A NOTE ON SOLUTION SENSITIVITY FOR KARUSH–KUHN–TUCKER SYSTEMS*

A. F. Izmailov[†] and M. V. Solodov[‡]

September 3, 2004

ABSTRACT

We consider Karush-Kuhn-Tucker (KKT) systems, which depend on a parameter. Our contribution concerns with the existence of solution of the directionally perturbed KKT system, approximating the given primal-dual base solution. To our knowledge, we give the first explicit result of this kind in the situation where the multiplier associated with the base primal solution may not be unique. The condition we employ can be interpreted as the 2-regularity property of a smooth reformulation of the KKT system. We also give a strictly sharper, compared to other statements in the literature, estimate for the contingent derivative of the KKT solution multifunction.

Key words. KKT system, perturbation, sensitivity analysis, 2-regularity.
AMS subject classifications. 90C30, 65K05.

---

* Research of the first author is supported by the Russian Foundation for Basic Research Grant 04-01-00341 and the RF President's Grant NS-1815.2003.1 for the support of leading scientific schools. The second author is supported in part by CNPq Grants 300734/95-6(RN) and 471780/2003-0, and by FAPERJ.

[†] Moscow State University, Faculty of Computational Mathematics and Cybernetics, Department of Operations Research, Leninskiye Gori, GSP-2, 119992 Moscow, Russia.
Email: `izmaf@ccas.ru`

[‡] Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ 22460-320, Brazil.
Email: `solodov@impa.br`

1 Introduction

Let $\Phi : \mathbf{R}^s \times \mathbf{R}^n \to \mathbf{R}^n$ and $G : \mathbf{R}^s \times \mathbf{R}^n \to \mathbf{R}^m$ be sufficiently smooth mappings. We consider the parametric *Karush–Kuhn–Tucker* (KKT) *system*: Find $(x,\, \mu) \in \mathbf{R}^n \times \mathbf{R}^m$ such that

$$\begin{gathered} \Phi(\sigma,\, x) - \left(\frac{\partial G}{\partial x}(\sigma,\, x)\right)^{\mathrm{T}} \mu = 0, \\ \mu \geq 0, \quad G(\sigma,\, x) \geq 0, \quad \langle \mu,\, G(\sigma,\, x)\rangle = 0, \end{gathered} \tag{1.1}$$

where $\sigma \in \mathbf{R}^s$ is a parameter, and $\langle \cdot, \cdot \rangle$ is the Euclidean inner product.

System (1.1) is a case of the mixed complementarity problem with a special (primal-dual) structure. If for some smooth function $f : \mathbf{R}^s \times \mathbf{R}^n \to \mathbf{R}$ it holds that

$$\Phi(\sigma,\, x) = \frac{\partial f}{\partial x}(\sigma,\, x), \quad \sigma \in \mathbf{R}^s,\, x \in \mathbf{R}^n, \tag{1.2}$$

then, as is well-known, (1.1) is the KKT optimality system for the parametric optimization problem

$$\begin{array}{ll} \text{minimize} & f(\sigma,\, x) \\ \text{subject to} & x \in D(\sigma), \end{array} \tag{1.3}$$

where

$$D(\sigma) = \{x \in \mathbf{R}^n \mid G(\sigma,\, x) \geq 0\}. \tag{1.4}$$

We note that all the developments given below extend in a straightforward manner to the case when equality constraints are present.

Let $\mathcal{KKT}$ be the set comprised by all triples $(\sigma,\, x,\, \mu) \in \mathbf{R}^s \times \mathbf{R}^n \times \mathbf{R}^m$ satisfying (1.1). We define the KKT solution multifunction by

$$KKT : \mathbf{R}^s \to 2^{\mathbf{R}^n \times \mathbf{R}^m}, \quad KKT(\sigma) = \{(x,\, \mu) \in \mathbf{R}^n \times \mathbf{R}^m \mid (\sigma,\, x,\, \mu) \in \mathcal{KKT}\}.$$

For a given (base) parameter value $\bar{\sigma} \in \mathbf{R}^s$, let $(\bar{x},\, \bar{\mu}) \in KKT(\bar{\sigma})$. The sensitivity theory is concerned with the local structure of the set $\mathcal{KKT}$ or, to put in other words, with the behavior of the multifunction KKT for the values of $\sigma \in \mathbf{R}^s$ close to the base value $\bar{\sigma}$.

There are two principal issues in stability/sensitivity analysis, which are to some extent independent of each other (and are typically considered separately in the sensitivity literature). One problem is that of *approximation* of the base solution by the solutions of the perturbed problems. It concerns with the properties of the map KKT, *assuming* that solutions exist, at least for some forms of perturbations. (This assumption is usually not made explicitly, but without it, the sensitivity statements become vacuous.) Such studies usually deal with the question whether the set $KKT(\sigma)$ approximates in some sense the set $KKT(\bar{\sigma})$ as $\sigma \to \bar{\sigma}$ ("stability"), and give some quantitative characterization of the approximation properties ("sensitivity"). Sensitivity information concerning the KKT multifunction can be presented in various (equivalent) forms. One relevant object is the *contingent cone* $C_{\mathcal{KKT}}(\bar{\sigma},\, \bar{x},\, \bar{\mu})$ to the set $\mathcal{KKT}$ at the point $(\bar{\sigma},\, \bar{x},\, \bar{\mu})$, or the (smaller, in general) *tangent cone* $T_{\mathcal{KKT}}(\bar{\sigma},\, \bar{x},\, \bar{\mu})$. In the terminology of [16], the multifunction from $\mathbf{R}^s$ to $2^{\mathbf{R}^n \times \mathbf{R}^m}$ whose graph coincides with $C_{\mathcal{KKT}}(\bar{\sigma},\, \bar{x},\, \bar{\mu})$, is called the *contingent* (*outer graphical*) *derivative* of KKT at $\bar{\sigma}$ for $(\bar{x},\, \bar{\mu})$. Moreover, KKT is said to be *protodifferentiable* at $\bar{\sigma}$ for $(\bar{x},\, \bar{\mu})$ if

