ISSN 1064-5624, Doklady Mathematics, 2016, Vol. 93, No. 2, pp. 166–169. © Pleiades Publishing, Ltd., 2016.
Original Russian Text © O.A. Brezhneva, A.A. Tret'yakov, 2016, published in Doklady Akademii Nauk, 2016, Vol. 467, No. 3, pp. 255–256.

MATHEMATICS

New Approach to Optimality Conditions for Degenerate Nonlinear Programming Problems[1]

O. A. Brezhneva[a] and A. A. Tret'yakov[b, c, d]

Presented by Academician Yu.G. Evtushenko July 6, 2015

Received July 8, 2015

Abstract—In the work, a new approach to constructing optimality conditions for degenerate smooth optimization problems with inequality constraints is proposed. The approach is based on the theory of p-regularity. A special case of degeneracy, when the first derivatives of some function-constraints are equal to zero up to some order, is considered. Optimality conditions for the general case of degeneracy with $p = 2$ are presented. Proposed constructions and optimality conditions are illustrated by an example. A general case of degeneracy is considered and optimality conditions for the case of $p \geq 2$ are proposed.

DOI: 10.1134/S1064562416020125

In this work, we propose a new approach to developing optimality conditions for degenerate smooth optimization problems with inequality constraints of the following form:

$$\begin{gathered}\underset{x \in S}{\text{minimize}} f(x),\\ S(x) = \{x \in \mathbb{R}^n \mid g(x) \leq 0\},\end{gathered} \tag{1}$$

where $g(x) = (g_1(x), \dots, g_m(x))$ and functions f, g_i: $\mathbb{R}^n \to \mathbb{R}$ are assumed to be sufficiently smooth. The approach is based on the theory of p-regularity, presented, for example, in [1–4].

1. REDUCTION OF THE GENERAL DEGENERATE CASE TO A SPECIAL ONE

Without loss of generality consider the case of $p = 2$, because it is most illustrative and demonstrates the main idea of the proposed approach. In this and the following parts of the paper, to simplify the notation, we assume that all constraints are active at the solution x^*, that is the set of indices of active constraints can be written as $I(x^*) = \{1, 2, ..., m\}$. Indeed, in the case of presence of inactive constraints, all results will hold with zero Lagrange multipliers corresponding to the inactive constraints.

When constructing optimality conditions for problem (1), it is convenient to assume that there is a number r such that

$$\begin{gathered} g_i'(x^*) \neq 0, \quad i = 1, 2, \dots, r,\\ g_i'(x^*) = 0, \quad i = r+1, r+2, \dots, m.\end{gathered} \tag{2}$$

At the same time, one of the main ideas of the proposed in the paper approach is to reduce degenerate problems to a new form for which (2) is satisfied.

Introduce an additional set:

$$H_g(x^*) = \{h \in \mathbb{R}^n \mid \langle g_i'(x^*), h\rangle \leq 0, i \in I(x^*)\}.$$

Fix some element $h \in H_g(x^*)$ and define a set of indices:

$$I_1(x^*, h) = \{i \in I(x^*) \mid \langle g_i'(x^*), h\rangle = 0\}.$$

We will describe a procedure for constructing special acute cones generated by the vectors $g_i'(x^*)$, $i \in I_1(x^*, h)$, under an assumption that $|I_1(x^*, h)| = m_1 \neq 0$. These special cones are required below for constructing optimality conditions.

We will require that, while constructing cones, each index contained in the set $I_1(x^*, h)$ was used for determining at least one cone, and all derived acute cones were different, so that their total number is $s \leq m_1$. Namely, for constructing the cone number k, where $k = 1, 2, ..., s$, indices $i_1, \dots, i_{r_k} \in I_1(x^*, h)$ are used and are selected in such a way that vectors $g_{i_1}'(x^*), \dots, g_{i_{r_k}}'(x^*)$ generate the largest possible acute cone and $i_j \neq i_l$ if $j \neq l$. As a result, there exists an element $\gamma_i \in \mathbb{R}^n$

[1] The article was translated by the authors.

[a] Department of Mathematics, Miami University, Oxford, USA
[b] Dorodnicyn Computing Center, Federal Research Center "Computer Science and Control", Russian Academy of Sciences, ul. Vavilova 40, Moscow, 119333 Russia
[c] System Research Institute, Polish Academy of Sciences, Warsaw
[d] University of Podlasie, Siedlce, Poland
e-mail: brezhnoa@miamioh.edu, tret@uph.edu.pl

such that $\langle g_j'(x^*), \gamma_i\rangle < 0, j = i_1, \dots, i_{r_k}$, and for all indices $j \in J_k(x^*, h) = I_1(x^*, h)\backslash\{i_1, \dots, i_{r_k}\}$, the following is satisfied:

$$-g_j'(x^*) = \alpha_{ji_1}g_{i_1}'(x^*) + \dots + \alpha_{ji_{r_k}}g_{i_{r_k}}'(x^*),$$

where $\alpha_{ji_1} \geq 0, \dots, \alpha_{ji_{r_k}} \geq 0$.

