† *This work was supported by the Office of Naval Research, Contract No. N00014-67-A-0321-0003 (NR047-095).*

* *Graduate School of Business Administration and Curriculum in Operations Research and Systems Analysis, University of North Carolina at Chapel Hill.*

** *Department of Statistics and Curriculum in Operations Research and Systems Analysis, University of North Carolina at Chapel Hill.*

STABILITY AND EXPONENTIAL PENALTY FUNCTION TECHNIQUES IN NONLINEAR PROGRAMMING†

by

J. P. Evans* *and* F. J. Gould**

Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 723

November, 1970

STABILITY AND EXPONENTIAL PENALTY FUNCTION TECHNIQUES
IN
NONLINEAR PROGRAMMING[†]

by

J. P. Evans[*] and F. J. Gould[**]

Abstract

In this paper, we develop the concept of a general exponential penalty function as a method of solving inequality constrained optimization problems. Relationships to the concept of stability in nonlinear programming are presented together with results on approximate solutions and bounds on the optimal value of a program. Conditions are stated which guarantee that sequences of penalty function maximizers will yield optimal solutions to the original programming problem.

I. Introduction

Consider the general inequality constrained non-linear programming problem,

$$(P_{\bar{b}})$$

$$\max f(x), \quad \text{subject to}$$

$$g_j(x) \le \bar{b}_j, \quad j = 1, \ldots, m$$

---

† *This work was supported by the Office of Naval Research, Contract No. N0014-67-A-0321-0003 (NR047-095).*

* *Graduate School of Business Administration and Curriculum in Operations Research and Systems Analysis, University of North Carolina at Chapel Hill.*

** *Department of Statistics and Curriculum in Operations Research and Systems Analysis, University of North Carolina at Chapel Hill.*

where f, g_j, $j = 1,\ldots,m$, are real-valued functions defined on R^n. Gould [4] explored the use of nonlinear multiplier functions in extended Lagrangians associated with problem $(P_{\bar{b}})$. Define

$$(1.1) \qquad L(x,\lambda) = f(x) - \sum_{j=1}^{m} \lambda_j(g_j(x))$$

where λ_j, $j = 1,\ldots,m$, is a monotonically nondecreasing function of a single variable with values in the extended reals such that $\lambda_j(\bar{b}_j) < \infty$, $j = 1,\ldots,m$. For such multiplier functions, the following result is valid:

Suppose $\bar{\lambda}$ is given and x^* satisfies

$$(1.2) \qquad \begin{aligned} &\text{i)} \quad x^* \text{ maximizes } L(x,\bar{\lambda}) \text{ over } R^n, \\ &\text{ii)} \quad g_j(x^*) \le \bar{b}_j, \quad j = 1,\ldots,m, \\ &\text{iii)} \quad \bar{\lambda}_j(g_j(x^*)) = \bar{\lambda}_j(\bar{b}_j), \quad j = 1,\ldots,m; \end{aligned}$$

then x^* is an optimal solution to $(P_{\bar{b}})$.

The monotonicity property of the λ_j's in (1.1) causes the second term in (1.1) to be larger for larger values of $g_j(x)$. Roughly speaking, the extended Lagrangian $L(x,\lambda)$ is a general penalty function which penalizes those x's feasible in $(P_{\bar{b}})$ less than those which are infeasible. In [2] Fiacco and McCormick cite a number of references to the use of an exponential type of penalty function, e.g. where the $\lambda_j(\cdot)$ in (1.1) has the form

$$\lambda_j(\xi) = \begin{cases} \xi^2, & \xi > 0 \\ 0, & \xi \le 0. \end{cases}$$

These references include Motzkin [5], Goldstein and Kripke [3], and Zangwill [6].

In this paper, we define a class of general exponential penalty functions and employ the properties of these functions in order to identify approximate solutions to mathematical programs and to obtain certain convergence results in the same context. Relationships to stability in nonlinear programming [1] and to the general exterior penalty function of Fiacco and McCormick [2] are explored.

In the remainder of the paper, we will employ the following definitions and notation:

a) $S_b = \{x \in R^n: \; g_j(x) \leq b_j, \; j = 1,\ldots,m\}$;

b) $B = \{b \in R^m: \; S_b \neq \phi\}$;

c) $f_{\mathit{sup}}(b): \; B \to R \cup \{+\infty\}$, the perturbation function defined by $\mathit{sup}[f(x): \; x \in S_b]$;

d) A vector $\bar{x} \in R^n$ is an ε-solution to $(P_{\bar{b}})$ if given $\varepsilon > 0$, $\bar{x}$ satisfies

i) $\bar{x} \in S_{\bar{b}}$,

ii) $f(\bar{x}) \geq f_{\mathit{sup}}(\bar{b}) - \varepsilon$;

e) $S: \; B \to 2^{R^n}$, a point-to-set correspondence defined by $S(b) = S_b$;

f) $z: \; B \to R$, the support function defined by

$$z(b) = \sum_{j=1}^{m} [\lambda_j(b_j) - \lambda_j(\bar{b}_j)] + f_{\mathit{sup}}(\bar{b}); \tag{1.3}$$

g) An m-vector of 1's is represented by e.

