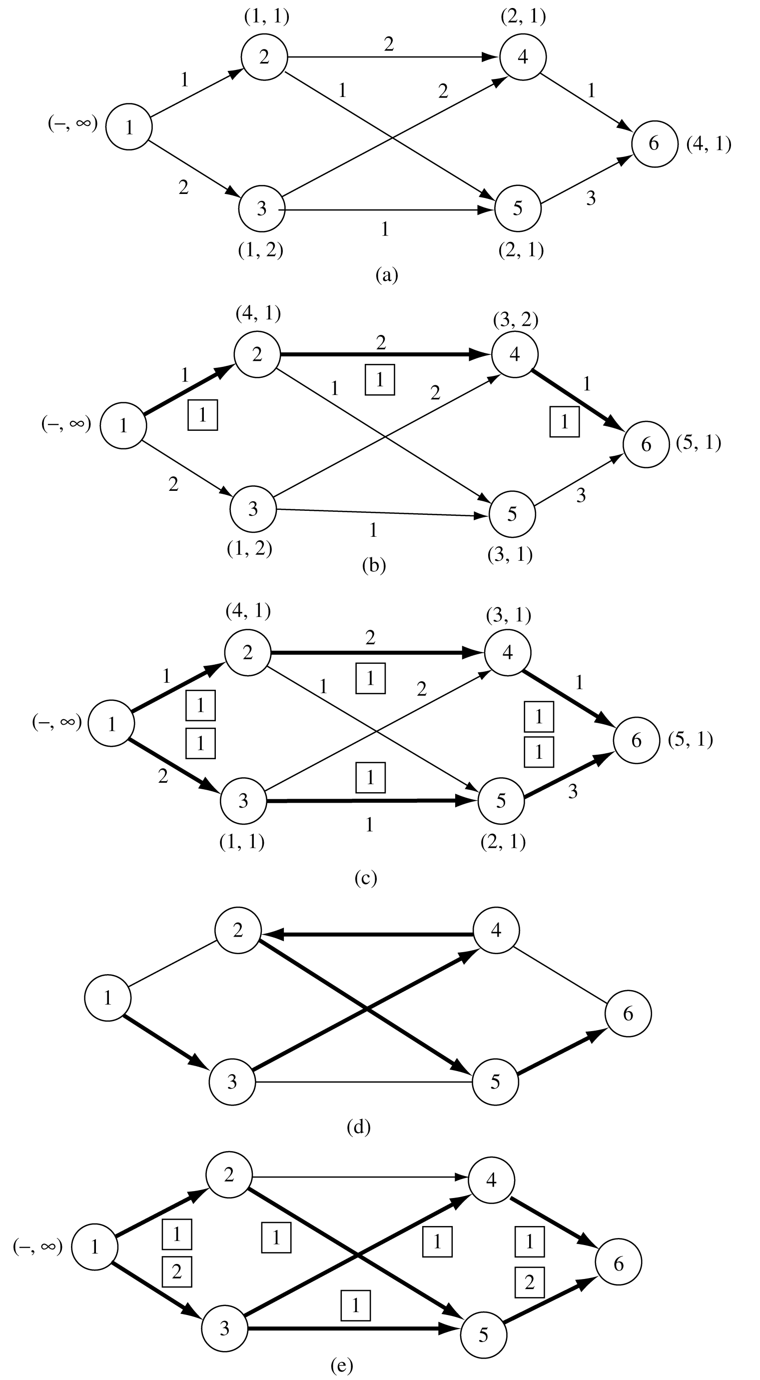

Fig. 4.4 Example of maximal flow problem

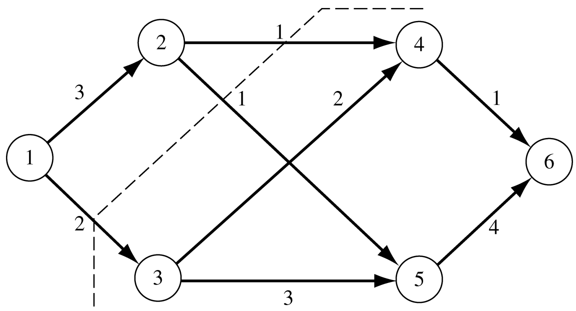

Fig. 4.5 A cut

is clear that the maximal amount of flow that can be sent through a cut is equal to its capacity. Thus each cut gives an upper bound on the value of the maximal flow problem. The max flow-min cut theorem states that equality is actually achieved for some cut. That is, the maximal flow is equal to the minimal cut capacity. It should be noted that the proof of the theorem also establishes the maximality of the flow obtained by the maximal flow algorithm.

Max Flow–Min Cut Theorem. *In a network the maximal flow between a source and a sink is equal to the minimal cut capacity of all cuts separating the source and sink.*

Proof. Since any cut capacity must be greater than or equal to the maximal flow, it is only necessary to exhibit a flow and a cut for which equality is achieved. Begin with a flow in the network that cannot be augmented by the maximal flow algorithm. For this flow find the effective arc capacities of all arcs for incremental flow changes as described earlier and apply the labeling procedure of the maximal flow algorithm. Since no augmenting path exists, the algorithm must terminate before the sink is labeled.

Let S and $\bar{S}$ consist of all labeled and unlabeled nodes, respectively. This defines a cut separating the source from the sink. All arcs originating in S and terminating in $\bar{S}$ have zero incremental capacity, or else a node in $\bar{S}$ could have been labeled. This means that each arc in the cut is saturated by the original flow; that is, the flow is equal to the capacity. Any arc originating in $\bar{S}$ and terminating in S, on the other hand, must have zero flow; otherwise, this would imply a positive incremental capacity in the reverse direction, and the originating node in $\bar{S}$ would be labeled. Thus, there is a total flow from S to $\bar{S}$ equal to the cut capacity, and zero flow from $\bar{S}$ to S. This means that the flow from source to sink is equal to the cut capacity. Thus the cut capacity must be minimal, and the flow must be maximal. ∎