For each index $j \in J_k(x^*, h)$, introduce new mappings

$$\tilde{g}_j(x) = g_j(x) + \alpha_{ji_1}g_{i_1}(x) + \dots + \alpha_{ji_{r_k}}g_{i_{r_k}}(x)$$

and, therefore, get s sets F_k, where $k = 1, 2, \dots, s$, and each set F_k consists of functions $g_{i_1}(x), \dots, g_{i_{r_k}}(x)$, $\tilde{g}_j(x), j \in J_k(x^*, h)$, and satisfies the conditions

$$\tilde{g}_j'(x^*) = 0, \quad j \in J_k(x^*, h). \tag{3}$$

In addition to the set S introduced in (1), for each $k = 1, 2, \dots, s$ define

$$S_k = \Big\{x \in \mathbb{R}^n \,\Big|\, g_{i_1}(x) \leq 0, \dots, g_{i_{r_k}}(x) \leq 0,$$

$$\tilde{g}_j(x) \leq 0, j \in J_j(x^*, h),$$

$$g_j(x) \leq 0, j \in I(x^*)\backslash I_1(x^*, h)\Big\}.$$

Lemma 1. *There exist sets of functions F_k, $k = 1$, $2, \dots, s$, defined above and such that* $S = \bigcap_{k=1}^{s} S_k$.

Lemma 1 implies that problem (1) can be written in the form

$$\text{minimize} f(x), \quad x \in \bigcap_{k=1}^{s} S_k.$$

Therefore, we need to write optimality conditions for each set S_k taking into account implementation of the conditions (3) for it, and then take the union of these optimality conditions for all values of the index k. Optimality conditions constructed in this way will give optimality conditions for the original problem (1) under some regularity assumption for the set $S = \bigcap_{k=1}^{s} S_k$. Notice that conditions (3) are similar to conditions (2) for the special case of problem (1). We will consider this special case in part 2.

Assume that there exists a vector $h \in H_g(x^*)$, satisfying the following inequalities

$$\langle \tilde{g}_j''(x^*)h, h\rangle \leq 0, \quad j \in J_k(x^*, h), \quad k = 1, 2, \dots, s.$$

Notice that if such h is not found, then x^* is an isolated feasible point for problem (1).

Recall that all indices under consideration belong to the set

$$I_1(x^*, h) = \{i \in I(x^*) | \langle g_i'(x^*), h\rangle = 0\},$$

where $h \in H_g(x^*)$, and define

$$I_0^{1k}(x^*, k) = \{i_1, \dots, i_{r_k}\},$$

$$I_0^{2k}(x^*, h) = \{i \in J_k(x^*, h) | \langle \tilde{g}_i''(x^*)h, h\rangle = 0\},$$

$$I_0^1(x^*, h) = \bigcup_{k=1}^{s} I_0^{1k}(x^*, h),$$

$$I_0^2(x^*, h) = \bigcup_{k=1}^{s} I_0^{2k}(x^*, h).$$

Definition 1. We say that mapping $g(x)$: $\mathbb{R}^n \to \mathbb{R}^m$ is 2-regular at the point $x^* \in \mathbb{R}^n$ along vector $h \in H_g(x^*)$, if there exists $\xi \in \mathbb{R}^n$, satisfying the inequalities

$$\langle g_{i_1}'(x^*), \xi\rangle < 0, \dots, \langle g_{i_{r_k}}'(x^*), \xi\rangle < 0,$$

$$\langle \tilde{g}_j''(x^*)h, \xi\rangle < 0, \quad j \in I_0^2(x^*, h),$$

$$k = 1, 2, \dots, s.$$

Notice that Definition 1 is a generalization of the Mangasarian-Fromovitz constraint qualification. However, there are examples where the definition does not hold (see Example 1 below). To analyze those problems, we introduce a more general definition.

