Reprinted from
Proceedings of the 2nd Symposium in Linear Programming, Washington, D.C., 1955, pp. 381–396.
Nat. Bur. Stand., Dept. of Commerce, and Directorate of Management Anal., DCS Comptroller, Headquarters, U.S. Air Force, Washington, D.C., 1955.

THE LINEAR TEAM: AN EXAMPLE OF LINEAR PROGRAMMING UNDER UNCERTAINTY

ROY RADNER

THE LINEAR TEAM:

AN EXAMPLE OF LINEAR PROGRAMMING UNDER UNCERTAINTY[1]

Roy Radner

0. Introduction[2]

In the typical linear programming problem, one maximizes a linear function

$$a'c$$

of a vector $\underline{a}$, subject to the constraint at $\underline{a}$ lie in some closed convex set K (usually polyhedral). The vector $\underline{c}$ of coefficients and the convex set K are assumed to be known with certainty. This paper will explore the consequences of assuming (1) that both $\underline{c}$ and K are subject to some probability distribution, (2) that in any instance $\underline{a}$ is determined on the basis of only incomplete information about $\underline{c}$ and K, and (3) that the values of different coordinates of $\underline{a}$ are determined on the basis of different information.

For example, suppose that $\underline{a}$ is a vector of activities in a firm, that $a'c$ is the profit function, and that there is a person in charge of each activity. Both $\underline{c}$ and K vary from day to day according to some known probability distribution, and at the beginning of each day each person $\underline{i}$ must set the level of his activity for

1. Research undertaken by the Cowles Commission for Research in Economics under contract Nonr-358(01), NR 047-006 with the Office of Naval Research.
2. Thanks to J. Marschak for having introduced me to this problem, and to both him and Mr. Beckmann for helpful discussions on the subject.

that day, knowing only the value of c_i for that day.

In general, if y_i denotes the information available to person *i* at the time he makes his decision, then the level of activity *i* will be determined according to some rule, or decision function, $_i$. Thus:

$$a_i = \alpha_i(y_i).$$

It will be assumed that the firm wants to maximize the expected profit

$$E \sum_i \alpha_i(y_i)\ c_i\ ,$$

subject to the constraint that $\alpha(y) = [\ \alpha_1(y_1), \ldots, \alpha_N(y_N)]$ always be in K.

J. Marschak [1] has called such a group of decision makers a *team*, to emphasize that, although they take different decisions, and take them on the basis of different information, the group members receive a common payoff, which is a function of their joint action and external random variables. Indeed, the fact that there are several persons in the above example is not an essential feature of the team, for the team problem is formally equivalent to an individual decision problem in which the action variable has several components, each of which may be made to depend upon different information. Thus, for example, a single person programming a sequence of activities in time, with uncertainty about the future, also constitutes a team.

The team being discussed in this paper is of course special in that the payoff function is linear in the "activities", or action variables; hence the name "linear team." In Section 1 it is shown that the problem of finding best decision functions for the linear team is a linear programming problem in the space of decision functions. Section 2 shows how the usual analysis by means of Langrangean multipliers (the duality theorem) leads to a system of probability distributions of implicit prices. In Section 3 a special type of information structure is discussed. Some remarks about the role of constraints in programming under uncertainty are made in the last section.

It should be noted that the statement that the probability distribution of $\underline{c}$ and K is "known" is open to two interpretations. The first is that this distribution is objectively known; the second that it represents, at least in part, a subjective "a priori" distribution of unknown parameters. In the latter case, this paper could be interpreted as a discussion of <u>Bayes solutions</u> of the linear team problem.

For other work on linear programming under uncertainty the reader is referred to an unpublished RAND memorandum by G.B.Dantzig [4], and to papers in this volume by G. Tintner and by D. F. Votaw.

Dantzig discusses a number of sequential problems of the type mentioned above. Tintner's paper is concerned with the determination of the probability distribution of the value of a linear programming problem in the case in which the parameters of the problem are random variables, but the problem is solved in each instance with complete knowledge of the parameter values.