II. General Exponential Penalty Functions

The relationship between penalty function techniques and the concept of stability will be developed in the following exposition. First we observe that given $\bar{\lambda}$ satisfying the monotonicity and finiteness conditions stated above, if some point $\bar{x}$ maximizes $L(x,\bar{\lambda})$ over R^n, then a point of the f_{sup} function has been obtained, namely $f_{\mathit{sup}}(g(\bar{x}))$. This follows by replacing $\bar{b}$ by $g(\bar{x})$ in the optimality conditions (1.2) above. But then it can be shown [4] that

$$z(b) = \sum_{j=1}^{m} [\bar{\lambda}_j(b_j) - \bar{\lambda}_j(g_j(\bar{x}))] + f_{sup}(g(\bar{x})) \geq f_{sup}(b), \quad \forall\, b \in B.$$

Thus after each maximization of the penalty function, the support function yields an upper bound on the value of f_{sup} at $\bar{b}$, the right-hand side in the original problem. Hence it is intuitively reasonable to employ the information from a succession of penalty function maximizations to choose new multiplier functions, λ_j, such that the next penalty function maximizer will in some sense be closer to an optimal solution for $(P_{\bar{b}})$. However, it should be clear that the success of this strategy will be sensitive to the behavior of the f_{sup} function at arguments near $\bar{b}$. Consider the situations shown in Figures 1 and 2 below.

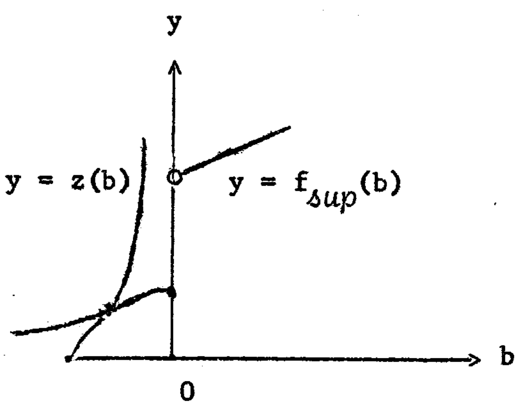

Figure 1

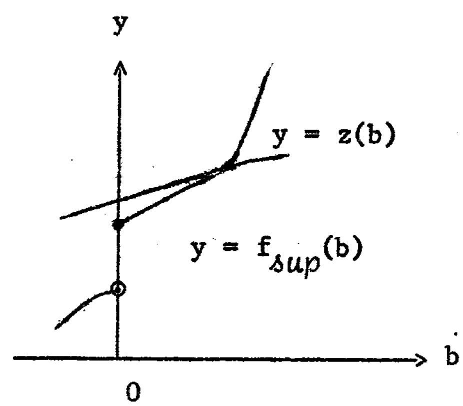

Figure 2

In Figure 1, the f_{sup} function is lower semi-continuous at 0. Intuitively we would like to "approach $\bar{b} = 0$ from below", because values of $b > 0$ yield poor estimates of $f_{sup}(0)$. This suggests that we would like to choose multiplier functions which increase rapidly for values of b near zero. In contrast, Figure 2 displays an f_{sup} function that is upper semi-continuous at 0. Consequently, we wish to approach $\bar{b} = 0$ from above since values of $b < 0$ will yield poor estimates of $f_{sup}(0)$.

The continuity properties of f_{sup} are closely related to properties of the point-to-set correspondence S defined in the preceding section. These properties were explored in [1] and can be summarized for our purposes here as follows. Assume $S_{\bar{b}}$ is compact and nonempty and that each constraint function, $g_1, \ldots, g_m$, is continuous on R^n.

a. The mapping S is upper semi-continuous at $\bar{b}$ if and only if there is a $\hat{b} > \bar{b}$ such that $S_{\hat{b}}$ is compact.[1]

b. If $I_{\bar{b}} = \{x: g(x) < \bar{b}\} \neq \phi$, the mapping S is lower semi-continuous at $\bar{b}$ if and only if $c\ell(I_{\bar{b}}) = S_{\bar{b}}$ (where $c\ell$ denotes closure).

c. If the mapping S is upper semi-continuous at $\bar{b}$ and $f(\cdot)$ is upper semi-continuous on R^n, then $f_{sup}(\cdot)$ is upper semi-continuous at $\bar{b}$.

d. Suppose $I_{\bar{b}} \neq \phi$. If $f(\cdot)$ is lower semi-continuous and S is lower semi-continuous at $\bar{b}$, then f_{sup} is lower semi-continuous at $\bar{b}$.