Definition 2. We say that mapping $g(x)$: $\mathbb{R}^n \to \mathbb{R}^m$ is tangent 2-regular at the point $x^* \in \mathbb{R}^n$ along vector $h \in H_g(x^*)$, if for any $\xi \in \mathbb{R}^n$ that satisfies the inequalities

$$\langle g_{i_1}'(x^*), \xi\rangle \leq 0, \dots, \langle g_{i_{r_k}}'(x^*), \xi\rangle \leq 0,$$

$$\langle \tilde{g}_j''(x^*)h, \xi\rangle \leq 0, \quad j \in I_0^2(x^*, h),$$

$$k = 1, 2, \dots, s,$$

there exist feasible elements $x \in S$ in the form

$$x = x^* + \alpha h + \omega(\alpha)\xi + \eta(\alpha),$$

where $\alpha > 0$ is sufficiently small, $\omega(\alpha) = o(\alpha)$, $\dfrac{\alpha^2}{\omega(\alpha)} \to 0$ as $\alpha \to 0$, and $\|\eta(\alpha)\| = o(\omega(\alpha))$.

If Definition 2 does not hold, then we need to analyze cases with $p > 2$, which are considered in part 3.

Introduce a generalized 2-factor Lagrange function

$$L_2(x, \lambda(h), h) = f(x) + \sum_{i \in I_0^1(x, h)} \lambda_i(h)g_i(x)$$

$$+ \sum_{i \in I_0^2(x, h)} \lambda_i(h)\langle \tilde{g}_i'(x), h\rangle.$$

Notice again that conditions (3) are satisfied.

Theorem 1. *Let x^* be a local minimizer for problem* (1), *$f \in C^1(\mathbb{R}^n)$ and $g \in C^2(\mathbb{R}^n)$. Assume that mapping $g(x)$ is tangent 2-regular at the point x^* along vector $h \in H_g(x^*)$ and $\langle f'(x^*), h\rangle = 0$. Then there exists*

$$\lambda^*(h) = (\lambda_i^*(h))_{i \in I_0^1(x^*, h) \cup I_0^2(x^*, h)}$$

such that $\lambda^*(h) \ge 0$ *and*

$$L'_{2x}(x^*, \lambda^*(h), h) = 0. \quad (4)$$

Notice that since $I_0^1(x^*, h) \subset I_1(x^*, h)$, $I_0^2(x^*, h) \subset I_1(x^*, h)$ and, for every index $j \in J_k(x^*, h)$, the following holds

$$\tilde{g}_j(x) = g_j(x) + \alpha_{ji_1} g_{i_1}(x) + \ldots + \alpha_{ji_{r_k}} g_{i_{r_k}}(x),$$

where $\alpha_{ji_1} \ge 0, \ldots, \alpha_{ji_{r_k}} \ge 0$, then conditions (4) can be written in the following equivalent form

$$f'(x^*) + \sum_{i \in I_1(x^*, h)} \lambda_i g'_i(x^*) + \sum_{i \in I_1(x^*, h)} \gamma_i g''_i(x^*)h = 0,$$

where $\lambda_i \ge 0$ and $\gamma_i \ge 0$.

2. SPECIAL DEGENERATE CASE WITH $p = 2$

In this part of the paper, we consider a special degenerate case, when for problem (1), conditions (2) are satisfied.

Introduce set $H_2(x^*)$ that consists of all $h \in \mathbb{R}^n$ such that the following inequalities hold:

$$\langle g'_i(x^*), h\rangle \le 0, \quad i = 1, 2, \ldots, r,$$

$$\langle g''_i(x^*)h, h\rangle \le 0, \quad i = r+1, r+2, \ldots, m.$$

For some $h \in H_2(x^*)$, also introduce additional sets:

$$I_1(h) = \{i \in \{1, 2, \ldots, r\} \mid \langle g'_i(x^*), h\rangle = 0\},$$

$$I_2(h) = \{i \in \{r+1, r+2, \ldots, m\} \mid \langle g''_i(x^*)h, h\rangle = 0\}.$$

Definition 3. Mapping $g(x)$ is called 2-regular at the point $x^* \in \mathbb{R}^n$ along the vector $h \in H_2(x^*)$, if either $I_1(h) = \emptyset$ and $I_2(h) = \emptyset$, or there exists $\xi = \xi(h) \in \mathbb{R}^n$ such that

$$\begin{aligned} &\langle g'_i(x^*), \xi\rangle < 0, \quad i \in I_1(h), \\ &\langle g''_i(x^*)h, \xi\rangle < 0, \quad i \in I_2(h). \end{aligned} \quad (5)$$

Definition 4. Mapping $g(x)$ is called 2-regular at the point $x^* \in \mathbb{R}^n$, if, for every $h \in H_2(x^*)$, either $I_1(h) = \emptyset$ and $I_2(h) = \emptyset$, or there exists $\xi = \xi(h)$ such that condition (5) is satisfied.

