DECOMPOSITION PRINCIPLE FOR LINEAR PROGRAMS†

George B. Dantzig and Philip Wolfe

The Rand Corporation, Santa Monica, California

(Received November 24, 1959)

A technique is presented for the decomposition of a linear program that permits the problem to be solved by alternate solutions of linear sub-programs representing its several parts and a coordinating program that is obtained from the parts by linear transformations. The coordinating program generates at each cycle new objective forms for each part, and each part generates in turn (from its optimal basic feasible solutions) new activities (columns) for the interconnecting program. Viewed as an instance of a 'generalized programming problem' whose columns are drawn freely from given convex sets, such a problem can be studied by an appropriate generalization of the duality theorem for linear programming, which permits a sharp distinction to be made between those constraints that pertain only to a part of the problem and those that connect its parts. This leads to a generalization of the Simplex Algorithm, for which the decomposition procedure becomes a special case. Besides holding promise for the efficient computation of large-scale systems, the principle yields a certain rationale for the 'decentralized decision process' in the theory of the firm. Formally the prices generated by the coordinating program cause the manager of each part to look for a 'pure' sub-program analogue of pure strategy in game theory, which he proposes to the coordinator as best he can do. The coordinator finds the optimum 'mix' of pure sub-programs (using new proposals and earlier ones) consistent with over-all demands and supply, and thereby generates new prices that again generates new proposals by each of the parts, etc. The iterative process is finite.

A VECTOR $\boldsymbol{P}$ interconnecting two parts of a program is viewed as obtained by linear transformations from linear sub-programs L and L' (more generally as drawn from convexes) defining the parts. $\boldsymbol{P}$ is represented as a convex combination of a finite set of possible $\boldsymbol{P}$ from each part and equated to a similar representation for the other—the selected vectors correspond to extreme points and certain homogeneous solutions of the sub-programs. Starting with some admissible m-component vector $\boldsymbol{P}=\boldsymbol{P}^0$, the total vectors (points) used in the two representations can be reduced

† The material of this paper appeared originally under the same title as Rand Corporation P-1544, November 10, 1958, and was presented at the Rand Symposium on Mathematical Programming in March, 1959. It was also presented before the American Mathematical Society, January 20, 1959, in Philadelphia, Pennsylvania (see *Am. Math. Society Notices*, Abstract No. 553-559, Vol. 5, No. 7, Issue No. 35, December, 1958, p. 811).

to $m+2$. This forms a basis whose Lagrange (simplex) multipliers π relative to a form being extremized are used to test optimality or for generating a better interconnecting vector. This is done by solving two independent linear programs for points in L and L' that minimize two linear forms dependent on π.

The entire procedure may be simulated as a *decentralized decision* process. Each independent part initially offers a possible bill of goods (a vector of outputs and supporting inputs including outside costs) to a central coordinating agency. As a set these are mutually feasible with each other and the given resources and demands from outside the system. The coordinator works out a system of 'prices' for paying for each component of the vector plus a special subsidy for each part that just balances the cost. A bonus or some other form of award is then offered the management of each part if he can offer, based on these prices, a new feasible program for his part with lower cost *without regard to whether it is feasible for any other part.* The coordinator, however, combines these new offers with the set of earlier offers so as to preserve mutual feasibility and consistency with exogenous demand and supply and to minimize cost. Using the improved over-all solution he generates a revised set of prices, subsidies, and new offers. *The essential idea is that old offers are never forgotten by the central agency* (unless using 'current' prices they are unprofitable); the former are mixed with the new offers to form new prices.

Computationally if $\boldsymbol{P}$ is an m-component vector and L is defined by k equations in n nonnegative variables and L' by k' equations in n' nonnegative variables, then each major simplex cycle consists in solving two $k\times n$ and $k'\times n'$ auxiliary linear programs after $m(m+n+n')$ multiplications to set up π and adjust the solution. The iterative procedure is finite. The principle is applied to decompose typical structures into several parts.

