Strictly Concave Parametric Programming, Part II: Additional Theory and Computational Considerations

Arthur M. Geoffrion

Management Science, Vol. 13, No. 5, Series A, Sciences (Jan., 1967), 359-370.

Stable URL:
http://links.jstor.org/sici?sici=0025-1909%28196701%2913%3A5%3C359%3ASCPPPI%3E2.0.CO%3B2-M

Management Science is currently published by INFORMS.

MANAGEMENT SCIENCE
Vol. 13, No. 5, January, 1967
Printed in U.S.A.

STRICTLY CONCAVE PARAMETRIC PROGRAMMING, PART II: ADDITIONAL THEORY AND COMPUTATIONAL CONSIDERATIONS*

ARTHUR M. GEOFFRION†

University of California, Los Angeles

The theory presented in Part I of this paper led to a Basic Parametric Procedure for a broad class of strictly concave parametric programs. In this part, additional theory is developed that facilitates efficient computational implementation. An illustrative graphical example is given, and some extensions are indicated.

1. Introduction

Part I of this paper presented and justified a Basic Parametric Procedure for solving concave parametric programs of the form:

$$(P\alpha) \quad \text{Maximize} \quad \alpha f_1(x) + (1 - \alpha)f_2(x) \qquad \text{subject to} \quad g_i(x) \geqq 0, \quad i = 1, \cdots, m,$$

for each α in the unit interval, where the functions are concave and satisfy certain regularity conditions and a solution of $(P\alpha)$ is available for some value of α. The importance of $(P\alpha)$ derives from the fact that it enables the computation of tradeoff curves between two criterion functions and the fact that it provides, by a simple device, a deformation method for ordinary (non-parametric) concave programming.

In this part we present additional theory that facilitates the development of efficient computational algorithms based on the Basic Parametric Procedure. Although an effort has been made to keep the discussion relatively self-contained by informally reviewing certain essential definitions and results, the reader should refer to Part I [3] for details. Even so, full details of several of the proofs are to be found only in [2].

1.1 *Informal Review*

The Lagrangian conditions associated with $(P\alpha)$ for fixed α when equality is required for a distinguished subset $S(S \subset \{1, \cdots, m\})$ of constraints are:

$$(=S)\alpha \qquad \nabla_x f(x; \alpha) + \textstyle\sum_{i=1}^{m} u_i \nabla_x g_i(x) = 0$$
$$g_i(x) = 0, \quad i \in S; \qquad u_i = 0, \quad i \notin S,$$

* Received June 1965 and revised June 1966.

† The author is pleased to express his sincere appreciation to Professor Harvey M. Wagner, under whose guidance this research was carried out as a portion of the author's doctoral dissertation [2] in the Operations Research Program at Stanford University. Thanks are also due to L. Breiman and J. B. Rosen for helpful comments. This work was supported partially by the Office of Naval Research under Task NR 047-041, Contract Nonr 233(75), The National Science Foundation under Grant 25064, and by the Western Management Science Institute under a grant from the Ford Foundation.

where $f(x; \alpha)$ denotes $\alpha f_1(x) + (1 - \alpha)f_2(x)$, ∇ the gradient operator, and 0 the null vector when appropriate. Let $(x^S(\alpha), u^S(\alpha))$ denote a solution of the equations $(=S)\alpha$. From the well-known results of Kuhn and Tucker, it follows that $x^S(\alpha)$ solves $(P\alpha)$ if $g_i(x^S(\alpha)) \geqq 0$, $i \not\varepsilon S$, and $u_i^S(\alpha) \geqq 0$, $i \varepsilon S$; in which case we write $(x^S(\alpha), u^S(\alpha)) = (x^*(\alpha), u^*(\alpha))$ and say that S is *valid*. The Basic Parametric Procedure of Part I is designed to solve $(P\alpha)$ on [0, 1] by maintaining the identity of a valid subset of constraints as α traverses the unit interval, the continuity properties of $(P\alpha)$ being exploited in an essential way. The values of α at which the identity of a valid set changes are called "points of change".

The Basic Parametric Procedure can be paraphrased as follows (we arbitrarily take 0 as the initial value of α):

Step 1: By any convenient method, find the optimal solution and dual variables $(x^*(0), u^*(0))$ of $(P0)$. Set $\alpha^0 = 0$, S equal to any set valid at $\alpha = 0$ (e.g., $S = \{i: u_i^*(0) > 0\}$), and $(x, u)^0 = (x^*(0), u^*(0))$.

Step 2: Solve $(=S)\alpha$ as α increases above α^0 for its unique continuous solution satisfying the left end-point condition $(x^S(\alpha^0), u^S(\alpha^0)) = (x, u)^0$, namely $(x^*(\alpha), u^*(\alpha))$, until either $\alpha = 1$ or a point α' is encountered to the right of which S is no longer valid. In the first case, terminate; in the second case, set $(x, u)^0 = (x^*(\alpha'), u^*(\alpha'))$ and go to Step 3.

