THE SEQUENTIAL UNCONSTRAINED MINIMIZATION TECHNIQUE (SUMT) WITHOUT PARAMETERS

Anthony V. Fiacco and Garth P. McCormick

Research Analysis Corporation, McLean, Virginia

(Received May 10, 1965)

An unconstrained minimization technique for solving nonlinear programming problems that involves no parameter selection is presented. The algorithm generates a sequence of interior feasible points with decreasing objective function values that converge to the optimum.

WE SHALL consider the problem of determining $\bar{x}$ that solves:

$$(A): \text{minimize } f(x)$$

subject to $$g_i(x) \geqq 0. \qquad (i=1, 2, \cdots, m)$$

The Sequential Unconstrained Minimization Technique (SUMT)[1,2,3,4] for solving (A) is based on minimizing

$$P(x, r_k) \equiv f(x) + r_k \sum_{i=1}^{i=m} 1/g_i(x)$$

in the interior of the feasible region over a monotonic decreasing sequence $\{r_k\}$, where $r_k > 0$, all k. A sequence of points $\{x(r_k)\}$, $k=0, 1, 2, \cdots$, are thus generated that respectively minimize $\{P(x, r_k)\}$. It follows that as $r_k \to 0$ $(k \to \infty)$, $x(r_k) \to \bar{x}$, the solution of (A), under conditions that correspond precisely to those that will be (required in the present modification and) specified in the next section.

It is reasonable to suppose that numerous variations of the P function could be constructed for converting (A) into a sequence of unconstrained problems, the solutions of which converge in the limit to $\bar{x}$. The class of such functions will be narrowed considerably if we insist further that the principal advantages of the P function be retained. These, briefly stated, are that: (1) P is convex over x satisfying the constraints of (A), if (A) is a convex programming problem; (2) a (Wolfe) dual-feasible point, as well as a primal-feasible point, is available with each minimization of $P(x, r)$, i.e., for each value of the parameter r; (3) a subproblem of (A) is solved with each minimization of P; and (4) the objective function $f(x)$ decreases monotonically from one minimum of P to the next, i.e., as $r \to 0$.

This is a fairly imposing list of requirements, but the computational efficacy of the P-function technique is due largely to these attributes. It is difficult to imagine that a 'penalty function' technique will be (computationally) competitive without fulfilling at least equations (1), (2), and (4).

The purpose of the present paper is to introduce another function that has the desired characteristics as mentioned. It is closely related to the P function, but is a nontrivial variant of that function. Furthermore, it has at least two features the P function does not possess and hence may warrant further development.

Define the function

$$Q(x, x^k) \equiv \{1/[f(x^k)-f(x)]\} + \sum_{i=1}^{i=m} [1/g_i(x)],$$

where x^0 is such that $g_i(x^0) > 0$, all i. The basic idea is to minimize $Q(x, x^0)$ over $\{x | f(x) \leqq f(x^0),\ g_i(x) \geqq 0, \text{ all } i\}$, assuming this set has a nonempty interior. Under suitable conditions, this will yield uniquely the point x^1. Replacing x^0 by x^1, the process is repeated. A sequence of points $\{x^k\}$ is thus defined, and it follows under conditions analogous to those required for the P function, that $x^k \rightarrow \bar{x}$ as $k \rightarrow \infty$.

A general statement of a class of techniques, called, "Methods of Centers," that proceed in a manner similar to that just stated is contained in reference 5.

The next section lists the required conditions and gives the proof of convergence for convex programming problems.

In the third section, the precise connection between the points $\{x^{k+1}\}$ respectively minimizing $\{Q(x, x^k)\}$, $k=0, 1, 2, \cdots$, and the points $\{x(r_k)\}$ respectively minimizing $\{P(x, r_k)\}$, is made explicit. It follows that the $\{x^{k+1}\}$ are a subsequence of the $\{x(r_k)\}$. Thereafter, of course, all of the development pertinent to the P function is directly applicable to the Q function.

The possible advantages of the Q function over the P function are indicated in the final section. These are that $f(x)$ is decreased monotonically *from the starting point*, whereas in the P function, the initial value of r is arbitrary, and hence the objective function may increase substantially from the starting point to the point minimizing P. The other advantage of the Q function is its independence of parameter adjustment: the process is determined in its entirety once the initial point is specified. Again, the rate of reduction of r in the P function is quite arbitrary. An algorithm based on the Q function has been applied with success, but the method has not as yet been fully exploited or adequately tested.