$C_{\mathcal{KKT}}(\bar{\sigma},\, \bar{x},\, \bar{\mu}) = T_{\mathcal{KKT}}(\bar{\sigma},\, \bar{x},\, \bar{\mu})$. The essence of this branch of sensitivity analysis can therefore be stated in terms of the contingent and tangent directions. In Section 3, using the the notion of 2-regularity [6, 4], we present an estimate of the contingent derivative of the KKT multifunction, which is sharper than other statements in the literature.

The second major issue of stability/sensitivity analysis is that of *existence* of solutions of the perturbed problems, i.e., whether $KKT(\sigma) \neq \emptyset$ for a given (or all) $\sigma \in \mathbf{R}^s$ close enough to $\bar{\sigma}$. This issue had been studied by many authors under quite mild assumptions (in particular, not implying the uniqueness of the multiplier associated with the base primal solution); see, for example, [20, 12, 23, 21, 13, 22, 17, 18, 8], and the recently published books [11, 3]. (Note that the existence results usually appear in conjunction with some kind of assertions on approximation properties, as discussed above.) In this paper, we are concerned with the following more specific question: we are looking for mild conditions guaranteeing, for *given* primal-dual base solution $(\bar{x},\, \bar{\mu})$, the existence of an arc of solutions of the form $(x(t),\, \mu(t)) = (\bar{x}+t\xi,\, \bar{\mu}+t\nu)+o(t)$ corresponding to the parameter values $\sigma(t) = \bar{\sigma}+td+\rho(t)$, $t \geq 0$, where $d \in \mathbf{R}^s$ and $(\xi,\, \nu) \in \mathbf{R}^n \times \mathbf{R}^m$. To our knowledge, this kind of analysis was previously known only in the context of Robinson's strong regularity, see [1, 16]. Strong regularity implies that $(\bar{x},\, \bar{\mu})$ is an isolated point of $KKT(\bar{\sigma})$, and in particular, $\bar{\mu}$ is the unique multiplier associated with $\bar{x}$. In Section 2, we prove *existence* results for directional perturbations under conditions which do not require the uniqueness of the multiplier. Those conditions can be interpreted in terms of the 2-regularity property [6, 4] of a certain smooth reformulation of the KKT system. The latter relation is discussed in Section 3.

A few words about our notation. Given a finite set I, $|I|$ stands for its cardinality. For $y \in \mathbf{R}^m$ and an index set $I \subset \{1,\, \ldots,\, m\}$, y_I stands for the vector with components y_i, $i \in I$. For a matrix (linear operator) Λ, $\mathrm{im}\, \Lambda$ is its range (image space), and $\ker \Lambda$ is its kernel (null space). For a directionally differentiable mapping $F : \mathbf{R}^n \to \mathbf{R}^m$, by $F'(x;\, d)$ we denote the usual directional derivative of F at $x \in \mathbf{R}^n$ in the direction $d \in \mathbf{R}^n$. Given a set D in $\mathbf{R}^n$, the contingent cone to D at a point $x \in D$ is given by $C_D(x) = \{\xi \in \mathbf{R}^n \mid \exists \{t_k\} \subset \mathbf{R}$ such that $\{t_k\} \to 0+,\ \mathrm{dist}(x + t_k\xi,\, D) = o(t_k)\}$. The tangent cone to D at x is defined as $T_D(x) = \{\xi \in \mathbf{R}^n \mid \mathrm{dist}(x + t\xi,\, D) = o(t),\ t \geq 0\}$, where $\mathrm{dist}(x,\, D) = \inf_{z \in D} \|x - z\|$.

For the base value $\bar{\sigma}$ and the given $\bar{x}$, we define the index sets associated with the active and inactive constraints in the usual way:

$$\begin{gathered} I = I(\bar{\sigma},\, \bar{x}) = \{i = 1,\, \ldots,\, m \mid G_i(\bar{\sigma},\, \bar{x}) = 0\}, \\ N = N(\bar{\sigma},\, \bar{x}) = \{1,\, \ldots,\, m\} \setminus I. \end{gathered} \tag{1.5}$$

For a given $\bar{\mu}$ such that $(\bar{\sigma},\, \bar{x},\, \bar{\mu}) \in \mathcal{KKT}$, the active constraints are further partitioned into the weakly and strongly active, as follows:

$$\begin{gathered} I_0 = I_0(\bar{\sigma},\, \bar{x},\, \bar{\mu}) = \{i \in I \mid \bar{\mu}_i = 0\}, \\ I_+ = I_+(\bar{\sigma},\, \bar{x},\, \bar{\mu}) = I \setminus I_0 = \{i \in I \mid \bar{\mu}_i > 0\}. \end{gathered} \tag{1.6}$$

We next state some constraint qualifications and second-order conditions that will be used in the paper. All those conditions are associated with the nonperturbed KKT system (i.e., for the base value $\bar{\sigma}$ of the parameter).

- The *linear independence constraint qualification* (LICQ): $\mathrm{rank}\, \frac{\partial G_I}{\partial x}(\bar{\sigma},\, \bar{x}) = |I|$.

- The *Mangasarian–Fromovitz constraint qualification* (MFCQ):
  $\exists\, h \in \mathbf{R}^n$ such that $\frac{\partial G_I}{\partial x}(\bar{\sigma},\, \bar{x})h > 0$.