From these statements, it is clear that if $f(\cdot)$ is continuous, we can characterize f_{sup} at $\bar{b}$ from properties of the constraint map S.

With this background, we will now define a class of general exponential multiplier functions which are designed to permit sufficient flexibility to deal with the situations in Figures 1 and 2.

Definition: λ: $R \to R$ is a general exponential multiplier function for $(P_{\bar{b}})$ if given parameters $\alpha \geq 1$, $\beta > 0$, and $\delta \geq 0$,

$$\lambda(\xi; \alpha, \beta, \delta) = \beta[h(\xi+\delta)]^{\alpha}$$

---

1 *The precise definitions of upper and lower semi-continuity for the mapping S need not concern us here; however, these are stated explicitly in* [1].

where $h: R \to R$ is a continuous, strictly increasing function satisfying

i) $\xi \to \infty$ implies $h(\xi) \to \infty$,

ii) $h(0) = 1$,

iii) $h(g_j(x)) > 0 \quad \forall\ x \in R^n, \quad j = 1,\ldots,m.$

Example 1: $h(\xi) = e^{\xi}$.

Example 2: Suppose $k_j = \inf_{x \in R^n} g_j(x) > -\infty, \quad j = 1,\ldots,m,$

and $K > |\min_j k_j|$. Then define $h(\xi) = (K+\xi)/K$,

and in the specification of λ, further restrict α to the odd integers.

Part of the motivation for general exponential penalty functions can be seen from the property that for an x such that $g_j(x) > \bar{b}_j$, as $\alpha \to \infty$,

$$\lambda(g_j(x)-\bar{b};\ \alpha,\ \beta,\ \delta) \to +\infty.$$

Thus an arbitrarily large penalty can be imposed on infeasible points. Correspondingly with reference to Figure 1, by increasing α, the support function can be given a large slope in the vicinity of $\bar{b}$.

III. Feasible and Near-feasible Penalty Function Maximizers

In this section, we show that under rather general conditions, given a value for the parameter β, the parameters α and δ can be manipulated so as to control the location of the penalty function maximizers. This leads to theorems on ε-solutions and convergence to optimal solutions. First, we make the following assumptions for the remainder of the exposition.

A-1. The multiplier function $\lambda(\cdot)$ has the general exponential form defined above.

A-2. In $(P_{\bar{b}})$, $\bar{b} = 0$, S_0 is nonempty and compact, and $f, g_1, \ldots, g_m$, are continuous on R^n.

A-3. There is a scalar α such that for each $j = 1, \ldots, m$, the ratio

$$\frac{f(x)}{\lambda(g_j(x); \alpha, 1, 0)}$$

is bounded from above on R^n.

The last assumption is employed to guarantee that for an appropriate choice of α the general exponential penalty function

$$P(x; \alpha, 1, 0) = f(x) - \sum_{j=1}^{m} \lambda(g_j(x); \alpha, 1, 0) \tag{3.1}$$

has a finite supremum. In the absence of some such condition, it is easy to construct examples for which the penalty function is unbounded on R^n. Now we present a collection of results related to the case portrayed in Figure 1 in which f_{sup} is, at worst, only lower semi-continuous.

LEMMA 3.1: Let $\beta > 0$ be given and suppose $\delta > 0$ is such that $S_{-\delta e} = \{x: g_j(x) \leq -\delta, j = 1, \ldots, m\}$ is nonempty. Then there exists a scalar $\bar{\alpha}$ (which depends on β and δ) such that $\alpha > \bar{\alpha}$ implies $P(x; \alpha, \beta, \delta)$ has a global maximum and each such maximizer belongs to S_0.[1]

PROOF: By assumption A-3, there is an M such that for some α_1 and for $j = 1, \ldots, m$

---

1 *Here, as in the entire developemtn,* α *is always subject to any restrictions assumed in the definition of* λ_j *as, for instance, in Example 2 of Section II.*

$$\frac{f(x)}{\lambda(g_j(x);\ \alpha_1,\ 1,\ 0)} = \frac{f(x)}{[h(g_j(x))]^{\alpha_1}} < M, \quad \text{each} \quad x \in R^n .$$

Suppose $x \epsilon R^n - S_0$; then for some j, $g_j(x) > 0$. Then we can choose α_2 such that $\alpha > \alpha_2$ and $x \epsilon R^n - S_0$ imply

$$\frac{f(x)}{[h(g_j(x)+\delta)]^{2\alpha}} < \frac{f(x)}{[h(\delta)]^{\alpha}[h(g_j(x))]^{\alpha}} < \frac{M}{[h(\delta)]^{\alpha}} < \beta/2 .$$