Credit is due to FORD AND FULKERSON for their proposal for solving multicommodity network problems as it served to inspire the present development.†

THE GENERAL PRINCIPLE

SUPPOSE we have a linear program expressed in vector notation in the form

$$\boldsymbol{P}_0 x_0+\boldsymbol{P}_1 x_1+\cdots+\boldsymbol{P}_n x_n=\boldsymbol{Q}, \qquad (1)$$

where $\boldsymbol{P}_j$, $\boldsymbol{Q}$ are given vectors and the problem is to choose Max x_0 and $x_j\geqq 0$ for $j=1, 2, \cdots, n$. Ordinarily the objective form of a linear-pro-

† L. R. FORD, JR., AND D. R. FULKERSON, "A Suggested Computation for Maximal Multicommodity Network Flow," *Management Sci.* 5, 97 (1958).

gramming problem is written as

$$c_1 x_1 + c_2 x_2 + \cdots + c_n x_n = z(\text{Min}), \tag{2}$$

which may be rewritten

$$x_0 + c_1 x_1 + c_2 x_2 + \cdots + c_n x_n = 0, \tag{3}$$

where x_0 is to be maximized. In this case $\boldsymbol{P}_0$ is a unit vector.

Let us now consider the general problem of solving the linear-programming problem of finding vectors $\boldsymbol{X} \geqq 0$, $\boldsymbol{Y} \geqq 0$ and Max x_0 satisfying

$$\boldsymbol{P}_0 x_0 + \bar{\boldsymbol{A}}_1 \boldsymbol{X} + \bar{\boldsymbol{A}}_2 \boldsymbol{Y} = \bar{\boldsymbol{b}} \tag{4}$$

subject to linear programs L and L' defined by

$$\begin{aligned} L: &\quad \boldsymbol{A}_1 \boldsymbol{X} = \boldsymbol{b}_1, &\quad (X \geqq 0) \\ L': &\quad \boldsymbol{A}_2 \boldsymbol{Y} = \boldsymbol{b}_2, &\quad (Y \geqq 0) \end{aligned} \tag{5}$$

where $\boldsymbol{A}_i, \bar{\boldsymbol{A}}_i$ are matrices, $\boldsymbol{P}_0$, $\bar{\boldsymbol{b}}$, $\boldsymbol{b}_i$ are vectors and $\bar{\boldsymbol{b}}$ has m components.† We shall show later how multistage models such as dynamic models with discrete time periods, angular systems multistage and block triangular models can be readily decomposed into several parts each exhibiting the above structure.

Let us define vectors $\boldsymbol{S}$ and $\boldsymbol{T}$ corresponding to $\boldsymbol{X}$ and $\boldsymbol{Y}$ by the linear transformations

$$\boldsymbol{S} = \bar{\boldsymbol{A}}_1 \boldsymbol{X} \quad \textit{for arbitrary} \quad \boldsymbol{X} \geqq 0 \quad \textit{and} \quad \boldsymbol{A}_1 \boldsymbol{X} = \boldsymbol{b}_1; \tag{6}$$

$$\boldsymbol{T} = \bar{\boldsymbol{A}}_2 \boldsymbol{Y} \quad \textit{for arbitrary} \quad \boldsymbol{Y} \geqq 0 \quad \textit{and} \quad \boldsymbol{A}_2 \boldsymbol{Y} = \boldsymbol{b}_2. \tag{7}$$

Starting Assumption I: Let us suppost first that a starting feasible solution is at hand: $\boldsymbol{X} = \boldsymbol{X}_0$, $\boldsymbol{Y} = \boldsymbol{Y}_0$, $x_0 = x_0^0$; that is to say one which satisfies all the constraints (4) and (5) except x_0^0 may not be maximal. Such a solution could have been generated for example, by the Phase I procedure of linear programming using artificial variables. If so the methods we are about to discuss would be applied to Phase I first.

Let $\boldsymbol{S}^0 = \bar{\boldsymbol{A}}_1 \boldsymbol{X}_0$ and $\boldsymbol{T}^0 = \bar{\boldsymbol{A}}_2 \boldsymbol{Y}_0$ so that

$$\boldsymbol{P}_0 x_0^0 + \boldsymbol{S}_0 + \boldsymbol{T}_0 = \boldsymbol{b}. \tag{8}$$