- The *strict Mangasarian–Fromovitz constraint qualification* (SMFCQ):
  rank $\frac{\partial G_{I_+}}{\partial x}(\bar{\sigma},\, \bar{x}) = |I_+|$ and $\exists\, h \in \mathbf{R}^n$ such that $\frac{\partial G_{I_0}}{\partial x}(\bar{\sigma},\, \bar{x})h > 0$, $\frac{\partial G_{I_+}}{\partial x}(\bar{\sigma},\, \bar{x})h = 0$.

- The *weak linear independence constraint qualification* (WLICQ):
  rank $\frac{\partial G_{I_+}}{\partial x}(\bar{\sigma},\, \bar{x}) = |I_+|$.

As is well known, SMFCQ is equivalent to the uniqueness of the multiplier.

For the sake of convenience, we define the mapping associated with the equations in (1.1) (leaving out the equation associated with the complementarity condition):

$$\Psi : \mathbf{R}^s \times \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^n, \quad \Psi(\sigma,\, x,\, \mu) = \Phi(\sigma,\, x) - \left(\frac{\partial G}{\partial x}(\sigma,\, x)\right)^{\mathrm{T}} \mu.$$

The second-order conditions have the form

$$\left\langle \frac{\partial \Psi}{\partial x}(\bar{\sigma},\, \bar{x},\, \bar{\mu})\xi,\, \xi \right\rangle \neq 0 \quad \forall\, \xi \in K \setminus \{0\},$$

with different choices of the cone $K \subset \mathbf{R}^n$:

- The *strong second-order sufficiency condition* (SSOSC) uses $K = \{\xi \in \mathbf{R}^n \mid \frac{\partial G_{I_+}}{\partial x}(\bar{\sigma},\, \bar{x})\xi = 0\}$.

- The *second-order condition* (SOC) uses $K = \{\xi \in \mathbf{R}^n \mid \frac{\partial G_I}{\partial x}(\bar{\sigma},\, \bar{x})\xi = 0\}$.

Note that the second-order conditions mean that $\langle \frac{\partial \Psi}{\partial x}(\bar{\sigma},\, \bar{x},\, \bar{\mu})\xi,\, \xi \rangle$ has the same sign for all ξ in the corresponding K.

2 Existence of Solutions under Directional Perturbations

Let the cone $\mathcal{L} = \mathcal{L}(\bar{\sigma},\, \bar{x},\, \bar{\mu})$ be the solution set (with respect to $(d,\, \xi,\, \nu) \in \mathbf{R}^s \times \mathbf{R}^n \times \mathbf{R}^m$) of the following "linearization" of the KKT system (1.1):

$$\begin{gathered} \frac{\partial \Psi}{\partial \sigma}(\bar{\sigma},\, \bar{x},\, \bar{\mu})d + \frac{\partial \Psi}{\partial x}(\bar{\sigma},\, \bar{x},\, \bar{\mu})\xi - \left(\frac{\partial G}{\partial x}(\bar{\sigma},\, \bar{x})\right)^{\mathrm{T}} \nu = 0, \\ \nu_i \geq 0,\, \langle G_i'(\bar{\sigma},\, \bar{x}),\, (d,\, \xi)\rangle \geq 0,\, \nu_i \langle G_i'(\bar{\sigma},\, \bar{x}),\, (d,\, \xi)\rangle = 0,\, i \in I_0, \\ \nu_N = 0, \quad G_{I_+}'(\bar{\sigma},\, \bar{x})(d,\, \xi) = 0. \end{gathered} \tag{2.1}$$

The following inclusion is well known:

$$C_{\mathcal{KKT}}(\bar{\sigma},\, \bar{x},\, \bar{\mu}) \subset \mathcal{L}. \tag{2.2}$$

For the special case of directional perturbations, this fact is stated in [1, Theorem 5.10] The same result, but in terms of the contingent derivative of KKT at $\bar{\sigma}$ for $(\bar{x},\, \bar{\mu})$, was given in [16, Proposition 2.5.1]. The early related references are [15, 19, 14], and some recent related statements can be found in [10, 8, 9, 11], and also [24, Lemma 4.1 and Theorem 5.1].

In this section, for a given triple $(d, \xi, \nu) \in \mathcal{L}$ and a given mapping $\rho : \mathbf{R}_+ \to \mathbf{R}^s$ such that $\rho(t) = o(t)$, we consider the arc of the form $\bar{\sigma} + td + \rho(t)$ in the space of parameters, and solutions of the form $(\bar{x} + t\xi, \bar{\mu} + t\nu) + o(t)$ of the corresponding perturbed KKT system. We are concerned with the existence and uniqueness of such solutions for the values $t \geq 0$ small enough.

As mentioned above, the question of the existence and uniqueness of solutions of perturbed KKT systems, approximating the given primal-dual base solution, was previously studied in the context of Robinson's strong regularity; see [1, 16]. Recall that in the optimization setting, strong regularity is equivalent to LICQ combined with SSOSC [1, Proposition 5.38]. In particular, it implies uniqueness of the multiplier associated with the given $\bar{x}$ for the base value $\bar{\sigma}$ of the parameter. The assumptions of the existence Theorem 2.1 below do not presume uniqueness of the multiplier. This issue will be further illustrated by examples in Section 4. We also note that the assumptions of Corollary 2.1 below are actually a certain 2-regularity property. We state the assumptions here in the algebraic form, leaving their conceptual interpretation until the next section.