In the network of Fig. 4.4, the minimal cut corresponds to the S consisting only of the source. That cut capacity is 3. Note that in accordance with the max flow–min cut theorem, this is equal to the value of the maximal flow, and the minimal cut is determined by the final labeling in Fig. 4.4(e). In Fig. 4.5 the cut shown is also minimal, and the reader should easily be able to determine the pattern of maximal flow.

Relation to Duality

The character of the max flow–min cut theorem suggests a connection with the Duality Theorem. We conclude this section by exploring this connection.

The maximal flow problem is a linear program, which is expressed formally by (9). The dual problem is found to be

$$\begin{aligned} \text{minimize} \quad & \mathbf{w}^T\mathbf{k} \\ \text{subject to} \quad & \mathbf{u}^T\mathbf{A} = \mathbf{w}^T \\ & \mathbf{u}^T\mathbf{e} = 1 \\ & \mathbf{w} \geq \mathbf{0}. \end{aligned} \tag{10}$$

When written out in detail, the dual is

$$\begin{aligned} \text{minimize} \quad & \sum_{ij} w_{ij}k_{ij} \\ \text{subject to} \quad & u_i - u_j = w_{ij} \\ & u_1 - u_m = 1 \\ & w_{ij} \geq 0. \end{aligned} \tag{11}$$

A pair i, j is included in the above only if (i, j) is an arc of the network.

A feasible solution to this dual problem can be found in terms of any cut set $(S, \bar{S})$. In particular, it is easily seen that

$$u_i = \begin{cases} 1 & \text{if} \quad i \in S \\ 0 & \text{if} \quad i \in \bar{S} \end{cases} \tag{12}$$
$$w_{ij} = \begin{cases} 1 & \text{if} \quad (i, j) \in (S, \bar{S}) \\ 0 & \text{otherwise} \end{cases}$$

is a feasible solution. The value of the dual problem corresponding to this solution is the cut capacity. If we take the cut set to be the one determined by the labeling procedure of the maximal flow algorithm as described in the proof of the theorem above, it can be seen to be optimal by verifying the complementary slackness conditions (a task we leave to the reader). The minimum value of the dual is therefore equal to the minimum cut capacity.

