p-regular nonlinearity: tangency at singularity in degenerate optimization problems

Ewa M. Bednarczuk\textsuperscript{1,4} · Alexey Tretyakov\textsuperscript{2,3}

Received: 30 November 2016 / Accepted: 15 September 2017 / Published online: 17 October 2017 © Springer-Verlag GmbH Germany 2017

Abstract  We investigate description of the tangent cone to the null set of a mapping $F$ at a given point $x^*$ in the case when $F$ is degenerate at $x^*$. To this aim we introduce the concept of modified 2-regular mappings, which generalizes the concept of $p$-regular mappings. Our main result provides the description of the tangent cone to the null set of modified 2-regular mappings. With the help of this result we derive new optimality conditions for a wide class of optimization problems with equality constraints.

Keywords  Tangent cone · Degenerate mapping · $p$-regularity · Nonlinear optimization problems · Optimality conditions

Mathematics Subject Classification  46N10 · 47J30 · 47N10 · 90C30

1 Introduction

The problem of local description of the solution set appears in formulation of optimality conditions and construction of solution methods for optimization problems. In the present paper we consider degenerate optimization problems

$$
\min \ \varphi(x) \ \text{subject to} \ F(x) = 0,
$$

\textsuperscript{1} Systems Research Institute, ul. Newelska 6, 01-447 Warszawa, Poland
\textsuperscript{2} Siedlce University of Natural Sciences and Humanities, Faculty of Sciences, Siedlce, Poland
\textsuperscript{3} Dorodnicyn Computing Center of the Russian Academy of Sciences, Moscow, Russia
\textsuperscript{4} Warsaw University of Technology, Koszykowa 75, 00-662 Warsaw, Poland

\textsuperscript{Ewa M. Bednarczuk}
Ewa.Bednarczuk@ibspan.waw.pl
where \( \varphi : X \to \mathbb{R} \) is defined on a Banach space \( X \) and the feasible solution set is described by a mapping \( F : X \to Y \), where \( Y \) is a Banach space, which is degenerate at the solution point \( x^* \), i.e. \( \text{Im} F'(x^*) \neq Y \).

Degenerate problems appear often in applications. It was shown in Marsden and Tretyakov (2003) that degeneracy (singularity) defined as

\[
\text{Im} F'(x^*) \neq Y \quad (1.1)
\]

at a given admissible point \( x^* \), \( F(x^*) = 0 \), is, in some sense, typical for nonlinear mappings \( F \). Degeneracy occurs in the calculus of variations and the optimal control problems with boundary value conditions, e.g. in Chaplygin problem. The development of optimality conditions for degenerate problems is an active research topic, see Byrd et al. (1995), Dmitruk (1987), Ledzewicz and Schättler (1995, 1998a, b).

Here we focus on the description of the tangent cone

\[
TM(x^*) = \{ h \in X \mid x^* + \alpha h + r(\alpha) \in M(x^*), \; \alpha \in [0, \varepsilon], \; \|r(\alpha)\| = o(\alpha) \},
\]

where \( M(x^*) = \{ x \in X \mid F(x) = F(x^*) = 0 \} \) and \( \varepsilon > 0 \) is small enough. In the nondegenerate (regular) case, i.e., when

\[
\text{Im} F'(x^*) = Y
\]

the problem of description of elements \( h \in X \) such that \( h \in TM(x^*) \) has been solved by the famous Lusternik’s theorem which states that the tangent cone \( TM(x^*) \) to the set \( M(x^*) \) at the point \( x^* \) coincides with the kernel of the derivative operator \( F'(x^*) \), i.e., we have

\[
TM(x^*) = \text{Ker} F'(x^*).
\]

The degenerate case has been already investigated e.g., in Brezhneva and Tretyakov (2007), Buchner et al. (1983), Ledzewicz and Schättler (1998a), Ledzewicz and Schätter (1998b), Tretyakov (1984), where the constructive descriptions of the tangent cone to the null set \( M(x^*) \) are given for some classes of degenerate mappings. However, the classes of mappings considered so far, do not contain many important degenerate mappings. To enlarge the class of degenerate (singular) mappings with the constructive description of the tangent cone \( TM(x^*) \) to the set \( M(x^*) \) at the point \( x^* \) we apply the tools of the \( p \)-regularity theory introduced and studied in Brezhneva and Tretyakov (2003), Marsden and Tretyakov (2003), Tretyakov (1984, 1983, 1987).

The main idea of the \( p \)-regularity theory is to replace the operator \( F'(x^*) \), which is not onto, with a linear operator \( \Psi_p(x^*, h) \), \( p \geq 2 \), related to the \( p \)-th order Taylor’s polynomial of \( F \) at \( x^* \), which is onto. The operator \( \Psi_p(x^*, h) \) contains the derivatives of \( F \) up to the \( p \)-th order, so in our considerations, \( F \) is assumed to be \( p \)-times continuously differentiable in a neighbourhood of \( x^* \). The order \( p \) is chosen as the smallest number for which the operator \( \Psi_p \) is regular.
Let us point out that mathematical programing problems with complementarity constraints
\[
\begin{align*}
\min & \quad \varphi(x) \text{ subject to} \\
& g_1(x) \geq 0, \ldots, g_n(x) \geq 0 \\
x_1 \geq 0, & \ldots, x_n \geq 0 \\
x_i g_i(x) = 0, & \quad i = 1, \ldots, n
\end{align*}
\] (1.2)
are degenerate. Indeed, the constraints \(x_i g_i(x) = 0, i = 1, \ldots, n\) are degenerate (nonregular) if \(x_i\) and \(g_i(x)\) are active at the solution point \(x^*\) (the strict complementarity conditions do not hold). Then we can not apply the classical optimality conditions and moreover, Newton-type methods are inapplicable.

