TWMS J. Pure Appl. Math. V.1, N.2, 2010, pp. 198-223

THE P-*th* ORDER OPTIMALITY CONDITIONS FOR DEGENERATE INEQUALITY CONSTRAINED OPTIMIZATION PROBLEMS

OLGA BREZHNEVA [1], ALEXEY A. TRET'YAKOV [2]

Abstract. In this paper, we present necessary and sufficient optimality conditions for optimization problems with inequality constraints in the finite dimensional spaces. We focus on the *degenerate* (*nonregular*) case when the linear independence constraint qualification (LICQ) and Mangasarian-Fromovitz constraint qualification (MFCQ) are not satisfied at the solution of the optimization problem. For the problems satisfying the p-regularity constraint qualification or p-regularity conditions, we present necessary and sufficient conditions that resemble the structure of the classical conditions and give new and nontrivial conditions for degenerate inequality constrained problems. We also present second-order necessary conditions and corresponding sufficient conditions. The optimality conditions can be applied to discretizations of calculus of variations and optimal control problems. In addition, we prove that the 2-regularity condition is weaker than the MFCQ.

Keywords: optimality conditions, constraint qualifications, degeneracy, Karush-Kuhn-Tucker conditions, Mangasarian-Fromovitz constraint qualification

AMS Subject Classification: 90C46, 90C30, 49K99, 49N60.

1. Introduction

We consider the nonlinear optimization problem with inequality constraints

$$\underset{x\in\mathbb{R}^n}{\text{minimize}}\ f(x) \quad \text{subject to} \quad g_1(x)\geq 0,\dots,g_m(x)\geq 0, \tag{1}$$

where the functions f and g_i, $i=1,\dots,m$, are sufficiently smooth, real–valued functions on $\mathbb{R}^n$.

Most optimality conditions obtained for problem (1) cover the case when a solution x^* is regular. The most common regularity conditions (or constraint qualifications) are the linear independence constraint qualification (LICQ) and Mangasarian-Fromovitz constraint qualification (MFCQ). The LICQ holds at a feasible point x^* if the active constraint gradients $(g_i'(x^*))^T$, $i\in I(x^*)=\{i=1,\dots,m \mid g_i(x^*)=0\}$, are linearly independent. The MFCQ holds at a feasible point x^* if there exists a vector ν such that $g_i'(x^*)\nu>0$ for all $i\in I(x^*)$. The MFCQ is a weaker condition than the LICQ in the sense that satisfaction of the LICQ implies the MFCQ, but not the reverse.

The classical optimality conditions are stated in terms of the *Lagrange function* defined by

$$L(x,\lambda)=f(x)-\sum_{i=1}^{m}\lambda_i g_i(x). \tag{2}$$

Namely, the Karush-Kuhn-Tucker conditions state that if x^* is a local minimum of problem (1), and the LICQ or MFCQ holds at x^*, then there exists a Lagrange multiplier vector $\lambda^*=$

[1] Department of Mathematics, Miami University,
e-mail: brezhnoa@muohio.edu
[2] Dorodnicyn Computing Center of the Russian Academy of Sciences, University of Podlasie in Siedlce, Poland,
e-mail: tret@ap.siedlce.pl
Manuscript received 11 June 2010.

$(\lambda_1^*, \ldots, \lambda_m^*)$ such that

$$\frac{\partial L}{\partial x}(x^*, \lambda^*) = f'(x^*) - \sum_{i=1}^{m} \lambda_i g_i'(x^*) = 0,$$

$$\lambda_i^* \geq 0, \quad \lambda_i^* g_i(x^*) = 0, \quad i = 1, \ldots, m.$$

The vector λ^* is unique if the LICQ holds.

The *primary goal of this paper* is to present necessary and sufficient optimality conditions for problem (1) without assuming LICQ or MFCQ. If the LICQ or MFCQ does not hold at the solution, we say that problem (1) is *degenerate* (*nonregular, abnormal*). For example, in a slightly modified version of the problem posted in [10],

$$\begin{array}{ll} \underset{x \in \mathbb{R}^2}{\text{minimize}} & f(x) \\ \text{subject to} & -x_1^4 + x_2 \geq 0, \\ & 3x_1^4 - x_2 \geq 0, \end{array}$$

both the LICQ and MFCQ do not hold at the point $x^* = 0$. We give optimality conditions for this example in Section 6.

There are several methods to overcome the difficulty of degeneracy, for instance, those in [1, 11, 12, 16, 18, 19, 21]. Here we pursue an approach based on the construction of p-regularity introduced earlier in [22, 23, 24, 25]. The main idea of the p-regularity approach is in using higher order derivatives of the constraints $g_i(x)$ to replace the gradients of the active constraints which are linearly dependent. To compare our approach with others, we would like to note that Ledzewicz and Schättler in [18] introduced a concept of the p-regular mapping, but in a different sense. A mapping is called p–regular at a point x^* with respect to an element h_1 in our sense if it is p–regular in the direction of the sequence $H_{p-1} = (h_1, 0, \ldots, 0)$ in the sense of Ledzewicz and Schättler [18]. However, both our definition and the definition from [18] reduce to the same definition of 2–regularity for p=2. Ledzewicz and Schättler [18] also analyze p–regular problems, but they require the functions to be $(2p-1)$-times continuously differentiable while we assume that the functions are $(p+1)$-times continuously differentiable. Other results of this type are obtained in the work of Izmailov [14, 15] and of Izmailov and Solodov [16]. We compare our results with ones obtained in [14, 15, 16] and other relevant work in Section 7.

This paper continues the series of publications devoted to optimality conditions for degenerate optimization problems. In [3], we considered problems with equality constraints given in the operator form as $F(x) = 0$. The focus of the paper was on the case when the constraints are not regular at the solution x^* in the sense that the operator $F'(x^*)$ is not surjective. In [3], we derived new sufficient conditions for p-regular problems and necessary optimality conditions for problems satisfying the generalized condition of p–regularity. In [5] and [6], we turned our attention to problems with inequality constraints in the finite dimensional spaces in the completely degenerate case (19) when the LICQ and MFCQ do not hold. In [5], we introduced a new constraint qualification, the *p-regularity constraint qualification* (PRCQ), and derived new necessary optimality conditions for problems satisfying the PRCQ. One of the assumptions in [5] is existence of a nonzero element in the set $\hat{H}_p(x^*)$ defined by (21). In [6], we considered the case when the set $\hat{H}_p(x^*)$ consists only of a zero vector and derived necessary conditions for degenerate optimization problems in that case.

In this paper, we extend our consideration to the general case and propose necessary conditions for degenerate problems that are not necessarily satisfy assumption (19), which is one of the main assumptions in [5] and [6]. We also derive new second-order necessary conditions and sufficient conditions for optimality for degenerate optimization problems. The presented optimality conditions resemble the structure of the classical optimality conditions. Necessary conditions given in this paper reduce to the KKT conditions in the regular case. The optimality conditions can be applied to discretizations of calculus of variations and optimal control problems. In the end, we show that our assumptions are weaker than the Mangasarian-Fromovitz constraint qualification.

We give a special consideration to the case of $p = 2$ and propose new necessary and sufficient conditions for this case. At the same time, there are some important applications that require $p > 2$ in such areas as geometric programming, quantum physics, singular optimal controls and others. Moreover, the examples given in the paper illustrate that there exist problems for which the proposed p-regularity concept works. The consideration in the paper does not intend to cover all degenerate (nonregular, abnormal) problems, but it covers some classes of the nonregular problems and allows us to get some theoretical results that were not obtained earlier for these classes of problems. The results presented in the paper can be viewed as continuation of the work published by A. A. Tret'yakov and J. E. Marsden in [25].

The organization of the paper is as follows. Section 2 is devoted to an overview of the main concepts of the p-regularity theory. Namely, in Section 2.1, we recall some definitions of the p–regularity theory [25] for finite dimensional spaces. We also give one of the main results, a generalization of the Lyusternik theorem, which we need for our consideration in the paper. In Section 2.2, we recall the p–order conditions for equality constrained optimization problems derived earlier in [3] and [17]. Then in Section 2.3, we give a modified version of the necessary conditions proposed originally in [6]. Namely, we show that the Lagrange multiplier is nonnegative (this proof was not given in [6]). In addition, the technique that is used to prove that result is new and has a potential to be used to derive other results in optimality conditions.

We propose new results in Sections 3, 4, and 5. In Section 3.1, we continue considering the completely degenerate case and present new first- and second-order necessary optimality conditions for that case. The main difference between the results presented in Section 2.3 and ones derived in Section 3.1 is in the approach that is used to obtain the new optimality conditions. Namely, we do not use slack variables in Section 3.1. Note that in Theorem 3.4, in addition to the first-order necessary conditions, we also derive the second-order necessary conditions. In Section 3.2, we consider the case of general degeneration and derive new necessary conditions, first, for the case of $p = 2$, and then for any $p > 2$. Section 4 presents new sufficient conditions of optimality. In Section 5, we present a new result stating that the 2-regularity condition is weaker than the MFCQ. In Section 6, we illustrate the obtained results by several examples, including an example arising from a discretization of an isoperimetrical problem in calculus of variations. Comparison with the other work in the area as well as some concluding remarks are given in Section 7.

Notation. The *active set* $I(x^*)$ at any feasible x^* is the set of indices of the active constraints; i.e., $I(x^*) = \{i = 1, \dots, m \,|\, g_i(x^*) = 0\}$. We denote by $g_I(x)$ the vector of functions $g_i(x)$, $i \in I(x)$.

Let p be a natural number and let $B : \mathbb{R}^n \times \mathbb{R}^n \times \ldots \times \mathbb{R}^n$ (with p copies of $\mathbb{R}^n$) $\to \mathbb{R}^m$ be a continuous symmetric p-multilinear mapping. The p-form associated to B is the map $B[\cdot]^p : \mathbb{R}^n \to \mathbb{R}^m$ defined by $B[x]^p = B(x, x, \dots, x)$, for $x \in \mathbb{R}^n$. If $f : \mathbb{R}^n \to \mathbb{R}$ is a differentiable function, the vector of its first-order partial derivatives at a point $x \in \mathbb{R}^n$ will be denoted by $f'(x) : \mathbb{R}^n \to \mathbb{R}^n$. If $F : \mathbb{R}^n \to \mathbb{R}^m$ is of class C^p, we let $F^{(p)}(x)$ be the pth derivative of F at the point x (a symmetric multilinear map of p copies of $\mathbb{R}^n$ to $\mathbb{R}^m$) and the associated p-form (also called the *p-order mapping*) is $F^{(p)}(x)[h]^p = F^{(p)}(x)(h, h, \dots, h)$. Further, $\operatorname{Ker} \Lambda = \{x \in \mathbb{R}^n \mid \Lambda x = 0\}$ denotes the null–space (kernel) of a given linear operator $\Lambda : \mathbb{R}^n \to \mathbb{R}^m$, and $\operatorname{Im} \Lambda = \{y \in \mathbb{R}^m \mid y = \Lambda x$ for some $x \in \mathbb{R}^n\}$ is its image space. Furthermore, we use the following key notation: $\operatorname{Ker}^p F^{(p)}(x) = \{h \in \mathbb{R}^n \,|\, F^{(p)}(x)\,[h]^p = 0\,\}$ is the *p-kernel (p–null-cone)* of the p-order mapping.

2. Factor-analysis of nonlinear mappings: p–regularity theory

In this section, we give an overview of basic definitions of the p–regularity theory [25] for finite dimensional spaces and recall optimality conditions for degenerate equality–constrained optimization problems derived in our earlier work. We also give one of the main results, a generalization of the Lyusternik theorem, which we need for our consideration in the paper. Then in Section 2.3, we give a modified version of the necessary conditions (Theorem 2.12) proposed originally in [6]. Namely, we show that the Lagrange multiplier is nonnegative (this proof was not given in [6]). In addition, the technique that is used to prove that result is new and has a potential to be used to derive other results in optimality conditions. Since the necessary conditions presented in Theorem 2.12 are formulated in terms of the auxiliary functions $F(x)$ given by (22), we introduce Theorem 2.13, which is a new formulation of Theorem 2.12 in terms of the original constraint function $g(x)$.

2.1. **Main definitions and some results of the p-regularity theory.** Consider a sufficiently smooth nonlinear mapping $F : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$. The mapping F can be represented as the m-vector of functions $F_i(x) : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$; i.e., $F(x) = (F_1(x), \dots, F_m(x))^T$.

The mapping F is called *regular* at some point $x^* \in \mathbb{R}^n$ if

$$\operatorname{Im} F'(x^*) = \mathbb{R}^m, \tag{3}$$

or, in other words,

$$\operatorname{rank} F'(x^*) = m.$$

The mapping F is called *nonregular (irregular, degenerate)* if the regularity condition (3) is not satisfied.

Assume that we can decompose the space $\mathbb{R}^m$ into the direct sum

$$\mathbb{R}^m = Y_1 \oplus \ldots \oplus Y_p, \tag{4}$$

where $Y_1 = \operatorname{Im} F'(x^*)$, the image of the first derivative of F evaluated at x^*, and the remaining spaces are defined as follows. Let Z_2 be a complementary subspace to Y_1, and let $P_{Z_2} : \mathbb{R}^m \to Z_2$ be the projection operator onto Z_2 along Y_1. Let Y_2 be the linear span of the image of the quadratic map $P_{Z_2} F''(x^*)[\cdot]^2$. More generally, define inductively,

$$Y_i = \operatorname{span} \operatorname{Im} P_{Z_i} F^{(i)}(x^*)[\cdot]^i \subseteq Z_i, \quad i = 2, \dots, p-1, \tag{5}$$

where Z_i is a choice of complementary subspace for $(Y_1 \oplus \ldots \oplus Y_{i-1})$ with respect to $\mathbb{R}^m$, $i = 2, \dots, p$, and $P_{Z_i} : \mathbb{R}^m \to Z_i$ is the projection operator onto Z_i along $(Y_1 \oplus \ldots \oplus Y_{i-1})$ with respect to $\mathbb{R}^m$, $i = 2, \dots, p$. Finally, let $Y_p = Z_p$. The order p is chosen as the minimum number for which (4) holds.