For any partition (I_1, I_2) of I_0 (i.e., a pair of index sets such that $I_1 \cup I_2 = I_0$, $I_1 \cap I_2 = \emptyset$), define the branch $\mathcal{KKT}_{(I_1, I_2)} = \mathcal{KKT}_{(I_1, I_2)}(\bar{\sigma}, \bar{x}, \bar{\mu})$ of the set $\mathcal{KKT}$, as the solution set of the following system:

$$\begin{array}{c} \Phi(\sigma, x) - \left(\frac{\partial G}{\partial x}(\sigma, x)\right)^{\mathrm{T}} \mu = 0, \\ \mu_{I_1} \geq 0, \quad G_{I_1}(\sigma, x) = 0, \\ \mu_{I_2} = 0, \quad G_{I_2}(\sigma, x) \geq 0, \\ \mu_N = 0, \quad G_{I_+}(\sigma, x) = 0. \end{array}$$

As is easy to see, near $(\bar{\sigma}, \bar{x}, \bar{\mu})$, the set $\mathcal{KKT}$ can be represented as the union of such branches for (the finite number of) all the possible partitions. Similarly, the cone $\mathcal{L}$ is the union of the branches $\mathcal{L}_{(I_1, I_2)} = \mathcal{L}_{(I_1, I_2)}(\bar{\sigma}, \bar{x}, \bar{\mu})$, given by

$$\begin{array}{c} \frac{\partial \Psi}{\partial \sigma}(\bar{\sigma}, \bar{x}, \bar{\mu})d + \frac{\partial \Psi}{\partial x}(\bar{\sigma}, \bar{x}, \bar{\mu})\xi - \left(\frac{\partial G}{\partial x}(\bar{\sigma}, \bar{x})\right)^{\mathrm{T}} \nu = 0, \\ \nu_{I_1} \geq 0, \quad G'_{I_1}(\bar{\sigma}, \bar{x})(d, \xi) = 0, \\ \nu_{I_2} = 0, \quad G'_{I_2}(\bar{\sigma}, \bar{x})(d, \xi) \geq 0, \\ \nu_N = 0, \quad G'_{I_+}(\bar{\sigma}, \bar{x})(d, \xi) = 0. \end{array}$$

Define the following index sets associated with the given triple (d, ξ, ν):

$$\begin{array}{l} I_0^0 = I_0^0(d, \xi, \nu) = \{i \in I_0 \mid \nu_i = 0, \ \langle G'_i(\bar{\sigma}, \bar{x}), (d, \xi)\rangle = 0\}, \\ I_0^+ = I_0^+(d, \xi, \nu) = \{i \in I_0 \mid \nu_i > 0, \ \langle G'_i(\bar{\sigma}, \bar{x}), (d, \xi)\rangle = 0\}, \\ I_0^N = I_0^N(d, \xi, \nu) = \{i \in I_0 \mid \nu_i = 0, \ \langle G'_i(\bar{\sigma}, \bar{x}), (d, \xi)\rangle > 0\}. \end{array} \tag{2.3}$$

Let (I_0^1, I_0^2) be any partition of I_0^0. Then (I_1, I_2), with $I_1 = I_0^1 \cup I_0^+$ and $I_2 = I_0^2 \cup I_0^N$, is a partition of I_0. Furthermore, $(d, \xi, \nu) \in \mathcal{L}_{(I_1, I_2)}$. The needed result makes use of Gollan's regularity condition [2] at $(\bar{\sigma}, \bar{x}, \bar{\mu})$ for the constraints defining the branch $\mathcal{KKT}_{(I_1, I_2)}$. After some computations, this condition can be expressed in the form

$$\det \begin{pmatrix} \frac{\partial \Psi}{\partial x}(\bar{\sigma}, \bar{x}, \bar{\mu}) & -\left(\frac{\partial G_{I_1 \cup I_+}}{\partial x}(\bar{\sigma}, \bar{x})\right)^{\mathrm{T}} \\ \frac{\partial G_{I_1 \cup I_+}}{\partial x}(\bar{\sigma}, \bar{x}) & 0 \end{pmatrix} \neq 0, \tag{2.4}$$

$$\exists\,(\bar{\xi},\,\bar{\nu}) \in \mathbf{R}^n \times \mathbf{R}^m \text{ such that } (d,\,\bar{\xi},\,\bar{\nu}_{I_1\cup I_+}) \in \ker \Lambda_{(I_1,\,I_2)},\ \bar{\nu}_{I_1} > 0,\ G'_{I_2}(\bar{\sigma},\,\bar{x})(d,\,\bar{\xi}) > 0, \tag{2.5}$$

where

$$\Lambda_{(I_1,\,I_2)} = \begin{pmatrix} \frac{\partial \Psi}{\partial \sigma}(\bar{\sigma},\,\bar{x},\,\bar{\mu}) & \frac{\partial \Psi}{\partial x}(\bar{\sigma},\,\bar{x},\,\bar{\mu}) & -\left(\frac{\partial G_{I_1\cup I_+}}{\partial x}(\bar{\sigma},\,\bar{x})\right)^{\mathrm{T}} \\ \frac{\partial G_{I_1\cup I_+}}{\partial \sigma}(\bar{\sigma},\,\bar{x}) & \frac{\partial G_{I_1\cup I_+}}{\partial x}(\bar{\sigma},\,\bar{x}) & 0 \end{pmatrix}. \tag{2.6}$$

The following theorem is now implied by [1, Theorem 3.4].