Step 3: Among all sets valid at α', find (by enumeration, if necessary) one which is valid to the right of α'. Call it S'. Set $\alpha^0 = \alpha'$, $S = S'$, and return to Step 2.

The Basic Theorem asserts, assuming four conditions hold, that this procedure is well-defined, that $(x^S(\alpha), u^S(\alpha)) = (x^*(\alpha), u^*(\alpha))$ on $[\alpha^0, \alpha']$ at each execution of Step 2, and that Step 3 will be executed only a finite number of times before a terminal state is reached. Condition 1 requires the analyticity of all functions and the concavity of the constraint functions, Condition 2 the nonemptiness and boundedness of the feasible region X of $(P\alpha)$, Condition 3 the local strict concavity (negative definite hessians) of the f_i on X, and Condition 4 the linear independence of the gradients of the binding constraints at $x^*(\alpha)$ for each $\alpha \varepsilon [0, 1]$. These conditions are assumed throughout this part as well.

In section 2, we ameliorate the enumeration seemingly required at Step 3. Section 3 hosts an illustrative graphical example. In section 4, we demonstrate the applicability of Newton's method in the appropriate sense for solving $(=S)\alpha$ as α varies in Steps 2 and 3, and discuss several aspects of computational implementation. Finally, extensions concerning linear equality constraints, nonstrictly concave f_i, and more general parametric programs are briefly noted.

2. Improvement of Step 3

Step 3 of the Basic Parametric Procedure involves a certain amount of trial and error—at a point of change $\alpha' < 1$, different subsets of constraints that are valid at α' must be tried until one is found that is valid to the right of α'. By [3, Cor. 1.2], a subset S is valid at α if and only if $A\alpha \subset S \subset B\alpha$, where $A\alpha \equiv \{i: u_i^*(\alpha) > 0\}$ and $B\alpha \equiv \{i: g_i(x^*(\alpha)) = 0\}$. Under our conditions, the constraints of $B\alpha' - A\alpha'$ can be shown to be degenerate in the sense that they are redundant and yet satisfied exactly at $x^*(\alpha')$. Thus, up to 2^k trials could be necessary, where k is the number of degenerate constraints. In most problems, k will be very small at each point of change, and hence the trial and error nature of Step 3 as it now stands is not unsatisfactory. When k is large, however, enumeration may be onerous. Faced with this possibility, one may follow two main courses of inquiry. One may attempt to construct methods of perturbing $(P\alpha)$ so as to ensure that $B\alpha - A\alpha$ consists of only one or two constraints at each point of change (*cf.* [6, p. 125] and [9, p. 156]). Alternatively, one may attempt to devise rules for deciding in what order the trials should be made so as to tend to keep the number of erroneous trials small. We choose to follow the second course of inquiry, because (a) this type of investigation is conspicuously lacking at present (for an exception in the context of a related problem see [8]), and (b) the second course of inquiry must be undertaken before the need for perturbation can be established.

We begin by establishing some terminology. Suppose that Step 2 has ended with the point of change $\alpha' < 1$. Let $\alpha'+$ be a point between α' and the next largest point of change. If S is valid at α' but not at $\alpha'+$, the unique continuous solution of $(=S)\alpha$ satisfying the left end-point value $(x^*(\alpha'), u^*(\alpha'))$ violates either the condition $g_i(x) \geqq 0$, $i \not\in S$, or the condition $u_i \geqq 0$, $i \in S$, or possibly both, as α increases above α'. In other words, S "causes an *alarm*" as α increases[1] above α'. A violation of the condition $g_i(x) \geqq 0$, $i \not\in S$, is called a *feasibility alarm*, while a violation of the condition $u_i \geqq 0$, $i \in S$, is called an *optimality alarm*. By continuity, the set of feasibility alarms must be contained in $B\alpha' - S$, and the set of optimality alarms must be contained in the set $S - A\alpha'$; hence, all alarms are from $B\alpha' - A\alpha'$. Since S is not valid at $\alpha'+$, by [3, Cor. 1.2] either $\{S - B\alpha'+\} \neq \phi$ or $\{A\alpha'+ - S\} \neq \phi$. The set $S - B\alpha'+$ will be called the *excess* of S at $\alpha'+$, and $A\alpha'+ - S$ will be called the *deficiency* of S at $\alpha'+$. Clearly, the smallest change in S which will result in a set which is valid at $\alpha'+$ is to delete its excess and add its deficiency. The number of constraint indices of $\{A\alpha'+ - S\} \cup \{S - B\alpha'+\}$ is therefore a measure of the minimum *distance*,[2] which we denote by $d(S)$, between S and the collection of all sets which are valid at $\alpha'+$.

An obvious conjecture (and one with some foundation [2, p. 97]) is that the feasibility alarms coincide with the deficiency, and that the optimality alarms coincide with the excess of S at $\alpha'+$. If this were so, then by deleting the constraints which yield optimality alarms and adding those which yield feasibility alarms, one could obtain from S a set which is valid at $\alpha'+$. Unfortunately, this conjecture of perfect coincidence is not necessarily true, as can easily be

[1] Since $x^S(\alpha)$ and $u^S(\alpha)$ are, under our assumptions, *analytic* functions, there is an $\epsilon > 0$ such that each component of $(g(x^S(\alpha)), u^S(\alpha))$ has constant sign on $(\alpha', \alpha' + \epsilon)$. It is in this sense that we define the alarms caused by S "as α increases above α'."

