JOURNAL OF COMPUTER AND SYSTEM SCIENCES: 2, 342–380 (1968)

Algorithms for Some Minimax Problems

VLADIMIR F. DEMJANOV*

Electrical Engineering Department, University of Southern California
Los Angeles, California 90007

Received March 4, 1968

ABSTRACT

The problem of minimizing differentiable functions on an entire vector space and on bounded subsets thereof has been studied by many authors. In this paper, we consider the problem of minimizing a nondifferentiable function φ of the form

$$\varphi(z) = \max_{i \in \{1, \cdots, N\}} f_i(z)$$

on the entire space E_n, or on a bounded set Ω in E_n, where the f_i are continuously differentiable functions. In Section 1 an expansion of φ is found, and continuous and discrete algorithms for finding a stationary point (that is, a point satisfying the necessary condition) are given. Some special cases are discussed.

Most of the results obtained here can be applied to minimax problems in function spaces and in particular to some time optimal control problems, optimal control problems in the presence of constraints on the phase coordinates, and some pursuit problems.

In this paper we shall consider the problem of minimizing the function

$$\varphi(z) = \max_{i \in \overline{1, N}} f_i(z)$$

on the entire space E_n (Sections 1–3) or on the bounded set Ω (Sections 4–8). For both cases necessary (and sufficient if possible) conditions for a minimum are proved and successive approximation methods for finding a stationary point (i.e., a point satisfying the necessary condition) are given.

The standard mathematical programming problem is a special case of this problem.

1. NOTATION

Let the $f_i(z)$, $1 \leqslant i \leqslant N$, be real, scalar-valued functions defined and of class C^l, $1 \leqslant l < \infty$, in some neighborhood $S \subset E_n$ of the point $y \in E_n$. Then we can write the

* The author's present address is Mathematics and Mechanics Department, Leningrad State University, Leningrad V 178, U.S.S.R.

following Taylor's series expansion of f_i in the direction g for any arbitrary $g \in E_n(\| g \| \leqslant M < \infty)$ and any real number α such that $(y + \alpha g) \in S$,

$$f_i(y + \alpha g) = f_i(y) + \sum_{k=1}^{l} \frac{\alpha^k}{k!} \frac{\partial^k f_i(y)}{\partial g^k} + o_i(|\alpha|^l) \tag{1.1}$$

where

$$y \equiv (y_1, \ldots, y_n), \qquad g \equiv (g_1, \ldots, g_n),$$

$$\frac{\partial^1 f_i(y)}{\partial g^1} \equiv \frac{\partial f_i(y)}{\partial g} \equiv \left(\frac{\partial f_i(y)}{\partial y}, g\right); \qquad \frac{\partial^2 f_i(y)}{\partial g^2} \equiv \left(g, \frac{\partial^2 f_i(y)}{\partial y^2} g\right), \ldots,$$

$$\frac{\partial^k f_i(y)}{\partial g^k} \equiv \sum_{j_1, \ldots, j_k = 1}^{n} \frac{\partial^k f_i(y)}{\partial y_{j_1} \partial y_{j_2}, \ldots, \partial y_{j_k}} g_{j_1} g_{j_2} \cdots g_{j_k}, \quad (1 \leqslant k \leqslant N), \quad \frac{o_i(|\alpha|^l)}{|\alpha|^l} \xrightarrow[\alpha \to 0]{} 0$$

for $1 \leqslant i \leqslant N$. The notation (A, B) designates the scalar product of vectors A and B, and the notation $\overline{1, N}$ designates the index set $\{1, 2, \ldots, N\}$.

Now let us consider the function

$$\varphi(z) \equiv \max_{i \in \overline{1,N}} f_i(z) \tag{1.2}$$

The function $\varphi(z)$ is not necessarily differentiable, but it will be shown later that it is directionally differentiable. We shall now obtain the expansion of $\varphi(z)$ in the given direction g.

Let us consider sets $R_k(y, g)$, $0 \leqslant k \leqslant N$, defined by the relations:

$$\begin{aligned} R_0 &\equiv R_0(y, g) \equiv \overline{1, N} \\ R_k &\equiv R_k(y, g) \equiv \left\{ i \mid i \in R_{k-1}(y, g), \frac{\partial^{k-1} f_i(y)}{\partial g^{k-1}} = \max_{j \in R_{k-1}(y,g)} \frac{\partial^{k-1} f_j(y)}{\partial g^{k-1}} \right\} \end{aligned} \tag{1.3}$$

It may be seen in (1.3) that $R_0 \supset R_1(y, g) \supset R_2(y, g) \supset \cdots$, and that R_0 does not depend on y and g. Since $R_1(y, g)$ does not depend on g, we may use the notation $R_1(y) \equiv R_1(y, g)$.

Let us consider sets $R_k(y, g, \epsilon_1, \epsilon_2, \ldots, \epsilon_k)$, $1 \leqslant k \leqslant N$, with $\epsilon_1 \geqslant 0, \epsilon_2 \geqslant 0, \ldots, \epsilon_k \geqslant 0$, given by the relation

$$\begin{aligned} R_1(y, \epsilon_1) &= \{ i \mid i \in \overline{1,N}, \varphi(y) - f_i(y) \leqslant \epsilon_1 \} \\ R_k(y, g, \epsilon_1, \epsilon_2, \ldots, \epsilon_k) &= \{ i \mid i \in R_{k-1}(y, g, \epsilon_1, \ldots, \epsilon_{k-1}), \\ &\quad \max_{j \in R_{k-1}(y, g, \epsilon_1, \ldots, \epsilon_{k-1})} \frac{\partial^{k-1} f_j(y)}{\partial g^{k-1}} - \frac{\partial^{k-1} f_i(y)}{\partial g^{k-1}} \leqslant \epsilon_k \} \end{aligned} \tag{1.4}$$

From (1.4), it may be shown that there exist $\epsilon_{1^0} > 0, \epsilon_{2^0} > 0, \ldots, \epsilon_{k^0} > 0$ such that for any $k \in \overline{1, N}$ and any $\epsilon_1, \epsilon_2, \ldots, \epsilon_k$ ($\epsilon_j \leqslant \epsilon_{j0}, j = 1, \ldots, k$) we have

$$R_k(y, g, \epsilon_1, \epsilon_2, \ldots, \epsilon_k) = R_k(y, g). \tag{1.5}$$

Before proceeding further, we shall establish the following functional inequality for arbitrary real, scalar-valued continuous on Ω functions $A(x)$ and $B(x)$, $x \in E_n$:

$$\max_{x \in \Omega}[A(x) + B(x)] \geqslant \max_{x \in \Omega} A(x) + \max_{x \in Q} B(x), \tag{1.6}$$

where Ω is a compact set,

$$Q \equiv \{x \mid x \in \Omega, A(x) = \max_{z \in \Omega} A(z)\}$$

Proof. For any $x' \in \Omega$, we have

$$\max_{x \in \Omega}[A(x) + B(x)] \geqslant A(x') + B(x')$$

So that (1.6) is true for $x' \in Q$. But for such x', $A(x') = \max_{x \in \Omega} A(x)$ and we obtain

$$\max_{x \in \Omega}[A(x) + B(x)] \geqslant \max_{x \in \Omega} A(x) + B(x')$$

Since this inequality is valid for all $x' \in Q$, the correctness of (1.6) is obvious. From (1.1), (1.2), and (1.6) we have, for $\alpha > 0$ and such that $(y + \alpha g) \in \Omega$,

$$\varphi(y + \alpha g) \geqslant \max_{i \in \overline{1, N}} f_i(y) + \max_{i \in R_1(y)} \left[\sum_{k=1}^{l} \frac{\alpha^k}{k!} \frac{\partial^k f_i(y)}{\partial g^k} + o_i(\alpha^l)\right]. \tag{1.7}$$

By repeatedly applying (1.6) to the second term of the right-hand side of (1.7), we obtain

$$\varphi(y + \alpha g) \geqslant \varphi(y) + \sum_{k=1}^{l} \frac{\alpha^k}{k!} \max_{i \in R_k(y, g)} \frac{\partial^k f_i(y)}{\partial g^k} + \mathbf{o}(\alpha^l) \tag{1.8}$$

where

$$\mathbf{o}(\alpha^l) \equiv \min_{i \in \overline{1, N}} o_i(\alpha.^l)$$

On the other hand if for any $\epsilon_1 > 0$ there exists $\alpha_1 > 0$ such that if $\alpha \in [0, \alpha_1]$, then

$$\begin{aligned}
\varphi(y + \alpha g) &= \max_{i \in \overline{1, N}} f_i(y + \alpha g) = \max_{i \in R_1(y, \epsilon_1)} f_i(y + \alpha g) \\
&= \max_{i \in R_1(y, \epsilon_1)} \left[f_i(y) + \sum_{k=1}^{l} \frac{\alpha^k}{k!} \frac{\partial^k f_i(y)}{\partial g^k} + o_i(\alpha^l)\right] \leqslant \max_{i \in R_1(y, \epsilon_1)} f_i(y) \\
&\quad + \max_{i \in R_1(y, \epsilon_1)} \left[\sum_{k=1}^{l} \frac{\alpha^k}{k!} \frac{\partial^k f_i(y)}{\partial g^k} + o_i(\alpha^l)\right] \\
&= \varphi(y) + \max_{i \in R_1(y, \epsilon_1)} \left[\sum_{k=1}^{l} \frac{\alpha^k}{k!} \frac{\partial^k f_i(y)}{\partial g^k} + o_i(\alpha^l)\right].
\end{aligned} \tag{1.9}$$

Let $\epsilon_1 > 0$, $\epsilon_2 > 0, \ldots, \epsilon_l > 0$ be fixed numbers. In a manner analogous to the derivation of (1.9) it follows that there exists $\alpha^* \equiv \min_{i \in \overline{1,l}} \alpha_i > 0$ such that if $\alpha \in [0, \alpha^*]$ then

$$\varphi(y + \alpha g) \leqslant \varphi(y) + \sum_{k=1}^{l} \max_{i \in R_k(y, g, \epsilon_1, \epsilon_2, \ldots, \epsilon_k)} \frac{\alpha^k}{k!} \frac{\partial^k f_i(y)}{\partial g^k} + \bar{o}(\alpha^l) \tag{1.10}$$

where $\bar{o}(\alpha^l) \equiv \max_{i \in \overline{1,N}} o_i(\alpha^l)$. It is obvious that $o(\alpha^l) \leqslant \bar{o}(\alpha^l)$ and that both $\mathbf{o}(\alpha^l)$ and $\bar{o}(\alpha^l)$ depend on g.

If $\epsilon_k (k \in \overline{1, N})$ are such that $\epsilon_k \leqslant \epsilon_{k0}$, then (1.5) is valid. Hence from (1.8) and (1.10), we obtain that if α is small enough then

$$\mathbf{o}(\alpha^l) \leqslant \varphi(y + \alpha g) - \varphi(y) - \sum_{k=1}^{l} \max_{i \in R_k(y, g)} \frac{\alpha^k}{k!} \frac{\partial^k f_i(y)}{\partial g^k} \leqslant \bar{o}(\alpha^l)$$

i.e.,

$$\varphi(y + \alpha g) = \varphi(y) + \sum_{k=1}^{l} \frac{\alpha^k}{k!} \frac{\partial^k \varphi(y)}{\partial g^k} + o(\alpha^l) \tag{1.11}$$

where

$$\frac{\partial^k \varphi(y)}{\partial g^k} = \max_{i \in R_k(y, g)} \frac{\partial^k f_i(y)}{\partial g^k} \qquad (k = 1, \ldots, N), \tag{1.12}$$

and $o(\alpha^l)$ depends on g and y, and

$$\mathbf{o}(\alpha^l) \leqslant o(\alpha^l) \leqslant \bar{o}(\alpha^l).$$

If the f_i have continuous $(l + 1) - st$ derivatives, then $o(\alpha^l)/\alpha^l \xrightarrow[\alpha \to 0]{} 0$ uniformly with respect to g ($\| g \| \leqslant 1$)

Thus, we have obtained that if α is a sufficiently small positive number such that $(y + \alpha g) \in S$, then the expansion (1.11) is valid. The quantity

$$\frac{\partial \varphi(y)}{\partial g} = \max_{i \in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, g \right) \tag{1.13}$$

shall be referred to as the first directional derivative of the function φ at the point y with respect to the direction g. In [1] it has been shown that

$$\lim_{\alpha \to +0} \frac{\varphi(y + \alpha g) - \varphi(y)}{\alpha} = \max_{i \in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, g \right) \equiv \frac{\partial \varphi(y)}{\partial g}.$$

Thus φ is a directionally differentiable function. Now it is also possible to obtain that

$$\lim_{\substack{\alpha \to +0 \\ q \to g}} \frac{\varphi(y + \alpha q) - \varphi(y)}{\alpha} = \frac{\partial \varphi(y)}{\partial g}$$

Note that $\partial\varphi(y)/\partial g$ is a continuous function of g, but that $\partial^k\varphi(y)/\partial g^k$ ($k = 2,..., l$) are not necessarily continuous in g. The function $\partial^k\varphi(y)/\partial g^k$ as a function of y (where g is fixed) is not continuous, in general.

2. Necessary and Sufficient Conditions for a Minimum

Let us consider the following problem. Suppose that $f_i(z)$, $1 \leqslant i \leqslant N$, are real-valued functions, continuous and continuously differentiable in E_n. It is required to find $\min_{z \in E_n} \varphi(z)$. In [1], the following theorem is proved:

Theorem 1. *In order that the point y ($\| y \| < \infty$) be a minimum point of $\varphi(y)$, it is necessary (and if φ is also convex it is sufficient) that*

$$\psi_1(y) = \min_{\|g\| \leqslant 1} \max_{i \in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, g \right) = 0 \tag{2.1}$$

We shall call a point y satisfying (2.1) a *stationary point* of φ on E_n.

Remark 1. Instead of condition (2.1) we can write

$$\psi_2(y) = \min_{\|g\| = 1} \max_{i \in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, g \right) \geqslant 0 \tag{2.2}$$

Let us note that conditions (2.1) and (2.2) are equivalent, i.e., if at the point y we have $\psi_1(y) = 0$, then also $\psi_2(y) \geqslant 0$, and conversely.

Proof. If (2.1) holds, then (2.2) holds necessarily, because otherwise there exists a vector $\bar{g}(\| \bar{g} \| = 1)$ such that

$$\partial\varphi(y)/\partial\bar{g} < 0,$$

whereupon we would obtain

$$\psi_1(y) = \min_{\|g\| \leqslant 1} \frac{\partial\varphi(y)}{\partial g} \leqslant \frac{\partial\varphi(y)}{\partial\bar{g}} < 0.$$

The above inequality contradicts the assumption that (2.1) holds. Conversely, if (2.2) does not hold, then for some $\bar{g}$, $\| \bar{g} \| \leqslant 1$, we would have $\partial\varphi(y)/\partial\bar{g} < 0$, since $\psi_1(y) \leqslant 0$, for all y. It means that $\| \bar{g} \| > 0$ and for $\tilde{g} = \| \bar{g} \|^{-1}\bar{g}$ we have

$$\frac{\partial\varphi(y)}{\partial\tilde{g}} = \frac{\partial\varphi(y)}{\partial\bar{g}} \| \bar{g} \|^{-1} < 0, \qquad \| \tilde{g} \| = 1$$

Moreover, $\psi_2(y) = \min_{\|g\|=1} \partial\varphi(y)/\partial g \leqslant \partial\varphi(y)/\partial\tilde{g} < 0$, which is again a contradiction.

