sciendo
Control and Cybernetics
vol. **50** (2021) No. 4
pages: 457- 471
DOI: 10.2478/candc-2021-0027

On the strong metric subregularity in mathematical programming* †

by

Nikolai P. Osmolovskii[1] **and Vladimir M. Veliov**[2]

[1]Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
osmolov@ibspan.waw.pl
[2]Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Austria,
vladimir.veliov@tuwien.ac.at

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 a finite number of inequalities. The conditions, under which SMSr is proven, assume that the data are twice continuously Fréchet differentiable, the strict Mangasarian-Fromovitz constraint qualification is satisfied, and the second-order sufficient optimality condition holds. The obtained result extends the one known for finite-dimensional problems. Although the applicability of the result is limited to the Banach space setting (due to the twice Fréchet differentiability assumptions and the finite number of inequality constraints), the paper can be valuable due to the self-contained exposition, and provides a ground for extensions. One possible extension was recently implemented in Osmolovskii and Veliov (2021).

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

1. Introduction

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

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

*Submitted: May 2021; Accepted: September 2021
†This research is supported by the Austrian Science Foundation (FWF) under grant No I4571.

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

$$\min\ f_0(x) \tag{1}$$
$$\text{subject to } g(x) = 0, \quad f_i(x) \le 0 \ \ (i = 1, \dots, k). \tag{2}$$

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

$$f_0'(x) + \sum_{i=1}^{k} \alpha_i f_i'(x) + (g'(x))^* y^* = 0, \tag{3}$$
$$g(x) = 0, \tag{4}$$
$$\alpha_i f_i(x) = 0, \quad i = 1, \dots, k, \tag{5}$$
$$f_i(x) \le 0, \quad \alpha_i \ge 0, \quad i = 1, \dots, k, \tag{6}$$

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

$$F(x, y^*, \alpha) := \begin{pmatrix} f_0'(x) + \sum_{i=1}^{k} \alpha_i f_i'(x) + (g'(x))^* y^* \\ g(x) \\ f(x) \end{pmatrix} - \{0\} \times \{0\} \times N_{\mathbb{R}^k_+}(\alpha) \ni 0. \tag{7}$$

Therefore, $F : S := X \times Y^* \times \mathbb{R}^k \rightrightarrows Z := X^* \times Y \times \mathbb{R}^k$ is called *optimality mapping*, while its inverse is called (in the case of a finite-dimensional space X, see Dontchev and Rockafellar, 1998 and 2014, p. 134) *KKT mapping*.

The regularity properties of the mapping F with respect to perturbations are of key importance in the qualitative analysis of optimization problems, such as (1)–(2), including convergence of numerical methods. In this paper we focus on the so-called *Strong Metric Subregularity* (SMSr) (see, e.g., Dontchev and Rockafellar, 2014, Chapter 3.9, and the recent paper by Cibulka, Dontchev and Kruger, 2018). We recall the definition for a set-valued mapping F acting from a metric space (S, d_S) to another metric space (Z, d_Z): F is called SMSr at $\bar{s} \in S$ for $\bar{z} \in Z$ if $\bar{z} \in F(\bar{s})$ and there is a constant λ along with neighborhoods U of $\bar{s}$ and V of $\bar{z}$ such that

$$d_S(s, \bar{s}) \leq \lambda \inf_{z \in F(s) \cap V} d_Z(z, \bar{z}) \quad \text{for all } s \in U.$$

The SMSr property is equivalent to isolated calmness of the inverse mapping F^{-1} at $\bar{z}$ (see Dontchev and Rockafellar, 2014, Chapter 3.9). We use this fact in the definition below, which is given in terms of the specific mapping F, and the spaces S and Z defined around (7).

Definition 1 *The mapping F is* strongly metrically 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*

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

The SMSr property was introduced under this name in Dontchev and Rockafellar (2004), but has also been used under several other names (see also Klatte and Kummer, 2002, Chapter 1, for the related but stronger property of (strong) upper regularity). A more detailed historical account can be found in Cibulka, Dontchev and Kruger (2018), Section 1.

