Linear Programming

It is one of several mathematical approaches to problems of optimal choice under constraint.

W. Allen Spivey

The goal of maximizing or minimizing an objective (say profits or costs), where the choice of means is not unrestricted but must be made under one or more constraints, is common to many different problems in the physical and social sciences, in industry, in agriculture, and in national defense. When the objective can be approximated satisfactorily by a linear function and the constraints can be expressed as linear qualities or inequalities, the problem can then be treated mathematically as a problem in linear programming.

Linear programming techniques, largely because of their relative simplicity and flexibility, have found increasingly wide application since systematic development of the theory began in 1948 with the work of George Dantzig and his associates, who were working at that time on programming problems of the U.S. Air Force. By 1955 there had been a remarkable development of the underlying mathematical theory (the work of A. W. Tucker, of Princeton University, in particular, and of a host of other brilliant mathematicians). Moreover, with the parallel development of data-processing and computer machines, it became possible to quickly solve large-scale linear programming problems, so that by 1960—only 12 years after the initial work—linear programming techniques had been successfully applied to the study of such diverse problems as production smoothing, traffic control at toll booths, investment scheduling in an electric-power industry, job assignment, transportation and warehousing of commodities, railway freight movements, blending of aviation gasoline, optimal crop rotation, Air Force contract bidding and the scheduling of aircraft maintenance, plastic limit analysis of structures, chemical composition at equilibrium, and many others.

The author is associate professor of statistics at the School of Business Administration, University of Michigan, Ann Arbor.

Geometrical Analogies

Before considering some linear programming applications in detail, let us look at a simple—and suggestive—geometry. We begin with a two-dimensional plane and a given coordinate system (Fig. 1). Each point in the plane is then represented by an ordered pair of numbers, called its coordinates, and a point can be thought of as an end point of a vector having its initial point at the origin and the given point as end point. By analogy with the corresponding operations on vectors, we define the sum of the points X [or (x_1, x_2)] and Y [or (y_1, y_2)] as

$$X + Y = (x_1 + y_1, x_2 + y_2)$$

and multiplication of a point X by a real number or scalar k as

$$kX = (kx_1, kx_2)$$

Combining these two operations on points, a *linear combination* of points $X_1, X_2, \ldots, X_n$ is defined to be the point X^*, where

$$X^* = k_1 X_1 + k_2 X_2 + \ldots + k_n X_n$$

In particular, if each of the scalars k_i is nonnegative, it is called a nonnegative linear combination, and if, in addition to each $k_i \geqq 0$, we have $k_1 + k_2 + \ldots + k_n = 1$, it is a *convex* linear combination or, more briefly, a convex combination. Some examples will help to make this clear. Suppose $X_1 = (2, 5)$, $X_2 = (5, 3)$; then the set of all nonnegative linear combinations $X^* = k_1 X_1 + k_2 X_2$ consists of all the points in the cone-shaped region in Fig. 2 (left), and the set of all convex combinations of X_1 and X_2 is the set of points on the line segment joining the points X_1 and X_2 (Fig. 2, right). A particularly interesting feature of convex combinations is their "linear" property: the set of all convex combinations of two points fills out the line segment joining the points. To see this still more clearly, let X_3 equal (6, 7) and consider the set of all convex combinations of the three points, X_1, X_2, and X_3. This is the set of points in the triangle in Fig. 3 (all points on the bounding segments and in the interior). If k_1 equals 0 and k_2 and k_3 vary over all permissible values, we obtain all the points on the line segment joining X_2 and X_3. Similar statements can be made concerning the line segments joining X_1 and X_3 and joining X_1 and X_2 for choices of $k_2 = 0$ and $k_3 = 0$, respectively. For each $k_i > 0$, one can verify that the convex combination is a point in the interior of the triangle.

Convex combinations can be used to define a concept of great importance in linear programming, that of a convex set. A set is said to be convex if, given any two points in the set, the line segment joining the points lies entirely inside the set. Since the line segment joining two points consists of the set of all convex combinations of the points, this can be stated as follows: a set is convex if, given any two points X_1, X_2 in the set, the point $k_1 X_1 + k_2 X_2$ is in the set for $k_1 + k_2 = 1$, $k_i \geqq 0$. Illustrations of convex sets are shown in Fig. 4. Points in a convex set which are not convex combinations of two other points in the set are called *extreme* points of the set. In Fig. 4*a* the five corner points are extreme points; in Fig. 4*b* each point on the circumference of the disk is an extreme point, and the set in Fig. 4*c* has no extreme points.

