

AD629505

FINDING EVERETT'S LAGRANGE MULTIPLIERS
BY LINEAR PROGRAMMING

R. Brooks &

A. Geoffrion

February 25, 1966

CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION			
Hardcopy	Microfiche		
\$1.10	\$0.50	9 pp	as
ARCHIVE COPY			

Code 1,

P-3315

FINDING EVERETT'S LAGRANGE MULTIPLIERS
BY LINEAR PROGRAMMING

R. Brooks &
A. Geoffrion^{*}

The RAND Corporation, Santa Monica, California

In an article [4] in this journal in 1963, Everett observed that if x^0 is optimal in

$$(1) \quad \begin{array}{ll} \text{Maximize} & f(x) - \sum_{i=1}^m u_i g_i(x), \\ x \in X & \end{array}$$

where the m constants u_i are non-negative "multipliers" and f and the g_i are arbitrary real-valued functions defined over an arbitrary set X , then x^0 also maximizes $f(x)$ over all $x \in X$ satisfying $g_i(x) \leq g_i(x^0)$ ($i=1, \dots, m$).^{1/} Thus to solve

$$(2) \quad \begin{array}{ll} \text{Maximize} & f(x) \text{ subject to } g_i(x) \leq b_i, i=1, \dots, m, \\ x \in X & \end{array}$$

where the b_i are given constants (it is convenient to think of the b_i as the amounts of available resources) it is sufficient to find non-negative multipliers u_i^0 such that -- and this we call Everett's Condition -- a corresponding optimal solution x^0 of (1) can be found

^{*} Any views expressed in this paper are those of the authors. They should not be interpreted as reflecting the views of The RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The RAND Corporation as a courtesy to members of its staff.

This paper was prepared for submission to Operations Research as a Letter to the Editor.

^{1/} This observation is Everett's "Main Theorem" [4, p. 401].

that satisfies $g_i(x^0) = b_i$ ($i=1, \dots, m$). If such multipliers exist and a convenient mechanism for finding them is available, then solving (2) by solving (1) may be computationally convenient. For many problems of practical interest, however, such multipliers do not exist; but there may still be multipliers for which the $g_i(x^0)$ approximate the b_i closely enough for x^0 to be a useful approximate solution to (2). This approach amounts to reducing (2) to a problem without the g_i constraints.

The k^{th} step ($k \geq 2$) of the iterative procedure implicitly suggested by Everett for finding an (approximate) solution to (2) is (here $u \equiv (u_1, \dots, u_m)$):

(k.1) Based on knowledge of $u^1, x^1, \dots, u^{k-1}, x^{k-1}$,

choose multipliers $u_i^k \geq 0$ ($i=1, \dots, m$) in an attempt to satisfy Everett's Condition.

(k.2) Solve (1) with $u = u^k$ for an optimal solution x^k .

(k.3) If $g_i(x^k)$ is "sufficiently near" b_i , $i=1,$

m , then stop; x^k is sufficiently near to

being optimal in (2). Otherwise, go to step $k+1$.

Step 1 is the same as the general step, except that it begins with an arbitrary u^1 (guessed on the basis of past experience with a similar problem, say). It is assumed that some method is available for performing substep (k.2).^{2/} How to perform substep (k.1) when $m \geq 2$ was

^{2/} Throughout this paper we assume, as Everett did implicitly in his, that (1) achieves its maximum for any set of non-negative multipliers. A sufficient condition for this when $X \subset R^n$ is that X be closed and bounded and f and the g_i be continuous. Similar sufficient conditions exist for more general spaces (e.g., [2, p. 69]). See also footnote 4.

left largely unresolved by Everett, and stimulated the present note.^{3/}

The main purpose of this note is to indicate how one might approximate the desired multipliers by means of linear programming. First, however, we weaken Everett's Condition slightly so that his approach can be applicable to problems with ineffectual constraints. A relation then becomes apparent to the saddle-point condition of Kuhn and Tucker [8] for nonlinear programming. Since Everett's approach seems most competitive with other known methods for certain discrete allocation problems, we consider this case in some detail. It will be seen that Everett's method, when the multipliers are found by linear programming, becomes essentially the Simplex method with a "column-generating" feature applied to an approximation of (2). Finally, we point out a relation to the so-called decomposition method of concave programming [3, 10] for continuous allocation problems.