In the present paper, SMSr of the optimality mapping is proven 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 Dontchev and Rockafellar (1998), Theorem 2.6, and Cibulka, Dontchev and Kruger (2018), Section 7.1. We mention that in the first of the quoted papers also local non-emptiness of F^{-1} is proven, as well as a number of related results that substantially use the finite dimensionality. More about the regularity properties of the problem (1)–(2) in the finite-dimensional case can be found in Klatte and Kummer (2002), Chapter 8, and Bonnans and Shapiro (2000), Chapter 5.2.

Various Lipschitz stability results, related to problem (1)–(2) (in Banach spaces), and the associated Lagrange multipliers are obtained in Bonnans and Shapiro (2000), Chapter 4, which, as far as we can see, do not imply the result, given in the present note.

The twice Fréchet differentiability assumption involved restricts the applicability of the result in infinite-dimensional problems. However, the purpose of this paper is to present a detailed, straightforward and self-contained proof of the SMSr property of the optimality map. It serves as the basis for our further study of the strong metric subregularity of the optimal system, already implemented in Osmolovskii and Veliov (2021) for the general problem of the calculus of variations and awaiting its implementation for nonlinear optimal control problems. In the present paper, we would like to highlight a narrower perspective than in Osmolovskii and Veliov (2021), and thus to provide shorter and simpler proofs.

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 Dubovitskii and Milyutin, 1965, and Dmitruk and Osmolovskii, 2018).

Theorem 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 $x_i^* \in \Omega_i^*$, $i = 1, \dots, k$ 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 $\alpha \geq 0$. 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

$$\|A^* y^*\| \geq a\|y^*\|.$$

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).

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

$$\langle l_i, x\rangle < 0, \quad i = 1, \dots, k, \quad Ax = 0. \tag{8}$$

The Dubovitskii - Milyutin theorem easily implies the following lemma.

LEMMA 1 *System (8) is inconsistent if and only if there are reals* $\alpha_1, \dots, \alpha_k$ *and a functional* $y^* \in Y^*$ *such that*

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

Recall that a system of linear functionals $l_i \in X^*$ $(i = 1, \dots, k)$ is said to be *positively independent on* X if

$$\alpha \in \mathbb{R}^k_+, \ \sum_{i=1}^{k} \alpha_i l_i = 0 \ \Rightarrow \ \alpha = 0.$$

For a system of linear functionals $l_i \in X^*$ $(i = 1, \dots, k)$ and a linear continuous operator $A : X \to Y$ with a closed image AX consider the following condition

$$\alpha \in \mathbb{R}^k_+, \ y^* \in Y^*, \ \sum_{i=1}^{k} \alpha_i l_i + A^* y^* = 0 \ \Rightarrow \ \alpha = 0, \ y^* = 0. \tag{9}$$

It can be easily realized that condition (9) is equivalent to the following one: $AX = Y$ and the functionals $l_i : \ker A \to Y$, $i = 1, \dots, k$ (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, $i = 1, \dots, k$ *are positively independent on* $\ker A$. Then, by Lemma 1, there is $\tilde{x} \in \ker A$ such that $\langle l_i, \tilde{x}\rangle < 0$, $i = 1, \dots, k$.

PROPOSITION 1 *Suppose that* $AX = Y$ *and* l_i, $i = 1, \dots, k$ *are positively independent on* $\ker A$*, that is – condition (9) is fulfilled. Then there exists a constant* $c > 0$ *such that*

$$\Big\| \sum_{i=1}^{k} \alpha_i l_i + A^* y^* \Big\| \geq c \left(\sum_{i=1}^{k} \alpha_i + \|y^*\| \right) \quad \forall \, \alpha \in \mathbb{R}^k_+, \ \forall \, y^* \in Y^*. \tag{10}$$

PROOF Since the condition (10) 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

$$\sum_{i=1}^{k} \alpha_{in} + \|y^*_n\| = 1 \text{ and } \| \sum_{i=1}^{k} \alpha_{in} l_i + A^* y^*_n \| \to 0 \quad (n \to \infty).$$

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,

$$\sum_{i=1}^k \alpha_i l_i + A^* y^* = 0 \text{ and } \sum_{i=1}^k \alpha_i + \|y^*\| = 1.$$

We came to a contradiction. □

Proposition 2 *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 (10). 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 $l_1, \dots, l_k, A$ *inequality (10) holds with* $c = \hat{c} - \varepsilon$.