Let $\bar{x} \epsilon S_{-\delta e}$, and let $c = f(\bar{x})$; then we can choose α_3 such that $\alpha > \alpha_3$ yields

$$\frac{|c-\beta m|}{[h(g_j(x)+\delta)]^{\alpha}} < \frac{|c-\beta m|}{[h(\delta)]^{\alpha}} < \beta/2.$$

Hence, for $\alpha > \max\ \{\alpha_1,\ \alpha_2,\ \alpha_3\}$ we have $\forall\ x \epsilon R^n - S_0$

$$\frac{f(x)}{[h(g_j(x)+\delta)]^{2\alpha}} - \frac{c - \beta m}{[h(g_j(x)+\delta)]^{2\alpha}} \leq \frac{f(x)}{[h(g_j(x)+\delta)]^{2\alpha}} + \frac{|c-\beta m|}{[h(g_j(x)+\delta)]^{2\alpha}} < \beta.$$

This implies there is an $\bar{\alpha} \ni \alpha > \bar{\alpha}$ yields

$$\frac{f(x)}{[h(g_j(x)+\delta)]^{\alpha}} < \beta + \frac{c - \beta m}{[h(g_j(x)+\delta)]^{\alpha}} ,$$

which can be rewritten as

$$f(x) < \beta([h(g_j(x)+\delta)]^{\alpha} - m) + c, \quad x \in R^n - S_0.$$

Since we can always choose $\bar{\alpha}$ large enough to insure that the coefficient of β above is positive, for all $\alpha > \bar{\alpha}$ we have

$$\frac{f(x) - c}{[h(g_j(x)+\delta)]^\alpha - m} < \beta, \quad \forall\, x \in R^n - S_0.$$

By the properties assumed for $h(\cdot)$, $h(g_j(x)) > 0$, $j = 1,\ldots,m$, each $x \epsilon R^n$, hence we have,

$$\frac{f(x) - c}{\sum_{j=1}^{m} [h(g_j(x)+\delta)]^\alpha - m} < \beta, \quad x \in R^n - S_0.$$

Thus we have shown that there is an $\bar{\alpha}$, which depends on β and δ, such that $\alpha > \bar{\alpha}$ implies

$$\text{(3.2)} \qquad f(x) - \sum_{j=1}^{m} \lambda(g_j(x);\, \alpha,\, \beta,\, \delta) < f(\bar{x}) - m\beta, \qquad x \in R^n - S_0.$$

Now since S_0 is compact, $P(x;\alpha,\beta,\delta)$ has a maximum on S_0 and since $\bar{x} \in S_{-\delta e} \subseteq S_0$, we have

$$\text{(3.3)} \qquad f(\bar{x}) - m\beta \leq f(\bar{x}) - \sum_{j=1}^{m} \lambda(g_j(\bar{x});\, \alpha,\, \beta,\, \delta) \leq \max_{x \epsilon S_0} P(x;\, \alpha,\, \beta,\, \delta).$$

This follows because $g_j(\bar{x}) \leq -\delta$ and $\delta > 0$ implies

$$\sum_{j=1}^{m} \lambda(g_j(\bar{x});\, \alpha,\, \beta,\, \delta) = \sum_{j=1}^{m} \beta[h(g_j(\bar{x})+\delta)]^\alpha \leq \sum_{j=1}^{m} \beta = m\beta.$$

Combining (3.2) and (3.3) we conclude that $\alpha > \bar{\alpha}$ implies

$$\max_{x \epsilon S_0} P(x;\, \alpha,\, \beta,\, \delta) = \max_{x \epsilon R^n} P(x;\, \alpha,\, \beta,\, \delta)$$

and furthermore, it is seen that (for $\alpha > \bar{\alpha}$) no maximizer of $P(x;\alpha,\beta,\delta)$ lies outside S_0. Q.E.D.

Now we apply Lemma 3.1 to obtain a theorem on ε-solutions to (P_0) for the case in which the mapping S defined in Section I is lower semi-continuous at 0.

Let $\{\beta_n\}$ be any sequence of positive numbers such that $\beta_n \to 0$, and assume there is a $\bar{d} > 0 \ni S_{-\bar{d}e}$ is nonempty (such a $\bar{d}$ will exist if S is lower semi-continuous). Let δ be a fixed positive scalar such that $0 < \delta < \bar{d}$. Then there is an $\bar{\alpha}_n$ for each n (which depends on β_n and δ) such that the conclusion of Lemma 3.1 is valid; hence $P(x;\bar{\alpha}_n,\beta_n,\delta)$ will have a maximum at some point $x^*_n(\delta) \in S_0$. Let x^* denote any solution to (P_0), let $x^*_{-\delta}$ denote any solution to $(P_{-\delta e})$, and let $\{\alpha_n\}$ be a sequence of scalars such that $\alpha_n > \bar{\alpha}_n$, each n.

