TECHNISCHE UNIVERSITÄT WIEN

Operations Research and Control Systems

SWM ORCOS

On the Strong Metric Subregularity in Mathematical Programming

N.P. Osmolovskii V.M. Veliov

Research Report 2020-09

June 2020

ISSN 2521-313X

Operations Research and Control Systems
Institute of Statistics and Mathematical Methods in Economics
Vienna University of Technology

Research Unit ORCOS
Wiedner Hauptstraße 8 / E105-04
1040 Vienna, Austria
E-mail: orcos@tuwien.ac.at

On the Strong Metric Subregularity in Mathematical Programming*

N.P. Osmolovskii† V.M. Veliov‡

Abstract

This note presents sufficient conditions for the property of strong metric subregularity (SMSr) of the system of first order optimality conditions for a mathematical programming problem in a Banach space (the Karush-Kuhn-Tucker conditions). The constraints of the problem consist of equations in a Banach space setting and finite number of inequalities. The conditions under which SMSr is proved involve Fréchet differentiability of the data, strict Mangasarian-Fromovitz constraint qualification, and second-order sufficient optimality condition. The obtained result extends the one known for finite-dimensional problems. Although the applicability of the result is limited in the truly Banach space setting (due to the Fréchet differentiability assumptions and the finite number of inequality constraints), the paper can be valuable due the self-contained exposition, and provides a ground for extensions that are applicable in calculus of variations and optimal control.

Keywords: optimization, mathematical programming, Karush-Kuhn-Tucker conditions, metric regularity

AMS Classification: 90C48

1 Introduction

Let X and Y be Banach spaces, and let the mappings

$$f_0 : X \to \mathbb{R}, \quad f_i : X \to \mathbb{R} \ (i = 1, \dots, k), \quad g : X \to Y$$

be twice continuously Fréchet differentiable. Consider the optimization (mathematical programming) problem

$$\min f_0(x) \tag{1}$$

$$\text{subject to} \qquad g(x) = 0, \quad f_i(x) \le 0 \ (i = 1, \dots, k). \tag{2}$$

*This research is supported by the Austrian Science Foundation (FWF) under grant No P31400-N32.

†Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland osmolov@ibspan.waw.pl

‡Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Austria, vladimir.veliov@tuwien.ac.at

The following system of equations and inequalities is known as Karush-Kuhn-Tucker (KKT) system associated with problem (1)–(2):

$$
\begin{aligned}
&f_0'(x) + \sum_{i=1}^{k} \alpha_i f_i'(x) + (g'(x))^* y^* = 0,\\
&g(x) = 0,\\
&\alpha_i f_i(x) = 0, \quad i = 1, \dots, k,\\
&f_i(x) \le 0, \quad \alpha_i \ge 0, \quad i = 1, \dots, k,
\end{aligned}
$$

where $x \in X$, $y^* \in Y^*$ (Y^* denotes the dual space to Y), and $\alpha := (\alpha_1, \dots, \alpha_k) \in \mathbb{R}^k$. Moreover, "primes" indicate Fréchet derivatives, and $(g'(x))^* : Y^* \to X^*$ is the adjoint of the continuous linear operator $g'(x) : X \to Y$.

Under additional conditions, usually referred to as (versions of) "Mangasarian-Fromovitz constraint qualification", the existence of a pair $(y^*, \alpha) \in Y^* \times \mathbb{R}^k$ such that the KKT system is fulfilled is a necessary condition for $x \in X$ to be a local solution of problem (1)–(2). The relations in the last two lines of the KKT system can be equivalently rewritten as

$$
f(x) \in N_{\mathbb{R}^k_+}(\alpha),
$$

where $f = (f_1, \dots, f_k)$, $\mathbb{R}^k_+$ is the set of all elements of $\mathbb{R}^k$ with non-negative components, and the normal cone to the set $\mathbb{R}^k_+$ is defined as usual:

$$
N_{\mathbb{R}^k_+}(\alpha) := \begin{cases} \{\lambda \in \mathbb{R}^k \,:\, \langle \lambda, \beta - \alpha \rangle \le 0 \text{ for all } \beta \in \mathbb{R}^k_+\} & \text{if } \alpha \in \mathbb{R}^k_+,\\ \emptyset & \text{if } \alpha \notin \mathbb{R}^k_+, \end{cases}
$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in $\mathbb{R}^k$. Consequently, one can reformulate the KKT system as

$$
F(x, y^*, \alpha) := \begin{pmatrix} f_0'(x) + \sum_{i=1}^k \alpha_i f_i'(x) + (g'(x))^* y^* \\ g(x) \\ f(x) \end{pmatrix} - \{0\} \times \{0\} \times N_{\mathbb{R}^k_+}(\alpha) \ni 0, \tag{3}
$$

Therefore, $F : X \times Y^* \times \mathbb{R}^k \rightrightarrows Z := X^* \times Y \times \mathbb{R}^k$ is called *optimality mapping*, while its inverse is called (in the case of a finite-dimensional space X, see [4] and [6, p. 134]) *KKT mapping.*

Regularity properties of the mapping F with respect to perturbations are of key importance in the qualitative analysis of optimization problems as (1)–(2), including convergence of numerical methods. In this paper we focus on the so-called *Strong Metric sub-Regularity* (SMSr) (see e.g. [6, Chapter 3.9] and the recent paper [2]).

We recall its definition in terms of the mapping F.

Definition 1.1 *The mapping F is* strongly metric subregular *at $(\hat{x}, \hat{y}^*, \hat{\alpha})$ for zero if $0 \in F(\hat{x}, \hat{y}^*, \hat{\alpha})$ and there exist a number λ and neighborhoods U of $(\hat{x}, \hat{y}^*, \hat{\alpha})$ and V of $0 \in Z$ such that for every $z \in V$ and for every $(x, y^*, \alpha) \in U$ satisfying $z \in F(x, y^*, \alpha)$, it holds that*

$$
\|x - \hat{x}\| + \|y^* - \hat{y}^*\| + \|\alpha - \hat{\alpha}\| \le \lambda \|z\|.
$$

The SMSr property was introduced under this name in [5], but has also been used under several other names (see also [11, Chapter 1] for the related but stronger property of (strong) upper regularity). A more detailed historical account can be found in [2, Section 1].

