MANAGEMENT SCIENCE
Vol. 21, No. 5, January 1975
Printed in U.S.A.

NOTE ON FINITE CONVERGENCE OF EXTERIOR PENALTY FUNCTIONS*

KLAUS TRUEMPER†

University of Texas, Dallas

It is shown that existence of a saddle point of the Lagrangian function in an optimization problem is sufficient to assure finite convergence of the linear exterior penalty function. Also, an estimate of the penalty weight is given that yields ϵ-convergence for the quadratic exterior penalty function.

Introduction

Consider the problem:

$$\max: f(x) \tag{1}$$

$$\text{ST}: g_i(x) \geqq 0, \qquad i = 1, 2, \cdots, m. \tag{2}$$

Let the set of feasible solutions to (2) be $F = \{x \in E^n \mid g_i(x) \geqq 0; \forall i\}$. Assume problem (1)–(2) has a finite optimal solution, and that there are $\bar{x}$ and i such that $g_i(\bar{x}) < 0$. If the latter assumption does not hold, (1)–(2) reduces to an unconstrained maximization problem to which subsequent developments are not applicable.

One method of obtaining an optimal solution to (1)–(2) is via penalty functions ([1], [3], [4], [10]). Here we will only consider the exterior penalty function

$$C(x, r) = f(x) + P(x, r), \quad \text{where} \quad \forall r \geqq 0, \tag{3}$$

$$\begin{aligned} P(x, r) &= 0, & x &\in F, \\ &< 0, & x &\notin F, \\ \lim_{r\to\infty} P(x, r) &= -\infty, & \forall x &\notin F. \end{aligned} \tag{4}$$

In the exterior penalty function method as described in [10], a sequence of unconstrained maximization problems max: $C(x, r_k)$ are solved, where $\{r_k\}$ is a positive, monotonically increasing, unbounded sequence of real numbers. The existence of the maximum for $C(x, r_k)$ will here be presumed for all values of k. For practical applications, continuity of $C(x, r_k)$ in x and restriction of search to a sufficiently large closed and bounded subset of E^n are sufficient to assure existence of the maximum for all values of k. If the optimal x^k of max: $C(x, r_k)$ is feasible for (2), then it is also optimal for (1)–(2). However, it may occur that convergence to an optimal solution of (1)–(2) is achieved only in the limit as $k \to \infty$.

In the following, it is shown for the linear exterior penalty function ($P(x, r) = r\sum_{i=1}^{m} \min(0, g_i(x))$) that existence of a saddle point of the Lagrangian function arising from (1)–(2) is sufficient to assure convergence with finite k, or equivalently, with finite $r = r_k$. Further, an estimate of r is given that yields ϵ-convergence for the quadratic exterior penalty function ($P(x, r) = -r\sum_{i=1}^{m} [\min(0, g_i(x))]^2$).

* Processed by Professor Willard I. Zangwill, Departmental Editor for Linear and Nonlinear Programming and Associate Editor Anthony V. Fiacco; received February 8, 1973, revised October 19, 1973. This paper has been with the author three months for revision.

† The author wishes to thank the referee for the careful review of the initial draft and for helpful suggestions.

Previous work on finite convergence of the linear exterior penalty function by Zangwill [10] and Pietrzykowski [9] is strongly related to results given here, and a detailed discussion of the relationships is given in the subsequent section. Recent results by Evans, Gould and Tolle [2] prove finite convergence for two classes of exterior penalty functions. The conditions for finite convergence are quite general, but it appears to be rather difficult to estimate the required value of r. In all of the above work, the exterior penalty functions are only piecewise continuously differentiable. Fletcher [5] has provided continuously differentiable exterior penalty functions for equality constrained problems, and in [6] for problems of type (1)–(2). In the latter case, the function requires solution of a concave program for any point.

Linear Exterior Penalty Function

This section presents convergence results concerning the linear exterior penalty function, where $P(x, r) = r\sum_{i=1}^{m} \min(0, g_i(x))$. It is shown that existence of a saddle point of the Lagrangian function arising from (1)–(2) is sufficient for convergence with finite penalty weight r.