Weakening Everett's Condition

In certain problems with ineffectual constraints, Everett's Condition is unnecessarily restrictive in that, when a multiplier is zero, it is not necessary to require that the corresponding constraint be satisfied with strict equality. All that is needed is to find x^0 and u^0 such that

- (i) x^0 is optimal in (1) with $u = u^0$ and
- (ii) $u^0 \geq 0$ and $u_1^0 > 0$ (resp. $= 0$) implies
 $g_i(x^0) = b_i$ (resp. $\leq b_i$), $i=1, \dots, m$.

It is easily shown that if these conditions are satisfied, then x^0 is optimal in (2). We shall henceforth deal with this slightly modified version of Everett's approach.

^{3/}We would like to thank David McGarvey for encouraging our interest in this question.

It is of interest to note that (i) and (ii) are equivalent to the requirement that (x^0, u^0) be a saddle-point of the Lagrangian

$$L(x, u) \equiv f(x) - \sum_{i=1}^m u_i (g_i(x) - b_i), \text{ i.e.,}$$

$$L(x, u^0) \leq L(x^0, u^0) \leq L(x^0, u) \text{ for all } x \in X \text{ and } u \geq 0.$$

Thus Everett's approach is seen to be essentially the attempt to construct a saddle-point for $L(x, u)$. Kuhn and Tucker [8] and others have given conditions on (2) which guarantee the existence of such a saddle-point. The basic condition for Euclidean spaces is that X be a convex set, f a concave function, and the g_i convex functions which satisfy any one of a number of mild qualifications [1]. Similar conditions for more general spaces are known (see, e.g., [7]). Unfortunately, such conditions do not cover the case in which X is discrete, the situation of greatest interest to Everett and perhaps the one in which his (modified) approach is most promising.

Finding the Multipliers by Linear Programming

When (2) is a linear programming problem, i.e., when X is the non-negative orthant of E^n and f and the g_i are linear functions, then it is not difficult to show that (x^0, u^0) satisfies conditions (i) and (ii) above if and only if x^0 solves (2) and u^0 solves the dual of (2). The u_i^0 are often interpreted as the "dual prices" associated with (2), and are produced as an automatic by-product of the computational solution of (2). Dropping the assumption of linearity now, and observing that the burden of substep (k.1) is to approximate such prices on the basis of the data $u^1, x^1, \dots, u^{k-1}, x^{k-1}$, it seems natural to use linear programming to compute the prices corresponding

to a linearized version of (2) over the convex hull of the grid $\langle x^1, \dots, x^{k-1} \rangle$. The resulting linear program, the dual prices of which are required at substep (k.1), is:

$$(3) \quad \begin{array}{ll} \text{Maximize} & \sum_{t=1}^{k-1} \lambda_t f(x^t) \text{ subject to } \sum_{t=1}^{k-1} \lambda_t = 1 \\ & \lambda_t \geq 0 \\ & \sum_{t=1}^{k-1} \lambda_t g_i(x^t) \leq b_i, \quad i=1, \dots, m. \end{array}$$

Substep (k.1 LP): Solve (3) for the dual prices

$u_0^k, u_1^k, \dots, u_m^k$ corresponding to the $m+1$ constraints.

By linear programming theory, $u_i^k \geq 0$ ($i=1, \dots, m$). The significance of u_0^k will become apparent below.

Discrete Case

If $X = \{\xi_1, \dots, \xi_N\}$, where N is a finite positive integer, then Everett's procedure using (k.1 LP) is very close to the Simplex method for the linear programming problem

$$(4) \quad \begin{array}{ll} \text{Maximize} & \sum_{j=1}^N \lambda_j f(\xi_j) \text{ subject to } \sum_{j=1}^N \lambda_j = 1 \\ & \lambda_j \geq 0 \\ & \sum_{j=1}^N \lambda_j g_i(\xi_j) \leq b_i, \quad i=1, \dots, m. \end{array}$$

The sub-problem (1), which now takes the form

$$(5) \quad \begin{array}{ll} \text{Maximize} & f(\xi) - \sum_{i=1}^m u_i^k g_i(\xi), \\ & \xi \in \{\xi_1, \dots, \xi_N\} \end{array}$$

does nothing more than determine (by the usual Simplex criterion) which new variable to bring into the basis at the k^{th} iteration.^{4/}

^{4/} In practice one probably would not solve (5) completely at every step, particularly in the early steps or when N is very large; from the theory of the Simplex method it is known that it is enough to find a ξ_j that gives a value greater than u_0^k to the maximand.