In the present paper, SMSr of the optimality mapping is proved under *strict Mangasarian-Fromovitz conditions* together with *second-order sufficient conditions* (formulated in Section 3). In the case of finite-dimensional spaces X and Y the result is known from [4, Theorem 2.6] and [2, Section 7.1]. We mention that in the first of the quoted papers also local non-emptiness of F^{-1} is proved, as well as a number of related results that substantially use the finite dimensionality. More about regularity properties of problem (1)–(2) in the finite-dimensional case can be found in [11, Chapter 8] and [1, Chapter 5.2].

Various Lipschitz stability results related to problem (1)–(2) (in Banach spaces) and the associated Lagrange multiplies are obtained in [1, Chapter 4], which, as far as we can see, do not imply the result in the present note.

As stated in the abstract, the Fréchet differentiability assumption involved restricts the applicability of the result in truly infinite-dimensional problems. However, the purpose of this research report is to present a detailed and self-contained proof of the SMSr property of the optimality map. It will provide a basis for our further investigation of the strong metric subregularity of the optimality system associated with problems of calculus of variations and optimal control.

The authors wish to express their gratitude to Asen Dontchev, who made numerous important suggestions for improving the exposition of the paper.

2 Preliminaries

In order to make the exposition more enlightening, in this section we recall some basic facts, mainly concerning systems of linear inequalities and equations.

Let X be a Banach space, X^* its dual space. If Ω is a cone in X, then Ω^* denotes its dual cone, consisting of all linear functionals $x^* \in X^*$ nonnegative on Ω. The following theorem is a simple consequence of the separation theorem (see [7], [3]).

Theorem 2.1 (Dubovitskii - Milyutin) *Let $\Omega_i \subset X$, i=1,...,k be nonempty open convex cones, $\Omega \subset X$ a nonempty convex cone. Then*

$$\left(\bigcap_{i=1}^{k} \Omega_i \right) \bigcap \Omega = \emptyset$$

if and only if there are functionals $x_i^* \in \Omega_i^*$, $i = 1, \ldots, k$ and $x^* \in \Omega^*$, not all equal zero and such that

$$\sum_{i=1}^{k} x_i^* + x^* = 0.$$

Let $l \in X^*$ be a nonzero functional, and $\Omega = \{x \in X : \langle l, x \rangle < 0\}$ an open half-space. It is easy to realize that $x^* \in \Omega^*$ if and only if $x^* = -\alpha l$ with some $\alpha \geq 0$.

Further, let $l_i \in X^*$, $i = 1, \ldots, k$ be nonzero linear functionals. Consider a system of linear inequalities

$$\langle l_i, x \rangle < 0, \quad i = 1, \ldots, k. \tag{4}$$

The following lemma easily follows from the Dubovitskii - Milyutin theorem.

Lemma 2.1 *System (4) is inconsistent if and only if there exist reals* $\alpha_1, \ldots, \alpha_k$ *such that*

$$\alpha_i \geq 0, \quad i = 1, \ldots, k, \quad \sum_{i=1}^{k} \alpha_i > 0, \quad \sum_{i=1}^{k} \alpha_i l_i = 0. \tag{5}$$

We say that l_i, $i = 1, \ldots, k$ are *positively independent* if

$$\alpha_i \geq 0, \quad i = 1, \ldots, k, \quad \sum_{i=1}^{k} \alpha_i l_i = 0 \quad \Rightarrow \quad \alpha_i = 0, \quad i = 1, \ldots, k. \tag{6}$$

Lemma 2.1 implies the following proposition.

Proposition 2.1 *The functionals* l_i, $i = 1, \ldots, k$, *are positively independent if and only if there is an element* $\tilde{x} \in X$ *such that*

$$\langle l_i, \tilde{x} \rangle < 0, \quad i = 1, \ldots, k. \tag{7}$$

We will use the following important estimate which holds for positively independent functionals.

Proposition 2.2 *The functionals* l_i, $i = 1, \ldots, k$, *are positively independent if and only if there is a constant* $c > 0$ *such that*

$$\| \sum_{i=1}^{k} \alpha_i l_i \| \geq c \sum_{i=1}^{k} \alpha_i \quad \forall\, \alpha = (\alpha_1, \ldots, \alpha_k) \in \mathbb{R}^k_+. \tag{8}$$

Proof. If the functionals $l_i \in X^*$, $i = 1, \ldots, k$ are positively dependent, then, obviously, there is no $c > 0$ such that estimate (8) holds.

Conversely, suppose that there is no $c > 0$ such that estimate (8) holds. Then there is a sequence $\alpha^n \in \mathbb{R}^k_+$ such that

$$\sum_{i=1}^{k} \alpha_i^n = 1 \quad \text{and} \quad \| \sum_{i=1}^{k} \alpha_i^n l_i \| \to 0 \quad \text{as} \quad n \to \infty. \tag{9}$$

Without loss of generality we can assume that $\alpha^n \to \alpha \in \mathbb{R}^k_+$ as $n \to \infty$. Passing to the limit in (9), we get $\sum_{i=1}^{k} \alpha_i = 1$ and $\sum_{i=1}^{k} \alpha_i l_i = 0$. The latter means that the system of functionals $l_i \in X^*$, $i = 1, \ldots, k$ is positively dependent. □

In the future, we will use the same notation c for various constants of this kind, hoping that this will not lead to confusion. We will also need the following proposition.

Proposition 2.3 *Let the functionals* $\hat{l}_i \in X^*$, $i = 1, \ldots, k$, *be positively independent and* $\hat{c} > 0$ *be the corresponding constant as in (8). Let the functionals* $l_i \in X^*$, $i = 1, \ldots, k$, *satisfy*

$$\|l_i - \hat{l}_i\| < \varepsilon, \quad i = 1, \ldots, k, \quad 0 < \varepsilon < \hat{c}.$$

Then, for the functionals $l_i \in X^*$, $i = 1, \ldots, k$, *inequality (8) holds with* $c = \hat{c} - \varepsilon$.