Define the following mappings ([24])

$$f_i(x) : \mathbb{R}^n \to Y_i, \quad f_i(x) = P_i F(x), \quad i = 1, \dots, p, \tag{6}$$

where $P_i : \mathbb{R}^m \to Y_i$ is the projection operator onto Y_i along $(Y_1 \oplus \ldots \oplus Y_{i-1} \oplus Y_{i+1} \oplus \ldots \oplus Y_p)$ with respect to $\mathbb{R}^m$, $i = 1, \dots, p$.

A p-factor operator plays the central role in the p-regularity theory. We give the following definition of the p-factor operator.

Definition 2.1. *The linear operator $\Psi_p(h) : \mathbb{R}^n \to \mathbb{R}^m$, defined by*

$$\Psi_p(h) = f_1'(x^*) + f_2''(x^*)h + f_3'''(x^*)[h]^2 + \ldots + f_p^{(p)}(x^*)[h]^{p-1}, \quad h \in \mathbb{R}^n, \tag{7}$$

is called the p-factor operator.

Observe that in Definition 2.1, the p-factor operator depends on h, however, there is *no* dependence of Y_i defined by (5) on h.

Now we are ready to introduce another important definition of the p-regularity theory.

Definition 2.2. *The mapping F is called p-regular at x^* with respect to an element h if*

$$\operatorname{Im}\Psi_p(h) = \mathbb{R}^m.$$

A set $H_p(x^*)$ introduced below is the key one for many results obtained in the p-regularity theory.

$$H_p(x^*) = \bigcap_{i=1}^{p} \operatorname{Ker}^i f_i^{(i)}(x^*) = \{h \in \mathbb{R}^n \mid f_i^{(i)}(x^*)\,[h]^i = 0,\, i = 1, \ldots, p\}. \tag{8}$$

To obtain results, it is important to assume that F is p-regular at x^* not only with respect to some element h, but rather with respect to any vector h from the set $H_p(x^*)$. Thus, the following definition is a significant one.

Definition 2.3. *The mapping F is called p-regular at x^* if either $H_p(x^*) = \{0\}$ or F is p-regular with respect to any $h \in H_p(x^*)\backslash\{0\}$.*

For our consideration, we need the following result, which is a generalization of the classical Lyusternik theorem for p-regular mappings. To state the theorem, we recall a definition of a tangent vector and a tangent cone (see, for instance, Ioffe and Tihomirov [13]).

Definition 2.4. *We call h a tangent vector to a set $M \subseteq \mathbb{R}^n$ at $x^* \in M$ if there exist $\varepsilon > 0$ and a function $r : [0, \varepsilon] \to \mathbb{R}^n$ with the property that for $t \in [0, \varepsilon]$ we have $x^* + th + r(t) \in M$ and $\|r(t)\| = o(t)$. The collection of all tangent vectors at x^* is called the tangent cone to M at x^* and it is denoted by $T_1M(x^*)$.*

Theorem 2.5 (Generalized Lyusternik Theorem)**.** *Let U be a neighborhood of a point $x^* \in \mathbb{R}^n$. Assume that $F : \mathbb{R}^n \to \mathbb{R}^m$ is a p–times continuously differentiable mapping in U and is p–regular at x^*. Then the tangent cone to the set $M(x^*) = \{x \in U \mid F(x) = F(x^*)\}$ is $T_1M(x^*) = H_p(x^*)$, where $H_p(x^*)$ is given by* (8)*.*

The proof of Theorem 2.5 is given in [22] for the completely degenerate case, in [9] for the case of $p = 2$ and in [17] for the general case.

In the two following subsections, we give a specific form of the p-factor operator (7) and of the set $H_p(x^*)$ defined in (8) for the completely degenerate case and for the case of $p = 2$. We need these specific forms for our consideration in the following sections of the paper.

2.1.1. *Completely degenerate case.* For some $p \geq 2$ and some x^*, we say that we have the *completely degenerate case* if

$$F^{(r)}(x^*) = 0, \quad r = 1, \ldots, p-1. \tag{9}$$

In this case, $Y_1 = \ldots = Y_{p-1} = 0$, and the p-factor operator (7) could be defined as

$$\Psi_p(h) = F^{(p)}(x^*)[h]^{p-1}, \quad h \in \mathbb{R}^n, \tag{10}$$

since $f_i^{(i)}(x^*) = 0$, $i = 1, \ldots, p-1$, and $f_p^{(p)}(x^*) = F^{(p)}(x^*)$. Moreover, the set $H_p(x^*)$ defined in (8) reduces to

$$H_p(x^*) = \operatorname{Ker}^p F^{(p)}(x^*) = \{h \in \mathbb{R}^n \mid F_i^{(p)}(x^*)\,[h]^p = 0,\, i = 1, \ldots, m\}. \tag{11}$$

Furthermore, in the completely degenerate case (9), Definition 2.3 of p–regularity of F at x^* is equivalent to the linear independence of the vectors $F_i^{(p)}(x^*)[h]^{p-1}$, $i = 1, \ldots, m$, for any h from the set $H_p(x^*)$ defined by (11).

2.1.2. *Case of* $p = 2$. The following definition is a specific form of Definition 2.1 for the case of $p = 2$.

Definition 2.6. *A linear operator* $\Psi_2(h) : \mathbb{R}^n \to \mathbb{R}^m$,

$$\Psi_2(h) = F'(x^*) + P^{\perp} F''(x^*)h, \quad h \in \mathbb{R}^n, \quad \|h\| = 1,$$

is said to be the 2–factor-operator, where $P^{\perp}$ *is a matrix of the orthoprojector onto* $(\operatorname{Im} F'(x^*))^{\perp}$, *which is an orthogonal complementary subspace to the image of the first derivative of* F *evaluated at* x^*.

The following definition is a specific case of Definition 2.2.

Definition 2.7. *The mapping* F *is called 2-regular at* x^* *with respect to an element* h *if*

$$\operatorname{Im} \Psi_2(h) = \mathbb{R}^m.$$

In the case of $p = 2$, the set $H_p(x^*)$ defined in (8) reduces to

$$H_2(x^*) = \{h \in \mathbb{R}^n \,|\, F'(x^*)h = 0, \quad P^{\perp} F''(x^*)[h]^2 = 0\}. \tag{12}$$

2.2. **The p-order optimality conditions for equality constrained optimization problems.** In this section, we recall the pth order optimality conditions for the following equality constrained problem

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad F(x) = (F_1(x), \dots, F_m(x))^T = 0, \tag{13}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a sufficiently smooth real valued function and $F : \mathbb{R}^n \to \mathbb{R}^m$ is a sufficiently smooth mapping. We use $\bar{x}$ to denote the solution to equality constrained optimization problem (13), while we use x^* to denote a solution to an inequality constrained optimization problem.

Consider some vector $h \in \mathbb{R}^n$ and define the *p-factor-Lagrange function* [3, 25] as

$$\mathcal{L}_p(x, h, \lambda_0(h), y(h)) = \lambda_0(h) f(x) - \sum_{i=1}^{p} \langle y_i(h),\, f_i^{(i-1)}(x)[h]^{i-1} \rangle, \quad x \in \mathbb{R}^n \quad y_i(h) \in Y_i, \tag{14}$$

where f_i, $i = 1, \dots, p$, are defined in (6), Y_i are defined in (5), and $\langle \cdot, \cdot \rangle$ denotes the *scalar* (*dot*) product of two vectors. Here, the function $\mathcal{L}_p$ plays the role of the Lagrange function, and $\lambda_0(h)$ and $y_i(h)$ play the role of higher-order Lagrange multipliers.

Theorem 2.8 (Necessary and Sufficient Conditions for Optimality)**.** *Let* $U \subset \mathbb{R}^n$ *be a neighborhood of the point* $\bar{x}$. *Suppose that* $f \in C^2(U, \mathbb{R})$ *and that* $F \in C^{p+1}(U, \mathbb{R}^m)$. *Suppose also that there exists* $h \in H_p(\bar{x}) \backslash \{0\}$, *where* $H_p(\bar{x})$ *is defined by* (8). *Let* $\Psi_p(h)$ *be defined by* (7).

(1) *If* $\bar{x}$ *is a local solution to problem* (13), *then there exist* $\lambda_0(h) \in \mathbb{R}$ *and multipliers* $y_i(h) \in Y_i$, $i = 1, \dots, p$, *such that they do not all vanish , and*

$$\frac{\partial \mathcal{L}_p}{\partial x}(\bar{x}, h, \lambda_0(h), y(h)) = \lambda_0(h)\, f'(\bar{x}) - \sum_{i=1}^{p} (f_i^{(i)}(\bar{x})[h]^{i-1})^T y_i(h) = 0.$$

If $\operatorname{Im} \Psi_p(h) = Y_1 \oplus \ldots \oplus Y_p$, *then* $\lambda_0(h) \neq 0$, *so*

$$\frac{\partial \mathcal{L}_p}{\partial x}(\bar{x}, h, 1, y(h)) = 0.$$

Moreover,

$$\frac{\partial^2 \mathcal{L}_p}{\partial x^2}(\bar{x}, h, 1, \tilde{y}(h))[h]^2 \geq 0, \tag{15}$$

where

$$\tilde{y}(h) = \left(y_1(h), \frac{1}{3} y_2(h), \dots, \frac{2}{i(i+1)} y_i(h), \dots, \frac{2}{p(p+1)} y_p(h) \right). \tag{16}$$

(2) *Suppose that* $\operatorname{Im}\Psi_p(h) = Y_1 \oplus \ldots \oplus Y_p$ *for any element* $h \in H_p(\bar{x})$.
If there exist $\alpha > 0$ *and multipliers* $y_i(h) \in Y_i$, $i = 1, \ldots, p$, *such that*

$$\frac{\partial \mathcal{L}_p}{\partial x}(\bar{x}, h, 1, y(h)) = 0 \tag{17}$$

and for all $h \in H_p(\bar{x})$

$$\frac{\partial^2 \mathcal{L}_p}{\partial x^2}(\bar{x}, h, 1, \tilde{y}(h))[h]^2 \geq \alpha \|h\|^2, \tag{18}$$

where $\tilde{y}(h)$ *is defined by* (16), *then* $\bar{x}$ *is an isolated solution to problem* (13).

Remark 2.1. (1) *The sufficient optimality conditions stated in part* (2) *complement the second-order necessary conditions given in part* (1). *In both parts* $\tilde{y}(h)$ *is defined by* (16).

(2) *The proof of statement* (1) *of Theorem* 2.8 *(the necessary conditions for optimality) is given in* [17], *and the proof of part* (2) *is given in* [3]. *The statement* (2) *in* [3] *has a error: the Lagrange multiplier* $\tilde{y}(h)$ *in the sufficiency part should be defined by* (16). *The error was corrected in* [4].

2.3. **Necessary optimality conditions for inequality constrained problems in the completely degenerate case.** In this section, we give a modified version of the necessary optimality conditions for problem (1), which were presented in our paper [6]. We consider the following completely degenerate case:

$$g_i^{(r)}(x^*) = 0, \quad r = 1, \ldots, p-1, \quad p \geq 2, \quad \forall i \in I(x^*), \tag{19}$$

where x^* is a solution to (1), and p denotes a number for which (19) holds.

The following constraint qualification was introduced in [5].

Definition 2.9 (p-regularity constraint qualification (PRCQ))**.** *Given a point* x^* *and the active set* $I(x^*)$, *we say that the p-regularity constraint qualification* (*PRCQ*) *holds at the feasible point* x^* *with respect to a vector* $h \in \mathbb{R}^n$ *if* x^* *is strictly feasible or if* (19) *holds and the set of vectors* $\{g_i^{(p)}(x^*)[h]^{p-1},\ i \in I(x^*)\}$ *is linearly independent.*

Let the rows of matrix $G_I(h)$ be composed of the vectors $g_i^{(p)}(x^*)[h]^{p-1}$, $i \in I(x^*)$. If the PRCQ holds at the point x^* with respect to the vector h, then the matrix $G_I(h)$ has full rank. In the regular case, the matrix $G_I(h)$ reduces to the Jacobian matrix of the active constraints. Thus, the PRCQ is a generalization of the LICQ.

Consider some vector $h \in \mathbb{R}^n$ and define the *p-factor–Lagrange function*:

$$L_p(x, h, \lambda(h)) = f(x) - \sum_{i=1}^{m} \lambda_i(h) g_i^{(p-1)}(x)[h]^{p-1}, \quad \lambda(h) = (\lambda_1(h), \ldots, \lambda_m(h))^T. \tag{20}$$

Similarly to the definition of the set $H_p(x^*)$ given by (11), we introduce set $\hat{H}_p(x^*)$, which is used in optimality conditions for problem (1),

$$\hat{H}_p(x^*) = \{h \in \mathbb{R}^n \mid g_i^{(p)}(x^*)[h]^p = 0, \quad i \in I(x^*)\}. \tag{21}$$

Consider the case when the set $\hat{H}_p(x^*)$ consists of the zero vector only and (19) holds with an even p. In this case, we can convert problem (1) to a problem with equality constraints by introducing slack variables s_i and replacing the inequalities $g_i(x) \geq 0$ with the equalities

$$F_i(x, s) = g_i(x) - s_i^p = 0, \quad s_i \in \mathbb{R}, \quad i = 1, \ldots, m. \tag{22}$$

The introduction of slack variables transforms problem (1) into the following one:

$$\underset{(x,\, s)}{\text{minimize}}\ f(x) \quad \text{subject to} \quad F_i(x, s) = g_i(x) - s_i^p = 0, \quad i = 1, \ldots, m, \tag{23}$$

where $F_i : \mathbb{R}^{n\times m} \to \mathbb{R}$, $i = 1, \dots, m$. This transformation of problem (1) into the equality constrained form (23) gives new possibilities for the analysis of problem (1). The slack variables are also used in [7] to obtain new methods for solving degenerate inequality constrained optimization problems. The following result stated in [6] is used in our consideration in the paper.