Starting Assumption II: We shall suppose further that $\boldsymbol{S}_0$ has been represented with weights $\lambda_i = \bar{\lambda}_i$ in the form

$$\begin{aligned} \boldsymbol{S}_0 &= \boldsymbol{S}_1 \bar{\lambda}_1 + \boldsymbol{S}_2 \bar{\lambda}_2 + \cdots + \boldsymbol{S}_k \bar{\lambda}_k, \qquad (\bar{\lambda}_i \geqq 0) \\ 1 &= \delta_1 \bar{\lambda}_1 + \delta_2 \bar{\lambda}_2 + \cdots + \delta_k \bar{\lambda}_k, \end{aligned} \tag{9}$$

† We assume for convenience that $\boldsymbol{b}_1 \neq 0$, $\boldsymbol{b}_2 \neq 0$. If $\boldsymbol{b}_i = 0$ the restriction requiring convex combinations of solutions of l and l' should be dropped in the development that follows [second equation of (9) or (12)].

where each
$$S_i = \bar{A}_1 X_i, \qquad (X_i \geqq 0) \quad (10)$$
where X_i is either an extreme point or homogeneous solution [*see* (11)] of L, and δ_i is defined by
$$\begin{Bmatrix} \delta_i = 1 \\ \delta_i = 0 \end{Bmatrix} \text{ if } X_i \geqq 0 \text{ is } \begin{Bmatrix} \text{extreme point} \\ \text{homogeneous} \end{Bmatrix} \text{ solution of } \begin{Bmatrix} A_1 X_i = b_i \\ A_1 X_i = 0 \end{Bmatrix}. \quad (11)$$
Similarly T_0 has been represented with weights $\mu_i = \bar{\mu}_i$ in the form
$$\begin{aligned} T_0 &= T_1 \bar{\mu}_1 + T_2 \bar{\mu}_2 + \cdots + T_l \bar{\mu}_l, \qquad (\bar{\mu}_i \geqq 0) \\ 1 &= \delta_1' \bar{\mu}_1 + \delta_2' \bar{\mu}_2 + \cdots + \delta_l' \bar{\mu}_l, \end{aligned} \quad (12)$$
where T_i corresponds to some solution Y_i of L'.
$$T_i = \bar{A}_2 Y_i, \qquad (Y_i \geqq 0) \quad (13)$$
such that
$$\begin{Bmatrix} \delta_i = 1 \\ \delta_i = 0 \end{Bmatrix} \text{ if } Y_i \text{ is an } \begin{Bmatrix} \text{extreme point} \\ \text{homogeneous} \end{Bmatrix} \text{ solution of } \begin{Bmatrix} A_i X_i = b_i \\ A_i X_i = 0 \end{Bmatrix}. \quad (14)$$
We now consider in place of (4) the obviously equivalent '*interconnecting linear program*' of determining weights $\lambda_i \geqq 0$, $\mu_i \geqq 0$, and Max x_0 satisfying
$$P_0 x_0 + S \lambda_0 + S_1 \lambda_1 + \cdots + S_k \lambda_k + T \mu_0 + T_1 \mu_1 + \cdots + T_l \mu_l = b, \quad (15)$$
$$\delta \lambda_0 + \delta_1 \lambda_1 + \cdots + \delta_k \lambda_k = 1, \quad (16)$$
$$\delta' \mu_0 + \delta_1' \mu_1 + \cdots + \delta_l' \mu_l = 1, \quad (17)$$
where S, T correspond to *arbitrary* extreme points $(\delta, \delta' = 1)$ or homogeneous solutions $(\delta, \delta' = 0)$ of (5). A particular feasible solution can be obtained by setting $\lambda = \mu = 0$ and $\lambda_i = \bar{\lambda}_i$, $\mu_j = \bar{\mu}_j$.

Starting Assumption III: Finally we suppose that the columns T_i, S_i, P_0 with added components for equations (16) and (17) form a *basis* B i.e., a $(m+2) \times (m+2)$ nonsingular matrix as in (18). Let the simplex multipliers associated with the rows of B be $(\pi; -s, -t)$ where $-s$, $-t$ are multipliers for the last two rows:
$$B = \begin{Vmatrix} P_0; & S_1 & \cdots & S_k; & T_1 & \cdots & T_l \\ 0 & \delta_1 & \cdots & \delta_k & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \delta_1' & \cdots & \delta_l' \end{Vmatrix}; \qquad \begin{Vmatrix} \pi \\ -s \\ -t \end{Vmatrix}. \quad (18)$$
Since we are maximizing x_0 the simplex multipliers by definition satisfy
$$\pi P_0 = 1, \qquad \pi S_i - \delta_i s = 0, \qquad \pi T_i - \delta_i' t = 0. \quad (19)$$
To test optimality of the proposed solution we now form the row vectors
$$\pi \bar{A}_1 = \gamma_1; \qquad \pi \bar{A}_2 = \gamma_2. \quad (20)$$