Remark 2. Let us consider functions $\psi_1(y)$ and $\psi_2(y)$ defined at every point of E_n. Let the set $H(y)$ be defined as

$$H(y) \equiv \left\{ x \in E_n \mid x = \frac{\partial f_i(y)}{\partial y}, i \in R_1(y) \right\}.$$

Now let us consider the convex hull $L(y)$ of $H(y)$. ($L(y)$ is a polyhedron in E_n). Let the function $h(g) \equiv \partial\varphi(y)/\partial g = (g, Q(g))$, where $Q(g) \in L(y)$, and such that

$$(g, Q(g)) = \max_{z \in L(y)} (g, z) \equiv \max_{z \in H(y)} (g, z) \qquad \text{(see Fig. 1)}$$

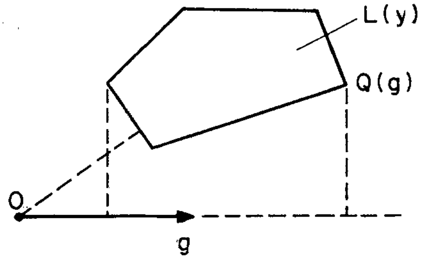

FIG. 1.

Note that $Q(y)$ is not necessarily unique. It is clear that the function $h(g)$ is continuous in g.

Let us show that $h(g)$ is also convex in g on E_n.

Proof. Let g_1 and g_2 be arbitrary vectors, and let $Q_1 \equiv Q(g_1)$ and $Q_2 \equiv Q(g_2)$. Then, for any $x \in L(y)$ and for $\alpha \in [0, 1]$, we have

$$(g_1, x) \leqslant (g_1, Q_1) = h(g_1),$$

and

$$(g_2, x) \leqslant (g_2, Q_2) = h(g_2).$$

$$(\alpha g_1 + (1 - \alpha) g_2, x) = \alpha(g_1, x) + (1 - \alpha)(g_2, x) \leqslant \alpha(g_1, Q_1) + (1 - \alpha)(g_2, Q_2).$$

Since the above expression is valid for an arbitrary $x \in L(y)$, it follows that

$$h(\alpha g_1 + (1 - \alpha) g_2) = \max_{x \in L(y)} (\alpha g_1 + (1 - \alpha) g_2, x) \leqslant \alpha h(g_1) + (1 - \alpha) h(g_2)$$

and hence $h(g)$ is convex.

Now, we claim that if $\psi_1(y) < 0$, then $h(g)$ has only one minimum point on the set $\|g\| \leqslant 1$.

Proof by Contradiction. Let $\| g_1 \| = \| g_2 \| \leqslant 1$, $g_1 \neq g_2$, and $h(g_1) = h(g_2) = \min_{\|g\| \leqslant 1} h(g)$. Then

$$h\left(\frac{g_1 + g_2}{2}\right) \leqslant \tfrac{1}{2}h(g_1) + \tfrac{1}{2}h(g_2) = \min_{\|g\| \leqslant 1} h(g) \equiv \psi_1(y).$$

Since

$$\beta^2 = \| \tfrac{1}{2}(g_1 + g_2)\|^2 \leqslant \tfrac{1}{4}\| g_1 \|^2 + \tfrac{1}{2}(g_1, g_2) + \tfrac{1}{4}\| g_2 \|^2 < 1((g, g_2) < 1 \quad \text{as} \quad g_1 \neq g_2)$$

and

$$h(\gamma g) = \gamma h(g) \qquad \text{for} \quad \gamma > 0,$$

it follows that

$$h(\bar{g}) \equiv h((1/2\beta)(g_1 + g_2)) = (1/\beta)\, h(\tfrac{1}{2}(g_1 + g_2)) \leqslant (1/\beta)\, \psi_1(y) < \psi_1(y)$$

since $(1/\beta) > 1$. This is contradiction, as $\| \bar{g} \| = 1$. (If $\beta = 0$, then instead of $\frac{1}{2}(g_1 + g_2)$ we could choose any point $\alpha g_1 + (1 - \alpha)\, g_2 (\alpha \in (0, 1), \alpha \neq \frac{1}{2})$). It is clear that if $\psi_1(y) < 0$, then $\| g(y) \| = 1$, where the vector $g(y)$ is such that $\psi_1(y) = \partial\varphi(y)/\partial g(y)$.

Thus we have proved that if the point y is not a stationary point, then there exists one and only one vector $g(y)$, $\| g(y) \| \leqslant 1$, such that $\psi_1(y) = \partial\varphi(y)/\partial g(y)$ and in addition $\| g(y) \| = 1$.

To find $g(y)$, it is possible to use the standard quadratic programming technique ([2]–[4]).

Geometrically, if y is not a stationary point then $\psi_1(y) = \psi_2(y) = -\rho < 0$, where ρ is the distance between the origin and $L(y)$. (See Fig. 1).

In fact, if for some g we define

$$\chi(g) \equiv \max_{i \in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, g\right) = -\rho' < 0$$

then $\min_{x \in L(y)} \| x \| \geqslant \rho'$ (see Fig. 2), i.e., there is no point of $L(y)$ inside the sphere $\| x \| \leqslant \rho'$. Now let $x_1 \in L(y)$ be such that

$$\| x_1 \| = \min_{x \in L(y)} \| x \|$$

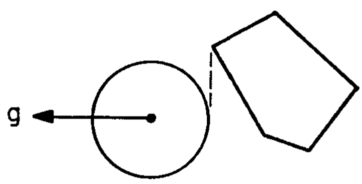

FIG. 2.

Then for $g_1 \equiv \| x_1 \|^{-1} x_1$ we have $\chi(g_1) = -\| x_1 \|$. Let us prove that $-\rho \equiv \psi_1(y) = -\| x_1 \|$. First of all

$$\psi_1(y) = \min_{\|g\| \leqslant 1} \chi(g) \leqslant \chi(g_1) = -\| x_1 \|.$$

Now suppose also that $\psi_1(y) \neq -\| x_1 \|$, i.e., for some $\bar{g} : \chi(\bar{g}) < -\| x_1 \|$. Then there is no point of $L(y)$ inside the sphere $\| x \| \leqslant -\chi(\bar{g})$. This is a contradiction, since $\| x_1 \| < -\chi(\bar{g})$.

Now let y be a stationary point, then $\psi_1(y) = 0$ but $\psi_2(y) = r$, where r is the radius of the largest sphere (with the origin as center) which can be inscribed in $L(y)$ (see Fig. 3). Geometrically, the necessary condition for a minimum is that at the minimum point y, the origin must belong to the convex hull $L(y)$.

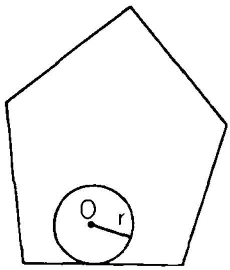

FIG. 3.

This is a generalization of the well-known necessary condition for a point y to be a minimum of a differentiable function. If φ is differentiable at y, then $\partial\varphi(y)/\partial y = 0$ necessarily at a minimum point of φ. In this case (i.e., where φ is differentiable) the sets $H(y)$ and $L(y)$ consist of one point $\partial\varphi(y)/\partial y$. For any point y the direction $G_1(y) = \partial\varphi(y)/\partial y$ is the direction of steepest ascent (this direction is called the gradient) and $-G_1(y)$ is the direction of steepest descent, and

$$\partial\varphi(y)/\partial g = (G_1(y), g). \tag{2.3}$$

In the case where φ is given by (1.2), then the steepest descent direction at y is the direction $-g(y)$.

Note that, in general, φ is not differentiable, so that the direction $g(y)$ is not necessarily the direction of the steepest ascent. In this case, we must use (1.13) instead of (2.3).

We now assert that if $\psi_1(y) = r > 0$, then

(1) the point y is a local minimum point, and

(2) the point y is a discontinuity point of the set function $R_1(y)$.

Proof (1). For all $i \in \overline{1, N}$, we have

$$f_i(x) \equiv f_i(y + (x - y)) = f_i(y) + \left(x - y, \frac{\partial f_i(y)}{\partial y}\right) + o_i(\| x - y \|).$$

Then

$$\varphi(x) = \max_{i \in \overline{1,N}} f_i(x) \geqslant \max_{i \in \overline{1,N}} f_i(y) + \max_{i \in R_1(y)} \left(x - y, \frac{\partial f_i(y)}{\partial y}\right) + \max_{i \in R_2(y, x-y)} o_i(\| x - y \|)$$

$$= \varphi(y) + \| x - y \| \max_{i \in R_1(y)} \left(g, \frac{\partial f_i(y)}{\partial y}\right) + \max_{i \in R_2(g)} o_i(\| x - y \|)$$

$$\geqslant \varphi(y) + \| x - y \| r + o(\| x - y \|), \qquad \text{where} \quad g = \| x - y \|^{-1}(x - y).$$

Since $r > 0$, it follows that there exists $\epsilon > 0$ such that

$$\varphi(x) > \varphi(y) \tag{2.4}$$

whenever $\| x - y \| \leqslant \epsilon$. But this implies that y is a local minimum point of φ.

Proof (2). Now let us prove that the point y is a discontinuity point of the set function R_1 along any direction g. The set function R_1 is said to be continuous at the point y if

$$\rho(R_1(x), R_1(y)) = \max_{z_2 \in R_1(x)} \min_{z_1 \in R_1(y)} (z_1 - z_2)^2 + \max_{z_2 \in R_1(y)} \min_{z_1 \in R_1(x)} (z_1 - z_2)^2 \xrightarrow[\|x-y\| \to 0]{} 0$$

The set function R_1 is said to be continuous at the point y along the direction $g(\| g \| < \infty)$ if

$$\rho(R_1(y + \alpha g), R_1(y)) \xrightarrow[\alpha \to 0^+]{} 0.$$

It is clear that the discontinuity of R_1 at y along any (and one is enough) direction g implies that R_1 is discontinuous at y.

Let us show that if $\psi_2(y) > 0$, then y is a discontinuity point of the set function R_1 along any direction g. We shall prove an even stronger statement: There exists no sequence $\{y_k\}$ such that $\| y_k - y \| = \epsilon_k \xrightarrow[k \to \infty]{} 0$ and $R_1(y_k) \xrightarrow[k \to \infty]{} R_1(y)$. This means that there exists a neighborhood S_1 of y such that for any $y' \in S_1$ $(y' \neq y)$

$$R_1(y') \neq R_1(y) \qquad (\text{and} \quad R_1(y') \subset R_1(y))$$

Proof by Contradiction. Suppose that the sequence $\{y_k\}$ is such that $\| y_k - y \| = \epsilon_k \xrightarrow[k \to \infty]{} 0$ and $R_1(y_k) \xrightarrow[k \to \infty]{} R_1(y)$. Since the range of $R_1(x)$ consists of a finite number of values (i.e., $R_1(x)$ is a "step-function") we have, that beginning with some k, $R_1(y_k) = R_1(y)$. Since the f_i, $1 \leqslant i \leqslant N$, are continuously differentiable then if k is large enough, we have $\psi_2(y_k) \geqslant \frac{1}{2}r$.

Now we have

$$\begin{aligned}\varphi(y) &\equiv \varphi(y_k + (y - y_k)) = \max_{i\in\overline{1,N}} f_i(y_k + (y - y_k)) \\ &\geqslant \max_{i\in\overline{1,N}} f_i(y_k) + \max_{i\in R_1(y_k)} \left(y - y_k, \frac{\partial f_i(y_k)}{\partial y}\right) + \max_{i\in R_2(y_k, y-y_k)} o_i(\|y - y_k\|) \\ &= \varphi(y_k) + \|y - y_k\| \max_{i\in R_1(y_k)} \left(\bar{g}, \frac{\partial f_i(y_k)}{\partial y}\right) + o(\|y - y_k\|) \\ &\geqslant \varphi(y_k) + \epsilon_k \frac{r}{2} + o(\epsilon_k),\end{aligned}$$

where $\bar{g} \equiv \|y - y_k\|^{-1}(y - y_k)$. For k large enough we obtain, recalling (2.4) and assuming that $\epsilon_k < \epsilon$,

$$\varphi(y) \geqslant \varphi(y_k) + \tfrac{1}{4} r\epsilon_k > \varphi(y_k) > \varphi(y)$$

which is absurd. This contradiction proves our assertion.

Remark 3. If the functions f_i are twice continuously differentiable on some neighborhood S of y, and if y is a stationary point of φ, then a sufficient condition for the point y to be a local minimum point is

$$\min_{\|g\|=1} \max_{i\in R_2(y,g)} \left(\frac{\partial^2 f(y)}{\partial y^2} g, g\right) > 0$$

If all the functions f_i, $i \in R_2(y)$ are strictly positive definite at y then y is a local minimum point, assuming of course that (2.1) holds. If the functions f_i, $i \in \overline{1, N}$, are convex, then φ is a convex function since

$$\begin{aligned}\varphi(\alpha x_1 + (1 - \alpha)x_2) &= \max_{i\in\overline{1,N}} f_i(\alpha x_1 + (1 - \alpha)x_2) \leqslant \max_{i\in\overline{1,N}} [\alpha f_i(x_1) + (1 - \alpha) f_i(x_2)] \\ &\leqslant \alpha \max_{i\in\overline{1,N}} f_i(x_1) + (1 - \alpha) \max_{i\in\overline{1,N}} f_i(x_2) = \alpha\varphi(x_1) + (1 - \alpha)\, \varphi(x_2).\end{aligned}$$

In this case, any stationary point y is a (global) minimum point of φ, so that if the f_i, $i \in R_1(y)$, are strictly convex, then y is the unique minimum point of φ.