Lemma 2.10. *Let x^* be a feasible point for problem* (1)*. Then x^* is a local minimizer to* (1) *if and only if $(x^*, 0)$ is a local minimizer to the following problem*

$$\underset{(x,\,s)}{\text{minimize}}\ f(x) \quad \text{subject to} \quad F_i(x,s) = g_i(x) - s_i^p = 0, \quad i \in I(x^*). \tag{24}$$

Consider some element $h = (h_x, h_s)$, where $h_x \in \mathbb{R}^n$ and $h_s \in \mathbb{R}^m$, and define the set

$$\widetilde{H}_p(x^*, 0) = \{h \in \mathbb{R}^{n\times m} \mid F_i^{(p)}(x^*, 0)[(h_x, h_s)]^p = 0, \quad i \in I(x^*)\}. \tag{25}$$

This set $\widetilde{H}_p(x^*, 0)$ replaces the set $\hat{H}_p(x^*)$ in this section.

We denote by $F(x, s)$, the vector of functions $F_i(x, s)$, $i \in I(x^*)$. Recall that we consider the case of $\hat{H}_p(x^*) = \{0\}$, where $\hat{H}_p(x^*)$ is defined in (21), i.e., only $h = 0$ satisfies $g_i^{(p)}(x^*)[h]^p = 0$ for all $i \in I(x^*)$. We assume that $\widetilde{H}_p(x^*, 0) \neq \{0\}$; otherwise, in accordance with the following lemma, x^* is an isolated feasible point for problem (1), and solving problem (1) reduces to solving the system of nonlinear inequalities which are constraints in (1).

Auxiliary properties of problem (1) derived in the following lemma were mentioned in [6] but without a formal statement and a proof, so we provide them here for completeness of our consideration.

Lemma 2.11. *Let x^* be a feasible point to* (1)*. Let the set $\widetilde{H}_p(x^*, 0)$ be defined by* (25)*, and let $F(x, s)$ be the vector of functions $F_i(x, s)$, $i \in I(x^*)$, with F_i defined by* (22)*.*

1. If the set $\widetilde{H}_p(x^*, 0)$ consists only of the zero vector, then x^* is an isolated feasible point for problem (1)*.*

2. Assume that $F(x, s)$ is p–regular at $(x^*, 0)$ and x^* is an isolated feasible point for problem (1)*. Then $\widetilde{H}_p(x^*, 0) = \{0\}$.*

Proof.

1. Since $\widetilde{H}_p(x^*, 0) = \{0\}$, then by Definition 2.3, the mapping $F(x, s)$ is p–regular at $(x^*, 0)$. Thus, by Theorem 2.5, $T_1M(x^*, 0) = \{0\}$, where $T_1M(x^*, 0)$ is the tangent cone to the set $M(x^*, 0) = \{(x, s) \in \bar{U} \mid F(x, s) = F(x^*, 0)\}$. This means that the point $(x^*, 0)$ is an isolated solution to the system $F(x, s) = 0$. Hence, by definition of F, x^* is an isolated solution to $g_i(x) = 0$, $i \in I(x^*)$. Together with the assumption that x^* is a feasible point to (1), this implies that x^* is an isolated feasible point for problem (1).

2. If x^* is an isolated feasible point for problem (1), then, the point $(x^*, 0)$ is an isolated solution to the system $F(x, s) = 0$. Hence, $T_1M(x^*, 0) = 0$. Since F is p-regular at $(x^*, 0)$, then by Theorem 2.5, $\widetilde{H}_p(x^*, 0) = \{0\}$.

In this subsection, $h = (h_x, h_s)$ denotes an element from the set $\widetilde{H}_p(x^*, 0)$, and h_{si} denotes the ith component of the vector h_s.

A slight modification of the following theorem was derived in [6], namely, we did not state the nonnegativity condition of the Lagrange multipliers.

We introduce a modification of the function $L_p(x, h, \lambda(h))$ given in (20) as follows:

$$\bar{L}_p(x, h, \lambda(h)) = f(x) - \sum_{i\in I(x^*)} \lambda_i(h) g_i^{(p-1)}(x)[h]^{p-1}. \tag{26}$$

Theorem 2.12 (The pth order necessary conditions for optimality in the case of an even p and $\hat{H}_p(x^*) = \{0\}$)**.** *Let x^* be a local minimum of problem* (1)*, let $U \subset \mathbb{R}^n$ be a neighborhood of the point x^*, and let* (19) *hold with an even p. Let the set $\widetilde{H}_p(x^*, 0)$ be defined by* (25)*, and let $F_i(x, s)$, $i = 1, \dots, m$, be defined by* (22)*. Suppose that $f \in C^2(U, \mathbb{R})$ and that $g \in$*

$C^{p+1}(U,\mathbb{R}^m)$. Suppose also that there exists $h=(h_x,h_s)$, $\|h\|=1$, $h\in \widetilde{H}_p(x^*,0)$, such that the vectors $F_i^{(p)}(x^*,0)[(h_x,h_s)]^{p-1}$, $i\in I(x^*)$, are linearly independent. Then there exists a Lagrange multiplier vector $\lambda^*(h)$, with components $\lambda_i^*(h)\geq 0$, $i\in I(x^*)$, such that

$$\frac{\partial \bar{L}_p}{\partial x}(x^*,h_x,\lambda^*(h)) = f'(x^*) - \sum_{i\in I(x^*)} \lambda_i^*(h) g_i^{(p)}(x^*)[h_x]^{(p-1)} = 0, \tag{27}$$

$$\lambda_i^*(h)h_{s_i}=0, \quad i\in I(x^*). \tag{28}$$

Proof. Eqs. (27) and (28) were derived in [6]. Now, we show that $\lambda_i^*(h)\geq 0$, $i\in I(x^*)$, The proof consists of several following parts. We consider the vector $h=(h_x,h_s)\in \widetilde{H}_p(x^*,0)$, such that the vectors $F_i^{(p)}(x^*,0)[h_x,h_s]^{p-1}$, $i\in I(x^*)$, are linearly independent.

(1) Assume on the contrary that there exists $j\in I(x^*)$ such that $\lambda_j^*(h)<0$. Since by the assumption, vectors $F_i^{(p)}(x^*,0)[h_x,h_s]^{p-1}$, $i\in I(x^*)$, are linearly independent, by the Kronecker-Capelli theorem, the following system with some $y_j>0$ and $y_i=0$, $i\neq j$, has a solution $\bar{\xi}=(\bar{\xi}_x,\bar{\xi}_s)$ such that

$$\begin{aligned} &(F_j^{(p)}(x^*,0)[h_x,h_s]^{p-1})\bar{\xi} = y_j>0,\\ &(F_i^{(p)}(x^*,0)[h_x,h_s]^{p-1})\bar{\xi} = y_i=0, \quad i\neq j, \quad i\in I(x^*). \end{aligned} \tag{29}$$

Since $\lambda_j^*<0$, we have $h_{s_j}=0$ by (28). Then by the definition of F we obtain

$$F_j{}^{(p)}(x^*,0)[h_x,h_s]^{p-1} = g_j^{(p)}(x^*)[h_x]^{p-1}, \tag{30}$$

and by (29) and (30),

$$(F_j^{(p)}(x^*,0)[h_x,h_s]^{p-1})\bar{\xi} = g_j^{(p)}(x^*)[h_x]^{p-1}\bar{\xi}_x = y_j>0. \tag{31}$$

Similarly, by (28), (29), and (30), for $i\neq j$, $i\in I(x^*)$, we have either $\lambda_i^*=0$ and/or $h_{s_i}=0$ and

$$g_i^{(p)}(x^*)[h_x]^{p-1}\bar{\xi}_x = y_i = 0. \tag{32}$$

(2) We will show that the arc $(\alpha h_x+\alpha^{3/2}\bar{\xi}_x+\omega(\alpha))$, $\|\omega(\alpha)\|=o(\alpha^3/2)$, is feasible at x^* for some sufficiently small $\alpha>0$, that is, for $i=1,\dots,m$,

$$g_i(x^*+\alpha h_x+\alpha^{3/2}\bar{\xi}_x+\omega(\alpha))\geq 0. \tag{33}$$

Since (19) holds, by Taylor expansion, we have for $i\in I(x^*)$,

$$g_i(x^*+\alpha h_x+\alpha^{3/2}\bar{\xi}_x+\omega(\alpha)) = \tag{34}$$

$$= \frac{1}{p!}\alpha^p g_i^{(p)}(x^*)[h_x]^p + \frac{1}{p!}\alpha^{p+1/2} g_i^{(p)}(x^*)[h_x^{p-1},\bar{\xi}_x] + \tilde{\omega}(\alpha)$$

,

where $\|\tilde{\omega}(\alpha)\|=o(\alpha^{p+1/2})$. Note that by the definition of F and $(h_x,h_s)\in \widetilde{H}_p(x^*,0)$, we have that

$$0 = F_i^{(p)}(x^*,0)[(h_x,h_s)]^p = g_i^{(p)}(x^*)[h_x]^p - p!h_{s_i}^p, \quad i\in I(x^*). \tag{35}$$

Assume that $\alpha>0$ is sufficiently small and show that (33) holds for $i=1,\dots,m$. If $i\in I(x^*)$ and $h_{s_i}\neq 0$, then since p is even, $g_i^{(p)}(x^*)[h_x]^p>0$ by (35), and (34) yields (33). If $i=j$ (j is defined as the index for which $\lambda^*<0$), then $h_{s_j}=0$ and (33) follows from (31), (34), and (35). For $i\in I(x^*)$ such that $g_i^{(p)}(x^*)[h_x]^p=0$ and $g_i^{(p)}(x^*)[h_x^{p-1},\bar{\xi}_x]=0$, the generalized Lyusternik Theorem [17] yields

$$g_i(x^*+\alpha h_x+\alpha^{3/2}\bar{\xi}_x+\omega(\alpha))=0,$$

and, hence, (33) holds. For $i\notin I(x^*)$, since $g_i(x^*)>0$, we can choose α sufficiently small so that (33) holds.

(3) Eq. (27) yields

$$f'(x^*)(\alpha h_x + \alpha^{3/2}\bar{\xi}_x + \omega(\alpha)) = \sum_{i\in I(x^*)} \lambda_i^*(h) g_i^{(p)}(x^*)[h_x]^{(p-1)}(\alpha h_x + \alpha^{3/2}\bar{\xi}_x + \omega(\alpha)).$$

As follows from consideration above, for $i \in I(x^*)$, either $\lambda_i^*(h) = 0$ or (31) and (32) hold. Hence,

$$f'(x^*)\bar{\xi}_x = \sum_{i\in I(x^*)} \lambda_i^*(h) g_i^{(p)}(x^*)[h_x]^{(p-1)}\bar{\xi}_x = \lambda_j^*(h) g_j^{(p)}(x^*)[h_x]^{(p-1)}\bar{\xi}_x < 0.$$

Moreover, for every $i \in I(x^*)$ either $\lambda_i^*(h) = 0$ or $g_i^{(p)}(x^*)[h_x]^p = 0$. Then using Taylor expansion of the function f gives

$$f(x^* + \alpha h_x + \alpha^{3/2}\bar{\xi}_x + \omega(\alpha)) < f(x^*),$$

which contradicts the assumption that x^* is a local minimizer.

Theorem 2.12 gives necessary conditions for problem (1), but its statement contains functions $F_i(x, s)$ introduced in (23) and the set $\widetilde{H}_p(x^*, 0)$ defined in (25). In the following theorem, we give a new reformulation of Theorem 2.12 in terms of the original constraint function $g(x)$. Recall that by definition of the functions $F_i(x, s)$ in (22), the set $\widetilde{H}_p(x^*, 0)$ given in (25) can be defined as

$$\bar{H}_p(x^*) = \{h \in \mathbb{R}^n \mid g_i^{(p)}(x^*)[h]^p \geq 0, \quad i \in I(x^*)\}. \tag{36}$$

The next theorem is a reformulation of Theorem 2.12 in terms of the function L_p defined in (20) and of the original constraint function $g(x)$.

Theorem 2.13 (The pth order necessary conditions for optimality)**.** *Let x^* be a local minimum of problem* (1)*, let $U \subset \mathbb{R}^n$ be a neighborhood of the point x^*, and let* (19) *hold with an even p. Let the set $\bar{H}_p(x^*)$ be defined by* (36)*. Suppose that $f \in C^2(U, \mathbb{R})$ and that $g \in C^{p+1}(U, \mathbb{R}^m)$. Suppose also that there exists $h = (h_x, h_s)$, $\|h\| = 1$, $h \in \bar{H}_p(x^*)$, such that the vectors $g_i^{(p)}(x^*)[h]^{p-1}$, $i \in I(x^*)$, are linearly independent, that is the PRCQ holds at x^* with respect to the vector h. Then there exists a Lagrange multiplier vector $\lambda^*(h)$, with components $\lambda_i^*(h) \geq 0$, $i \in I(x^*)$, such that*

$$\frac{\partial L_p}{\partial x}(x^*, h, \lambda^*(h)) = f'(x^*) - \sum_{i=1}^{m} \lambda_i^*(h) g_i^{(p)}(x^*)[h]^{p-1} = 0, \tag{37}$$

$$\lambda_i^*(h) g_i^{(p)}(x^*)[h]^p = 0, \quad i = 1, \ldots, m, \tag{38}$$

$$\lambda_i^*(h) \geq 0, \quad \lambda_i^*(h) g_i(x^*) = 0, \quad i = 1, \ldots, m. \tag{39}$$

3. New necessary optimality conditions for degenerate optimization problems

The theorems derived in this section can be viewed as extension of our results presented in Section 2.3 for inequality constrained optimization problems for the case of $\hat{H}_p(x^*) = \{0\}$ and an arbitrary p. In Section 3.1, we continue considering the completely degenerate case when condition (19) holds. The main difference between the results presented in Section 2.3 and ones derived in Section 3.1 is in the approach that is used to obtain the new optimality conditions. Namely, we do not use slack variables in Section 3.1. Note that in Theorem 3.4, in addition to the first-order necessary conditions, we also derive the second-order necessary conditions. Then in Section 3.2 we obtain new necessary optimality conditions for the case when a solution x^* is degenerate, but (19) does not hold for any p. Note that some results presented in the previous section hold for even or odd values of p only. We do not make additional requirements on p in this section, so the new necessary conditions hold for any p.