Definition 1 ([2]). *$C(x, r)$ of (3) is exact if there is a finite r_0 such that* $\max C(x, r_0) = f(x^0)$, *where x^0 solves (1)–(2).*

Proposition 1. *$C(x, r)$ of (3) is exact iff there is a point (x^*, r_*) for which*

$$C(x^*, r) = C(x^*, r_*) \geqq C(x, r_*); \qquad \forall r \geqq 0; \qquad \forall x \in E^n.$$

Proof. For the "if" part, note that $C(x^*, r) = C(x^*, r_*)$, $\forall r \geqq 0$, implies that x^* is feasible. By $C(x^*, r_*) \geqq C(x, r_*)$, $\forall x \in E^n$, x^* must be optimal. Choosing $r_0 = r_*$, $\max C(x, r_0) = f(x^*) = f(x^0)$, where x^0 is any optimal solution of (1)–(2). Conversely, let $(x^*, r_*) = (x^0, r_0)$. Since x^0 is feasible, $C(x^*, r) = C(x^*, r_*)$, $\forall r \geqq 0$. Then $f(x^*) = C(x^*, r_*) = \max C(x, r_*)$, $\forall x \in E^n$, which completes the "only if" part. Q.E.D.

Definition 2. *$C(x, r)$ of (3) exhibits finite convergence if it is exact, and if any x^* which maximizes $C(x, r_0)$, with r_0 given in Definition 1, also solves (1)–(2).*

Clearly exactness (finite convergence) for some r_0 implies exactness (finite convergence) $\forall r \geqq r_0$. Moreover, exactness implies finite convergence under a very weak condition, as shown in

Proposition 2. *Let $C(x, r)$ of (3) be exact, and suppose r_0 assures* $\max C(x, r_0) = f(x^0)$, *where x^0 solves (1)–(2). If there is $r_* > r_0$ such that $P(x, r_*) < P(x, r_0)$, $\forall x \notin F$, then $C(x, r)$ has finite convergence.*

Proof. With the above definitions of r_0, r_*, x^0, $C(x^0, r_*) = C(x^0, r_0) \geqq C(x, r_0) > C(x, r_*)$, $\forall x \notin F$. So any solution of $\max C(x, r_*)$ must be in F, hence it must solve (1)–(2). Q.E.D.

An immediate consequence of Definition 2 is

Proposition 3. *Finite convergence with $r = r_0$ implies that the set of points which maximize $C(x, r_0)$ coincides with the set of solutions to (1)–(2).*

Proof. Let x^* maximize $C(x, r_0)$, and suppose x^0 solves (1)–(2). By Definition 2, x^* also solves (1)–(2). Since then

$$C(x^*, r_0) = f(x^*) = f(x^0) = C(x^0, r_0),$$

x^0 must also maximize $C(x, r_0)$. Q.E.D.

Before providing an equivalent characterization of finite convergence, we introduce

Assumption 1. $P(x, r)$ meets

(a) $P(x, 0) = 0; \forall x \in E^n$.

(b) $P(x, r)$ is twice continuously differentiable in r, and $\forall r \geqq 0$ and $x \notin F$

$$\frac{\delta P(x, r)}{\delta r} \equiv P'(x, r) < 0;$$

$$\frac{\delta^2 P(x, r)}{\delta r^2} \equiv P''(x, r) \leqq 0.$$

The above assumption is met by most if not all exterior penalty functions which are used in practical applications.

PROPOSITION 4. *Suppose Assumption* 1 *holds. If there exists* $y \in F$ *such that* $\sup_{x \notin F} [f(y) - f(x)]/P'(x, 0)$ *is finite, then finite convergence is assured for* $C(x, r)$. *Conversely, assume finite convergence of* $C(x, r)$ *and* $P''(x, r) = 0, \forall r \geqq 0$ *and* $x \notin F$. *Then* $\sup_{x \notin F} [f(x^0) - f(x)]/P'(x, 0)$ *is finite, where* x^0 *is any optimal solution of* (1)–(2).