This permits the economy of carrying explicitly at one time no more than $m+1$ of the N columns corresponding to the ξ_j . The usual Simplex termination signal occurs at the first step k_0 such that

$$(6) \quad \begin{array}{ll} \text{Maximize} & [f(\xi) - \sum_{i=1}^m u_i^k g_i(\xi)] \leq u_0^k \\ \xi \in \{\xi_1, \dots, \xi_N\} & \end{array}$$

(actually the maximum will be u_0^k). Thus in the finite discrete case Everett's procedure becomes precisely the Simplex method applied to (4) with a "column-generation" feature if substep (k.3) is replaced by

Substep (k.3 LP): If (6) holds, stop.

Otherwise, go to step $k + 1$.

Since (4) is a finite linear program, Everett's procedure with substeps (k.1 LP) and (k.3 LP) is finitely convergent to the optimal solution λ_j^* , $j=1, \dots, N$, of (4).

This method has been used to advantage by Gilmore and Gomory [6]. In their problem a ξ_j was a cutting pattern, and the subproblem a knapsack problem.

The question arises regarding the relation of the optimal solution of (4) to the original problem (2). Harking back to Everett's discussion of his method in terms of "payoff-constraint space," we see that if the points $(f(\xi_j), g_1(\xi_j), \dots, g_m(\xi_j)) \in R^{m+1}$ ($j=1, \dots, N$) are sufficiently dense near the boundary of their convex hull, then some of the policies ξ_j corresponding to $\lambda_j^* > 0$ (and there will be no more than $m+1$ of these) will be good approximate solutions to (2).

Note that it is not necessary to store all of the ξ_j corresponding to the basic λ_j as the calculations proceed, but only the corresponding $f(\xi_j)$ and $g_i(\xi_j)$, $i=1, \dots, m$. After termination, the "basic" ξ_j

can be recovered if desired by utilizing the fact that they "price out" to 0. That is, they achieve the maximum $u_0^* \equiv u_0^{k_0}$ in (6). In fact all of the ξ_j that achieve u_0^* in (6) are "used" by some optimal solution of (4). If it is desired to examine (7.j) for the ξ_j used in near-optimal solutions of (4), then one should recover the ξ_j that satisfy

$$(8) \quad f(\xi_j) - \sum_{i=1}^m u_i^* g_i(\xi_j) \geq u_0^* - \epsilon$$

for some suitably small $\epsilon > 0$.

A possibly useful interpretation of (4) is the following: it is the extension of (2) from pure to mixed (randomized) strategies with f and the g_i replaced by their expectations. In this interpretation, λ_j^* is the probability of utilizing allocation ξ_j . When mixed strategies have a legitimate and acceptable interpretation, then (2) should have been written as (4) in the first place.^{5/}

Continuous Case

If X is not a finite discrete set, then the analysis of the previous case is complicated by the fact that there are an infinite number of variables in (4). Nevertheless, Everett's procedure using substep (k.1 LP) is almost exactly the so-called decomposition procedure for nonlinear programming [3, 10]. When X is a bounded convex set and f is concave and the g_i are convex functions, then the sequence $\langle \sum_{t=1}^{k-1} \lambda_t x^t \rangle$ converges [3, 9] to an optimal solution of (2) as $k \rightarrow \infty$.

^{5/} Cf. Fromovitz [5], and Gilmore and Gomory [6].

REFERENCES

1. Arrow, K. J., L. Hurwicz, and H. Uzawa, "Constraint Qualifications in Maximization Problems," Naval Research Logistics Quarterly, 8, No. 2 (June 1961), 175-190.
2. Berge, C., Topological Spaces, MacMillan, 1963.
3. Dantzig, G., Linear Programming and Extensions, Princeton University Press, 1963.
4. Everett, H., "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources," Operations Research, 11, No. 3 (May-June 1963), 399-417.
5. Fromovitz, S., "Nonlinear Programming with Randomization," Management Science, 11, No. 9 (July 1965), 831-846.
6. Gilmore, P., and R. Gomory, "A Linear Programming Approach to the Cutting-Stock Problem," Operations Research, 9, No. 6 (Nov.-Dec., 1961), 849-859.
7. Hurwicz, L., "Programming in Linear Spaces," Studies in Linear and Nonlinear Programming (K. J. Arrow, L. Hurwicz, and H. Uzawa, eds.), Stanford University Press, 1958.
8. Kuhn, H. K., and A. W. Tucker, "Nonlinear Programming," Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 1951, 481-492.
9. Roubault, M., "De la dualité des programmes mathématiques convexes," Abstract, 1964 International Symposium on Mathematical Programming, London, July 6-10, 1964.
10. Wolfe, P., "Methods of Nonlinear Programming," in Recent Advances in Mathematical Programming (R. Graves and P. Wolfe, eds.), McGraw-Hill, 1963.