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

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

The following lemma has the spirit of the so called Hoffman's lemma, originally proven in Hoffman (1952) in the case when X is finite dimensional.

Lemma 2 (Ioffe, 2017, Theorem 3) *Let* $A : X \to Y$ *be a surjective linear continuous operator and* $l_i \in X^*$, $i = 1, \dots, k$. *Then, there is a constant* $C_H > 0$ *such that, for any* $\xi = (\xi_1, \dots, \xi_k) \in \mathbb{R}^k$, $\eta \in Y$ *and* $x_0 \in X$ *satisfying*

$$\langle l_i, x_0 \rangle \leq \xi_i, \quad Ax_0 = \eta,$$

there is a solution x' *to the system*

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

such that

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

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

3. Statement of the problem

In this section we formulate the assumptions needed and provide some basic facts concerning problem (1)–(2).

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

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

Assumption 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, in which X and Y are finite dimensional, these conditions are often called the *Mangasarian-Fromovitz constraint qualification*.

According to Lemma 1, Assumption 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 1 (see, e.g., Theorem 4 in Chapter 1 of Ioffe and Tikhomirov, 1974).

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

$$\alpha \geq 0, \quad \alpha_i f_i(\hat{x}) = 0, \quad i = 1, \ldots, k, \tag{11}$$

$$f_0'(\hat{x}) + \sum_{i=1}^{k} \alpha_i f_i'(\hat{x}) + (g'(\hat{x}))^* y^* = 0. \tag{12}$$

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 (11) and (12) 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 1.

Assumption 2 *The following implication holds:*

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

$$\Rightarrow \quad \alpha_i = 0 \;(i \in I), \quad y^* = 0. \tag{13}$$

We emphasize that on the left-hand side of the implication (13) the signs of α_i for $i \in I_1$ are not prescribed. In the finite dimensional case this condition is known as *strict Mangasarian-Fromovitz condition*.

Condition (13) 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-Fromovitz condition is equivalent to single-valuedness of Λ, see, e.g., Kyparisis (1985). This fact is also valid in the Banach space setting.

Lemma 3 *Under Assumption 2, the set* Λ *is the singleton* $\{(\hat{\alpha}, \hat{y}^*)\}$.

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

$$f_0'(\hat{x}) + \sum_{i=1}^{k} \hat{\alpha}_i f_i'(\hat{x}) + (g'(\hat{x}))^* \hat{y}^* = 0. \tag{14}$$

Take any other pair $(\alpha, y^*) \in \Lambda$. It satisfies conditions (11) and (12). Subtracting (14) from (12) 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 (2), it follows that

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

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*

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

The following second-order sufficient condition for local optimality is well known (Levitin, Milyutin and Osmolovskii, 1978):

Assumption 3 *There exists $c_0 > 0$ such that*

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

Theorem 3 *Suppose that for an admissible point $\hat{x}$ the set Λ is nonempty and Assumption 2 is fulfilled (in this case, Λ is a singleton). Let also Assumption 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 this, along with the original *unperturbed system of optimality conditions* (3)–(6), we consider the *perturbed system of optimality conditions*:

$$f_i(x) \leq \xi_i, \quad i = 1, \ldots, k, \tag{15}$$

$$g(x) = \eta, \tag{16}$$

$$\alpha_i(f_i(x) - \xi_i) = 0, \quad i = 1, \ldots, k, \tag{17}$$

$$\alpha_i \geq 0, \quad i = 1, \ldots, k, \tag{18}$$

$$f_0'(x) + \sum_{i=1}^{k} \alpha_i f_i'(x) + (g'(x))^* y^* = \zeta, \tag{19}$$

where $\xi \in \mathbb{R}^k$, $\eta \in Y$, $\zeta \in X^*$. Note that the original system (3)–(6) corresponds to zero values of the parameters ξ, η, ζ in (15)–(19).

Theorem 4 *Let $(\hat{x}, \hat{\alpha}, \hat{y}^*)$ be a solution of the unperturbed optimality system (15)–(19) (that is, with $\xi = 0$, $\eta = 0$ and $\zeta = 0$) and let Assumptions 1–3 be fulfilled. Then, there are reals $\varepsilon > 0$, $\delta > 0$ and $\lambda > 0$ such that if $|\xi| < \varepsilon$, $\|\eta\| < \varepsilon$, and $\|\zeta\| < \varepsilon$, and then, for any solution (x, α, y^*) of the perturbed system (15)-(19), such that $\|x - \hat{x}\| < \delta$, the following estimate holds:*

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

We shall reformulate the above theorem in terms of SMSr (Definition 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 (7).