PROOF. For the first part, assume $\sup_{x \notin F} [f(y) - f(x)]/P'(x, 0) < r_0$ for some $y \in F$; also let x^0 be optimal for (1)–(2). Then

$$f(x^0) + P(x^0, r_0) = f(x^0) \geqq f(y) > f(x) + r_0 P'(x, 0), \qquad \forall x \notin F.$$

Using Taylor Series expansion about $r = 0$ and Assumption 1,

$$f(x) + r_0 P'(x, 0)$$
$$\geqq f(x) + P(x, 0) + r_0 P'(x, 0) + (r_0^2/2) P''(x, \theta \cdot r_0) = C(x, r_0); \, \forall x \notin F$$

where $0 < \theta < 1$. In conclusion, $C(x^0, r_0) > C(x, r_0)$; $\forall x \notin F$, and maximization of $C(x, r_0)$ must yield an optimal solution to (1)–(2). As for the second part, suppose finite convergence exists with $r = r_0$. By Proposition 3, any optimal solution x^0 of (1)–(2) also maximizes $C(x, r_0)$; i.e., $f(x^0) = f(x^0) + P(x^0, r_0) \geqq f(x) + P(x, r_0)$; $\forall x \in E^n$. Since $P(x, 0) = P''(x, r) = 0, \forall r \geqq 0$ and $x \notin F$, Taylor Series expansion about $r = 0$ yields $P(x, r_0) = r_0 P'(x, 0), \forall x \notin F$. Combining results,

$$[f(x^0) - f(x)]/P'(x, 0) \leqq r_0, \qquad \forall x \notin F.$$

Q.E.D.

Obviously, verification of the above condition for finite convergence is not easy for nontrivial problems. Before turning to the special case with

$$P(x) = \textstyle\sum_{i=1}^{m} \min (0, g_i(x)),$$

we need

DEFINITION 3. *A point* $(x^0, u^0), u^0 \geqq 0$, *is a saddle point of the Lagrangian function*

$$L(x, u) = f(x) + \textstyle\sum_{i=1}^{m} u_i g_i(x) \tag{5}$$

if

$$L(x^0, u) \geqq L(x^0, u^0) \geqq L(x, u^0); \quad \forall u \geqq 0; \quad \forall x \in E^n. \tag{6}$$

A list of sufficient conditions for the existence of saddle points in concave programs may be found in [8, pp. 81, 87].

The main result concerning linear exterior penalty functions can now be stated.

LEMMA 1. *Let* (x^0, u^0) *be a saddle point of* $L(x, u)$ *of* (5). *Then finite convergence exists if* $P(x, r) = r \cdot \sum_{i=1}^{m} \min (0, g_i(x))$.

PROOF. First note that $P(x, r)$ meets Assumption 1, and that the saddle point (x^0, u^0) exists only if x^0 is optimal for (1)–(2), and $u_i^0 g_i(x^0) = 0, \forall i$. For proof, see [8, p. 84].

Then, choosing

$$r_0 > \max_i u_i^0, \tag{7}$$

$$f(x^0) + \sum_{i=1}^m u_i^0 g_i(x^0) = f(x^0) \geqq f(x) + \sum_{i=1}^m u_i^0 g_i(x) > f(x) + P(x, r_0); \quad \forall x \notin F,$$

where the first inequality arises from (6), and the second (strict) inequality from (7) and the fact that $g_i(x) \geqq \min(0, g_i(x))$; $\forall x \in E^n$, with x being restricted to infeasible values.

Therefore, we get

$$[f(x^0) - f(x)]/P'(x, 0) < r_0; \quad \forall x \notin F,$$

and by Proposition 4, finite convergence is assured. Q.E.D.