3.1. **New necessary optimality conditions in the completely degenerate case for any $p \geq 2$.** We consider the case when $\hat{H}_p(x^*) = \{0\}$ and (19) holds. Introduce the following notation

$$\begin{aligned} A_f(x^*) &= \{h \in \mathbb{R}^n \mid \langle f'(x^*), h\rangle \geq 0\}, \\ A_g(x^*) &= \{h \in \mathbb{R}^n \mid g_i^{(p)}(x^*)[h]^p \geq 0, \quad i \in I(x^*)\} \\ I_0(h) &= \{i \in I(x^*) \mid g_i^{(p)}(x^*)[h]^p = 0, \quad h \in A_g(x^*)\}, \\ K(x^*) &= A_g(x^*) \bigcap \mathrm{Ker} f'(x^*). \end{aligned}$$

Definition 3.1. *The mapping $g(x)$ is called p–regular at x^* with respect to h, if the vectors $\{g_i^{(p)}(x^*)[h]^{p-1},\ i \in I_0(h)\}$, are linearly independent.*

Definition 3.2. *The mapping $g(x)$ is called p–regular at x^* if g is p-regular with respect to any h such that $I_0(h) \neq \emptyset$.*

In the proof of the following theorem, we use Farkas' lemma:

Lemma 3.3 (Farkas' lemma)**.** *Let $a_1, \ldots, a_r, c \in \mathbb{R}^n$. Then*

$$c^T\, y \geq 0$$

for all y such that $a_j^T\, y \leq 0$, $j = 1, \ldots, r$, if and only if there exist nonnegative scalars $\mu_1, \ldots, \mu_r$ such that

$$c + \mu_1 a_1 + \ldots + \mu_r a_r = 0.$$

Theorem 3.4. *Let U be some neighborhood of $x^* \in \mathbb{R}^n$, $f \in C^2(U)$, $g \in C^{p+1}(U)$, and (19) hold. Assume that mapping $g(x)$ is p–regular at x^*. If x^* is a local minimizer to (1), then*

$$A_g(x^*) \subset A_f(x^*). \tag{40}$$

Moreover, for any $h \in K(x^*)$, $\|h\| = 1$, there exists $\lambda^*(h) = (\lambda_i^*(h))_{i \in I_0(h)}$, $\lambda_i^*(h) \geq 0$, such that

$$\frac{\partial L_p}{\partial x}(x^*, h, \lambda^*(h)) = f'(x^*) - \sum_{i \in I_0(h)} \lambda_i^*(h) g_i^{(p)}(x^*)[h]^{p-1} = 0 \tag{41}$$

and

$$\langle \frac{\partial^2 L_p}{\partial x^2}(x^*, h, \tilde{\lambda}^*(h))h, h\rangle \geq 0, \quad \tilde{\lambda}^*(h) = \frac{2\lambda^*(h)}{p(p+1)}. \tag{42}$$

Remark. If there exists $h \in K(x^*)$ such that $I_0(h) = \emptyset$, then $f'(x^*) = 0$.

Proof. To prove (40), assume on the contrary that x^* is a local minimizer to (1), and there exists $\bar{h} \in A_g(x^*)$ with $\langle f'(x^*), \bar{h}\rangle < 0$. Then, by Taylor expansion, since $g(x)$ is p-regular at x^*, there exists a sufficiently small $t > 0$ such that $f(x^* + t\bar{h}) < f(x^*)$ and $g(x^* + t\bar{h}) \geq 0$, which contradicts the assumption that x^* is a local minimizer. Hence, (40) holds.

Now, consider an element $h \in K(x^*)$, and divide our consideration into two following cases:

(1) If $I_0(h) = \emptyset$, then $g_i^{(p)}(x^*)[h]^p > 0$ for all $i \in I(x^*)$. Moreover, $h \in K(x^*)$ implies $f'(x^*)h = 0$. If we assume that $f'(x^*) \neq 0$, then there exists $\bar{h}$ such that $\langle f'(x^*), \bar{h}\rangle < 0$. Then for a sufficiently small $t > 0$, we have $g(x^* + th + t^{3/2}\bar{h}) \geq 0$ and

$$f(x^* + th + t^{3/2}\bar{h}) < f(x^*),$$

which contradicts the assumption that x^* is a local minimizer. Hence,

$$f'(x^*) = 0 \tag{43}$$

and (41) holds with $\lambda_i(h) = 0$ for all $i \in I_0(h)$. Moreover, using the Taylor expansion of $f(x^* + th)$, (43), the assumption that x^* is a local minimizer, the definition of function $\mathcal{L}$, and the property $\lambda_i(h) = 0$ for all $i \in I_0(h)$ we get

$$0 \leq f(x^* + th) - f(x^*) = \frac{1}{2}\left\langle f''(x^*)th, th\right\rangle + o(t^2) = \frac{1}{2}t^2\langle \mathcal{L}''_{p_{xx}}(x^*, h, \tilde{\lambda}^*(h))h, h\rangle + o(t^2),$$

which yields (42).

(2) If $I_0(h) \neq \emptyset$, then one can show that $\langle f'(x^*), z\rangle \geq 0$ for any z such that $g_i^{(p)}(x^*)[h]^{p-1}z \geq 0$, $i \in I_0(h)$, which implies (41) by Farkas Lemma 3.3.

Now, we will show that (42) holds in this case. Consider any h such that $I_0(h) \neq \emptyset$. Then $g_i^{(p)}(x^*)[h]^p = 0$, $i \in I_0(h)$. Then by the generalized Lyusternik Theorem, there exists $r(t)$ and a sufficiently small $\varepsilon > 0$ such that for $t \in (0, \varepsilon)$,

$$g_i(x^* + th + r(t)) = 0, \quad \|r(t)\| = o(t), \quad i \in I_0(h).$$

Since x^* is a local minimizer, $f'(x^*)h = 0$, and $\|r(t)\| = o(t)$, we get by using the Taylor expansion

$$0 \leq f(x^* + th + r(t)) - f(x^*) = f'(x^*)r(t) + \frac{1}{2}f''(x^*)(th)^2 + o(t^2). \tag{44}$$

Moreover, we have for every $i \in I_0(h)$ and $t \in (0, \varepsilon)$,

$$\begin{aligned} 0 &= g_i(x^* + th + r(t)) = \\ &= \tfrac{1}{p!}g_i^{(p)}(x^*)[th + r(t)]^p + \tfrac{1}{(p+1)!}g_i^{(p+1)}(x^*)[th + r(t)]^{p+1} + o(t^{p+1}) = \\ &= \tfrac{p}{p!}g_i^{(p)}(x^*)[th]^{p-1}r(t) + \tfrac{1}{(p+1)!}g_i^{(p+1)}(x^*)[th]^{p+1} + o(t^{p+1}). \end{aligned} \tag{45}$$

Multiplying both parts of (45) by $\frac{(p-1)!}{t^{p-1}}\lambda_i^*(h)$ for every $i \in I_0(h)$ and subtracting the resulting equations from (44) we get

$$0 \leq \left\langle f'(x^*) - \sum_{i \in I_0(h)} \lambda_i^*(h) g_i^{(p)}(x^*)[h]^{p-1}, r(t)\right\rangle +$$

$$+\frac{1}{2}f''(x^*)(th)^2 + \frac{(p-1)!}{(p+1)!}\sum_{i \in I_0(h)} \lambda_i^*(h)g_i^{(p+1)}(x^*)[h]^{p-1}(th)^2 + o(t^2).$$

Hence, by (41),

$$f''(x^*)(th)^2 + \sum_{i \in I_0(h)} \frac{2}{p(p+1)}\lambda_i^*(h)g_i^{(p+1)}(x^*)[h]^{p-1}(th)^2 + o(t^2) \geq 0.$$

Dividing the last relation by t^2 and taking the limit as $t \to 0$ we get (42), which finishes the proof.

3.2. **New necessary optimality conditions in the general case.** In this section, we consider the general case. We assume that the vectors $g_i'(x^*)$ are linearly dependent, $i \in I(x^*)$, and there exists an index $j \in I(x^*)$ such that $g_j'(x^*) \neq 0$, where x^* is a solution to (1). To make it simpler, we first consider the case of $p = 2$.

The case of $p = 2$. Convert problem (1) to a problem with equality constraints by introducing slack variables s_i^2 and replacing the inequalities $g_i(x) \geq 0$, $i = 1, \dots, m$, by

$$g_i(x) - s_i^2 = 0, \quad s_i \in \mathbb{R}, \quad i = 1, \dots, m.$$

By Lemma 2.10, if x^* is a local minimizer to problem (1), $(x^*, 0)$ is a local minimizer to the following problem

$$\begin{array}{ll} \underset{(x,\,s)}{\text{minimize}} & f(x) \\ \text{subject to} & F_i(x, s) = g_i(x) - s_i^2 = 0, \quad i \in I(x^*). \end{array} \tag{46}$$

Without loss of generality, assume that $I(x^*) = \{1, \dots, k\}$. Moreover, assume that the vectors $g_i'(x^*)$, $i = 1, \dots, r$, $r < k$, are linearly independent for some $r \geq 1$. Then $g_i'(x^*)$, $i = r+1, \dots, k$, can be represented as a linear combination of $g_1'(x^*)$, …, $g_r'(x^*)$ with some coefficients α_{1i}, …, α_{ri}.

By analogy with the procedure in [2], transform problem (46) into the following problem:

$$\begin{array}{ll} \underset{(x,\, s)}{\text{minimize}} & f(x) \\ \text{subject to} & F_i(x,s) = 0, \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad i = 1, \dots, r, \\ & \widetilde{F}_i(x,s) = F_i(x,s) - \alpha_{1i}F_1(x,s) - \ldots - \alpha_{ri}F_r(x,s) = 0 \quad i = r+1, \dots, k, \end{array} \tag{47}$$

where $\frac{\partial \widetilde{F}_i}{\partial x}(x^*, 0) = 0$, $i = r+1, \dots, k$. As was shown in [6], $(x^*, 0)$ is a local minimizer to problem (46) if and only if $(x^*, 0)$ is a local minimizer to (47).

Introduce the mapping

$$\bar{F}(x,s) = \begin{pmatrix} F_1(x,s) \\ \dots \\ F_r(x,s) \\ \widetilde{F}_{r+1}(x,s) \\ \dots \\ \widetilde{F}_k(x,s) \end{pmatrix}. \tag{48}$$

We use $P^{\perp}$ to denote the orthoprojector onto $(\operatorname{Im} \bar{F}'(x^*))^{\perp}$ in $\mathbb{R}^k$, where $(\operatorname{Im}\bar{F}\,'(x^*))^{\perp}$ is the orthogonal complementary subspace of $\operatorname{Im}\bar{F}'(x^*)$ in $\mathbb{R}^k$. By definition of $\bar{F}$, matrix $P^{\perp}$ is a diagonal matrix, namely, $P^{\perp} = \operatorname{diag}(p_j)_{j=1}^k$, where

$$p_j = \begin{cases} 0, & j = 1, \dots, r, \\ 1, & j = r+1, \dots, k \end{cases}. \tag{49}$$

We also use the following notation

$$\widetilde{g}_i(x) = g_i(x) - \alpha_{1i}g_1(x) - \ldots - \alpha_{ri}g_r(x), \quad i = r+1, \dots, k \tag{50}$$

where α_{ji} are defined in (47). Moreover, (47) implies that $\widetilde{g}_i'(x^*) = 0$, $i = r+1, \dots, k$.

For $h \in \mathbb{R}^n$, define the 2-*factor-Lagrange function* as

$$L_2(x, h, \lambda(h), \gamma(h)) = f(x) - \sum_{i=1}^{r} \lambda_i(h) g_i(x) - \sum_{i=r+1}^{k} \gamma_i(h)(\widetilde{g}_i'(x)h),$$

where $\lambda(h) = (\lambda_1(h), \dots, \lambda_r(h))$ and $\gamma(h) = (\gamma_{r+1}(h), \dots, \gamma_k(h))$.

Theorem 3.5 (Necessary conditions for optimality in the general case)**.** *Let x^* be a local minimum of problem (1), and let $U \subset \mathbb{R}^n$ be a neighborhood of the point x^*. Let the mapping $\bar{F}$ be given by (48). Suppose that $f \in C^2(U, \mathbb{R})$ and that $g \in C^2(U, \mathbb{R}^m)$. Assume that $I(x^*) = \{1, \dots, k\}$. Moreover, assume that the vectors $g_i'(x^*)$, $i = 1, \dots, r$, $r < k$, are linearly independent for some $r \geq 1$. Suppose that there exists $h = (h_x, h_s)$, $\|h\| = 1$, such that*

$$\bar{F}'(x^*, 0)h = 0, \quad P^{\perp}\bar{F}''(x^*, 0)[h]^2 = 0, \tag{51}$$

and $\bar{F}(x,s)$ is 2-regular at $(x^*, 0)$ with respect to the element h. Then there exist multipliers $\lambda^*(h) = (\lambda_1^*(h), \dots, \lambda_r^*(h))$ and $\gamma^*(h) = (\gamma_{r+1}^*(h), \dots, \gamma_k^*(h))$ such that

$$\frac{\partial L_2}{\partial x}(x^*, h, \lambda^*(h), \gamma^*(h)) = f'(x^*) - \sum_{i=1}^{r} \lambda_i^*(h) g_i'(x^*) - \sum_{i=r+1}^{k} \gamma_i^*(h) \widetilde{g}_i''(x^*) h_x = 0, \tag{52}$$

and

$$\sum_{i=r+1}^{k} \gamma_i^*(h) A^i h_s = 0, \tag{53}$$

where A^i, $i = r+1, \dots, k$, is a diagonal $k \times k$ matrix given as $A^i = diag(a^i_j)^k_{j=1}$,

$$a^i_j = \begin{cases} 2\alpha_{j\,i}, & j = 1, \dots, r \\ -2, & j = i \\ 0, & otherwise \end{cases}, \tag{54}$$

where α_{ji} are defined in (47).