1. The Linear Team Problem as a Linear Programming Problem in the Space of Decision Functions.

Let X be a given probability space, with a known probability measure. The elements x of X are called the states of the world. Let $Y_1, \ldots, Y_N$ be N other probability spaces with elements denoted by $y_1, \ldots, y_N$, respectively, and for each i let η_i be a function from X to Y_i. Thus

$$y_i = \eta_i(x)$$

The y_i are called the information variables, and $\eta = (\eta_1, \ldots, \eta_N)$ and $Y = (Y_1, \ldots, Y_N)$ taken together are called the information structure.

For any given information structure, a team decision function is an N-tuple $\alpha = (\alpha_1, \ldots, \alpha_N)$ of real-valued functions α_i on Y_i; A will denote the set of all decision functions. The constraints on the decision functions are defined as follows: for every x in X let K(x) be a closed convex set in N-dimensional Euclidean space

R^N, and let $\mathcal{X}$ be the set of decision functions for which $\alpha(\eta(x))$ is in $K(x)$ for every x; then α is constrained to lie in .

The random coefficients in the linear payoff function are determined by a given function γ from X to R^N; so that for any decision function α, the corresponding expected payoff is:

$$E \sum_i \alpha_i(y_i)\, \gamma_i(x).$$

Defining $\bar{\gamma}_i(y_i)$ as the conditional expectation $E[\gamma_i(x) \mid y_i]$, the expected payoff can also be written as

$$(\alpha, \bar{\gamma}) = E \sum_i \alpha_i(y_i)\, \bar{\gamma}_i(y_i) .$$

The linear team problem can be summarized as follows:

The Problem: Given the information structure, maximize $(\alpha, \bar{\gamma})$ with respect to α, subject to the constraint that α be in $\mathcal{X}$.

Under the usual definitions of addition and scalar multiplication of functions, the set A of decision functions is a linear space. The dimension of A is the sum of the dimensions of the spaces A_i of real-valued functions on Y_i. In order to keep A finite dimensional, and thus keep the mathematics simple, the following assumption will be made:

Assumption: X has only a finite number of elements.

In the light of this assumption, each Y_i also need have only a finite number of elements, and the dimension of A_i will be the number of elements in Y_i.[1]

It is easily verified that if K(x) is closed and convex in R^N for every x, then $\mathcal{K}$ is closed and convex in A. Typically, K(x) is a convex polyhedron defined by inequalities of the form

$$(2) \quad \begin{cases} a \geq 0 \\ T(x)\ a \leq \beta(x), \end{cases}$$

where for each x, T(x) is a MxN matrix, and β is a given function from X to R^M. The random matrix T induces a linear transformation τ from A to B, the space of all functions from X to R^M, so that (2) can be rewritten

$$(3) \quad \begin{matrix} \alpha \geq 0 \\ \tau \alpha \leq \beta . \end{matrix}$$

This particular formulation will be used in the next section, which is on the application of the duality theorem. The <u>adjoint</u> of τ will be denoted by τ^*.

1. This assumption has been made chiefly in order to avoid certain mathematical difficulties associated with the application of the duality theorem in the infinite dimensional case. The author hopes to present a discussion of this particular point in the near future

2. Application of the Duality Theorem: Distributions of Implicit Prices.

Associated with a linear programming problem there is a non-negative saddle point problem involving the use of "Lagrangean multipliers;" furthermore these multipliers can be interpreted as prices. This section will show how the multipliers for the team problem can be interpreted as random prices.

2.1 The Duality Theorem

Consider the problem formulated in the previous section, with the constraint set $\mathcal{X}$ defined by (3).