Let us denote the minimal value of φ on E_n by φ^*. If the f_i, $i \in \overline{1, N}$, are convex, and if for some set $\underset{\sim}{Q} \in \overline{1, N}$ we have at the point y

$$\min_{\|g\|\leqslant 1} \max_{i\in Q} \left(\frac{\partial f_i(y)}{\partial y}, g\right) \equiv \min_{\|g\|\leqslant 1} \chi(g) = 0 \tag{2.5}$$

then

$$\min_{i\in Q} f_i(y) \leqslant \varphi^* \leqslant \varphi(y). \tag{2.6}$$

Since the right hand inequality is obvious, let us prove the left hand inequality. Really assuming the contrary, if for some x we have $\varphi(x) < \min_{i\in Q} f_i(y)$ then $f_i(x) < f_i(y)$, $i \in Q$. Since the f_i are convex, we have

$$\left(\frac{\partial f_i(y)}{\partial y}, x - y\right) < 0, \qquad i \in Q.$$

It follows that

$$\max_{i\in Q} \left(\frac{\partial f_i(y)}{\partial y}, x - y\right) < 0$$

and

$$\chi(\bar{g}) = \max_{i\in Q} \left(\frac{\partial f_i(y)}{\partial y}, \bar{g}\right) < 0$$

where

$$\bar{g} = \| x - y \|^{-1}(x - y), \| \bar{g} \| = 1,$$

hence, a fortiori,

$$\min_{\|g\|\leqslant 1} \chi(g) \leqslant \chi(\bar{g}) < 0$$

which contradicts (2.5). Thus the inequality (2.6) is valid. This inequality enables us to stop searching for the minimum point after achieving the desired precision. If it turns out that $Q \subset R_1(y\ \epsilon)$, i.e., if we can find the corresponding $\epsilon > 0$ for any set $Q \subset \overline{1, N}$, then

$$\varphi(y) - \varphi^* \leqslant \epsilon.$$

3. Some Successive Approximation Methods

The problem being considered is a generalization of the standard mathematical programming problem, for which there exist many methods of successive approximations. (See [5]–[13].) Here, we shall discuss some methods of successive approximations which can be obtained from the minimax approach and which are useful not only for solving this problem but also for finding minimum points of more complicated functions (arising, for example, in optimal control problems).

The main difficulty in developing methods of successive approximations arises because of the discontinuity of the set function $R_1(y)$. To show this difficulty, let us consider the "obvious" generalization of the gradient method. Since $g(y)$ is the direction of steepest descent, we can use the following procedure: Let y_1 be an arbitrary point of E_n, and let $g_1 = g(y_1)$. If $\psi_1(y_1) = 0$, then y_1 is a stationary point, and the process is finished. If $\psi_1(y_1) < 0$, then let us consider the ray

$$y_{1\alpha} = y_1 + \alpha g_1 \qquad (\alpha \geqslant 0)$$

and find $\alpha_1 \in (0, \infty)$ such that

$$\varphi(y_{1\alpha_1}) = \min_{\alpha \in [0, \infty)} \varphi(y_{1\alpha}).$$

At this point we set $y_2 = y_{1\alpha_1}$ and continue in the same manner. This "obvious" method fails, in general, to lead us to a stationary point because of the discontinuity of $R_1(y)$. One of the methods for overcoming this difficulty has been described in [*1*]. At the point y we can construct sets $Q_0(y), Q_1(y), \ldots, Q_m(y), Q_k(y) \subset \overline{1, N}$ and with $f_i(y) - \varphi(y) \equiv a_k$ for all $i \in Q_k$ and such that $a_0(y) > a_1(y) > \cdots > a_m(y)$. It is clear that $m = m(y) \leqslant N - 1$. Let $R_1{}^k(y) = \bigcup_{i=0}^{k} Q_i(y)$, $0 \leqslant k \leqslant m$. Note that $R_1{}^m(y) = R_1(y)$. Now let us consider some methods.

Method 1. Let $y_1 \in E_n$ be an arbitrary point. Suppose that y_l has been found. If $\delta_l{}^0 = \psi_1(y_l) = 0$, then y_l is a stationary point and the process is finished. If $\delta_l{}^0 < 0$, then let $R_{1l}^k = R_1{}^k(y_l)$, and the $g_l{}^k$, $0 \leqslant k \leqslant m_l$, where $m_l = m(y_l)$ satisfy the following relation

$$\delta_l{}^k = \max_{i \in R_{1l}^k} \left(\frac{\partial f_i(y_l)}{\partial y}, g_l{}^k \right) = \min_{\|g\| \leqslant 1} \max_{i \in R_{1l}^k} \left(\frac{\partial f_i(y_l)}{\partial y}, g_l{}^k \right). \tag{3.1}$$

Note that $0 < \delta_l{}^0 \leqslant \delta_l{}^1 \leqslant \delta_l{}^2 \leqslant \cdots \leqslant \delta_l^{m_l}$. Denote $a_{kl} \equiv a_k(y_l)$, $0 \leqslant k \leqslant m_l$, and $a_{0l} = 0 < a_{il} < \cdots < a_{ml}$. Let us consider the rays

$$y_{l\alpha}^k = y_l + \alpha g_l{}^k (\alpha \in [0, \infty))(k \in \overline{0, m_l})$$

and find $\underline{a_{kl}} \in [0, \infty)$ such that $\varphi(y_{l\alpha_k}^k) = \min_{\alpha \in [0,\infty)} \varphi(y_{l\alpha}^k)(k \in 0, m_l)$. Let k_l be such that $k_l \in \overline{0, m_l}$ and

$$\varphi(y_{l\alpha_{kl}}^k) = \min_{k \in \overline{0, m_l}} \varphi(y_{l\alpha_k}^k) \tag{3.2}$$

(if there exist several points of this kind, choose any of them). At this point we set

$$y_{l+1} = y_{l\alpha_{k_l}}^{k_l}$$

and continue in the same manner. Note that

$$\varphi(y_1) > \varphi(y_2) > \cdots > \varphi(y_l) > \cdots$$

Thus we construct the sequence $\{y_l\}$. If this is a finite sequence, then the extreme right point is a stationary point of φ. If it is an infinite sequence (i.e., if it consists of an infinite number of points y_l) then the following theorem is valid:

THEOREM 2. *If the set*

$$D \equiv \{x \in E_n \mid \varphi(x) \leqslant \varphi(y_1)\}$$

is bounded and if all the functions $f_i(x)(i \in \overline{1, N})$ are continuously differentiable on D, then any limit point y^* of the sequence $\{y_l\}$ is a stationary point of φ.

Proof. Let

$$\lim_{l\to\infty} \varphi(y_l) \equiv \varphi^*, \qquad -\infty < \varphi^* < \infty, \qquad \varphi(y_l) > \varphi^*. \tag{3.3}$$

This limit exists because the sequence $\{\varphi(y_l)\}$ is monotone and bounded from below (all the y_i, $1 \leqslant i < \infty$, belong to D, and φ is continuous on D). Let

$$\lim_{s\to\infty} y_{l_s} \equiv y^* \in D\,(l_s \xrightarrow[s\to\infty]{} \infty).$$

Since φ is continuous, we have $\varphi(y^*) = \varphi^*$. We shall show that y^* is a stationary point. Let us prove this by contradiction: Suppose that $\psi_1(y^*) = -\delta^* < 0 (\delta^* > 0)$. Then if l_s is large enough, $\psi_1(y_l) \leqslant -\frac{1}{2}\delta^*$, since $R_1(y_l) \subset R_1(y^*)$ for $l_s \geqslant l^*$ (where l^* is large enough) and

$$\max_{i\in R_1(y_l)} \left(\frac{\partial f_i(y_l)}{\partial y}, g\right) \leqslant \max_{i\in R_1(y^*)} \left(\frac{\partial f_i(y_l)}{\partial y}, g\right)$$

There exist $\epsilon_1 > 0$ and $\epsilon_2 > 0$, $\epsilon_1 < \epsilon_2$, such that for any $l_s \geqslant \bar{l}$ (where $\bar{l} \geqslant l^*$ is large enough) we can find k_s such that

$$a_{k_s l_s} \leqslant \epsilon_1, \qquad a_{k_s+1, l_s} \geqslant \epsilon_2 \tag{3.4}$$

and

$$R_{1l_s}^{k_s} = R_1(y^*). \tag{3.5}$$

As long as

$$R_{1l_s}^{k} \subset R_{1l_s}^{k_s} \qquad \text{for} \quad k \in \overline{1, k_s} \qquad \text{and for} \quad l_s \geqslant \bar{l}$$

we have

$$\delta_{l_s}^k \leqslant -\tfrac{1}{2}\,\delta^*. \tag{3.6}$$

Now for $i \in R_{1l_s}^{k}$ we have

$$f_i(y_{l_s\alpha}^{k_s}) = f_i(y_{l_s} + \alpha g_{l_s}^{k_s}) = f_i(y_{l_s}) + \alpha \left(\frac{\partial f_i(y_{l_s})}{\partial y}, g_{l_s}^{k_s}\right) + o_{il_sk_s}(\alpha).$$

From (3.4) and (3.5) we have, for $\alpha \in (0, \alpha^*)$, where α^* is the same for $l_s \geqslant \bar{l}$,

$$\varphi(y_{l_s\alpha}^{k_s}) = \max_{i\in R_1(y^*)} f_i(y_{l_s\alpha}^{k_s}) \leqslant \varphi(y_{l_s}) + \alpha \max_{i\in R_1(y_{l_s})} \left(\frac{\partial f_i(y_{l_s})}{\partial y}, g_{l_s}^{k_s}\right) + \max_{i\in R_2(y_{l_s})} o_{il_sk_s}(\alpha).$$

Recalling (3.6), we obtain

$$\varphi(y_{l_s\alpha}^{k_s}) \leqslant \varphi(y_{l_s}) - \tfrac{1}{2}\alpha\delta^* + o(\alpha)$$

where $o(\alpha)/\alpha \xrightarrow[\alpha\to+0]{} 0$ uniformly on l_s and k_s. For $\alpha \in (0, \bar{\alpha})$ ($\bar{\alpha} \leqslant \alpha^*$) we have

$$\varphi(y_{l_s\alpha}^{k_s}) \leqslant \varphi(y_{l_s}) - \tfrac{1}{4}\alpha\delta^*.$$

Choose some $\tilde{\alpha} \in (0, \bar{\alpha})$. As long as $\varphi(y_{l_s}) = \varphi^* + \epsilon_{l_s}$, $\epsilon_{l_s} > 0$, $\epsilon_{l_s} \xrightarrow[l_s\to\infty]{} 0$, for l_s large enough (and, of course, $l_s \geqslant l^*$) we have

$$\varphi(y_{l_s\tilde{\alpha}}^{k_s}) \leqslant \varphi^* - \tfrac{1}{8}\tilde{\alpha}\delta^* < \varphi^*,$$

and then, certainly,

$$\varphi(y_{l_s+1}) = \min_{x\in\overline{1,m_{l_s}}} \min_{\alpha\in[0,\infty)} \varphi(y_{l_s\alpha}^{k}) \leqslant \varphi(y_{l_s\tilde{\alpha}}^{k_s}) < \varphi^*,$$

which contradicts (3.3). Thus, Theorem 2 is proved.

Remark 4. It is clear from the preceding proof that the method can be modified in the following way.

Method 1a. Let us choose some $\epsilon > 0$ and some k_l in (3.2) such that $k_l \in \overline{0, m_l'}$ and

$$\varphi(y_{l\alpha_{k_l}}^{k_l}) = \min_{k\in\overline{0,m_l'}} \varphi(y_{l\alpha_k}^{k}),$$

where m_l' is such that

$$a_{m_l'l} \leqslant \epsilon < a_{m_l'+1,l}\,.$$

Remark 5. Note that by increasing l we cannot change ϵ_2 (see (3.4) to make ϵ_1 as small as desired. However, we can apply the following method:

Method 1b. Let $\{\beta_i\}$ be a decreasing sequence such that $\beta_i > 0$, $\beta_i > \beta_{i-1}$, and

$$\beta_i \xrightarrow[i\to\infty]{} 0.$$

As usual, choose an arbitrary $y_1 \in E_n$. Let y_l have already been found. If $\delta_l^0 = \psi_1(y_l) = 0$, then y_l is a stationary point and the process is finished. If $\delta_l^0 < 0$, then let us find a subsequence $\{\beta_{l_i}\}$ of $\{\beta_i\}$, consisting of a finite set of numbers, such that

$$a_{k_{l_i+1}l} < \beta_{l_i} \leqslant a_{k_{l_i}l} < a_{k_{l_{i+1}}l}\,.$$

Now let k_l be such that $k_l \in \{k_l\}$ and

$$\varphi(y^{k_l}_{l\alpha_{k_l}}) = \min_{k \in \{k_{l_i}\}} \varphi(y^k_{l\alpha_k}).$$

This method enables us to reduce the labor involved in finding a minimum point. Applying Methods 1, 1a, and 1b, each step requires minimization of φ on several rays. If this is too laborious, we can use other methods.

Method 2. For the first approximation, we choose an arbitrary vector $y_1 \in E_n$. Let y_l have already been found. Suppose that $\delta_l^0 < 0$, because otherwise y_l would be a stationary point. In accordance with (3.1) we have

$$0 > \delta_l^0 \leqslant \delta_l^1 \leqslant \delta_l^2 \leqslant \cdots \leqslant \delta_l^{m_l}$$

$$0 < a_{1l} < a_{2l} < \cdots < a_{m_l l}$$

Find the largest $k_l \in \overline{0, m_l}$ such that

$$\delta^{k_l}_{l_k} \leqslant -\mu a_{k_l l}, \tag{3.7}$$

where μ is some fixed number. Note that if $\delta_l^k \leqslant -\mu a_{kl}$ (and if $k \geqslant i$) then a fortiori $\delta_l^{k-1} < -\mu a_{kl}$. In fact, if $\delta_l^{k-1} = \delta_l^k - \rho_{kl}$, $\rho_{kl} \geqslant 0$, $a_{k-1,l} = a_{kl} - \chi_{kl}$, $\chi_{kl} > 0$, then

$$\delta_l^{k-1} = \delta_l^k - \rho_{kl} \leqslant -\mu a_{kl} - \rho_{kl} = -\mu a_{k-1,l} - \mu\chi_{kl} - \rho_{kl} < -\mu a_{k-1,l}$$

which is the desired result. And now, certainly, recalling (3.7) we have

$$\delta_l^k < -\mu a_{kl} \qquad \text{for} \quad k \in \overline{0, k_l - 1}$$

Let us form the ray

$$y^{k_l}_{l\alpha} = y_l + \alpha g_l^{k_l} (\alpha \in [0, \infty))$$

and find $\alpha_l \in [0, \infty)$ such that

$$\varphi(y^{k_l}_{l\alpha_l}) = \min_{\alpha \in [0,\infty)} \varphi(y^{k_l}_{l\alpha}). \tag{3.7}$$

Now we may set

$$y_{l+1} = y^{k_l}_{l\alpha_{k_l}}$$

By construction

$$\varphi(y_{l+1}) < \varphi(y_l), \tag{3.8}$$

and we may continue in the same manner. The sequence $\{y_l\}$ which we have thus constructed tends to a stationary point of φ. This statement may be given as the following theorem:

THEOREM 3. *If the hypotheses of Theorem 2 are satisfied, then the sequence* $\{y_l\}$ *formed in accordance with* (3.7) *converges to a stationary point of* φ. *This theorem can be proved in just the same way as one of the theorems in* [*1*].

Remark 6. Theoretically, μ can be chosen arbitrarily, but there exist more or less reasonable values of μ in particular cases.