We also introduce the 2-factor-Lagrange function:

$$L_2(x, \lambda(h), h) = f(x) + \sum_{i \in I_1(h)} \lambda_i(h) g_i(x)$$

$$+ \sum_{i \in I_2(h)} \lambda_i(h) \langle g'_i(x)h\rangle,$$

where $x, h \in \mathbb{R}^n$ and $\lambda(h) = (\lambda_i(h))_{i \in I_1(h) \cup I_2(h)}$.

For problem (1) the following optimality conditions hold.

Theorem 2. *Let* $f(x) \in C^1(\mathbb{R}^n)$ *and* $g(x) \in C^2(\mathbb{R}^n)$. *Assume that mapping* $g(x)$ *is 2-regular at the point* x^*.

1\. *If* x^* *is a local minimizer for problem* (1), *then for any* $h \in H_2(x^*)$ *either*

$$\langle f'(x^*), h\rangle > 0 \quad (6)$$

or there exists vector $\lambda^*(h) = (\lambda_i^*(h))_{i \in I_1(h) \cup I_2(h)}$ *such that*

$$L'_{2_x}(x^*, \lambda^*(h), h) = 0, \quad \lambda^*(h) \ge 0. \quad (7)$$

2\. *If for any* $h \in H_2(x^*)$ *either condition* (6) *is satisfied or there exists* $\beta > 0$ *such that* (7) *holds and*

$$L''_{2_{xx}}(x^*, \tilde{\lambda}^*(h), h)[h]^2 \ge \beta\|h\|^2,$$

where $\tilde{\lambda}_i^*(h) = \lambda_i^*(h)$ *for* $i \in I_1(h)$, *and* $\tilde{\lambda}_i^*(h) = \dfrac{\lambda_i^*(h)}{3}$ *for* $i \in I_2(h)$, *then* x^* *is an isolated local minimizer for problem* (1).

We will illustrate the introduced concepts and Theorem 1 by the following example.

Example 1. Let in problem (1),

$$f(x) = x_1 + x_2 - x_3 + x_1^2 + x_2^2 + x_3^2$$

and the constraints are defined in the following way:

$$g_1(x) = -x_1,$$

$$g_2(x) = -x_2,$$

$$g_3(x) = x_2 - x_1^2 + x_2^2 + x_3^2.$$

In this example, $x \in \mathbb{R}^3$ and $x^* = (0, 0, 0)^{\mathrm{T}}$. Consider $h = (1, 0, 1)^{\mathrm{T}}$.

Then $m = 3$, $m_1 = 2$, $s = 2$, $I_1(0, h) = \{2, 3\}$. Constructing the cone number $k = 1$, we have $r_1 = 1$, $g_{1_1}(x) = g_2(x)$, $J_1(0, h) = \{3\}$, and $\tilde{g}_3(x) = -x_1^2 + x_2^2 + x_3^2$. Constructing the cone number $k = 2$ we have $r_2 = 1$, $g_{2_1}(x) = g_3(x)$, $J_2(0, h) = \{2\}$, and $\tilde{g}_2(x) = -x_1^2 + x_2^2 + x_3^2$.

We get two systems of inequalities

$$\begin{aligned} g_1(x) &= -x_1 \le 0, \\ g_2(x) &= -x_2 \le 0, \\ \tilde{g}_3(x) &= -x_1^2 + x_2^2 + x_3^2 \le 0; \end{aligned} \quad (\mathrm{A})$$

$$\begin{aligned} g_1(x) &= -x_1 \le 0, \\ g_3(x) &= x_2 - x_1^2 + x_2^2 + x_3^2 \le 0, \\ \tilde{g}_2(x) &= -x_1^2 + x_2^2 + x_3^2 \le 0. \end{aligned} \quad (\mathrm{B})$$

Using the notation introduced above, we have $S_1 = \{x \in \mathbb{R}^n \mid (\mathrm{A})\}$, $S_2 = \{x \in \mathbb{R}^n \mid (\mathrm{B})\}$.

Notice that in this example, $\langle f'(x^*), h\rangle = 0$, $\tilde{g}''_2(x^*)[h] = \tilde{g}''_3(x^*)[h] = (-2, 0, 2)^{\mathrm{T}}$, and mapping $g(x)$ is tangent 2-regular at the point x^* along the vector h. Therefore, all conditions of Theorem 1 are satisfied and there exist $\lambda_i^*(h) \ge 0$, $i = 1, 2, 3, 4$, such that condition (4) is satisfied:

$$f'(0) + \lambda_1^*(h)g'_{1_1}(0) + \lambda_2^*(h)g'_{2_1}(0) + \lambda_3^*(h)\tilde{g}''_3(0)[h] + \lambda_4^*(h)\tilde{g}''_2(0)[h]$$

$$= \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{2} \cdot \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix} + 0 \cdot \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