THEOREM 1: *The solution* $(\boldsymbol{X}_0, \boldsymbol{Y}_0, x_0^0)$ *is maximal if for all* $\boldsymbol{X} \geqq 0$, $\boldsymbol{Y} \geqq 0$, z_1, z_2 *satisfying*

$$A_1 \boldsymbol{X} = \boldsymbol{b}_1, \qquad \gamma_1 \boldsymbol{X} = z_1, \qquad (\boldsymbol{X} \geqq 0) \quad (21)$$

$$A_2 \boldsymbol{Y} = \boldsymbol{b}_2, \qquad \gamma_2 \boldsymbol{Y} = z_2, \qquad (\boldsymbol{Y} \geqq 0) \quad (22)$$

it is true that

$$\text{Min } z_1 = s, \qquad \text{Min } z_2 = t. \qquad (23)$$

Proof. Multiplication of (9) and (12) on the left by $\boldsymbol{\pi}$ yields, by (19) and (20),

$$s = \gamma_1 \boldsymbol{X}_0, \qquad t = \gamma_2 \boldsymbol{Y}_0. \qquad (24)$$

On the other hand multiplication of (4) on the left by $\boldsymbol{\pi}$ yields, by (20), (21), and (22)

$$x_0 + \gamma_1 \boldsymbol{X} + \gamma_2 \boldsymbol{Y} = \boldsymbol{\pi}\bar{\boldsymbol{b}}, \qquad (25)$$

or

$$x_0 + z_1 + z_2 = \boldsymbol{\pi}\bar{\boldsymbol{b}}. \qquad (26)$$

In particular for $\boldsymbol{X} = \boldsymbol{X}_0$, $\boldsymbol{Y} = \boldsymbol{Y}_0$, $x = x_0^0$ by (24):

$$x_0^0 + s + t = \boldsymbol{\pi}\bar{\boldsymbol{b}}. \qquad (27)$$

Subtracting (26) from (27),

$$x_0^0 - x_0 = (z_1 + z_2) - (s + t). \qquad (28)$$

The right-hand side is always nonnegative by (23); hence $x_0^0 \geqq x_0$ so that the solution is maximal.

If the solution $(\boldsymbol{X}_0, \boldsymbol{Y}_0, x_0^0)$ fails to satisfy the test for optimality (23), say Min $z_1 < s$, then there exists *either* an *extreme point solution* $\boldsymbol{X} = \boldsymbol{X}^*$ to (21) such that

$$\text{Min } z = \gamma_1 \boldsymbol{X}^* < s, \qquad (29)$$

or there exists a *homogeneous solution* $\boldsymbol{X}^*$ satisfying

$$A_1 \boldsymbol{X}^* = 0 \quad \text{and} \quad \gamma_1 \boldsymbol{X}^* < 0. \qquad (30)$$

The latter possibility warrants some discussion. It may happen that the convex $A_1 \boldsymbol{X} = b_1$, $\boldsymbol{X} \geqq 0$ is *unbounded* and that there is no lower bound for z_1. In this case *using the simplex method* a homogeneous solution, $A_1 \boldsymbol{X} = 0$, $\boldsymbol{X} \geqq 0$, will be obtained with the above property. It is not difficult to show that the set of possible homogeneous solutions generated in this manner, excluding multiples, is finite.