It turns out, however, that the constraints of problem (1.2) are 2-regular along some element \(h \in X\) in the sense defined below, and thus, we are able to provide meaningful optimality conditions for (1.2) and to construct efficient Newton-type solution methods.

There are many papers devoted to the investigation of deformations and perturbations of optimization problems, see e.g. Jongen et al. (1986, 1983, 1986), Jongen et al. (1990), Klatte and Kummer (2002), Rückmann (1993) and the references therein. Small perturbations of data of degenerate optimization problems may lead to large changes in solutions and/or to nonexistence of approximate solutions. It turns out that the existence of the \(p\)-regular structure of the problem under investigation entails stability of approximate solutions.

For these reasons, \(p\)-regularity theory is a valuable and adequate tool for providing optimality conditions and solution methods for large classes of degenerate optimization problems.

2 Elements of the \(p\)-regularity theory

Consider the equation
\[
F(x) = 0,
\] (2.1)
where \(F : X \to Y\) \(X, Y\) are Banach spaces, \(F \in C_{p+1}(X)\). Let us assume that \(F(x^*)\) is degenerate (singular) at a given point \(x^* \in M(x^*)\). In this section we recall basic constructions of \(p\)-regularity theory as developed in Brezhneva and Tretyakov (2003), Marsden and Tretyakov (2003), Tretyakov (1984), Tretyakov (1983), Tretyakov (1987) to investigate singular mappings.

We assume that the space \(Y\) is decomposed into the direct sum
\[
Y = Y_1 \oplus Y_2 \oplus \cdots \oplus Y_p,
\] (2.2)
where \(Y_1 := \text{cl Im} F'(x^*), Z_1 = Y\). Let \(Z_2\) be a closed complementary subspace to \(Y_1\) (we assume that such closed complement subspace exists), and let \(P_{Z_2} : Y \to Z_2\) be the projection operator onto \(Z_2\) along \(Y_1\). By \(Y_2\) we mean the closed linear span of the image of the map \(P_{Z_2} F^{(2)}(x^*)[\cdot]^2\). More generally, we define inductively \(Y_i := \text{cl}(\text{Span Im} P_{Z_i} F^{(i)}(x^*)[\cdot]^i) \subset Z_i, i = 2, \ldots, p - 1\), where \(Z_i\) is a chosen closed complementary subspace for \((Y_1 \oplus Y_2 \cdots \oplus Y_{i-1})\) with respect to \(Y, i = 2, \ldots, p\).
and $P_{Z_i} : Y \rightarrow Z_i$ is the projection operator onto $Z_i$ along $(Y_1 \oplus Y_2 \cdots \oplus Y_{i-1})$ with respect to $Y$, $i = 2, \ldots, p$. Finally, $Y_p = Z_p$. The order $p$ is the smallest number for which (2.2) holds.

Let $F_i : X \rightarrow Y_i, i = 1, \ldots, p$ be defined as $F_i(x) := P_i F(x)$, where $P_i := P_{Y_i} : Y \rightarrow Y_i$ is the projection operator onto $Y_i$ along $(Y_1 \oplus Y_2 \cdots \oplus Y_{i-1} \oplus Y_{i+1} \oplus \cdots \oplus Y_p)$ with respect to $Y$, $i = 1, \ldots, p$.

**Definition 2.1** The linear operator $\Psi_p(x^*, h) \in L(X, Y_1 \oplus Y_2 \cdots \oplus Y_p)$ for $h \in X$, $h \neq 0$,

$$\Psi_p(x^*, h) := F_1'(x^*) + F_2''(x^*)[h] + \cdots + F_p^{(p)}(x^*)[h]^{p-1}$$

(2.3) is called the $p$-factor operator of the mapping $F$ at the point $x^*$.

**Definition 2.2** We say that the mapping $F$ is $p$-regular at $x^*$ along (on) an element $h \in X$ if

$$\text{Im} \Psi_p(x^*, h) = Y.$$  

(2.4)

**Definition 2.3** We say that the mapping $F$ is $p$-regular at $x^*$ if it is $p$-regular along any $h \in X$ from the set

$$H_p(x^*) := \bigcap_{k=1}^p \text{Ker}^k F_k^{(k)}(x^*) \setminus \{0\},$$

where $\text{Ker}^k F_k^{(k)}(x^*) = \{\xi \in X | F_k^{(k)}(x^*)[\xi]^k = 0\}$ is the $k$-kernel of the $k$-order mapping $F_k^{(k)}(x^*)[\xi]^k$.

For the linear surjective operator $\Psi_p(x^*, h) : X \rightarrow Y$, by $\{\Psi_p(x^*, h)\}^{-1}$ we denote its right inverse, $\{\Psi_p(x^*, h)\}^{-1} : Y \rightarrow 2^X$, and we have

$$\{\Psi_p(x^*, h)\}^{-1} y = \{x \in X | \Psi_p(x^*, h)x = y\}.$$

We define the norm of $\{\Psi_p(x^*, h)\}^{-1}$ by the formula

$$\|\{\Psi_p(x^*, h)\}^{-1}\| = \sup_{\|y\|=1} \inf_{x \in \{\Psi_p(x^*, h)\}^{-1}y} \|x\|.$$  

We say that $\{\Psi_p(x^*, h)\}^{-1}$ is bounded if $\|\{\Psi_p(x^*, h)\}^{-1}\| < +\infty$.