Remark 7. Method 2 can be modified in the following way:

Method 2a. Let $\mu > 0$ and $\epsilon > 0$ be fixed. For the first approximation, we choose an arbitrary vector $y_1 \in E_n$. Let y_l have been found. Suppose that $\delta_l^0 < 0$ (because otherwise if $\delta_l^0 = 0$, y_l is a stationary point of φ and the process is finished). Let $\epsilon_{l1} = \epsilon$. If

$$\delta_{l1} = \min_{\|g\| \leqslant 1} \max_{i \in R_{1\epsilon_{l1}}(y_l)} \left(\frac{\partial f_i(y_l)}{\partial y}, g\right) = \max_{i \in R_{1\epsilon_{l1}}(y_l)} \left(\frac{\partial f_i(y_l)}{\partial y}, g_{l1}\right) \leqslant -\mu\epsilon_{l1}$$

where

$$R_{1\epsilon_{l1}}(y_l) = \{i \mid i \in \overline{1, N}, \varphi(y_l) - f_i(y_l) \leqslant \epsilon_{l1}\},$$

then we set $\epsilon_l = \epsilon_{l1}$, $g_l = g_{l1}$ and form the ray $y_{l\alpha} = y_l + \alpha g_l (\alpha \in [0, \infty))$ and find $\alpha_l \in (0, \infty)$ such that

$$\varphi(y_{l\alpha_l}) = \min_{\alpha \in [0, \infty)} \varphi(y_{l\alpha}).$$

Now we may set

$$y_{l+1} = y_{l\alpha_l}.$$

If $\delta_{l1} > -\mu\epsilon_{l1}$, then we repeat the same process beginning with $\epsilon_{l2} = \frac{1}{2}\epsilon_{l1}$ until we obtain $\delta_{lk} \leqslant -\mu\epsilon_{lk}$. Note that in this case, the set $R_{1_\epsilon}(y_l)$ coincides with one of the sets $R_1^k(y_l)$, $0 \leqslant k \leqslant m_l$.

Remark 8. At each step of all these methods it is required to find minima or mini-maxima of comparatively simple functions. It may be shown that it is possible to obtain approximate solution of these auxiliary extremal problems. For example it is possible, instead of finding $\min_{\alpha \in [0, \infty)} \varphi(y_{l\alpha})$, to try to find $\min_{\alpha \in [0, A]} \varphi(y_{l\alpha})$, where A, $0 < A < \infty$, is fixed and does not depend on l. (For other details, see [*14*], pp. 284-285).

Remark 9. Let $y(t) \in E_n$ be a vector-valued function, continuous on $[0, T]$, $0 < T < \infty$, such that $y(t) \in S$ for $t \in [0, T]$, where $S \subset E_n$ is a bounded closed set, and whose derivative satisfies the following equation

$$\dot{y}_+(t) = g(t), \qquad y(o) = y_0$$

where $\dot{y}_+(t) = \lim_{\alpha\to 0+} [y(t+\alpha) - y(t)]/\alpha$, and where the vector-valued function $g(t) = g(y(t))$ is piecewise continuous and bounded on $[0, T]$. Suppose that the set function $R_1(t) \equiv R_1(y(t))$ is such that meas $\omega = 0$, where $\omega \equiv \{\tau \in [0, T] \mid \tau$ is a dicontinuity point of $R_1\}$. Note that if $R_1(t)$ is continuous on $[t', t'']$ it means that $R_1(t)$ is constant on $[t', t'']$. Then we shall prove that

$$\varphi(y(t)) = \varphi(y_0) + \int_0^t \frac{\partial\varphi(y(\tau))}{\partial g(\tau)}\,d\tau \tag{3.9}$$

where as usual

$$\frac{\partial\varphi(y(\tau))}{\partial g(\tau)} = \max_{i\in R_1(\tau)} \left(\frac{\partial f_i(y(\tau))}{\partial y}, g(\tau)\right).$$

Proof. Let $\Delta = t/m$, $t_0 = 0$, $t_k = k\Delta$, $t_m = t$, $k = 0, 1,..., m$. Since

$$f_i(y(t)) = f_i(y_0) + \int_0^t \left(\frac{\partial f_i(y(\tau))}{\partial y}, g(\tau)\right) d\tau,$$

and recalling (1.6), we have for an arbitrary $\Delta > 0$

$$\max_{i\in\overline{1,N}} f_i(t) \geqslant \max_{i\in\overline{1,N}} \left[f_i(y_0) + \int_0^{t-\Delta} \chi_i(\tau)\,d\tau\right] + \max_{i\in R_1(t-\Delta)} \int_{t-\Delta}^t \chi_i(\tau)\,d\tau, \tag{3.10}$$

where

$$f_i(t) \equiv f_i(y(t)), \qquad \chi_i(t) = \left(\frac{\partial f_i(y(t))}{\partial y}, g(t)\right), \qquad R_1(t) = R_1(y(t)).$$

By repeatedly applying (1.6) we obtain from (3.10)

$$\varphi(y(t)) \geqslant \varphi(y_0) + \sum_{k=0}^{m-1} \max_{i\in R_1(t_k)} \int_{t_k}^{t_{k+1}} \chi_i(\tau)\,d\tau \tag{3.11}$$

On the other hand, for the same Δ, we have

$$\max_{i\in\overline{1,N}} \left[f_i(y_0) + \int_0^t \chi(\tau)\,d\tau\right] = \max_{i\in\overline{1,N}} \left[f_i(t-\Delta) + \int_{t-\Delta}^t \chi_i(\tau)\,d\tau\right]$$

$$= \max_{i\in R_1(t)} \left[f_i(t-\Delta) + \int_{t-\Delta}^t \chi_i(\tau)\,d\tau\right] \leqslant \max_{i\in R_1(t)} f_i(t-\Delta) + \max_{i\in R_1(t)} \int_{t-\Delta}^t \chi_i(\tau)\,d\tau$$

$$\leqslant \max_{i\in\overline{1,N}} f_i(t-\Delta) + \max_{i\in R_1(t)} \int_{t-\Delta}^t \chi_i(\tau)\,d\tau.$$

Continuing in the same manner we obtain

$$\varphi(t) \leqslant \varphi(y_0) + \sum_{k=0}^{m-1} \max_{i \in R_1(t_{k+1})} \int_{t_k}^{t_{k+1}} \chi_i(\tau)\, d\tau \tag{3.12}$$

Let

$$K_m = \{k \mid k \in \overline{0, m-1}, R_1(\tau) = R_1(t_k) \text{ for } \tau \in [t_k, t_{k+1}]\} .$$

Then $R_1(t_k) = R_1(t_{k+1})$ for $k \in K_m$ and we can rewrite (3.11) and (3.12) as follows

$$\varphi(t) \equiv \varphi(y(t)) \geqslant \varphi(y_0) + \sum_{k \in K_m} \max_{i \in R_1(t_k)} \int_{t_k}^{t_{k+1}} \chi_i(\tau)\, d\tau$$
$$+ \sum_{\substack{k \in \overline{0, m-1} \\ k \notin K_m}} \max_{i \in R_1(t_k)} \int_{t_k}^{t_{k+1}} \chi_i(\tau)\, d\tau \tag{3.13}$$

$$\varphi(t) \leqslant \varphi(y_0) + \sum_{k \in K_m} \max_{i \in R_1(t_k)} \int_{t_k}^{t_{k+1}} \chi_i(\tau)\, d\tau$$
$$+ \sum_{\substack{k \in \overline{0, m-1} \\ k \notin K_m}} \max_{i \in R_1(t_{k+1})} \int_{t_k}^{t_{k+1}} \chi_i(\tau)\, d\tau \tag{3.14}$$

Note that

$$\int_{t_k}^{t_{k+1}} \chi_i(\tau)\, d\tau = \chi_i(\theta_{ik}) \cdot \Delta \qquad (i \in \overline{1, N}, k \in \overline{0, m-1}) \tag{3.15}$$

where

$$\theta_{ik} \in [t_k, t_{k+1}]$$

Since the $f_i(t)$ are continuously differentiable on S then for any $\epsilon > 0$ we can find M such that for $m > M$

$$| \max_{i \in A} \chi_i(\theta_i) - \max_{i \in A} \chi_i(t_k)| \leqslant \epsilon \tag{3.16}$$

whenever $\theta_i \in [t_k, t_{k+1}]$, $k \in K_m$; $A \subset \overline{1, N}$.

Now from (3.15) and (3.16) we have

$$\sum_{k \in K_m} \max_{i \in R_1(t_k)} \int_{t_k}^{t_{k+1}} \chi_i(\tau)\, d\tau = \sum_{k \in K_m} \max_{i \in R_1(t_k)} \chi_i(\theta_{ik}) \cdot \Delta$$
$$= \sum_{k=0}^{m-1} \max_{i \in R_1(t_k)} \chi_i(t_k)\Delta + \sum_{k \in K_m} \left[\max_{i \in R_1(t_k)} \chi_i(\theta_{ik}) - \max_{i \in R_1(t_k)} \chi_i(t_k)\right] \Delta$$
$$- \sum_{\substack{k \in \overline{0, m-1} \\ k \notin K_m}} \max_{i \in R_1(t_k)} \chi_i(t_k)\Delta. \tag{3.17}$$

Let

$$\max_{\substack{i\in\overline{1,N}\\ x\in S}} |\chi_i(x)| = H \qquad (H < \infty)$$

then from (3.13), (3.14) and (3.17) we have

$$\left| \varphi(t) - \varphi(o) - \sum_{k=0}^{m-1} \max_{i\in R_1(t_k)} \chi_i(t_k)\Delta \right| \leqslant \epsilon t + 3Hp\Delta$$

where $p = p(m)$ is the number of k such that $k \in \overline{0,\ m}$, $k \notin K_m$. Since $p\Delta \xrightarrow[m\to\infty]{}$ means $\omega = 0$ and since $\epsilon \xrightarrow[m\to\infty]{} 0$, $\Delta \equiv \Delta_m \xrightarrow[m\to\infty]{} 0$, we obtain

$$\begin{aligned}\varphi(t) &= \varphi(o) + \lim_{m\to\infty} \sum_{k=0}^{m-1} \max_{i\in R_1(t_k)} \chi_i(t_k)\Delta \\ &\equiv \varphi(o) + \int_0^t \max_{i\in R_1(\tau)} \chi_i(\tau)\, d\tau \equiv \varphi(o) + \int_0^t \frac{\partial\varphi(\tau)}{\partial g(\tau)}\, d\tau \end{aligned} \tag{3.18}$$

In addition we have proved that the limit in (3.18) exists. Thus (3.9) is proved.

Note that the formula

$$\varphi(y(t)) = \varphi(y_0) + t\,\frac{\partial\varphi(y(\theta))}{\partial g(\theta)}, \qquad \theta \in [0, t],$$

is not valid in this case since $\partial\varphi(y(\tau))/\partial g(\tau)$ is not a continuous function. In this case we will have

$$\varphi(y(t)) = \varphi(y_0) + t\theta$$

where

$$\theta \in \left[\inf_{\tau\in[0,t]} \frac{\partial\varphi(y(\tau))}{\partial g(\tau)},\ \sup_{\tau\in[0,t]} \frac{\partial\varphi(y(\tau))}{\partial g(\tau)}\right].$$

Now let us consider the following system of differential equations

$$\dot{y}_+(t) = g(t), \qquad y(o) = y_0 \tag{3.19}$$

where $g(t)$ is given by

$$\max_{i\in R_1(t)} \left(\frac{\partial f_i(y(t))}{\partial y}, \bar{g}(t)\right) = \min_{\|g\|\leqslant 1} \max_{i\in R_1(t)} \left(\frac{\partial f_i(y(t))}{\partial y}, g\right) \equiv -\rho(t),$$

$$g(t) \equiv \rho(t)\,\bar{g}(t).$$

We shall assume that there exists a solution of system (3.13) for any $t \in [0, \infty)$.

Suppose that the set $M(y_0) \equiv \{x \mid \varphi(x) \leqslant \varphi(y_0)\}$ is bounded, and the $f_i(x)$ are as described above (in Section 1), and $R_1(x)$ is a piecewise continuous set function on $M(y_0)$. Then any limit point of $y(t)$, given by (3.19), is a stationary point of φ, i.e., if $\{t_k\}$ is a sequence such that

$$t_k \xrightarrow[k\to\infty]{} \infty, \qquad y(t_k) \xrightarrow[k\to\infty]{} y^*,$$

then

$$\rho(y^*) \equiv \psi_1(y^*) = 0.$$

Note that since $y(t) \in M(y_0)$ for $t \in [0, \infty)$, then there exists at least one limit point. First of all, for any $\epsilon > 0$, we must have meas $\omega(\epsilon) \leqslant M_\epsilon < \infty$, where $\omega(\epsilon) = \{t \mid \rho(t) \geqslant \epsilon\}$. For otherwise (see (3.9)), $\varphi(y(t)) \xrightarrow[t\to\infty]{} -\infty$, which is impossible by assumption.

Thus

$$\operatorname{ess} \lim_{t\to\infty} \psi(y(t)) = 0.$$

We shall prove that $\psi(y^*) = 0$ where $y^* = \lim_{t_k\to\infty} y(t_k)$.

First of all let us prove that if $\psi\ (y(t_k)) \to 0$, $y(t_k) \to y^*$, then $\psi(x^*) = 0$. In fact, even though the function ψ_1 is not necessarily continuous, it turns out that as $\lim R_1(t_k) \subset R_1(y^*)$, then for $k \geqslant k_l$

$$R_1(t_k) \subset R_1(y^*).$$

Since $\| g \| \leqslant 1$ then for any $\epsilon > 0$ there exists $k_2(\epsilon)$ such that if $k > k_2(\epsilon)$, then

$$\left(\frac{\partial f_i(y(t_k))}{\partial y}, g\right) \leqslant \left(\frac{\partial f_i(y^*)}{\partial y}, g\right) + \epsilon \tag{3.20}$$

uniformly in g and $i \in \overline{1, N}$.