Proof. Let $\alpha \in \mathbb{R}^k_+$. Then

$$\|\sum_{i=1}^k \alpha_i l_i\| \geq \|\sum_{i=1}^k \alpha_i \hat{l}_i\| - \|\sum_{i=1}^k \alpha_i (l_i - \hat{l}_i)\| \geq \hat{c}\sum_{i=1}^k \alpha_i - \varepsilon \sum_{i=1}^k \alpha_i.$$

□

Now, let Y be a Banach space, $A : X \to Y$ a surjective linear continuous operator, that is $AX = Y$. In this case the adjoint operator $A^* : Y^* \to X^*$ is injective and has a closed image $A^*Y^* \subset X^*$. By the Banach open mapping theorem the inverse operator $(A^*)^{-1} : A^*Y^* \to Y^*$ is bounded, and hence there is a constant $a > 0$ such that

$$\|A^* y^*\| \geq a\|y^*\|. \tag{10}$$

Each functional of the form $x^* = A^*y^*$ vanishes on $\ker A$. The opposite is also true: if x^* vanishes on $\ker A$, that is $x^* \in (\ker A)^*$, then there exists a uniquely defined functional $y^* \in Y^*$ such that $x^* = A^*y^*$. (Hereafter, for a subspace $L \subset X$, we denote by L^* the set of all linear functionals vanishing on L).

Again, let $l_i \in X^*$, $i = 1, \ldots, k$ be nonzero linear functionals and $A : X \to Y$ a surjective linear continuous operator. Consider a system of linear inequalities and equality

$$\langle l_i, x\rangle < 0, \quad i = 1, \ldots, k, \quad Ax = 0. \tag{11}$$

The Dubovitskii - Milyutin theorem easily implies the following lemma.

Lemma 2.2 *System (11) is inconsistent if and only if there are reals $\alpha_1, \ldots, \alpha_k$ and a functional $y^* \in Y^*$ such that*

$$\alpha_i \geq 0, \quad i = 1, \ldots, k, \quad \sum_{i=1}^k \alpha_i > 0, \quad \sum_{i=1}^k \alpha_i l_i + A^* y^* = 0. \tag{12}$$

For the system of functionals l_i and the operator A consider the following condition

$$\alpha \in \mathbb{R}^k_+,\ y^* \in Y^*,\ \sum_{i=1}^k \alpha_i l_i + A^* y^* = 0 \ \Rightarrow\ \alpha_i = 0,\ i = 1, \ldots, k,\ y^* = 0. \tag{13}$$

It can be easily realized that condition (13) is equivalent to the following one: $AX = Y$ and the functionals $l_i : \ker A \to Y$, $i = 1, \ldots, k$ (the restrictions of the functionals l_i to the subspace $\ker A$) are positively independent. In this case we say that A *is surjective and* l_i, $i = 1, \ldots, k$ *are positively independent on* $\ker A$; then there is $\tilde{x} \in \ker A$ such that $\langle l_i, \tilde{x}\rangle < 0$, $i = 1, \ldots, k$.

The following lemma has the spirit of the so called Hoffman's lemma originally proved in [8] in the case when X is finite dimensional.

The following lemma has the spirit of the so called Hoffman's lemma, originally proved in [8] in the case when X is finite dimensional.

Lemma 2.3 *Let $A : X \to Y$ be a surjective linear continuous operator and $l_i \in X^*$, $i = 1, \ldots, k$, be positively independent on $\ker A$. Then there is a constant $C_H > 0$ such that, for any $\xi = (\xi_1, \ldots, \xi_k) \in \mathbb{R}^k$, $\eta \in Y$ and $x_0 \in X$ satisfying*

$$\langle l_i, x_0\rangle \leq \xi_i, \quad Ax_0 = \eta, \tag{14}$$

there is a solution x' to the system

$$\langle l_i, x_0 + x' \rangle \leq 0, \quad A(x_0 + x') = 0 \tag{15}$$

such that

$$\|x'\| \leq C_H \left(\max\{\xi_1^+, \dots, \xi_k^+\} + \|\eta\| \right), \tag{16}$$

where $\xi_i^+ = \max\{\xi_i, 0\}$.

Proof. Unlike the Hoffman lemma, which does not suppose the positive independence of l_i, the proof of this lemma is rather simple. It is enough to show that there exists a solution x' to the system

$$\langle l_i, x' \rangle + \xi_i \leq 0, \quad i = 1, \dots, k, \quad Ax' = -\eta, \tag{17}$$

satisfying estimate (16). Then $x_0 + x'$ satisfies (15).

Since $AX = Y$, then by the Banach theorem there is a constant $a > 0$ such that for any $y \in Y$ there exists $x \in X$ such that $Ax = y$ and $\|x\| \leq a\|y\|$. Let x_η be such that $Ax_\eta = -\eta$ and $\|x_\eta\| \leq a\|\eta\|$. Let us find $x'' \in \ker A$ such that

$$\langle l_i, x_\eta + x'' \rangle + \xi_i \leq 0, \quad i = 1, \dots, k. \tag{18}$$

Then we can put $x' = x_\eta + x''$, because $A(x_\eta + x'') = Ax_\eta = -\eta$.

Since l_i, $i = 1, \dots, k$ are positively independent on $\ker A$, then, in view of Proposition 2.1, there exists $\tilde{x} \in \ker A$ such that $\langle l_i, \tilde{x} \rangle < -1$, $i = 1, \dots, k$. Set $x'' = \lambda \tilde{x}$ with

$$\lambda = a\|\eta\| \max_i \|l_i\| + \max_i \xi_i^+.$$

Then

$$\langle l_i, x_\eta + x'' \rangle + \xi_i \leq \|l_i\| a \|\eta\| - \lambda + \xi_i^+ \leq 0, \quad i = 1, \dots, k.$$

Moreover, $\|x'\| \leq \|x_\eta\| + \|x''\| \leq a\|\eta\| + \lambda\|\tilde{x}\|$. Consequently, estimate (16) holds with $C_H = a + a\max_i \|l_i\| \|\tilde{x}\| + \|\tilde{x}\|$. □

Now let us prove a proposition similar to Proposition 2.2.