THE CONVERGENCE THEOREM

WE DEFINE $R \equiv \{x | g_i(x) \geqq 0,\ i=1, \cdots, m\}$ and $R_k \equiv \{x | f(x) \leqq f(x^k)\}$, $k=0, 1, 2, \cdots$. The interior of these sets will be denoted respectively by R^0 and R_k^0.

The following conditions will be assumed:

C1. $R^0 \neq \phi$.

C2. $f(x), g_1(x), \cdots, g_m(x)$ are continuously differentiable.

C3. For every finite k, $R \cap R_k$ is bounded.

NOTE. Conditions C1–C3 imply the existence of a finite number v_0, where $v_0 = \inf_{x \in R} f(x) = \min_{x \in R} f(x)$.

C4. $f(x)$ is convex.

C5. $g_1(x), \cdots, g_m(x)$ are concave.

NOTE. Conditions C4 and C5 imply the convexity of $Q(x, x^k)$ in $R^0 \cap R_k^0$. This is the direct result of a fact easily proved that if $h(x)$ is concave, then $1/h(x)$ is convex over $\{x | h(x) > 0\}$.

The following condition is assumed merely for convenience, to assure the uniqueness of any existing minimum of $Q(x, x^k)$ in $R^0 \cap R_k^0$.

C6. The function $Q(x, x^k) \equiv \{1/[f(x^k) - f(x)]\} + \sum_{i=1}^{i=m} [1/g_i(x)]$ is, for $k = 0, 1, 2, \cdots$, strictly convex for $x \in R^0 \cap R_k^0$.

NOTE. The strict convexity follows if, for example, f or any $-g_i(x)$ is strictly convex, or if any set of n independent linear constraints, such as the non negativity restrictions, is included in the problem.

THEOREM 1. *If* $x^k \in R^0$, *then* C1–C6 *imply that*

(a) *if* $R^0 \cap R_k^0 \neq \phi$, $Q(x, x^k)$ *is minimized over* $R^0 \cap R_k^0$ *at a unique point* x^{k+1} *where*

$$\nabla Q(x^{k+1}, x^k) = 0, \tag{1}$$

and

(b) $\lim_{k \to \infty} f(x^k) = \min_{x \in R} f(x) = v_0$; *and*

(c) *if* $R^0 \cap R_k^0 = \phi$ *for some finite* k, $f(x^k) = \min_{x \in R} f(x) = v_0$,

where

$$\nabla f(x^k) = 0. \tag{2}$$

Proof. (a) Assume $x^* \in R^0 \cap R_k^0$ and let $Q_* = Q(x^*, x^k)$. Define the set $S \equiv \{x | f(x) \leqq f(x^k) - (1/Q_*)$, and $g_i(x) \geqq (1/Q_*), i = 1, \cdots, m\}$. It follows immediately that $\inf_{x \in S} Q(x, x^k) = \inf_{x \in R^0 \cap R_k^0} Q(x, x^k) \geqq v_0 > -\infty$. But $S \neq \phi$, since $x^* \in S$, S contains no boundary points of R, and S is compact. Also, $Q(x, x^k)$ is continuous in S by C2.

Since the greatest lower bound of a continuous function bounded on a compact set is attained at a point in that set, at least one minimum x^{k+1} exists. The strict convexity of $Q(x, x^k)$ in $R^0 \cap R_k^0$ implies that x^{k+1} is the unique minimum of $Q(x, x^k)$ in $R^0 \cap R_k^0$.

By condition C2 and the necessity of the vanishing of the first partial derivatives of $Q(x, x^k)$ at an unconstrained minimum, equation (1) follows and part (a) of the theorem is proved.

(b) From part (a), $Q(x, x^k)$ is minimized uniquely over $R^0 \cap R_k^0$ at x^{k+1}. Since $R_k^0 \equiv \{x | f(x) < f(x^k)\}$, it follows that $f(x^{k+1}) < f(x^k)$. Since part (a) of the theorem can now be applied to x^{k+1} it follows inductively that a strictly monotonic decreasing sequence $\{f(x^k)\}$, $k=0, 1, 2, \cdots$, has been defined, where the corresponding points $\{x^k\}$ are all in R^0. Because conditions C1–C3 assure that $f(x)$ is bounded below in R by v_0, we can conclude that $\lim_{k\to\infty} f(x^k) = f(\bar{\bar{x}}) \geqq v_0$. We shall assume that $f(\bar{\bar{x}}) > v_0$ and force a contradiction.

