JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 27, No. 2, FEBRUARY 1979

An Exponential Penalty Method for Nondifferentiable Minimax Problems with General Constraints

J. J. STRODIOT[1] AND V. H. NGUYEN[2]

Communicated by A. V. Fiacco

Abstract. A well-known approach to constrained minimization is via a sequence of unconstrained optimization computations applied to a penalty function. This paper shows how it is possible to generalize Murphy's penalty method for differentiable problems of mathematical programming (Ref. 1) to solve nondifferentiable problems of finding saddle points with constraints. As in mathematical programming, it is shown that the method has the advantages of both Fiacco and McCormick exterior and interior penalty methods (Ref. 2). Under mild assumptions, the method has the desirable property that all trial solutions become feasible after a finite number of iterations. The rate of convergence is also presented. It should be noted that the results presented here have been obtained without making any use of differentiability assumptions.

Key Words. Exponential penalty method, saddle points, convex analysis, minimax problems, general constraints.

1. Introduction

Problems of finding saddle points form a large class of problems encountered both in various types of game situations and also in intrinsically mathematical problems, for example, in the problem of nonlinear programming under Lagrangian formulations.

Let C_0 and D_0 be subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, and let f be a real function defined on $\mathbb{R}^n \times \mathbb{R}^m$. We will be interested in finding (if it exists) a saddle

[1] Premier Assistant, Department of Mathematics, Facultés Universitaires de Namur, Namur, Belgium.

[2] Chargé de Cours, Department of Mathematics, Facultés Universitaires de Namur, Namur, Belgium.

0022-3239/79/0200-0205$03.00/0

point of f with respect to $C_0 \times D_0$, i.e., a point $(\bar{x}, \bar{y})$ in $C_0 \times D_0$ such that, for every (x, y) in $C_0 \times D_0$,

$$f(\bar{x}, y) < f(\bar{x}, \bar{y}) < f(x, \bar{y}).$$

Dem'yanov and Pevnyi (Ref. 3) give an excellent survey of different numerical methods for finding saddle points. One of these methods is the penalty method. The strategy is to transform the constrained problem into a sequence of unconstrained problems which are considerably easier to solve than the original problem. The penalty functions are constructed so that all convergent subsequences of solutions of unconstrained problems converge satisfactorily to a saddle point of f with respect to $C_0 \times D_0$.

Sasai (Ref. 4) proposed an interior penalty method, and Auslender (Ref. 5) proposed an exterior penalty method. The disadvantages of these methods are the same as in nonlinear programming: the search of an interior starting point for interior methods and the unnecessary feasibility of the sequence of trial solutions for exterior methods. Fundamental results dealing with mathematical programming problems and a survey of the literature can be found in the monograph of Fiacco and McCormick (Ref. 2), which basically deals with the penalty method.

In this paper, we propose and analyze some aspects of an exponential penalty method. The proposed method is a generalization of Murphy's penalty method for differentiable problems in nonlinear programming (Ref. 1) to nondifferentiable minimax problems.

Contrary to interior penalty methods, the method does not require an interior point (to the feasible inequality-constrained region) to initialize computations. On the other hand, we can guarantee that, after a finite number of iterations, the method does produce intermediate solutions that are feasible under mild assumptions. Hence, the proposed method has the advantages of the interior and exterior penalty methods without having their drawbacks.

In Section 2, we formulate the problem; in Section 3, we show that, without the well-known Slater constraint qualification (Ref. 6, Section 5.4.3), *almost all* trial solutions remain in a compact subset and prove a convergence theorem. Section 4 contains the main result: trial solutions become feasible after a finite number of iterations when the Slater condition is satisfied. The key of the proof is based on a boundedness property for the set of subgradients of a family of convex functions. This property enables us to give the proof without differentiability assumptions. Section 5 gives an estimation of the rate of convergence of the algorithm. The final section of the paper contains proofs of theorems and lemmas.

All the definitions concerning convex analysis are those of Rockafellar (Ref. 7).

2. Problem Formulation and Assumptions

Let f be a real function defined on $\mathbb{R}^n \times \mathbb{R}^m$, and let g_i, $i = 1, \dots, p$, and h_j, $j = 1, \dots, q$, be real functions defined on $\mathbb{R}^n$ and on $\mathbb{R}^m$, respectively. We introduce the following notation:

$$\tilde{g}(x) \equiv \max_{1 \leq i \leq p} g_i(x), \qquad \tilde{h}(y) \equiv \max_{1 \leq j \leq q} h_j(y), \tag{1}$$

$$C_\delta \equiv \{x \in \mathbb{R}^n \mid \tilde{g}(x) \leq \delta\}, \tag{2}$$

$$D_\delta \equiv \{y \in \mathbb{R}^m \mid \tilde{h}(y) \leq \delta\}, \tag{3}$$

for $\delta \geq 0$. Then we consider the following problem.