Theorem 2.1 *For $(\bar{\sigma},\,\bar{x},\,\bar{\mu}) \in \mathcal{KKT}$ and a given $(d,\,\xi,\,\nu) \in \mathcal{L}$, assume that Gollan's condition holds at $(\bar{\sigma},\,\bar{x},\,\bar{\mu})$ for the constraints defining the branch $\mathcal{KKT}_{(I_0^1\cup I_0^+,\,I_0^2\cup I_0^N)}$ for some partition $(I_0^1,\,I_0^2)$ of I_0^0.*

Then for every mapping $\rho : \mathbf{R}_+ \to \mathbf{R}^s$ such that $\rho(t) = o(t)$, there exists a mapping $r : \mathbf{R}_+ \to \mathbf{R}^n \times \mathbf{R}^m$ such that for $t \geq 0$ it holds that $(\bar{x}+t\xi,\,\bar{\mu}+t\nu)+r(t) \in KKT(\bar{\sigma}+td+\rho(t))$, $r(t) = o(t)$.

Under the additional assumption that $I_0^0 = \emptyset$, the existence result in Theorem 2.1 can be complemented by the following uniqueness result.

Corollary 2.1 *For $(\bar{\sigma},\,\bar{x},\,\bar{\mu}) \in \mathcal{KKT}$ and a given $(d,\,\xi,\,\nu) \in \mathcal{L}$, assume that $I_0^0 = \emptyset$ (see (2.3)) and*

$$\det \begin{pmatrix} \frac{\partial \Psi}{\partial x}(\bar{\sigma},\,\bar{x},\,\bar{\mu}) & -\left(\frac{\partial G_{I_0^+\cup I_+}}{\partial x}(\bar{\sigma},\,\bar{x})\right)^{\mathrm{T}} \\ \frac{\partial G_{I_0^+\cup I_+}}{\partial x}(\bar{\sigma},\,\bar{x}) & 0 \end{pmatrix} \neq 0. \tag{2.7}$$

Then for every mapping $\rho : \mathbf{R}_+ \to \mathbf{R}^s$ such that $\rho(t) = o(t)$, and for every $t > 0$ small enough, there exists the unique element $r(t) \in \mathbf{R}^n \times \mathbf{R}^m$ such that $(\bar{x} + t\xi,\,\bar{\mu} + t\nu) + r(t) \in KKT(\bar{\sigma} + td + \rho(t))$, $r(t) = o(t)$.

Proof. The assumptions of Theorem 2.1 are certainly satisfied here. Indeed, the equality $I_0^0 = \emptyset$ implies that the only suitable partition of I_0 is $(I_1,\,I_2) = (I_0^+,\,I_0^N)$, and for this partition (2.4) reduces to (2.7). Moreover, from (2.1), (2.3) and (2.6), it follows that (2.5) holds, e.g., with $\bar{\xi} = \xi$, $\bar{\nu} = \nu$.

It remains to show that for all $t > 0$ small enough, the element $r(t) = (x(t),\,\mu(t)) \in \mathbf{R}^n \times \mathbf{R}^m$ defined according to Theorem 2.1 is unique. From (1.5), (1.6) and (2.3) it follows that this element satisfies

$$(\bar{\mu} + t\nu + \mu(t))_{I_0^+\cup I_+} > 0,$$
$$G_{I_0^N\cup N}(\bar{\sigma} + td + \rho(t),\,\bar{x} + t\xi + x(t)) > 0,$$

and hence, by necessity,

$$\Psi(\bar{\sigma} + td + \rho(t),\,\bar{x} + t\xi + x(t),\,\bar{\mu} + t\nu + \mu(t)) = 0,$$
$$G_{I_0^+\cup I_+}(\bar{\sigma} + td + \rho(t),\,\bar{x} + t\xi + x(t)) = 0,$$
$$\mu_{I_0^N\cup N}(t) = 0.$$

Condition (2.7) evidently means that the Jacobian of the latter system of equations with respect to (x, μ) is nonsingular. This implies the needed uniqueness. ■

The equality $I_0^0 = \emptyset$ can be interpreted as the strict complementarity condition at the solution (d, ξ, ν) of the "linearized" KKT system (2.1) defining $\mathcal{L}$. Under this condition, the following two "limit cases" can be pointed out:

- If $I_0^+ = I_0$ (i.e., $I_0^N = \emptyset$), then (2.7) implies LICQ, and is implied by LICQ combined with SOC.
- If $I_0^N = I_0$ (i.e., $I_0^+ = \emptyset$), then (2.7) implies WLICQ, and is implied by WLICQ combined with SSOSC.

We omit the proofs, as they are quite direct.

3 Connections with 2-regularity and the contingent derivative

In this section, we exhibit the connections between some of the key conditions which appeared above and the property of 2-regularity of a nonlinear mapping [6, 4]. We also obtain a sharper estimate for the contingent derivative of the KKT multifunction, see Proposition 3.1.

When introducing 2-regularity, we simplify the setting to what is needed in the context of this paper. In particular, we state everything in finite dimensions. Let the following hypotheses be satisfied:

(H1) Z and W are (finite-dimensional) Euclidean spaces, $L(Z, W)$ is the space of linear operators from Z to W, V is a neighborhood of a point $\bar{z}$ in Z.

(H2) $F : V \to W$ is Fréchet-differentiable on V, and the mapping $F' : V \to L(Z, W)$ is continuous at $\bar{z}$.

(H3) $W_1 = \operatorname{im} F'(\bar{z})$, W_2 is some complementary subspace of W_1 in W, P is the projector in W onto W_2 parallel to W_1.

(H4) The mapping $PF' : V \to L(Z, W)$ is Lipschitz-continuous on V and directionally differentiable at $\bar{z}$ with respect to every direction in Z.

Definition 3.1 The mapping F is referred to as *2-regular at the point $\bar{z}$ with respect to a direction $\zeta \in Z$*, if

$$\operatorname{im}(F'(\bar{z}) + (PF')'(\bar{z}; \zeta)) = W.$$

Furthermore, F is said to be *2-regular at the point $\bar{z}$*, if it is 2-regular at this point with respect to every direction $\zeta \in T_2 \setminus \{0\}$, where

$$T_2 = \{\zeta \in \ker F'(\bar{z}) \mid (PF')'(\bar{z}; \zeta)\zeta = 0\}.$$

Among other things, we have the following.