Proposition 2.4 *Suppose that $AX = Y$ and l_i, $i = 1, \dots, k$ are positively independent on* $\ker A$, *that is condition (13) is fulfilled. Then there exists a constant $c > 0$ such that*

$$\| \sum_{i=1}^k \alpha_i l_i + A^* y^* \| \geq c \, (\sum_{i=1}^k \alpha_i + \|y^*\|) \quad \forall \, \alpha \in \mathbb{R}^k_+, \; \forall \, y^* \in Y^*. \tag{19}$$

Proof. Since the condition (19) is positively homogeneous, it suffices to prove it for pairs $(\alpha, y^*) \in \mathbb{R}^k_+ \times Y^*$ such that $\sum_{i=1}^k \alpha_i + \|y^*\| = 1$. Suppose that the proposition is not true. Then there is a sequence $(\alpha_n, y_n^*) \in \mathbb{R}^k_+ \times Y^*$ such that $\sum_{i=1}^k \alpha_{in} + \|y_n^*\| = 1$ and $\| \sum_{i=1}^k \alpha_{in} l_i + A^* y_n^* \| \to 0 \quad (n \to \infty)$. Without loss of generality we can assume that $\alpha_n \to \alpha \in \mathbb{R}^k_+$. Then $\| \sum_{i=1}^k \alpha_i l_i + A^* y_n^* \| \to 0 \quad (n \to \infty)$. Consequently $A^* y_n^*$ strongly converges to some $x^* \in (\ker A)^*$. The latter implies that $x^* = A^* y^*$ with some $y^* \in Y^*$. Then we have that $\|A^* y_n^* - A^* y^*\| \to 0$, whence $\|y_n^* - y^*\| \to 0$ as $n \to \infty$. Consequently, $\sum_{i=1}^k \alpha_i l_i + A^* y^* = 0$ and $\sum_{i=1}^k \alpha_i + \|y^*\| = 1$. We came to a contradiction. □

The following proposition is similar to Proposition 2.3.

Proposition 2.5 *Suppose that $\hat{A}X = Y$ and $\hat{l}_i$, $i = 1, \dots, k$ are positively independent on* $\ker \hat{A}$. *Let $\hat{c} > 0$ be the constant for the system $\hat{l}_1, \dots, \hat{l}_k, \hat{A}$ as in (19). Let the functionals $l_i \in X^*$, $i = 1, \dots, k$ and operator $A : X \to Y$ satisfy*

$$\|l_i - \hat{l}_i\| < \varepsilon, \quad i = 1, \dots, k, \quad \|A - \hat{A}\| < \varepsilon, \quad 0 < \varepsilon < \hat{c}.$$

Then, for the system $l_1, \dots, l_k$, inequality (19) holds with $c = \hat{c} - \varepsilon$.

Proof. Let $\alpha \in \mathbb{R}^k_+$, $y^* \in Y^*$. Then

$$\begin{aligned} \|\sum_{i=1}^k \alpha_i l_i + A^* y^*\| &\geq \|\sum_{i=1}^k \alpha_i \hat{l}_i + \hat{A}^* y^*\| - \|\sum_{i=1}^k \alpha_i (l_i - \hat{l}_i)\| - \|(A^* - \hat{A}^*) y^*\| \\ &\geq \hat{c}(\sum_{i=1}^k \alpha_i + \|y^*\|) - \varepsilon(\sum_{i=1}^k \alpha_i + \|y^*\|). \qquad \square \end{aligned}$$

3 Statement of the problem

In this section we formulate the needed assumptions and some basic facts concerning problem (1)–(2) as describes in the first lines of the Introduction.

Let $\hat{x}$ be an admissible point. Define the *set of active indices*

$$I = \{i \in \{1, \dots, k\} : \ f_i(\hat{x}) = 0\}.$$

Assumption 3.1 (a) $g'(\hat{x})X = Y$. (b) There exists $\tilde{x} \in X$ such that $g'(\hat{x})\tilde{x} = 0$ and $\langle f_i'(\hat{x}), \tilde{x} \rangle < 0 \ \forall\, i \in I$.

In the case where X and Y are finite dimensional, these conditions are often called the *Mangasarian-Fromovitz constraint qualification.*

According to Lemma 2.2, Assumption 3.1 is equivalent to the condition:

$$\alpha_i \geq 0 \ (i \in I), \ y^* \in Y^*, \ \sum_{i \in I} \alpha_i f_i'(\hat{x}) + (g'(\hat{x}))^* y^* = 0 \ \Rightarrow \ \alpha_i = 0 \ (i \in I), \ y^* = 0.$$

We formulate the well-known first-order necessary optimality condition in problem (1)–(2) under Assumption 3.1 (see e.g. Theorem 4 in Chapter 1, [10]).

Theorem 3.1 *If $\hat{x}$ is a local minimum in problem (1)–(2) such that Assumption 3.1 is fulfilled, then there are multipliers $\alpha \in \mathbb{R}^k$ and $y^* \in Y^*$ such that*

$$\alpha \geq 0, \quad \alpha_i f_i(\hat{x}) = 0, \quad i = 1, \dots, k, \tag{20}$$

$$f_0'(\hat{x}) + \sum_{i=1}^k \alpha_i f_i'(\hat{x}) + (g'(\hat{x}))^* y^* = 0. \tag{21}$$

Now fix an admissible point $\hat{x}$ and denote by Λ the set of pairs $(\alpha, y^*) \in \mathbb{R}^k \times Y^*$ such that conditions (20) and (21) hold ($\hat{x}$ is not necessarily assumed to be a solution of (1)–(2)). Assume that $\Lambda \neq \emptyset$ and let us fix an element $(\hat{\alpha}, \hat{y}^*) \in \Lambda$. Define two sets of indices

$$I_0 = \{i \in I : \ \hat{\alpha}_i = 0\}, \qquad I_1 = \{i \in I : \ \hat{\alpha}_i > 0\}.$$

Note that $\hat{\alpha}_i = 0$ for any $i \notin I$. Now we make a stronger assumption than Assumption 3.1.