If we now define

$$\bar{A}_1 \boldsymbol{X}^* = \boldsymbol{S}^*, \qquad (31)$$

and $\delta^* = 1,0$ depending on whether $\boldsymbol{X}^*$ is an extreme or homogeneous solution and substitute $\boldsymbol{S} = \boldsymbol{S}^*$, $\delta = \delta^*$, the interconnecting program has one more variable than the number of equations. The basic solution $\lambda = 0$, $\lambda_i = \bar{\lambda}_i$, $\mu_i = \bar{\mu}_i$ does not satisfy the test for optimality when the simplex multipliers $(\pi,-s,-t)$ are multiplied scalarly by the column of coefficients of λ_0; indeed by (31), then (20), (29), and (30), this yields

$$\pi S^* - \delta^* s = \pi A_1 X^* - \delta^* s = \gamma_1 X^* - \delta^* s < 0. \qquad (32)$$

Hence, assuming nondegeneracy by perturbing $\bar{b}$ if necessary, an improved solution can be obtained by introducing S^* from the auxiliary program into the basis of the interconnecting problem and dropping some other S_i or T_i. Multipliers with respect to the new basis can then be determined by (19); these will generate new values γ_1 and γ_2 by (20) and new definitions for z_1 and z_2; finally the two auxiliary linear programs (21) and (22) must be resolved to determine new points where z_1 and z_2 are minimized. The algorithm is finite because of the finiteness of the set of possible bases derived from extreme points and homogeneous solutions generated by the simplex method when applied to the auxiliary problems, and because no basis of the interconnecting program can repeat itself since x_0 is monotonically increasing (when Q is perturbed to avoid degeneracy).

If a record has been maintained of the solution vectors X_i and Y_i to the auxiliary problems, corresponding to vectors S_i and T_i in the final basis, then an optimal solution $\hat{X}$ and $\hat{Y}$ can be obtained by means of the relations

$$\hat{X} = \sum \lambda_i X_i, \qquad \hat{Y} = \sum \mu_i Y_i. \qquad (33)$$

From iteration to iteration it is not necessary, however, to maintain a record of the solutions X_i and Y_i to the auxiliary problems, but only a record of the S_i, T_i in the current basis and whether or not they correspond to extreme or homogeneous solutions. When an optimum has been achieved, the final selection of S_i and λ_i and T_i and μ_i are used to compute

$$\hat{S} = \sum \lambda_i S_i, \qquad \hat{T} = \sum \mu_i T_i, \qquad (34)$$

and to determine an optimum solution for $X = \hat{X}$ and $Y = \hat{Y}$ by solving the auxiliary problems

$$\begin{array}{ll} A_1 X = b_1, & A_2 Y = b_2, \\ \bar{A}_1 X = \hat{S}, & A_2 Y = \hat{T}, \quad (X \geqq 0,\ Y \geqq 0) \end{array} \qquad (35)$$

where any feasible solution is an optimum solution. This is greatly simplified by observing that the choice of components of X and Y, which are basic and nonbasic, are further restricted to the set which minimize z_1 and z_2 respectively.

A GENERALIZED PROGRAMMING PROBLEM

CONSIDER a generalized linear-programming problem in which each P_j for $j \neq 0$ may be freely chosen to be any $P_j \in C_j$ where C_j is a convex set defined by linear inequalities. By a simple extension, Q and P_0 could also be replaced by any Q and P_0 in convex sets; however, for our purposes these are held constant. We shall assume C_j is a convex polyhedron so that a general P_j can be represented by a convex linear combination of a finite set of extreme points of C_j and a nonnegative linear combination of homogeneous solutions (in case C_j is unbounded). Two easy theorems are:

THEOREM 2: *A solution* (x_j^*, P_j^*) *for* $j=0, 1, 2, \cdots, n$, *is optimal if there exists a* π, *such that* $\pi P_0 = 1$, $\pi P_j \geqq 0$ *for all* $P_j \in C_j$ *and* $\pi P_j^* = 0$ *for all* $x_j^* > 0$ $(j \neq 0)$.

THEOREM 3: *Only a finite number of iterations of the simple algorithm is required if each basic feasible solution is improved by introducing into the basis either an extreme point* $P_{j^*} \in C_{j^*}$ *chosen so that*

$$\pi P_{j^*} = \text{Min}_{P_j \in C_j}\ \pi P_j < 0, \quad (j = 1, 2, \cdots, n) \quad (36)$$

where π *are the simplex multipliers of the basis or any homogeneous solution* $\bar{P}_{j^*}$ *from the finite set such that* $\pi P_{j^*} < 0$.

The reader will recognize that our Theorem 1 is a special case of Theorem 2. Also that we followed the procedure implicit in Theorem 3 except we introduced vectors into the basis of the interconnecting problem *other than* extreme points and homogeneous rays of the convex (for the extreme points and homogeneous solutions of the auxiliary problems need not transform into similar solutions of C and C' under the mappings $\bar{A}_1 X$ and $\bar{A}_2 Y$).