Problem (P). Find a saddle point of f with respect to $C_0 \times D_0$.

We impose the following assumptions:

(A1) f is a convex–concave function (i.e., convex in the first argument and concave in the second argument);

(A2) g_i and h_j are convex functions for all $i = 1, \dots, p$ and $j = 1, \dots, q$;

(A3) C_0 and D_0 are nonempty bounded subsets.

Note that Assumptions (A1) and (A2) imply that $f(\cdot, y)$ and $f(x, \cdot)$ are continuous functions for all x, y and that g_i and h_j are continuous functions for all i and j (Ref. 7, Corollary 10.1.1). Hence, C_0 and D_0, being bounded subsets, are compact; and, by the classical minimax theorem (Ref. 8, Chapter 6, Proposition 2.1), Problem (P) has at least one solution. Moreover, observe that C_δ and D_δ are bounded subsets for all $\delta > 0$. This is a direct consequence of (A2), (A3), and Ref. 7, Corollary 8.7.1.

For solving Problem (P), we propose a penalty method. We replace the original problem by a sequence of problems which are easier to solve. More precisely, for each integer $k \geq 1$, we solve a problem of the following form.

Problem ($\mathbf{P}_k$). Find a saddle point of F_k with respect to $\mathbb{R}^n \times \mathbb{R}^m$, where

$$F_k : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$$

is the function given by the relation

$$F_k(x, y) = f(x, y) + (1/s(k)) \left\{ \sum_{i=1}^{p} \exp[r(k) g_i(x)] - \sum_{j=1}^{q} \exp[r(k) h_j(y)] \right\}, \tag{4}$$

where $r(k)$ and $s(k)$ are real parameters satisfying

$$r(k) \to +\infty \text{ as } k \to +\infty \qquad \text{and} \qquad r(k) \geq s(k) \geq 1 \qquad \text{for all } k.$$

For example, one can take

$$r(k)=k, \qquad s(k)=1,$$

or

$$r(k)=s(k)=a^k,$$

or

$$r(k)=b^k, \qquad s(k)=a^k,$$

where a, b are real numbers such that $1<a<b$. By generating solutions of the sequence of Problems (P_k) which converge to a solution of the Problem (P), we obtain a new algorithm for solving a minimax problem subject to inequality constraints.

3. Existence of Solutions for Problem $(\mathbf{P}_k)$ and Convergence Theorems

Before giving convergence conditions for the algorithm, it is important to see that each Problem (P_k) is well defined. This is stated precisely by the following two theorems.

Theorem 3.1. If there exist $x_0 \in \mathbb{R}^n$ and $y_0 \in \mathbb{R}^m$ such that

$$\lim_{\|y\|\to+\infty} f(x_0, y) < +\infty, \qquad \lim_{\|x\|\to+\infty} f(x, y_0) > -\infty, \tag{5}$$

then Problem (P_k) has at least one solution for all k.

The next theorem shows that Assumption (5) can be dropped for k sufficiently large.

Theorem 3.2. Let $\delta>0$ be given. If $r(k)/s(k)$ tends to infinity, then, for k sufficiently large, F_k has at least a saddle point with respect to $\mathbb{R}^n \times \mathbb{R}^m$ and all unconstrained saddle points of F_k are in $C_\delta \times D_\delta$.

Comment. As C_δ and D_δ are compact subsets, the function F_k has a saddle point with respect to $C_\delta \times D_\delta$ by the classical minimax theorem. We first prove that this constrained saddle point is an unconstrained one and afterward we show that all unconstrained saddle points of F_k are in $C_\delta \times D_\delta$. The key of the proof is based on the following lemma.

Lemma 3.1. If C and D are compact subsets of $\mathbb{R}^n$ and R^m, then $f(C \times D)$ is a bounded subset of $\mathbb{R}$ and

$$\cup\{\partial f(x, y) \mid x \in C, y \in D\}$$

is a bounded subset of $\mathbb{R}^n$, where $\partial f(x, y)$ is the subgradient of the convex function $f(\cdot, y)$ at x, i.e.,

$$\partial f(x, y) = \{x^* \in \mathbb{R}^n \mid f(z, y) \geq f(x, y) + (x^*, z - x), \forall z \in \mathbb{R}^n\}.$$

This lemma enables us to give the proof of Theorem 3.2 without differentiability assumptions. Observe also that Theorem 3.2 requires no Slater condition.

Introduce now the following subsets:

$$C_0^0 \equiv \{x \in \mathbb{R}^n \mid g_i(x) < 0, i = 1, \ldots, p\},$$
$$D_0^0 \equiv \{y \in \mathbb{R}^m \mid h_j(y) < 0, j = 1, \ldots, q\}.$$

Observe that, by Ref. 7, Corollary 7.6.1, if C_0^0 and D_0^0 are nonempty, then

$$\overline{C_0^0} = C_0, \qquad \overline{D_0^0} = D_0. \tag{6}$$

We are now in a position to present a convergence theorem.