Let $k(\epsilon) = \max\{k_1, k_2(\epsilon)\}$. For $k \geqslant k(\epsilon)$, we have

$$\max_{i\in R_1(t_k)} \left(\frac{\partial f_i(y(t_k))}{\partial y}, g\right) \leqslant \max_{i\in R_1(y^*)} \left(\frac{\partial f_i(y(t_k))}{\partial y}, g\right).$$

Hence,

$$\psi_1(y(t_k)) = \min_{\|g\|\leqslant 1} \max_{i\in R_1(t_k)} \left(\frac{\partial f_i(y(t_k))}{\partial y}, g\right) \leqslant \min_{\|g\|\leqslant 1} \max_{i\in R_1(y^*)} \left(\frac{\partial f_i(y(t_k))}{\partial y}, g\right). \tag{3.21}$$

From (3.20) we obtain

$$\min_{\|g\|\leqslant 1} \max_{i\in R_1(y^*)} \left(\frac{\partial f_i(y(t_k))}{\partial y}, g\right) \leqslant \min_{\|g\|\leqslant 1} \max_{i\in R_1(y^*)} \left(\frac{\partial f_i(y^*)}{\partial y}, g\right) + \epsilon = \psi_1(y^*) + \epsilon. \tag{3.22}$$

Making use of (3.21) and (3.22), we have

$$\psi_1(y(t_k)) \leqslant \psi_1(y^*) + \epsilon,$$

and therefore

$$0 = \lim_{k\to\infty} \psi_1(y(t_k)) \leqslant \psi_1(y^*) + \epsilon.$$

Since ϵ is as small as desired,

$$0 \leqslant \psi_1(y^*). \tag{3.23}$$

On the other hand, by the property of the function ψ_1, we have

$$\psi_1(y^*) \leqslant 0. \tag{3.24}$$

From (3.23) and (3.24) we finally obtain

$$\psi_1(y^*) = 0.$$

Now let us prove that $\psi_1(y^*) = 0$ for any limit point y^*. Suppose that our assumption is false, i.e., $\psi_1(y^*) = -\rho^* < 0$. Since for y close enough to y^*, $R_1(y) \subset R_1(y^*)$, then there exists $\delta > 0$ such that

$$\psi_1(y) \leqslant -\tfrac{1}{2}\rho^* \qquad \text{whenever} \quad \| y - y^* \| \leqslant \delta.$$

Let $\{t_{k_i}\}$ be a subsequence of $\{t_k\}$ such that $\| y_{k_i} - y^* \| \leqslant \frac{1}{2}\delta$. Then, since $g(t)$ is bounded, there exists $\delta_1 > 0$ such that $\| y(t) - y(t_{k_i})\| \leqslant \frac{1}{2}\delta$ whenever $| t - t_{k_i} | \leqslant \delta_1$. This means that for t such that $| t - t_{k_i} | \leqslant \delta_1$, we have $\| y(t) - y^* \| \leqslant \delta$, i.e., $\psi_1(y(t)) \leqslant -\frac{1}{2}\rho^*$. In the sequence $\{t_{i_k}\}$ we leave only terms for which

$$| t_{k_i} - t_{k_{i-1}} | \geqslant \delta_1$$

(Assume that the t_{k_i} have been obtained and that $t_{k(i)}$ is the first t_{k_i} such that $t_{k(i)} - t_{k_i} \geqslant \delta_1$, so that we may put $t_{k_{i+1}} = t_{k(i)}$). Using (3.9), and remembering that $\psi_1(\tau) \leqslant 0$, we have

$$\varphi(y(t)) = \varphi(y_0) + \int_0^t \psi_1(\tau)\, d\tau \leqslant \varphi(y_0) + \sum_{i=1}^{m(t)} \int_{t_{k_i}-\delta_1}^{t_{k_i}+\delta_1} \psi_1(\tau)\, d\tau$$

$$\leqslant \varphi(y_0) - \sum_{i=1}^{m(t)} \rho_i \xrightarrow[t\to\infty]{} -\infty \quad \text{since} \quad \rho_i = -\int_{t_{k_i}-\delta_1}^{t_{k_i}+\delta_1} \psi_1(\tau)\, d\tau \geqslant \rho^* > 0,$$

where $m(t)$ is such that $t_{k_{m(t)}} + \delta_1 \leqslant t < t_{k_{m(t)+1}} + \delta_1$.

This result contradicts the boundedness assumed for $M(y_0)$ and the continuity of φ on E_n. Thus we have obtained that if (3.19) has a solution (we are not discussing here the question of existence and uniqueness of the solution of (3.19)) and if $M(y_0)$ is a bounded set, then this "continuous" method converges to a stationary point of φ. Generally speaking, (3.19) is not stable. This is why some of the well-known numerical methods of solving (3.19) failed to give us a stationary point of φ, and, hence, we were forced to use various alternate methods of successive approximations.

4. Minimization on Bounded Sets

Let $\Omega \subset E_n$ be a compact set (not necessarily convex or connected). An element $g \in E_n$ will be called an admissible direction at the point $y \in \Omega$ if there exists a sequence $\{g_s\}(g_s \in E_n)$ and a number sequence $\{\alpha_s\}$ such that

1) $y + \alpha_s g_s \in \Omega$
2) $g_s \to g$
3) $\alpha_s > 0, \qquad \alpha_s \to 0$

We shall denote the cone of admissible directions by M_y, M_y is a closed set [15]. Now let $\varphi(y) \equiv \min_{i \in \overline{1,N}} f_i(y)$, where the $f_i(y)$ $(i \in \overline{1, N})$ are continuously differentiable on Ω_ϵ where $\Omega_\epsilon = \{x \mid \| x - y \| \leqslant \epsilon, y \in \Omega, \epsilon > 0\}$. We are interested in the minimization of the function $\varphi(y)$ on Ω. The following theorem is valid:

Theorem 4. *In order that a point $y \in \Omega$ be a minimum point of φ on Ω, it is necessary (and in case where the set Ω and the function φ are convex it is also sufficient) that*

$$\psi_1(y) = \min_{\substack{\|g\| \leqslant 1 \\ g \in M_y}} \max_{i \in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, g \right) = 0. \tag{4.1}$$

As usual, we shall call a point y satisfying (4.1) a stationary point of the function φ on the set Ω.

Proof. Necessity. We shall argue by contradiction. Note that for all $y \in \Omega$, we have $\psi_1(y) \leqslant 0$. Let y be a minimum point and suppose that (4.1) is violated, i.e., there exists $\bar{g}$, $\| \bar{g} \| \leqslant 1$, $\bar{g} \in M_y$ such that

$$\max_{i \in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, \bar{g} \right) = -\rho < 0. \tag{4.2}$$

Then from (1.11) we have

$$\varphi(y + \bar{\alpha}_s \bar{g}_s) = \varphi(y) + \bar{\alpha}_s \frac{\partial \varphi(y)}{\partial \bar{g}_s} + o(\bar{\alpha}_s),$$

where $\{\bar{\alpha}_s\}$, $\{\bar{g}_s\}$ are sequences, corresponding to $\bar{g}$ in the definition of M_y. If s is large enough we have from (4.2) that

$$\varphi(y + \bar{\alpha}_s \bar{g}_s) \leqslant \varphi(y) - \tfrac{1}{2}\rho \bar{\alpha}_s + o(\bar{\alpha}_s).$$

Since $o(\bar{\alpha}_s)/\bar{\alpha}_s \xrightarrow[s\to\infty]{} 0$ uniformly with respect to $\bar{g}_s$, we would have for s large enough

$$\varphi(y + \bar{\alpha}_s\bar{g}_s) \leqslant \varphi(y) - \tfrac{1}{4}\rho\bar{\alpha}_s < \varphi(y)$$

which is absurd, since $y + \bar{\alpha}_s\bar{g}_s \in \Omega$ and y is a minimum point of φ on Ω by assumption. Thus necessity is proved.

Sufficiency. Let (4.1) hold at y and let the set Ω and the functions $\varphi(y)$ be convex. We shall prove that y is a minimum point. Assume the contrary. Let $\bar{y} \in \Omega$ be such that

$$\varphi(\bar{y}) < \varphi(y). \tag{4.3}$$

Now let us consider the line segment $[y, \bar{y}]$ entirely contained in Ω since Ω is convex. Since φ is convex we have

$$\varphi(y + \alpha(\bar{y} - y)) = \varphi(\alpha\bar{y} + (1 - \alpha)y) \leqslant \alpha\varphi(\bar{y}) + (1 - \alpha)\varphi(y).$$

From (4.3), we obtain for $g \equiv \bar{y} - y$ that

$$\frac{\partial\varphi(y)}{\partial g} = \lim_{\alpha\to+0} \frac{\varphi(y + \alpha(\bar{y} - y)) - \varphi(y)}{\alpha} \leqslant \varphi(\bar{y}) - \varphi(y) < 0.$$

Then, certainly, for $\bar{g} = \| g \|^{-1} g = \| \bar{y} - y \|^{-1}(\bar{y} - y)$, we have $\partial\varphi(y)/\partial\bar{g} < 0$ which contradicts (4.1). Thus the theorem is proved.

Remark 10. Suppose that M_y is a convex set. Let us form the projection of the convex hull $-L(y)$ onto the come M_y at the point y. We shall denote the convex hull of the projection by $P(y)$. The necessary condition for a minimum means geometrically that the set $P(y)$ must contain the origin. To see this, let (4.1) hold at y and let us prove that the origin indeed belongs to $P(y)$. Let us assume the contrary then there exists $x_1 \in P(y)$ such that

$$\rho(y) = \min_{x\in P(y)} x^2 = x_1{}^2 > 0. \tag{4.4}$$

Let $g_1 = \| x_1 \|^{-1} x_1$. Then we assert that

$$\max_{i\in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, g_1\right) = \max_{z\in L(y)} (z, g_1) = -\min_{z\in -L(y)} (z, g_1) \leqslant -(x_1, g_1) < 0. \tag{4.5}$$

Proof. (by Contradiction). Suppose the contrary. Let $z_1 \in -L(y)$ be such that

$$(z_1, g_1) < (x_1, g_1)$$

The projection $\bar{x}_1$ of z_1 onto M_y is such that (1): $(\bar{x}_1, g_1) \geqslant (x_1, g_1)$ (for otherwise one can find a point x^1 (belonging to the line segment $[\bar{x}_1, x_1]$) such that $x^{12} < x_1{}^2$ which is impossible, because of (4.4).

(2): $(\bar{z}_1 - z_1)^2 < (z_1 - z_1)^2$, (see Fig. 4), because otherwise $\bar{z}_1$ is the projection of z_1 and (4.4) is again violated.

Now from the triangle $\bar{X}_1X_10$ (the angle $\bar{X}_1X_10$ is obtuse) we have that $(Z_1 - \tilde{Z}_1)^2 > (Z_1 - \bar{X}_1)^2$ which is impossible, since $\bar{X}_1$ is the projection of Z_1 onto M_y.

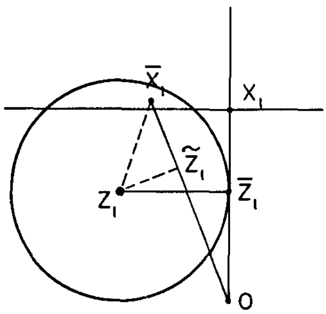

FIG. 4.

If the angle $\bar{X}_10Z_1$ is obtuse then we have again contradiction since $\| Z_1 \| < \| Z_1 - \bar{X}_1 \|$.

Thus we have proved that if (4.5) holds, then (4.1) is violated. The contradiction so obtained proves that the set $P(y)$ contains the origin.

Conversely, let the origin belong to $P(y)$. We shall prove that (4.1) holds at y. Assume the contrary: Let $g_1(\| g_1 \| = 1, g_1 \in M_y)$ be such that

$$\max_{i \in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, g_1 \right) = \max_{x \in L(y)} (x, g_1) = -\rho < 0.$$

This implies

$$\min_{x \in -L(y)} (x, g_1) = -\rho > 0$$

i.e.,

$$\min_{x \in -L(y)} x^2 \geqslant \rho^2 > 0.$$

For every $x \in -L(y)$ we have $(P_x - x)^2 \leqslant (\bar{x} - x)^2$, where P_x is the projection of x onto M_y and $\bar{x}$ is the point of the ray $\{x \mid x = \alpha g_1, \alpha > 0\}$ which is nearest to x. Then (see Figure 5)

$$\| P_x \| \geqslant [\bar{x}^2 + (x - \bar{x})^2]^{1/2} - \| x - \bar{x} \| > 0.$$

Since $L(y)$ is a bounded and closed set, it follows that $\min_{x \in -L(y)} \| P_x \| > 0$, which contradicts the assumption that the origin belongs to $P(y)$.

Thus we have obtained that at a stationary point the origin belongs to the convex hull of the projection of the set $-L(y)$ onto the cone M_y. Note that the projection itself is not a convex set, even though M_y and $-L(y)$ are convex.

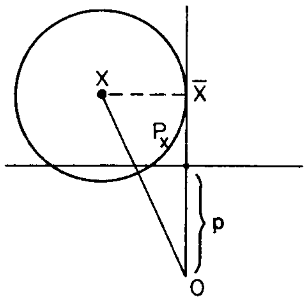

FIG. 5.

Remark 11. Condition (4.1) is equivalent to the condition

$$\psi_2(y) = \min_{\substack{\|g\|=1 \\ g\in M_y}} \max_{i\in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, g\right) \geqslant 0 \tag{4.6}$$

Using the same reasoning as in Section 1, we can obtain that if $\psi_2(y) = r > 0$, then

1) the point y is a local minimum point of φ, and
2) y is a discontinuity point of the set function $R_1(y)$, and moreover, there exists no sequence $\{y_k\}$ such that

$$y_k \in \Omega, \qquad y_k \xrightarrow[k\to\infty]{} y, \qquad R_1(y_k) \longrightarrow R_1(y).$$

Remark 12. If Ω is convex, then (4.1) can be written as

$$\min_{x\in\Omega} \max_{i\in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, x - y\right) = 0. \tag{4.7}$$

Remark 13. If the functions f_i are twice continuously differentiable on $\Omega_\epsilon \cap S$ (where S is some neighborhood of y) and if y is a stationary point of φ, then a sufficient condition for the point y to be a local minimum point of φ on Ω is

$$\min_{\substack{\|g\|=1 \\ g\in M_y}} \max_{i\in R_2(y,g)} \left(\frac{\partial^2 f(y)}{\partial y^2} g, g\right) > 0. \tag{4.8}$$

One can prove this statement by using (1.11) for $l = 2$ and taking into consideration that if (4.8) holds, then for some $\epsilon > 0$

$$\min_{\substack{\|g\|=1 \\ g\in \bar{M}_{\epsilon y}}} \max_{i\in R_2(y,g)} \left(\frac{\partial^2 f(y)}{\partial y^2} g, g\right) > 0$$

where

$$M_{\epsilon y} = \{g \mid y + \alpha g \in \Omega \quad \text{for some} \quad \alpha \in (0, \epsilon], \epsilon > 0\}.$$

It is clear that $M_y \subset \overline{M}_{\epsilon y}$ for any $\epsilon > 0$, but is is not necessarily the case that $M_y \subset M_{\epsilon y}$.