Assumption 3.2 *The following implication holds for the fixed triple* $(\hat{x}, \hat{\alpha}, \hat{y}^*) \in X \times \Lambda$*:*

$$\alpha \in \mathbb{R}^k, \quad y^* \in Y^*, \quad \alpha_i \geq 0 \ (i \in I_0), \quad \sum_{i \in I} \alpha_i f_i'(\hat{x}) + (g'(\hat{x}))^* y^* = 0$$

$$\Rightarrow \quad \alpha_i = 0 \ (i \in I), \quad y^* = 0. \tag{22}$$

We emphasize that in (22) the signs of α_i for $i \in I_1$ are arbitrary. In the finite dimensional case this condition is known as *strict Mangasarian-Fromovtz condition.*

Condition (22) means that a) $g'(\hat{x})X = Y$, b) the functionals $f_i'(\hat{x})$, $i \in I_1$ are linearly independent on $\ker g'(\hat{x})$, and c) the functionals $f_i'(\hat{x})$, $i \in I_0$ are positively independent on the subspace

$$\{x \in X : \ f_i'(\hat{x})x = 0, \ i \in I_1, \ g'(\hat{x})x = 0\}.$$

It is known that in the finite dimensional case the strict Mangasarian-Fromovtz condition is equivalent to single-valuedness of Λ, see e.g. [12]. This fact is also valid in the Banach space setting.

Lemma 3.1 *Under Assumption 3.2, the set* Λ *is the singleton* $\{(\hat{\alpha}, \hat{y}^*)\}$.

Proof. For $(\hat{\alpha}, \hat{y}^*) \in \Lambda$, we have

$$f_0'(\hat{x}) + \sum_{i=1}^{k} \hat{\alpha}_i f_i'(\hat{x}) + (g'(\hat{x}))^* \hat{y}^* = 0. \tag{23}$$

Take any other pair $(\alpha, y^*) \in \Lambda$. It satisfies conditions (20) and (21). Subtracting (23) from (21) and taking into account the definitions of I_0 and I_1, we get

$$\sum_{i \in I_0} \alpha_i f_i'(\hat{x}) + \sum_{i \in I_1} (\alpha_i - \hat{\alpha}_i) f_i'(\hat{x}) + (g'(\hat{x}))^*(y^* - \hat{y}^*) = 0.$$

In view of (22), it follows that

$$\alpha_i = 0, \quad i \in I_0, \quad \alpha_i - \hat{\alpha}_i = 0, \quad i \in I_1, \quad y^* - \hat{y}^* = 0.$$

□

So, $(\hat{\alpha}, \hat{y}^*)$ is the only element of the set Λ. Introduce the *Lagrange function*

$$L(x, \alpha, y^*) = f_0(x) + \sum_{i=1}^{k} \alpha_i f_i(x) + \langle y^*, g(x) \rangle.$$

We have $L_x(\hat{x}, \hat{\alpha}, \hat{y}^*) = 0$. Taking into account the definitions of I_0 and I_1, define the *critical cone*

$$K = \{\delta x \in X : \ \langle f_i'(\hat{x}), \delta x \rangle \leq 0, \ i \in I_0; \ \langle f_i'(\hat{x}), \delta x \rangle = 0, \ i \in I_1; \ g'(\hat{x})\delta x = 0\}.$$

The following second-order sufficient condition for local optimality is well known, see [13, Corollary 12.1].

Assumption 3.3 *There exists $c_0 > 0$ such that*

$$\Omega(\delta x) := \langle L_{xx}(\hat{x}, \hat{\alpha}, \hat{y}^*)\delta x, \delta x\rangle \geq c_0\|\delta x\|^2 \quad \forall\, \delta x \in K. \tag{24}$$

Theorem 3.2 *Suppose that for an admissible point $\hat{x}$ the set Λ is nonempty and Assumption 3.2 is fulfilled (in this case, Λ is a singleton). Let also Assumption 3.3 be fulfilled. Then the following quadratic growth condition for the cost function f_0 holds at $\hat{x}$: there exist $c > 0$ and $\varepsilon > 0$ such that $f_0(x) - f_0(\hat{x}) \geq c\|x - \hat{x}\|^2$ for all admissible x such that $\|x - \hat{x}\| < \varepsilon$. Hence $\hat{x}$ is a strict local minimizer in problem (1)–(2).*

4 Strong Metric sub-Regularity

In this section we prove strong metric subregularity of the optimality mapping associated with problem (1)–(2) under assumptions formulated below. For that we consider the *perturbed system of optimality conditions*:

$$f_i(x) \leq \xi_i, \quad i = 1, \dots, k, \tag{25}$$

$$g(x) = \eta, \tag{26}$$

$$\alpha_i(f_i(x) - \xi_i) = 0, \quad i = 1, \dots, k, \tag{27}$$

$$\alpha_i \geq 0, \quad i = 1, \dots, k, \tag{28}$$

$$f_0'(x) + \sum_{i=1}^{k} \alpha_i f_i'(x) + (g'(x))^* y^* = \zeta, \tag{29}$$

where $\xi \in \mathbb{R}^k$, $\eta \in Y$, $\zeta \in X^*$.

Theorem 4.1 *Let $(\hat{x}, \hat{\alpha}, \hat{y}^*)$ be a solution of the unperturbed optimality system (25)–(29) (that is, with $\xi = 0$, $\eta = 0$ and $\zeta = 0$) and let assumptions 3.1 –3.3 be fulfilled. Then there are reals $\varepsilon > 0$, $\delta > 0$ and $\lambda > 0$ such that if $|\xi| < \varepsilon$, $\|\eta\| < \varepsilon$, and $\|\zeta\| < \varepsilon$, then for any solution (x, α, y^*) of the perturbed system (25)-(29) such that $\|x - \hat{x}\| < \delta$, the following estimates hold:*

$$\|x - \hat{x}\| \leq \lambda(|\xi| + \|\eta\| + \|\zeta\|),$$

$$|\alpha - \hat{\alpha}| + \|y^* - \hat{y}^*\| \leq \lambda(|\xi| + \|\eta\| + \|\zeta\|).$$

We shall reformulate the above theorem in terms or SMSr (Definition 1.1). To shorten the notation, we denote $\Xi := X \times Y^* \times \mathbb{R}^k$, $\hat{s} = (\hat{x}, \hat{y}^*, \hat{\alpha})$. We also remind that the definition of the optimality mapping F is given in (3).