Theorem 3.3. For each k, let (x_k, y_k) be a saddle point of F_k with respect to $\mathbb{R}^n \times \mathbb{R}^m$. If $r(k)/s(k)$ tends to infinity, then each accumulation point (and there exists at least one) of the sequence $\{(x_k, y_k)\}_k$ belongs to $C_0 \times D_0$. Moreover, if $s(k)$ tends to infinity or if C_0^0 and D_0^0 are nonempty, then all accumulation points of $\{(x_k, y_k)\}_k$ are saddle points of f with respect to $C_0 \times D_0$.

Observe that, if $s(k)$ tends to infinity, then all accumulation points of $\{(x_k, y_k)\}_k$ are solutions of Problem (P) without requiring a Slater condition.

We close this section by stating the following corollary, which is a direct consequence of Theorem 3.3.

Corollary 3.1. Let the conditions stated in Theorem 3.3 be satisfied. Then, $\{f(x_k, y_k)\}_k$ and $\{F_k(x_k, y_k)\}_k$ converge to the saddle value $f(\bar{x}, \bar{y})$ of f with respect to $C_0 \times D_0$. Moreover, if the saddle point $(\bar{x}, \bar{y})$ of f with respect to $C_0 \times D_0$ is unique, then the whole sequence $\{(x_k, y_k)\}_k$ converges to $(\bar{x}, \bar{y})$.

4. Main Result

In this section, we prove that the exponential penalty method has the advantage of interior penalty methods: after a finite number of iterations, the trial solutions are feasible. For this purpose, as for interior methods, we suppose that the feasible domain has a nonempty interior.

Theorem 4.1. If C_0^0 and D_0^0 are nonempty and if $r(k)/s(k)$ tends to infinity, then the trial solutions (x_k, y_k) become feasible for k sufficiently large.

Comment. By Theorem 3.2, the trial solutions (x_k, y_k) belong to $C_\delta \times D_\delta$ for k sufficiently large, where δ is a fixed positive scalar. Hence, x_k is a solution of the following convex problem

$$\text{minimize } F_k(x, y_k), \qquad \text{subject to } x \in C_\delta. \tag{7}$$

We show that the solution of (7) belongs to C_0, provided k is large enough. The result was given in Ref. 1, Theorem 3, for nonlinear programming problems under differentiability assumptions. Differentiability of f is not assumed here.

5. Rate of Convergence

Note that, since the trial solutions (x_k, y_k) belong to $C_0 \times D_0$ after a finite number of iterations, we do need only to establish the rate of convergence for feasible trial solutions.

Theorem 5.1. If (x_k, y_k) is an unconstrained saddle point of F_k and if (x_k, y_k) belongs to $C_0 \times D_0$, then

$$|f(x_k, y_k) - \bar{\delta}| \leq (p+q)/s(k),$$

where $\bar{\delta}$ is the saddle value of f with respect to $C_0 \times D_0$.

Observe that the rate of convergence of the sequence $\{f(x_k, y_k)\}_k$ depends heavily on the rate of convergence of the sequence $\{s(k)\}_k$ to infinity. A similar result is obtained for the sequence $\{(x_k, y_k)\}_k$, but under assumptions which are stronger than those of Theorem 5.1.

We make use of a definition that was given in Ref. 9.

Definition 5.1. A real-valued function K is uniformly convex (respectively, concave) on a convex set T if there exists a nondecreasing function $\delta_1(t) > 0$ (respectively, $\delta_2(t) > 0$) on $(0, \infty)$ such that, for $x, y \in T$:

$$K((x+y)/2) \leq (1/2)K(x) + (1/2)K(y) - \delta_1(\|x-y\|)$$

respectively, $K((x+y)/2) \geq (1/2)K(x) + (1/2)K(x) + \delta_2(\|x-y\|)]$.

Then, we obtain the following theorem which gives the convergence rate of the algorithm.

Theorem 5.2. Let the conditions of Theorem 5.1 be satisfied. If f has a unique saddle point $(\bar{x}, \bar{y})$ with respect to $C_0 \times D_0$, if $f(\cdot, \bar{y})$ is uniformly convex (with function δ_1), and if $f(\bar{x}, \cdot)$ is uniformly concave (with function δ_2), then

$$0 \leq \delta_1(\|\bar{x} - x_k\|) \leq (p+q)/s(k),$$

$$0 \leq \delta_2(\|\bar{y} - y_k\|) \leq (p+q)/s(k).$$

For example, if

$$\delta_1(t) = \delta_2(t) = t^2,$$

then, for all k,

$$\max(\|\bar{x} - x_k\|, \|\bar{y} - y_k\|) \leq ((p+q)/s(k))^{1/2}.$$

Remark 5.1. Finally, by Theorems 4.1, 5.1, and 5.2, it seems that a good choice for the parameters $r(k)$ and $s(k)$ is as follows:

$$\{r(k)\}_k, \{r(k)/s(k)\}_k, \text{ and } \{s(k)\}_k \text{ tend to infinity.}$$

