Докл. Акад. Наук СССР
Том 278 (1984), N° 1, 24–27

Soviet Math. Dokl.
Vol. 30 (1984), N° 2, 313–316

SUFFICIENT CONDITIONS FOR A MINIMUM FOR NONLINEAR PROGRAMMING PROBLEMS

UDC 519.8

YU.G. EVTUSHENKO AND V.A. PURTOV

Revised 18 June 2002

$\boxed{\textbf{1.}}$ Consider the general nonlinear programming problem

$$\min_{x\in X} f(x); \tag{1}$$

here $X \subset X_0 \subset \mathbb{E}^n$, $\mathbb{E}^i$ is the i-dimensional Euclidean space, and the function $f(x)$ is defined on X_0 and takes on real values. Denote by X_* the set of all solutions of problem (1). In what follows we always assume that X_* is nonempty.

Consider a function $H(x,y) = f(x) + \xi(x,y)$, where $y \in \mathbb{E}^m$ and $\xi(x,y) : X_0 \times \mathbb{E}^m \to \mathbb{E}^1$, and define point-set mappings

$$X(y) = \operatorname*{Argmin}_{x\in X_0} H(x,y); \qquad Y(x) = \{y \in \mathbb{E}^m : \xi(z,y) \le \xi(x,y),\ \ \forall\, z \in X\}.$$

As a rule, the set X_0 is convex and has a relatively simple structure; therefore, the problem of finding points from $X(y)$ is easier than problem (1). In particular, X_0 may be the space $\mathbb{E}^n$. In many cases H may be chosen in such a way that for some $y \in \mathbb{E}^m$

$$X(y) = X_*. \tag{2}$$

As examples let us take

$$H(x,y) = |f(x) - y|^p + S(x), \qquad H(x,y) = (f(x) - y)_+^p + S(x), \tag{3}$$

where $S(x) : \mathbb{E}^n \to \mathbb{E}^1$, $S(x) = 0$ for all $x \in X$, $S(x) > 0$ for any $x \notin X$, $\varphi_+ = \max[0, \varphi]$ and $p > 0$. If $y = f(x_*)$, $x_* \in X_*$, then (2) holds. However, the value of $f(x_*)$ is usually unknown, which makes it difficult to use (3) for numerical calculations. Let us give other sufficient conditions for a minimum in problem (1).

Definition 1. *A point $x,y) \in \mathbb{E}^n \times \mathbb{E}^m$ will be called a* **singular point** *of the function $H(x,y)$ if $x \in X \cap X(y)$, $y \in Y(x)$.*

Theorem 1. *Let there exist a singular point (x,y) of the function $H(x,y)$. Then $x \in X_* \subset X(y)$.*

Theorem 2. *Let there exist a vector $y \in \mathbb{E}^m$ such that the set $X(y)$ is nonempty, $X(y) \subset X$, and the function $\xi(x,y)$ has a constant value for all $x \in X$. Then (2) holds.*

Theorem 3. *Let $f(x)$ be continuous on X_0, and let there exist a sequence $\{x_k, y_k\}$ such that $x_k \in X(y_k)$, $y_k \in Y(x_k)$ and $\lim_{k\to\infty} x_k = x_* \in X$. Then $x_* \in X_*$.*

Theorem 1 is a new formulation of Theorem 1.7.1 in [1]. The vector y may not be present in the function H; then, denoting

$$\bar{X} = \operatorname*{Argmin}_{x\in X_0} H(x), \qquad H(x) = f(x) + \xi(x), \qquad \xi(x) : \mathbb{E}^n \to \mathbb{E}^1,$$

we obtain the following assertion from Theorem 1.