Proof. All conditions of Theorem 2.8 hold for problem (47). By Theorem 2.8, we have

$$f'_{(x,s)}(x^*) = \sum_{i=1}^{r} \lambda^*_i(h) F'_i(x^*, 0) + \sum_{i=r+1}^{k} \gamma^*_i(h)(\widetilde{F}''_i(x^*, 0)[h]).$$

The last relation is equivalent to the two following equalities:

$$\begin{aligned} f'_x(x^*) \quad &= \sum_{i=1}^{r} \lambda^*_i(h) \left(\frac{\partial F_i}{\partial x}(x^*, 0) \right) + \sum_{i=r+1}^{k} \gamma^*_i(h) \left(\frac{\partial^2 \widetilde{F}_i}{\partial x^2}(x^*, 0) h_x \right) = \\ &= \sum_{i=1}^{r} \lambda^*_i(h) g'_i(x^*) + \sum_{i=r+1}^{k} \gamma^*_i(h)(\widetilde{g}''_i(x^*) h_x) \end{aligned} \tag{55}$$

and

$$0 = f'_s(x^*) = \sum_{i=1}^{r} \lambda^*_i(h) \left(\frac{\partial F_i}{\partial s}(x^*, 0) \right) + \sum_{i=r+1}^{k} \gamma^*_i(h) \left(\frac{\partial^2 \widetilde{F}_i}{\partial s^2}(x^*, 0) h_s \right). \tag{56}$$

We have for $i = 1, \dots, r$, $\frac{\partial \widetilde{F}_i}{\partial s}(x^*, 0) = -2s_i|_{s_i=0} = 0$, and for $i = r+1, \dots, k$,

$$\frac{\partial^2 \widetilde{F}_i}{\partial s^2} = \begin{array}{c} \\ \\ \\ \\ \\ \\ (i) \\ \\ \end{array} \begin{pmatrix} & & & & (i) & & & \\ 2\alpha_{1i} & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 2\alpha_{2i} & \dots & 0 & 0 & 0 & \dots & 0 \\ & \dots & & \dots & & \dots & & \\ 0 & 0 & \dots & 2\alpha_{ri} & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & -2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix} = A^i,$$

where A^i is defined in (54). Then from (55) we obtain (52), and from the last equation and (56) we get

$$0 = \sum_{i=r+1}^{k} \gamma^*_i(h) \left(\frac{\partial \widetilde{F}_i}{\partial s^2}(x^*, 0) h_s \right) = \sum_{i=r+1}^{k} \gamma^*_i(h) \left(A^i h_s \right),$$

that is (53).

The following theorem is a specific case of Theorem 3.5.

Theorem 3.6 (Necessary conditions for optimality)**.** *Let x^* be a local minimum of problem* (1), *and let $U \subset \mathbb{R}^n$ be a neighborhood of the point x^*. Assume that $f \in C^2(U, \mathbb{R})$, $g \in C^2(U, \mathbb{R}^m)$, and that $I(x^*) = \{1, \dots, k\}$. Assume also that there exists $\|h\| = 1$, such that the vectors $g'_1(x^*)$,..., $g'_r(x^*)$, $\widetilde{g}''_{r+1}(x^*)h$, ..., $\widetilde{g}''_k(x^*)h$ are linearly independent, and*

$$\begin{aligned} &g'_i(x^*)h = 0, && i = 1, \dots, r, \\ &\widetilde{g}''_i(x^*)[h]^2 = 0, && i = r+1, \dots, k, \end{aligned} \tag{57}$$

where $\widetilde{g}_i$ are defined in (50). *Then there exist multipliers $\lambda^*(h) = (\lambda^*_1(h), \dots, \lambda^*_r(h))$ and $\gamma^*(h) = (\gamma^*_{r+1}(h), \dots, \gamma^*_k(h))$ such that*

$$\frac{\partial L_2}{\partial x}(x^*, h, \lambda^*(h), \gamma^*(h)) = f'(x^*) - \sum_{i=1}^{r} \lambda^*_i(h) g'_i(x^*) - \sum_{i=r+1}^{k} \gamma^*_i(h) \widetilde{g}''_i(x^*) h = 0. \tag{58}$$

Proof. All conditions of Theorem 3.5 are satisfied with $h_s = 0$ and $h_x = h$, where h is defined in (57). Then (52) yields (58), and (53) transforms into $0 \equiv 0$.

Introduce the following definitions. Recall that, to simplify the notation, we assume that $I(x^*) = \{1, \dots, k\}$. Moreover, we assume that the first r row vectors of the partial derivatives of the active constraints $g_i'(x^*)$, $i = 1, \dots, r$, $r < k$, are linearly independent for some $r \geq 1$.

Definition 3.7. *We say that the vector of constraints $g(x) = (g_1(x), \dots g_m(x))$ is 2-regular at the feasible point x^* with respect to the vector $h = (h_x, h_s) \in \mathbb{R}^{n+k}$, $\|h\| = 1$, satisfying (51), if for $I(x^*) = \{1, \dots, k\}$, the mapping $\bar{F}(x, s)$ is 2-regular at $(x^*, 0)$ with respect to the vector h, where $\bar{F}(x, s)$ is defined in (48).*

In the case when $h_s = 0$, Definition 3.7 reduces to the following one.

Definition 3.8 (2-regularity constraint qualification)**.** *Given a point x^* and the active set $I(x^*)$, we say that the 2-regularity constraint qualification holds at the feasible point x^* with respect to a vector $h \in \mathbb{R}^n$ if x^* is strictly feasible or if the row vectors $g_1'(x^*)$, $\dots$, $g_r'(x^*)$, $\widetilde{g}_{r+1}''(x^*)h$, $\dots, \widetilde{g}_k''(x^*)h$ are linearly independent.*

Definition 3.8 can be viewed as the p-regularity constraint qualification (PRCQ) in the general case with $p = 2$.

Definition 3.9. *We say that problem (1) is 2-regular if $\bar{F}(x, s)$ is 2-regular at $(x^*, 0)$ with respect to any element h satisfying (51), where $\bar{F}(x, s)$ is defined in (48).*

Using Definitions 3.7, 3.8, and 3.9, we can reformulate the theorems presented in this section in terms of 2-regular problems.

The case of $p > 2$. Without loss of generality, assume that $I(x^*) = \{1, \dots, k\}$. Then similarly to the transformation described in the beginning of this section, in the case of $p > 2$, we can introduce slack variables s_i^{2q}, $q = [p + 1]/2$, and transform problem (1) into the equivalent form

$$\begin{array}{ll} \underset{(x,\,s)}{\text{minimize}} & f(x) \\ \text{subject to} & F_i(x, s) = g_i(x) - s_i^{2q} = 0, \quad i = 1, \dots, k. \end{array}$$

Assume that $g_1'(x^*)$, $\dots$, $g_{r_1}'(x^*)$ are linearly independent and transform $F_i(x, s)$, $i = r_1 + 1, \dots, m$, in such a way that for some $r_2 \leq r_3 \leq \dots \leq r_{p-1} \leq \dots \leq k$ the following holds

$$\begin{array}{l} \widetilde{F}_i'(x^*, 0) = 0, \quad i = r_1 + 1, \dots, k, \\ \widetilde{F}_i''(x^*, 0) = 0, \quad i = r_2 + 1, \dots, k, \\ \dots \\ \widetilde{F}_i^{(p-1)}(x^*, 0) = 0, \quad i = r_{p-1} + 1, \dots, k. \end{array} \tag{59}$$

The following theorem is a generalization of Theorem 3.5.

Theorem 3.10 (Necessary conditions for optimality)**.** *Let x^* be a local minimum of problem (1), and let $U \subset \mathbb{R}^n$ be a neighborhood of the point x^*. Assume that $f \in C^p(U, \mathbb{R})$, $g \in C^p(U, \mathbb{R}^m)$, and that $I(x^*) = \{1, \dots, k\}$. Assume also that there exist $r_1 \leq r_2 \leq \dots \leq r_{p-1} \leq \dots \leq k$ such that $g_1'(x^*)$, $\dots$, $g_{r_1}'(x^*)$ are linearly independent and (59) holds. Moreover, assume that there is $\|h\| = 1$ such that the vectors $F_1'(x^*, 0), \dots, F_{r_1}'(x^*, 0)$, $\widetilde{F}_{r_1+1}{}''(x^*, 0)h$, $\dots, \widetilde{F}_{r_2}{}''(x^*, 0)h, \dots$ $\widetilde{F}_{r_{p-1}+1}^{(p)}(x^*, 0)[h]^{p-1}, \dots, \widetilde{F}_k^{(p)}(x^*, 0)[h]^{p-1}$ are linearly independent, and*

$$\begin{array}{ll} F_i'(x^*, 0)h = 0, & i = 1, \dots, r_1, \\ \widetilde{F}_i{}''(x^*, 0)[h]^2 = 0, & i = r_1 + 1, \dots, r_2, \\ \dots & \\ \widetilde{F}_i^{(p)}(x^*, 0)[h]^p = 0, & i = r_{p-1} + 1, \dots, k. \end{array} \tag{60}$$

Then there exist multipliers $\lambda_1^*(h), \ldots, \lambda_k^*(h)$ *such that*

$$\begin{pmatrix} f'(x^*) \\ 0 \end{pmatrix} = \sum_{i=1}^{r_1} \lambda_i^*(h) F_i'(x^*, 0) + \sum_{i=r_1+1}^{r_2} \lambda_i^*(h) \widetilde{F}_i''(x^*, 0) h + \ldots + \sum_{i=r_{p-1}+1}^{k} \lambda_i^*(h) \widetilde{F}_i^{(p)}(x^*, 0)[h]^{p-1}. \tag{61}$$

Proof. The proof of this theorem is similar to that of Theorem 3.5.

In the next theorem, we present a generalization of Theorem 3.6 for the case of $p > 2$.

Consider problem (1) and without loss of generality, assume again that $I(x^*) = \{1, \ldots, k\}$. Moreover, assume that the first r_1 row vectors $g_i'(x^*)$, $i = 1, \ldots, r_1$, $r_1 < k$, are linearly independent for some $r_1 \geq 1$.

Consider a special case of problem (1), when functions $g_i(x)$ satisfy the relations (63) similar to ones given in (59):

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \ f(x) \quad \text{subject to} \quad g(x) = (g_1(x), \ldots, g_{r_1}(x), g_{r_1+1}(x), \ldots, g_m(x)) \geq 0, \tag{62}$$

where

$$\begin{array}{ll} g_i'(x^*) = 0, & i = r_1 + 1, \ldots, k, \\ g_i''(x^*) = 0, & i = r_2 + 1, \ldots, k, \\ \ldots & \\ g_i^{(p-1)}(x^*) = 0, & i = r_{p-1} + 1, \ldots, k, \end{array} \tag{63}$$

and $r_1 \leq r_2 \leq \ldots \leq r_{p-1} \leq \ldots \leq k$.

Theorem 3.11 (Necessary conditions for optimality)**.** *Let x^* be a local minimum of problem* (1)*, and let $U \subset \mathbb{R}^n$ be a neighborhood of the point x^*. Assume that $f \in C^p(U, \mathbb{R})$, $g \in C^p(U, \mathbb{R}^m)$, and that $I(x^*) = \{1, \ldots, k\}$. Assume also that there exists $\|h\| = 1$, such that the vectors*

$$g_1'(x^*), \ldots, g_{r_1}'(x^*), \ g_{r_1+1}{}''(x^*)h, \ \ldots, g_{r_2}{}''(x^*)h, \ldots g_{r_{p-1}+1}^{(p)}(x^*)[h]^{p-1}, \ldots, g_k^{(p)}(x^*)[h]^{p-1}$$

are linearly independent, and

$$\begin{array}{ll} g_i'(x^*)h = 0, & i = 1, \ldots, r_1, \\ g_i{}''(x^*)[h]^2 = 0, & i = r_1 + 1, \ldots, r_2, \\ \ldots & \\ g_i^{(p)}(x^*)[h]^p = 0, & i = r_{p-1} + 1, \ldots, k. \end{array} \tag{64}$$

Then there exist multipliers $\lambda_1^*(h), \ldots, \lambda_k^*(h)$, $\lambda_i^*(h) \geq 0$, $i = 1, \ldots, k$, such that

$$f'(x^*) = \sum_{i=1}^{r_1} \lambda_i^*(h) g_i'(x^*) + \sum_{i=r_1+1}^{r_2} \lambda_i^*(h) g_i''(x^*) h + \ldots + \sum_{i=r_{p-1}+1}^{k} \lambda_i^*(h) g_i^{(p)}(x^*)[h]^{p-1}. \tag{65}$$

Proof. The proof of this theorem is similar to that of Theorem 3.5.

Remark. In general case, we cannot guarantee that the Lagrange multipliers $\lambda_i^*(h)$, $i = 1, \ldots, k$, are nonnegative. As an example, consider the problem:

$$\underset{x \in \mathbb{R}^2}{\text{minimize}} \ x_2 \quad \text{subject to} \quad x_1^3 - x_2 \geq 0, \quad x_2^3 - x_1^{12} \geq 0,$$

One can verify that $x^* = (0, 0)$ is a minimizer with $\lambda^* = (0, -1)$.

4. Sufficient conditions for optimality

In this section, we derive new sufficient conditions for optimality.

4.1. **The completely degenerate case for an arbitrary $p \geq 2$.** In this section we propose new sufficient optimality conditions for the completely degenerate case (19). The necessary conditions for the case presented in Theorem 4.1 are given in Theorem 2.12.

Theorem 4.1 (The pth order sufficient conditions for optimality in the case of an even p and $\hat{H}_p(x^*) = \{0\}$). *Let x^* be a feasible point for problem* (1)*, let $U \subset \mathbb{R}^n$ be a neighborhood of the point x^*, and let* (19) *hold with an even p. Let the set $\widetilde{H}_p(x^*, 0)$ be defined by* (25)*, and let $F_i(x, s)$, $i = 1, \ldots, m$, be defined by* (22)*. Suppose that $f \in C^2(U, \mathbb{R})$ and that $g \in C^{p+1}(U, \mathbb{R}^m)$. Suppose also that for any $h = (h_x, h_s) \in \widetilde{H}_p(x^*, 0)$, $h \neq 0$, the vectors $F_i^{(p)}(x^*, 0)[(h_x, h_s)]^{p-1}$, $i \in I(x^*)$, are linearly independent. Moreover, assume that for any $h = (h_x, h_s) \in \widetilde{H}_p(x^*, 0)$, $h \neq 0$, there exist $\alpha > 0$ and a Lagrange multiplier vector $\lambda^*(h)$, with components $\lambda_i^*(h) \geq 0$, $i \in I(x^*)$, such that the necessary conditions* (27)–(28) *hold and*

$$\frac{\partial^2 \bar{L}_p}{\partial x^2}(x^*, h, \tilde{\lambda}^*(h))[h_x]^2 \geq \alpha \|h_x\|^2, \quad \tilde{\lambda}^*(h) = \frac{2\lambda^*(h)}{p(p+1)}. \tag{66}$$

Then x^* is an isolated solution to problem (1).