**Definition 2.4** The mapping $F$ is called strongly $p$-regular at the point $x^*$ if there exists $\gamma > 0$ such that

$$\sup_{h \in H_\gamma} \|\{\Psi_p(x^*, h)\}^{-1}\| < +\infty,$$

where

$$H_\gamma = \{h \in X | \|F_k^{(k)}(x^*)[h]^{k}\|_{Y_k} \leq \gamma, \ k = 1, \ldots, p, \ \|h\| = 1\}.$$  

Springer
The following two theorems describe the tangent cone $TM(x^*)$ to the set $M(x^*)$ at the point $x^*$ and the null sets $M(x^*)$ of $p$-regular and strongly $p$-regular mappings, respectively.

**Theorem 2.5** (Brezhneva and Tretyakov 2007) Let $X$, $Y$ be Banach spaces. Let $F : X \to Y$, $F \in C^{p+1}(X)$, $F(x^*) = 0$ and let $F$ be $p$-regular at $x^*$. Then

$$TM(x^*) = H_p(x^*).$$

**Theorem 2.6** (Prusińska and Tretyakov 2016) Let $X$, $Y$ be Banach spaces. Let $F : X \to Y$, $F \in C^{p+1}(X)$, $p \geq 2$, $F(x^*) = 0$, and let $F$ be strongly $p$-regular at $x^*$. Then there exists a neighbourhood $U(x^*)$, a mapping $\xi \to x(\xi) : U(x^*) \to X$ and a constant $\delta > 0$ such that

$$F(\xi + x(\xi)) = 0 \quad \|x(\xi)\| \leq \delta \sum_{k=1}^{p} \|F_k(\xi)\|_{Y_k}$$

$$\text{and} \quad \|x(\xi)\| \leq \delta \sum_{k=1}^{p} \|F_k(\xi)\|_{Y_k}^{1/k}$$

for all $\xi \in U(x^*)$.

The proof of Theorem 2.6 can be found in Prusińska and Tretyakov (2016). The importance of this result in the degenerate case is analogous to the importance of the classical implicit function theorem in nondegenerate case. In particular, Theorem 2.6 is used in proving optimality conditions for degenerate constrained optimization problems (Brezhneva and Tretyakov 2003) (see Sect. 4).

### 3 Generalization of $p$-regularity and description of the $p$-th order tangent cone

In Theorem 2.5, the crucial assumption which allows the constructive description of the tangent cone $TM(x^*)$ to the set $M(x^*)$ at $x^*$ is condition (2.4), i.e., the $p$-regularity of $F$ along any element $h \in H_p(x^*)$. However, condition (2.4) may fail. For $p = 2$, there are examples such that $F''(x^*)h \cdot X \neq Y$ ($F'(x^*) = 0$), where $h \in \text{Ker}^2 F''(x^*)$.

**Example 3.1** Consider the mapping $F : \mathbb{R}^4 \to \mathbb{R}^2$,

$$F(x) := \begin{pmatrix} x_1x_2 - x_3^2 \\ x_3x_4 \end{pmatrix}, \quad h = (1, 0, 0, 0)^T \in \text{Ker}^2 F''(0).$$

At $x^* = 0$, the 2-factor-operator $F''(0)h = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ is singular. Therefore, Theorem 2.5 does not apply. However, $h \in TM(0)$, i.e. $th + \omega(t) \in M(0) \Rightarrow F(th + \omega(t)) = 0$, $\|\omega(t)\| = o(t)$, $t \in [0, \varepsilon]$, $\varepsilon > 0$ sufficiently small. Here $\omega(t) = (0, 0, 0, 0)^T$.

In the sequel, we consider separately the cases where $p = 2$ and $p \geq 3$. 

© Springer
3.1 Case \( p = 2 \)

Suppose that the space \( Y \) is decomposed into the direct sum

\[
Y = Y_1 \oplus Y_2^2 \oplus \cdots \oplus Y_q^2,
\]

where \( Y_1 = \text{cl Im } F'(x^*) \), \( Z_1 = Y \), \( Z_1^2 \) is a closed complementarity subspace to \( Y_1 \) and let \( P_{Z_1}^2 : Y \to Z_1^2 \) be the projection operator onto \( Z_1^2 \) along \( Y_1 \). By \( Y_1^2 \) we mean \( Y_1 \cap \text{cl Im } F''(x^*) \).

More generally, define inductively \( Y_k^2 := \text{cl Im } P_{Z_k}^2 F''(x^*) h_k \), \( k = 2, \ldots, q - 1 \), where \( Z_k^2 \) is a chosen closed complementarity subspace for \( (Y_1 \oplus Y_1^2 \oplus \cdots \oplus Y_{k-1}^2) \) with respect to \( Y \), and \( k_1^2 := P_{Z_1}^2 : Y \to Z_1^2 \) is the projection operator onto \( Z_1^2 \) along \( (Y_1 \oplus Y_1^2 \oplus \cdots \oplus Y_{k-1}^2) \) with respect to \( Y \), \( k = 2, \ldots, q \). Finally \( Y_q^2 = Z_2^2 \).

The order \( q \) is chosen as the smallest number for which condition (3.1) holds.

Let us define the mappings

\[
F_1 : X \to Y_1, \quad F_1(x) := P_{Y_1} \cdot F(x),
\]

\[
F_{2,k} : X \to Y_k^2, \quad F_{2,k}(x) := P_{Y_k^2} \cdot F(x), \quad k = 1, \ldots, q,
\]

where \( P_{Y_k^2} : Y \to Y_k^2 \) is the projection operator onto \( Y_k^2 \) along

\[
\left( Y_1 \oplus Y_1^2 \oplus \cdots \oplus Y_{k-1}^2 \oplus Y_{k+1}^2 \oplus \cdots \oplus Y_q^2 \right).
\]

**Definition 3.2** The linear operator

\[
\Psi_q^2(x^*; h_1, \ldots, h_q) \in \mathcal{L} \left( X, Y_1 \oplus Y_1^2 \oplus \cdots \oplus Y_q^2 \right)
\]

\( h_k \neq 0, k = 1, \ldots, q \)

\[
\Psi_q^2(x^*; h_1, \ldots, h_q)) = F_1'(x^*) + F_{2,1}''(x^*) h_1 + \cdots + F_{2,q}''(x^*) h_q
\]

(3.2)

is called the modified 2-factor-operator.