Theorem 3.1 *[6, Theorem 2.1] Under the hypotheses (H1)–(H4), the following statements hold:*

(a) $C_{F^{-1}(F(\bar z))}(\bar z) \subset T_2$, *where* $F^{-1}(F(\bar z)) = \{z \in Z \mid F(z) = F(\bar z)\}$.

(b) If $\zeta \in T_2$, *and the mapping* F *is 2-regular at* $\bar z$ *with respect to* ζ, *then* $\zeta \in T_{F^{-1}(F(\bar z))}(\bar z)$.

In particular, if the mapping F *is 2-regular at* $\bar z$, *then* $C_{F^{-1}(F(\bar z))}(\bar z) = T_{F^{-1}(F(\bar z))}(\bar z) = T_2$.

Note that since the tangent cone is always a closed set, in the last assertion of Theorem 3.1 it is sufficient to assume that F is 2-regular at $\bar z$ with respect to every element in some dense subset of T_2.

As is well known (e.g., [7]) and easy to see, the set $\mathcal{KKT}$ can be equivalently represented as

$$\mathcal{KKT} = \{z \in \mathbf{R}^s \times \mathbf{R}^n \times \mathbf{R}^m \mid F(z) = 0\},$$

where $z = (\sigma,\, x,\, \mu)$,

$$F : \mathbf{R}^s \times \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^n \times \mathbf{R}^m, \quad F(z) = \begin{pmatrix} \Psi(z) \\ S(z) \end{pmatrix},$$

$$S : \mathbf{R}^s \times \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^m, \quad S_i(z) = \mu_i G_i(\sigma,\, x) - (\min\{0,\, G_i(\sigma,\, x) + \mu_i\})^2/2,\ i = 1,\, \ldots,\, m.$$

By direct computation, for $\bar z = (\bar\sigma,\, \bar x,\, \bar\mu)$ we have that

$$S_i'(\bar z) = \begin{cases} (\bar\mu_i \frac{\partial G_i}{\partial \sigma}(\bar\sigma,\, \bar x),\, \bar\mu_i \frac{\partial G_i}{\partial x}(\bar\sigma,\, \bar x),\, 0), & i \in I_+, \\ (0,\, 0,\, G_i(\bar\sigma,\, \bar x)e^i), & i \in N, \\ 0, & i \in I_0, \end{cases}$$

where e^i is the i-th vector of the canonic basis in $\mathbf{R}^m$. Therefore (possibly after the rearrangement of the indices), we have that

$$F'(\bar z) = \begin{pmatrix} \Lambda \\ 0 \end{pmatrix},$$

where the matrix Λ of dimension $(n + |I_+| + |N|) \times (s + n + m)$ is given by

$$\Lambda = \begin{pmatrix} \frac{\partial \Psi}{\partial \sigma}(\bar\sigma,\, \bar x,\, \bar\mu) & \frac{\partial \Psi}{\partial x}(\bar\sigma,\, \bar x,\, \bar\mu) & -\left(\frac{\partial G}{\partial x}(\bar\sigma,\, \bar x)\right)^{\mathrm{T}} \\ A & B & 0 \\ 0 & 0 & C \end{pmatrix},$$

and A, B and C are matrices of dimensions $|I_+| \times s$, $|I_+| \times n$ and $|N| \times m$, respectively, with the following rows:

$$A_i = \bar\mu_i \frac{\partial G_i}{\partial \sigma}(\bar\sigma,\, \bar x),\ B_i = \bar\mu_i \frac{\partial G_i}{\partial x}(\bar\sigma,\, \bar x),\ i \in I_+, \quad C_i = G_i(\bar\sigma,\, \bar x)e^i,\ i \in N.$$

Taking into account that $\bar\mu_i > 0\ \forall\, i \in I_+$, and $G_i(\bar\sigma,\, \bar x) > 0\ \forall\, i \in N$, it is easy to see that

$$\ker F'(\bar z) = \ker \Lambda = \{\zeta = (d,\, \xi,\, \nu) \in \mathbf{R}^s \times \mathbf{R}^n \times \mathbf{R}^m \mid (d,\, \xi,\, \nu_I) \in \ker \Gamma,\ \nu_N = 0\}, \tag{3.1}$$

where

$$\Gamma = \left(\begin{array}{ccc} \frac{\partial \Psi}{\partial \sigma}(\bar{\sigma},\, \bar{x},\, \bar{\mu}) & \frac{\partial \Psi}{\partial x}(\bar{\sigma},\, \bar{x},\, \bar{\mu}) & -\left(\frac{\partial G_I}{\partial x}(\bar{\sigma},\, \bar{x})\right)^{\mathrm{T}} \\ \frac{\partial G_{I_+}}{\partial \sigma}(\bar{\sigma},\, \bar{x}) & \frac{\partial G_{I_+}}{\partial x}(\bar{\sigma},\, \bar{x}) & 0 \end{array}\right). \tag{3.2}$$

In particular, from (2.1) it follows that $\mathcal{L} \subset \ker F'(\bar{z})$.

