Aplikace matematiky

Jan Palata

About the relation between some optimality conditions

Aplikace matematiky, Vol. 29 (1984), No. 3, 189–193

Persistent URL: http://dml.cz/dmlcz/104084

ABOUT THE RELATION BETWEEN SOME OPTIMALITY CONDITIONS

JAN PALATA

(Received February 11, 1983)

Let us consider a nonlinear programming problem

(1) $$\min_{\mathbf{x}\in N} F(\mathbf{x}),$$

where

$$N = \{\mathbf{x} \in E_n \mid G_i(\mathbf{x}) \leqq 0,\ i = 1, \ldots, m\}$$

and let us assume the functions $F(\mathbf{x})$, $G_i(\mathbf{x})$, $i = 1, \ldots, m$ to be continuously differentiable with a nonzero gradient at a point ${}_0\mathbf{x}$. In [1] and [2] we have established optimality conditions for the point ${}_0\mathbf{x}$ to be a local optimal solution of the problem under consideration, in terms of contact cone approximations. We have also showed that it is convenient to formulate our problem of finding a local extremum in another manner:

Find a point ${}_0\mathbf{x}$ such that the sets

$$M = \{\mathbf{x} \in E_n \mid F(\mathbf{x}) \leqq 0\}$$

and N are locally disjoint at that point.

If we put $I_1 = \{i \in \{1, \ldots, m\} \mid G_i({}_0\mathbf{x}) = 0\}$, $N_i = \{\mathbf{x} \in E_n \mid G_i(\mathbf{x}) \leqq 0\}$, $i = 1, \ldots, m$ and denote by $S({}_0\mathbf{x}; A)$ the contact cone of a set A at a point ${}_0\mathbf{x} \in \bar{A}$, we have

Theorem 1. (See [1].) *Let $M \cap N \neq \emptyset$. Then the following implication is valid: ${}_0\mathbf{x} \in M \cap N$, $\operatorname{int} S({}_0\mathbf{x}; M) \cap S({}_0\mathbf{x}, N) \neq \emptyset \Rightarrow U({}_0\mathbf{x}) \cap \operatorname{int} M \cap N \neq \emptyset$ for any neighbourhood $U({}_0\mathbf{x})$ of the point ${}_0\mathbf{x}$.*

Theorem 2. (See [2].) *Assume that $\bigcap_{i\in I_1} S({}_0\mathbf{x}; N_i) \cap S({}_0\mathbf{x}; M) = \{{}_0\mathbf{x}\}$. Then there exists a neighbourhood $U({}_0\mathbf{x})$ with the property $U({}_0\mathbf{x}) \cap M \cap N = \{{}_0\mathbf{x}\}$.*

As simple examples show, the necessary condition given in Theorem 1 is not sufficient generally. Nevertheless, the assumptions of the following theorem already guarantee its sufficiency.

Theorem 3. *In the problem* (1) *let, moreover, the function* $F(\mathbf{x})$ *be pseudo-convex and the functions* $G_i(\mathbf{x})$, $i \in I_1$ *quasi-convex in* E_n. *Then there exists a neighbourhood* $U({}_0\mathbf{x})$ *of the point* ${}_0\mathbf{x}$ *such that* $U({}_0\mathbf{x}) \cap \operatorname{int} M \cap N = \emptyset$ *if and only if* $\operatorname{int} S({}_0\mathbf{x}; M) \cap \cap S({}_0\mathbf{x}; N) = \emptyset$.

Proof. We only have to prove sufficiency of the condition. Suppose that for every neighbourhood $U({}_0\mathbf{x})$ there exists a point $\mathbf{x} \in U({}_0\mathbf{x}) \cap \operatorname{int} M \cap N$, i.e. a point $\mathbf{x} \in N$ with the property $F(\mathbf{x}) < F({}_0\mathbf{x})$. With respect to the pseudo-convexity of $F(\mathbf{x})$ we have

$$\sum_{\alpha=1}^{n} \frac{\partial F}{\partial x^\alpha}({}_0\mathbf{x})(x^\alpha - {}_0x^\alpha) < 0$$

and therefore $\mathbf{x} \in \operatorname{int} S({}_0\mathbf{x}; M)$. However, the quasi-convexity of the functions $G_i(\mathbf{x})$, $i \in I_1$ implies $N \subset S({}_0\mathbf{x}; N)$. Thus $\operatorname{int} S({}_0\mathbf{x}; M) \cap S({}_0\mathbf{x}; N) \neq \emptyset$ and this contradicts our condition.

In view of the fact that any local minimum of pseudo-convex function is also the global one, we have a global minimum criterion here.