**Definition 3.3** We say that the mapping \( F \) is modified 2-regular at \( x^* \) along \( h_1, h_2, \ldots, h_q \) if \( \text{Im } \Psi_q^2(x^*; h_1, \ldots, h_q) = Y \).

**Example 3.4** (Continuation of Example 3.1)

Let \( h_1 := (1, 0, 0, 0)^T, q = 2, h_2 := (0, 0, 1, 0)^T \). Then \( P_{Y_1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \), \( P_{Y_2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \), \( P_{Y_2^2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \), \( F_1(x) = \begin{pmatrix} 0 \end{pmatrix} \), \( F_{2,1}(x) := \begin{pmatrix} x_1 x_2 - x_3^2 \\ 0 \end{pmatrix} \), \( F_{2,2}(x) := \begin{pmatrix} 0 \\ x_3 x_4 \end{pmatrix} \), \( F_{2,1}(0) h_1 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \), \( F_{2,1}(0) h_2 = \begin{pmatrix} 0 & 0 \end{pmatrix} \).

\( \mathcal{S} \) Springer
Moreover, among the points \(w\) satisfying (3.4) there exists a point such that

\[
\|\tilde{w} - w_0\| \leq \frac{2}{1-\alpha} \text{dist}(w_0, \Lambda(w_0)).
\]
Here $H(\Lambda_1, \Lambda_2)$ is the Hausdorff distance between sets $\Lambda_1$ and $\Lambda_2$,

$$H(\Lambda_1, \Lambda_2) := \max\{ \sup_{x \in \Lambda_1} \text{dist}(x, \Lambda_2), \sup_{y \in \Lambda_2} \text{dist}(y, \Lambda_1)\}.$$ 

**Proof of Theorem 3.5** For the sake of simplicity consider the case $F'(x^*) = 0$. Set $\gamma := \frac{1}{2}q$ and introduce the following operators

$$A_k := F_{2,k}'(x^*)[h_k], \quad k = 1, \ldots, q$$

and

$$A := (tA_1, \ldots, t^{1+(q-1)\gamma}Aq).$$

Define $h(t) := th_1 + \cdots + t^{1+(q-1)\gamma}h_q$ and consider the mapping

$$\Lambda(x) := x - A^{-1}(F_{2,1}(x^* + h(t) + x), \ldots, F_{2,q}(x^* + h(t) + x)).$$

We show that all assumptions of $(MCMP)$ are satisfied for $\Lambda(x)$ with some ball $U_{r(t)}(0)$, where $r_1 := r(t) = o(t^{1+(q-1)\gamma})$.

We start by checking the assumption 2. of $(MCMP)$. By the definition of mapping $F_{2,k}$, we have

$$F_{2,k}'(x^*)[h_1 + \cdots + t^{1+(q-1)\gamma}h_q] = 0,$$

where $\|\omega_k(t)\| \leq ct^3$. By definition of mapping $F_{2,k}$, we have

$$F_{2,k}'(x^*)[h_i, h_j] = 0, \text{ for } i < k \text{ and } j \leq k.$$  

By (3.7) and (3.6), we obtain

$$\frac{\|F_{2,k}(x^* + h(t))\|}{t^{1+(k-1)\gamma}} = o(t^{1+(q-1)\gamma}) \text{ for } k = 1, \ldots, q.$$
Then, by (3.5) and the latter relation with \( \gamma = \frac{1}{2q} \) we obtain
\[ \| \Lambda(0) \| = o(t^{1+(q-1)\gamma}) = o(t^{1+(q-1)/2q}). \]

For sufficiently small \( t \) with some \( \alpha \in (0, 1) \) and \( r_1 := r(t) = o(t^{1+(q-1)\gamma}) \) we get
\[ \| \Lambda(0) \| < (1 - \alpha)r_1 \]
which proves assumption 2. of \((MCMP)\).

Now we show that assumption 1. of \((MCMP)\) holds for all \( x_1, x_2 \in U_{r(t)}(0) \) that is
\[ H(\Lambda(x_1), \Lambda(x_2)) \leq \alpha \| x_1 - x_2 \|, \quad 0 < \alpha < 1. \]

By the definition of \( \Lambda(x) \), with \( \bar{x}(t) := x^* + h(t) \), we have
\[
H(\Lambda(x_1), \Lambda(x_2)) = \inf \{ \| z_1 - z_2 \| \mid z_i \in \Lambda(x_i), \ i = 1, 2 \}
= \inf \{ \| z_1 - z_2 \| \mid Az_i = Ax_i - F(\bar{x}(t) + x_i), \ i = 1, 2 \}
= \inf \{ \| z \| \mid Az = A(x_1 - x_2) - F(\bar{x}(t) + x_1) + F(\bar{x}(t) + x_2) \}
\leq \| A^{-1}(A(x_1 - x_2) - F(\bar{x}(t) + x_1) + F(\bar{x}(t) + x_2)) \|.
\]

From this we deduce that
\[
H(\Lambda(x_1), \Lambda(x_2)) \leq c \left( \frac{\| F''_{x^*}(x^*)[t h_1](x_1-x_2)-F_{2,1}(\bar{x}(t)+x_1)+F_{2,1}(\bar{x}(t)+x_2)}{t^{1+\gamma(q-1)}} + \cdots + \frac{\| F''_{x^*}(x^*)[t^{1+\gamma(q-1)}]h_q](x_1-x_2)-F_{2,q}(\bar{x}(t)+x_1)+F_{2,q}(\bar{x}(t)+x_2)}{t^{1+\gamma(q-1)}} \right).
\]