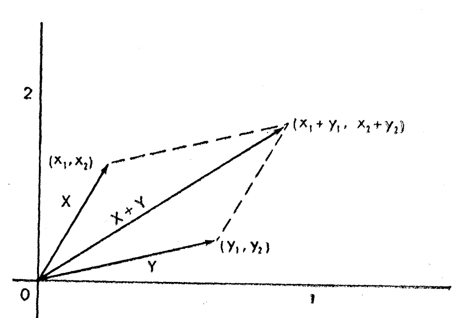

Fig. 1. The sum of two vectors.

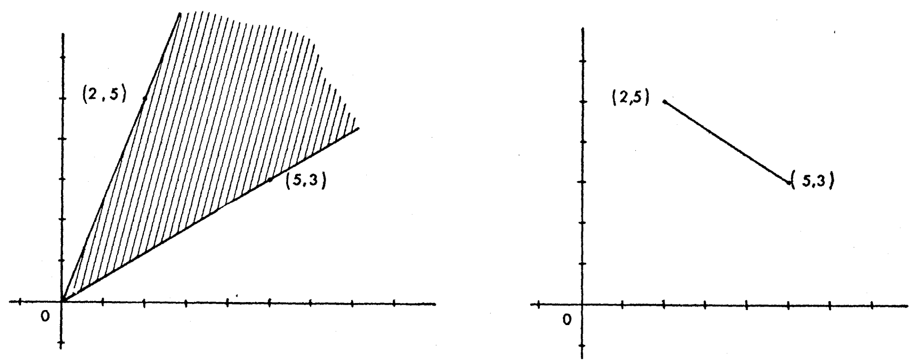

Fig. 2. (Left) Set of all nonnegative linear combinations of two points. (Right) Set of all convex combinations of two points.

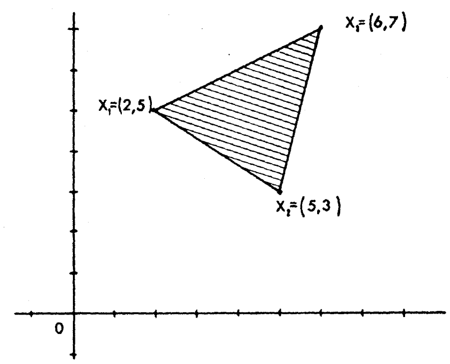

Fig. 3. Set of all convex combinations of three points.

Suppose we now consider the set of points (x_1, x_2) satisfying a linear inequality—say,

$$5x_1 + 6x_2 \leqq 30 \qquad (1)$$

The points whose coordinates satisfy inequality 1 are either on the line which is the graph of the equation $5x_1 + 6x_2 = 30$ (the boundary of the set) or below it. The set defined by inequality 1 is called a half plane and is a special case of a more general concept called a half space. A half space in n-space is the set of all points $(x_1, x_2, \ldots, x_n)$ whose coordinates satisfy a linear inequality of the form

$$a_{11}\, x_1 + \ldots + a_{1n}\, x_n \leqq b \qquad (2)$$

The graph of the equation $a_{11}\, x_1 + \ldots + a_{1n}\, x_n = b$ is called a hyperplane in n-space (generalization of a straight line), and it is the boundary of the half space defined by inequality 2.

Returning to the two-dimensional example, suppose we determine the set of points (x_1, x_2) satisfying inequality 1 and also the inequalities

$$3x_1 + 2x_2 \leqq 12 \qquad (3)$$
$$x_1 \geqq 0 \qquad (4)$$
$$x_2 \geqq 0 \qquad (5)$$

Each of these inequalities defines a half plane, and the points whose coordinates simultaneously satisfy inequality 1 and inequalities 3 to 5 lie in the common part or intersection of the sets. Since it can be shown that a half space is a convex set and that the intersection of any collection of convex sets is a convex set, the intersection shaded in Fig. 5 is a convex set.