Now we shall describe how our theory is connected with the known results in the special case just mentioned (see the assumptions of Theorem 3). As usual, we set $N = \{\mathbf{x} \in E_n \mid G_i(\mathbf{x}) \leqq 0,\ i = 1, \ldots, m;\ x^\alpha \geqq 0,\ \alpha = 1, \ldots, n\}$.

Theorem 4. *Let the Slater condition hold. Then the necessary and sufficient condition for the existence of a local minimum from the preceding theorem is equivalent to the Kuhn-Tucker conditions: Defining a function* $\Phi(\mathbf{x}, \mathbf{u}) = F(\mathbf{x}) + \sum_{i=1}^{m} u_i G_i(\mathbf{x})$, *there exists a point* ${}_0\mathbf{u} = ({}_0u^1, \ldots, {}_0u^m)$ *such that*

(2) 1) $\dfrac{\partial \Phi}{\partial x^\alpha}({}_0\mathbf{x}, {}_0\mathbf{u}) \geqq 0\,, \quad \alpha = 1, \ldots, n\,,$

2) $\dfrac{\partial \Phi}{\partial x^\alpha}({}_0\mathbf{x}, {}_0\mathbf{u})\, {}_0x^\alpha = 0\,, \quad \alpha = 1, \ldots, n\,,$

3) ${}_0x^\alpha \geqq 0\,, \quad \alpha = 1, \ldots, n\,,$

4) $\dfrac{\partial \Phi}{\partial u^i}({}_0\mathbf{x}, {}_0\mathbf{u}) \leqq 0\,, \quad i = 1, \ldots, m\,,$

5) $\dfrac{\partial \Phi}{\partial u^i}({}_0\mathbf{x}, {}_0\mathbf{u})\, {}_0u^i = 0\,, \quad i = 1, \ldots, m\,,$

6) ${}_0u^i \geqq 0\,, \quad i = 1, \ldots, m\,.$

Proof. Introduce an index set

$$I_2 = \{\alpha \in \{1, \ldots, n\} \mid {}_0x^\alpha = 0\}$$

and a set

$$N^*_\beta = \{\mathbf{x} \in E_n \mid x^\beta \geqq 0\}\,, \quad \beta \in I_2\,.$$

First we show that

$$S({}_0\mathbf{x}; N) = \bigcap_{i \in I_1} S({}_0\mathbf{x}; N_i) \cap \bigcap_{\beta \in I_2} S({}_0\mathbf{x}; N^*_\beta)\,. \tag{3}$$

Choosing a point $\mathbf{x}^* \geqq 0$ which satisfies $G_i(\mathbf{x}^*) < 0$, $i = 1, \ldots, m$ (the Slater condition), we distinguish two possibilities:

a) ${}_0\mathbf{x} = \mathbf{x}^*$. Then $S({}_0\mathbf{x}; N_i) = E_n$, $i \in I_1$ and in virtue of Theorem 4 in [1] $S({}_0\mathbf{x}; N) = S({}_0\mathbf{x}; \bigcap_{\beta \in I_2} N^*_\beta) = \bigcap_{\beta \in I_2} S({}_0\mathbf{x}; N^*_\beta)$, seeing that the system of contact cones $S({}_0\mathbf{x}, N^*_\beta)$, $\beta \in I_2$ is not separable in E_n.

b) ${}_0\mathbf{x} \neq \mathbf{x}^*$. From the quasi-convexity of the functions $G_i(\mathbf{x})$, $i \in I_1$ it follows that a half-open line segment $({}_0\mathbf{x}, \mathbf{x}^*\rangle$ belongs to the set $\bigcap_{i \in I_1} \operatorname{int} N_i \cap \bigcap_{\beta \in I_2} N^*_\beta$. Thus $\bigcap_{i \in I_1} \operatorname{int} N_i \cap \bigcap_{\beta \in I_2} \operatorname{int} N^*_\beta \neq \emptyset$ and Theorem 4 in [1] implies (3) again, for the relations

$$\bigcap_{\beta \in I_2} \operatorname{int} N^*_\beta \subset \bigcap_{\beta \in I_2} \operatorname{int} S({}_0\mathbf{x}; N^*_\beta)\,,$$

$$\bigcap_{i \in I_1} \operatorname{int} N_i \subset \bigcap_{i \in I_1} \operatorname{int} S({}_0\mathbf{x}; N_i)$$

obviously hold in this case.