By using Mean Value Theorem and Taylor’s expansion for \( k = 1, \ldots, q \), there exists \( c_k > 0 \) such that
\[
\frac{\| F''_{x^*}(x^*)[t^{1+\gamma(k-1)}]h_k](x_1-x_2)-F_{2,k}(\bar{x}(t)+x_1)+F_{2,k}(\bar{x}(t)+x_2)}{t^{1+\gamma(k-1)}} \leq c_k \sup_{\xi \in U_{r_1}(0)} \| \xi \| \frac{\| x_1-x_2 \|}{t^{1+\gamma(k-1)}}.
\]

Hence, with \( r_1 := r(t) = o(t^{1+\gamma(q-1)}) \) we get
\[ H(\Lambda(x_1), \Lambda(x_2)) \leq \alpha \| x_1 - x_2 \| \]
which proves assumption 1. of \((MCMP)\).

By \((MCMP)\), there exists \( \omega(t) \) such that \( \omega(t) \in \Lambda(\omega(t)) \) which is equivalent to
\[ 0 \in A^{-1}(F(x^* + h(t) + \omega(t))). \]
Hence,

$$F(x^* + h(t) + \omega(t)) = 0$$

with \(\|\omega(t)\| \leq c\|\Lambda(0)\| = o(t^{1+(q-1)/2q})\). One can easily show that, by (3.5), and by the inequality \(\|\omega(t)\| \leq c\|\Lambda(0)\|\), we obtain the following estimate

$$\|\omega(t)\| \leq c \left( \|F_1(x^* + h(t))\| + \sum_{k=1}^{q} \|F_{2,k}(x^* + h(t))\|^{\frac{1}{p-k+2}} \right)$$

which finishes the proof. \(\square\)

### 3.2 Case \(p \geq 3\)

We seek an element \(x(t) \in M(x^*)\) in the form

$$x(t) := x^* + th_1 + t^{1+\alpha_2}h_2 + \cdots + t^{1+\alpha_q}h_q + \omega(t)$$

where \(\|\omega(t)\| = o(t^{1+\alpha_q}),\ 0 < \alpha_2 < \cdots < \alpha_q < 1\).

For the sake of simplicity we assume that

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

As previously,

$$Y = Y_1 \oplus \cdots \oplus Y_p$$

where \(Y_1 = \text{cl Im } F^{(p)}(x^*)[h_1]^{p-1}\), \(Z_1 := Y_1, Z_2\) is closed complementary subspace to \(Y_1\) and \(P_{Z_1} : Y \rightarrow Z_1, P_{Z_2} : Y \rightarrow Z_2\) are projection operators onto \(Z_1\) and \(Z_2\), respectively, \(Y_2 := \text{cl Im } P_{Z_2} F^{(p)}(x^*)[h_1]^{p-2}[h_2]\).

More generally, we define inductively

$$Y_k := \text{cl Im } P_{Z_k} F^{(p)}(x^*)[h_1]^{p-k}[h_2]^{k-1},\ k = 3, \ldots, p$$

where \(P_{Z_k} : Y \rightarrow Z_k\) is the projection operator onto \(Z_k\) along \((Y_1 \oplus \cdots \oplus Y_{k-1})\) with respect to \(Y, k = 3, \ldots, p\). Finally \(Y_p := Z_p\).

Let us define \(F_k(x) := P_k F(x)\), where \(P_k : Y \rightarrow Y_k\) is the projection operator onto \(Y_k\) along

\((Y_1 \oplus \cdots \oplus Y_{k-1} \oplus Y_{k+1} \oplus \cdots \oplus Y_p),\ k = 1, \ldots, p.\)
**Definition 3.8** The linear operator $\Psi_p(x^*; h_1, h_2) \in \mathcal{L}(X, Y_1 \oplus \cdots \oplus Y_p)$ defined as

$$\Psi_p(x^*; h_1, h_2) := F_1^{(p)}(x^*)[h_1]^{p-1} + F_2^{(p)}(x^*)[h_1]^{p-2}[h_2] + \cdots + F_p^{(p)}(x^*)[h_2]^{p-1}$$

is called the modified $p$-factor operator.

**Remark 3.9** In case $q \geq 3$ the modified $p$-factor operator $\Psi_p(x^*; h_1, \ldots, h_q)$ has the following form

$$\Psi_p(x^*; h_1, \ldots, h_q) := \sum_{k=1}^{N(p,q)} F_k^{(p)}(x^*)[h_i]^{q_1}, \ldots, [h_{i_{p-1}}]^{q_{p-1}},$$

where $N(p,q) := \left(\frac{q-1}{p+q-1}\right), i_j \in \{1, \ldots, p\}, i_k \neq i_j, \text{ for } k \neq j \text{ and } q_k \in \{0, \ldots, p-1\}$ such that $q_1 + \ldots, q_{p-1} = p-1$ and mappings $F_k$ are defined in the same way as in case $q = 2$ and, obviously, depend on $(p-1)-$ tuple $q_1, \ldots, q_{p-1}.$

**Definition 3.10** The mapping $F$ is modified $p$-regular at $x^*$ along $h_1, h_2$ if $\text{Im } \Psi_p(x^*; h_1, h_2) = Y \text{ (or } Y_1 \oplus \cdots \oplus Y_p).$