Note that the existence of the saddle point is sufficient, but not necessary. Lemma 1 covers the case of $f(x)$, $g_i(x)$ concave $\forall i$, and existence of $\bar{x} \ni g_i(\bar{x}) > 0$, $\forall i$, for which a finite convergence proof is given in [10], since these conditions are sufficient for the existence of a saddle point [7, p. 239].

A result by Pietrzykowski [9] can also be proved by Lemma 1. Defining $P(x, r)$ as in Lemma 1, he shows that for sufficiently large finite r_0 a strong local maximum x^0 of (1)–(2) is also a strong local maximum of $C(x, r)$, $\forall r \geqq r_0$, provided $f(x)$ and the $g_i(x)$ are continuously differentiable in a neighborhood of x^0, and the gradients of the binding constraints are linearly independent. If in addition $f(x)$ and the $g_i(x)$ are concave, then finite convergence is assured to the strong local maximum x^0 of (1)–(2). The first result cannot be derived from Lemma 1 since under the assumptions given a saddle point cannot be established. However, with the added condition of concavity Fiacco and McCormick [4, pp. 19, 22, 90] show that there must exist a vector $u^0 \geqq 0$ such that (x^0, u^0) is a saddle point of the Lagrangian function $L(x, u)$. Lemma 1 then guarantees finite convergence to the strong global maximum x^0.

Estimation of r_0 of (7) is by no means easy. However, if there exist a point $\bar{x} \ni g_i(\bar{x}) > 0$, $\forall i$, and an upper bound z of $f(x^0)$, where x^0 is optimal for (1)–(2), then an estimate of r_0 is provided by

$$\bar{r} > \frac{z - f(\bar{x})}{\min_i g_i(\bar{x})}. \tag{8}$$

This same estimate is also used in [10]. To see that $\bar{r} > \max_i u_i^0$, note that

$$f(\bar{x}) + \sum_{i=1}^m u_i^0 g_i(\bar{x}) \leqq f(x^0) \leqq z < f(\bar{x}) + \bar{r} \min_i g_i(\bar{x}) \tag{9}$$

where the leftmost inequality arises from (6), and the remaining ones from (8) and the fact that z is an upper bound of $f(x^0)$. From (9),

$$\bar{r} > \sum_{i=1}^m u_i^0 a_i \geqq \sum_{i=1}^m u_i^0 \geqq \max_i u_i^0,$$

where $a_i = g_i(\bar{x})/\min_j g_j(\bar{x}) \geqq 1$, $\forall i$, proving that $\bar{r}$ of (8) may be used to achieve finite convergence.

Quadratic Exterior Penalty Function

In this section results concerning the convergence of the quadratic exterior penalty function, where $P(x, r) = -r\sum_{i=1}^m [\min(0, g_i(x))]^2$, are discussed. For arbitrary $\epsilon > 0$, an estimate of the penalty weight, r_ϵ, is provided so that any x^* which maximizes $C(x, r_\epsilon)$ meets all constraints of (2) within the specified ϵ-tolerance.

For practical applications, penalty terms of the form

$$P(x, r) = -r\sum_{i=1}^m |\min(0, g_i(x))|^{1+\alpha}, \qquad 0 < \alpha \leqq 1,$$

are often used. Such penalty terms have the advantage of yielding a differentiable $C(x, r)$, provided $f(x)$ and $g_i(x)$ are differentiable. On the other hand, finite convergence of Definition 2 is then frequently lost, even if $L(x, u)$ of (5) has a saddle point.