Proof. Let $z = (x, s)$ and $z^* = (x^*, 0)$. Introduce the function:

$$\begin{aligned} \mathfrak{L}_p(z, h, \lambda(h)) &= f(x) - \langle \lambda(h), F^{(p-1)}(z)[h_x, h_s]^{p-1} \rangle = \\ &= f(x) - \sum_{i \in I(x^*)} \lambda_i(h) F_i^{(p-1)}(z)[h_x, h_s]^{p-1}. \end{aligned}$$

By the assumption of the theorem (27) and (28) hold, which implies that

$$\frac{\partial \mathfrak{L}_p}{\partial z}(z^*, h, \lambda^*(h)) = 0$$

or that (17) is satisfied. Moreover, it follows from (28) and (66) that

$$\frac{\partial^2 \mathfrak{L}_p}{\partial z^2}\left(z^*, h, \tilde{\lambda}^*(h)\right)[h]^2 = \begin{bmatrix} (\bar{L}_p)''_{xx} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} h_x \\ h_s \end{bmatrix}^2 = \frac{\partial^2 \bar{L}_p}{\partial x^2}[h_x]^2 \geq \alpha \|h_x\|^2.$$

Since for any $h = (h_x, h_s) \in \widetilde{H}_p(x^*, 0)$, $h \neq 0$, the vectors $F_i^{(p)}(x^*, 0)[(h_x, h_s)]^{p-1}$, $i \in I(x^*)$, are linearly independent, there exists α_1 such that

$$\alpha \|h_x\|^2 \geq \alpha_1 \|h\|^2 \quad \text{for all} \quad h = (h_x, h_s) \in \widetilde{H}_p(x^*, 0).$$

Thus, the sufficient conditions given in Theorem 2.8 hold for problem (24) at z^*.

Then

$$\frac{\partial^2 \mathfrak{L}_p}{\partial z^2}\left(z^*, h, \tilde{\lambda}^*(h)\right)[h]^2 \geq \alpha_1 \|h\|^2,$$

and by sufficient conditions given in Theorem 2.8, $z^* = (x^*, 0)$ is a local minimizer to problem (24). Hence, by Lemma 2.10, x^* is an isolated solution to problem (1).

Theorem 4.1 gives new sufficient conditions for problem (1), but its statement contains functions $F_i(x, s)$ introduced in (23) and the set $\widetilde{H}_p(x^*, 0)$ defined in (25).

The next theorem is a reformulation of Theorem 4.1.

Theorem 4.2 (The pth order sufficient conditions for optimality). *Let x^* be a feasible point for problem* (1)*, let $U \subset \mathbb{R}^n$ be a neighborhood of the point x^*, and let* (19) *hold with an even p. Suppose that $f \in C^2(U, \mathbb{R})$ and that $g \in C^{p+1}(U, \mathbb{R}^m)$. Suppose also that for any $h \in \bar{H}_p(x^*) \backslash \{0\}$, vectors $g_i^{(p)}(x^*)[h]^{p-1}$, $i \in I(x^*)$, are linearly independent, and there exists a vector $\lambda^*(h) \geq 0$ and $\alpha > 0$ such that* (37)–(39) *hold and*

$$\frac{\partial^2 L_p}{\partial x^2}(x^*, h, \tilde{\lambda}^*(h))[h]^2 \geq \alpha \|h\|^2, \quad \tilde{\lambda}^*(h) = \frac{2\lambda^*(h)}{p(p+1)}. \tag{67}$$

Then x^* is an isolated solution to problem (1).

The following theorem complements the necessary conditions stated in Theorem 3.4. We use the notation that is introduced in Section 3.1.

Theorem 4.3. *Let U be some neighborhood of $x^* \in \mathbb{R}^n$, $f \in C^2(U)$, $g \in C^{p+1}(U)$, and* (19) *and* (40) *hold. Assume that $g(x)$ is p–regular at x^*. Assume also that for any $h \in K(x^*)$, there exist $\beta > 0$ and $\lambda^*(h) \geq 0$ such that* (41) *holds and*

$$\langle \frac{\partial^2 L_p}{\partial x^2}(x^*, h, \tilde{\lambda}^*(h))h, h\rangle \geq \beta \|h\|^2, \quad \tilde{\lambda}^*_p(h) = \frac{2\lambda^*_p}{p(p+1)}.$$

Then x^* is an isolated local minimizer of (1)*.*

Proof. Assume on the contrary that x^* is not a local minimizer. Then there exists a sequence $\{x_k\} \to x^*$ such that $f(x_k) < f(x^*)$ and $g(x_k) \geq 0$. Consider a sequence (we use the same notation for the sequence and its convergent subsequence) $\left\{\frac{x_k - x^*}{\|x_k - x^*\|}\right\}$ that converges to some element $\tilde{h}$. Then

$$x_k = x^* + \|x_k - x^*\|\tilde{h} + w(x^k) = x^* + t\tilde{h} + \xi_k,$$

$\|w(x^k)\| = o(\|x_k - x^*\|)$.

Observe that $\tilde{h} \in A_g(x^*)$ and consider two cases:

(1) If $\langle f'(x^*), \tilde{h}\rangle > 0$, then

$$f(x_k) = f(x^*) + \left\langle f'(x^*), \|x_k - x^*\|\tilde{h}\right\rangle + \xi(x_k) > f(x^*),$$

which is a contradiction, so this case does not hold.

(2) If $\langle f'(x^*), \tilde{h}\rangle = 0$, then there exists $\lambda(\tilde{h}) \geq 0$ such that

$$\begin{aligned}
f(x_k) - f(x^*) \quad &\geq \quad f(x_k) - f(x^*) - \frac{(p-1)!}{t^{p-1}} \sum_{i \in I_0(\tilde{h})} \lambda^*_i(\tilde{h}) g_i(x_k) = \\
&= \quad f'(x^*)(x_k - x^*) + \frac{1}{2} f''(x^*)(x_k - x^*)^2 - \\
&\quad -\frac{1}{pt^{p-1}} \sum_{i \in I_0(\tilde{h})} \lambda^*_i(\tilde{h}) g_i^{(p)}(x^*)[x_k - x^*]^p - \\
&\quad -\frac{1}{p(p+1)t^{p-1}} \sum_{i \in I_0(\tilde{h})} \lambda^*_i(\tilde{h}) g_i^{(p+1)}(x^*)[x_k - x^*]^{p+1} + o((x_k - x^*)^2) = \\
&= \quad \left\langle f'(x^*) - \sum_{i \in I_0(\tilde{h})} \lambda^*_i(\tilde{h}) g_i^{(p)}(x^*)[\tilde{h}]^{p-1}, \xi_k \right\rangle + \\
&\quad +\frac{1}{2}\left\langle f''(x^*) - \frac{2}{p(p+1)} \sum_{i \in I_0(\tilde{h})} \lambda^*_i(\tilde{h}) g_i^{(p+1)}(x^*)[\tilde{h}]^{p-1}, (t\tilde{h})^2 \right\rangle + o(t^2) \geq \\
&\geq \quad \frac{\alpha}{2}\|t\tilde{h}\|^2 + o(t^2) > 0,
\end{aligned}$$

which contradicts the assumption $f(x_k) < f(x^*)$. Hence, x^* is a strict local minimizer.

4.2. **General case.** In the next theorem, we present new sufficient conditions for optimality. The first order necessary conditions for the corresponding classes of nonregular problems are given in Theorem 3.5.

Theorem 4.4 (Sufficient conditions for optimality)**.** *Let x^* be a feasible point for problem* (1)*, and let $U \subset \mathbb{R}^n$ be a neighborhood of the point x^*. Suppose that $f \in C^2(U, \mathbb{R})$ and that $g \in C^2(U, \mathbb{R}^m)$. Suppose also that for any $h = (h_x, h_s)$, $\|h\| = 1$, satisfying* (51)*, $\bar{F}(x, s)$ is 2-regular at $(x^*, 0)$ with respect to the element h. Moreover, assume that for any $h = (h_x, h_s)$,*

$\|h\| = 1$, $h_x \neq 0$, satisfying (51)*, there exist multipliers $\lambda^*(h) = (\lambda_1^*(h), \ldots, \lambda_r^*(h))$ and $\gamma^*(h) = (\gamma_{r+1}^*(h), \ldots, \gamma_k^*(h))$ such that* (52)–(53) *hold, and*

$$\frac{\partial^2 L_2}{\partial x^2}\left(x^*, h, \lambda^*(h), \frac{\gamma^*(h)}{3}\right)[h_x]^2 \geq \alpha \|h_x\|^2. \tag{68}$$

Then x^* is an isolated local minimum of problem (1).

Proof. The proof of this theorem is similar to that of Theorem 4.1. Namely, the sufficient conditions given in Theorem 2.8 hold for problem (47) at $(x^*, 0)$. Thus x^* is an isolated solution to (1).

The following theorem is a specific case of Theorem 4.4.

Theorem 4.5 (Sufficient conditions for optimality)**.** *Let x^* be a feasible point for problem* (1)*, and let $U \subset \mathbb{R}^n$ be a neighborhood of the point x^*. Suppose that $f \in C^2(U, \mathbb{R})$ and that $g \in C^2(U, \mathbb{R}^m)$. Suppose also that for any h, $\|h\| = 1$, satisfying* (57)*, the vectors $g_1'(x^*), \ldots, g_r'(x^*)$, $g_{r+1}''(x^*)h, \ldots, g_k''(x^*)h$ are linearly independent. Moreover, assume that for any h, $\|h\| = 1$, $h_x \neq 0$, satisfying* (51)*, there exist multipliers λ_i^* and $\gamma_i^*(h)$ such that* (58) *holds, and*

$$\frac{\partial^2 L_2}{\partial x^2}\left(x^*, h, \lambda^*(h), \frac{\gamma^*(h)}{3}\right)[h]^2 \geq \alpha \|h\|^2. \tag{69}$$

Then x^* is an isolated local minimum of problem (1).

Remark 4.1. Observe that we cannot drop the assumption $h_x \neq 0$ in both Theorem 4.4 and Theorem 4.5. For example, consider the problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}}\ f(x) = -x_1^2 - x_2^2 \quad \text{subject to} \quad -x_1 \geq 0, \quad -x_2 \geq 0.$$

Observe that assumptions of Theorem 4.4 and Theorem 4.5 are *not* satisifed with $h_x \neq 0$ at $x^* = 0$, and indeed 0 is not a local minimizer.

In the next theorem, we present new sufficient conditions for optimality in the general case of $p \geq 2$. The first order necessary conditions for the corresponding classes of nonregular problems are given in Theorem 3.11.

Recall a specific form of the optimization problem that we introduce earlier in (62):

$$\underset{x \in \mathbb{R}^n}{\text{minimize}}\ f(x) \quad \text{subject to} \quad g(x) = (g_1(x), \ldots, g_{r_1}(x), g_{r_1+1}(x), \ldots, g_m(x)) \geq 0,$$

where $g(x)$ satisfies (63):

$$\begin{array}{l} g_i'(x^*) = 0, \quad i = r_1 + 1, \ldots, k, \\ g_i''(x^*) = 0, \quad i = r_2 + 1, \ldots, k, \\ \ldots \\ g_i^{(p-1)}(x^*) = 0, \quad i = r_{p-1} + 1, \ldots, k, \end{array}$$

and $r_1 \leq r_2 \leq \ldots \leq r_{p-1} \leq \ldots \leq k$.

Introduce the following notation: $\quad I(x^*) = \{1, \ldots, k\}$,

$$\begin{array}{rcl} A_g(x^*) = \{h \in \mathbb{R}^n & | & g_i'(x^*)[h] \geq 0, \quad i = 1, \ldots, r_1, \\ & & g_i''(x^*)[h]^2 \geq 0, \quad i = r_1 + 1, \ldots, r_2, \\ & & \vdots \\ & & g_i^{(p)}(x^*)[h]^p \geq 0, \quad i = r_{p-1} + 1, \ldots, k\}, \end{array}$$

$$\begin{array}{rcl} I_0^j(h) & = & \{i \in \{r_{j-1} + 1, \ldots, r_j\} \mid g_i^{(j)}(x^*)[h]^j = 0, \quad h \in A_g(x^*)\}, \quad j = 1, \ldots, p, \\ K(x^*) & = & A_g(x^*) \bigcap \mathrm{Ker} f'(x^*). \end{array}$$

Definition 4.6. *Mapping $g(x)$ is called p-regular at x^* with respect to vector h if the vectors*

$$\{g_i'(x^*), i \in I_0^1(h); g_i''(x^*)[h], i \in I_0^2(h); \ldots, g_i^{(p)}(x^*)[h]^{p-1}, i \in I_0^p(h)\}$$

are linearly independent.

Definition 4.7. *Mapping $g(x)$ is called p-regular at x^* if $g(x)$ is p-regular at x^* with respect to any vector h such that*

$$\{I_0^1(h) \cup I_0^2(h) \cup \ldots \cup I_0^p(h)\} \neq \emptyset.$$

Introduce a generalization of the Lagrange function:

$$L_p(x,h,\lambda(h)) = f(x) - \sum_{i \in I_0^1(h)} \lambda_i^1(h) g_i(x) - \sum_{i \in I_0^2(h)} \lambda_i^2(h) g_i'(x)[h] - \ldots - \sum_{i \in I_0^p(h)} \lambda_i^p(h)(g_i^{(p-1)}(x)[h]^{p-1},$$

where $\lambda(h) = (\lambda^1(h), \ldots, \lambda^p(h))$ and $\lambda^j(h) = (\lambda_i(h))_{i \in I_0^j(h)}$, $j = 1, \ldots, p$.