Theorem 5 *Let $0 \in F(\hat{s})$, and let Assumptions 1–3 be fulfilled for $\hat{s}$. Then, the mapping $F : \Xi \rightrightarrows Z$ is strongly metrically subregular at $\hat{s}$ for zero. Moreover, the neighborhood U in Definition 1 can be taken of the form $\mathbb{B}_X(\hat{x}; \delta) \times Y^* \times \mathbb{R}^k$, where $\mathbb{B}_X(\hat{x}; \delta)$ is the ball in X centered at $\hat{x}$ with radius $\delta > 0$.*

Remark 1 The optimality map F, defined in (7), 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 Cibulka, Dontchev and Kruger (2018), 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 1–3 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 Cibulka, Dontchev and Kruger (2018). This allows for obtaining the last claim of Theorem 5. Moreover, the proof remains self-contained.

Proof Let us analyze the perturbed system (15)-(19). 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 (17) we have: there exist $\delta > 0$ and $\varepsilon > 0$ 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 $\|\Delta x\| < \delta$ and $|\xi| < \varepsilon$.

Assumption 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 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 $\|x - \hat{x}\| \le \delta$

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

This and equality (19) implies that there exists a constant C such that

$$|\alpha| + \|y^*\| \le C$$

whenever $\|\Delta x\| < \delta$ and $\|\zeta\| < \varepsilon$. Therefore, $\Delta\alpha = \alpha - \hat{\alpha}$ and $\Delta y^* = y^* - \hat{y}^*$ are also bounded.

Subtracting (14) from (19) we obtain

$$f_0'(x) - f_0'(\hat{x}) + \sum_{i=1}^{k} (\alpha_i f_i'(x) - \hat{\alpha}_i f_i'(\hat{x})) + (g'(x))^* y^* - (g'(\hat{x}))^* \hat{y}^* = \zeta. \quad (20)$$

Set

$$x - \hat{x} = \Delta x, \quad y^* - \hat{y}^* = \Delta y^*.$$

Then

$$f_0'(x) - f_0'(\hat{x}) = \langle f_0''(\hat{x}), \Delta x \rangle + o(\|\Delta x\|),$$

$$\begin{aligned} \alpha_i f_i'(x) - \hat{\alpha}_i f_i'(\hat{x}) &= \hat{\alpha}_i (f_i'(x) - f_i'(\hat{x})) + (\Delta\alpha_i) f_i'(\hat{x}) + (\Delta\alpha_i)(f_i'(x) - f_i'(\hat{x})) \\ &= \hat{\alpha}_i f_i''(\hat{x}) \Delta x + (\Delta\alpha_i) f_i'(\hat{x}) + (\Delta\alpha_i) f_i''(\hat{x}) \Delta x + o(\|\Delta x\|). \end{aligned}$$

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 (20), we get

$$L_{xx}(\hat{x}, \hat{\alpha}, \hat{y}^*)\Delta x + \sum_{i=1}^{k}(\Delta\alpha_i) f_i'(\hat{x}) + (g'(\hat{x}))^*(\Delta y^*)$$
$$+ \sum_{i=1}^{k}(\Delta\alpha_i) f_i''(\hat{x})\Delta x + (g''(\hat{x})\Delta x)^*(\Delta y^*) + o(\|\Delta x\|) = \zeta. \tag{21}$$

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

$$\sum_{i\in I_0}(\Delta\alpha_i) f_i'(\hat{x}) + \sum_{i\in I_1}(\Delta\alpha_i) f_i'(\hat{x}) + (g'(\hat{x}))^*(\Delta y^*)$$
$$= -L_{xx}(\hat{x}, \hat{\alpha}, \hat{y}^*)\Delta x - \sum_{i=1}^{k}(\Delta\alpha_i) f_i''(\hat{x})\Delta x - (g''(\hat{x})\Delta x)^*(\Delta y^*) - o(\|\Delta x\|) + \zeta. \tag{22}$$

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 the number of elements of I_1. Due to Assumption 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 1 to this system of functionals and operator and taking into account that all $\Delta\alpha_i$ and Δy^* are bounded, we obtain from (22) that

$$\sum_{i=1}^{k}|\Delta\alpha_i| + \|\Delta y^*\| \le c_1(\|\Delta x\| + \|\zeta\|) \tag{23}$$

with some $c_1 > 0$. Hereinafter we assume that $\|\Delta x\| < \delta$ and $\|\zeta\| < \varepsilon$.