[2] Let the distance between a set C and a set D, where C and D are both sets of integers, be defined as the number of elements in the symmetric difference set $\{C - D\} \cup \{D - C\}$.

demonstrated by counterexamples.[3] What can be shown, however, is that at least one of the alarms given during a failure is from the deficiency or excess at $\alpha' +$ of the trial set. After proving this fact, we use it to derive an ordering of trials at Step 3.

2.1 A Theorem

In the following, we refer to $(x^S(\alpha), u^S(\alpha))$ in an interval about a point of change α' at which S is valid. From Theorems 2 and 3 of [3], it is clear that by making such an interval suitably small we may assume that $(x^S(\alpha), u^S(\alpha))$ is uniquely defined and continuous there.

Lemma 4.1: Let $\alpha' \varepsilon [0, 1]$ be a point of change; let S be valid at α', and assume that Conditions 1 through 4 hold. Then there exists a convex set $X' \supset X$ and an open interval containing α' such that, for each fixed value of α in this interval, $x^S(\alpha)$ is the optimal solution of

$$\text{Maximize}_{x \varepsilon X'} \; f(x; \alpha) \quad \text{subject to} \quad \begin{aligned} g_i(x) &= 0, & i &\varepsilon \{S - S^+\alpha\} \\ g_i(x) &\geqq 0, & i &\varepsilon S^+\alpha, \end{aligned}$$

where $S^+\alpha \subset \{i \varepsilon S: u_i^S(\alpha) \geqq 0\}$.

The proof [2, p. 100] of this technical lemma is omitted. An easy proof can, however, be constructed from the Kuhn-Tucker Theorem when all constraints are linear; in this case, X' may be taken to be E^n, and $(=S)\alpha$ are necessary and sufficient conditions for a maximum of $f(x; \alpha)$ subject to $g_i(x) = 0$, $i \varepsilon S$.

Theorem 4: Let $\alpha' \varepsilon [0, 1]$ be a point of change; let S be valid at α', and assume that Conditions 1 through 4 hold.

There exists an open interval containing α' such that, for each fixed value of α in this interval, if S is *not* valid at α then either

$$g_i(x^S(\alpha)) < 0 \qquad \text{for some } i \varepsilon \{A\alpha - S\} \text{ or}$$
$$u_i^S(\alpha) < 0 \quad \text{for some } i \varepsilon \{S - B\alpha\}, \text{ or both.}$$

Proof: It is sufficient to show that the theorem holds on the interval mentioned in Lemma 4.1. We suppose the conclusion to be false and derive the contradiction $(x^S(\alpha), u^S(\alpha)) = (x^*(\alpha), u^*(\alpha))$.

Assume that the conclusion is false for some fixed value of α in the interval mentioned in Lemma 4.1. Then $u_i^S(\alpha) \geqq 0$, $i \varepsilon \{S - B\alpha\}$, and applying Lemma 4.1 with $S^+\alpha = \{S - B\alpha\}$ one may assert the existence of a convex set $X' \supset X$ such that $x^S(\alpha)$ is an optimal solution of

$$\text{(1)} \quad \text{Maximize}_{x \varepsilon X'} \; f(x; \alpha) \quad \text{subject to} \quad \begin{aligned} g_i(x) &= 0, & i &\varepsilon \{B\alpha \cap S\} \\ g_i(x) &\geqq 0, & i &\varepsilon \{S - B\alpha\}. \end{aligned}$$

Using the supposition $g_i(x^S(\alpha)) \geqq 0$, $i \varepsilon \{A\alpha - S\}$, it follows that $x^S(\alpha)$ is

[3] In [2, Appendix *B*] counterexamples are presented for one of the simplest classes of problems subsumed under the present theory: f_1 and f_2 diagonally quadratic and the constraints linear. Although there are some "contrived" aspects to these counterexamples, the fact that they exist for such a simple class of problems seems to render it unlikely that perfect coincidence should obtain for more general classes of problems.

feasible in

$$\text{(2)} \quad \text{Maximize}_{x \varepsilon X'} \ f(x; \alpha) \quad \text{subject to} \quad g_i(x) = 0, \qquad i \ \varepsilon \ \{B\alpha \cap S\}$$
$$g_i(x) \geqq 0, \ i \ \varepsilon \ \{S - B\alpha\} \cup \{A\alpha - S\}.$$

Since the feasible region of (2) is included in that of (1), $x^S(\alpha)$ must be an optimal solution of (2).