DECOMPOSING SPECIAL SYSTEMS

Angular Systems

This will be a direct application to an important class of linear programs. It will serve as a review of the decomposition principle given earlier in which no new concepts will be introduced. Consider for convenience a special case of an 'angular system with three block diagonal terms'

$$\begin{aligned} A_1 X_1 \qquad\qquad\qquad\qquad &= b_1, \\ +A_2 X_2 \qquad\qquad &= b_2, \\ +A_3 X_3 \quad &= b_3, \\ +\bar{A}_1 X_1 + \bar{A}_2 X_2 + \bar{A}_3 X_3 + P_0 x_0 &= \bar{b}, \end{aligned} \qquad (X_t \geqq 0) \quad (37)$$

where A_t is an $m_t \times n_t$ matrix, X_t a column vector of n_t components and b_t, $\bar{b}$ column vectors with m_t and m components.

We consider the linear programming problem in the general form

$$P_1 + P_2 + P_3 + P_0\, x_0 = b, \qquad (\lambda_t = 1) \quad (38)$$

where $\boldsymbol{P}_t$ is defined by the linear transformation

$$\bar{\boldsymbol{A}}_t \boldsymbol{X}_t = \boldsymbol{P}_t, \qquad (39)$$

and $\boldsymbol{X}_t$ satisfies the sub-program L_t:

$$\boldsymbol{A} \boldsymbol{X}_t = \boldsymbol{b}_t. \qquad (\boldsymbol{X}_t \geqq 0) \quad (40)$$

It will be convenient to assume each L_t is bounded; adjustments for the unbounded case can be made by forming nonnegative linear combinations instead of convex combinations for points corresponding to homogeneous solutions.

We shall require however, that each $\boldsymbol{P}_t$ be represented as a convex combination (positive weights that sum to unity) of points $\boldsymbol{P}_{ti}$ of $\boldsymbol{P}_{ti} = \bar{\boldsymbol{A}}\boldsymbol{X}_{ti}$ where $\boldsymbol{X}_t = \boldsymbol{X}_{ti} \geqq 0$ are extreme point solutions of L_t.

Let us consider in place of (38), the linear-programming problem of finding $\lambda_i \geqq 0$, $\lambda_{ij} \geqq 0$, Max x_0 satisfying

$$\begin{gathered} \boldsymbol{P}_0 x_0 + \textstyle\sum_{k=1}^{k=k_1} \boldsymbol{P}_{1k} \lambda_{1k} + \boldsymbol{P}_1 \lambda_1 + \sum_{k=1}^{k=k_2} \boldsymbol{P}_{2k} \lambda_{2k} + \boldsymbol{P}_2 \lambda_2 \\ + \textstyle\sum_{k=1}^{k=k_3} \boldsymbol{P}_{3k} \lambda_{3k} + \boldsymbol{P}_3 \lambda_3 + \boldsymbol{P}_0 x_0 = \bar{\boldsymbol{b}}, \qquad (41) \\ \textstyle\sum \lambda_{1k} + \lambda_1 = 1, \qquad \sum \lambda_{2k} + \lambda_2 = 1, \qquad \sum \lambda_{3k} + \lambda_3 = 1. \end{gathered}$$

These two problems are equivalent because the convexity of L_t implies that a convex combination of vectors $\boldsymbol{P}_{ti}$ and $\boldsymbol{P}_t$ corresponds to a feasible $\boldsymbol{X}$ in L_t. Suppose at some stage of the simplex process several points $\boldsymbol{P}_{t1}, \boldsymbol{P}_{t2}, \cdots, \boldsymbol{P}_{tk_t}$ have been introduced into the basis so that a feasible solution $\lambda_{ik} \geqq 0$ can be obtained by setting $\lambda_1 = \lambda_2 = \lambda_3 = 0$ subject to $\sum \lambda_{ik} = 1$ for $i = 1, 2, 3$.