A simple linear programming problem can now be formulated. Determine the largest possible value of the linear function

$$f(x_1, x_2) = x_1 + 5x_2 \qquad (6)$$

not for all real numbers x_1 and x_2, but for those satisfying the following inequality constraints:

$$5x_1 + 6x_2 \leqq 30$$
$$3x_1 + 2x_2 \leqq 12$$
$$x_1 \geqq 0$$
$$x_2 \geqq 0$$

The set of points (x_1, x_2) satisfying these inequalities, it has been observed, is the intersection of the sets defined by the inequalities. This set is called the feasible solution set of the linear programming problem. Equation 6 defines the objective function of the problem; it can be used to specify a family of straight lines in the plane, each one of which is obtained by an appropriate selection of a value for $f(x_1, x_2)$. A geometry for the problem appears in Fig. 6, and the dashed lines represent graphs of Eq. 6 for $f(x_1, x_2)$ equal to 12, 16, and 25. The maximization of Eq. 6 subject to the constraints can be represented geometrically by moving the straight line specified by Eq. 6 across the feasible-solution set S until we reach a point of S lying on the line which is "most distant" from the origin. The coordinates of this point will yield a maximum value for $f(x_1, x_2)$. From Fig. 6 it may be seen that the point (0, 5) maximizes $f(x_1, x_2)$, since the value of $f(x_1, x_2)$ can be made larger only by moving the line still further from the origin, in which case it would no longer intersect the solution set. We have for this point

$$\max f(x_1, x_2) = f(0, 5) = 0 + 5(5) = 25$$

If, on the other hand, our objective had been to minimize $f(x_1, x_2)$, we would have moved the line as far down as possible, so that it would intersect S in such a way as to give the smallest permissible value to $f(x_1, x_2)$. In this case, the point would have been the origin itself, and minimum $f(x_1, x_2) = f(0, 0) = 0$.

A feasible solution which either maximizes or minimizes the linear function is said to be an optimal solution, and it can be shown that if a linear program-

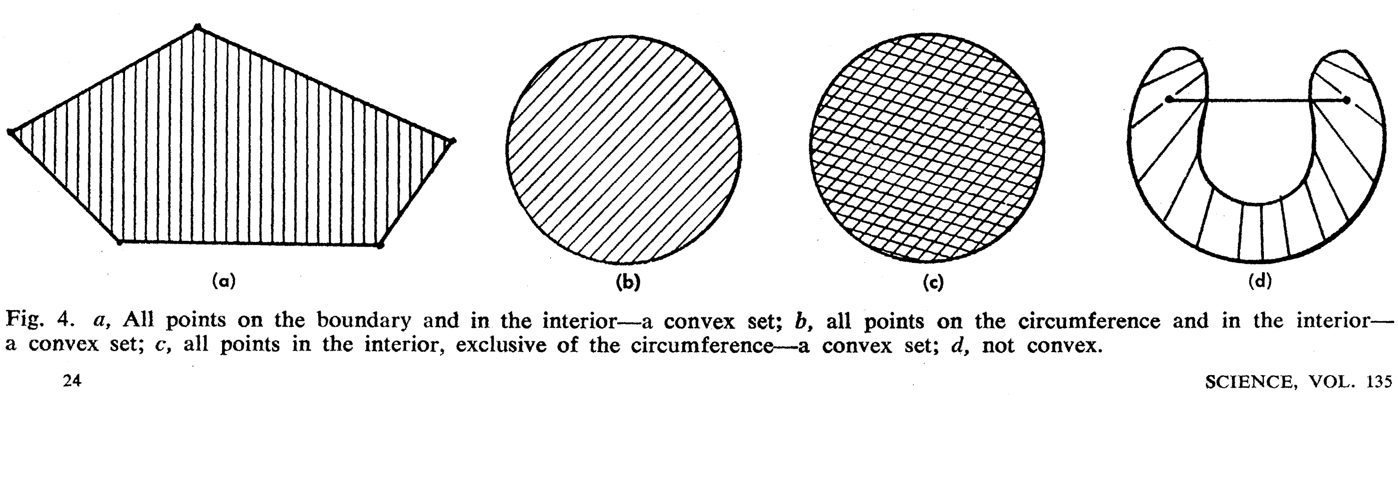

Fig. 4. *a*, All points on the boundary and in the interior—a convex set; *b*, all points on the circumference and in the interior—a convex set; *c*, all points in the interior, exclusive of the circumference—a convex set; *d*, not convex.

ming problem has an optimal solution, then the linear function assumes its greatest or least value over an extreme point of the solution set [the points (0, 5) and (0, 0) in the preceding example are extreme points of the solution set]. This is an "economizing" result from the computational point of view, since we need not search for optimal solutions from among all points of the solution set but only among the extreme points (for some problems, however, the number of extreme points is quite large).