4.6 THE DUAL SIMPLEX METHOD

Often there is available a basic solution to a linear program which is not feasible but which prices out optimally; that is, the simplex multipliers are feasible for the dual problem. In the simplex tableau this situation corresponds to having no negative elements in the bottom row but an infeasible basic solution. Such a situation may arise, for example, if a solution to a certain linear programming problem is calculated and then a new problem is constructed by changing the vector $\mathbf{b}$. In such situations a basic feasible solution to the dual is available and hence it is desirable to pivot in such a way as to optimize the dual.

Rather than constructing a tableau for the dual problem (which, if the primal is in standard form; involves m free variables and n nonnegative slack variables), it is more efficient to work on the dual from the primal tableau. The complete technique based on this idea is the dual simplex method. In terms of the primal problem, it operates by maintaining the optimality condition of the last row while working toward feasibility. In terms of the dual problem, however, it maintains feasibility while working toward optimality.

Given the linear program

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geqslant \mathbf{0}, \end{array} \tag{13}$$

suppose a basis $\mathbf{B}$ is known such that $\mathbf{y}$ defined by $\mathbf{y}^T = \mathbf{c}_\mathbf{B}^T\mathbf{B}^{-1}$ is feasible for the dual. In this case we say that the corresponding basic solution to the primal, $\mathbf{x}_\mathbf{B} = \mathbf{B}^{-1}\mathbf{b}$, is *dual feasible*. If $\mathbf{x}_B \geqslant \mathbf{0}$ then this solution is also primal feasible and hence optimal.

The given vector $\mathbf{y}$ is feasible for the dual and thus satisfies $\mathbf{y}^T\mathbf{a}_j \leqslant c_j$, for $j = 1, 2, \ldots, n$. Indeed, assuming as usual that the basis is the first m columns of $\mathbf{A}$, there is equality

$$\mathbf{y}^T\mathbf{a}_j = c_j, \text{ for } j = 1, 2, \ldots, m, \tag{14a}$$

and (barring degeneracy in the dual) there is inequality

$$\mathbf{y}^T\mathbf{a}_j < c_j, \text{ for } j = m+1, \ldots, n. \tag{14b}$$

To develop one cycle of the dual simplex method, we find a new vector $\bar{\mathbf{y}}$ such that one of the equalities becomes an inequality and one of the inequalities becomes equality, while at the same time increasing the value of the dual objective function. The m equalities in the new solution then determine a new basis.

Denote the ith row of $\mathbf{B}^{-1}$ by $\mathbf{u}^i$. Then for

$$\bar{\mathbf{y}}^T = \mathbf{y}^T - \varepsilon\mathbf{u}^i, \tag{15}$$

we have $\bar{\mathbf{y}}^T\mathbf{a}_j = \mathbf{y}^T\mathbf{a}_j - \varepsilon\mathbf{u}^j\mathbf{a}_j$. Thus, recalling that $z_j = \mathbf{y}^T\mathbf{a}_j$ and noting that $\mathbf{u}^i\mathbf{a}_j = y_{ij}$, the ijth element of the tableau, we have

$$\bar{\mathbf{y}}^T\mathbf{a}_j = c_j, \qquad j = 1, 2, \ldots, m, \quad i \neq j \tag{16a}$$

$$\bar{\mathbf{y}}^T\mathbf{a}_i = c_i - \varepsilon \tag{16b}$$

$$\bar{\mathbf{y}}^T\mathbf{a}_j = z_j - \varepsilon y_{ij}, \quad j = m+1, \quad m+2, \ldots, n. \tag{16c}$$

Also,

$$\bar{\mathbf{y}}^T\mathbf{b} = \mathbf{y}^T\mathbf{b} - \varepsilon\mathbf{x}_{\mathbf{B}i}. \tag{17}$$

These last equations lead directly to the algorithm:

Step 1. Given a dual feasible basic solution $\mathbf{x_B}$, if $\mathbf{x_B} \geqslant \mathbf{0}$ the solution is optimal. If $\mathbf{x_B}$ is not nonnegative, select an index i such that the ith component of $\mathbf{x_B}$, $\mathbf{x_B}_i < 0$.