To establish another type of convergence for the above mentioned penalty functions, we need first

DEFINITION 4. *ϵ-convergence of $C(x, r)$ of (3) occurs whenever, for any $\epsilon > 0$, a finite r_ϵ can be specified a priori such that $g_i(x^*) \geqq -\epsilon$, $\forall i$, for any x^* which maximizes $C(x, r_\epsilon)$.*

It should be noted that the x^* of the above definition need not be the optimal solution to

$$\text{(10)} \qquad \max: f(x)$$

$$\text{(11)} \qquad \text{ST:} \quad g_i(x) \geqq -\epsilon, \qquad i = 1, 2, \cdots, m.$$

If that was the case, one could incorporate the ϵ into the original problem (1)–(2), and Definition 4 would be superfluous. However, it is true that the x^* of Definition 4 is the optimal solution to

$$\text{(12)} \qquad \max: f(x)$$

$$\text{(13)} \qquad \text{ST:} \quad g_i(x) \geqq \min\ (0, g_i(x^*)) \geqq -\epsilon, \qquad i = 1, 2, \cdots, m.$$

To see this, let $\bar{x}$ be feasible for (13), and assume $\alpha \geqq 0$. Then

$$\textstyle\sum_{i=1}^m |\min\ (0, g_i(x^*))|^{1+\alpha} \geqq \sum_{i=1}^m |\min\ (0, g_i(\bar{x}))|^{1+\alpha},$$

and since

$$f(x^*) - r_\epsilon \textstyle\sum_{i=1}^m |\min\ (0, g_i(x^*))|^{1+\alpha} \geqq f(\bar{x}) - r_\epsilon \sum_{i=1}^m |\min\ (0, g_i(\bar{x}))|^{1+\alpha},$$

we have $f(x^*) \geqq f(\bar{x})$, thus proving the above assertion.

Subsequently, we will restrict our discussion to penalty terms with $\alpha = 1$, since that case has been researched extensively, yielding many useful theoretical results and computational approaches (for details, see [4]).

THEOREM 1. *Assume that $L(x, u)$ of (5) has a saddle point (x^0, u^0), and that $P(x, r) = -r\sum_{i=1}^m [\min\ (0, g_i(x))]^2$. Then ϵ-convergence occurs if r_ϵ is chosen as*

$$\text{(14)} \qquad r_\epsilon = r_0/(\epsilon/h),$$

where r_0 is defined by (7), and

$$\text{(15)} \qquad h = (m^{1/2} + 1)/2.$$

PROOF. Presume x^* maximizes $C(x, r_\epsilon)$ (with the quadratic penalty term), and finite ϵ-convergence is not achieved. Let

$$I = \{i \mid g_i(x^*) < -\epsilon\},$$
$$J = \{i \mid -\epsilon \leqq g_i(x^*) \leqq 0\}.$$

By assumption $I \neq \emptyset$. Now

$$\text{(16)} \qquad \begin{aligned} &f(x^*) - r_\epsilon \textstyle\sum_{i=1}^m [\min\ (0, g_i(x^*))]^2 \\ &\qquad = f(x^*) - r_\epsilon \textstyle\sum_{i\in I} [g_i(x^*)]^2 - r_\epsilon \sum_{i\in J} [g_i(x^*)]^2. \end{aligned}$$

Examining the first penalty term on the r.h.s. of (16), we see that

$$\text{(17)} \quad r_\epsilon \textstyle\sum_{i\in I} [g_i(x^*)]^2 = r_0 \sum_{i\in I} |g_i(x^*)| + r_0 \sum_{i\in I} [(g_i(x^*))^2/(\epsilon/h) - |g_i(x^*)|].$$

To simplify notation, we let $y = | g_i(x^*) |$ for some $i \in I$. Then define

$$d(y) = y^2/(\epsilon/h) - y > 0; \quad \forall y > \epsilon/h.$$

Moreover, $d(y)$ is monotone increasing $\forall y > \epsilon/h$, so

$$d(y) > \epsilon^2/(\epsilon/h) - \epsilon = \epsilon(h - 1); \quad \forall y > \epsilon \geqq \epsilon/h$$

since $h \geqq 1$ by (15). Using this result in (17) yields

$$r_\epsilon \textstyle\sum_{i \in I} [g_i(x^*)]^2 > r_0 \sum_{i \in I} | g_i(x^*) | + r_0 \epsilon (h - 1). \tag{18}$$