These concepts can be readily generalized to n-space. In a problem corresponding to the maximization of Eq. 6 we seek a point $X = (x_1, \ldots, x_n)$ which will maximize the linear function

$$f = c_1x_1 + \ldots + c_nx_n = \sum_{j=1}^{n} c_jx_j \qquad (7)$$

subject to

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &\leqq b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &\leqq b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \qquad (8) \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &\leqq b_m \\ x_i \geqq 0 (i = 1, \ldots n) & \end{aligned}$$

where each inequality defines a half space in n-space and where the a_{ij}, c_j, b_i are assumed to be known. The intersection of the half spaces, as before, is the convex set of feasible solutions, and Eq. 7 defines a family of hyperplanes in n-space. The maximization of Eq. 7 subject to inequalities 8 can be thought of as moving the hyperplane defined by Eq. 7 across the solution set S until a point in S most distant from the origin and lying in the hyperplane is reached (if it exists). This point will yield a maximum for f and will be an optimal solution to the problem. (Unfortunately, space does not permit consideration of the conditions under which feasible and optimal solutions always exist; extensive discussions of this can be found in Boulding and Spivey and in Gass.)

Primal and Dual Problems

Associated with every linear programming problem is a closely related problem—also of linear programming—called its dual. The original problem (a maximization problem in the illustration above) is called the primal, and

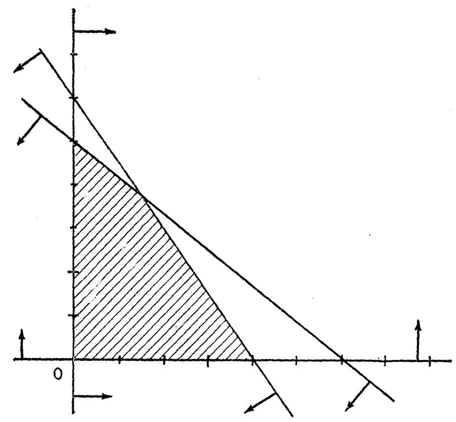

Fig. 5. Geometry of a convex set of points whose coordinates simultaneously satisfy inequalities 1, 3, 4, and 5.

its dual is a minimization problem. Optimal solutions to both problems are intimately related, and if one is given, the other can easily be obtained. Specifically, the dual of the maximization problem of Eq. 7 and of inequalities 8 is the following problem. Minimize

$$g = b_1u_1 + \ldots + b_mu_m = \sum_{i=1}^{m} b_iu_i \qquad (9)$$

subject to

$$\begin{aligned} a_{11}u_1 + a_{21}u_2 + \ldots + a_{m1}u_m &\geqq c_1 \\ a_{12}u_1 + a_{22}u_2 + \ldots + a_{m2}u_m &\geqq c_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \qquad (10) \\ a_{1n}u_1 + a_{2n}u_2 + \ldots + a_{mn}u_m &\geqq c_n \\ u_i \geqq 0 (i = 1, \ldots, m) & \end{aligned}$$

A comparison with the problem in Eq. 7 and in inequalities 8 shows that the coefficients of the variables in Eq. 9 appear as constant terms on the right in inequalities 8 and that coefficients of the variables in Eq. 7 are the constant terms of the right in inequalities 10. The u_i are called dual variables, and the reader acquainted with matrix algebra will observe that the matrix of the system of inequalities 10 is the transpose of the matrix of the system of inequalities 8. Since it can be shown that the dual of the dual problem is the original problem, either problem can be considered the primal and the other the dual.

A number of interesting relationships exist between a primal problem and its dual. It can be proved, for example, that

$$f = \sum_{j=1}^{n} c_jx_j \leqq \sum_{i=1}^{m} b_iu_i = g$$

that for optimal solutions $(x_1^*, \ldots, x_n^*)$ and $(u_1^*, \ldots, u_m^*)$,

$$max\, f = \sum_{j=1}^{n} c_jx_j^* = \sum_{i=1}^{m} b_iu_i^* = min\, g$$

and that a necessary and sufficient condition that both primal and dual problems have optimal solutions is that both problems have feasible solutions. Aside from their great theoretical importance, the intimate relationships between optimal solutions to the primal and dual problems can be utilized in many practical ways. For example, since an optimal solution to one can be obtained from an optimal solution to the other, we are free to solve the simpler of the two and utilize these relationships to determine an optimal solution to the other.