THEOREM 3.2: Assume the mapping S is lower semi-continuous at 0 and let $\varepsilon > 0$ be given. Then there is a scalar $d_\varepsilon > 0$ such that $0 < \delta < d_\varepsilon$ implies the sequence $\{x^*_n(\delta)\}$ has a convergent sebsequence, and if v_δ is a subsequential limit

i) $v_\delta \in S_0$, and

ii) $f_{\sup}(0) - \varepsilon \le f(v_\delta) \le f_{\sup}(0)$.

PROOF: Since f is continuous and S is lower semi-continuous at 0, $f_{\sup}$ is lower semi-continuous at 0. Thus given $\varepsilon > 0$, there is a scalar $d > 0$ such that $0 < \delta < d$ implies $f_{\sup}(0) - \varepsilon \le f_{\sup}(-\delta e) \le f_{\sup}(0)$, and thus $f(x^*) - \varepsilon \le f(x^*_{-\delta}) \le f(x^*)$. Of course as noted above, d can be chosen sufficiently small that $S_{-\delta e}$ is nonempty. Thus for $\delta \in (0,d)$ we have

$$(3.4) \quad f(x^*) - \varepsilon - m\beta_n \le f(x^*_{-\delta}) - m\beta_n \le f(x^*_{-\delta}) - \sum_{j=1}^{m} \lambda(g_j(x^*_{-\delta});\ \alpha_n,\ \beta_n,\ \delta)$$

because $g_j(x^*_{-\delta}) + \delta \le 0$, $j = 1,\ldots,m$, which implies $h(g_j(x^*_{-\delta})+\delta) \le 1$, which in turn implies

$$\sum_{j=1}^{m} \lambda(g_j(x^*_{-\delta});\ \alpha_n,\ \beta_n,\ \delta) = \sum_{j=1}^{m} \beta_n[h(g_j(x^*_{-\delta})+\delta)]^{\alpha_n} \le \sum_{j=1}^{m} \beta_n = m\beta_n.$$

Since $x^*_n(\delta)$ maximizes $P(x;\ \alpha_n,\ \beta_n,\ \delta)$, we have

(3.5)
$$f(x^*_{-\delta}) - \sum_{j=1}^{m} \lambda(g_j(x^*_{-\delta});\alpha_n,\beta_n,\delta) \le f(x^*_n(\delta)) - \sum_{j=1}^{m} \lambda(g_j(x^*_n(\delta));\alpha_n,\beta_n,\delta)$$
$$< f(x^*_n(\delta)) \le f(x^*)$$

in which the last inequality follows from the fact that $x^*_n(\delta)\epsilon S_0$, each n.

Now since the sequence $\{x^*_n(\delta)\}\epsilon S_0$, and S_0 is compact, there is a convergent subsequence $\{x^*_{n_k}(\delta)\}$ converging to $v_\delta \epsilon S_0$. From (3.4) and (3.5), it follows that for any such convergent subsequence

$$f(x^*) - \varepsilon - m\beta_{n_k} \le f(x^*_{n_k}(\delta)) \le f(x^*), \text{ each } k.$$

Hence we conclude

$$f(x^*) - \varepsilon \le f(v_\delta) \le f(x^*)$$

Q.E.D.

The thrust of the above theorem is that holding the parameter δ constant, we can employ a sequence of exponents α_n so as to insure convergence of the sequence of penalty function maximizers to an ε-solution for (P_0). It is to be noted that with $\delta > 0$, we cannot guarantee convergence to a true solution to (P_0), whereas unless $\delta > 0$ we cannot insure feasibility of the penalty function maximizers. At the expense of a somewhat more complicated algorithm, we can obtain both of these properties; this is indicated in the following result.

THEOREM 3.3: Suppose the assumptions of Theorem 3.2 hold and let $\{\delta_n\}$ be a sequence of positive scalars such that $\delta_n < d$, each n and $\delta_n \to 0$. Then a sequence of exponents $\{\alpha_n\}$ can be chosen such that

i) $P(x;\ \alpha_n,\ \beta_n,\ \delta_n)$ has a maximum at $x^*_n \in S_0$;

ii) $\lim_{n\to\infty} f(x^*_n) = f(x^*)$;

iii) $\{x^*_n\}$ has a convergent subsequence and each subsequential limit is a solution to (P_0).