5. Consideration of Some Special Cases

If the set Ω is given by the inequalities

$$\Omega = \{x \mid g_i(x) \leqslant 0, i \in \overline{1, N_1}\} \tag{5.1}$$

where the $g_i(x)$ are continuously differentiable functions on Ω_ϵ, then Theorem 4 can be rewritten as follows:

THEOREM 4'. *In order that a point y, where y is such that $g_i(y) \leqslant 0, i \in \overline{1, N}$ be a minimum point of φ on Ω (given by (5.1)) it is necessary (and in case where $g_i(x)$ are convex and where*

$$\min_{x \in E_n} \psi(x) < 0, \qquad \textit{where} \quad \psi(x) = \max_{j \in \overline{1, N_1}} g_j(x)$$

and the function φ is convex it is also sufficient) that

$$\psi(y) = \min_{\|g\| \leqslant 1} \max \left\{ \max_{i \in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, g \right), \max_{j \in Q(y)} \left(\frac{\partial g_j(y)}{\partial g}, g \right) \right\} = 0, \tag{5.2}$$

where

$$Q(y) = \{j \mid j \in \overline{1, N_1}, g_j(y) = 0\}.$$

Geometrically, condition (5.2) means that at a stationary point the origin must belong to the convex hull $H(y)$ of the vectors

$$\frac{\partial f_i(y)}{\partial y} (i \in R_1(y)) \qquad \text{and} \qquad \frac{\partial g_j(y)}{\partial y} (j \in Q(y)).$$

From the necessary condition, it is also true that if:

There is no stationary point of the function $\max_{j \in \overline{1, N_1}} g_j(x)$ on the set $\{x \mid \max_{j \in \overline{1, N_1}} g_j(x) = 0\}$ (5.3)

then there exist multipliers (so-called Lagrange multipliers) $\lambda_{1i} \geqslant 0$ $(i \in \overline{1, N})$ and $\lambda_{2j} \geqslant 0$ $(j \in \overline{1, N_1})$, where the λ_{1i} are not all zero, such that

$$\sum_{i=1}^{N} \lambda_{1i} \frac{\partial f_i(y)}{\partial y} + \sum_{t=1}^{N_1} \lambda_{2j} \frac{\partial g_j(y)}{\partial y} = 0.$$

Conversely, if there exist real numbers $\lambda_{1i} \geqslant 0 \; (i \in \overline{1, N})$, $\lambda_{2j} \geqslant 0 \; (j \in \overline{1, N_1})$ not all equal to zero, such that (5.4) holds, then y is a stationary point of φ on Ω.

If the convex hull $H(y)$ is a simplex, then the vector $(\lambda_{11}, ..., \lambda_{1N}, \lambda_{21}, ..., \lambda_{2N_1})$ satisfying (5.4) is unique. It is clear, from geometrical reasoning, that the set of vectors in $(N + N_1)$-th dimensional Euclidean space satisfying (5.4) is convex.

Note that (5.4) holds not only if (5.3) is valid, but also in the case where $\lambda_{2j} = 0$ for all $j \in \overline{1, N_1}$.

If $\min_{x \in E_n} \varphi(x) < \min_{x \in \Omega} \varphi(x)$ then in (5.4) the multipliers λ_{2j} are not all zero.

Let us consider the case

$$\Omega = \{x \mid g_i(x) = 0, i = \overline{1, N_1}\} \tag{5.5}$$

where the $g_i(x)$ $(i \in \overline{1, N_1})$ satisfy the above conditions. Generally speaking, Ω is not a convex set (but it is convex if, for example, the $g_i(x)$ are linear functions). The following theorem is valid:

THEOREM 4″. *In order that a point y (where y is such that $g_i(y) = 0$ for all $i \in \overline{1, N_1}$ and $[\partial g_i(y)/\partial y]$ $(i \in \overline{1, N_1})$ are linearly independent) be a minimum point of φ on Ω (given by (5.5)) it is necessary that*

$$\min_{g \in M_y} \max_{i \in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, g \right) = 0, \tag{5.6}$$

where

$$M_y = \left\{ g \mid \| g \| \leqslant 1, \left(\frac{\partial g_j(y)}{\partial y}, g \right) = 0 \text{ for all } j \in \overline{1, N_1} \right\}.$$

The linear independence required in the statement of Theorem 4″ is essential. This may be seen from the following examples: Let

$$\Omega_1 = \{x \mid x \in E_n, (x - A)^2 - 1 = 0, (x + A)^2 - 1 = 0, A \in E_n, \| A \| = 1\}$$

$$\Omega_2 = \{x \mid x \in E_n, (x - A)^2 - 1 = 0, (x - 2A)^2 - 4 = 0, A \in E_n, \| A \| = 1\}$$

In both cases Ω_1 and Ω_2 consist of only one point (the origin), and the vectors $\partial g_j(y)/\partial y$ $(j = 1, 2)$ are linearly dependent. Any function achieves its minimum value on Ω (where Ω is Ω_1 or Ω_2) at 0, but condition (5.6) is not valid.

Finally, let us consider the case where

$$\Omega = \{x \mid g_{1i}(x) \leqslant 0, i \in \overline{1, N_1}, g_{2j}(x) = 0 \; j \in \overline{1, N_2}\} \tag{5.7}$$

and where the $g_i(x)$ and $g_{2j}(x)$ are continuously differentiable on Ω_ϵ. Suppose that at $y \in \Omega$ the vectors $[\partial g_{2j}(y)/\partial y]$ $(j \in \overline{1, N_2})$ are linearly independent. Then the following theorem is valid:

THEOREM 4‴. *In order that the point* $y \in \Omega$ *be a minimum point of* φ *on* Ω *(given by* (5.7)) *it is necessary that*

$$\min_{g \in M_y} \max \left\{ \max_{i \in R_1(y)} \left(\frac{\partial f^i(y)}{\partial y}, g \right), \max_{j \in Q(y)} \left(\frac{\partial g_{1j}(y)}{\partial y}, g \right) \right\} = 0 \tag{5.8}$$

where

$$Q(y) = \{ j \mid j \in \overline{1, N_1}, g_{1j}(y) = 0 \},$$

$$M_y = \left\{ g \mid \| g \| \leqslant 1, \left(\frac{\partial g_{2j}(y)}{\partial y}, g \right) = 0 \text{ for all } j \in \overline{1, N_2} \right\}.$$

Proof. Suppose the contrary. Let y be a minimum point of φ on Ω and let $\tilde{g} \in M_y$ be such that

$$h(\tilde{g}) = \max \left\{ \max_{i \in R_1(y)} \left(\frac{\partial f_i(y)}{\partial y}, \tilde{g} \right), \max_{j \in Q(y)} \left(\frac{\partial g_{1j}(y)}{\partial y}, \tilde{g} \right) \right\} = -\rho < 0.$$

We can assume that $\| \tilde{g} \| = 1$ since $\tilde{g} \neq 0$.

Let Ω' be the intersection of surfaces $g_{2j}(x) = 0$, then every $g \in M_y$ belongs to the tangent plane to Ω' at y.

Now let us consider the ray $y + \alpha\tilde{g} = y_\alpha (\alpha > 0)$ and let $\bar{y}_\alpha \in \Omega'$ be such that

$$\| \bar{y}_\alpha - y \| = \alpha, \qquad \| \bar{y}_\alpha - y_\alpha \| = \min_{\substack{z \in \Omega' \\ \|z - y\| = \alpha}} \| z - y_\alpha \|$$

Note that $h(\alpha\tilde{g}) = -\alpha\rho$.

There exists $\alpha_1 > 0$ such that for any $\alpha \in (0, \alpha_1]$ there is at least one $\bar{y}_\alpha$. We can find $\bar{\alpha} > 0$ such that $\bar{\alpha} < \alpha_1$ and $h(\bar{y}_\alpha - y) \leqslant -\frac{1}{2}\alpha\rho$, then for α sufficiently small $y_\alpha \in \Omega$ and $\varphi(y_\alpha) < \varphi(y)$ which is a contradiction. Thus the theorem is proved.

Necessary conditions for a minimum in different problems have been considered in ([*16*], [*17*]).

6. Methods of Successive Approximations

Let Ω be a compact set of E_n and let the f_i be as described above. Also let

$$R_{1\epsilon}(y) = \{ i \mid i \in \overline{1, N}, \varphi(y) - f_i(y) \leqslant \epsilon, \epsilon > 0 \}.$$

For any g, $\| g \| = 1$, and for $\rho > 0$, we may define the set

$$S(g, \rho) = \left\{ q \mid \| q \| = 1, (g, q) = \cos(\gamma - \beta), \cos\gamma = \frac{\rho}{2M}, \cos\beta = \frac{\rho}{M} \right\}$$

$$M = \sup_{y \in \Omega} \max_{i \in \overline{1, N}} \left\| \frac{\partial f_i(y)}{\partial y} \right\|$$

If for some g, $\| g \| = 1$, we were to have

$$\max_{i \in R_{1\epsilon}(y)} \left(\frac{\partial f_i(y)}{\partial y}, g \right) = -\rho < 0$$

then for $q \in S(g, \rho)$ we have

$$\max_{i \in R_{1\epsilon}(y)} \left(\frac{\partial f_i(y)}{\partial y}, q \right) \leqslant -\tfrac{1}{2}\rho.$$

Let

$$\bar{S}(g, \rho, y) \equiv \{ q \mid q - y = \alpha \bar{q}, q \in S(g, \rho), \alpha > 0 \},$$

so that $\bar{S}$ is a cone. Now suppose that the following condition holds:

Condition A. For every $g \in M_y$, $(\| g \| = 1)$, and for $\rho > 0$, there exists $\delta_1(g, \rho, y) > 0$ such that for any $\delta' \in (0, \delta_1(g, \rho, y))$ we can find $x' \in \Omega \cap \bar{S}(g, \rho, y)$ such that $(g, x' - y) = \delta'$. By $\delta(g, \rho, y)$ we shall denote the greatest such $\delta_1(g, \rho, y)$. In many practical problems, the above condition is realized automatically. Let

$$\chi(g, \epsilon, y) = \max_{i \in R_{1\epsilon}(y)} \left(\frac{\partial f^i(y)}{\partial y}, g \right)$$

Let us define the following set (which may be empty for some δ):

$$M_y(\epsilon, \delta) = \{ g \mid \| g \| = 1, g \in M_y, \delta(g, \chi(g, \epsilon, y), y) \geqslant 0 \}$$

where

$$\bar{\epsilon} > 0, \qquad \bar{\rho} > 0, \qquad \bar{\delta} > 0.$$

For the first approximation, we choose any $y_1 \in \Omega$.

Let y_k have already been found and let $\psi_1(y_k) = -h_k$. If $h_k = 0$, then y_k is a stationary point of φ on Ω and our process is finished. If $h_k > 0$, then we set $\epsilon_{k1} = \bar{\epsilon}$, $\rho_{k1} = \bar{\rho}$, and $\delta_{k1} = \bar{\delta}$. Let $g_k \in M$ be a vector such that

$$g_{kj_k} \in M_{y_k}(\epsilon_{kj_k}, \delta_{kj_k}) \equiv M_{kj_k}, \qquad \bar{\delta}_{kj_k} = \delta(g_{kj_k}, \bar{\rho}_{kj_k}, y_k),$$

$$-\bar{\rho}_{kj_k} = \max_{i \in R_{1\epsilon_{kj_k}}(y_k)} \left(\frac{\partial f_i(y_k)}{\partial y}, g_{kj_k} \right) = \min_{g \in M_{kj_k}} \chi(g, \epsilon_{kj_k}, y_k).$$

If

$$\bar{\rho}_{kj_k} > \rho_{kj_k} \tag{6.1}$$

then we set

$$\delta_k = \delta_{kj_k}, \qquad \bar{\delta}_k = \bar{\delta}_{kj_k}, \qquad \rho_k = \rho_{kj_k}, \qquad \bar{\rho}_k = \bar{\rho}_{kj_k}, \qquad \epsilon_k = \epsilon_{kj_k},$$

and $\bar{g}_k = g_{kj_k}$. If (6.1) is not satisfied, then we set $\epsilon_{kj_k+1} = \frac{1}{2}\epsilon_{kj_k}$, $\rho_{kj_k+1} = \frac{1}{2}\rho_{kj_k}$, $\delta_{kj_k+1} = \frac{1}{2}\delta_{kj_k}$ and continue in the same manner until δ_k, $\bar{\delta}_k$, ρ_k, $\bar{\rho}_k$, ϵ_k, and $\bar{g}_k$ are found. Since $\rho_k > 0$, there do indeed exist such $\delta_k > 0$, $\bar{\delta}_k > 0$, $\rho_k > 0$, $\bar{\rho}_k > 0$, $\epsilon_k > 0$. Now let us consider the set

$$\{y_k(\alpha)\} = \{x \mid x \in \Omega \cap \bar{S}(\bar{g}_k, \bar{\rho}_k, y_k), (g_k, x - y) = \alpha\}$$

where $\{y_k(\alpha)\}$ is not empty, by the above assumption, for any $\alpha \in [0, \bar{\delta}_k]$. For $\alpha > \bar{\delta}_k$ it may turn out that $\{y_k(\alpha)\}$ is an empty set. Let $y_{k\alpha}$ be the point of $\{y_k(\alpha)\}$ which is nearest to the point $(y_k + \alpha g_k)$.

Let us find $\alpha_k \in [0, \infty)$ such that

$$\varphi(y_{k\alpha_k}) = \min_{\alpha \in [0, \infty)} \varphi(y_{k\alpha}).$$

Now if we set $y_{k+1} = y_{k\alpha_k}$, then clearly $\varphi(y_{k+1}) < \varphi(y_k)$. We may continue in the same manner. Thus we have developed a sequence $\{y_k\}$, such that

$$\varphi(y_{k+1}) < \varphi(y_k). \tag{6.2}$$

If this sequence is finite (i.e., contains a finite number of vectors) then the rightmost element of this sequence is a stationary point of φ on Ω. If the sequence $\{y_k\}$ contains infinite number of terms, then the following theorem holds:

THEOREM 5.

$$\lim_{k \to \infty} \rho_k = 0, \qquad \lim_{k \to \infty} \bar{\rho}_k = 0 \tag{6.3}$$

and any limit point of the sequence $\{y_k\}$ is a stationary point of the function φ on Ω.

Proof. Let y^* be a limit point of $\{y_k\}$, i.e., there exists a subsequence $\{y_{k_l}\}$ of $\{y_k\}$ such that $y_{k_l} \xrightarrow[k_l \to \infty]{} y^*$. Let C denote $\lim_{k\to\infty} \varphi(y_k)$. From (6.2) we have

$$\varphi(y_k) \geqslant C. \tag{6.4}$$

Since Ω is a compact set, and since the f_i are continuously differentiable, we have $C > -\infty$.