Applications

It will be well at this point to consider a few of the many applications of linear programming.

Bid evaluation. Suppose there are n depots and m separate bidders. Each bidder wishes to produce an amount of a commodity not exceeding a_i $(i = 1, \ldots, m)$, and the demands at the n depots are known to be the numbers b_j $(j = 1, \ldots, n)$. The cost of delivering a unit of the commodity from the ith bidder to the jth depot is c_{ij}. If x_{ij} denotes the quantity purchased from the ith manufacturer for shipment to the jth destination, then the problem is to minimize

$$\sum_{i,j} c_{ij}\, x_{ij}$$

subject to

$$\sum_{j=1}^{n} x_{ij} \leqq a_i \; (i = 1, \ldots, m)$$

$$\sum_{i=1}^{m} x_{ij} \leqq b_j \; (j = 1, \ldots, n)$$

$$x_{ij} \geqq 0 (\text{all } i,j).$$

Production smoothing. A commodity is to be produced, over a given number of time periods, which must satisfy certain requirements in each of the periods. The objective is to minimize total costs (consisting of costs of production and storage and costs of changes

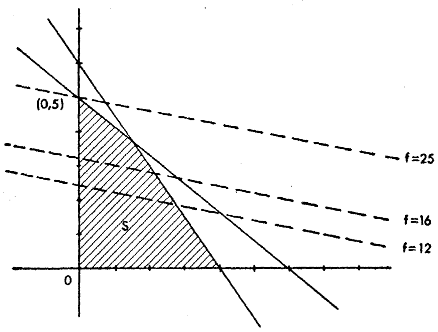

Fig. 6. Geometry of the linear programming problem defined by Eq. 6 and inequalities 1, 3, 4, and 5.

in the rate of production) subject to the condition that shipping requirements be met. Let T be the total number of time periods, let r_t be the known shipping requirement at time period t, let x_t be the quantity produced in period t, and let

$$y_t = x_{t+1} - x_t \geqq 0$$

be the increase in production at time t. Then

$$X_i = \sum_{t=0}^{i} x_t$$

equals total production from period 0 to period i, and

$$R_i = \sum_{t=0}^{i} r_i$$

equals the total requirement from period 0 to period i, where we define R_0 as 0. The excess of accumulated production over accumulated requirements up to time period i is given by

$$u_i = u_0 + X_i - R_i \geqq 0 \qquad (11)$$

where u_0 is a known constant, the excess of production at the beginning of the process. To get an expression for costs, let c_i represent the cost of producing a unit in period $i - 1$ to i; d_i, the cost of storing each unit of excess u_i for one period; and e_i, the cost of increasing production rate one unit time at period i. Total costs then are

$$\sum_{i=0}^{T} (c_i x_i + d_i u_i + e_i y_i) \qquad (12)$$

and the linear programming problem is then to minimize Eq. 12 subject to

$$\sum_{t=1}^{i} x_t - \sum_{t=1}^{i} r_t \geqq 0$$

$$x_i, y_i, u_i \geqq 0$$

for u_i as defined in inequality 11.

A transportation problem. There are m origins or supply centers and n destinations or markets to which a given (homogeneous) commodity is to be shipped. The ith origin has an amount s_i of the commodity ($i = 1, \ldots, m$), and the requirements are such that the jth destination is to receive the amount r_j ($j = 1, \ldots, n$). Let x_{ij} be the quantity of the commodity that is to be shipped from origin i to destination j, and let c_{ij} be the cost of shipping one unit of the commodity from origin i to destination j. Total shipping cost is then

$$\sum_{i,j} c_{ij} x_{ij} \qquad (13)$$

if we also require that destination demands be fulfilled from supplies available at the origins, then the constraints become

$$\sum_{j} x_{ij} = s_i \; (i = 1, \ldots, m) \qquad (14)$$

$$\sum_{i} x_{ij} = r_j \; (j = 1, \ldots, n) \qquad (15)$$

$$x_{ij} \geqq 0 \text{ (all } i, j) \qquad (16)$$

where s_i, r_i, and c_i are given nonnegative integers and supply is assumed to be equal to demand, $\Sigma r_j = \Sigma s_i$. The problem is to minimize Eq. 13 subject to Eqs. 14 and 15 and inequality 16.