Theorem 4.2 *Let $0 \in F(\hat{s})$, and let assumptions 3.1 –3.3 be fulfilled for $\hat{s}$. Then the mapping $F : \Xi \rightrightarrows Z$ is strongly metrically subregular at $\hat{s}$. Moreover, the neighborhood U in Definition 1.1 can be taken of the form $I\!\!B_X(\hat{x}; \delta) \times Y^* \times \mathbb{R}^k$, where $I\!\!B_X(\hat{x}; \delta)$ is the ball in X centered at $\hat{x}$ with radius $\delta > 0$.*

Remark 4.1 The optimality map F defined in (3) is a sum of a Fréchet differentiable function $\varphi(s)$ and a normal cone, call it $N(s)$. According to Corollary 2.2 and Remark 2.4 in [2] the SMSr of this mapping at a point $\hat{s}$ for zero is equivalent to the same property for the partially linearized mapping, $\varphi(\hat{s}) + \varphi'(\hat{s})(s - \hat{s})$. Notice that if assumptions 3.1–3.3 are fulfilled for the mapping F, they also hold for the corresponding linearized functions. Dealing with linear functions only makes the proof easier. However, we do not make use of this fact and present a direct proof working with the possibly nonlinear functions, partly repeating in this way the argument behind Corollary 2.2 in [2]. This allows to obtain the last claim of Theorem 4.2. Moreover, the proof remains self-contained.

Proof. Let us analyze the perturbed system (25)-(29). Let x, α, y^* be a solution to this system for given ξ, η, ζ. Set $\Delta x = x - \hat{x}$. Since $f(x) \to f(\hat{x})$ as $\|\Delta x\| \to 0$, by complementary slackness conditions (27) we have: there exists $\delta > 0$ such that $\alpha_i = 0$ for all $i \notin I$, and hence $\Delta\alpha_i := \alpha_i - \hat{\alpha}_i = 0$ for all $i \notin I$, whenever $\|\Delta x\| < \delta$.

Assumption 3.2 implies that the functionals $f_i'(\hat{x})$, $i \in I$, are positively independent on $\ker g'(\hat{x})$. Then, according to Proposition 2.5 and the continuity of f_i' and g', there exists $\delta > 0$ and a constant $c > 0$ such that for every $x \in X$ with $\|x - \hat{x}\| \leq \delta$

$$\Big\| \sum_{i \in I} \alpha_i f_i'(x) + (g'(x))^* y^* \Big\| \geq c\Big(\sum_{i \in I} \alpha_i + \|y^*\| \Big).$$

Hence, there exists a constant C such that

$$|\alpha| + \|y^*\| \leq C \tag{30}$$

whenever $\|\Delta x\| < \delta$. This implies that $\Delta\alpha = \alpha - \hat{\alpha}$ and $\Delta y^* = y^* - \hat{y}^*$ are also bounded.

Subtracting (23) from (29) we obtain

$$f_0'(x) - f_0'(\hat{x}) + \sum_{i=1}^{k} (\alpha_i f_i'(x) - \hat{\alpha}_i f_i'(\hat{x})) + (g'(x))^* y^* - (g'(\hat{x}))^* \hat{y}^* = \zeta. \tag{31}$$

Here

$$f_0'(x) - f_0'(\hat{x}) = \langle f_0''(\hat{x}), \Delta x \rangle + o(\|\Delta x\|),$$
$$\alpha_i f_i'(x) - \hat{\alpha}_i f_i'(\hat{x}) = \hat{\alpha}_i (f_i'(x) - f_i'(\hat{x})) + (\Delta\alpha_i) f_i'(\hat{x}) + (\Delta\alpha_i)(f_i'(x) - f_i'(\hat{x}))$$
$$= \hat{\alpha}_i f_i''(\hat{x})\Delta x + (\Delta\alpha_i) f_i'(\hat{x}) + (\Delta\alpha_i) f_i''(\hat{x})\Delta x + o(\|\Delta x\|).$$

Similarly,

$$(g'(x))^* y^* - (g'(\hat{x}))^* \hat{y}^* = (g''(\hat{x})\Delta x)^* \hat{y}^* + (g'(\hat{x}))^* (\Delta y^*) + (g''(\hat{x})\Delta x)^* (\Delta y^*) + o(\|\Delta x\|).$$

Using these relations in (31), we get

$$L_{xx}(\hat{x}, \hat{\alpha}, \hat{y}^*)\Delta x + \sum_{i=1}^{k} (\Delta\alpha_i) f_i'(\hat{x}) + (g'(\hat{x}))^* (\Delta y^*)$$

$$+ \sum_{i=1}^{k} (\Delta\alpha_i) f_i''(\hat{x})\Delta x + (g''(\hat{x})\Delta x)^* (\Delta y^*) + o(\|\Delta x\|) = \zeta. \tag{32}$$

If $i \notin I$, then $\alpha_i = \hat{\alpha}_i = 0$ and $\Delta\alpha_i = 0$. Using that $I = I_0 \cup I_1$, we represent this equation in the form

$$\sum_{i \in I_0} (\Delta\alpha_i) f_i'(\hat{x}) + \sum_{i \in I_1} (\Delta\alpha_i) f_i'(\hat{x}) + (g'(\hat{x}))^*(\Delta y^*)$$

$$= -L_{xx}(\hat{x}, \hat{\alpha}, \hat{y}^*)\Delta x - \sum_{i=1}^{k} (\Delta\alpha_i) f_i''(\hat{x})\Delta x - (g''(\hat{x})\Delta x)^*(\Delta y^*) - o(\|\Delta x\|) + \zeta. \tag{33}$$

Let A be the operator which takes each $x \in X$ to the tuple

$$\big(\langle f_i'(\hat{x}), x\rangle, i \in I_1, g'(\hat{x})x\big) \in \mathbb{R}^{|I_1|} \times Y,$$

where $|I_1|$ is a number of elements of I_1. Due to Assumption 3.2, this operator is surjective, and the functionals $l_i = f_i'(\hat{x})$, $i \in I_0$ are positively independent on its kernel. Applying Proposition 2.4 to this system of functionals and operator and taking into account that all $\Delta\alpha_i$ and Δy^* are bounded, we obtain from (33) that

$$\sum_{i=1}^{k} |\Delta\alpha_i| + \|\Delta y^*\| \le c_1(\|\Delta x\| + \|\zeta\|) \tag{34}$$

with some $c_1 > 0$.