If we suppose that $g^* = \lim_{k_l\to\infty} g_{k_l}$, $\bar{g}^* = \lim_{k_l\to\infty} \bar{g}_{k_l}$, then two cases are possible:

Case (1). There exist as many k_l as may be desired such that

$$\epsilon_{k_l} \geqslant \epsilon^* > 0, \qquad \bar{\rho}_{k_l} \geqslant \rho^* > 0, \qquad \bar{\delta}_{k_l} \geqslant \delta^* > 0 \tag{6.5}$$

i.e.,

$$\max_{i \in R_{1\epsilon_{k_l}}(y_{k_l})} \left(\frac{\partial f_i(y_{k_l})}{\partial y}, \bar{g}_{k_l}\right) \leqslant -\rho^*.$$

For k_l large enough and satisfying (6.5)

$$R_{1\epsilon_{k_l}}(y_{k_l}) \supset R_{1(\epsilon^*/2)}(y^*)$$

and

$$\max_{i \in R_{1(\epsilon^*/2)}(y^*)} \left(\frac{\partial f_i(y^*)}{\partial y}, g^*\right) \leqslant -\tfrac{1}{2}\rho^*, \qquad \varphi(y^*) = C,$$

$\delta = \delta(g^*, \rho^*, y^*) \geqslant \delta^*$. Let $y(\alpha) \in \Omega \cap \bar{S}(g^*, \rho^*, y^*)$ be a point such that $(g^*, y(\alpha) - y) = \alpha$. Then for any $\alpha \in [0, \delta]$ there exists such a $y(\alpha)$. By definition of $\delta(g, \rho, y)$, we have

$$\max_{i \in R_{1(\epsilon^*/2)}(y^*)} \left(y(\alpha) - y^*, \frac{\partial f_i(y^*)}{\partial y}\right) \leqslant -\tfrac{1}{4}\rho^*$$

We can find $\alpha^* \in (0, \bar{\delta})$ such that

$$\varphi(y(\alpha^*)) \leqslant C - \mu < C, \qquad \mu > 0$$

For k_l large enough (and satisfying (6.5)) we have

$$\varphi(y_{k_l}(\alpha^*)) \leqslant C - \tfrac{1}{2}\mu < C$$

where $y_{k_l}(\alpha^*) \in \Omega \cap \bar{S}(\bar{g}_{k_l}, \bar{\rho}_{k_l}, y_{k_l})$ is such that $(\bar{g}_{k_l}, y_{k_l}(\alpha^*) - y_{k_l}) = \alpha^*$. The inequality (6.6) is valid for all such $y_{k_l}(\alpha^*)$. Moreover,

$$\varphi(y_{k_l+1}) = \min_{\alpha \in [0, \infty)} \varphi(y_{k_l\alpha}) \leqslant \varphi(y_{k_l\alpha^*}) = \varphi(y_{k_l}(\alpha^*)) < C$$

which contradicts (6.4). Thus, we have established the correctness of (6.3).

Case (2). $\epsilon_{k_l} \to 0$, $\bar{\rho}_{k_l} \to 0$, $\bar{\delta}_{k_l} \to 0$ (these sequences either tend to zero or do not tend to zero simultaneously). We claim in this case that y^* is a stationary point, i.e., $\psi_1(y^*) = 0$.

Proof by Contradiction. Assume the contrary. Let $g(y^*) \in M_{y^*}$, $\|g(y^*)\| = 1$, be such that

$$\max_{i \in R_1(y^*)} \left(\frac{\partial f_i(y^*)}{\partial y}, g(y^*)\right) = -h < 0, \qquad \text{let} \quad \delta(g(y^*), h, y^*) = \delta_1 > 0.$$

Then for $\epsilon \in (0, \epsilon_0]$, $\epsilon_0 > 0$, we have

$$\max_{i \in R_{1\epsilon}(y^*)} \left(\frac{\partial f_i(y^*)}{\partial y}, g(y^*)\right) \leqslant -\tfrac{1}{2}h < 0.$$

Choose another such ϵ so that for k_l large enough we have

$$R_{1\epsilon}(y^*) \supset R_{1\epsilon'}(y_{k_l}) \qquad \text{for any} \quad \epsilon' \in (0, \tfrac{1}{2}\epsilon]$$

and

$$\max_{i\in R_{1\epsilon'}(y_{k_l})}\left(\frac{\partial f_i(y^*)}{\partial y}, g(y^*)\right) \leqslant \max_{i\in R_{1\epsilon}(y^*)}\left(\frac{\partial f_i(y^*)}{\partial y}, g(y^*)\right),$$

and

$$\max_{i\in R_{1\epsilon'}(y_{k_l})}\left(\frac{\partial f_i(y_{k_l})}{\partial y}, g(y^*)\right) \leqslant \max_{i\in R_{1\epsilon'}(y_{k_l})}\left(\frac{\partial f_i(y^*)}{\partial y}, g(y^*)\right) + \tfrac{1}{4}h$$

$$\leqslant \max_{i\in R_{1\epsilon}(y^*)}\left(\frac{\partial f_i(y^*)}{\partial y}, g(y^*)\right) + \tfrac{1}{4}h < -\tfrac{1}{4}h < 0.$$

Since $\delta_1 > 0$ for k_l large enough, then for any $\epsilon' \in (0, \bar{\epsilon}]$, where $0 < \bar{\epsilon} \leqslant \epsilon_0$ and where $\bar{\epsilon}$ does not depend on k_l, it turns out that $g(y^*) \in M_{y_{k_l}}(\epsilon', \tfrac{1}{2}\delta_1)$. Now choose another such ϵ'. Then

$$\min_{g\in M_{y_{k_l}}(\epsilon', \frac{1}{2}\delta_1)} \chi(g, \epsilon') \leqslant \max_{i\in R_{1\epsilon'}(y_{k_l})}\left(\frac{\partial f_i(y_{kl})}{\partial y}, g(y^*)\right) \leqslant -\tfrac{1}{4}h$$

which contradicts the assumption that $\epsilon_{k_l} \to 0$, $\bar{\rho}_{k_l} \to 0$, $\bar{\delta}_{k_l} \to 0$. This contradiction completes the proof of Theorem 5.

Remark 14. Instead of Condition A, we could assume that for every $g \in M_y$, $\|g\| = 1$, and for every $\rho > 0$ there exists $\delta_1(g, \rho, y) \in \{\delta_k\}$, where $\delta_k > 0$, $\delta_k \xrightarrow[k\to\infty]{} 0$, $\delta_{k+1} < \delta_k$, i.e., $\delta_1(g, \rho, y) = \delta_{k_1}$ is such that for any $\delta_k(k > k_1)$ we can find at least one $x' \in \Omega \cap \bar{S}(g, \rho, y)$ such that $(g, x' - y) \in [\delta_k, \delta_{k-1})$ where $\delta(g, \rho, y)$ denotes the largest $\delta_1(g, \rho, y)$.

7. Special Case

Consider the case where Ω is convex. Then (4.7) is a necessary condition for a minimum. Let $Q_\alpha(y) \equiv \{x \in \Omega \mid \|x - y\| \leqslant \alpha > 0\}$. Suppose $x(y, \alpha, \epsilon)$ is a point such that $x(y, \alpha, \epsilon) \in Q_\alpha(y)$ and

$$J_1(y, \alpha, \epsilon) = \min_{x\in Q_\alpha(y)} \max_{i\in R_{1\epsilon}(y)}\left(\frac{\partial f_i(y)}{\partial y}, x - y\right) = \max_{i\in R_{1\epsilon}(y)}\left(\frac{\partial f_i(y)}{\partial y}, x(y, \alpha, \epsilon) - y\right). \quad (7.1)$$

Let us choose $\bar{\epsilon} > 0, \bar{\rho} > 0$. For the first approximation, we select any $y_1 \in \Omega$ and begin the iterations. Suppose we have arrived at y_k and found $\psi_1(y_k) = -h_k$. If $h_k = 0$, then y_k is a stationary point, and we stop. Otherwise, if $h_k > 0$, then we set $\epsilon_{k_1} = \bar{\epsilon}$ and $\bar{\rho}_{k_1} = \bar{\rho}$. Let $\tilde{y}_k \in M$ be such that

$$-h_k = \max_{i\in R_1(y_k)}\left(\frac{\partial f_i(y_k)}{\partial y}, \tilde{y}_k - y_k\right) = \min_{i\in\Omega} \max_{i\in R_1(y_k)}\left(\frac{\partial f_i(y_k)}{\partial y}, x - y_k\right).$$

Suppose that

$$-\bar{h}_{k_1} = \max_{i \in R_{1\epsilon_{k_1}}(y_k)} \left(\frac{\partial f_i(y_k)}{\partial y}, \bar{y}_{k_1} - y_k \right) = \min_{x \in \Omega} \max_{i \in R_{1\epsilon_{k_1}}(y_k)} \left(\frac{\partial f_i(y_k)}{\partial y}, x - y_k \right).$$

If $\bar{h}_{k_1} \geqslant \rho_{k_1}$, then we set $\epsilon_k = \epsilon_{k_1}$, $\rho_k = \rho_{k_1}$, $\bar{h}_k = \bar{h}_{k_1}$, $\bar{y}_k = \bar{y}_{k_1}$, and $R_{1\epsilon_k} = R_{1\epsilon_{k_1}}(y_k)$. Otherwise if $\bar{h}_{k_1} < \rho_{k_1}$, we set $\epsilon_{k_2} = \frac{1}{2}\epsilon_{k_1}$ and continue in the same manner until ϵ_k, ρ_k, $\bar{h}_k$, $\bar{y}_k$, $R_{1\epsilon_k}$ are found (such that $\bar{h}_k > \rho_k$).

The next approximation y_{k+1} can be chosen by one of the following methods.

Method 1. Let us consider the linear segment $y_{k\alpha} = y_k + \alpha(\bar{y}_k - y_k)$, $\alpha \in [0, 1]$, where $y_{k\alpha} \in \Omega$ (since Ω is convex), and let us find $\alpha_k \in [0, 1]$ such that $\varphi(y_{k\alpha_k}) = \min_{\alpha \in [0,1]} \varphi(y_{k\alpha})$, and then set $y_{k+1} = y_{k\alpha_k}$. In addition, we have $\varphi(y_{k+1}) < \varphi(y_k)$. We may continue this process to subsequent steps.

Method 2. Let $y_{k\alpha} = x(y_k, \alpha, \epsilon_k)$ (see (7.1)) and let us find $\alpha_k \in [0, \infty)$ such that $\varphi(y_{k\alpha_k}) = \min_{\alpha \in [0,\infty)} \varphi(y_{k\alpha})$ and set $y_{k+1} = y_{k\alpha_k}$.

Method 3. Fix any $\alpha^* > 0$ (not depending on k) and let

$$y_{k\alpha} = y_k + \alpha(x(y_k, \alpha^*, \epsilon_k) - y_k), \qquad \alpha \in [0, 1], \qquad y_{k\alpha} \in \Omega,$$

and let $\alpha_k \in [0, 1]$ be such that

$$\varphi(y_{k\alpha_k}) = \min_{\alpha \in [0,1]} \varphi(y_{k\alpha}).$$

Then $y_{k+1} = y_{k\alpha_k}$, and so on. Generally speaking, Method 3 becomes Method 1 when α^* tends to infinity.

The convergence of the sequence $\{y_k\}$ (constructed according to Method 1) to a stationary point has been proved in [*1*]. We can prove this fact in a similar manner for Method 3. For Method 2, the following theorem holds:

THEOREM 6. *Any limit point of the sequence $\{y_k\}$ constructed according to Method 2 is a stationary point of φ on Ω.*

Proof. Let

$$C = \lim_{k \to \infty} \varphi(y_k), \qquad \varphi(y_k) \geqslant C \tag{7.2}$$

Let y^* be a limit point of $\{y_k\}$, i.e., there exists a subsequence of $\{y_{k_l}\}$ of $\{y_k\}$ such that $y_{k_l} \xrightarrow[k_l \to \infty]{} y^*$. We can assume that $\bar{y}_{k_l} \to \bar{y}, \tilde{y}_{k_l} \to \tilde{y}$. Two cases are possible:

Case (1). There exist as many k_l as may be desired such that

$$\epsilon_{k_l} \geqslant \epsilon^* > 0, \qquad \rho_{k_l} \geqslant \rho^* > 0 \tag{7.3}$$

For such k_l,

$$\| \bar{y}_{k_l} - y_{k_l} \| \geqslant \rho_{k_l} M^{-1} \geqslant \frac{\rho^*}{M} \quad \left(\text{where } M = \max_{x \in \Omega} \max_{i \in \overline{1,N}} \left\| \frac{\partial f_i(x)}{\partial x} \right\|\right).$$

Then we have for $\alpha \in (0, M^{-1}\rho^*)$

$$h(y_{k_l}, \alpha, \epsilon_{k_l}) \leqslant \max_{i \in R_{1\epsilon}(y)} \left(\frac{\partial f_i(y)}{\partial y}, \frac{\alpha(\bar{y}_{kl} - y_{kl})}{\| \bar{y}_{k_l} - y_{k_l} \|}\right) \leqslant \frac{-\alpha \rho_{kl}}{\| \bar{y}_{k_l} - y_{k_l} \|} \leqslant -\frac{\alpha \rho^*}{\mathscr{D}},$$

where $\mathscr{D}$ is the diameter of Ω (since the point $z_{k_l\alpha} = \alpha(\bar{y}_{k_l} - y_{k_l}) \| \bar{y}_{k_l} - y_{k_l} \|^{-1} \in Q_\alpha(y)$). Thus for k_l large enough and satisfying (7.3), and for $\alpha \in (0, M^{-1}\rho^*)$, we have

$$f_i(y_{k_l\alpha}) \leqslant f_i(y_{k_l}) - \alpha\rho^* \mathscr{D}^{-1} + o_i(\alpha)$$

$$\varphi(y_{k_l\alpha}) \leqslant \varphi(y_{k_l}) - \alpha\rho^* \mathscr{D}^{-1} + o(\alpha)$$

where $o(\alpha)$ does not depend on k_l. For α sufficiently small,

$$| o(\alpha) | \leqslant (1/2\mathscr{D})\, \alpha\rho^*, \qquad \varphi(y_{k_l\alpha}) \leqslant \varphi(y_{k_l}) - (1/2\mathscr{D})\, \rho^*\alpha.$$

Let us choose such an α, which we shall designate as α^*. Since $\varphi(y_{k_l}) \to C$, for k_l large enough, we obtain

$$\varphi(y_{k_l\alpha^*}) \leqslant C - (1/4\mathscr{D})\, \rho^*\alpha^*.$$

Furthermore,

$$\varphi(y_{k_l+1}) = \min_{\alpha \in [0, \infty)} \varphi(y_{k_l\alpha}) \leqslant \varphi(y_{k_l\alpha^*}) \leqslant C - (1/4\mathscr{D})\, \rho^*\alpha^* < 0,$$

which contradicts (7.2). Thus (7.3) is impossible.

Case (2). $\epsilon_{k_l} \to 0$, $\rho_{k_l} \to 0$. Then we claim that y^* is a stationary point. We shall argue by contradiction. Suppose that for $x^* \in \Omega$ we have

$$\max_{i \in R_1(y^*)} \left(\frac{\partial f_i(y^*)}{\partial y}, x^* - y^*\right) = -h < 0.$$

Then there exists $\epsilon^* > 0$ such that for $\epsilon \in (0, \epsilon^*]$

$$\max_{i \in R_{1\epsilon}(y^*)} \left(\frac{\partial f_i(y^*)}{\partial y}, x^* - y^*\right) \leqslant -\tfrac{1}{2}h < 0.$$

Choose some such ϵ. For k_l large enough, we have

$$\max_{i \in R_{1(\epsilon/2)}(y_{k_l})} \left(\frac{\partial f_i(y_{kl})}{\partial y}, x^* - y_{k_l}\right) \leqslant -\tfrac{1}{4}h.$$

Moreover,

$$\min_{x\in\Omega}\max_{i\in R_{1(\epsilon/2)}(y_{k_l})}\left(\frac{\partial f_i(y_{kl})}{\partial y}, x-y_{k_l}\right)\leqslant -\tfrac{1}{4}h.$$

But now it is obvious that if $\epsilon_{k_j}\leqslant\frac{1}{2}\epsilon$, then $\bar{h}_{k_j}\geqslant\frac{1}{4}h$, and for $\rho_{k_j}\leqslant\frac{1}{4}h$ we have $\bar{h}_{k_j}\geqslant\rho_{k_j}$, i.e., neither ϵ_k nor ρ_k tend to zero, which contradicts our assumption. This contradiction completes the proof of Theorem 6.