This problem was first proposed by Hitchcock, a physicist, in 1941, and independently by the Russian mathematician Kantorovich (1942) and by Koopmans (1944). With the development of linear programming the mathematical properties of the problem were worked out (it was seen to be a problem in linear programming), a general statement of this class of problems was developed, and general conditions for the existence of solutions came to be known.

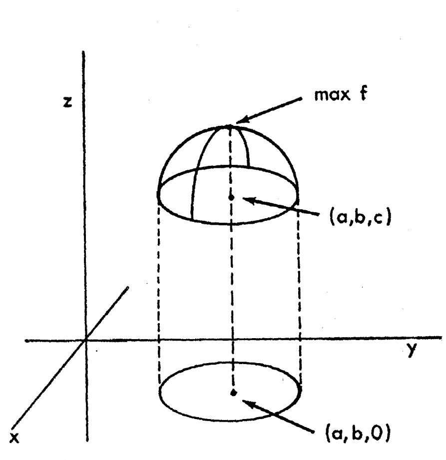

Fig. 7. Graph of Eq. 17 and the set of points satisfying inequality 18.

Table 1. Optimal assignment problem (see text). [After Koopmans and Beckmann]

Worker	Job productivity			
	No. 1	No. 2	No. 3	No. 4
1	25	*20*	5	19
2	*18*	3	*0*	12
3	*22*	4	2	12
4	16	7	−2	*10*

An optimal assignment problem. This is a special case of the transportation problem and can be obtained from it by letting $s_i = r_j = 1$ and $m = n$. There are n jobs to be done and n workers to fill them. The "value" (determined perhaps by a psychological test) or cost of having worker i perform job j is assumed to be known, and the problem is to determine an assignment of the n workers to n jobs which maximizes total value or minimizes total cost. Suppose the value of worker i in job j is as shown in Table 1. Italics indicate an optimal assignment (worker 2 to job 1, worker 1 to job 2, and so on), and it is interesting to observe that an optimal assignment in this case is not one in which each worker is assigned to the job which "he can do best" (where he has the highest value). Worker 1 is most productive in job 1, but this is not in the optimal assignment, whereas worker 2, on the other hand, is least efficient at job 3, yet this does appear in the optimal assignment.

The mathematics of programming problems has undergone further extensions and refinements, primarily in the directions of nonlinear, dynamic, and stochastic (or probabilistic) programming. A nonlinear programming problem is one in which either the objective function or at least one of the constraints has a nonlinear term; a dynamic programming problem is one in which time plays an explicit and fundamental role; and a stochastic problem is one in which at least one of the following in Eq. 7 and in inequality 8 is a random variable: a_{ij}, b_i, or c_j. The existence of a nonlinearity, a time dimension, or a random influence usually creates difficult theoretical and computational problems—much more difficult than those in linear programming—and only a few special classes of these problems can be solved at the present time (although approximation methods are available for some nonlinear problems). Space does not permit an examination of these problems, but a simple example illustrating the kind of difficulties that arise is easily constructed. Consider the nonlinear programming problem of maximizing the function

$$f = c + [1 - (x-a)^2 - (y-b)^2]^{\frac{1}{2}} \quad (17)$$

subject to

$$(x-a)^2 + (y-b)^2 \leq 1 \quad (18)$$

The graph of Eq. 17 is taken to be the half-sphere with center at the point (a, b, c), and the set of points satisfying inequality 18 lie either on the circumference or in the interior of a circle in the x, y plane with center at the point $(a, b, 0)$. The problem is shown in Fig. 7, and the maximum value of f is assumed over the point $(a, b, 0)$, which is an interior point of the feasible solution set. This illustrates a basic difficulty in nonlinear programming problems: it is not possible to confine one's attention, as in the linear case, to extreme points of the solution set nor even to points on the boundary, for an optimal solution can be any point in the feasible set. Clearly, more powerful analytical methods are needed to deal with such problems.

At the present time, research in many aspects of mathematical programming is continuing at a rapid rate, and the greatest prospects for widening the areas of application still more appear to lie in the further development of the theory of dynamic and stochastic programming, and, of course, in the continuing development of computer technology. Since difficult and unsolved problems are found to be perennially attractive to mathematicians, it is eminently reasonable to be optimistic about the future development of the mathematics of optimal choice under conditions of constraint.