Theorem 4.8 (Sufficient conditions for optimality)**.** *Let x^* be a feasible point for problem* (1)*, and let $U \subset \mathbb{R}^n$ be a neighborhood of the point x^*. Suppose that $f \in C^2(U,\mathbb{R})$ and that $g \in C^{p+1}(U,\mathbb{R}^m)$. Assume that $A_g(x^*) \subset A_f(x^*)$. Assume also that $g(x)$ is p-regular at x^* and that for any $h \in K(x^*)$ there exists $\lambda^*(h)$, $\lambda^*(h) \geq 0$, and $\beta > 0$ such that*

$$\frac{\partial L_p}{\partial x}(x^*, h, \lambda^*(h)) = 0$$

and

$$\left\langle \frac{\partial^2 L_p}{\partial x^2}(x^*, h, \tilde{\lambda}^*(h))h, h \right\rangle \geq \beta \|h\|^2,$$

where

$$\tilde{\lambda}^*(h) = \left(\lambda^{1*}(h), \frac{1}{3}\lambda^{2*}(h), \ldots, \frac{2}{i(i+1)}\lambda^{i*}(h), \ldots, \frac{2}{p(p+1)}\lambda^{p*}(h)\right).$$

Then x^* is an isolated local minimum of problem (1)*.*

The proof is similar to one of Theorem 4.3.

5. Mangasarian-Fromovitz constraint qualification (MFCQ) and 2-regularity condition

In this section, we present a new result stating that the 2-regularity condition given in Definition 3.7 is weaker than the MFCQ. Namely, there are examples in which the MFCQ does not hold, while the 2-regularity does. One such example is Example 1 given in the next section. On the other hand, any problem that satisfies the MFCQ is 2-regular with respect to some vector h in the sense of Definition 3.7. To prove this statement, we will construct the vector h satisfying Definition 3.7.

Theorem 5.1. *Assume that the Mangasarian-Fromovitz constraint qualification holds at a feasible point x^* for the vector of constraints $g(x) = (g_1(x), \ldots g_m(x))$ in problem* (1)*. Then there exists a vector h such that $g(x)$ is 2-regular at x^* with respect to the vector h in the sense of Definition 3.7.*

Proof. Consider constraints of problem (1)

$$g_i(x) \geq 0, \quad i = 1, \ldots, m, \tag{70}$$

and assume that the MFCQ holds at x^*, that is, there exists a vector ν such that

$$g_i'(x^*)\nu > 0 \quad \text{for all} \quad i \in I(x^*). \tag{71}$$

Conditions (71) imply that the vectors $g_i'(x^*) \neq 0$, $i = 1, \ldots, m$.

Without loss of generality, assume that $I(x^*) = \{1, \ldots, k\}$. Moreover, assume that the vectors $g_i'(x^*)$, $i = 1, \ldots, r$, $r < k$, are linearly independent for some $r \geq 1$.

Since the vectors $g_1'(x^*), \ldots, g_r'(x^*)$ are linearly independent, then the vectors $g_i'(x^*)$, $i = r+1, \ldots, k$, can be represented as a linear combination of $g_1'(x^*), \ldots, g_r'(x^*)$ with some coefficients $\alpha_{1i}, \ldots, \alpha_{ri}$, which are not all equal to zero. By using the constraint transformation described in the beginning of Section 3.2 we get

$$\begin{aligned} F_i(x,s) &= g_i(x) - s_i^2, \qquad\qquad i = 1, \ldots, r, \\ \widetilde{F}_i(x,s) &= F_i(x,s) - \alpha_{1i}F_1(x,s) - \ldots - \alpha_{ri}F_r(x,s) = \\ &= g_i(x) - s_i^2 - \alpha_{1i}(g_1(x) - s_1^2) - \ldots - \alpha_{ri}(g_r(x) - s_r^2) = \\ &= g_i(x) - s_i^2 - (\alpha_{1i}g_1(x) + \ldots + \alpha_{ri}g_r(x)) + (\alpha_{1i}s_1^2 + \ldots + \alpha_{ri}s_r^2), \\ & i = r+1, \ldots, k \end{aligned} \tag{72}$$

where $\frac{\partial \widetilde{F}_i}{\partial x}(x^*, 0) = 0, \quad i = r+1, \ldots, k.$

Introduce the mapping

$$\bar{F}(x,s) = \begin{pmatrix} F_1(x,s) \\ \ldots \\ F_r(x,s) \\ \widetilde{F}_{r+1}(x,s) \\ \ldots \\ \widetilde{F}_k(x,s) \end{pmatrix}. \tag{73}$$

Define the vector $h = (h_x, h_s)$, $h_x \in \mathbb{R}^n$, $h_s = (h_{s1}, \ldots, h_{sk})$, as follows: $h_x = 0$,

$$(h_{s_i})^2 = \begin{cases} g_i'(x^*)\nu, & i = 1, \ldots, r \\ \alpha_{1i}h_{s1}^2 + \ldots + \alpha_{ri}h_{sr}^2, & i = r+1, \ldots, k \end{cases}, \tag{74}$$

where ν is defined in (71). We will show that the vector h satisfies (51). First, we observe that $(h_{s_i})^2 > 0$ for all $i = 1, \ldots, k$. Namely, by (71), $g_i'(x^*)\nu > 0$, $i = 1, \ldots, r$. Moreover, for $i = r+1, \ldots, k$, (71), (72), (74), and $\widetilde{F}_i'(x^*, 0) = 0$ yield that

$$\alpha_{1i}h_{s1}^2 + \ldots + \alpha_{ri}h_{sr}^2 = \alpha_{1i}g_1'(x^*)\nu + \ldots + \alpha_{ri}g_r'(x^*)\nu = g_i'(x^*)\nu > 0.$$

Thus $(h_{s_i})^2 > 0$ for all $i = 1, \ldots, k$ and, hence, $h \neq 0$.

Since $h_x = 0$ and $\bar{F}_s'(x^*, 0) = 0$, then $h \in \operatorname{Ker} \bar{F}'(x^*, 0)$. By definition of the vector h_s given in (74), $h \in \operatorname{Ker}^2 P^\perp \bar{F}''(x^*, 0)$. Thus

$$h \in \{\operatorname{Ker} \bar{F}'(x^*, 0) \cap \operatorname{Ker}^2 P^\perp \bar{F}''(x^*, 0)\}.$$

Hence, the vector h satisfies (51).

Consider the 2-factor-operator, which is defined by (2.6),

$$\begin{aligned} \Psi_2(x^*, 0) &= \bar{F}'(x^*, 0) + P^\perp \bar{F}''(x^*, 0)h \\ &= \begin{pmatrix} g_1'(x^*) & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ \ldots & & \ldots & & \ldots & & & \\ g_r'(x^*) & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ 0 & 2\alpha_{1r+1}h_{s_1} & \ldots & 2\alpha_{rr+1}h_{s_r} & -2h_{s_{r+1}} & 0 & \ldots & 0 \\ 0 & 2\alpha_{1r+2}h_{s_1} & \ldots & 2\alpha_{rr+2}h_{s_r} & 0 & -2h_{s_{r+2}} & \ldots & 0 \\ \ldots & & \ldots & & \ldots & & & \\ 0 & 2\alpha_{1k}h_{s_1} & \ldots & 2\alpha_{rk}h_{s_r} & 0 & \ldots & 0 & -2h_{s_k} \end{pmatrix} \end{aligned}$$

Since $(h_{s_i})^2 > 0$, $i = r+1, \ldots, k$, and the vectors $g_1'(x^*), \ldots, g_r'(x^*)$ are linearly independent, then the matrix $\Psi_2(x^*, 0)$ has the full rank. Thus, by Definition 2.7, the mapping $\bar{F}(x,s)$ is 2-regular at $(x^*, 0)$ with respect to the element h and hence, $g(x)$ is 2-regular at x^* with respect to the vector h in the sense of Definition 3.7.

6. Examples

Example 6.1. Consider the following problem

$$\begin{array}{ll} \underset{x\in\mathbb{R}^2}{\text{minimize}} & f(x) \\ \text{subject to} & -x_1^4 + x_2 \geq 0, \\ & 3x_1^4 - x_2 \geq 0. \end{array} \tag{75}$$

Assume that $x^* = (0,0)^T$ is a local minimum of problem (75). We will show that all conditions of Theorem 3.10 are satisfied for (75). In our notation, $r_1 = 1$, $p = 4$, $r_2 = r_3 = 1$ and $k = 2$.

By introduction of the slack variables we obtain

$$\begin{array}{l} F_1(x,s) = -x_1^4 + x_2 - s_1^4 = 0 \\ F_2(x,s) = 3x_1^4 - x_2 - s_2^4 = 0 \end{array}$$

Transform the mapping $F(x,s) = (F_1(x,s), F_2(x,s))$ into the following one

$$\begin{array}{l} F_1(x,s) = -x_1^4 + x_2 - s_1^4 = 0 \\ \widetilde{F}_2(x,s) = F_2(x,s) - \alpha_{12}F_1(x,s) = 2x_1^4 - s_1^4 - s_2^4 = 0 \end{array}$$

where $\alpha_{12} = -1$. Observe that the conditions of Theorem 3.10 hold with $h = (h_1, h_2, h_3, h_4) = \frac{1}{1+\sqrt{2}}(1, 0, \sqrt[4]{2}, 0)$. Hence, by Theorem 3.10, we get that there exist multipliers $\lambda_1^*(h)$ and $\lambda_2^*(h)$ such that

$$\begin{pmatrix} f'(x^*) \\ 0 \end{pmatrix} = \lambda_1^*(h)F_1'(x^*) + \lambda_2^*(h)(\widetilde{F}_2^{(4)}(x^*)h^3)^T \tag{76}$$

or

$$\begin{pmatrix} f'_{x_1}(x^*) \\ f'_{x_2}(x^*) \\ 0 \\ 0 \end{pmatrix} = \lambda_1^*(h)\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \frac{1}{1+\sqrt{2}}\lambda_2^*(h)\begin{pmatrix} 48 \\ 0 \\ -24\sqrt[4]{2^3} \\ 0 \end{pmatrix}.$$

The last equality yields $\lambda_2^*(h) = 0$.

Next example illustrates the sufficient conditions given in Theorem 4.2.

Example 6.2. Isoperimetrical problem in Calculus of Variations. A slightly modified version of this example was used in [6] to illustrate the necessary optimality conditions. Here, we verify that the sufficient conditions given in Theorem 4.2 hold with $p = 4$ and $x^* = (0,0)$.

Consider the problem of minimizing the functional

$$J_0[y] = \int_{-3/2}^{3/2} y^2 dx \tag{77}$$

subject to the constraints

$$\int_{-3/2}^{3/2} ((y')^4 - y^4)dx \geq 0, \quad y(-3/2) = y(3/2) = 0. \tag{78}$$

It can be shown that $y^*(x) = 0$ is a 4–regular solution of the problem (77)–(78).

We consider a discretization of the problem (77)–(78):

$$\underset{x\in\mathbb{R}^2}{\text{minimize}}\, f(x_1,x_2) = x_1^2 + x_2^2 \quad \text{subject to} \quad g(x_1,x_2) = 2(x_1 - x_2)^4 - x_2^4 \geq 0. \tag{79}$$

We have

$$\bar{H}_4(x^*) = \{h \in \mathbb{R}^2 \,|\, g^{(4)}(x^*)[h]^4 \geq 0\} = \{h \in \mathbb{R}^2 \,|\, (\sqrt[4]{2}h_1 - (\sqrt[4]{2}+1)h_2)(\sqrt[4]{2}h_1 - (\sqrt[4]{2}-1)h_2) \geq 0\}.$$

For any vector h in $\bar{H}_4(x^*)\backslash\{0\}$, the vector $g^{(4)}(x^*)[h]^3$ is nonzero. Moreover, in this example, (37)–(39) hold with $\lambda^*(h) = 0$ for any h. Furthermore, since

$$\frac{\partial^2 L_2}{\partial x^2}(x)[h]^2 = f''(x)[h]^2 = 2h_1^2 + 2h_2^2 \geq 2\|h\|^2,$$

then (67) holds at $x^* = (0,0)^T$ with any h. Thus, by Theorem 4.2, $(0,0)^T$ is the isolated solution to problem (79).

The following example illustrates an application of Theorem 3.11 and Theorem 4.8.

Example 6.3. Consider the problem:

$$\begin{array}{ll} \underset{x\in\mathbb{R}^2}{\text{minimize}} & f(x) = x_2 \\ \text{subject to} & g_1(x) = x_1 + x_2 - x_3 - x_3^2 + x_2^3 \geq 0, \\ & g_2(x) = (-x_1 + x_2 + x_3)(x_1^4 + x_2^4 + x_3^4) + x_1^6 - |x_1|^{13/2} \geq 0 \end{array} \tag{80}$$

The point $x^* = (0,0,0)$ is a local minimum of (80).

We will show that all the conditions of Theorem 3.11 hold at x^* with $p = 5$.

Both constraints, $g_1(x)$ and $g_2(x)$, are active at x^*, so $I(x^*) = \{1,2\}$, $k = 2$, and $r_1 = 1$. Moreover,

$$g_2{}'(x^*) = 0, \quad g_2{}''(x^*) = 0, \quad g_2'''(x^*) = 0, \quad g_2^{(4)}(x^*) = 0,$$

so (63) hold at x^* with $\widetilde{g}_2 = g_2$ and with $r_1 = r_2 = r_3 = r_4 = 1$.