It follows from the fact that $(x^*(\alpha), u^*(\alpha))$ satisfies $(=A\alpha)\alpha$ that $x^*(\alpha)$ is optimal in

$$\text{(3)} \quad \text{Maximize}_{x \varepsilon X'} \ f(x; \alpha) \quad \text{subject to} \quad g_i(x) \geqq 0, \qquad i \ \varepsilon \ A\alpha.$$

Since the feasible region of (2) is included in that of (3), and since $x^*(\alpha)$ is feasible in (2), $x^*(\alpha)$ must be optimal in (2). That is, both $x^*(\alpha)$ and $x^S(\alpha)$ are optimal in (2); thus, $f(x^*(\alpha); \alpha) = f(x^S(\alpha); \alpha)$. Now the optimality of $x^S(\alpha)$ for (3) follows from its feasibility for (3). Because of the strict concavity of $f(x; \alpha)$, (3) must have a unique optimal solution; therefore, $x^S(\alpha) = x^*(\alpha)$. This implies, by Condition 4, that $u^S(\alpha) = u^*(\alpha)$.

2.2 Rules for Determining the Order of Trials

Suppose that Step 2 has ended with the point of change $\alpha' < 1$. Designate the set of alarms which are given by S^0 (the set used during Step 2) as α increases above α' by T. Applying Theorem 4 at α', we know that at least one of the alarms is from the excess or deficiency of S^0 at $\alpha'+$. Unfortunately, we do not know which one. A logical way of proceeding at Step 3 is to modify S^0 by one constraint at a time for each constraint in T, i.e., try the sets $S^0 \pm i$ for each $i \ \varepsilon \ T$, where the symbol $S^0 \pm i$ means $S^0 \cup i$ if $i \not\varepsilon \ S^0$ and $S^0 - i$ if $i \ \varepsilon \ S^0$. This notation is designed to avoid having to distinguish between feasibility and optimality alarms. In other words, add the constraints which were feasibility alarms to S^0 and delete constraints which were optimality alarms from S^0 one at a time until each alarm has been heeded individually. Note that $S^0 \pm i$, $i \ \varepsilon \ T$, is valid at α' since all alarms caused by a set which is valid at α' must be from $B\alpha' - A\alpha'$.

When T has been exhausted by this *first generation* of trials, at least one trial set, say $S^0 \pm i_0$, is one unit of distance closer to a valid set at $\alpha'+$. If $d(S^0) = 1$, then $S^0 \pm i_0$ is valid at $\alpha'+$ (it yields no alarms), and Step 3 has been successfully completed. If $d(S^0) > 1$ then $d(S^0 \pm i_0) = d(S^0) - 1 > 0$, and a *second generation* of trials is necessary. At each first generation trial, let T_i denote the alarms due to $S^0 \pm i$, $i \ \varepsilon \ T$. At the second generation one should try $S^0 \pm i \pm j$ for each $i \ \varepsilon \ T$ and all $j \ \varepsilon \ T_i$. The symbol $S^0 \pm i \pm j$ means $\{S^0 \pm i\} \cup j$ if $j \not\varepsilon \ S^0 \pm i$ and $\{S^0 \pm i\} - j$ if $j \ \varepsilon \ S^0 \pm i$. Applying Theorem 4 at α' with $S = S^0 \pm i_0$, we see that at least one of the alarms due to $S^0 \pm i_0$ is from the excess or deficiency of $S^0 \pm i_0$ at $\alpha'+$, but we do not know which one. Hence at least one of the sets $S^0 \pm i_0 \pm j$, $j \ \varepsilon \ T_{i_0}$, is one unit of distance closer to a set which is valid at $\alpha'+$. Designate one such set by $S^0 \pm i_0 \pm j_0$. If $d(S^0) = 2$, then $S^0 \pm i_0 \pm j_0$ is valid at $\alpha'+$ (it yields no alarms), and Step

3 has been successfully completed. If $d(S^0) > 2$, then $d(S^0 \pm i_0 \pm j_0) = d(S^0) - 2 > 0$, and a *third generation* of trials is necessary.

The third generation of trials is constructed in a manner analogous to the preceding generations, and so on for the higher order generations. If, at any trial, a set is encountered which has been tried before, it may, of course, be discarded.

At each generation the distance from *some* trial set, and perhaps from several, to the collection of all sets which are valid at $\alpha' +$ is decreased by one unit. Since $d(S^0)$ is finite (it is bounded by the number of constraints in $B\alpha' - A\alpha'$ minus the number of constraints in $B\alpha'+ - A\alpha'+$), after a finite number of generations of trials a set which is valid at $\alpha'+$ will be obtained—after exactly $d(S^0)$ generations, in fact, The nearest valid set is, it will be recalled, S^0 plus its deficiency at $\alpha'+$ minus its excess at $\alpha'+$. In the author's experience $d(S^0)$ is usually very small, so that only a few trials are apt to be required.

Since the only modification of Step 3 being suggested here is a more complete specification of the order in which the trial sets are to be considered, and since this order has been shown to lead to a successful completion of Step 3, the assertions of the Basic Theorem still apply to the Basic Parametric Procedure with Step 3 modified as above.