The second penalty term on the r.h.s. of (16) can be written as

$$r_\epsilon \textstyle\sum_{i \in J} [g_i(x^*)]^2 = r_0 \sum_{i \in J} | g_i(x^*) | + r_0 \sum_{i \in J} [(g_i(x^*))^2/(\epsilon/h) - | g_i(x^*) |]. \tag{19}$$

Now let $y = | g_i(x^*) |$ for some $i \in J$. Then use again

$$d(y) = y^2/(\epsilon/h) - y \geqq -\epsilon/(4h), \qquad 0 \leqq y \leqq \epsilon,$$

the lower bound being achieved at $y = \epsilon/(2h)$. Applying this to (19), we get, since the cardinality of J is $\leqq m - 1$,

$$r_\epsilon \textstyle\sum_{i \in J} [g_i(x^*)]^2 \geqq r_0 \sum_{i \in J} | g_i(x^*) | - r_0(m - 1)\epsilon/(4h). \tag{20}$$

Equations (18) and (20) may now be used in (16). Then

$$\begin{aligned} f(x^*) - r_\epsilon \textstyle\sum_{i=1}^m [\min (0, g_i(x^*))]^2 &< f(x^*) + r_0 \textstyle\sum_{i=1}^m | \min (0, g_i(x^*)) | \\ &\quad + r_0[(m-1)\epsilon/(4h) - \epsilon(h-1)] \\ &< f(x^*) + r_0 \textstyle\sum_{i=1}^m | \min (0, g_i(x^*)) | \\ &< f(x^0); \quad x^0 \text{ is optimal for (1)–(2)}. \end{aligned}$$

The second inequality is justified by $(m - 1)\epsilon/(4h) - \epsilon(h - 1) = 0$ using (15), and the last inequality is just an application of the preceding Lemma 1. Since x^* maximizes $C(x, r_\epsilon)$, the assumption of $I \neq \varnothing$ must then be false. Q.E.D.

The estimation procedure for r_0 is also useful to get an estimate $\bar{r}_\epsilon$ of r_ϵ. That is,

$$\bar{r}_\epsilon = \bar{r}/(\epsilon/h),$$

where $\bar{r}$ is defined by (8).

An *a priori* estimate of the difference between $f(x^*)$ and $f(x^0)$, where x^* maximizes $C(x, r_\epsilon)$ and x^0 solves (1)–(2), can also be obtained, provided r_0 of (7) or $\bar{r}$ of (8) is known. Let x^ϵ be an optimal solution to (10)–(11). Since x^* solves (12)–(13),

$$f(x^\epsilon) \geqq f(x^*) \geqq f(x^0).$$

We now make the reasonable assumption that for small $\epsilon > 0$ the variables u_i^0 of the saddle point (x^0, u^0) can be used to describe the change of optimal solution value when problem (1)–(2) is changed to (10)–(11) (for a detailed discussion, see, e.g., [4, p. 34]). Then by (7)

$$\begin{aligned} f(x^*) - f(x^0) \leqq f(x^\epsilon) - f(x^0) \simeq \textstyle\sum_{i=1}^m u_i^0 \epsilon &\leqq m r_0 \epsilon \\ &\leqq m \bar{r} \epsilon \end{aligned}$$

where $\bar{r}$ is defined by (8).

It remains to be shown how z in (8) may be obtained. Suppose that x^* maximizes $C(x, r)$, where r is an initially selected penalty weight (usually based on previous computational experience). Then, with x^0 solving (1)–(2), $C(x^*, r) \geqq C(x^0, r) = f(x^0)$, so $z = C(x^*, r)$ may be used. If $f(x)$ and the $g_i(x)$ are concave continuously differentiable functions, and if the set of points solving (1)–(2) is compact, then $z =$

$f(x^*) - 2r\sum_{i=1}^m [\min(0, g_i(x^*))]^2$ is a tighter upper bound on $f(x^0)$ than $C(x^*, r)$ with quadratic penalty term. Details and justification for this bound may be found in [4, p. 105]. In conclusion then, knowledge of a point $\bar{x}$ for which $g_i(\bar{x}) > 0$, $\forall_i$, and at most one maximization of $C(x, r)$ are required to compute $\bar{r}$ or $\bar{r}_\epsilon$.