We are now interested in testing whether or not the basic solution maximizes x_0. To find a better solution we determine the simplex multipliers associated with the basis

$$\boldsymbol{B} = \left\| \begin{matrix} \boldsymbol{P}_{11} & \boldsymbol{P}_{12} & \cdots & \boldsymbol{P}_{1k_1}; & \boldsymbol{P}_{21} & \boldsymbol{P}_{22} & \cdots & \boldsymbol{P}_{2k_2}; & \boldsymbol{P}_{31} & \boldsymbol{P}_{32} & \cdots & \boldsymbol{P}_{3k_3}; & \boldsymbol{P}_0 \\ 1 & 1 & \cdots & 1 & & & & & & & & & \\ & & & & 1 & 1 & \cdots & 1 & & & & & \\ & & & & & & & & 1 & 1 & \cdots & 1 & \end{matrix} \right\| \qquad (42)$$

by the relations

$$\pi \boldsymbol{P}_0 = 1, \qquad \pi \boldsymbol{P}_{ij} + p_j = 0, \qquad (j = 1, 2, \cdots, k_i; i = 1, 2, 3) \quad (43)$$

where $\pi = \|\bar{\pi}; p_1, p_2, p_3\|$ and p_1, p_2, p_3 are the multipliers associated with the last three equations. To test optimality or to find a better vector to introduce into the basis, we solve for each t the auxiliary problems

$$\begin{aligned} \boldsymbol{A}_t \boldsymbol{X}_t &= b_t, & (\boldsymbol{X}_t \geqq 0) \\ (\pi \bar{\boldsymbol{A}}_t) \boldsymbol{X}_t + p_t &= z_t (\text{Min}), & \end{aligned} \qquad (44)$$

where $\pi \bar{A}_t$ is a constant row vector. If Min $z_t \geqq 0$ for all t, the solution is optimal. If not, $P_t = P_t^*$ is introduced into the basis corresponding to that t, such that $X_t = X_t^*$ yields

$$\text{Min}_t \text{ Min}_{X_t} z_t. \tag{45}$$

To summarize, the linear-programming problem consists of iterative cycling between two parts:

PART I: Determine new multipliers by a simple change of basis for the set of $m+3$ equations (40) and (41).

PART II: Solve several auxiliary $m_j \times n_j$ linear programs (44), for testing optimality of the solution and the introduction of new vectors $\bar{A}_t X_t^*$ into the basis of Part I by (45).

Multi-Stage Systems

Consider a system with structure

$$\begin{aligned} A_1 X_1 \qquad\qquad\qquad\qquad &= e_1, \\ +\bar{A}_1 X_1 + A_2 X_2 \qquad\qquad &= e_2, \\ +\bar{A}_2 X_2 + A_3 X_3 \qquad &= e_3, \qquad (X_j \geqq 0) \quad (46) \\ +\bar{A}_3 X_3 + A_4 X_4 &= e_4, \\ c_4 X_4 &= z_4(\text{Min}), \end{aligned}$$

where X_t are vectors, A_t matrices, e_t and c_4 vectors.

We replace this formally by

$$\begin{aligned} +P_3 + A_4 X_4 &= e_4, \\ c_4 X_4 &= z_4(\text{Min}), \end{aligned} \qquad (X_4 \geqq 0) \quad (47)$$

where $P_3 = \bar{A}_3 X_3$ is any vector corresponding to (X_1, X_2, X_3) satisfying the first three relations of (46) and $X_j \geqq 0$. We shall require again however, that P_3 be represented by a convex linear combination with weights $\lambda_{3i} = \bar{\lambda}_{31} P_{3i}$ of C_3 derived from extreme points. (This assumes boundedness; if not appropriate changes discussed earlier should be made for homogeneous solutions.)

$$\sum \bar{\lambda}_{3i} = 1. \qquad (\bar{\lambda}_{3i} \geqq 0) \quad (48)$$

The interconnecting 'stage 4' linear program is to determine λ_3, λ_{3i}, X_4 and Min z_4 satisfying

$$\begin{aligned} \sum \lambda_{3i} P_{3i} + \lambda_3 P_3 + A_4 X_4 &= e_4, \\ \sum \lambda_{3i} \quad + \lambda_3 \qquad\qquad &= 1, \qquad (\lambda_{3i} \geqq 0,\ \lambda_3 \geqq 0,\ X_4 \geqq 0) \quad (49) \\ c_4 X_4 &= z_4(\text{Min}). \end{aligned}$$