6. Proofs of Theorems, Corollaries, and Lemmas

Proof of Theorem 3.1. For each $y \in \mathbb{R}^m$ we have, using (1) and (4),

$$F_k(x_0, y) \leq f(x_0, y) + (1/s(k)) \sum_{i=1}^{p} \exp[r(k) g_i(x_0)] - (1/s(k)) \exp[r(k)\tilde{h}(y)]. \tag{8}$$

As D_δ is a bounded subset for each $\delta \geq 0$, $\tilde{h}(y)$ tends to infinity when $\|y\| \to +\infty$. Hence, using (5), the right-hand side of (8) converges to $-\infty$ when $\|y\| \to +\infty$ and

$$\lim_{\|y\| \to +\infty} F_k(x_0, y) = -\infty. \tag{9}$$

Following a similar reasoning, one obtains

$$\lim_{\|x\| \to +\infty} F_k(x, y_0) = +\infty. \tag{10}$$

So, by (9), (10), and Ref. 8, Chapter 6, Proposition 2.2, Problem (P_k) has at least one solution. □

Proof of Theorem 3.2. Let $\delta > 0$ be given. We shall first prove that F_k has an unconstrained saddle point belonging to $C_\delta \times D_\delta$ for k sufficiently large (Claims 6.1, 6.2, 6.3). Afterward, we show that all unconstrained saddle points of F_k are in $C_\delta \times D_\delta$ (Claim 6.4).

Claim 6.1. There exists an integer k_1 such that, for all $x \in \mathbb{R}^n \backslash C_\delta$ and all $y \in D_\delta$,

$$-f(x, y) \leq r(k_1)\tilde{g}(x). \tag{11}$$

Proof. Since C_0 is nonempty, we can find a point x_0 such that $\tilde{g}(x_0) < \delta$. It then follows from Ref. 10, Lemma 3.1, that there exists $\epsilon > 0$ such that, for all $x \in \mathbb{R}^n \backslash C_\delta$, there is a z on the boundary of C_δ such that

$$\tilde{g}(x) \geq \epsilon \|x - z\|. \tag{12}$$

Let $x \in \mathbb{R}^n \backslash C_\delta$, $y \in D_\delta$, and z be the point corresponding to x. As $f(\cdot, y)$ is a convex function, for each $c(z, y) \in \partial f(z, y)$,

$$f(x, y) \geq f(z, y) + \langle x - z, c(z, y) \rangle, \tag{13}$$

where $\partial f(z, y)$ is the set of subgradients of $f(\cdot, y)$ at z (see Refs. 7 and 8, for example). Hence, using Lemma 3.1 with $C = C_\delta$ and $D = D_\delta$, (13) and (12), we see that there exist $p_1, p_2 > 0$ such that, for all $x \in \mathbb{R}^n \backslash C_\delta$ and all $y \in D_\delta$,

$$\begin{aligned} -f(x, y) &\leq p_1 + \|x - z\| p_2 \\ &\leq p_1 + \tilde{g}(x) p_2/\epsilon \\ &\leq K\tilde{g}(x), \end{aligned}$$

where

$$K = (p_1/\delta) + (p_2/\epsilon).$$

Since $r(k)$ tends to infinity, we easily obtain (11). □

Claim 6.2. If $r(k)/s(k)$ tends to infinity, then there exists k_2 such that, for all $k \geq k_2$, all $x \in \mathbb{R}^n \backslash C_\delta$, and all $y \in D_\delta$,

$$F_k(x, y) \geq \min_{z \in C_\delta} F_k(z, y). \tag{14}$$

Proof. Let $\bar{x}$ be a point in C_0. As

$$\sum_{i=1}^{p} \exp[r(k) g_i(\bar{x})] \leq p,$$

for all k and all $y \in D_\delta$ we have

$$\min_{z \in C_\delta} F_k(z, y) \leq f(\bar{x}, y) + p/s(k) - (1/s(k)) \sum_{j=1}^{q} \exp[r(k) h_j(y)]. \tag{15}$$

Since $r(k)/s(k)$ tends to infinity and $f(\bar{x}, y) + p/s(k)$ is bounded for all k and for all $y \in D_\delta$, we can choose k' such that $k \geq k'$, $x \in \mathbb{R}^n \backslash C_\delta$, and $y \in D_\delta$ imply that

$$f(\bar{x}, y)+p/s(k) \leq r(k)\delta/2s(k) \leq r(k)\tilde{g}(x)/2s(k). \tag{16}$$