Now we seek an element $x(t) \in M(x^*)$ in the form

$$x(t) := x^* + th_1 + t^{1+\alpha} h_2 + \omega(t), \quad (3.8)$$

where $\alpha := \frac{1}{x^p}$ and $\|\omega(t)\| = o(t^{1+\alpha}).$ The proof of the theorem below remains the same when $\alpha$ assumes any value from a given interval $\alpha \in (0, \varepsilon), \varepsilon < 1.$

**Theorem 3.11** Let $X, Y$ be Banach spaces. Let $F : X \to Y, F \in C^{p+1}(X), F(x^*) = 0, F^{(k)}(x^*) = 0, k = 1, \ldots, p-1.$

Assume that $F$ is modified $p$-regular at $x^*$ along $h_1, h_2$ and for the linear operator $\Psi_p(x^*; h_1, h_2)$

$$\Psi_p(x^*; h_1, h_2) \cdot h_1 = 0, \quad \Psi_p(x^*; h_1, h_2) \cdot h_2 = 0.$$

Then $h_1 \in TM(x^*)$ and there exists $\omega(t), \|\omega(t)\| = o(t^{1+\alpha})$ such that

$$F(x^* + th_1 + t^{1+\alpha} h_2 + \omega(t)) = 0$$

and

$$\|\omega(t)\| \leq c \sum_{k=1}^{p} \|F_k(x^* + th_1 + t^{1+\alpha} h_2)\| \frac{1+\alpha}{p+\alpha k},$$

where $c > 0$ is an independent constant.

**Proof** The proof of this theorem is analogous to the proof of Theorem 3.5 and therefore we omit it.  

\(\Box\)
4 Degenerate optimization problems

We consider the nonlinear optimization problem

\[
\begin{align*}
\min & \quad \varphi(x) \\
\text{subject to} & \quad F(x) = 0,
\end{align*}
\]  

(4.1)

where \( \varphi : X \to \mathbb{R} \) is a sufficiently smooth function and \( F : X \to Y \) is a sufficiently smooth mapping from a Banach space \( X \) into a Banach space \( Y \). Let us consider the case where the mapping \( F \) is degenerate at the solution \( x^* \) of problem (4.1) that is, when the derivative \( F'(x^*) \) is not onto. In our previous works (Bednarczuk et al. 2011; Brezhneva and Tretyakov 2003) we derived optimality conditions for constrained optimization problems (4.1) that are \( p \)-regular at \( x^* \), i.e., when \( F \) is \( p \)-regular at \( x^* \).

Now we use the results of the previous sections to prove optimality conditions for problems with mappings \( F \) which are strongly \( p \)-regular or modified 2-regular. Let us define \( p \)-factor Lagrange function

\[
L_p(x, \lambda, h) := \varphi(x) + \left( \sum_{k=1}^{p} F_k^{(k-1)}(x)[h]^{k-1}, \lambda \right),
\]

where \( \lambda \in Y^* \), \( F_1^{(0)}(x) := F(x) \).

The following optimality conditions for \( p \)-regular and strongly \( p \)-regular mappings \( F \) were proved in Brezhneva and Tretyakov (2003).

**Theorem 4.1** Let \( X \) and \( Y \) be Banach spaces. Let \( \varphi : X \to \mathbb{R}, \varphi \in C^2(X) \), \( F : X \to Y \), \( F \in C^{p+1}(X) \). Suppose that \( h \in H_p(x^*) \) and \( F \) is \( p \)-regular along \( h \) at the point \( x^* \).

If \( x^* \) is a solution to problem (4.1), then there exist multipliers \( \lambda^*(h) \in Y^* = Y_1^* \times Y_2^* \times \cdots \times Y_p^* \) such that

\[
L'_p(x^*, \lambda^*(h), h) = 0 \iff \varphi'(x^*) + \left( F'_1(x^*) + \cdots + F'_p(x^*)h^{p-1} \right)^* \lambda^*(h) = 0
\]

(4.2)

Assume that \( F \) is strongly \( p \)-regular at \( x^* \).

If there exists \( \alpha > 0 \) and multipliers \( \lambda^*(h) \in Y_1^* \times Y_2^* \times \cdots \times Y_p^* \) such that

\[
L'_p(x^*, \lambda^*(h), h) = 0
\]

and

\[
L''_p(x^*, \lambda^*(h), h)[h]^2 \geq \alpha \|h_p\|^2 \quad \forall \ h_p \in H_p(x^*),
\]

where \( \lambda^*(h) := (\lambda_1^*(h), 1/3\lambda_2^*(h), \ldots, 2/p(p+1)\lambda_p^*(h)) \), then \( x^* \) is a strict local minimizer of problem 4.1.

Springer
The proof of this theorem is based on Theorem 2.6 and can be found in Brezhneva and Tretyakov (2003). It turns out that there exist numerous problems for which the assumption of $p$-regularity of the mapping $F$ fails at the solution $x^*$.

**Example 4.2** Consider the following problem

$$\min x_3 \quad (4.3)$$

subject to

$$F(x) := \left( \frac{x_1 x_2 - x_3^2}{x_3 x_4} \right) = 0. \quad (4.4)$$

We investigate optimality of $x^* := (0, 0, 0, 0)^T$. The mapping $F$ is not 2-regular at $x^*$ and we cannot apply Theorem 4.1.

However, for modified 2-regular mappings $F$ the following result holds. Let us introduce the modified 2-factor Lagrange function

$$L_{2,q}(x, \lambda, h_1, \ldots, h_q) := \varphi(x) + \left(F(x) + \sum_{k=1}^q F'_2(x)h_k, \lambda \right).$$

**Theorem 4.3** (Case $p = 2$) Let $X$ and $Y$ be Banach spaces, $F : X \to Y$, $\varphi : X \to \mathbb{R}$, $\varphi \in C^2(X)$ and $F \in C^3(X)$.