The fact that $\{f(x^k)\}$ decreases to $f(\bar{\bar{x}})$ implies that there exists an N such that for $k \geqq N$, $0 < f(x^k) - f(x^{k+1}) \leqq \delta$, where δ can be made as small as desired by appropriate choice of N.

Since $f(x^k) > f(x^{k+1}) > f(\bar{\bar{x}}) > v_0$, we can assume that $f(\bar{\bar{x}}) \geqq v_0 + 2\epsilon$ for $\epsilon > 0$ small enough. Also, since $f(x)$ is continuous and R is the closure of R^0, it is possible to select $x^* \epsilon R$ such that $f(x^*) \leqq v_0 + \epsilon$. Combining these inequalities, it follows that $f(\bar{\bar{x}}) - f(x^*) \geqq \epsilon$ for ϵ small enough and for suitably selected $x^* \epsilon R^0$. Note that we also have $x^* \epsilon R_k^0$.

Fix ϵ and x^* to satisfy the last inequality. It follows that

$$Q(x^*, x^k) = \{1/[f(x^k) - f(x^*)]\} + \textstyle\sum_{i=1}^{i=m} [1/g_i(x^*)] < (1/\epsilon) + \sum_{i=1}^{i=m} [1/g_i(x^*)],$$

since $f(x^k) - f(x^*) > f(\bar{\bar{x}}) - f(x^*) \geqq \epsilon$, for all k.

Select δ such that

$$1/\delta \geqq (1/\epsilon) + \textstyle\sum_{i=1}^{i=m} [1/g_i(x^*)].$$

From above, there exists an N such that $k \geqq N$ implies that $0 < f(x^k) - f(x^{k+1}) \leqq \delta$, giving that

$$Q(x^{k+1}, x^k) \equiv [1/f(x^k) - f(x^{k+1})] + \textstyle\sum_{i=1}^{i=m} [1/g_i(x^{k+1})] \geqq (1/\delta) \geqq (1/\epsilon) + \sum_{i=1}^{i=m} [1/g_i(x^*)], \text{ so}$$

$$Q(x^{k+1}, x^k) > Q(x^*, x^k).$$

But since $x^* \epsilon R^0 \cap R_k^0$, this inequality contradicts the fact that $Q(x, x^k)$ is minimized over $R^0 \cap R_k^0$ at x^{k+1}. Therefore, the assumption that $f(\bar{\bar{x}}) > v_0$ is false. Since we must have $f(\bar{\bar{x}}) \geqq v_0$, it follows that we must have equality and hence that $\lim_{k\to\infty} f(x^k) \equiv \min_{x\epsilon R} f(x) = v_0$.

(c) If $x^k \epsilon R^0$ and $R^0 \cap R_k^0 = \phi$ for some finite k, then this means there exist no $x \epsilon R^0$ such that $f(x) < f(x^k)$. In this case, since $f(x)$ is continuous and R is the closure of R^0, this immediately implies that

$$f(x^k) = \min_{x\epsilon R} f(x) = v_0.$$

Equation (2), the vanishing of the gradient of $f(x)$ at x^k, is a necessary consequence of C2 and the fact that x^k is an unconstrained minimum of $f(x)$.

RELATION BETWEEN THE P AND Q FUNCTIONS

As REMARKED in the first section, the function utilized in SUMT is $P(x, r) \equiv f(x) + r\sum_{i=1}^{i=m} [1/g_i(x)]$. The following theorem gives the explicit connection between the P function and the Q function, via their minimizing points.

THEOREM 2. *If* C1–C6 *hold, and if* x^{k+1} *minimize* $Q(x, x^k)$ *in* $R^0 \cap R_k^0$, *then* x^{k+1} *minimizes* $P(x, r_k)$ *in* R, *for* $r_k = \alpha_k^2$, *where* $\alpha_k \equiv f(x^k) - f(x^{k+1})$.