On the other hand, using Claim 6.1, we can choose k'' such that $k \geq k''$, $x \in \mathbb{R}^n \backslash C_\delta$, and $y \in D_\delta$ imply that

$$-f(x, y) \leq r(k)\tilde{g}(x)/2s(k). \tag{17}$$

Gathering (16) and (17) and using (1), we obtain, for $k \geq k_2 = \sup(k', k'')$,

$$\begin{aligned} -f(x, y)+f(\bar{x}, y)+p/s(k) &\leq r(k)\tilde{g}(x)/s(k) \leq \exp[r(k)\tilde{g}(x)]/s(k) \\ &\leq (1/s(k)) \sum_{i=1}^{p} \exp[r(k)g_i(x)]. \end{aligned} \tag{18}$$

Finally, from (18) and (15), we immediately deduce (14). □

Claim 6.3. For k sufficiently large, F_k has an unconstrained saddle point belonging to $C_\delta \times D_\delta$.

Proof. Let (x_k, y_k) be a saddle point of F_k with respect to $C_\delta \times D_\delta$. Then, by the definition of a saddle point and Claim 6.2, we have, for k sufficiently large and for $x \in \mathbb{R}^n \backslash C_\delta$,

$$F_k(x_k, y_k) \leq \min_{z \in C_\delta} F_k(z, y_k) \leq F_k(x, y_k).$$

Hence, for all $x \in \mathbb{R}^n$,

$$F_k(x_k, y_k) \leq F_k(x, y_k). \tag{19}$$

By a similar argument, for k sufficiently large and for $y \in \mathbb{R}^m$,

$$F_k(x_k, y) \leq F_k(x_k, y_k). \tag{20}$$

Then, Claim 6.3 follows from (19) and (20). □

Claim 6.4. If F_k has an unconstrained saddle point belonging to $C_\delta \times D_\delta$, then all unconstrained saddle points of F_k belong to $C_\delta \times D_\delta$.

Proof. Let (x_1, y_1) and (x_2, y_2) be two unconstrained saddle points of F_k. Suppose that $(x_1, y_1) \in C_\delta \times D_\delta$ and that $x_2 \notin C_\delta$. Then, there exists $i \in \{1, \ldots, p\}$ such that

$$g_i(x_2) > \delta \geq g_i(x_1). \tag{21}$$

Since g_i is convex and $\exp[r(k)x]$ is a strictly convex function, we have, using (21),

$$\exp[r(k)g_i(z_1)] \leq \exp[r(k)(\lambda g_i(x_1)+(1-\lambda)g_i(x_2))]$$
$$< \lambda \exp[r(k)g_i(x_1)]+(1-\lambda)\exp[r(k)g_i(x_2)],$$

where

$$z_1=\lambda x_1+(1-\lambda)x_2, \qquad 0<\lambda<1.$$

Hence,

$$F_k(z_1, z_2)<\lambda F_k(x_1, z_2)+(1-\lambda)F_k(x_2, z_2),$$

where

$$z_2=\lambda y_1+(1-\lambda)y_2;$$

and, (x_i, y_i) being saddle points, $i=1,2$,

$$F_k(z_1, z_2)<\lambda F_k(x_1, y_1)+(1-\lambda)F_k(x_2, y_2). \tag{22}$$

On the other hand, since $F_k(z_1, \cdot)$ is a concave function, we have

$$F_k(z_1, z_2) \geq \lambda F_k(z_1, y_1)+(1-\lambda)F_k(z_1, y_2)$$
$$\geq \lambda F_k(x_1, y_1)+(1-\lambda)F_k(x_2, y_2), \tag{23}$$

contrary to relation (22). It follows that $x_2 \in C_\delta$, and similarly $y_2 \in D_\delta$. □

Proof of Lemma 3.1. For each $y \in D$, we denote

$$B_y = \cup\{\partial f(x, y) \mid x \in C\}.$$

By Ref. 7, Theorem 24.7, B_y is a bounded subset of $\mathbb{R}^n$. Suppose now that $\cup\{B_y \mid y \in D\}$ is unbounded. Then,

$$\forall n, \exists y_n \in D \text{ and } c_n \in B_{y_n} \text{ such that } \|c_n\| \geq n. \tag{24}$$

So we obtain a sequence $\{y_n\}_n$ contained in D. Since D is compact, there exists a subsequence $\{y_{n_k}\}_k$ of $\{y_n\}$ converging to a point $\bar{y} \in D$. If we show that

$$\exists k_0 \text{ such that, } \forall k \geq k_0 \text{ and } x \in C,\ \partial f(x, y_{n_k}) \subseteq B_{\bar{y}}+B, \tag{25}$$

where B is the Euclidean unit ball of $\mathbb{R}^n$, then we obtain a contradiction with (24). Indeed, as $B_{\bar{y}}+B$ is a bounded subset of $\mathbb{R}^n$, there exists $\delta>0$ such that $\|z\|<\delta$ for all $z \in B_{\bar{y}}+B$. Let k be an integer such that $k \geq k_0$ and $n_k \geq \delta$. Then, by (24), there exist $x \in C$ and $c_{n_k} \in \partial f(x, y_{n_k})$ such that $\|c_{n_k}\| \geq \delta$. On the other hand, by (25), we have $c_{n_k} \in B_{\bar{y}} + B$, and thus $\|c_{n_k}\| < \delta$. So, (24) contradicts (25).