The basic idea in Theorem 1 can be slightly generalized to penalty functions of the form

$$P(x, r) = r\sum_{i=1}^m \{q_i[\min(0, g_i(x))]\}^2. \tag{21}$$

For the functions $q_i(\cdot)$ of (21) we need

Assumption 2. *$q_i(\cdot)$ is strictly monotone increasing, and $q_i(0) = 0$, $\forall i$.*

Theorem 2. *Suppose $\bar{C}(x, r) = f(x) + r\sum_{i=1}^m q_i[\min(0, g_i(x))]$ achieves finite convergence with $r = r_0$, and Assumption 2 holds. Then $C(X, r)$, with $P(X, r)$ defined by (21), has ϵ-convergence when r_ϵ is chosen as*

$$r_\epsilon = r_0/(\min_i \epsilon_i/h)$$

where h is defined by (15) and

$$\epsilon_i = -q_i(-\epsilon); \qquad \forall i.$$

Proof. Let x^* maximize $C(x, r_\epsilon)$ (with $P(x, r_\epsilon)$ of (21)). Under Assumption 2, $q_i[\min(0, g_i(x^*))] = \min(0, q_i[g_i(x^*)])$ for any i. So replacing "$g_i(x^*)$" by "$q_i[g_i(x^*)]$" in the proof of Theorem 1, and assuming finite convergence for $\bar{C}(x, r)$ instead of using Lemma 1, we must have

$$q_i[\min(0, g_i(x^*))] = \min(0, q_i[g_i(x^*)]) \geqq -\min_i \epsilon_i \geqq -\epsilon_i.$$

The strict monotonicity of the $q_i(\cdot)$ then implies $g_i(x^*) \geqq -\epsilon$; $\forall i$. Q.E.D.

A remark about the value of Lemma 1 and Theorems 1 and 2 seems to be in order. In planning the sequence $\{r_k\}$, it certainly is useful to know an upper bound r_0 (r_ϵ) for which finite convergence (ϵ-convergence) is assured. Also, the relationship for r_ϵ shows how severely tight ϵ-requirements may increase the r_ϵ value that assures acceptable feasibility.

References

1. Carroll, C. W., "The Created Response Surface Technique for Optimizing Nonlinear Restraint Systems," *Operations Research*, Vol. 9, No. 2 (1961), pp. 169–184.
2. Evans, J. P., Gould, F. J. and Tolle, J. W., "Exact Penalty Functions in Nonlinear Programming," *Mathematical Programming*, Vol. 4, No. 2 (1973), pp. 72–97.
3. Fiacco, A. V. and McCormick, G. P., "The Sequential Unconstrained Minimization Technique for Nonlinear Programming. A Primal-Dual Method," *Management Science*, Vol. 10, No. 2 (1964), pp. 360–366.
4. —— and ——, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley and Sons, Inc., New York, 1968.
5. Fletcher, R., "A Class of Methods for Nonlinear Programming with Termination and Convergence Properties" in *Integer and Nonlinear Programming*, J. Abadie (ed.), North-Holland, Amsterdam, 1970.
6. ——, "An Exact Penalty Function for Nonlinear Programming with Inequalities," T. P. 478 Atomic Energy Research Establishment, Harwell, February 1972.
7. Karlin, S., *Mathematical Methods and Theory in Games, Programming and Economics*, Vol. 1, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1959.
8. Lasdon, L. S., *Optimization Theory for Large Systems*, Macmillan Company, New York, 1970.
9. Pietrzykowski, T., "An Exact Potential Method for Constrained Maxima," *SIAM J. Numer. Anal.*, Vol. 6, No. 2 (1968), pp. 299–304.
10. Zangwill, W. I., "Non-Linear Programming via Penalty Functions," *Management Science*, Vol. 13, No. 5 (1967), pp. 344–358.