Bibliography

C. E. Bishop, R. A. King, J. G. Sutherland, "Programming resource use and capital investment in agriculture," *Management Sci.* **3**, 173 (1957).

K. Boulding and W. A. Spivey, *Linear Programming and the Theory of the Firm* (Macmillan, New York, 1960).

A. Charnes, W. W. Cooper, B. Mellon, "Blending aviation gasolines—a study in programming interdependent activities in an integrated oil company," *Econometrica* **20**, 135 (1952).

A. Charnes and M. H. Miller, "A model for the optical programming of railway freight train movements," *Management Sci.* **3**, 74 (1956).

G. Dantzig, S. Johnson, W. White, "A linear programming approach to the chemical equilibrium problem," *ibid.* **5**, 38 (1958).

W. S. Dorn and H. S. Greenberg, "Linear programming and plastic limit analysis of structures," *Quart. Appl. Math.* **10**, 155 (1957).

L. C. Edie, "Traffic delays at toll booths," *Operations Research* **2**, 107 (1954).

M. M. Flood, "On the Hitchcock distribution problem," *Pacific J. Math.* **3**, 369 (1953).

L. Gainen, D. P. Honig, E. D. Stanley, "Linear programming in bid evaluations," *Naval Research Logistics Quart.* **1**, 48 (1954).

W. W. Garvin, *Introduction to Linear Programming* (McGraw-Hill, New York, 1960).

S. I. Gass, *Linear Programming* (McGraw-Hill, New York, 1958), pp. 37–49.

A. J. Hoffman and W. Jacobs, "Smooth patterns of production," *Management Sci.* **1**, 86 (1954).

T. C. Koopmans and M. J. Beckmann, "Assignment problems and the location of economic activities," *Econometrica* **25**, 53 (1957).

H. W. Kuhn and A. W. Tucker, "Non-linear programming," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (Univ. of California Press, Berkeley, 1951).

P. Massé and R. Gibrat, "Applications of linear programming to investments in the electric power industry," *Management Sci.* **3**, 149 (1957).

V. Riley and S. I. Gass, *Linear Programming and Associated Techniques* (Johns Hopkins Univ. Press, Baltimore, 1958) (a comprehensive bibliography).

News and Comment

The Fallout Booklet: It Did Not Aim for a Passionate Response; Decline of the Test Ban

The promised civil defense booklet is now generally available, and despite its weaknesses, indeed partly because of them, it may turn out to be generally acceptable.

The pamphlet has been criticized by supporters of civil defense for being written in too pedestrian a way to get its message across to the general public as clearly as might have been done, and by opponents of the program for not making clear enough the real meaning of a nuclear attack.

Both criticisms are valid enough in the view of Administration officials—it is easy to see how the pamphlet could have been written in a way calculated to arouse more enthusiastic preparation by the public, although hard to see how this could have been done without, for example, giving the impression that war soon is likely if not inevitable, or that the whole business would be a grand adventure to be looked forward to by every red-blooded citizen, or both. On the other side, the pamphlet could have given a more graphic account of what a nuclear war would be like, but at the risk of making people feel there is really not much point in doing anything.

The pamphlet is therefore aimed at a very modest goal, but one achievable, the Administration hopes, without either misleading or terrifying the public. This involves winning public support for the Administration's program of marking and provisioning shelter areas in existing buildings, encouraging the preparation of more elaborate shelter areas in new buildings, encouraging (but not urging) the preparation of modest home shelters, and perhaps most important, giving the public some elementary information about what it should do if an attack should come.

The whole program is on a modest scale—about equal to next year's federal budget for health research, and less than 1/50 of the defense budget. The pamphlet, and the program generally, is based on the probability that the nation could emerge in significantly less bad shape if modest precautions were taken before an attack and if the citizens were supplied with a general idea of how to behave. The Administration's defense against its critics is principally to argue that among a range of alternatives, all of which are far from satisfactory, the best that can be done is to choose the least unsatisfactory.

Advertising of Shelters

An important sidelight on the shelter question is the advertising for fallout shelters, which has generally tended to give an impression quite different from what Civil Defense officials have been giving. Very often the shelters have been advertised as sort of homey playrooms in peacetime, which would provide nearly complete protection in the