It remains to prove (25). For notational simplicity, we denote again by $\{y_n\}_n$ the sequence $\{y_{n_k}\}_k$. Then, (25) becomes

$$\exists n_0 \text{ such that}, \forall n \geq n_0 \text{ and } x \in C,\ \partial f(x, y_n) \subseteq B_{\bar{y}} + B. \tag{26}$$

To establish this result, suppose that the conclusion (26) is false. Then, there exists a subsequence $\{y_{n_k}\}_k$ of $\{y_n\}$ and a sequence $\{x_k\}_k$ in C such that, for all k, we have

$$\partial f(x_k, y_{n_k}) \not\subseteq B_{\bar{y}} + B. \tag{27}$$

Since C is compact, there exists a subsequence $\{x_{k_j}\}_j$ of $\{x_k\}_k$ converging to $\bar{x} \in C$, and consider the subsequence $\{y_{n_{k_j}}\}_j$ of $\{y_{n_k}\}_k$. Then, since $\{x_{k_j}\}_j$ converges to $\bar{x}$ and $\{f(\cdot\,, y_{n_{k_j}})\}_j$ is a sequence of finite convex functions on $\mathbb{R}^n$ converging pointwise to $f(\cdot\,, \bar{y})$, we obtain, by Ref. 7, Theorem 24.5, for j sufficiently large,

$$\partial f(x_{k_j}, y_{n_{k_j}}) \subseteq \partial f(\bar{x}, \bar{y}) + B \subseteq B_{\bar{y}} + B, \tag{28}$$

contrary to relation (27). It follows that (26) is true. This completes the proof of Lemma 3.1. □

Proof of Theorem 3.3. Since C_δ and D_δ are compact subsets and, by Theorem 3.2, (x_k, y_k) belongs to $C_\delta \times D_\delta$ for k sufficiently large, the sequence $\{(x_k, y_k)\}_k$ has at least one accumulation point. Let $(\bar{x}, \bar{y})$ be such a point. We denote again by $\{(x_k, y_k)\}_k$ the subsequence of $\{(x_k, y_k)\}_k$ converging to $(\bar{x}, \bar{y})$. Then, $(\bar{x}, \bar{y}) \in C_0 \times D_0$. Indeed, if $\bar{x} \notin C_0$, there exists $\delta > 0$ such that $\bar{x} \notin C_\delta$.

On the other hand, we have $x_k \to \bar{x}$. Since $x_k \in C_\delta$ for large k (Theorem 3.2) and C_δ is closed, $\bar{x} \in C_\delta$. Hence, $\bar{x} \in C_0$. Similarly, $\bar{y} \in D_0$. We now prove that $(\bar{x}, \bar{y})$ is a saddle point of f with respect to $C_0 \times D_0$. For this purpose, let (x, y) be a point of $C_0 \times D_0$. Then, (x_k, y_k) being a saddle point of F_k, we have

$$f(x_k, y) - (1/s(k)) \sum_{j=1}^{q} \exp[r(k)h_j(y)]$$
$$\leq F_k(x_k, y_k) \leq f(x, y_k) + (1/s(k)) \sum_{i=1}^{p} \exp[r(k)g_i(x)].$$

If $s(k)$ tends to infinity, we obtain immediately

$$f(\bar{x}, y) \leq f(x, \bar{y}). \tag{29}$$

If C_0^0 and D_0^0 are nonempty, then (29) is satisfied for each $(x, y) \in C_0^0 \times D_0^0$. Using (6) and the continuity of f in each argument, we obtain (29) for each

$(x, y) \in C_0 \times D_0$. Hence, in the two cases, $(\bar{x}, \bar{y})$ is a saddle point of f with respect to $C_0 \times D_0$. □

Proof of Theorem 4.1. Let δ be a fixed positive scalar. We must show that the solution of (7) belongs to C_0 for k sufficiently large, i.e., that for each $x \in C_\delta \backslash C_0$, one can find a point $x_B \in C_0$ such that

$$F_k(x_B, y_k) < F_k(x, y_k).$$

Let x_0 be a point in C_0^0.