3. GENERAL DEGENERATE CASE WITH $p > 2$

In this part, we consider the general degenerate case, namely, the results given in this part, will be formulated for some $p \geq 2$. Similarly to the approach described in the previous parts of the paper, the constraints of the problem (1) can be reduced to equivalent ones in such a way that, without change of the notation, the constraints satisfy the following relations:

$$g'_i(x^*) \neq 0, \quad i = 1, \ldots, r_1,$$
$$g'_i(x^*) = 0, \quad i = r_1 + 1, \ldots, m,$$
$$g''_i(x^*)h \neq 0, \quad i = r_1 + 1, \ldots, r_2,$$
$$g''_i(x^*)h = 0, \quad i = r_2 + 1, \ldots, m,$$
$$\ldots$$
$$g_i^{(p-1)}(x^*)[h]^{p-2} \neq 0, \quad i = r_{p-2} + 1, \ldots, r_{p-1},$$
$$g_i^{(p-1)}(x^*)[h]^{p-2} = 0, \quad i = r_{p-1} + 1, \ldots, m,$$
$$\langle g_i^{(p)}(x^*)[h]^{p-1}, h\rangle \leq 0, \quad i = r_p, \ldots, m.$$

Introduce the sets:

$$I_1(x^*, h) = \{i \in I(x^*) | \langle g'_i(x^*), h\rangle = 0\},$$
$$I_2(x^*, h) = \{i \in I(x^*) | g''_i(x^*)[h]^2 = 0\},$$
$$\ldots$$
$$I_p(x^*, h) = \{i \in I(x^*) | g_i^{(p)}(x^*)[h]^p = 0\}.$$

Definition 5. We say that mapping $g(x)$: $\mathbb{R}^n \to \mathbb{R}^m$ is p-regular at the point $x^* \in \mathbb{R}^n$ along the vector $h \in H_g(x^*)$, if there exists $\xi \in \mathbb{R}^n$, satisfying the following inequalities:

$$\langle g'_i(x^*), \xi\rangle < 0, \quad i \in I_1(x^*, h),$$
$$\langle g''_i(x^*)h, \xi\rangle < 0, \quad i \in I_2(x^*, h),$$
$$\ldots$$
$$\langle g_i^{(p)}(x^*)[h]^{p-1}, \xi\rangle < 0, \quad i \in I_p(x^*, h).$$

Theorem 3. *Let x^* be a local minimizer of problem* (1), *$f \in C^1(\mathbb{R}^n)$ and $g \in C^{p+1}(\mathbb{R}^n)$. Assume that mapping $g(x)$ is p-regular at the point x^* along the vector $h \in H_g(x^*)$ and $\langle f'(x^*), h\rangle = 0$.*

Then there exists

$$\lambda^*(h) = (\lambda_i^*(h))_{i \in I_1(x^*, h) \cup I_2(x^*, h) \cup \ldots \cup I_p(x^*, h)}$$

such that $\lambda^*(h) \geq 0$ and

$$f'(x^*) + \sum_{i \in I_1(x^*, h)} \lambda_i g'_i(x^*) + \sum_{i \in I_2(x^*, h)} \lambda_i g''_i(x^*)h \ldots + \sum_{i \in I_p(x^*, h)} \lambda_i g_i^{(p)}(x^*)[h]^{p-1} = 0.$$

Notice that similarly to the Definition 2 given above, one can introduce a definition of the mapping $g(x)$: $\mathbb{R}^n \to \mathbb{R}^m$, which is tangent p-regular at the point $x^* \in \mathbb{R}^n$ along the vector $h \in H_g(x^*)$ for $p > 2$, and formulate more general optimality conditions similar to ones given in Theorem 3.

ACKNOWLEDGMENTS

The work of the second author is partially financially supported by the Russian Foundation for Basic Research, grant no. 14-07-00805, by the Leading Scientific Schools, grant no. 4640.2014.1, and by the program of Presidium RAN P-8.

REFERENCES

1. A. A. Tret'yakov, USSR Comput. Math. Math. Phys. **24** (1), 123–127 (1984).
2. A. A. Tret'yakov, Russ. Math. Surv. **42** (5), 179–180 (1987).
3. A. A. Tret'yakov and J. E. Marsden, Commun. Pure Appl. Anal. **2**, 425–445 (2003).
4. O. A. Brezhneva, Yu. G. Evtushenko, and A. A. Tret'yakov, Comput. Math. Math. Phys. **46** (11), 1896–1909 (2006).