2.3 An Alternative Ordering of Trials

The above rules for determining the order of trials at Step 3 have the advantage, when $d(S^0)$ is small (as is usually the case), of leading quite directly to the valid set nearest S^0. The disadvantage, however, is that they are cumbersome to program for a digital computer. This consideration, plus the fact that $d(S^0)$ can be relatively large when there are more than a half-dozen or so degenerate constraints at a point of change, has led the author to favor the use in this situation of a Markovian rather than deterministic scheme for ordering the trials: heed an alarm at random to determine the next trial set. This rule is very easy to program and is more economical in terms of computer storage. By a basic property of finite Markov chains it can easily be shown from Th. 4 that the number of trials before success will be finite with probability 1, and by an argument similar to the one in [4], one expects the average number of trials to be quite small. This expectation has been borne out in computational experience.

3. A Graphical Example

We shall illustrate the Basic Parametric Procedure and the application of Theorem 4 for the problem of maximizing, for $0 \leqq \alpha \leqq 1$, $\alpha \sum_{i=1}^{2} - (x_i - c'_i)^2 + (1 - \alpha)\sum_{i=1}^{2} - (x_i - c_i)^2$ subject to $a_i^t x + b_i \geqq 0$, $i = 1, \cdots, 4$. For convenience, the discussion is in graphical terms with extensive use being made of the fact that, since the constraints are linear, $(=S)\alpha$ comprises the first order conditions for a maximum of the objective function over $X_S \equiv \{x\colon a_i^t x + b_i = 0, i \varepsilon S\}$. From the circularity of the level curves of this particular objective function, it is evident that this constrained maximum is precisely the point of X_S nearest to the unconstrained maximum $x(\alpha) \equiv \alpha c' + (1 - \alpha)c$.

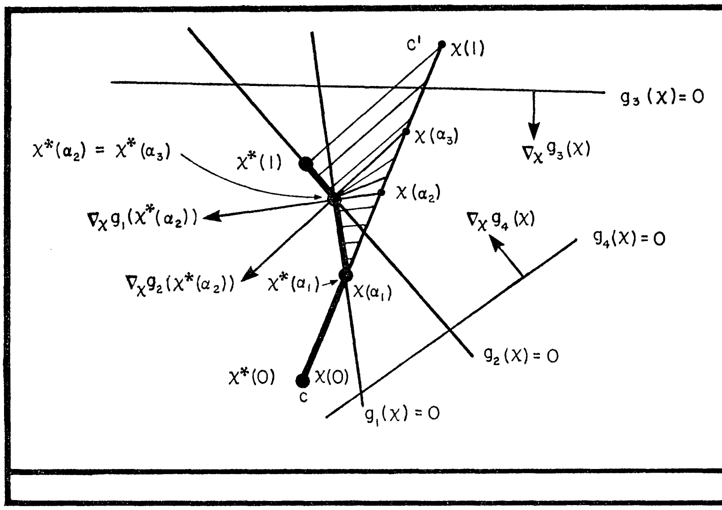

FIGURE 1

Figure 1 is drawn in x-space $(n = 2)$. The loci of $g_i(x) = 0, i = 1, \cdots, 4$, the unconstrained maximum $x(\alpha)$, and the constrained maximum $x^*(\alpha)$ (the heavy line) are drawn, as well as certain features pertaining to the points of change. Light lines representing the projection of $x(\alpha)$ onto the feasible region are also drawn; in view of the circularity of the level curves of the objective function for fixed α, these lines are in the direction of the gradient of the objective function at $x^*(\alpha)$. The gradients of the constraints point into the feasible region.

From $(=S)\alpha$ we see that the dual variables express minus the gradient of the objective function at $x^S(\alpha)$ as a linear combination of the gradients of the constraints in S. The signs of $u_i^S(\alpha)(i \varepsilon S)$ are easily determined by visual inspection of the figures.

At $\alpha = 0$, the unconstrained maximum $x(0)$ is interior to the feasible region. Thus, the constrained maximum $x^*(0)$ equals $x(0)$ and $B0 = \phi$, which implies that $A0 = \phi$ since $A\alpha \subset B\alpha$ for all α. We are obliged to let $S = \phi$, for the empty set is the only valid set at $\alpha = 0$. (Recall that S is valid at α if and only if $A\alpha \subset S \subset B\alpha$.) Step 1 is complete. Step 2 requires that we solve $(=\phi)\alpha$ as α increases above 0 until an alarm is given, i.e., until $x^\phi(\alpha)$ leaves the feasible region or $u_i^\phi(\alpha)$ becomes negative for some i. The last alternative (an optimality alarm) cannot occur for $S = \phi$, for $(=\phi)\alpha$ requires $u^\phi(\alpha) = 0$. Only the first alternative (a feasibility alarm) can occur. Equations $(=\phi)\alpha$ are easily seen to be the conditions for an unconstrained maximum. Since $x(0)$ is interior to the feasible region for $0 \leqq \alpha < \alpha_1$, no alarms are given on $[0, \alpha_1)$; $(x^\phi(\alpha), u^\phi(\alpha)) = (x^*(\alpha), u^*(\alpha)) = (x(\alpha), 0)$ and $A\alpha = B\alpha = \phi$ on $[0, \alpha_1)$. At α_1 the unconstrained maximum happens to be on the boundary of the feasible region, but beyond α_1 it violates the first constraint, i.e. $(=\phi)\alpha$ leads to a feasibility alarm for g_1 just above α_1. Thus, α_1 is the point of change which completes Step 2, and $(x^\phi(\alpha_1), u^\phi(\alpha_1)) = (x^*(\alpha_1), u^*(\alpha_1)) = (x(\alpha_1), 0)$,