PROOF: The set $S_{-\delta_n e}$ is nonempty and compact by the assumptions of lower semi-continuity of S, continuity of g, and compactness of S_0. Let $\hat{x}_n$ be an optimal solution to $(P_{-\delta_n e})$; $\hat{x}_n$ exists by the continuity of f. The assumptions also guarantee that Lemma 3.1 is valid, hence there is an α_n such that conclusion i) of the theorem holds. Then we have

$$\begin{aligned} f(\hat{x}_n) - m\beta_n &\leq f(\hat{x}_n) - \sum_{j=1}^{m} \lambda(g_j(\hat{x}_n); \alpha_n, \beta_n, \delta_n) \\ &\leq f(x^*_n) - \sum_{j=1}^{m} \lambda(g_j(x^*_n); \alpha_n, \beta_n, \delta_n) \\ &< f(x^*_n) \leq f(x^*) = f_{sup}(0). \end{aligned} \tag{3.6}$$

Since f_{sup} is lower semi-continuous at 0, we have $\lim_{n\to\infty} f(\hat{x}_n) = f_{sup}(0)$. Combining this with (3.6) yields

$$\lim_{n\to\infty} f(x^*_n) = f_{sup}(0) ,$$

which proves ii).

Since S_0 is compact and α_n was chosen to insure $x^*_n \in S_0$, $\{x^*_n\}$ has a convergent subsequence, $x^*_{n_k} \to x_0 \in S_0$. The argument at the conclusion of Theorem 3.2 permits us to conclude that $f(x_0) = f_{sup}(0)$ which proves iii).

Q.E.D.

Now we turn to a collection of results paralleling those above which permit us to deal with the case pictured in Figure 2 in which f_{sup} is upper semi-continuous at 0. We continue to assume that conditions A-1, A-2, and A-3 hold.

LEMMA 3.4: Let $\beta > 0$, and suppose $\delta > 0$ is such that $S_{\delta e}$ is compact. Then there exists a scalar $\bar{\alpha}$ dependent on β and δ such that $\alpha > \bar{\alpha}$ implies $P(x;\alpha,\beta,0)$ has a global maximum and each such maximizer belongs to $S_{\delta e}$.

PROOF: Let $\bar{x} \in S_0$; then for $\alpha > 1$ by the compactness of $S_{\delta e}$ we have

$$(3.7) \qquad \max_{x \in S_{\delta e}} P(x;\, \alpha,\, \beta,\, 0) \;\geq\; f(\bar{x}) - \sum_{j=1}^{m} \lambda(g_j(\bar{x});\, \alpha,\, \beta,\, 0) \;\geq\; f(\bar{x}) - m\beta.$$

Now suppose $x \in R^n - S_{\delta e}$; then for some j_0, $1 \leq j_0 \leq m$, $g_{j_0}(x) > \delta$. By assumption A-3 there is an α_1 and a constant M such that for $\alpha > \alpha_1$, $f(x)/[h(g_{j_0}(x))]^\alpha < M$, which implies

$$\frac{f(x)}{[h(g_{j_0}(x))]^{2\alpha}} < \frac{M}{[h(g_{j_0}(x))]^{\alpha}} < \frac{M}{[h(\delta)]^{\alpha}} < \beta\,.$$

By property iii) of the function h, $h(g_j(x)) > 0$ for each $j = 1,\ldots,m$; hence we can choose α_2 such that for $\alpha > \alpha_2$

$$\frac{f(x)}{\sum_{j=1}^{m} [h(g_j(x))]^{\alpha}} < \beta, \quad \text{each } x \in R^n - S_{\delta e}\,.$$

With an argument similar to that employed in the proof of Lemma 3.1, we can conclude that there is an α_3 such that $\alpha > \alpha_3$ implies

$$(3.8) \qquad \frac{f(x) - f(\bar{x})}{\sum_{j=1}^{m} [h(g_j(x))]^{\alpha} - m} < \beta, \quad \text{each } x \in R^n - S_{\delta e};$$

since the denominator on the left side of (3.8) can be assumed to be positive, we have for $\alpha > \bar{\alpha} = \max\,(\alpha_1,\alpha_2,\alpha_3)$,

$$(3.9) \qquad f(x) - \sum_{j=1}^{m} \lambda(g_j(x);\, \alpha,\, \beta,\, 0) = f(x) - \sum_{j=1}^{m} \beta[h(g_j(x))]^{\alpha} < f(x) - m\beta$$

for each $x \epsilon R^n - S_{\delta e}$. The conclusion of the lemma follows from (3.8) and (3.9).

Q.E.D.