We further obtain that

$$W_1 = \operatorname{im} F'(\bar{z}) = \operatorname{im} \Lambda \times \{0\}.$$

Hence, we can take W_2 as follows:

$$W_2 = (\operatorname{im} F'(\bar{z}))^{\perp} = (\operatorname{im} \Lambda)^{\perp} \times \mathbf{R}^{|I_0|}.$$

With this choice, P is the orthogonal projector in $\mathbf{R}^n \times \mathbf{R}^m$ onto W_2:

$$Pw = (\Pi(u,\, v_{I_+ \cup N}),\, v_{I_0}), \quad w = (u,\, v) \in \mathbf{R}^n \times \mathbf{R}^m,$$

where Π is the orthogonal projector onto $(\operatorname{im} \Lambda)^{\perp}$.

Observe that for $i \in I_+ \cup N$, we have that $S_i(z) = \mu_i G_i(\sigma,\, x)$ for all $z = (\sigma,\, x,\, \mu)$ close to $\bar{z}$. In particular, $S_{I_+ \cup N}$ is sufficiently smooth (say, twice differentiable at $\bar{z}$). Hence, for any $\zeta = (d,\, \xi, \nu) \in \mathbf{R}^s \times \mathbf{R}^n \times \mathbf{R}^m$, we have that

$$(PF')'(\bar{z};\, \zeta) = \left(\begin{array}{c} \Pi \left(\begin{array}{c} \Psi''(\bar{z})\zeta \\ S''_{I_+ \cup N}(\bar{z})\zeta \end{array}\right) \\ (S'_{I_0})'(\bar{z};\, \zeta) \end{array}\right). \tag{3.3}$$

For $i \in I_0$, we obtain

$$\begin{aligned} (S'_i)'(\bar{z};\, \zeta) \;&=\; (\nu_i - \min\{0,\, \langle G'_i(\bar{\sigma},\, \bar{x}),\, (d,\, \xi)\rangle + \nu_i\}) \left(\frac{\partial G_i}{\partial \sigma}(\bar{\sigma},\, \bar{x}),\, \frac{\partial G_i}{\partial x}(\bar{\sigma},\, \bar{x}),\, 0\right) \\ &\quad + (\langle G'_i(\bar{\sigma},\, \bar{x}),\, (d, \xi)\rangle - \min\{0,\, \langle G'_i(\bar{\sigma},\, \bar{x}),\, (d,\, \xi)\rangle + \nu_i\})(0,\, 0,\, e^i), \end{aligned} \tag{3.4}$$

and

$$(S'_i)'(\bar{z};\, \zeta)\zeta = 2\nu_i \langle G'_i(\bar{\sigma},\, \bar{x}),\, (d,\, \xi)\rangle - (\min\{0,\, \langle G'_i(\bar{\sigma},\, \bar{x}),\, (d,\, \xi)\rangle + \nu_i\})^2.$$

In particular,

$$(S'_i)'(\bar{z};\, \zeta)\zeta = 0 \quad \Longleftrightarrow \quad \nu_i \geq 0,\ \langle G'_i(\bar{\sigma},\, \bar{x}),\, (d,\, \xi)\rangle \geq 0,\ \nu_i \langle G'_i(\bar{\sigma},\, \bar{x}),\, (d,\, \xi)\rangle = 0.$$

Hence, by (2.1), (3.1), (3.2) and (3.3), we conclude that

$$T_2 = \mathcal{L} \cap \mathcal{Q},$$

where

$$\mathcal{Q} = \left\{ \zeta \in \mathbf{R}^s \times \mathbf{R}^n \times \mathbf{R}^m \;\middle|\; \Pi \left(\begin{array}{c} \Psi''(\bar{z})[\zeta,\, \zeta] \\ S''_{I_+ \cup N}(\bar{z})[\zeta,\, \zeta] \end{array}\right) = 0 \right\}.$$

In particular, taking into account assertion (a) of Theorem 3.1, we obtain the following estimate for the contingent derivative of the KKT multifunction, sharper than (2.2) (see also Example 3.1).

Proposition 3.1 *Under the hypotheses (H1)–(H4), for* $(\bar{\sigma}, \bar{x}, \bar{\mu}) \in \mathcal{KKT}$*, the following inclusion holds:*

$$C_{\mathcal{KKT}}(\bar{\sigma}, \bar{x}, \bar{\mu}) \subset \mathcal{L} \cap \mathcal{Q}. \tag{3.5}$$

In the special case when $\operatorname{rank} \Lambda = n + |I_+| + |N|$, or equivalently,

$$\operatorname{rank} \Gamma = n + |I_+|, \tag{3.6}$$

it holds that $\Pi = 0$, so that $\mathcal{Q} = \mathbf{R}^s \times \mathbf{R}^n \times \mathbf{R}^m$, and thus $T_2 = \mathcal{L}$. In this case, inclusions (2.2) and (3.5) are the same.

We next show that (3.6), and thus $T_2 = \mathcal{L}$, hold in the special case of *canonical perturbations*. The KKT system is said to be canonically perturbed if the parameterization includes arbitrary right-hand side perturbations of Φ and G:

$$\begin{gathered} \Phi(\sigma, x) = \Phi_0(\sigma_0, x) + \sigma_1, \quad G(\sigma, x) = G_0(\sigma_0, x) + \sigma_2, \\ \sigma = (\sigma_0, \sigma_1, \sigma_2) \in \mathbf{R}^{s_0} \times \mathbf{R}^n \times \mathbf{R}^m, \; x \in \mathbf{R}^n, \end{gathered} \tag{3.7}$$

with $\bar{\sigma} = (\bar{\sigma}_0, 0, 0)$, $\bar{\sigma}_0 \in \mathbf{R}^{s_0}$. Here, $\Phi_0 : \mathbf{R}^{s_0} \times \mathbf{R}^n \to \mathbf{R}^n$, $G_0 : \mathbf{R}^{s_0} \times \mathbf{R}^n \to \mathbf{R}^m$ are sufficiently smooth mappings. It is easy to see that in this case

$$\operatorname{rank} \begin{pmatrix} \frac{\partial \Psi}{\partial \sigma}(\bar{\sigma}, \bar{x}, \bar{\mu}) \\ \frac{\partial G}{\partial \sigma}(\bar{\sigma}, \bar{x}) \end{pmatrix} = n + m, \tag{3.8}$$

and hence, (3.6) holds (see (3.2)). Thus, for the case of canonical perturbations, $T_2 = \mathcal{L}$, and (3.5) coincides with (2.2).