$A\alpha_1 = \phi$, $B\alpha_1 = \{1\}$. Since $\alpha_1 < 1$, we go to Step 3. By Th.4, the first constraint must be in the deficiency of S; so we try $S = \{1\}$ and find that $\{1\}$ is valid above α_1. Control is now returned to Step 2 with $S = \{1\}$.

To execute Step 2 for the second time, we must solve $(=\{1\})\alpha$ as α increases above α_1 until an alarm obtains. These equations are the conditions for a maximum of the objective function subject to the first constraint being exactly satisfied. As α increases above α_1, $x^1(\alpha)$ moves along the portion of the boundary determined by the first constraint; since minus the gradient of the objective function at $x^1(\alpha)$ is expressed as $u_1^1(\alpha)$ times the gradient of g_1, it is geometrically clear that $u_1^1(\alpha)$ grows increasingly positive as α increases. Hence no alarms are given until α_2 is passed, when the second constraint begins to be violated. We have $x^1(\alpha) = x^*(\alpha)$, $u_1^1(\alpha) = u_1^*(\alpha) > 0$, $u_2^1(\alpha) = u_2^*(\alpha) = 0$, $A\alpha = B\alpha = \{1\}$ on (α_1, α_2). Since $\alpha_2 < 1$ is the point of change at which Step 2 is completed, we go to Step 3. By Th. 4, the second constraint must be in the deficiency of $\{1\}$ above α_2. Trying $S = \{1, 2\}$, we find that it is valid above α_2. Control is returned to Step 2 again, this time with $S = \{1, 2\}$.

Step 2 now requires that $(= \{1, 2\})\alpha$ be solved as α increases above α_2 until an alarm occurs. These equations are the conditions for a maximum of the objective function subject to both the first and second constraints being satisfied exactly. Since their intersection determines a unique point, $x^{1,2}(\alpha)$ is constant for all α. The projection lines of $x(\alpha)$ onto the feasible region and the interpretation of the dual variables make it clear that $u^{1,2}(\alpha) > 0$ on (α_2, α_3), $u_1^{1,2}(\alpha_3) = 0$, $u_2^{1,2}(\alpha_3) > 0$, and $u_1^{1,2}(\alpha) < 0$, $u_2^{1,2}(\alpha) > 0$ for $\alpha > \alpha_3$. In other words, an optimality alarm occurs for the first constraint just above α_3, so that Step 2 is complete at that point of change. Going to Step 3, we see from Th. 4 that the first constraint must be in the excess of $\{1, 2\}$ above α_3. Trying $S = \{2\}$, we find that it is valid above α_3 and return to Step 2 with $S = \{2\}$.

At Step 2, $(=\{2\})\alpha$ must be solved as α increases above α_3. Reasoning as before, we see that $\{2\}$ remains valid on $[\alpha_3, 1]$. Hence $x^2(\alpha) = x^*(\alpha)$, $A\alpha = B\alpha = \{2\}$, $u_1^2(\alpha) = 0$, and $u_2^2(\alpha) > 0$ on $(\alpha_3, 1]$.

This completes the solution of the example. A summary appears in Table 1.

4. Solving $(=S)\alpha$

In order to implement the Basic Parametric Procedure, it is necessary to have a method of actually solving $(=S)\alpha$ as α changes parametrically. Only in certain simple cases is it possible or economical to solve these equations analytically. Usually, numerical methods must be used. Variants of Newton's method [5, 7] can be an efficient means of solving $(=S)\alpha$ on a digital computer as α changes by small discrete jumps, although a number of other methods are also suitable for digital (or even analog) computation. We shall content ourselves with pointing out that not only does Newton's method apply in the usual sense, but also that the maximum permissible step size in α is bounded away from zero uniformly on [0, 1]. This is important in establishing a minimum step size for repeated applications of Newton's method.