Recall that $\langle L''(\hat{x}, \hat{\alpha}, \hat{y}^*)\Delta x, \Delta x\rangle =: \Omega(\Delta x)$. Then, 'multiplying' (33) by Δx, we get

$$\Omega(\Delta x) + \sum_{i=1}^{k} (\Delta\alpha_i)\langle f_i'(\hat{x}), \Delta x\rangle + \langle (g'(\hat{x}))^*(\Delta y^*), \Delta x\rangle$$

$$+ \sum_{i=1}^{k} (\Delta\alpha_i)\langle f_i''(\hat{x})\Delta x, \Delta x\rangle + \langle (\Delta y^*) g''(\hat{x})\Delta x, \Delta x\rangle$$

$$+ o(\|\Delta x\|^2) = \langle \zeta, \Delta x\rangle. \tag{35}$$

Now we use conditions (25) and (27). Let us show that if $\varepsilon > 0$ and $\delta > 0$ are small enough and $\|\Delta x\| < \delta$, $\|\zeta\| < \varepsilon$, then

$$(\Delta\alpha_i)(f_i(x) - \xi_i) = 0 \tag{36}$$

for all $i = 1, \ldots, k$. It is enough to prove this equalities for $i \in I = I_0 \cup I_1$, because for $i \notin I$ we have $\Delta\alpha_i = 0$.

First let us show this for $i \in I_1$. In view of (34) and the conditions $\|\Delta x\| < \delta$ and $\|\zeta\| < \varepsilon$, the vector $\Delta\alpha$ can be regarded as arbitrary small. Then, for $i \in I_1$ we have: $\alpha_i := \hat{\alpha}_i + \Delta\alpha_i > 0$ (because $\hat{\alpha}_i > 0$ and $\Delta\alpha_i$ is arbitrary small). Then the complementary slackness conditions (27) implies

$$f_i(x) - \xi_i = 0, \quad i \in I_1. \tag{37}$$

whenever $\varepsilon > 0$ and $\delta > 0$ are small enough. Hence (36) follows for all $i \in I_1$ and for $\varepsilon > 0$ and $\delta > 0$ small enough.

For $i \in I_0$ we have: $\hat{\alpha}_i = 0$, consequently, $\alpha_i = \Delta\alpha_i$ and then (27) implies (36). Thus, (36) is proved for all $i = 1, \ldots, k$, provided that $\varepsilon > 0$ and $\delta > 0$ are small enough.

Consequently,

$$\begin{aligned}\sum_{i=1}^{k}(\Delta\alpha_i)\Big(\langle f_i'(\hat{x}),\Delta x\rangle-\xi_i\Big) &= \sum_{i=1}^{k}(\Delta\alpha_i)(f_i(x)-f_i(\hat{x})-\xi_i)+|\Delta\alpha|O(\|\Delta x\|^2)\\ &= \sum_{i=1}^{k}(\Delta\alpha_i)f_i(\hat{x})+|\Delta\alpha|O(\|\Delta x\|^2)=|\Delta\alpha|O(\|\Delta x\|^2),\end{aligned}$$

hence,

$$\sum_{i=1}^{k}(\Delta\alpha_i)\langle f_i'(\hat{x}),\Delta x\rangle=\langle\Delta\alpha,\xi\rangle+|\Delta\alpha|O(\|\Delta x\|^2). \tag{38}$$

Using (38) in (35), we get

$$\begin{aligned}&\Omega(\Delta x)+\langle\Delta\alpha,\xi\rangle+|\Delta\alpha|O(\|\Delta x\|^2)+\langle(g'(\hat{x}))^*(\Delta y^*),\Delta x\rangle\\ &+\sum_{i=1}^{k}(\Delta\alpha_i)\langle f_i''(\hat{x})\Delta x,\Delta x\rangle+\langle(\Delta y^*)g''(\hat{x})\Delta x,\Delta x\rangle\\ &+o(\|\Delta x\|^2)=\langle\zeta,\Delta x\rangle.\end{aligned} \tag{39}$$

Equalities (37), the inequalities $f_i(x)\le\xi_i$, $i\in I_0$, and equality (26) imply, respectively,

$$\begin{aligned}\langle f_i'(\hat{x}),\Delta x\rangle &= \xi_i+O(\|\Delta x\|^2), \quad i\in I_1,\\ \langle f_i'(\hat{x}),\Delta x\rangle &\le \xi_i+O(\|\Delta x\|^2), \quad i\in I_0,\\ g'(\hat{x})\Delta x &= \eta+O(\|\Delta x\|^2).\end{aligned}$$

Then, by Lemma 2.3, there exist a constant $C_H>0$ and a correction x' such that

$$\langle f_i'(\hat{x}),\Delta x+x'\rangle=0, \quad i\in I_1, \tag{40}$$

$$\langle f_i'(\hat{x}),\Delta x+x'\rangle\le 0, \quad i\in I_0, \tag{41}$$

$$g'(\hat{x})(\Delta x+x')=0, \tag{42}$$

and, moreover,

$$\begin{aligned}\|x'\| &\le C_H\left(\sum_{i\in I_0}\xi_i^+ +\sum_{i\in I_1}|\xi_i|+\|\eta\|\right)+O(\|\Delta x\|^2)\\ &\le C_H(|\xi|+\|\eta\|)+O(\|\Delta x\|^2).\end{aligned} \tag{43}$$

Relations (40)-(42) imply that

$$\delta x:=\Delta x+x'\in K, \tag{44}$$

and then, by Assumption 3.3, $\Omega(\delta x)\ge c_0\|\delta x\|^2$. Let us compare $\|\delta x\|^2$ with $\|\Delta x\|^2$ and $\Omega(\delta x)$ with $\Omega(\Delta x)$, respectively. We have

$$\|\delta x\|^2=\|\Delta x\|^2+r, \tag{45}$$

where $|r| \le 2\|\Delta x\|\|x'\| + \|x'\|^2$. According to (43),

$$\|\Delta x\|\|x'\| \le \|\Delta x\|\Big(C_H(|\xi| + \|\eta\|) + O(\|\Delta x\|^2)\Big)$$

$$= C_H\|\Delta x\|(|\xi| + \|\eta\|) + o(\|\Delta x\|^2),$$

$$\|x'\|^2 \le 2C_H^2(|\xi| + \|\eta\|)^2 + o(\|\Delta x\|^2)$$

(here we used: $(a+b)^2 \le 2a^2 + 2b^2$). Consequently, there is $c_r > 0$ such that

$$|r| \le c_r(|\xi| + \|\eta\|)(\|\Delta x\| + |\xi| + \|\eta\|) + o(\|\Delta x\|^2). \tag{46}$$