Applying the well-known saddle point theorem of linear programming to this problem (see, for example [2], p. 487), one immediately gets the following result:

The decision rule α^0 maximizes

$$(\alpha, \bar{\gamma})$$

subject to the constraints (3) if and only if there is a linear functional $\delta^0 \geq 0$ on B such that the pair (α^0, δ^0) constitutes a non-negative saddle point for the bilinear functional

$$\phi(\alpha, \delta) = (\alpha, \bar{\gamma}) + (\beta, \delta) - (\tau\alpha, \delta) \tag{4}$$

$$= (\alpha, \bar{\gamma}) + (\beta, \delta) - (\alpha, \tau^* \delta);$$

i.e. for all $\alpha \geq 0$ and $\delta \geq 0$

$$\phi(\alpha,\delta^{o}) \leq \phi(\alpha^{o},\delta^{o}) \leq \phi(\alpha^{o},\delta).$$

Any linear functional δ on B can of course be represented as a function from X to R^M.

The <u>dual problem</u> is: choose a linear functional δ on B so as to minimize

$$E \sum_i \bar{\beta}_i(x)\ \delta_i(x)\ ,$$

subject to the constraints

$$\begin{cases} \delta \geq 0 \\ \tau^* \delta \geq \ . \end{cases}$$

If α^o is a solution of the original problem, and δ^o is a solution of the dual problem, then (α^o,δ^o) is a saddle point for ϕ; i.e., a solution of the dual problem is a set of Lagrangean functions for the original problem, and, conversely, a solution of the original problem (best team decision function) is a set of Lagrangean functions for the dual problem.

The form of the adjoint of τ is settled in the following lemma.

<u>Lemma</u>. The adjoint τ^* of τ is given by:[1]

---

1. If v is a vector, $[v]_j$ denotes the j'th coordinate of v. Thus if v is a vector valued function $[v]_j(x)$ denotes the j'th coordinate evaluated at x.

$$[\tau^*\beta]_j(y_j) = E\left\{\sum_i t_{ij}(x)\beta_i(x) \mid y_j\right\}$$

$$= E\left\{[T'(x)\beta(x)]_j \mid y_j\right\},$$

where β is any function in B.

<u>Proof</u>. For any α in A and β in B, τ^* must satisfy

$$(\alpha, \tau^*\beta) = (\tau\alpha, \beta),$$

or $$E\sum_j \alpha_j(y_j)\,[\tau^*\beta]_j(y_j) = E\sum_i\sum_j t_{ij}(x)\,\alpha_j(y_j)\,\beta_i(x)$$

$$= E\sum_j \alpha_j(y_j)\sum_i t_{ij}(x)\beta_i(x).$$

Thus for any α_j,

$$E\,\alpha_j(y_j)\,[\tau^*\beta]_j(y_j) = E\,\alpha_j(y_j)\sum_i t_{ij}(x)\beta_i(x)$$

$$= E(\alpha_j(y_j)E\left\{\sum_i t_{ij}(x)\beta_i(x) \mid y_j\right\}).$$

Hence for every y_j,

$$[\tau^*\beta]_j(y_j) = E\left\{\sum_i t_{ij}(x)\beta_i(x) \mid y_j\right\}. \qquad \text{QED}$$

2.2 <u>Distributions of Implicit Prices</u>.

Just as in the case of certainty, the Lagrangean multipliers may be thought of as prices;[2] but being functions of x, they will be random prices. Thus suppose that for each x, $\sum \gamma_j(x)\, a_j$ represents the profit to a firm for operating the N activities at the levels a_j, and that $t_{ij}(x)$ represents the added amount of input i ($i = 1,\ldots,m$) needed to increase the level of the j'th activity by one unit. Let $\delta_i(x)$ be interpreted as the "shadow price" of the i'th input, for each value of x. The conditional expected "net shadow profit" for team member j, given y_j, is:

$$\alpha_j(y_j)\; E\,[\gamma_j(x) - \sum_i t_{ij}(x)\; \delta_i(x) \mid y_j].$$

Thus he will choose an α_j with the property that:

$$\alpha_j(y_j) \begin{Bmatrix} > \\ = \end{Bmatrix} 0, \text{ according as } E[\gamma_j(x) - \sum_i t_{ij}(x)\delta_i(x) \mid y_j] \begin{Bmatrix} > \\ \leq \end{Bmatrix} 0.$$

This condition provides the basis for a probabilistic analogue of the set of rules Koopmans (ibid.) gives for the maintenance of a best activity vector by a price mechanism under conditions of certainty. It should be emphasized that it is the price function δ_j which is used by the j'th team member in making his decision, and not the value $\delta_j(x)$ of the price function in any given instance.