Under Conditions 1 through 4, it is not difficult to show from standard results

TABLE 1

α	Valid Sets at α: S	Feasibility and Optimality Alarms Due to S Just Above α		Deficiency and Excess of S Just Above α	
		Feasibility	Optimality	Deficiency	Excess
$[0, \alpha_1)$	$\emptyset$	----	----	----	----
α_1	$\emptyset$	$\{1\}$	None	$\{1\}$	None
	$\{1\}$	None	None	None	None
(α_1, α_2)	$\{1\}$	----	----	----	----
α_2	$\{1\}$	$\{2\}$	None	$\{2\}$	None
	$\{1, 2\}$	None	None	None	None
(α_2, α_3)	$\{1, 2\}$	----	----	----	----
α_3	$\{1, 2\}$	None	$\{1\}$	None	$\{1\}$
	$\{2\}$	None	None	None	None
$(\alpha_3, 1]$	$\{2\}$	----	----	----	----

on the convergence of Newton's method [5, p. 136] that for each $\alpha_0 \epsilon [0, 1]$, Newton's method applied to $(=S)\alpha_0$ is well-defined and quadratically convergent to $(x^*(\alpha_0), u^*(\alpha_0))$ if S is valid at α_0 and if the starting point is in a sufficiently small neighborhood of $(x^*(\alpha_0), u^*(\alpha_0))$. Since $(x^*(\alpha), u^*(\alpha))$ is continuous, by taking $\Delta\alpha$ small enough $(x^*(\alpha_0 - \Delta\alpha), u^*(\alpha_0 - \Delta\alpha))$ is such a starting point. In other words, Newton's method is applicable point by point. It is not quite so clear that the size of the "neighborhoods of convergence" may be taken to be bounded away from zero uniformly on [0, 1]. Fortunately, it can be shown from continuity and compactness considerations that the neighborhoods do not become vanishingly small. The details of the proof [2, Sec. 3.1 of Ch. III, and App. C.1] are tedious, and are omitted here.

4.1 An Exemplary Algorithm

Figure 2 is a rudimentary flow chart of one possible implementation of the Basic Parametric Procedure. Its justification rests on the above remarks, the Basic Theorem, and the fact that for $\Delta\alpha$ and ϵ sufficiently small the results of section 2 apply each time a point of change is encountered.

We say "rudimentary" because, in practice, one would want to incorporate (i) a variable step-size feature in order to accelerate progress between points of change (assuming that optimal solutions for a coarser grid of values for α is acceptable), and (ii) partitioning, bordering, and refinement methods for the inverse matrix required by Newton's method (see [2, Appendix C] and [1, pp. 99–107]). The latter devices, especially, should be incorporated into any computer routine for this algorithm; without the substantial computational econo-

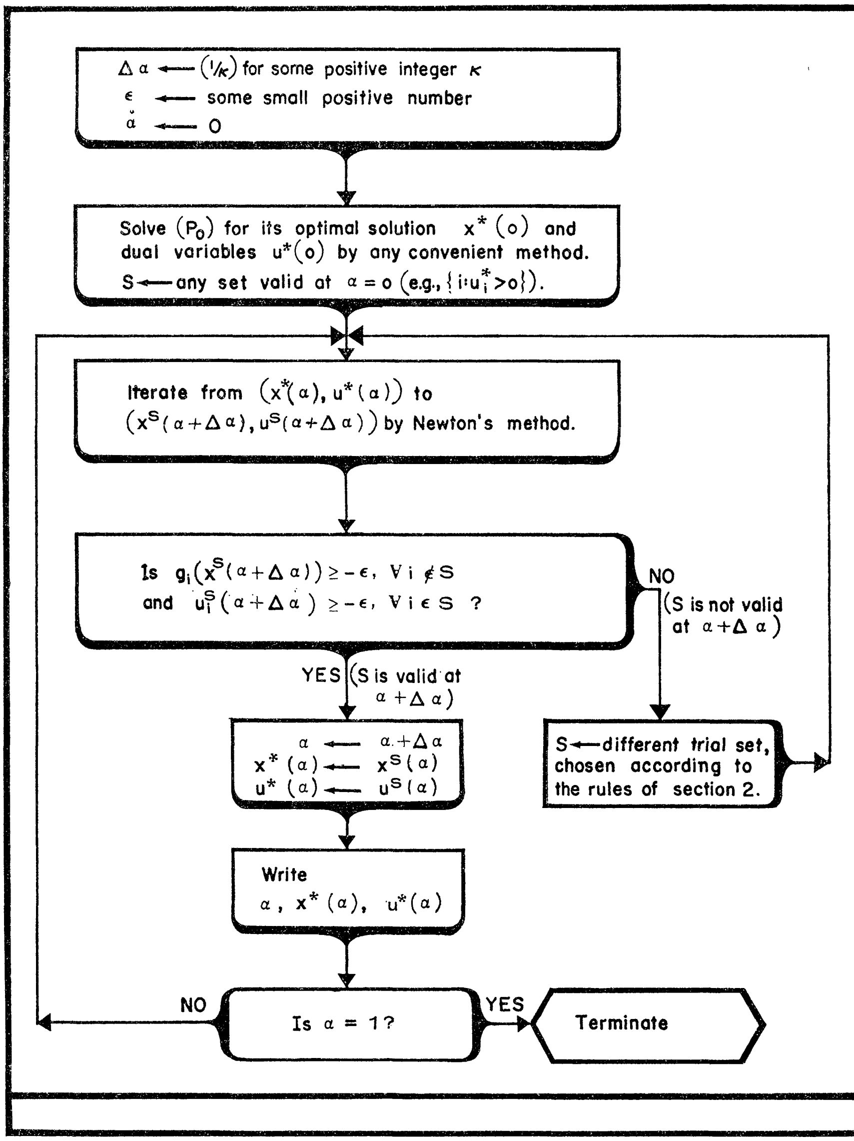

FIG. 2. Flow chart of a rudimentary computational algorithm

mies they make possible, the number and size of matrix inversions would preclude the use of Newton's method.