We have

$$g_1{}'(x^*) = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad \widetilde{g}_2^{(5)}(x^*)[h]^4 = 24 \begin{pmatrix} -5h_1^4 + 4h_1^3h_2 + 4h_1^3h_3 - h_2^4 - h_3^4 \\ h_1^4 - 4h_1h_2^3 + 5h_2^4 + 4h_2^3h_3 + h_3^4 \\ h_1^4 + h_2^4 - 4h_1h_3^3 + 4h_2h_3^3 + 5h_3^4 \end{pmatrix} \tag{81}$$

As is easy to verify, the vector $h = (h_1,h_2,h_3)^T = (1,0,1)^T$ satisfies equations (64), which, in this example, are the following:

$$\begin{array}{rl} h_1 + h_2 - h_3 & = 0 \\ (-5h_1^4 + 4h_1^3h_2 + 4h_1^3h_3 - h_2^4 - h_3^4)h_1 + & \\ +(h_1^4 - 4h_1h_2^3 + 5h_2^4 + 4h_2^3h_3 + h_3^4)h_2 + & \\ +(h_1^4 + h_2^4 - 4h_1h_3^3 + 4h_2h_3^3 + 5h_3^4)h_3 & = 0 \end{array}$$

For $h = (h_1,h_2,h_3)^T = (1,0,1)^T$, vectors

$$g_1'(x^*) = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \quad \text{and} \quad \widetilde{g}_2^{(5)}(x^*)[h]^4 = 48 \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

are linearly independent. Thus, all conditions of Theorem 3.11 are satisfied and, by Theorem 3.11, there exist multipliers $\lambda_1^* = 1/2$ and $\lambda_2^* = 1/96$ such that (65) holds, i.e.,

$$f'(x^*) = \frac{1}{2}g_1'(x^*) + \frac{1}{96}\widetilde{g}_2^{(5)}(x^*)[h]^4, \qquad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + \frac{48}{96}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

Now, we will illustrate sufficient conditions stated in Theorem 4.8.

In this example,

$$\begin{array}{rcl} A_g(x^*) & = & \{h \in \mathbb{R}^3 \mid g_1'(x^*)[h] \geq 0,\ g_2^{(5)}(x^*)[h]^5 \geq 0\} = \\ & = & \{h \in \mathbb{R}^3 \mid h_1 + h_2 - h_3 \geq 0,\ -h_1 + h_2 + h_3 \geq 0\} \subset \\ & \subset & A_f = \{h \in \mathbb{R}^3 \mid h_2 \geq 0\}. \end{array}$$

Moreover,

$$\mathrm{Ker} f'(x^*) \cap A_g(x^*) = \{h \in \mathbb{R}^3 \mid h_2 = 0,\ h_1 = h_3\}.$$

Then

$$f'(x^*) = \lambda_1 g_1'(x^*) + 48\lambda_2 h_1^4 \tilde{g}_2^{(5)}(x^*)[h]^4, \qquad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + 48\lambda_2 h_1^4 \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

where $\lambda_1 = 1/2$ and $\lambda_2 = 1/(96h_1^4)$.

Furthermore, there exists $\beta \in (0, 1/8)$ such that

$$\left\langle \frac{\partial^2 L_5}{\partial x^2}(x^*, h, \tilde{\lambda}^*(h))h, h \right\rangle = 1/4 h_1^2 \geq \beta \|h\|^2.$$

Hence, all assumptions of Theorem 4.8 hold and x^* is a local minimizer of problem (80).

The following example illustrates an application of Theorem 3.10.

Example 6.4. Consider the problem:

$$\begin{array}{ll} \underset{x\in\mathbb{R}^3}{\text{minimize}} & f(x) = x_2^2 + x_3 \\ \text{subject to} & g_1(x) = -x_1^2 + x_2^2 - x_3^2 - x_2^5 - x_1^3 x_3^2 + |x_1|^{13/2} \geq 0, \\ & g_2(x) = -x_1^2 + x_2^2 - x_3^2 - 2x_2x_3 - x_3^5 - x_2^2 x_3^3 + |x_2|^{13/2} \geq 0 \\ & g_3(x) = 2x_1^2 - 2x_2^2 + 2x_3^2 + 2x_2x_3 + x_1^4 + x_2^4 + x_3^4 - x_1^5 - x_2 x_3^6 + |x_3|^{13/2} \geq 0 \end{array} \tag{82}$$

The point $x^* = (0, 0, 0)$ is a local minimum of (82) and $I(x^*) = \{1, 2, 3\}$. By introducing s_i^4, $i = 1, 2, 3$, we reduce the inequality constraints to the equality ones:

$$\begin{array}{l} F_1(x, s) = -x_1^2 + x_2^2 - x_3^2 - x_2^5 - x_1^3 x_3^2 - s_1^4 = 0, \\ F_2(x, s) = -x_1^2 + x_2^2 - x_3^2 - 2x_2x_3 - x_3^5 - x_2^2 x_3^3 - s_2^4 = 0, \\ F_3(x, s) = 2x_1^2 - 2x_2^2 + 2x_3^2 + 2x_2x_3 + x_1^4 + x_2^4 + x_3^4 - x_1^5 - x_2 x_3^6 - s_3^4 = 0. \end{array}$$

By applying the transformation described in Section 3.2 we can reduce the constraints to the equivalent form:

$$\begin{array}{l} F_1(x, s) = -x_1^2 + x_2^2 - x_3^2 - x_2^5 - x_1^3 x_3^2 - s_1^4 = 0, \\ F_2(x, s) = -x_1^2 + x_2^2 - x_3^2 - 2x_2x_3 - x_3^5 - x_2^2 x_3^3 - s_2^4 = 0, \\ \tilde{F}_3(x, s) = x_1^4 + x_2^4 + x_3^4 - x_1^5 - x_2 x_3^6 - x_2^5 - x_1^3 x_3^2 - x_3^5 - x_2^2 x_3^3 - s_1^4 - s_2^4 - s_3^4 = 0. \end{array} \tag{83}$$

In the notation of Theorem 3.10, we have $k = 3$, $r_1 = 0$, $r_2 = 2$, $r_3 = 2$, and $p = 4$. Moreover, for $h = (1, 1, 0, 1, 1, 0)^T$, the vectors

$$F_1''(x^*, 0)h, \quad F_2''(x^*, 0)h, \quad \text{and} \quad \tilde{F}_3^{(4)}(x^*, 0)[h]^3$$

are linearly independent. In addition, the equalities (59) and (60) hold. Hence, all assumptions of Theorem 3.10 are satisfied and there exist multipliers $\lambda_1^* = 1/2$, $\lambda_2^* = -1/2$ and $\lambda_3^* = 0$ such that (61) holds.

7. Comparison with other work and concluding remarks

In this paper we extended the results from [5] and [6] to new classes of nonregular optimization problems. The closest results to ours are those obtained in the work of Izmailov [14, 15] and of Izmailov and Solodov [16]. Papers [15] and [16] consider the case of $p = 2$ only, while we are considering the case of $p \geq 2$. Even for the case of $p = 2$, our results are derived under assumptions that are weaker than those in [14, 15, 16]. For example, our Theorems 2.13 can be viewed as a generalization of the results obtained in [14]. Namely, there is an additional restrictive assumption in [14] that the objective function and its derivatives up to some order are equal to zero, i.e.,

$$f^{(k)}(x^*) = 0, \quad k = 1, 2, \ldots, q-1.$$

As follows from consideration in [14], k has to be greater than or equal to 1. We do not make this assumption, so our theorems cover classes of problems that are not subsumed by the theorems proposed in [14]. Another additional assumption in [14] is one on the constraint functions, which is

$$g_i^{(k)}(x^*) = 0, \quad k = 1, \ldots, p_i - 1.$$

Having such an assumption would restrict classes of nonregular problems under consideration and would not cover the general case that is considered in our paper.

Moreover, papers [14, 15, 16] present only necessary conditions for optimality and do not consider sufficient ones. In addition, Theorem 3.4 covers the case of any $p \geq 2$ and also subsumes the case when the set $I_0(h)$ is empty.

The optimality conditions given in papers [1, 11, 12] can be used to analyze some degenerate optimization problems in case $p = 2$. However, those optimality conditions cannot be applied in the case of $p > 2$, which is the main focus of this paper. In addition, necessary conditions given in [12] allow the coefficient λ_0 of the objective function to be zero. In the contrast, the conditions given in our paper provide $\lambda_0 \neq 0$. Paper [18] considers the case of $p > 2$. However, it requires the functions to be $(2p-1)$-times continuously differentiable in the case of degeneracy of order p. At the same time, in this paper, we only require the functions to be $(p+1)$-times continuously differentiable.

Our results will also be true in the Banach spaces, but under some additional assumptions (see, for example, [8]).

References

[1] Ben-Tal, A., (1980), Second order and related extremality conditions in nonlinear programming, J. Optim. Theory Appl., 31, pp.143–165.

[2] Brezhneva, O.A., Izmailov, A.F., Tret'yakov, A.A. and Khmura, A., (2000), An approach to finding singular solutions to a general system of nonlinear equations, Comput. Math. and Math. Phys., 40, pp. 347–358.

[3] Brezhneva, O.A. and Tret'yakov, A.A., (2003), Optimality conditions for degenerate extremum problems with equality constraints, SIAM J. Control Optim., 42, pp. 729-745.

[4] Brezhneva, O.A. and Tret'yakov, A.A., (2010), Corrigendum: Optimality conditions for degenerate extremum problems with equality constraints, SIAM J. Control Optim., 48, pp. 3670-3673.

[5] Brezhneva, O.A. and Tret'yakov, A.A., (2003), The p-th order necessary optimality conditions for inequality–constrained optimization problems, in Lecture Notes in Computer Science, Springer-Verlag Heidelberg, 2667/2003, pp.903–911.

[6] Brezhneva, O.A. and Tret'yakov, A.A.,(2005), The pth order optimality conditions for inequality constrained optimization problems, Nonlinear Analysis, 63, pp. e1357-e1366.

[7] Brezhneva, O.A. and Tret'yakov, A.A., (2007), The p-factor-Lagrange methods for degenerate nonlinear programming, Numerical Functional Analysis and Optimization, 28, pp. 1051-1086.

[8] Brezhneva, O.A. and Tret'yakov, A.A., (2008), The pth order optimality conditions for nonregular optimization problems, Doklady Mathematics, 77, pp. 1-3.

[9] Buchner, M., Marsden, J.E. and Schecter, S., (1983), Applications of the blowing-up construction and algebraic geometry to bifurcation problems, J. Differential Equations, 48, pp. 404–433.

[10] Forsgren, A., Gill, P.E., Wright, M.H., (2002), Interior methods for nonlinear optimization, SIAM Review, 44, pp. 525-597.

[11] Gfrerer, H., (2006), Second-order optimality conditions for scalar and vector optimization problems in Banach spaces, SIAM J. Control Optim., 45, pp. 972-997.

[12] Ioffe, A.D., (1979), Necessary and sufficient conditions for a local minimum 3: Second order conditions and augmented duality, SIAM J. Control Optim., 17, pp. 266–288.

[13] Ioffe, A.D. and Tihomirov, V.M., (1979), Theory of extremal problems, North–Holland, Amsterdam, The Netherlands.

[14] Izmailov A.F., (1992), Degenerate extremum problems with inequality-type constraints, Comput. Math and Math Phys., 32 pp. 1413-1421.

[15] Izmailov, A.F., (1994), Optimality conditions for degenerate extremum problems with inequality-type constraints, Comput. Math and Math Phys., 34, pp. 723-736.

[16] Izmailov, A.F. and Solodov, M.V., (2001), Optimality conditions for irregular inequality-constrained problems, SIAM J. Control Optim., 40, pp. 1280-1295.

[17] Izmailov, A.F. and Tret'yakov, A.A., (1994), Factor-analysis of nonlinear mappings, Nauka, Moscow. (in Russian).

[18] Ledzewicz, U. and Schättler, H., (1999), High-order approximations and generalized necessary conditions for optimality, SIAM J. Control Optim., 37, pp. 33-53.
[19] Levitin, E.S., Milyutin, A.A., Osmolovskii, N.P., (1978), Conditions of higher order for a local minimum in problems with constraints, Russian Math Surveys, 33, pp.97-168.
[20] Nocedal, J. and Wright, S.J., (1999), Numerical optimization, Springer-Verlag New York Inc.
[21] Penot, J.P., (1999), Second-order conditions for optimization problems with constraints, SIAM J. Control Optim., 37, pp.303-318.
[22] Tret'yakov, A. A., (1983), Necessary conditions for optimality of pth order, Control and Optimization, Moscow, MSU, pp. 28-35. (in Russian).
[23] Tret'yakov, A.A., (1984), Necessary and sufficient conditions for optimality of pth order, USSR Comput. Math. and Math. Phys., 24, pp. 123-127.
[24] Tret'yakov, A. A., (1987), The implicit function theorem in degenerate problems, Russ. Math. Surv., 42, pp. 179-180.
[25] Tret'yakov, A. A. and Marsden, J. E., (2003), Factor-analysis of nonlinear mappings: p–regularity theory, Communications on Pure and Applied Analysis, 2, pp. 425–445.

Olga Brezhneva received her Master's Degree in Applied Mathematics from Moscow State University and her Ph.D. in Mathematics from the Russian Academy of Sciences in 2000. She is an associate professor in the department of mathematics at Miami University in Oxford, Ohio. She joined the Miami faculty in 2004. Prior to Miami, she was a postdoctoral associate at the Institute for Mathematics and its Applications at the University of Minnesota. Olga's primary area of research is optimization. She is a co-author of 32 research publications in the fields of optimization, differential equations, and numerical analysis. She has given talks and organized sessions at national and international conferences.

Alexey A. Tret'yakov received the M.S. and Ph.D. degrees from Moscow State University, and Doctor of Sciences degree from the Russian Academy of Sciences. He works as a Full Professor at the University of Natural and Humanistic Sciences in Siedlce (formerly known as University of Podlasie in Siedlce) and as a leading research scientist at the Russian Academy of Sciences. He is an author and co-author of more than 60 articles and 4 books. He has given many invited talks at various international conferences and academic institutions. He has also supervised many M.S. and Ph.D. students. His primary areas of research are optimization, nonlinear analysis, numerical analysis and calculus of variations.

The main accomplishment is creation of a new theory, Factor-Analysis of Nonlinear Mappings or P-regularity Theory, which can be used to describe the structure of nonlinear singular mappings. The theory can be applied to different fields of mathematics including numerical analysis, calculus of variations and optimal control, integral and differential equations, mathematical physics, optimization, ill-posed problems and others.