Claim 6.5. Let $\mathrm{bd}(C_0)$ denote the boundary of C_0. There is an integer k_1 such that, for all $k \geq k_1$, one has

$$F'_k(x_B, y_k; x_0 - x_B) < 0 \qquad \text{for all } x_B \in \mathrm{bd}(C_0),$$

where $F'_k(x_B, y_k; x_0 - x_B)$ is the directional derivative of the convex function $F_k(\cdot, y_k)$ at x_B in the direction $x_0 - x_B$.

Proof. Let x_B be a point in $\mathrm{bd}(C_0)$. From the definition of F_k [see (4)],

$$F'_k(x_B, y_k; x_0 - x_B)$$
$$= f'(x_B, y_k; x_0 - x_B) + \sum_{i=1}^{p} \exp[r(k) g_i(x_B)] g'_i(x_B; x_0 - x_B) r(k)/s(k), \tag{30}$$

where $f'(x_B, y_k; x_0 - x_B)$ and $g'_i(x_B; x_0 - x_B)$ denote the directional derivatives of the convex function $f(\cdot, y_k)$ and g_i at x_B in the direction $x_0 - x_B$. Then, using Ref. 7, Theorems 23.4 and 24.7, and Lemma 3.1, we obtain that

$$g'_i(x; x_0 - x_B) \qquad \text{is bounded on } C_0 \text{ for each } i, \tag{31}$$

$$f'(x, y_k; x_0 - x_B) \qquad \text{is bounded on } C_0, \text{ uniformly in } k. \tag{32}$$

We now examine the second term of the right-hand side of (30). For this purpose, we introduce the notations

$$2\gamma \equiv \tilde{g}(x_0), \tag{33}$$

and, for $x \in C_0$,

$$I(x) \equiv \{i \mid g_i(x) = 0\}, \tag{34}$$

$$J(x) \equiv \{i \mid \gamma < g_i(x) < 0\}, \tag{35}$$

$$K(x) \equiv \{i \mid g_i(x) \leq \gamma\}. \tag{36}$$

Then, for $x \in C_0$ with $K(x) \neq \phi$ and for $i \in K(x)$, we have

$$\exp[r(k)g_i(x)]|g_i'(x;x_0-x)|r(k)/s(k) \leq \exp[r(k)\gamma]|g_i'(x;x_0-x)|r(k)/s(k). \quad (37)$$

Since $s(k) \geq 1$ and $\exp[r(k)\gamma]r(k)$ tends to zero when $k \to +\infty$, the right-hand side of (37) converges to zero. Moreover, by (31), the convergence is uniform for $x \in C_0$ with $K(x) \neq \phi$. Hence, there exists k' such that, for any $k \geq k'$ and any $x \in C_0$ with $K(x) \neq \phi$, we have

$$\exp[r(k)g_i(x)]|g_i'(x;x_0-x)|r(k)/s(k) < 1 \qquad \text{for all } i \in K(x). \quad (38)$$

On the other hand, by (33), we have, for each $x_i^* \in \partial g_i(x)$ and each $x \in C_0$, that

$$0 > 2\gamma \geq g_i(x_0) \geq g_i(x) + (x_i^*, x_0 - x), \quad (39)$$

since g_i is a convex function. Hence, for $x \in C_0$ with $J(x) \neq \phi$ and for $i \in J(x)$, we obtain

$$g_i'(x;x_0-x) \leq 2\gamma - g_i(x) < \gamma, \quad (40)$$

$$\exp[r(k)g_i(x)]g_i'(x;x_0-x)r(k)/s(k) \leq \gamma \exp[r(k)g_i(x)]r(k)/s(k) < 0, \quad (41)$$

where Ref. 7, Theorem 23.4, and (39) have been used. Finally, for any $x \in C_0$ with $I(x) \neq \phi$ and any $i \in I(x)$, the inequality (40) holds again, and hence

$$g_i'(x;x_0-x)r(k)/s(k) \leq \gamma r(k)/s(k). \quad (42)$$

To finish the proof, we observe that $I(x)$ is nonempty for each $x \in \text{bd}(C_0)$. We can now invoke (32), (38), (41), (42) to deduce that

$$F_k'(x_B, y_k; x_0 - x_B) \leq M + p + \gamma r(k)/s(k) \quad (43)$$

for $k \geq k_0$ and $x_B \in \text{bd}(C_0)$. Since $r(k)/s(k)$ tends to infinity and $\gamma < 0$, the conclusion of the claim is evident. □

Claim 6.6. x_k belongs to C_0 for k sufficiently large.