Variants of the algorithm have been programmed for a large digital computer, with gratifying results. Details of computational experience will be reported elsewhere.

4.2 Extrapolation

From $(=S)\alpha$, convenient expressions can easily be derived [2, p. 135] for the first derivatives of $(x^S(\alpha),\ u^S(\alpha))$ in terms of the inverse matrix required by

Newton's method. These can be used to facilitate the convergence of Newton's method, when large step sizes are desired, by extrapolating $(x^S(\alpha_0), u^S(\alpha_0))$ to a better approximation of $(x^S(\alpha_0 + \Delta\alpha), u^S(\alpha_0 + \Delta\alpha))$.

Although expressions for higher order derivatives can be obtained that permit $(x^S(\alpha), u^S(\alpha))$ to be approximated to any desired degree of accuracy in an interval about α_0, the use of second and higher order derivatives is probably computationally onerous.

5. Some Extensions

We briefly consider three extensions.

5.1 Linear Equality Constraints

If $(P\alpha)$ has L linear equality constraints, Condition 4 may not hold if they are expressed as $L + 1$ linear inequality constraints. This difficulty is easily overcome by a simple modification of the Basic Parametric Procedure: always include the linear equality constraints in S at Step 2 and in the trial sets at Step 3, and ignore any optimality alarms that such constraints may give. If *all* of the constraints happened to be linear equalities, Step 3 would disappear entirely.

5.2 Non-Strictly Concave f_i

In order to permit one of the f_i to be linear or concave but not *strictly* concave, it is of interest to observe that if Condition 3 is weakened to assert the local strict concavity of $\alpha f_1(x) + (1 - \alpha)f_2(x)$ on X only for α in some sub-interval I of [0, 1], then all of the results of this paper hold for α traversing any closed interval included by I. In this case, Condition 4 need only hold on I, of course.

5.3 More General Parametric Problems

With appropriate modifications of the four conditions, it can be shown that many of the results of this paper apply to any one-dimensional perturbation of

$$(Pp) \qquad \text{Maximize}_x \; f(x, p) \quad \text{subject to} \quad g(x, p) \geqq 0,$$

where the parameter $p = (p_1, \cdots, p_k)$ varies over a convex set P in E^k, $f(x, p)$ is continuous in (x, p) and strictly concave in x for each $p \varepsilon P$, and $g_i(x, p)$ $(i = 1, \cdots, m)$ is concave in (x, p). By a one-dimensional perturbation of (Pp), we mean a parametric problem of the form

$$\text{Maximize}_x \; f(x, p' + \alpha(p'' - p')) \quad \text{subject to} \quad g(x, p' + \alpha(p'' - p')) \geqq 0$$

for each value of $\alpha \varepsilon [0, 1]$, where $p', p'' \varepsilon P$.

Evidently, (Pp) is general enough to include many of the parametric problems of interest in sensitivity analysis of concave programming problems.

References

1. Faddeeva, V., *Computational Methods of Linear Algebra*, (C. D. Benster, translator), Dover Publications, New York, 1959.
2. Geoffrion, A., "A Parametric Programming Solution to the Vector Maximum Problem, with Applications to Decisions Under Uncertainty," Ph.D. Dissertation, Operations Research Program, Stanford University, February, 1965. Also issued as Working

Paper No. 68, Western Management Science Institute, University of California, Los Angeles.

3. ———, "Strictly Concave Parametric Programming, Part I: Basic Theory," *Management Science*, Vol. 13, No. 3 (Nov. 1966).
4. ———, "A Markovian Procedure for Strictly Concave Programming with Some Linear Constraints," to appear in the Proceedings of the Fourth International Conference on Operational Research at Cambridge, Massachusetts, Aug. 29–Sept. 2, 1966.
5. HOUSEHOLDER, A., *Principles of Numerical Analysis*, McGraw-Hill, New York, 1953.
6. MARKOWITZ, H., "The Optimization of a Quadratic Function Subject to Linear Constraints," *Naval Research Logistics Quarterly*, Vol. 3 (1956), pp. 111–133.
7. OSTROWSKI, A., *Solution of Equations and Systems of Equations*, Academic Press, New York, 1960.
8. THEIL, H. AND VAN DE PANNE, C., "Quadratic Programming as an Extension of Classical Quadratic Maximization," *Management Science*, Vol. 7 (1960), pp. 1–20.
9. ZAHL, S. "A Deformation Method for Quadratic Programming," *Journal of the Royal Statistical Society*, Series B, Vol. 26 (1964), pp. 141–160.