Similarly, there is $c_\Omega > 0$ such that

$$\Omega(\delta x) = \Omega(\Delta x) + r_\Omega, \tag{47}$$

where

$$|r_\Omega| \le c_\Omega(|\xi| + \|\eta\|)(\|\Delta x\| + |\xi| + \|\eta\|) + o(\|\Delta x\|^2). \tag{48}$$

Hence, the inequality $c_0\|\delta x\|^2 \le \Omega(\delta x)$ implies

$$c_0(\|\Delta x\|^2 + r) \le \Omega(\Delta x) + r_\Omega. \tag{49}$$

Moreover, the relations $g'(\hat{x})\delta x = 0$ and $\delta x = \Delta x + x'$ imply

$$\langle (g'(\hat{x}))^*(\Delta y^*), \Delta x\rangle = -\langle (g'(\hat{x}))^*(\Delta y^*), x'\rangle,$$

whence

$$|\langle (g'(\hat{x}))^*(\Delta y^*), \Delta x\rangle| \le \|g'(\hat{x})\|\|\Delta y^*\|\Big(C_H(|\xi| + \|\eta\|) + O(\|\Delta x\|^2)\Big). \tag{50}$$

Using (49) and (50) in (39) and estimating from above the norm of each term, we obtain

$$\begin{aligned} c_0\|\Delta x\|^2 \le\ & c_0|r| + |r_\Omega| \\ &+|\Delta\alpha||\xi| + |\Delta\alpha| O(\|\Delta x\|^2) \\ &+\|g'(\hat{x})\|\|\Delta y^*\|\Big(C_H(|\xi| + \|\eta\|) + O(\|\Delta x\|^2)\Big) \\ &+|\Delta\alpha|\Big(\sum_{i=1}^k \|f_i''(\hat{x})\|\Big)\|\Delta x\|^2 + \|\Delta y^*\|\|g''(\hat{x})\|\|\Delta x\|^2 \\ &+\|\zeta\|\|\Delta x\| + o(\|\Delta x\|^2), \end{aligned} \tag{51}$$

Using (34), (46) and (48) in this inequality, we get: there exists $c_0' > 0$ such that

$$c_0'\|\Delta x\|^2 \le (|\xi| + \|\eta\|)(\|\Delta x\| + |\xi| + \|\eta\|)$$

$$+(\|\Delta x\| + \|\zeta\|)(|\xi| + \|\eta\| + \|\Delta x\|^2) + \|\zeta\|\|\Delta x\|.$$

Set $\omega = (\xi, \eta, \zeta)$, $\|\omega\| = |\xi| + \|\eta\| + \|\zeta\|$. From the previous inequality we easily obtain: there exist $\varepsilon > 0$, $\delta > 0$, and $c_0'' > 0$ such that if $\|\omega\| < \varepsilon$ and $\|\Delta x\| < \delta$, then $c_0''\|\Delta x\|^2 \le \|\Delta x\|\|\omega\| + \|\omega\|^2$, whence

$$\hat{c}\|\Delta x\| \le \|\omega\| \tag{52}$$

with $\hat{c} = \frac{1}{2}\left(\sqrt{4c_0'' + 1} - 1\right)$. Together with (34) this completes the proof of the theorem. $\square$

References

[1] J.F. Bonnans and A Shapiro. *Perturbation analysis of optimization problems.* Springer, 2000.

[2] R. Cibulka, A.L. Dontchev, and A.Y. Kruger. Strong metric subregularity of mappings in variational analysis and optimization. *Journal of Mathematical Analysis and Application*, **457**:1247–1282, 2018.

[3] A.V. Dmitruk and N.P. Osmolovskii A General Lagrange Multipliers Theorem and Related Questions, Control Systems and Mathematical Methods in Economics (Feichtinger et al. eds.), Lecture Notes in Economics and Mathematical Systems, **687**:165–194, Berlin, Springer, 2018.

[4] A. L. Dontchev, R. T. Rockafellar. Characterizations of Lipschitz stability in nonlinear programming *Mathematical programming with data perturbations*, 65–82, Lecture Notes in Pure and Appl. Math., 195, Dekker, New York, 1998.

[5] A. L. Dontchev, R. T. Rockafellar. Regularity and conditioning of solution mappings in variational analysis. *Set-Valued Analysis*, **12**:79–109, 2004.

[6] A.L. Dontchev and R.T. Rockafellar. *Implicit Functions and Solution Mappings: A View from Variational Analysis. Second edition.* Springer, New York , 2014.

[7] A. Ya. Dubovitskii, and A.A. Milyutin. Extremum problems in the presence of restrictions, *USSR Comput. Math. and Math. Phys.,* **5**(3):1–80, 1965.

[8] A.J. Hoffman, On approximate solutions of systems of linear inequalities. *J. Res. Nat'l Bur. Standarts* **49**:263–265, 1952.

[9] A.D. Ioffe. *Variational analysis and regular mappings.* Springer, 2017.

[10] A.D Ioffe and V.M. Tikhomorov. *Theory of extremal problems.* Nauka, Moscow, 1974 (in Russian); English translation: North Holland, 1979.

[11] D. Klatte and B. Kummer. *Nonsmooth equations in optimization.* Kluwer Academic Publisher, 2002.

[12] J. Kyparisis. On uniqueness of Kuhn-Tucker multiplies in non-linear programming. *Math. Programming*, **32**:242–246, 1985.

[13] E.S. Levitin, A.A. Milyutin and N.P. Osmolovskii, Higher-order local minimum conditions in problems with constraints. *Uspekhi Mat. Nauk*, **33**:8–148, 1978; English translation in *Russian Math. Surveys*, **33**:9–168, 1978.