We assume a basic solution is at hand using columns $\boldsymbol{P}_{3i}$ and columns from A_4. Let $\pi=(\bar{\pi}, p_4)$ be the simplex multipliers associated with the basis where p_4 is the multiplier associated with the last equation. The solution is optimal if

$$\pi \boldsymbol{P}_3+p_4 \geqq 0, \tag{50}$$

for all $\boldsymbol{P}_3$ in C_3. To ascertain this we consider the subproblem to find $\boldsymbol{X}_j \geqq 0$, Min z_3 satisfying

$$\begin{array}{llll} A_1 \boldsymbol{X}_1 & & & \\ \bar{A}_1 \boldsymbol{X}_1 + A_2 \boldsymbol{X}_2 & & & \\ & \bar{A}_2 \boldsymbol{X}_2 + A_3 \boldsymbol{X}_3 & & \\ & & \boldsymbol{c}_3 \boldsymbol{X}_3 + p_4 = z_3(\text{Min}), & \end{array} \tag{51}$$

where

$$\boldsymbol{c}_3 = \pi \bar{A}_3. \tag{52}$$

If Min $z_3 \geqq 0$ the solution is optimal. With respect to this sub-problem, in an analogous manner we set up a 'stage 3' linear-programming problem

$$\begin{aligned} \sum \lambda_{2i} \boldsymbol{P}_{2i} + \lambda_2 \boldsymbol{P}_2 + A_3 \boldsymbol{X}_3 &= \boldsymbol{e}_3, \\ \sum \lambda_{2i} \quad + \lambda_2 \qquad &= 1, \\ c_3 X_3 + p_4 &= z_3(\text{Min}). \end{aligned} \tag{53}$$

This in turn induces a 'stage 2' and a 'stage 1' linear-programming problem.

This nested set of programming problems is not as ideal as may appear at first because there could be a great deal of jockeying up and down the various stages seeking improved solutions. One procedure could be

1. Optimize *stage 4* over the columns $\boldsymbol{P}_{3i}$ and A_4 with $\lambda_3=0$ and use simplex multipliers (π_4, ρ_4) to determine $\boldsymbol{c}_3$.

2. Use $\boldsymbol{c}_3$ to optimize *stage 3* over the columns $\boldsymbol{P}_{2i}$ with $\lambda_2=0$ and use simplex multipliers $(\pi_3, \boldsymbol{P}_3)$ to determine $\boldsymbol{c}_2$.

3. Use $\boldsymbol{c}_2$ to optimize *stage 2* over the columns $\boldsymbol{P}_{1i}$ with $\lambda_1=0$ and use simplex multipliers (π_2, p_2) to determine $\boldsymbol{c}_1$.

4. Use $\boldsymbol{c}_1$ to optimize *stage 1* over the columns of A_1 to determine $\boldsymbol{X}_1=\boldsymbol{X}_1^*$.

5. Substitute $\boldsymbol{P}_1=\bar{A}_1\boldsymbol{X}_1^*$ into *stage 2* and continue optimization of *stage 2* allowing λ_1 to vary to determine $\boldsymbol{X}_2=\boldsymbol{X}_2^*$.

6. Substitute $\boldsymbol{P}_2=\bar{A}_2\boldsymbol{X}_2^*$ into *stage 3* and continue optimization of *stage 3* allowing λ_2 to vary to determine $\boldsymbol{X}_3=\boldsymbol{X}_3^*$.

7. Substitute $\boldsymbol{P}_3=\bar{A}_3\boldsymbol{X}_3^*$ into *stage 4* and continue optimization of *stage 4* allowing λ_3 to vary to determine $\boldsymbol{X}_4=\boldsymbol{X}_4^*$.

8. *Recycle*, treating each $\boldsymbol{P}_i$ generated above as a new $\boldsymbol{P}_{ij}$. All $\boldsymbol{P}_{ij}$ are preserved for use in step **9.**

9. When no new P_{ij} are generated, the fourth stage is optimized, any feasible solution to the third stage is optimal subject to $\bar{A}_3 X_3 = \sum \lambda_{3i}^* P_{3i}$ where $\lambda_{3i} = \lambda_{3i}^*$, $\lambda_3 = 0$ are the optimal fourth-stage weights. This in turn permits optimal solution of the second stage via the best third-stage weights, etc.

Block-Triangular Systems

The treatment of the block-triangular case is similar to the multistage case just considered.