Remark 15. In Method 2, it is not necessary to choose $\alpha_k\in[0,\infty)$. Rather, it is sufficient to find $\alpha_k\in[0,\alpha^*]$ (where $\alpha^*>0$ is fixed and does not depend on k) such that

$$\varphi(y_{k\alpha_k})=\min_{\alpha\in[0,\alpha^*]}\varphi(y_{k\alpha})\qquad\text{and set}\quad y_{k+1}=y_{k\alpha_k}.$$

In all these methods it is necessary to solve some auxiliary optimization problems. In just the same way as in [*14*] one can solve all these problems approximately.

8. Nonlinear Mathematical Programming Problems

Let the set Ω be given by (5.1). To obtain a stationary point (satisfying (5.2)), we can apply the following modification to the above algorithms. Let us define

$$Q_\epsilon(y)\equiv\{j\in\overline{1,N_1}\mid -\epsilon\leqslant g_j(y)\leqslant 0,\epsilon>0\}\tag{8.1}$$

Let us choose any $\epsilon>0,\rho>0,\mu>0$.

For the first approximation, let us choose any arbitrary $y_1\in\Omega$ and begin the iterations. Suppose we have arrived at y_k. Let ψ_k, R_{1k}, Q_k denote $\psi(y_k)$, $R_1(y_k)$, $Q(y_k)$, respectively. If $\psi_k=0$ then y_k is a stationary point, and the process is finished. But if $\psi_k<0$, then we set $\epsilon_{k_1}=\epsilon$, $\rho_{k_1}=\rho$, $\mu_{k_1}=\mu$ and find

$$\psi_{k\epsilon_{k1}}=\min_{\|g\|\leqslant 1}\max\left\{\max_{i\in R_{1\epsilon_{k1}}(y_k)}\left(\frac{\partial f_i(y_k)}{\partial y},g\right), h\cdot\max_{j\in Q_{\mu_{k1}}(y_k)}\left(\frac{\partial g_j(y_k)}{\partial y},g\right)\right\}\tag{8.2}$$

where $h>0$ is in general an arbitrary number not depending on k. Let g_{k1} be a vector such that $\|g_{k1}\|\leqslant 1$ and

$$\psi_{k\epsilon_{k1}}=\max\left\{\max_{i\in R_{1\epsilon_{k1}}(y_k)}\left(\frac{\partial f_i(y_k)}{\partial y},g_{k1}\right), h\cdot\max_{j\in Q_{\mu_{k1}}(y_k)}\left(\frac{\partial g_j(y_k)}{\partial y},g_{k1}\right)\right\}.\tag{8.3}$$

If $-\psi_{k\epsilon_{k1}}\geqslant\rho_{k1}$, we set $\epsilon_k=\epsilon_{k1}$, $\rho_k=\rho_{k1}$, $\mu_k=\mu_{k1}$, and $g_k=g_{k1}$. However, if $-\psi_{k\epsilon_{k1}}<\rho_{k1}$, we let $\epsilon_{k2}=\frac{1}{2}\epsilon_{k1}$, $\rho_{k2}=\frac{1}{2}\rho_{k1}$, $\mu_{k2}=\frac{1}{2}\mu_{k1}$, and we continue in an analogous manner until r_k is found such that $-\psi_{k\epsilon_{kr_k}}\geqslant\rho_{kr_k}$, and then we set $\epsilon_k=\epsilon_{kr_k}$,

$\rho_k = \rho_{kr_k}$, $\mu_k = \mu_{kr_k}$, and $g_k = g_{kr_k}$. Such an r_k will necessarily be found since $\psi_k < 0$. Let us now consider the ray $y_{k\alpha} = y_k + \alpha g_k (\alpha > 0)$. (Note that for α small enough $y_{k\alpha} \in \Omega$ because $\psi_{k\epsilon_k} \leqslant -\rho_k < 0$.) Find $\alpha_k \in A_k$ such that

$$\varphi(y_{k\alpha_k}) = \min_{\alpha \in A_k} \varphi(y_{k\alpha})$$

where

$$A_k = \{\alpha \mid \alpha > 0, y_k + \alpha g_k \in \Omega\}$$

and set $y_{k+1} = y_{k\alpha_k}$. It is obvious that $\varphi(y_{k+1}) < \varphi(y_k)$. We continue in the same manner. The sequence $\{y_k\}$ constructed in this fashion converges to a stationary point. This statement can be proved by arguments similar to those used for proving the above theorems.

In the above method, we supposed that $y_1 \in \Omega$. The following method is free of this assumption. Let the f_i and g_i be defined on the entire space. Choose any $\epsilon > 0$, $\rho > 0$, $\mu > 0$. For the first approximation we choose an arbitrary $y_1 \in E_n$. Let y_k have already been found. Let us denote $h_k = \max_{j \in \overline{1, N_1}} g_j(y_k)$. If $h_k \leqslant 0$ and $\psi(y_k) = 0$ (see 5.2), then y_k is a stationary point of φ on Ω and our process is finished. Otherwise we take $\epsilon_{k1} = \epsilon$, $\rho_{k1} = \rho$, and $\mu_{k1} = \mu$. Three cases are possible.

Case (1). $h_k > \mu_{k1}$. Then we set $R_{1\epsilon_{k1}} = \Lambda$ (an empty set) and

$$Q_{\mu_{k1}} = \{j \mid j \in \overline{1, N_1}, h_k - g_j(y_k) \leqslant \mu_{k1}\}.$$

Case (2). $0 < h_k \leqslant \mu_{k1}$. Then we set

$$R_{1\epsilon_{k1}} = \{i \mid i \in \overline{1, N}, \max_{j \in \overline{1, N}} f_j(y_k) - f_i(y_k) \leqslant \epsilon_{k1}\}$$

$$Q_{\mu_{k1}} = \{j \mid j \in \overline{1, N_1}, h_k - g_j(y_k) \leqslant \mu_{k1}\}.$$

Case (3). $h_k \leqslant 0$. Then we set

$$R_{1\epsilon_{k1}} = \{i \mid i \in \overline{1, N}, \max_{j \in \overline{1, N}} f_j(y_k) - f_i(y_k) \leqslant \epsilon_{k1}\}$$

$$Q_{\mu_{k1}} = \{j \mid j \in \overline{1, N_1}, -\mu_{k1} \leqslant g_j(y_k) \leqslant 0\}$$

(clearly, if $h(y_k) < -\mu_{k1}$, then $Q_{\mu_{k1}} = \Lambda$.)

Now we find $\psi_{k\epsilon_{k1}}$ (see 8.2) and the vector g_{kl} (see 8.3). Then we find g_k, ϵ_k, ρ_k, and μ_k as above. Then we form the ray $y_{k\alpha} = y_k + \alpha g_k (\alpha > 0)$ and find $\alpha_k > 0$ such that:

1) If $h(y_k) > 0$ and $R_{1\epsilon_k} = \Lambda$ then

$$\max_{j\in\overline{1,N_1}} g_j(y_{k\alpha_k}) = \min_{\alpha\in[0,\infty)} g_j(y_{k\alpha}).$$

2) If $h(y_k) > 0$ and $R_{1\epsilon_k} \neq \Lambda$ then

$$\max\{\varphi(y_{k\alpha_k}) - \varphi(y_k), \max_{j\in\overline{1,N_1}} g_j(y_{k\alpha_k}) - h(y_k)\}$$
$$= \min_{\alpha\in[0,\infty)} \max\{\varphi(y_{k\alpha}) - \varphi(y_k), \max_{j\in\overline{1,N_1}} g_j(y_{k\alpha}) - h(y_k)\}.$$

3) If $h(y_k) \leqslant 0$, then

$$\varphi(y_{k\alpha_k}) = \min_{\alpha\in A_k} \varphi(y_{k\alpha})$$

where

$$A_k = \{\alpha \mid \alpha \geqslant 0, \max_{j\in\overline{1,N_1}} g_j(y_{k\alpha}) \leqslant 0\}.$$

Now we set $y_{k+1} = y_{k\alpha_k}$ and continue in the same manner. Note that if y_k is far enough from Ω, the next y_{k+1} is chosen as if we were minimizing $h(y) \equiv \max_{j\in\overline{1,N_1}} g_j(y)$. If $y_k \in \Omega$, then all succeeding y_k are chosen as in the previous method. We have some difficulties only if y_k does not belong to Ω but is close enough to it.

The sequence $\{y_k\}$ thus constructed converges to a stationary point of φ on Ω in the sense of the following theorem:

THEOREM 7. *Let y^* be any limit point of $\{y_k\}$*, i.e., *there exists a subsequence $\{y_{k_l}\}$ of $\{y_k\}$ such that $y_{k_l} \xrightarrow[k_l\to\infty]{} y^*$, $y_{k_l} \in \{y_k\}$. The following statements are true*:

1) *If $y^* \notin \Omega$, then*

$$\min_{\|g\|\leqslant 1} \max_{j\in\mathscr{D}(y^*)} \left(\frac{\partial g_j(y^*)}{\partial y}, g\right) = 0$$

where

$$\mathscr{D}(y^*) = \{j \mid j \in \overline{1, N_1}, g_j(y^*) = \max_{k\in\overline{1,N_1}} g_k(y^*)\}$$

2) *If $y^* \in \Omega$ then*

$$\min_{\|g\|\leqslant 1} \max \left\{\max_{i\in R_1(y^*)} \left(\frac{\partial f_i(y^*)}{\partial y}, g\right), \max_{j\in Q(y)} \left(\frac{\partial f_j(y^*)}{\partial y}, g\right)\right\} = 0,$$

where $Q(y)$ is the same as in (5.2).

Remark 16. We can choose h, ϵ, ρ and μ in Section 8, taking into consideration previous computational experience. For each particular class of problems, reasonable auxiliary coefficients h, ϵ, ρ and μ may be found.

Remark 17. If $N = 1$, then $\varphi(X) \equiv f(X)$, where f is a continuously differentiable function. Then Ω is given by (5.1), and we have a standard mathematical programming problem. Our methods for this particular case are similar to those of G. Zoutendijk ([*5*], [*8*]) and Zuchovitskii ([*5*], [*1*]).

In (5.2) we have $\|g\| \leqslant 1$. Instead of the unit sphere we can consider an arbitrary closed bounded set with the origin as an interior point. By choosing different sets one can obtain new methods for solving nonlinear mathematical programming problems. These correspond to different Zoutendijk normalizations.

Remark 18. For each of the methods in Section 2, we can form systems of differential equations similar to that in Section 1 whose solutions, under suitable assumptions, tend to a stationary point. However, we have no assurance that these solutions are stable. This instability is a reason for the so-called zigzagging effects of numerical methods for solving these differential equations.

Remark 19. The methods discussed above are applicable if the auxiliary linear problem arising therein can easily be solved. In many cases this problem is simple enough (for example if Ω is given by (5.11)). In other cases it is necessary to try various methods of successive approximation for solving the auxiliary linear problem. This problem has been discussed in detail in [*18*] for some particular cases of Ω.

Remark 20. Other methods of successive approximation can be obtained by using modifications of the algorithms in Section 3.

Acknowledgment

This paper was written during the author's visit to the University of Southern California, Los Angeles. The author considers it a pleasant duty to express his gratitude to Lucien W. Neustadt for his useful comments during some discussions, and to James R. La Frieda and David Wallis who checked and corrected the author's English.

References

1. V. F. Dem'yanov. On the solution of certain minimax problems, I, II. *Kybernetica*. **2**, 58–66 (1966); **3**, 62–66 (1967).
2. M. Frank and P. Wolfe. An algorithm for quadratic programming. *Naval Res. Logist. Quart.* **3**, 95–110 (1956).
3. E. M. L. Beale. On quadratic programming. *Naval Res. Logist. Quart.* **6**, 227–243 (1959).
4. G. B. Dantzig. "Linear Programming and Extensions." Princeton University Press, Princeton, 1963.

5. G. Zontendijk. "Methods of Feasible Directions." Elsevier, Amsterdam, 1960.
6. S. I. Zuchovitskiy, G. A. Polyak, and M. E. Primak. An algorithm for solving convex chebyshev approximation Problems. *Dokl. Akad. Nauk SSSR*, **151**, 27–30 (1963). [English transl. *Soviet Math.* **4**, 901–904 (1963).
7. S. I. Zuchovitskiy, G. A. Polyak, and M. E. Primak. A Numerical Method for Solving a Convex Programming Problem in Hilbert Space. *Dokl. Akad. Nauk SSSR.* **163**, 282–284 (1965). [English transl. *Soviet Math.* **6**, 903–905 (1965).
8. G. Zontendijk. Nonlinear programming: a numerical survey. *SIAM J. Control.* **4**, 194–210 (1966).
9. J. E. Kelley, Jr. The cutting-plane method for solving convex programs. *J. Soc. Indust. Appl. Math.* **8**, 703–712 (1960).
10. J. B. Rosen. The gradient projection method for nonlinear programming, I. *J. Soc. Indust. Appl. Math.* **8**, 181–217 (1960).
11. P. Wolfe. "Methods of Nonlinear Programming, Recent Advances in Mathematical Programming." R. L. Graves and P. Wolfe, eds. McGraw-Hill, New York, 1963, 67–86.
12. E. S. Levitin and B. T. Poliak. Methods for constrained minimization. *Zh. Vychisl. Mat. i Mat. Fiz.* **6**, 787–823 (1966).
13. N. E. Kirin. On programming optimization of linear systems in the presence of constraints on phase coordinates. Proceedings of III All-Union Conference on Control (Technical Cybernetics), Optimal Systems and Statistics Methods, Nauka, Moscow, 1967, 92–98.
14. V. F. Dem'yanov and A. M. Rubinov. Minimizing a smooth convex functional on a convex set. Transl. in *SIAM J. Control.* **5**, 280–294 (1967). *Vestnik Leningrad Univ.* **19**, 5–17 (1964).
15. V. F. Dem'yanov and A. M. Rubinov. On necessary conditions for an extremum. *Economika i Matematicheskie Metody.* **2**, 406–417 (1966).
16. A. Ya. Dubovitzkii and A. A. Milintin. Extremum problems with constraints. *Zh. Vychsl. Mat. i Mat. Fiz.* **5**, 395–453 (1965).
17. L. W. Neustadt. An abstract variational theory with applications to a broad class of optimization problems, I and II. *SIAM J. Control.* **4**, 505–525 (1966); **5**, 90–137 (1967).
18. B. N. Pshenichniy. A dual method in extremal problems I and II. *Kybernetica.* **1**, no. 3, pp. 89–95 (1965); no. 4, pp. 64–69.