2. See Koopmans, [3], pp. 93-95, for the certainty case.

The problem of how to use such a probabilistic price mechanism as the basis for an optimum-seeking model is an interesting one, but will not be discussed in this paper.

3. Information Variables with Independent Ranges

This section will consider the consequences of the assumptions that the ranges of variation of the different information variables y_i are independent, and that the constraint set K is not random. It will be seen that these assumptions generally result in a reduction of the "size" of the problem.

Assumption: (1) For every i, the conditional range of y_i given all the other $y_j (j \neq i)$ is independent of the values of the other y_j.
(2) $K(x) = K$ independent of x.

For any α, and for each i, let $[\underline{a}_i, \bar{a}_i]$ be the smallest closed interval such that

$$\text{Prob}\left\{\alpha_i(y_i) \in [\underline{a}_i, \bar{a}_i]\right\} = 1.$$

It is easy to see that the effect of the above assumption is that the constraint that $\alpha(y)$ be in K for all y (K is a closed convex set) is equivalent to the requirement that the Cartesian product

$$I(\alpha) = \prod_i [\underline{a}_i, \bar{a}_i]$$

be contained in K.

Given any particular rectangle I in R^N (i.e. a Cartesian product of intervals) one may ask: what is the best α such that $I(\alpha) = I$? If $\hat{\alpha}$ is the best decision rule for which $I(\alpha) = I$, then for every i and every y_i, $\hat{\alpha}_i(y_i)$ must maximize

$$\alpha_i(y_i)E(\gamma_i \mid y_i) + \sum_{j \neq i} E(\hat{\alpha}_j(y_j)\gamma_j \mid y_i)$$

subject to $\underline{a}_i \leq \alpha_i(y_i) \leq \bar{a}_i$. This gives:

$$\hat{\alpha}_i(y_i) = \begin{Bmatrix} \bar{a}_i \\ \underline{a}_i \end{Bmatrix} \quad \text{as } E(\gamma_i \mid y_i) \begin{Bmatrix} > \\ \leq \end{Bmatrix} 0.$$

Thus the best payoff corresponding to the rectangle I is

$$U(I) = \sum_i (\bar{a}_i \, \bar{d}_i + \underline{a}_i \, \underline{d}_i)$$

where $\bar{d}_i = E[E(\gamma_i \mid y_i) \mid E(\gamma_i \mid y_i) > 0] \text{ Prob } [E(\gamma_i \mid y_i) > 0]$

$\underline{d}_i = E[E(\gamma_i \mid y_i) \mid E(\gamma_i \mid y_i) \leq 0] \text{ Prob } [E(\gamma_i \mid y_i) \leq 0]$

[Note that $\bar{d}_i \geq 0$ and $\underline{d}_i \leq 0$.]

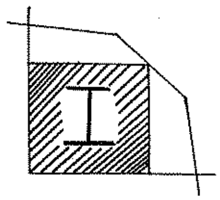
Thus the original problem has been reduced to one of maximizing U(I), which is linear in the $\bar{a}_i$ and $\underline{a}_i$, subject to the condition that I be contained in K. If K is defined by M linear constraints (besides the constraints $a_i \geq 0$), the reduced problem is a linear programming problem in 2N variables (the $\bar{a}_i$ and $\underline{a}_i$) with $(2^N)M$

linear constraints (besides the 2N positivity constraints), independent of the dimension of A. The number 2^N of course, gets very large, very quickly!