Proof. Since x_0 belongs to C_0^0, for each $x \in C_\delta \backslash C_0$, there exists $x_B \in \text{bd}(C_0)$ such that

$$x = x_B + q(x_0 - x_B) \qquad \text{with } q < 0. \quad (44)$$

Let k be an integer greater than k_1. Since $F_k(\cdot, y_k)$ is a convex function, we see, using Claim 6.5 and (44), that for all $x \in C_\delta \backslash C_0$, there exists $x_B \in C_0$ such that

$$F_k(x_B, y_k) < F_k(x, y_k). \quad (45)$$

Since $F_k(\cdot, y_k)$ is continuous and C_0 is compact, the minimum of $F_k(\cdot, y_k)$ over C_0 exists and, by (45), is the global minimum of $F_k(\cdot, y_k)$ on C_δ. That is, x_k belongs to C_0 for $k \geq k_1$. □

By a similar argument, we prove that y_k belongs to D_0 for k sufficiently large. Hence, the trial solutions become feasible after a finite number of iterations.

Proof of Theorem 5.1. Let $(\bar{x}, \bar{y})$ be a saddle point of f with respect to $C_0 \times D_0$. Then,

$$f(\bar{x}, \bar{y}) = \bar{\delta}.$$

Moreover, as $x_k \in C_0$, we easily obtain

$$\begin{aligned} f(\bar{x}, \bar{y}) &\leq f(x_k, \bar{y}) \\ &= F_k(x_k, \bar{y}) - (1/s(k))\Big\{\sum_i \exp[r(k) g_i(x_k)] - \sum_j \exp[r(k) h_j(\bar{y})]\Big\}. \end{aligned} \tag{46}$$

Since

$$h_j(\bar{y}) \leq 0, \qquad g_i(x_k) \leq 0,$$

and (x_k, y_k) is a saddle point of F_k, it follows that

$$\begin{aligned} f(\bar{x}, \bar{y}) &\leq F_k(x_k, \bar{y}) + q/s(k) \leq F_k(x_k, y_k) + q/s(k) \\ &\leq f(x_k, y_k) + (p+q)/s(k). \end{aligned} \tag{47}$$

By a similar reasoning, we obtain the other inequality. □

Proof of Theorem 5.2. Since $(1/2)(\bar{x} + x_k) \in C_0$ and $f(\cdot, \bar{y})$ is uniformly convex, we have

$$f(\bar{x}, \bar{y}) \leq f((1/2)(\bar{x} + x_k), \bar{y}) \leq (1/2) f(\bar{x}, \bar{y}) + (1/2) f(x_k, \bar{y}) - \delta_1(\|\bar{x} - x_k\|);$$

consequently,

$$\delta_1(\|\bar{x} - x_k\|) \leq (1/2) f(x_k, \bar{y}) - (1/2) f(\bar{x}, \bar{y}). \tag{48}$$

On the other hand, using (46) and (47), we obtain

$$f(x_k, \bar{y}) \leq f(x_k, y_k) + (p+q)/s(k).$$

Hence, (48) becomes

$$\delta_1(\|\bar{x} - x_k\|) \leq (1/2) f(x_k, y_k) + (p+q)/s(k) - (1/2) f(\bar{x}, \bar{y}).$$

Applying now Theorem 5.1, we obtain the first part of Theorem 5.2. A similar reasoning gives us the second part. □

7. Conclusions

This paper presents a new penalty method for solving nondifferentiable saddle-point problems with constraints. Our main result is Theorem 4.1. As in nonlinear programming, the exponential penalty method has the advantages of both the interior and exterior penalty methods without having their drawbacks. This theorem enables us to give a rate-of-convergence estimate.

References

1. MURPHY, F., *A Class of Exponential Penalty Functions*, SIAM Journal on Control, Vol. 12, pp. 679–687, 1974.
2. FIACCO, A. V., and MCCORMICK, G. P., *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley, New York, New York, 1968.
3. DEM'YANOV, V., and PEVNYI, A., *Numerical Methods for Finding Saddle Points*, USSR Computational Mathematics and Mathematical Physics, Vol. 12, pp. 11–52, 1972.
4. SASAI, H., *An Interior Penalty Method for Minimax Problems with Constraints*, SIAM Journal on Control, Vol. 12, pp. 643–649, 1974.
5. AUSLENDER, A., *Problèmes de Minimax via l'Analyse Convexe et les Inégalités Variationnelles: Théorie et Algorithmes*, Springer-Verlag, Berlin, Germany, 1972.
6. MANGASARIAN, O., *Nonlinear Programming*, McGraw Hill Book Company, New York, New York, 1969.
7. ROCKAFELLAR, R., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
8. EKELAND, I., and TEMAM, R., *Convex Analysis and Variational Problems*, North-Holland Publishing Company, Amsterdam, Holland, 1975.
9. LEVITIN, E. S., and POLYAK, B. T., *Constrained Minimization Methods*, USSR Computational Mathematics and Mathematical Physics, Vol. 6, pp. 1–50, 1966.
10. EVANS J., GOULD, F., and TOLLE, J., *Exact Penalty Functions in Nonlinear Programming*, Mathematical Programming, Vol. 4, pp. 72–97, 1973.