Step 2. If all $y_{ij} \geqslant 0$, $j = 1, 2, \ldots, n$, then the dual has no maximum (this follows since by (16) $\bar{\boldsymbol{\lambda}}$ is feasible for all $\varepsilon > 0$). If $y_{ij} < 0$ for some j, then let

$$\varepsilon_0 = \frac{z_k - c_k}{y_{ik}} = \min_j \left\{ \frac{z_j - c_j}{y_{ij}} : y_{ij} < 0 \right\}. \tag{18}$$

Step 3. Form a new basis $\mathbf{B}$ by replacing $\mathbf{a}_i$ by $\mathbf{a}_k$. Using this basis determine the corresponding basic dual feasible solution $\mathbf{x_B}$ and return to Step 1.

The proof that the algorithm converges to the optimal solution is similar in its details to the proof for the primal simplex procedure. The essential observations are: (a) from the choice of k in (18) and from (16a, b, c) the new solution will again be dual feasible; (b) by (17) and the choice $\mathbf{x}_{\mathbf{B}_i} < 0$, the value of the dual objective will increase; (c) the procedure cannot terminate at a nonoptimum point; and (d) since there are only a finite number of bases, the optimum must be achieved in a finite number of steps.

Example. A form of problem arising frequently is that of minimizing a positive combination of positive variables subject to a series of "greater than" type inequalities having positive coefficients. Such problems are natural candidates for application of the dual simplex procedure. The classical diet problem is of this type as is the simple example below.

$$\begin{array}{rl}
\text{minimize} & 3x_1 + 4x_2 + 5x_3 \\
\text{subject to} & x_i + 2x_2 + 3x_3 \geqslant 5 \\
& 2x_1 + 2x_2 + x_3 \geqslant 6 \\
x_1 \geqslant 0, & x_2 \geqslant 0, \quad x_3 \geqslant 0.
\end{array}$$

By introducing surplus variables and by changing the sign of the inequalities we obtain the initial tableau

−1	−2	−3	1	0	−5
−②	−2	−1	0	1	−6
3	4	5	0	0	0

Initial tableau

The basis corresponds to a dual feasible solution since all of the $c_j - z_j$'s are nonnegative. We select any $\mathbf{x}_{\mathbf{B}_i} < 0$, say $x_5 = -6$, to remove from the set of basic variables. To find the appropriate pivot element in the second row we compute the ratios $(z_j - c_j)/y_{2j}$ and select the minimum positive ratio. This yields the pivot indicated. Continuing, the remaining tableaus are

0	−①	−5/2	1	−1/2	−2
1	1	1/2	0	−1/2	3
0	1	7/2	0	3/2	9

Second tableau

0	1	5/2	−1	1/2	2
1	0	−2	1	−1	1
0	0	1	1	1	11

Final tableau

The third tableau yields a feasible solution to the primal which must be optimal. Thus the solution is $x_1 = 1,\ x_2 = 2,\ x_3 = 0$.

*4.7 *THE PRIMAL-DUAL ALGORITHM

In this section a procedure is described for solving linear programming problems by working simultaneously on the primal and the dual problems. The procedure begins with a feasible solution to the dual that is improved at each step by optimizing an *associated restricted primal* problem. As the method progresses it can be regarded as striving to achieve the complementary slackness conditions for optimality. Originally, the primal-dual method was developed for solving a special kind of linear program arising in network flow problems, and it continues to be the most efficient procedure for these problems. (For general linear programs the dual simplex method is most frequently used). In this section we describe the generalized version of the algorithm and point out an interesting economic interpretation of it. We consider the program