Recall that $\langle L''(\hat{x}, \hat{\alpha}, \hat{y}^*)\Delta x, \Delta x\rangle =: \Omega(\Delta x)$. Then, 'multiplying' (22) by Δx, we get

$$\Omega(\Delta x) + \sum_{i=1}^{k}(\Delta\alpha_i)\langle f_i'(\hat{x}), \Delta x\rangle + \langle (g'(\hat{x}))^*(\Delta y^*), \Delta x\rangle$$
$$+ \sum_{i=1}^{k}(\Delta\alpha_i)\langle f_i''(\hat{x})\Delta x, \Delta x\rangle + \langle (\Delta y^*) g''(\hat{x})\Delta x, \Delta x\rangle$$
$$+ o(\|\Delta x\|^2) = \langle \zeta, \Delta x\rangle. \tag{24}$$

Now we use conditions (15) and (17). Let us show that if $\varepsilon > 0$ and $\delta > 0$ are small enough and $\|\Delta x\| < \delta$, $\|\zeta\| < \varepsilon$, then

$$(\Delta\alpha_i)(f_i(x) - \xi_i) = 0 \tag{25}$$

for all $i = 1, \dots, k$. It is enough to prove these 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 (23) and the conditions $\|\Delta x\| < \delta$ and $\|\zeta\| < \varepsilon$, the vector $\Delta\alpha$ can be regarded as arbitrarily 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 arbitrarily small). Then, the complementary slackness conditions (17) imply

$$f_i(x) - \xi_i = 0, \quad i \in I_1, \tag{26}$$

whenever $\varepsilon > 0$ and $\delta > 0$ are small enough. Hence, (25) 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 (17) implies (25). Thus, (25) is proven for all $i = 1, \dots, k$, 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,

$$\sum_{i=1}^{k}(\Delta\alpha_i)\langle f_i'(\hat{x}), \Delta x\rangle = \langle \Delta\alpha, \xi\rangle + |\Delta\alpha| O(\|\Delta x\|^2). \tag{27}$$

Using (27) in (24), we get

$$\begin{aligned}&\Omega(\Delta x) + \langle \Delta\alpha, \xi\rangle + |\Delta\alpha| O(\|\Delta x\|^2) + \langle (g'(\hat{x}))^*(\Delta y^*), \Delta x\rangle \\ &+ \sum_{i=1}^{k}(\Delta\alpha_i)\langle f_i''(\hat{x})\Delta x, \Delta x\rangle + \langle (\Delta y^*) g''(\hat{x})\Delta x, \Delta x\rangle + o(\|\Delta x\|^2) \\ &= \langle \zeta, \Delta x\rangle.\end{aligned} \tag{28}$$

Equalities (26), the inequalities $f_i(x) \le \xi_i$, $i \in I_0$, and equality (16) 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, there exist a constant $C_H > 0$ and a correction x' such that

$$\langle f_i'(\hat{x}), \Delta x + x' \rangle = 0, \quad i \in I_1, \tag{29}$$

$$\langle f_i'(\hat{x}), \Delta x + x' \rangle \le 0, \quad i \in I_0, \tag{30}$$

$$g'(\hat{x})(\Delta x + x') = 0, \tag{31}$$

and, moreover,

$$\begin{aligned} \|x'\| &\le C_H\Big(\sum_{i\in I_0}\xi_i^+ + \sum_{i\in I_1}|\xi_i| + \|\eta\|\Big) + O(\|\Delta x\|^2) \\ &\le C_H\big(|\xi| + \|\eta\|\big) + O(\|\Delta x\|^2). \end{aligned} \tag{32}$$