Corollary 1. *Let $\bar{x} \in X \cap \bar{X}$, and let the function $\xi(x)$ be such that $\xi(x) \leq \xi(\bar{x})$ for all $x \in X$. Then $\bar{x} \in X_* \subset \bar{X}$.*

Consider the functions

$$R(x,y) = f(x) + \tilde{\xi}(x,y), \qquad \tilde{\xi}(x,y) = \xi(x,y) - \gamma(y), \qquad \gamma(y) = \sup_{x \in X} \xi(x,y).$$

Denote by Y the domain of $\gamma(y)$. Clearly, $Y \supset \bigcup_{x \in X_0} Y(x)$. Consider the problem of finding

$$\sup_{y \in Y} \inf_{x \in X_0} R(x,y) \tag{4}$$

and put on R the constraint

($\mathbf{A}_0$). For every $x \in X$ there exists a point $y \in Y$ such that $\tilde{\xi}(x,y) = 0$; for every $x \in X_0 \backslash X$ there exists a sequence $\{y_k\}$ such that $y_1, y_2, \ldots \in Y$ and $\lim_{k \to \infty} \tilde{\xi}(x, y_k) = \infty$.

Theorem 4. *For a point (x,y) to be a singular point of $R(x,y)$ it is necessary, and if condition* **($\mathbf{A}_0$)** *holds also sufficient, that it be a saddle point in problem* (4).

$\boxed{\mathbf{2.}}$ Consider reducing the original nonlinear programming problem to the problem of finding points (x,y) that satisfy the conditions

$$G(x,y) = 0, \qquad x \in X(y); \tag{5}$$

here $G(x,y) : X_0 \times \mathbb{E}^m \to \mathbb{E}^t$. We recall a definition from [2].

Definition 2. *A pair $\{H, G\}$ is* **consistent with problem** (1) *if the solution set of* (5) *is nonempty and any point (x,y) satisfying* (5) *is such that $x \in X_*$.*

Put the following constraint on the functions H and G.

($\mathbf{A}_1$). If there exist x and y satisfying the equation $G(x,y) = 0$, then $x \in X$ and $y \in Y(x)$.

Theorem 5. *For every nonlinear programming problem* (1) *with a nonempty solution set there exists a pair that is consistent with it. If for a pair $\{H, G\}$ condition* **($\mathbf{A}_1$)** *holds and the solution set of* (5) *is nonempty, then this pair is consistent with problem* (1).

Consider an auxiliary majorizing function $\eta(y)$ that is defined on some set $Y_H \subset Y$ and satisfies on it the inequality $\gamma(y) \leq \eta(y)$. Define a mapping

$$W(x) = \{y \in Y_H : \eta(y) \leq \xi(x,y)\}.$$

Clearly, $W(x) \subset Y(x)$ for any $x \in X_0$. Then in the formulation of Theorem 5 condition **($\mathbf{A}_1$)** can be replaced by

($\mathbf{A}_2$). If the point (x,y) is such that $G(x,y) = 0$, then $x \in X$ and $y \in W(x)$.

With additional assumptions about the function ξ condition **($\mathbf{A}_2$)** can be weakened. Let, for instance, $\xi(x,y) = \eta(y)$ for all $(x,y) \in X \times Y_H$; then instead of **($\mathbf{A}_2$)** one may use the condition

($\mathbf{A}_3$). If the point (x,y) is such that $G(x,y) = 0$, then $x \in X$ and $y \in Y_H$.

If, in addition, $X(y) \subset X$ for any $y \in Y_H$, then **($\mathbf{A}_2$)** is replaced by

($\mathbf{A}_4$). If the point (x,y) is such that $G(x,y)=0$, then $y\in Y_H$.