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geqslant \mathbf{0} \end{array} \tag{19}$$

and the corresponding dual program

$$\begin{array}{ll} \text{maximize} & \mathbf{y}^T\mathbf{b} \\ \text{subject to} & \mathbf{y}^T\mathbf{A} \leqslant \mathbf{c}^T. \end{array} \tag{20}$$

Given a feasible solution $\mathbf{y}$ to the dual, define the subset P of 1, 2, . . . , n by $i \in P$ if $\mathbf{y}^T\mathbf{a}_i = c_i$ where $\mathbf{a}_i$ is the ith column of $\mathbf{A}$. Thus, since $\mathbf{y}$ is dual feasible, it follows that $i \notin P$ implies $\mathbf{y}^T\mathbf{a}_i < c_i$. Now corresponding to $\mathbf{y}$ and P, we define the *associated restricted primal* problem

$$\begin{array}{lll} \text{minimize} & \mathbf{1}^T\mathbf{y} & \\ \text{subject to} & \mathbf{A}\mathbf{x} + \mathbf{y} = \mathbf{b} & \\ & \mathbf{x} \geqslant \mathbf{0}, & x_i = 0 \text{ for } i \notin P \\ & \mathbf{y} \geqslant \mathbf{0}, & \end{array} \tag{21}$$

where $\mathbf{1}$ denotes the m-vector (1, 1, . . ., 1).

The dual of this associated restricted primal is called the *associated restricted dual*. It is

$$\begin{aligned} &\text{maximize} \quad \mathbf{u}^T\mathbf{b} \\ &\text{subject to} \quad \mathbf{u}^T\mathbf{a}_i \leqslant \mathbf{0}, \quad i \notin P \\ &\qquad\qquad\quad \mathbf{u} \leqslant \mathbf{1}. \end{aligned} \tag{22}$$

The condition for optimality of the primal-dual method is expressed in the following theorem.

Primal-Dual Optimality Theorem. *Suppose that* $\mathbf{y}$ *is feasible for the dual and that* $\mathbf{x}$ *and* $\mathbf{y} = 0$ *is feasible (and of course optimal) for the associated restricted primal. Then* $\mathbf{x}$ *and* $\mathbf{y}$ *are optimal for the original primal and dual programs, respectively.*

Proof. Clearly $\mathbf{x}$ is feasible for the primal. Also we have $\mathbf{c}^T\mathbf{x} = \mathbf{y}^T\mathbf{A}\mathbf{x}$, because $\mathbf{y}^T\mathbf{A}$ is identical to $\mathbf{c}^T$ on the components corresponding to nonzero elements of $\mathbf{x}$. Thus $\mathbf{c}^T\mathbf{x} = \mathbf{y}^T\mathbf{A}\mathbf{x} = \mathbf{y}^T\mathbf{b}$ and optimality follows from Lemma 1, Section 4.2. ∎

The primal–dual method starts with a feasible solution to the dual and then optimizes the associated restricted primal. If the optimal solution to this associated restricted primal is not feasible for the primal, the feasible solution to the dual is improved and a new associated restricted primal is determined. Here are the details:

Step 1. Given a feasible solution $\mathbf{y}_0$ to the dual program (20), determine the associated restricted primal according to (21).

Step 2. Optimize the associated restricted primal. If the minimal-value of this problem is zero, the corresponding solution is optimal for the original primal by the Primal-Dual Optimality Theorem.

Step 3. If the minimal value of the associated restricted primal is strictly positive, obtain from the final simplex tableau of the restricted primal, the solution $\mathbf{u}_0$ of the associated restricted dual (22). If there is no j for which $\mathbf{u}_0^T\mathbf{a}_j > 0$ conclude the primal has no feasible solutions. If, on the other hand, for at least one j, $\mathbf{u}_0^T\mathbf{a}_j > 0$, define the new dual feasible vector

$$\mathbf{y} = \mathbf{y}_0 + \varepsilon_0\mathbf{u}_0$$

where

$$\varepsilon_0 = \frac{c_k - \mathbf{y}_0^T\mathbf{a}_k}{\mathbf{u}_0^T\mathbf{a}_k} = \min_j \left\{ \frac{c_j - \mathbf{y}_0^T\mathbf{a}_j}{\mathbf{u}_0^T\mathbf{a}_j} \; : \; \mathbf{u}_0^T\mathbf{a}_j > 0 \right\}.$$

Now go back to Step 1 using this $\mathbf{y}$.