In a proposition similar to Theorem 3.2, we can now state a result concerning what might be termed near-feasible, near-optimal solutions for P_0. Let $\{\beta_n\}$ be as in Theorem 3.2 and assume the mapping S is upper semi-continuous at 0; from the discussion in Section II above there must be a $\bar{\delta} > 0$ such that $S_{\bar{\delta}e}$ is compact. Then for each n there is an $\bar{\alpha}_n$ which depends on β_n and δ $(0 < \delta \leq \bar{\delta})$ such that for $\alpha_n > \bar{\alpha}_n$ the conclusion of Lemma 3.4 is valid; hence $P(x;\alpha_n,\beta_n,0)$ will have a maximum at some point $x^*_n(\delta) \epsilon S_{\delta e}$. Let x^*_δ be optimal in $(P_{\delta e})$.

THEOREM 3.5: Let $\varepsilon > 0$, $0 < \delta < \bar{\delta}$ be given, where $S_{\bar{\delta}e}$ is assumed to be compact. Then there is a scalar $d < \bar{\delta}$ such that $0 < \delta < d$ implies the sequence $\{x^*_n(\delta)\}$ has a convergent subsequence and if v_δ is a subsequential limit,

i) $v_\delta \in S_{\delta e}$;

ii) $f_{sup}(0) \leq f(v_\delta) \leq f_{sup}(0) + \varepsilon$.

PROOF: Since f is continuous and $S_{\delta e}$ is compact, f_{sup} is upper semi-continuous at 0. Thus, given $\varepsilon > 0$ there is a scalar $d > 0$, which can be assumed to be less than $\bar{\delta}$, such that $0 < \delta < d$ implies

$$f_{sup}(0) \leq f_{sup}(\delta e) \leq f_{sup}(0) + \varepsilon,$$

and hence

(3.10) $$f(x^*) \leq f(x^*_\delta) \leq f_{sup}(0) + \varepsilon.$$

Now since $S_{\delta e}$ is compact for $\delta < d$, and since the conclusion of Lemma 3.4 is valid, we can choose α_n such that the maximizer of $P(x;\alpha_n,\beta_n,0)$ belongs to

$S_{\delta e}$ for each n. The compactness of $S_{\delta e}$ guarantees that there is a convergent subsequence $x^*_{n_k}(\delta) \to v_\delta \in S_{\delta e}$.

Now $x^* \in S_0$, hence $g_j(x^*) \le 0$, $j = 1,\ldots,m$. Thus for each $k = 1,2,\ldots,$ we have

(3.11) $$f(x^*) - \beta_{n_k} m \le f(x^*) - \sum_{j=1}^{m} \lambda(g_j(x^*);\, \alpha_{n_k},\, \beta_{n_k},\, 0),$$

and since $x^*_{n_k}(\delta)$ is a penalty function maximizer,

(3.12) $$f(x^*) - \sum_{j=1}^{m} \lambda(g_j(x^*);\alpha_{n_k},\beta_{n_k},0) \le f(x^*_{n_k}(\delta)) - \sum_{j=1}^{m} \lambda(g_j(x^*_{n_k}(\delta));\alpha_{n_k},\beta_{n_k},0)$$

$$\le f(x^*_{n_k}(\delta)).$$

By the upper semi-continuity of S, $f(v_\delta) \le f(x^*)+\varepsilon$; thus combining the fact that $x^*_{n_k}(\delta) \to v_\delta$ with (3.11), and (3.12), we conclude $f(x^*) \le f(v_\delta) \le f(x^*)+\varepsilon$.

Q.E.D.

Observe that in Lemma 3.4 and Theorem 3.5 we do not guarantee the feasibility of any penalty function maximizer. In addition, we cannot assert that the limit point v_δ is optimal or even feasible in (P_0).

COROLLARY 3.6: Assume the conditions of Theorem 3.5 hold. Then a sequence $\{\alpha_n\}$ can be chosen such that the sequence of penalty function maximizers has a subsequential limit $v \in S_0$ and v is optimal in (P_0).

PROOF: For each n we can choose α_n to insure that $P(x;\alpha_n,\beta_n,0)$ has its maximum value at $x^*_n \in S^n$ where

$$S^n = \{x\colon\ g_j(x) \le \delta/n, \quad j = 1,\ldots,m\};$$

the conclusion follows from Theorem 3.5.

Q.E.D.

IV. Bounds for $f_{\delta up}$

In the discussion in Section II, it was observed that each maximization of a penalty function (Lagrangian) as in (1.1) yields a point of the $f_{\delta up}$ function and that via the support function (1.3), an upper bound can be computed for $f_{\delta up}(0)$. In this section, we present a result on these upper bounds for general exponential penalty functions.