Assume that there exist elements $h_1, \ldots, h_q \in X$ such that the mapping $F$ is modified 2-regular at $x^*$ along $h_1, \ldots, h_q$ and assumption 3.3 of Theorem 3.5 is fulfilled.

Then there exists a multiplier $\lambda^* \in Y^*$ such that

$$L'_{2,q}(x^*, \lambda^*, h_1, \ldots, h_q) = 0 \iff \varphi'(x^*) + \left(F'(x^*) + F''_2(x^*)h_1 + \cdots + F''_q(x^*)h_q \right)\lambda^* = 0 \quad (4.5)$$

The proof of Theorem 4.3 is similar to the proof of Theorem 3.3 of Brezhneva and Tretyakov (2003) and we omit it here.

Let us note that for $h_1 := (1, 0, 0, 0)$ and $h_2 := (0, 0, 1, 0)$ the mapping $F$ from Example 4.2 defined by (4.4) is modified 2-regular along $h_1$ and $h_2$ at $x^* = 0$ and

$$\Psi^2_q(0; h_1, h_2) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$ 

If $x^* = (0, 0, 0, 0)^T$ would solve the problem (4.3), (4.4), then according to Theorem 4.3, there would be a multiplier $\lambda^* \in \mathbb{R}^2$ such that

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}\lambda^* = 0$$
which is impossible.

Consider the case \( p \geq 3 \) and \( q = 2 \) when \( F'(x^*) = 0, \ldots, F^{(p-1)}(x^*) = 0 \). Let us introduce the modified \( p \)-factor Lagrange function

\[
L_{p,2}(x, \lambda, h_1, h_2) := \varphi(x) + \left( \sum_{k=1}^{p} F_k^{(p-1)}(x)[h_1]^{p-k}[h_2]^{k-1}, \lambda \right).
\]

**Theorem 4.4 (Case \( p \geq 3 \))** Let \( X \) and \( Y \) be Banach spaces, \( F: X \to Y, \varphi: X \to \mathbb{R} \), \( \varphi \in C^2(X) \) and \( F \in C^{p+1}(X) \), and let \( x^* \) be a solution to optimization problem (4.1).

Assume that there exist elements \( h_1, h_2 \in X \) such that the mapping \( F \) is modified \( p \)-regular at \( x^* \) along \( h_1, h_2 \) and

\[
\Psi_p(x^*; h_1, h_2) \cdot h_k = 0, \quad k = 1, 2.
\]

Then there exists a multiplier \( \lambda^* \in Y^* \) such that

\[
L_{p,2}'(x^*, \lambda^*, h_1, h_2) = 0 \iff 
\varphi'(x^*) + \left( (F_1^{(p)}(x^*)[h_1]^{p-1} + F_2^{(p)}(x^*)[h_1]^{p-2}[h_2] + \cdots + F_p^{(p)}(x^*)[h_2]^{p-1})^* \right) \lambda^* = 0
\]

(4.6)

**Proof** We will show that for any \( z \in \text{Ker} \Psi(x^*; h_1, h_2) \) the equality \( \langle \varphi'(x^*), z \rangle = 0 \) holds. By annihilator lemmas [ATF], it means that

\[
\varphi'(x^*) \in \text{Im} \Psi(x^*; h_1, h_2) = \text{Im} \left( F_1^{(p)}(x^*)[h_1]^{p-1} + F_2^{(p)}(x^*)[h_1]^{p-2}[h_2] + \cdots + F_p^{(p)}(x^*)[h_2]^{p-1} \right)^*
\]

or, there exists \( \lambda^* \in Y^* \) such that

\[
\varphi'(x^*) + \left( (F_1^{(p)}(x^*)[h_1]^{p-1} + F_2^{(p)}(x^*)[h_1]^{p-2}[h_2] + \cdots + F_p^{(p)}(x^*)[h_2]^{p-1})^* \right) \lambda^* = 0
\]

Let \( z \in \text{Ker} \Psi(x^*; h_1, h_2) \). It means that

\[
(F_1^{(p)}(x^*)[h_1]^{p-1} + F_2^{(p)}(x^*)[h_1]^{p-2}[h_2] + \cdots + F_p^{(p)}(x^*)[h_2]^{p-1}) \cdot z = 0.
\]

Taking into account the last inequality we will show that there exists \( \omega(z, t) \) such that

\[
x(z, t) = x^* + th_1 + t^{1+\alpha}h_2 + t^{1+\alpha+\varepsilon}z + \omega(z, t) \in M(x^*),
\]

\( \square \) Springer
\( t \in [0, \delta], \| \omega(z, t) \| = o(t^{1+\alpha+\varepsilon}), \alpha + \varepsilon < 1 \) and \( \delta > 0 \) sufficiently small. To this aim we introduce mapping

\[
\Lambda_p(x) := x - \Psi(x^*; th_1, t^{1+\alpha}h_2)^{-1} (F_1(x^* + th_1 + t^{1+\alpha}h_2 + t^{1+\alpha+\varepsilon}z + t^{1+\alpha+\varepsilon}z + x) + \ldots + F_p(x^* + th_1 + t^{1+\alpha}h_2 + t^{1+\alpha+\varepsilon}z + x)).
\]

Based on the fact that

\[
\| F_k(x^* + th_1 + t^{1+\alpha}h_2 + t^{1+\alpha+\varepsilon}z) \| \leq O(t^2)
\]

for \( k = 1, \ldots, p \) we obtain, analogously as in the proof of Theorem 3.5, that mapping \( \Lambda_p(x) \) satisfies all the assumptions of \((MCMP)\) with some ball \( B(\bar{r}(t)(0)) \), where \( \bar{r}(t) = o(t^{1+\alpha+\varepsilon}) \).