To prove convergence of this method a few simple observations and explanations must be made. First we verify the statement made in Step 3 that $\mathbf{u}_0^T\mathbf{a}_j \leqslant 0$ for all j implies that the primal has no feasible solution. The vector $\mathbf{y}_\varepsilon = \mathbf{y}_0 + \varepsilon\mathbf{u}_0$ is feasible for the dual problem for all positive ε, since $\mathbf{u}_0^T\mathbf{A} \leqslant 0$. In addition, $\mathbf{y}_\varepsilon^T\mathbf{b} = \mathbf{y}_0^T\mathbf{b} + \varepsilon\mathbf{u}_0^T\mathbf{b}$ and, since $\mathbf{u}_0^T\mathbf{b} = \mathbf{1}^T y > 0$, we see that as ε is increased we obtain an unbounded solution to the dual. In view of the Duality Theorem, this implies that there is no feasible solution to the primal.

Next suppose that in Step 3, for at least one j, $\mathbf{u}_0^T\mathbf{a}_j > 0$. Again we define the family of vectors $\mathbf{y}_\varepsilon = \mathbf{y}_0 + \varepsilon\mathbf{u}_0$. Since $\mathbf{u}_0$ is a solution to (22) we have $\mathbf{u}_0^T\mathbf{a}_i \leqslant 0$ for $i \in$

P, and hence for small positive ε the vector $\mathbf{y}_\varepsilon$ is feasible for the dual. We increase ε to the first point where one of inequalities $\mathbf{y}_\varepsilon^T \mathbf{a}_j < c_j,\ j \notin P$ becomes an equality. This determines $\varepsilon_0 > 0$ and k. The new $\mathbf{y}$ vector corresponds to an increased value of the dual objective $\mathbf{y}^T \mathbf{b} = \mathbf{y}_0^T \mathbf{b} + \varepsilon \mathbf{u}_0^T \mathbf{b}$. In addition, the corresponding new set P now includes the index k. Any other index i that corresponded to a positive value of x_i in the associated restricted primal is in the new set P, because by complementary slackness $\mathbf{u}_0^T \mathbf{a}_i = 0$ for such an i and thus $\mathbf{y}^T \mathbf{a}_i = \mathbf{y}_0^T \mathbf{a}_i + \varepsilon_0 \mathbf{u}_0^T \mathbf{a}_i = c_i$. This means that the old optimal solution is feasible for the new associated restricted primal and that $\mathbf{a}_k$ can be pivoted into the basis. Since $\mathbf{u}_0^T \mathbf{a}_k > 0$, pivoting in $\mathbf{a}_k$ will decrease the value of the associated restricted primal.

In summary, it has been shown that at each step either an improvement in the associated primal is made or an infeasibility condition is detected. Assuming nondegeneracy, this implies that no basis of the associated primal is repeated—and since there are only a finite number of possible bases, the solution is reached in a finite number of steps.

The primal-dual algorithm can be given an interesting interpretation in terms of the manufacturing problem in Example 2, Section 2.2. Suppose we own a facility that is capable of engaging in n different production activities each of which produces various amounts of m commodities. Each activity i can be operated at any level $x_i \geqslant 0$, but when operated at the unity level the ith activity costs c_i dollars and yields the m commodities in the amounts specified by the m-vector $\mathbf{a}_i$. Assuming linearity of the production facility, if we are given a vector $\mathbf{b}$ describing output requirements of the m commodities, and we wish to produce these at minimum cost, ours is the primal problem.