$\boxed{\mathbf{3.}}$ Let us give some examples of consistent pairs. We first specify the set X. Let

$$X=\{x\in\mathbb{E}^n: g(x)=0,\ \ h(x)\le 0\};\tag{6}$$

here $g(x):\mathbb{E}^n\to\mathbb{E}^\ell$ and $h(x):\mathbb{E}^n\to\mathbb{E}^c$. Denote by $\mathbb{E}^i_+$ the nonnegative orthant of the space $\mathbb{E}^i$. Define the set $Y_L=\mathbb{E}^\ell\times\mathbb{E}^c_+$. Let us take for H the Lagrange function

$$L(x,y)=f(x)+\xi(x,y),\qquad \xi(x,y)=\sum_{i=1}^{\ell}u^ig^i(x)+\sum_{j=1}^{c}v^jh^j(x),\qquad y=(u,v)\in\mathbb{E}^{\ell+c}.$$

For any problem (1) with admissible set (6) we have $Y_L\subset Y$ and $\gamma(y)\le 0$ on Y_L, i.e. the function $\gamma(y)\equiv 0$ majorizes $\gamma(y)$ on Y_L. Construct the function G in the form

$$G(x,y)=\{\alpha(g^1(x),u^1),\dots,\alpha(g^\ell(x),u^\ell),\quad \beta(h^1(x),v^1),\dots,\beta(h^c(x),v^c)\}.\tag{7}$$

Put the following conditions on $\alpha(a,b)$ and $\beta(a,b)$ that map $\mathbb{E}^2$ into $\mathbb{E}^1$:

($\mathbf{B}_1$). If the equation $\alpha(a,b)=0$ has a solution, then $a=0$.

($\mathbf{B}_2$). If the equation $\beta(a,b)=0$ has a solution, then $a\le 0$, $b\ge 0$ and $ab=0$.

It is easy to construct functions $\alpha(a,b)$ satisfying $(\mathbf{B}_1)$: a, $\sin a$, $\arctan a$, etc. Some examples of functions $\beta(a,b)$ satisfying $(\mathbf{B}_2)$ are $(a+b)_+-b$, $ab+a_+^2+b_-^2$, $ab+(a-b)_+^2$, $b\arctan a+a_+^2+b_-^2$, $a\arctan b+a_+^2+b_-^2$, $(a+b)_+^3-b^3-(a_-^2b)/(1+a^2)$.

If the function G is determined from (7) and conditions $(\mathbf{B}_1)$ and $(\mathbf{B}_2)$ are satisfied, then $(\mathbf{A}_2)$ holds. If, in addition, in problem (1) there exists a saddle point of the Lagrange function, then the pair $\{L,G\}$ is consistent with problem (1).

Now take for H the function

$$L_2(x,y)=f(x)+\sum_{i=1}^{\ell}u^ig^i(x)+\sum_{j=1}^{c}(v^j)^2h^j(x).$$

For any problem (1), $Y=\mathbb{E}^{\ell+c}$ and $\gamma(y)\le 0$; therefore, $\eta(y)\equiv 0$ majorizes $\gamma(y)$ on $\mathbb{E}^{\ell+c}$. The function G will as before be constructed in the form (7). Impose on $\beta(a,b)$ the condition

($\mathbf{B}_3$). If the equation $\beta(a,b)=0$ has a solution, then $a\le 0$ and $ab=0$.

Clearly, if for $G(x,y)$ conditions $(\mathbf{B}_1)$ and $(\mathbf{B}_3)$ hold, then $(\mathbf{A}_2)$ holds. On the other hand, if in problem (1) there exists a saddle point of the function L_2, then the pair $\{L_2,G\}$ is consistent with this problem. Examples of functions $\beta(a,b)$ that satisfy $(\mathbf{B}_3)$ are $ae^{-b}-a_-$, $a_+^p+ab^2$ and $a_+^p+a_-^pb$; here $p>0$ is a natural number. Note that $Y_L\subset\mathbb{E}^{\ell+c}$; therefore, when constructing a consistent pair $\{L_2,G\}$ one can use functions β that satisfy $(\mathbf{B}_2)$.