Proof. From Theorem 1, x^{k+1} is the unique minimum of $Q(x, x^k)$ in $R^0 \cap R_k^0$ and, from equation (1),

$$\nabla Q(x^{k+1}, x^k) = \{1/[f(x^k) - f(x^{k+1})]^2\} \nabla f(x^{k+1}) - \sum_{i=1}^{i=m} [1/g_i^2(x^{k+1})] \nabla g_i(x^{k+1}) = \bar{0}.$$

This gives

$$\alpha_k^2 \nabla Q(x^{k+1}, x^k) = \nabla f(x^{k+1}) - \alpha_k^2 \sum_{i=1}^{i=m} [1/g_i^2(x^{k+1})] \nabla g_i(x^{k+1}) = \bar{0}, \quad (3)$$

where α_k is defined above.

Consider now the gradient of the P function,

$$\nabla P(x, r_k) = \nabla f(x) - r_k \sum_{i=1}^{i=m} [1/g_i^2(x)] \nabla g_i(x). \quad (4)$$

It is clear, comparing (3) and (4), that

$$\nabla P(x^{k+1}, \alpha_k^2) = \alpha_k^2 \nabla Q(x^{k+1}, x^k) = \bar{0}.$$

But it follows from results proved in reference 2 that conditions C1–C6 also guarantee the strict convexity of P in R, and assure that $P(x, r)$, with $r > 0$, be minimized by a unique point $x(r)$ in R^0. The convexity of P in R is enough to conclude that P is minimized at any point in R where its gradient vanishes. The conclusion of the theorem follows.

Since it was shown in reference 3 that $f[x(r_k)]$, where $x(r_k)$ minimizes $P(x, r_k)$ in R, is a monotonic increasing function of r_k, it follows a fortiori that the sequence $\{\alpha_k^2\}$ is likewise monotonic decreasing. It can be shown directly that $\{\alpha_k^2\}$ is a strictly monotonic decreasing sequence. This result is proved in lemma 2, which utilizes the conclusion of the following lemma. Conditions C1–C6 as well as the conclusion of theorems 1 and 2 are assumed.

LEMMA 1. $$\sum_{i=1}^{i=m} [1/g_i(x^{k+1})] < \sum_{i=1}^{i=m} [1/g_i(x^{k+2})].$$

Proof. Since $\{f(x^k)\}$ is a strictly decreasing monotonic sequence, we know that $f(x^k) > f(x^{k+1}) > f(x^{k+2})$. This gives

$$0 < f(x^k) - f(x^{k+1}) < f(x^k) - f(x^{k+2}). \quad (5)$$

From the fact that $Q(x, x^k)$ is minimized uniquely over $R^0 \cap R_k^0$ by

x^{k+1}, it follows that $Q(x^{k+1}, x^k) < Q(x^{k+2}, x^k)$, i.e., that

$$\{1/[f(x^k)-f(x^{k+1})]\}+\sum_{i=1}^{i=m}[1/g_i(x^{k+1})]$$
$$<\{1/[f(x^k)-f(x^{k+2})]\}+\sum_{i=1}^{i=m}[1/g_i(x^{k+2})]$$
$$<\{1/[f(x^k)-f(x^{k+1})]\}+\sum_{i=1}^{i=m}[1/g_i(x^{k+2})],$$

using the above inequality (5).

The conclusion of the lemma follows immediately on deleting the common first term on both sides of the inequality.

LEMMA 2. $\alpha_k \equiv f(x^k)-f(x^{k+1}) > f(x^{k+1})-f(x^{k+2}) \equiv \alpha_{k+1}$.

Proof. Since $\min_{x\epsilon R} P(x, \alpha_k^2) = P(x^{k+1}, \alpha_k^2)$, and since the point minimizing $P(x, r)$ in R for any specified value of $r>0$ is unique, it follows from the definition of P that,

$$f(x^{k+1})+\alpha_k^2\sum_{i=1}^{i=m} 1/g_i(x^{k+1}) < f(x^{k+2})+\alpha_k^2\sum_{i=1}^{i=m} 1/g_i(x^{k+2}), \quad \text{and}$$
$$f(x^{k+2})+\alpha_{k+1}^2\sum_{i=1}^{i=m} 1/g_i(x^{k+2}) < f(x^{k+1})+\alpha_{k+1}^2\sum_{i=1}^{i=m} 1/g_i(x^{k+1}).$$

Combining these two inequalities and rearranging gives

$$(\alpha_k^2-\alpha_{k+1}^2)\left[\sum_{i=1}^{i=m} 1/g_i(x^{k+2})-\sum_{i=1}^{i=m} 1/g_i(x^{k+1})\right]>0.$$

From lemma 1, the second factor of this inequality is positive. Therefore, we must have $\alpha_k^2>\alpha_{k+1}^2$, or $\alpha_k>\alpha_{k+1}$, as asserted.