By \((MCMP)\), there exists \( \omega(z, t) \in \Lambda_p(\omega(z, t)) \) which is equivalent to

\[
0 \in \Psi(x^*; th_1, t^{1+\alpha}h_2)^{-1} F(x^* + th_1 + t^{1+\alpha}h_2 + t^{1+\alpha+\varepsilon}z + \omega(z, t)),
\]

\( t \in [0, \delta] \) with

\[
\| \omega(z, t) \| \leq c(\Lambda_p(0)) = o(t^{1+\alpha+\varepsilon}).
\]

It means that

\[
x(z, t) := x^* + th_1 + t^{1+\alpha}h_2 + t^{1+\alpha+\varepsilon}z + \omega(z, t) \in M(x^*)
\]

for all \( z \in \text{Ker} \Psi_p(x^*; h_1, h_2) \).

Now we finish the proof by observing that it must be

\[
\langle \varphi'(x^*), h_1 \rangle = \langle \varphi'(x^*), h_2 \rangle = \langle \varphi'(x^*), z \rangle = 0
\]

for all \( z \in \text{Ker} \Psi_p(x^*; h_1, h_2) \) since otherwise \( x^* \) is not a minimizer of our problem.

\[ \square \]

5 Conclusions

In this paper we derived new optimality conditions for problem with degenerate equality constraints. Our approach is based on constructions of \( p \)-regularity theory and on the modification of the concept of \( p \)-regularity. In Sect. 3 we proved a new theorem on the null set description and investigate the structure of the tangent cone for modified \( p \)-regular mappings. These results generalize the tangent cone descriptions obtained so far. Let us note that Theorems 3.5 and 3.11 do not give a complete description of the tangent cone \( TM(x^*) \) to the set \( M(x^*) \) at the point \( x^* \) of the mapping \( F \) but knowing a single element \( h \in TM(x^*) \) is enough to prove optimality conditions for optimization problems (4.1) with the modified \( p \)-regular mappings \( F \).
In Sect. 4 we derived new optimality conditions for modified $p$-regular constrained optimization problems. These results generalize necessary optimality conditions obtained for $p$-regular problems. The presented results can be considered as a part of the $p$-regularity theory.

Acknowledgements For the second author this work was supported by the Russian Foundation for Basic Research under the Grant No. 11-01-00786-a, by the Leading Scientific School, Grant No. 5264.2012.1 and by the Russian Academy of Sciences Presidium Program P-18.

References

Bednarczuk E, Prusińska A, Tretyakov A (2011) High order stability conditions for degenerate optimization problems; elements of $p$-regularity theory. Nonlinear Anal Theory Appl 74:836–846
Brezhneva O, Tretyakov A (2007) Implicit function theorems for nonregular mappings in Banach spaces. Exit from singularity. In: Banach spaces and their applications in analysis, deGruyter pp 285–302
Brezhneva O, Tretyakov A (2003) Optimality conditions for degenerate extremum problems with equality constraints. SIAM J Control Optim 42:729–745
Buchner M, Marsden JE, Schechter S (1983) Applications of the blowing-up construction and algebraic geometry to bifurcation problems. J Differ Equ 48:404–433
Byrd RH, Feng D, Schnabel RB (1995) On optimality conditions for singular optimization problems, Tech.rep.95.03, Research Institute for Advanced Computer Science, NASA Ames Research Center, Moffett Field, CA
Dmitruk AV (1987) Quadratic conditions for a Pontryagin minimum in an optimal control problem linear with respect to control. II Theorems on the relaxing of constraints n the equality, Math. USSR-Izv. 31, pp 121–141
Jongen HT, Jonker P, Twilt F (1986) Critical sets in parametric optimization. Math Program 34:333–353
Jongen HT, Jonker P, Twilt F (1983) Nonlinear Optimization in $\mathbb{R}^n$. I.Morse Theory, Chebyshev Approximation, Peter Lang, Frankfurt
Jongen HT, Jonker P, Twilt F (1986) Nonlinear Optimization in $\mathbb{R}^n$. II. Transversality, Flows, Parametric Aspects, Peter Lang, Frankfurt
Jongen HT, Klatte D, Tammer K (1990) Implicit functions and stationary points. Math Program 49:123–138
Klatte D, Kummer B (2002) Nonsmooth equations in optimization. Regularity, calculus, methods and applications. Kluwer, Dordrecht
Ledzewicz U, Schättler H (1995) Second-order conditions for extremum problems with non-regular equality constraints. J Optim Theory Appl 86:113–144
Ledzewicz U, Schättler H (1998) A higher order generalization of the Lusternik theorem. Nonlinear Anal 34:793–815
Ledzewicz U, Schättler H (1998) High order approximations and generalized necessary conditions for optimality. SIAM J Control Optim 37:33–53
Marsden J, Tretyakov A (2003) Factor-analysis of nonlinear mappings: $p$-regularity theory. Commun Pure Appl Anal 2:425–445
Prusińska A, Tretyakov AA (2016) $p$-regularity theory. Tangent cone description in the singular case. Ukr Math J 67:1236–1246
Rückmann JJ (1993) Stability of noncompact feasible sets in nonlinear optimization. In: Guddat J et al (eds) Parametric optimization and related topics III. Peter Lang, Frankfurt, pp 467–502
Tretyakov A (1984) Necessary and sufficient conditions for optimality of $p$-th order. Comput Math Math Phys 24:123–127
Tretyakov A (1983) Necessary conditions for optimality of $p$-th order, control and optimization, MGU, pp 28–35, (in Russian)
Tretyakov A (1987) The implicit function theorem in degenerate case problems. Russ Math Surv 42:179–180