Lemma 4.1: Let $\alpha > 1$, $\beta > 0$, $\delta \geq 0$ be given and assume $\bar{x}$ maximizes $P(x;\alpha,\beta,\delta)$. Then

$$f_{\delta up}(0) \leq f(\bar{x}) + \beta\{m[h(\delta)]^{\alpha} - \sum_{j=1}^{m} [h(g_j(\bar{x})+\delta)]^{\alpha}\}.$$

Proof: Let x^* be an optimal solution to (P_0). Then $g_j(x^*) \leq 0$, $j = 1,\ldots,m$, hence

$$f(x^*) - m\beta[h(\delta)]^{\alpha} \leq f(x^*) - \beta \sum_{j=1}^{m} [h(g_j(x^*)+\delta)]^{\alpha}$$

$$\leq f(\bar{x}) - \beta \sum_{j=1}^{m} [h(g_j(\bar{x})+\delta)]^{\alpha} .$$

The first inequality above holds because of the monotonicity of h, and the second inequality holds because $\bar{x}$ maximizes $P(x;\alpha,\beta,\delta)$. The conclusion follows.

Q.E.D.

V. Relationship to other Penalty Function Methods

As mentioned earlier much attention has been devoted to penalty function methods in general, and several references exist to various kinds of exponential penalties. First we mention a relationship to what Fiacco-McCormick [2] called a general exterior penalty function. This is a class of one-parameter functions which satisfy several conditions, among them being the following, restated in the form for the maximization required in (P_0).

$$\text{If} \quad x \in S_0, \quad \lim_{t \to \infty} V(x;t) = f(x)$$

$$\text{where} \quad V(x;t) = f(x) - U(x;t).$$

In this formulation, V is the penalty function and U corresponds to the sum of the multiplier functions. A relationship to the general exponential penalty function of equation (3.1) can now be stated. Consider the special case

$$P(x;\, \alpha,\, 1/\alpha,\, 0) = f(x) - 1/\alpha \sum_{j=1}^{m} [h(g_j(x))]^{\alpha} .$$

If $x \in S_0$, $g_j(x) \leq 0$, $j = 1,\ldots,m$; thus by property ii) of the function h, $h(g_j(x)) \leq 1$, $j = 1,\ldots,m$. Hence for feasible points

$$\lim_{\alpha \to \infty} P(x;\, \alpha,\, 1/\alpha,\, 0) = f(x) .$$

Thus the Fiacco-McCormick condition above is satisfied and it can be shown that the other conditions for a general exterior penalty function are also satisfied by $P(x;\, \alpha,\, 1/\alpha,\, 0)$.

If we retain the condition $\beta = 1/\alpha$, but select $\delta > 0$, then for any point on the boundary of S_0

$$\lim_{\alpha \to \infty} P(x;\, \alpha,\, 1/\alpha,\, \delta) = -\infty$$

hence $\delta > 0$ does not yield a general exterior penalty function.

The following analogues of Lemmas 3.1 and 3.4 can be stated; they are given here without proofs. Assume conditions A-1, A-2 hold and that A-3 holds with $\beta = 1/\alpha$.

LEMMA 5.1: Suppose $\delta > 0$ is such that $S_{-\delta e}$ is nonempty. Then there exists a scalar $\bar{\alpha}$ such that $\alpha > \bar{\alpha}$ implies $P(x;\alpha,1/\alpha,\delta)$ has a global maximizer and each such maximizer belongs to S_0.

LEMMA 5.2: Suppose $\delta > 0$ is such that $S_{\delta e}$ is compact. Then there is a scalar $\bar{\alpha}$ such that $\alpha > \bar{\alpha}$ implies that $P(x;\alpha,1/\alpha,0)$ has a global maximizer and each such maximizer belongs to $S_{\delta e}$.

These results differ from Lemma 3.1 and 3.4 in that here β is no longer fixed.

REFERENCES

[1] Evans, J.P. and F.J. Gould, "Stability in Nonlinear Programming," *Operations Research*, Vol. 18: 107-118 (1970).

[2] Fiacco, A.V. and G.P. McCormick, *Nonlinear Programming: Sequential Uncontrained Minimization Techniques*, John Wiley and Sons, Inc. (1968).

[3] Goldstein, A.A. and B.R. Kripke, "Mathematical Programming by Minimizing Differentiable Functions," *Numer. Math.*, Vol. 6: 47-48 (1964).

[4] Gould, F.J., "Extensions of Lagrange Multipliers in Nonlinear Programming," *SIAM J. Appl. Math.*, Vol. 19: 1280-1297 (1969).

[5] Motzkin, T.S., "New Techniques for Linear Inequalities and Optimization," in *Project SCOOP, Symposium on Linear Inequalities and Programming*, Planning Research Division, Director of Management Analysis Service, U.S. Air Force, Washington, D.C., No. 10, 1952.

[6] Zangwill, W.A., "Nonlinear Programming via Penalty Functions," *Mgmt. Sci.*, Vol. 13: 344-358 (1967).