$\boxed{\mathbf{4.}}$ One obtains a wide class of consistent pairs if one takes for H

$$H(x,y)=f(x)+\sum_{i=1}^{\ell}\varphi(g^i(x),u^i)+\sum_{j=1}^{c}\psi(h^j(x),v^j),\tag{8}$$

and G is determined, as before, from (7). Here $\varphi(a,b):\mathbb{E}^2\to\mathbb{E}^1$ and $\varphi(a,b):\mathbb{E}^2\to\mathbb{E}^1$. In general it makes sense to choose φ and ψ in such a way that it becomes possible to choose a sufficiently representative subset $Y_H\subset Y$ and a majorant $\eta(y)$ for all problems (1) with an admissible set in the form (6), and afterwards to determine the mapping $W(x)$ and to construct functions α and β such that for G one of the conditions $(\mathbf{A}_2)$ – $(\mathbf{A}_4)$ is satisfied (or directly condition $(\mathbf{A}_1)$). Let us make use of the functions α and β constructed in § 3. For this we put conditions on $\psi(a,b)$. Denote $\delta(b) = \sup\limits_{a\le 0} \psi(a,b)$.

$(\mathbf{C}_1)$. $\delta(b)$ takes finite values on $\mathbb{E}^1_+$, and if $a \le 0$ $b \ge 0$ and $ab = 0$, then $\psi(a,b) = \delta(b)$.

$(\mathbf{C}_2)$. $\delta(b)$ is defined everywhere on $\mathbb{E}^1$, and if $a \le 0$ and $ab = 0$, then $\psi(a,b) = \delta(b)$.

Let $H(x,y)$ be defined by (8), and let (a,b) satisfy condition $(\mathbf{C}_1)$ (or $(\mathbf{C}_2)$). Then the function

$$\eta(y) = \sum_{i=1}^{\ell} \varphi(0,u^i) + \sum_{j=1}^{c} \delta(v^j)$$

majorizes $\gamma(y)$ on $Y_L(\mathbb{E}^{\ell+c})$ for any problem (1). If, in addition, G is determined from (7) and for α and β conditions $(\mathbf{B}_1)$ and $(\mathbf{B}_2)$ (or $(\mathbf{B}_1)$ and $(\mathbf{B}_3)$) are satisfied, then the pair $\{H,G\}$ satisfies condition $(\mathbf{A}_2)$. By Theorem 5, this pair is consistent with problem (1), provided the solution set of (5) is nonempty.

Examples of $\psi(a,b)$ for which condition $(\mathbf{C}_1)$ holds are $r(a+b)^p_+$, $ab + ra^p_+$ and $be^a + ra^p_+$. The following functions $\psi(a,b)$ satisfy condition $(\mathbf{C}_2)$: $ab^2 + ra^p_+$ and $a_+^{2p-1}b + 1/2a_-^{2p-1}b^2$. Here $p > 0$ is a natural number and $r > 0$ is a real number; p and r are chosen so as to guarantee a solution of the minimization problem for $H(x,y)$ on X_0, if this is possible. Examples of functions $\varphi(a,b)$ are $ab + ra^p$, $ab + r\cosh a$ and $a(b-1) + re^a$.

5. Let us remark in conclusion that the transition from problem (1) to problem (5) makes it possible, in solving the original nonlinear programming problem, to use a variety of numerical methods for solving systems of nonlinear equations coupled with methods for finding a minimum on sets of simple structure. For example, if the dimensions of G and y are the same then for solving (5) one may use the Jacobi iteration

$$y_{k+1} = y_k + \alpha G(x_k, y_k), \qquad x_k \in X(y_k),$$

where α is the coefficient that guarantees the convergence of the process. At the same time, the known sufficient conditions for the convergence of the Jacobi iteration method will be reformulated in terms of the functions H and G. One may proceed similarly when using other methods for solving systems of equations.

REFERENCES

1. Yu.G. Evtushenko, *Methods for solving extremal problems and their application in optimization systems*, "Nauka", Moscow, 1982; English transl., Optimization Software, New York (distributed by Springer-Verlag), 1984.

2. Yu.G. Evtushenko and V.G. Zhadan, Izv. Akad. Nauk SSSR Tekhn. Kibernet. 1983, no. 1, 47–59; English transl. in Engrg. Cybernetics **21** (1983).