According to the Farkas lemma the condition $\operatorname{int} S({}_0\mathbf{x}; M) \cap S({}_0\mathbf{x}; N) = \emptyset$ is equivalent to the existence of numbers ${}_0\bar{u}^i \geqq 0$ $(i \in I_1)$ ${}_0v^\beta \geqq 0 (\beta \in I_2)$ with

$$-\frac{\partial F}{\partial x^\alpha}({}_0\mathbf{x}) = \sum_{i \in I_1} \frac{\partial G_i}{\partial x^\alpha}({}_0\mathbf{x})\, {}_0\bar{u}^i - \sum_{\beta \in I_2} \delta_{\alpha\beta}\, {}_0v^\beta \quad (\alpha = 1, \ldots, n)\,. \tag{4}$$

We claim that (4) is true if and only if the Kuhn-Tucker conditions (2) are fulfilled (with ${}_0\mathbf{u} = ({}_0u^1, \ldots, {}_0u^m)$ where ${}_0u^i = {}_0\bar{u}^i\, (i \in I_1)$, ${}_0u^i = 0$ $(i \notin I_1)$). This is now to be verified.

a) Using the relation (4) we obtain

1) $$\frac{\partial \Phi}{\partial x^\alpha}({}_0\mathbf{x}, \mathbf{u}_0) = \frac{\partial F}{\partial x^\alpha}({}_0\mathbf{x}) + \sum_{i=1}^{m} {}_0u^i \frac{\partial G_i}{\partial x^\alpha}({}_0\mathbf{x}) = \frac{\partial F}{\partial x^\alpha}({}_0\mathbf{x}) + \sum_{i \in I_1} {}_0\bar{u}^i \frac{\partial G_i}{\partial x^\alpha}({}_0\mathbf{x}) =$$

$$\sum_{\beta \in I_2} \delta_{\alpha\beta}\, {}_0v^\beta \geqq 0 \quad (\alpha = 1, \ldots, n)\,,$$

2) $$\frac{\partial \Phi}{\partial x^\alpha}({}_0\mathbf{x}, {}_0\mathbf{u})\, {}_0x^\alpha = \frac{\partial F}{\partial x^\alpha}({}_0\mathbf{x})\, {}_0x^\alpha + \sum_{i=1}^{m} {}_0u^i \frac{\partial G_i}{\partial x_\alpha}({}_0\mathbf{x})\, {}_0x^\alpha = \sum_{\beta \in I_2} \delta_{\alpha\beta_0}\, {}_0v^\beta x^\alpha -$$

$$-\sum_{i\in I_1} {}_0u^i \frac{\partial G_i}{\partial x^\alpha}({}_0\mathbf{x})\, {}_0x^\alpha + \sum_{i=1}^{m} {}_0u^i \frac{\partial G_i}{\partial x^\alpha}({}_0\mathbf{x})\, {}_0x^\alpha = \sum_{\beta\in I_2} \delta_{\alpha\beta}\, {}_0v^\beta\, {}_0x^\alpha = 0 \quad (\alpha = 1, \ldots, n),$$

3) ${}_0x^\alpha \geqq 0 \quad (\alpha = 1, \ldots, n)$.

4) $\dfrac{\partial \Phi}{\partial u^i}({}_0\mathbf{x}, {}_0\mathbf{u}) = G_i({}_0\mathbf{x}) \leqq 0 \quad (i = 1, \ldots, m)$,

5) $\dfrac{\partial \Phi}{\partial u^i}({}_0\mathbf{x}, {}_0\mathbf{u})\, {}_0u^i = G_i({}_0\mathbf{x})\, {}_0u^i = 0 \quad (\mathrm{i} = 1, \ldots, m)$, because

$G_i({}_0\mathbf{x}) = 0$ for $i \in I_1$ and ${}_0u^i = 0$ for $i \in \{1, \ldots, m\} \setminus I_1$,

6) ${}_0u^i \geqq 0 \quad (i = 1, \ldots, m)$.

This means that the Kuhn-Tucker conditions are fulfilled.

b) Let the Kuhn-Tucker conditions hold. We establish the existence of numbers ${}_0v^\alpha \geqq 0$ $(\alpha = 1, \ldots, n)$ ensuring the validity of (4). Setting

$${}_0v^\alpha = \frac{\partial \Phi}{\partial x^\alpha}({}_0\mathbf{x}, {}_0\mathbf{u}) \quad (\alpha = 1, \ldots, n),$$

we have

$${}_0v^\alpha = \frac{\partial F}{\partial x^\alpha}({}_0\mathbf{x}) + \sum_{i=1}^{m} {}_0u^i \frac{\partial G_i}{\partial x^\alpha}({}_0\mathbf{x}) \quad (\alpha = 1, \ldots, n),$$

which yields

$$(5) \qquad -\frac{\partial F}{\partial x^\alpha}({}_0\mathbf{x}) = \sum_{i=1}^{m} {}_0u^i \frac{\partial G_i}{\partial x^\alpha}({}_0\mathbf{x}) - \sum_{\beta=1}^{n} \delta_{\alpha\beta}\, {}_0v^\beta \quad (\alpha = 1, \ldots, n).$$