If in the original constraints,

$$T\ a \leq b,$$

the elements of T and b are all ≥ 0, then the above problem is greatly simplified. In that case it is not hard to show that all $\underline{a}_i$ must be zero. (This diagram will convince the reader immediately:,



Thus the problem reduces to: "choose $\bar{a}_i$ so as to maximize

$$\sum \bar{d}_i\ \bar{a}_i\ ,$$

subject to $(\bar{a}_i,\ldots,\bar{a}_n)$ in K." This is a problem of the same "size" as the original, with the same constraints, and with coefficients which can be calculated from the distribution of $\gamma(x)$.

4. <u>The Role of the Constraints in an Uncertain Program</u>

Because of the requirement that the given constraints be satisfied with certainty, even though the information on which the activity levels are based is random, the convex set χ of decision rules may be quite small in many instances. This corresponds to our intuitive feeling that if different decisions are based on <u>different</u> (random) information, then it may be very difficult to make sure that the constraints are always satisfied, without unduly restricting the range over which the decisions can vary. This will be illustrated by a two variable example.

Let the linear function to be maximized be:

$$E[\gamma_1(x)\alpha_1(y_1) + \gamma_2(x)\alpha_2(y_2)],$$

and the constraints be:

$$\alpha_1(y_1) + \alpha_2(y_2) \leq 1, \; \alpha_i(y_i) \geq 0.$$

Suppose that the conditional ranges of y_1 and y_2 are independent, as in Sec. 3. The result at the end of that section applies here, giving

$$\alpha_i(y_i) = \begin{Bmatrix} \bar{a}_i \\ 0 \end{Bmatrix} \text{ as } E(\gamma_i \mid y_i) \begin{Bmatrix} > \\ \leq \end{Bmatrix} 0.$$

where $\bar{a}_1$ and $\bar{a}_2$ must be chosen to maximize

$$\bar{d}_1 \bar{a}_1 + \bar{d}_2 \bar{a}_2 ,$$

subject to

$$\bar{a}_1 + \bar{a}_2 \leq 1, \; \bar{a}_i \geq 0,$$

where $\bar{d}_i = E[E(\gamma_i \mid y_i) \quad E(\gamma_i \mid y_i) > 0]$ Prob $[E(\gamma_i \mid y_i) > 0]$.

It is easy to see that for the best choice, $\bar{a}_j = 1$ for that j for which $\bar{d}_j$ is the largest, and the other $\bar{a}_i = 0$. Thus the level of only one of the activities actually depends upon the information received.

Since the constraints in a linear programming problem often arise as approximations in a situation in which the loss incurred by violating the constraints rises sharply with the magnitude of the violation, the disturbing situation exemplified above might be corrected by dropping the constraints and modifying the payoff function appropriately. This of course would change the problem into an ordinary maximum problem, without the linear programming "flavor". In many cases, however, the loss incurred by violating the constraints might be realistically represented as the cost of some linear program needed to remedy the violation. Such a situation would be covered by the present formulation, by dividing the activities in two groups: those which are based on incomplete information, but are not subject to constraints; and those which are based on complete information (the remedial activities) and are subject to constraints.

References

[1] J. Marschak, "Elements for a Theory of Teams," in Management Science, Vol. I, No. 2, pp. 127-137 (January 1955).

[2] H. Kuhn and A. Tucker, "Nonlinear Programming," in Second Berkeley Symposium on Mathematical Statistics and Probability (Neyman, ed.), U. of California Press, Berkeley, 1951 (p. 481-492).

[3] T. C. Koopmans, "Analysis of Production as an Efficient Combination of Activities," Ch. III of Activity Analysis of Production and Allocation, Wiley, N. Y., 1951.

[4] G. B. Dantzig, "Linear Programming under Uncertainty" (Notes on Linear Programming: Part XVII), RAND Research Memorandum 1374, November 16, 1954 (unpublished).

Cowles Commission for Research in Economics