The precise relation between the points minimizing the P function and the Q function in the feasible region is now clear. $P(x, r_k)$ must be minimized in R for a strictly monotonic decreasing sequence $\{r_k\}$, where $r_k>0$, all k, and $k=0, 1, 2, \cdots$. The point $x(r_k)$ minimizing $P(x, r_k)$ in R coincides with the point x^{k+1} minimizing $Q(x, x^k)$ in $R^0 \cap R_k^0$, when we take $r_k=\alpha_k^2 \equiv [f(x^k)-f(x^{k+1})]^2$. From lemma 2 it follows that $\{\alpha_k^2\}$ is a strictly monotonic decreasing sequence and, from theorem 1, the sequence decreases to 0.

Since $x(r_k)\rightarrow\bar{x}$, where $\{r_k\}$ need only be a strictly monotonic sequence of positive number decreasing to 0, but is otherwise an arbitrary sequence, it follows that $\{\alpha_k^2\}$ is a particular realization of such a sequence. In short, the sequence of points minimizing $P(x, r_k)$ and $Q(x, x^k)$ coincide when $r_k=\alpha_k^2$, all k.

Another way of stating the relation is to consider the trajectory of minima $x(r)$ of $P(x, r)$, which is well-defined and continuous,[4] when r is permitted to decrease strictly and continuously to 0. Then, it follows that the points minimizing $Q(x, x^k)$, $k=0, 1, 2, \cdots$, all lie on this trajectory.

We shall not pursue the analogy further, except to remark that all of the theoretical and computational results that apply to the P function technique, can now be readily translated to apply to the Q function technique. The next section points out two advantages of the Q function that may warrant further comparative analysis with the P function, particularly with regard to computational efficiency.

POSSIBLE ADVANTAGES OF THE Q FUNCTION

THE Q FUNCTION is defined in such a manner that forces a decrease in the value of the objective function $f(x)$, from the initial starting point. The decrease in $f(x)$ is thereafter monotonic, from one minimizing point to the next, a characteristic shared by the P function process. However, in the latter technique, the initial value of r is, by and large, arbitrary. If r is chosen too small, the minimization problem is usually expensive computationally; if too large, then the first minimizing point may be far inside the region, where it is quite possible that $f(x)$ be increased substantially over its value at the point of departure. The Q function implicitly and with no arbitrariness 'sets' the initial value of r in such a manner that the objective function must decrease. The remaining question is whether the initial minimization problem is 'difficult' computationally, compared with initial minimizations typically encountered via the P function and the aforementioned arbitrariness in selecting r initially. This is not presently known.

In utilizing the P function, a decreasing sequence of values of the parameter r must be specified. Again, the choice is quite arbitrary, although it has been found expedient to reduce r by a constant factor in order to realize a great simplification in extrapolation. The Q function allows no such arbitrariness, since there are no controllable parameters involved in the technique utilizing it. Once the initial point is selected, the entire process is well determined. Again, it is an open question as to whether the minimization problems that are defined by the Q function procedure are, in some sense, computationally more or less difficult than the sequence of minimization problems typically encountered in utilizing the P function, allowing for the arbitrariness in fixing the rate of decrease of r. Another question that must be answered is whether an extrapolation technique can be devised for the Q function that is as efficient as that developed for the P function, where the decrease in r can be accommodated to make the extrapolation calculations quite simple.

In conclusion, it may be of interest to note that a few preliminary experiments using an algorithm based on the Q function indicated correct convergence to the solution of a convex problem. The algorithm is not sufficiently developed, nor is the data adequate to speculate as to the comparative efficiency of this approach.

REFERENCES

1. C. W. Carroll, "The Created Response Surface Technique for Optimizing Nonlinear Restrained Systems," *Opns. Res.* **9,** 169–184 (1961).
2. A. V. Fiacco and G. P. McCormick, "Programming Under Nonlinear Constraints by Unconstrained Minimization: A Primal-Dual Method," The Research Analysis Corporation, RAC-TP-96, Sept. 1963.
3. ——— and ———, "The Sequential Unconstrained Minimization Technique for Nonlinear Programming, a Primal-Dual Method," *Management Sci.* **10,** 360–366 (1964).
4. ——— and ———, "Computational Algorithm for the Sequential Unconstrained Minimization Technique for Nonlinear Programming," *Management Sci.* **10,** 601–617 (1964).
5. P. Huard, "Résolution des P. M. à Constraintes Non Linéares par la Méthode des Centres," Note E. D. F. no. HR 5.690 du 6 Mai 1964.
6. R. E. Stong, "A Note on the Sequential Unconstrained Minimization Technique for Non-Linear Programming," *Management Sci.* **12,** 142–144 (1965).