Nonnegativity of ${}_0v^\alpha$ $(\alpha = 1, \ldots, n)$ is caused by 1) in (2). We know that $G_i({}_0\mathbf{x}) < 0$ for $i \in \{1, .., m\} \setminus I_1$ and thus by 5) in (2) we conclude ${}_0u^i = 0$ $(i \in \{1, \ldots, m\} \setminus I_1)$. For $\alpha \in \{1, \ldots, n\} \setminus I_2$ we have ${}_0x^\alpha > 0$. Therefore

$$\frac{\partial \Phi}{\partial x^\alpha}({}_0\mathbf{x}, {}_0\mathbf{u})\, {}_0x^\alpha = {}_0v^\alpha\, {}_0x^\alpha = 0 \quad (\alpha = 1, \ldots, n)$$

(see 2) in (2)) leads to ${}_0v^\alpha = 0$ $(\alpha \in \{1, \ldots, n\} \setminus I_2)$. In this way (5) is reduced to (4) with nonnegative multipliers ${}_0u^i (i \in I_1)$, ${}_0v^\beta (\beta \in I_2)$ q.e.d.

Let us go back to our original assumptions concerning the functions $F(\mathbf{x})$, $G_i(\mathbf{x})$ $(i = 1, \ldots, m)$ again and let us ask what can be said about the multipliers u^i in connection with the necessary condition (Theorem 1) and with the sufficient condition (Theorem 2). The set N is supposed to be defined as in (1). We shall assume the validity of the relation $S({}_0\mathbf{x}; N) = \bigcap_{i\in I_1} S({}_0\mathbf{x}, N_i)$ (holding e.g. when the system of contact cones $S({}_0\mathbf{x}; N_i)$, $i \in I_1$ is not separable in E_n).

From the Farkas lemma it follows that the necessary condition for a local extremum contained in Theorem 1 yields the existence of nonnegative multipliers u^i which fulfil

$$\nabla F({}_0\mathbf{x}) + \sum_{i \in I_1} u^i \nabla G_i({}_0\mathbf{x}) = 0 , \tag{6}$$

where $\sum_{i \in I_1} u^i > 0$ (provided $\nabla F({}_0\mathbf{x}) \neq 0$).

By the sufficient condition stronger requirements are imposed on the position of the contact cones $S({}_0\mathbf{x}; M)$ and $S({}_0\mathbf{x}, N)$, so that it could seem that we shall also get some additional information on the possible positivity of the multipliers u^i. However, this is impossible if we admit linear dependence of the gradients $\nabla G_i({}_0\mathbf{x})$, $i \in I_1$. In this case each coefficient in the relation (6) can vanish after adding a convenient zero linear combination $\sum_{i \in I_1} t^i \nabla G_i({}_0\mathbf{x})$. If the gradients $\nabla G_i({}_0\mathbf{x})$, $i \in I_1$ are linearly independent in E_n, we easily get the result as follows:

There exist $u^i > 0$, $i \in I$, with $\nabla F({}_0\mathbf{x}) + \sum_{i \in I_1} u^i \nabla G_i({}_0\mathbf{x}) = 0$.

References

[1] *Jan Palata:* First-order necessary condition for the existence of optimal point in nonlinear programming problem. Aplikace matematiky 25 (1980), 257—266.

[2] *Jan Palata:* Eine hinreichende Bedingung für die Existenz eines eindeutigen lokalen Extremums. Math. Operationsforsch. u. Statist. 11 (1980), No. 4, 531—536.

Souhrn

SOUVISLOST NĚKTERÝCH PODMÍNEK OPTIMALITY

JAN PALATA

V článku je ukázán vztah mezi obecnými podmínkami optimality odvozenými pomocí aproximací styčnými kuželi a známými Kuhnovými-Tuckerovými podmínkami ve speciálním případě pseudokonvexních a kvazikonvenčních funkcí i jejich důsledek pro Lagrangeovy multiplikátory.

Author's address: RNDr. *Jan Palata*, CSc., matematicko-fyzikální fakulta UK, Malostranské nám. 25, 118 00 Praha 1.