Imagine that an entrepreneur *not knowing* the value of our requirements vector $\mathbf{b}$ decides to sell us these requirements directly. He assigns a price vector $\mathbf{y}_0$ to these requirements such that $\mathbf{y}_0^T \mathbf{A} \leqslant \mathbf{c}$. In this way his prices are competitive with our production activities, and he can assure us that purchasing directly from him is no more costly than engaging activities. As owner of the production facilities we are reluctant to abandon our production enterprise but, on the other hand, we deem it not frugal to engage an activity whose output can be duplicated by direct purchase for lower cost. Therefore, we decide to engage only activities that cannot be duplicated cheaper, and at the same time we attempt to minimize the total business volume given the entrepreneur. Ours is the associated restricted primal problem.

Upon receiving our order, the greedy entrepreneur decides to modify his prices in such a manner as to keep them competitive with our activities but increase the cost of our order. As a reasonable and simple approach he seeks new prices of the form

$$\mathbf{y} = \mathbf{y}_0 + \varepsilon \mathbf{u}_0,$$

where he selects $\mathbf{u}_0$ as the solution to

$$\begin{aligned} \text{maximize} \quad & \mathbf{u}^T \mathbf{y} \\ \text{subject to} \quad & \mathbf{u}^T \mathbf{a}_i \leqslant \mathbf{0}, \quad i \in P \\ & \mathbf{u} \leqslant \mathbf{1}. \end{aligned}$$

The first set of constraints is to maintain competitiveness of his new price vector for small ε, while the second set is an arbitrary bound imposed to keep this subproblem bounded. It is easily shown that the solution $\mathbf{u}_0$ to this problem is identical to the solution of the associated dual (22). After determining the maximum ε to maintain feasibility, he announces his new prices.

At this point, rather than concede to the price adjustment, we recalculate the new minimum volume order based on the new prices. As the greedy (and shortsighted) entrepreneur continues to change his prices in an attempt to maximize profit he eventually finds he has reduced his business to zero! At that point we have, with his help, solved the original primal problem.

Example. To illustrate the primal-dual method and indicate how it can be implemented through use of the tableau format consider the following problem:

$$\begin{array}{ll} \text{minimize} & 2x_1 + x_2 + 4x_3 \\ \text{subject to} & x_1 + x_2 + 2x_3 = 3 \\ & 2x_1 + x_2 + 3x_3 = 5 \\ & x_1 \geqslant 0, \quad x_2 \geqslant 0, \quad x_3 \geqslant 0. \end{array}$$

Because all of the coefficients in the objective function are nonnegative, $\mathbf{y} = (0,\ 0)$ is a feasible vector for the dual. We lay out the simplex tableau shown below

	$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\cdot$	$\cdot$	$\mathbf{b}$
	1	1	2	1	0	3
	2	1	3	0	1	5
	-3	-2	-5	0	0	-8
$c_i - \mathbf{y}^T\mathbf{a}_i \rightarrow$	2	1	4	$\cdot$	$\cdot$	$\cdot$

First tableau

To form this tableau we have adjoined artificial variables in the usual manner. The third row gives the relative cost coefficients of the associated primal problem—the same as the row that would be used in a phase I procedure. In the fourth row are listed the $c_i - \mathbf{y}^T\mathbf{a}_i$'s for the current $\mathbf{y}$. The allowable columns in the associated restricted primal are determined by the zeros in this last row.

Since there are no zeros in the last row, no progress can be made in the associated restricted primal and hence the original solution $x_1 = x_2 = x_3 = 0,\ y_1 = 3,\ y_2 = 5$ is optimal for this $\mathbf{y}$. The solution $\mathbf{u}_0$ to the associated restricted dual is $\mathbf{u}_0 = (1, 1)$, and the numbers $-\mathbf{u}_0^T\mathbf{a}_i,\ i = 1, 2, 3$ are equal to the first three elements in the third row. Thus, we compute the three ratios $\frac{2}{3}, \frac{1}{2}, \frac{4}{5}$ from which we find $\varepsilon_0 = \frac{1}{2}$. The new values for the fourth row are now found by adding ε_0 times the (first three) elements of the third row to the fourth row.