Note on 'The Sequential Maximization Technique'

John M. Danskin

Center for Naval Analyses, Arlington, Virginia

(Received March 29, 1967)

This note presents a generalization of the main theorem of the sequential maximization technique given in Fiacco and McCormick's paper just preceding this note.

THE present author refereed the preceding paper.[1] He pointed out to the authors that their Theorem 1 in the second section could be proved under much weaker conditions and for wider spaces. They, however, wished to leave their exposition intact. It is at their suggestion that the generalization is presented here.

Consider the problem of minimizing $f(x)$ subject to the side conditions

$$g_j(x) \geqq 0. \qquad (j=1, \cdots, m) \quad (1)$$

Here x ranges on a metric space, on which the conditions (1) define a compact set K. We suppose that the interior K^0 of K is not empty, and that on K^0 the function $f(x)$ assumes values arbitrarily close to the minimum. We suppose concerning $f(x)$ and $g_j(x), j=1, \cdots, m$, only that they are continuous.

We form the function, for $k \geqq 0$.

$$Q_k(x)=1/[f(x^k)-f(x)]+\sum_j [1/g_j(x)] \quad (2)$$

defined for $x \epsilon K^0$ and $f(x)<f(x^k)$ provided there are such points. Here x^0 is any point of K^0 and x^{k+1} is defined as a point yielding the minimum to $Q_k(x)$ over all $x \epsilon K^0$ with $f(x)<f(x^k)$. It is possible that the set $f(x)<f(x^k)$ in K^0 is empty for some k. Then x^k is the minimum. Otherwise we have an infinite sequence $\{x^k\}$. The main theorem is the following.

THEOREM 1. *Any subsequence of $\{x^k\}$ that converges to a limit converges to a minimum for $f(x)$ on K.*

Proof. Let $c=\min_{x \epsilon K} f(x)$. Suppose that the subsequence $\{x^{k'}\}$ of $\{x^k\}$ converges to a point x^*. Suppose that $f(x^*)=c+2\rho$, where $\rho>0$. By hypothesis there is a point $\bar{x} \epsilon K^0$ at which $f(\bar{x})<c+\rho$. Put

$$\alpha_k=f(x^k)-f(x^{k+1}).$$

Then $\alpha_k>0$ and $\alpha_k \rightarrow 0$ as $k \rightarrow \infty$. Take k_0 so large that if $k \geqq k_0$

$$\alpha_k \sum [1/g_j(\bar{x})]<\tfrac{1}{2}, \quad (3)$$

and also so that $2\alpha_k<\rho$. Then, since $f(x^k)-f(\bar{x})>\rho$,

$$2\alpha_k<f(x^k)-f(\bar{x}). \quad (4)$$

Then

$$\begin{aligned} Q_k(x^{k+1}) &= 1/[f(x^k)-f(x^{k+1})]+\textstyle\sum_j [1/g_j(x^{k+1})] \\ &= 1/\alpha_k[1+\alpha_k \textstyle\sum [1/g_j(x^{k+1})]] \\ &> 1/\alpha_k = (1/2\alpha_k)+(1/2\alpha_k) > 1/[f(x^k)-f(\bar{x})]+\textstyle\sum [1/g_j(\bar{x})] \\ &= Q_k(\bar{x}), \end{aligned}$$

where at the next-to-last step we have used (4) and (3). It follows that x^{k+1} does not minimize $Q_k(x)$ on the set $x \epsilon K^0$, $f(x) < f(x^k)$. This is a contradiction; hence $f(x^*) = c$ as desired. The theorem is proved.

It is perfectly possible that the whole series $\{x^k\}$ does not converge. But if the minimum is unique, the theorem implies that it converges.

REFERENCE

1. F. V. Fiacco and G. P. McCormick, "The Sequential Unconstrained Minimization Technique (SUMT) without Parameters," *Opns. Res.* **15,** 820–827 (1967).