Relations (29)-(31) imply that

$$\delta x := \Delta x + x' \in K,$$

and then, by Assumption 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

$$\|\delta x\|^2 = \|\Delta x\|^2 + r,$$

where $|r| \le 2\|\Delta x\|\|x'\| + \|x'\|^2$. According to (32),

$$\begin{aligned} \|\Delta x\|\|x'\| &\le \|\Delta x\|\Big(C_H(|\xi| + \|\eta\|) + O(\|\Delta x\|^2)\Big) \\ &= C_H\|\Delta x\|\big(|\xi| + \|\eta\|\big) + o(\|\Delta x\|^2), \\ \|x'\|^2 &\le 2C_H^2(|\xi| + \|\eta\|)^2 + o(\|\Delta x\|^2) \end{aligned}$$

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

$$|r| \le c_r(|\xi| + \|\eta\|)(\|\Delta x\| + |\xi| + \|\eta\|) + o(\|\Delta x\|^2). \tag{33}$$

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

$$\Omega(\delta x) = \Omega(\Delta x) + r_\Omega,$$

where

$$|r_\Omega| \le c_\Omega(|\xi| + \|\eta\|)(\|\Delta x\| + |\xi| + \|\eta\|) + o(\|\Delta x\|^2). \tag{34}$$

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

$$c_0(\|\Delta x\|^2 + r) \le \Omega(\Delta x) + r_\Omega. \tag{35}$$

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

$$|\langle (g'(\hat{x}))^*(\Delta y^*), \Delta x\rangle| \le \|g'(\hat{x})\| \|\Delta y^*\| \Big(C_H(|\xi| + \|\eta\|) + O(\|\Delta x\|^2) \Big). \quad (36)$$

Using (35) and (36) in (28) and estimating from above the norm of each term, we obtain

$$\begin{aligned} c_0\|\Delta x\|^2 \quad \le \quad & c_0|r| + |r_\Omega| + |\Delta\alpha||\xi| + |\Delta\alpha| O(\|\Delta x\|^2) \\ & + \|g'(\hat{x})\| \|\Delta y^*\| \Big(C_H(|\xi| + \|\eta\|) + O(\|\Delta x\|^2) \Big) \\ & + |\Delta\alpha| \Big(\sum_{i=1}^{k} \|f_i''(\hat{x})\| \Big) \|\Delta x\|^2 + \|\Delta y^*\| \|g''(\hat{x})\| \|\Delta x\|^2 \\ & + \|\zeta\| \|\Delta x\| + o(\|\Delta x\|^2). \end{aligned} \quad (37)$$

Using (23), (33) and (34) in this inequality, we get: there exists $c_0' > 0$ such that

$$\begin{aligned} c_0'\|\Delta x\|^2 \quad \le \quad & \big(|\xi| + \|\eta\| \big) \big(\|\Delta x\| + |\xi| + \|\eta\| \big) \\ & + (\|\Delta x\| + \|\zeta\|)(|\xi| + \|\eta\| + \|\Delta x\|^2) + \|\zeta\| \|\Delta x\|. \end{aligned}$$

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

$$\hat{c}\|\Delta x\| \le \|\omega\| \quad (38)$$

with $\hat{c} = \frac{1}{2}\left(\sqrt{4c_0'' + 1} - 1 \right)$. Together with (23), this completes the proof of the theorem. □

References

Bonnans, J.F. and Shapiro, A. (2000) *Perturbation Analysis of Optimization Problems*. Springer.

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

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

Dontchev, A. L. and Rockafellar, R. T. (1998) 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.

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

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

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

Hoffman, A.J. (1952) On approximate solutions of systems of linear inequalities. *J. Res. Nat'l Bur. Standards* **49**: 263–265.

Ioffe, A.D. (1979) Regular points of Lipschitz functions. *Trans. Am. Math. Soc.* 251, 61–69.

Ioffe, A.D. (2017) *Variational Analysis and Regular Mappings*. Springer.

Ioffe, A.D and Tikhomorov, V.M. (1974) *Theory of Extremal Problems.* Nauka, Moscow (in Russian); English translation: North Holland, 1979.

Klatte, D. and Kummer, B. (2002) *Nonsmooth Equations in Optimization.* Kluwer Academic Publisher.

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

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

Osmolovskii, N.P. and Veliov, V.M. (2021) On the strong subregularity of the optimality mapping in mathematical programming and calculus of variations. *J. Math. Anal. Appl.* **500**(1), Article 125077.