$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	·	·	$\mathbf{b}$
1	①	2	1	0	3
2	1	3	0	1	5
−3	−2	−5	0	0	−8
1/2	0	3/2	·	·	·

Second tableau

Minimizing the new associated restricted primal by pivoting as indicated we obtain

$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	·	·	$\mathbf{b}$
1	1	2	1	0	3
1	0	1	−1	1	2
−1	0	−1	2	0	−2
−1/2	0	3/2	·	·	·

Now we again calculate the ratios $\frac{1}{2}, \frac{3}{2}$ obtaining $\varepsilon_0 = \frac{1}{2}$, and add this multiple of the third row to the fourth row to obtain the next tableau.

$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	·	·	$\mathbf{b}$
1	1	2	1	0	3
①	0	1	−1	1	2
−1	0	−1	2	0	−2
0	0	1	·	·	·

Third tableau

optimizing the new restricted primal we obtain the tableau:

$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	·	·	$\mathbf{b}$
0	1	1	2	−1	1
1	0	1	−1	1	2
0	0	0	1	1	0
0	0	1	·	·	·

Final tableau

Having obtained feasibility in the primal, we conclude that the solution is also optimal: $x_1 = 2,\ x_2 = 1,\ x_3 = 0$.

4.8 SUMMARY

There is a corresponding dual linear program associated with every (primal) linear program. Both programs share the same underlying cost and constraint coefficients. We have demonstrated rich theorems to relate the pair. The variables of the dual problem can be interpreted as prices associated with the constraints of the original (primal) problem, and through this association it is possible to give an economically meaningful characterization to the dual whenever there is such a characterization for the primal.

Mathematically, the pair also establish an optimality certificate to each other: one cannot claim an optimal objective value unless you find an solution for the dual to achieve the same value of the dual objective. This also leads to the set of optimality conditions, including the complementarity conditions, that we would see many times in the rest of the book.

4.9 EXERCISES

1. Verify in detail that the dual of a linear program is the original problem.
2. Show that if a linear inequality in a linear program is changed to equality, the corresponding dual variable becomes free.
3. Find the dual of

$$\begin{array}{rl} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geqslant \mathbf{a} \\ \text{where} & \mathbf{a} \geqslant \mathbf{0}. \end{array}$$

4. Show that in the transportation problem the linear equality constraints are not linearly independent, and that in an optimal solution to the dual problem the dual variables are not unique. Generalize this observation to any linear program having redundant equality constraints.
5. Construct an example of a primal problem that has no feasible solutions and whose corresponding dual also has no feasible solutions.
6. Let $\mathbf{A}$ be an $m \times n$ matrix and $\mathbf{b}$ be an n-vector. Prove that $\mathbf{A}\mathbf{x} \leqslant \mathbf{0}$ implies $\mathbf{c}^T\mathbf{x} \leqslant \mathbf{0}$ if and only if $\mathbf{c}^T = \mathbf{y}^T\mathbf{A}$ for some $\mathbf{y} \geqslant \mathbf{0}$. Give a geometric interpretation of the result.
7. There is in general a strong connection between the theories of optimization and free competition, which is illustrated by an idealized model of activity location. Suppose there are n economic activities (various factories, homes, stores, etc.) that are to be individually located on n distinct parcels of land. If activity i is located on parcel j that activity can yield s_{ij} units (dollars) of value.
   If the assignment of activities to land parcels is made by a central authority, it might be made in such a way as to maximize the total value generated. In other words, the assignment would be made so as to maximize $\sum_i \sum_j s_{ij}x_{ij}$ where

$$x_{ij} = \begin{cases} 1 & \text{if activity } i \text{ is assigned to parcel } j \\ 0 & \text{otherwise} \end{cases}$$

   More explicitly this approach leads to the optimization problem