But beyond the case of canonical perturbations, T_2 can be a strictly sharper estimate of the contingent derivative than $\mathcal{L}$, as illustrated by the following example, which is obtained by introducing the (non-canonical) perturbation in [5, Example 2].

Example 3.1 Let $s = 1$, $n = m = 2$, $\Phi(\sigma, x) = (\sigma + x_1, x_2^2)$, $G(\sigma, x) = (x_1 - x_2^2/2, x_1 + x_2^2/2)$, $\bar{\sigma} = 0$, $\bar{x} = \bar{\mu} = 0$.

By direct computations it can be shown that $\mathcal{L}$ is the solution set of the following system of equations in variables $(d, \xi_1, \xi_2, \nu_1, \nu_2)$:

$$\begin{aligned} d + \xi_1 - \nu_1 - \nu_2 &= 0, \\ \min\{\xi_1, \nu_1\} &= 0, \\ \min\{\xi_1, \nu_2\} &= 0. \end{aligned}$$

In particular, ξ_2 is arbitrary.

At the same time, elements of T_2 should satisfy the additional equation

$$\xi_2(\xi_2 - \nu_1 + \nu_2) = 0,$$

so that ξ_2 is no longer arbitrary. This shows that the estimate (3.5) is sharper than the estimate (2.2).

We next consider the conditions which ensure that a given $\zeta = (d,\, \xi,\, \nu) \in \mathcal{L}$ belongs to $T_{\mathcal{KKT}}(\bar{\sigma},\, \bar{x},\, \bar{\mu})$. Recall that $I_1 = I_0^1 \cup I_0^+$, $I_2 = I_0^2 \cup I_0^N$, where $(I_0^1,\, I_0^2)$ is a partition of I_0^0, and I_0^0, I_0^+, I_0^N are defined in (2.3).

Assume that MFCQ holds at $(\bar{\sigma},\, \bar{x},\, \bar{\mu})$ for the constraints defining the branch $\mathcal{KKT}_{(I_1,\, I_2)}$:

$$\begin{gathered} \operatorname{rank} \Lambda_{(I_1,\, I_2)} = n + |I_1| + |I_+|, \\ \exists\, (\bar{d},\, \bar{\xi},\, \bar{\nu}_{I_1 \cup I_+}) \in \ker \Lambda_{(I_1,\, I_2)} \text{ such that } \bar{\nu}_{I_1} > 0,\ G'_{I_2}(\bar{\sigma},\, \bar{x})(\bar{d},\, \bar{\xi}) > 0, \end{gathered} \tag{3.9}$$

where $\Lambda_{(I_1,\, I_2)}$ is defined in (2.6). Take $\bar{\zeta} = (\bar{d},\, \bar{\xi},\, \bar{\nu})$, where $\bar{\nu}_{I_2 \cup N} = 0$. Note that the first condition in (3.9) implies (3.6). Hence, by (3.3),

$$F'(\bar{z}) + (PF')'(\bar{z};\, \bar{\zeta}) = \begin{pmatrix} \Lambda \\ (S'_{I_0})'(\bar{z};\, \bar{\zeta}) \end{pmatrix}. \tag{3.10}$$

By the second condition in (3.9), it can be seen that $\bar{\zeta} \in \mathcal{L}_{(I_1,\, I_2)}$, and that the strict complementarity condition holds in (2.1) at $\bar{\zeta}$. In particular,

$$\bar{\nu}_i - \min\{0,\, \langle G'_i(\bar{\sigma},\, \bar{x}),\, (\bar{d},\, \bar{\xi})\rangle + \bar{\nu}_i\} = \begin{cases} \bar{\nu}_i > 0, & i \in I_1, \\ 0, & i \in I_2, \end{cases}$$

$$\langle G'_i(\bar{\sigma},\, \bar{x}),\, (\bar{d},\, \bar{\xi})\rangle - \min\{0,\, \langle G'_i(\bar{\sigma},\, \bar{x}),\, (\bar{d},\, \bar{\xi})\rangle + \bar{\nu}_i\} = \begin{cases} 0, & i \in I_1, \\ \langle G'_i(\bar{\sigma},\, \bar{x}),\, (\bar{d},\, \bar{\xi})\rangle > 0, & i \in I_2. \end{cases}$$

By (2.6), (3.4) and (3.10), it is now evident that (3.9) implies 2-regularity of F at $\bar{z}$ with respect to $\bar{\zeta}$. Moreover, it can be seen that (3.9) actually implies that F is 2-regular at $\bar{z}$ with respect to every direction in some dense subset of $\mathcal{L}_{(I_1,\, I_2)}$. In particular, piecewise MFCQ (that is, MFCQ (3.9) for every partition $(I_1,\, I_2)$ of I_0) implies that F is 2-regular at $\bar{z}$ with respect to every direction in some dense subset of $\mathcal{L}$. From assertion (b) of Theorem 3.1 we now obtain

Proposition 3.2 *For $(\bar{\sigma},\, \bar{x},\, \bar{\mu}) \in \mathcal{KKT}$ and a given $(d,\, \xi,\, \nu) \in \mathcal{L}$, assume that MFCQ (3.9) holds at $(\bar{\sigma},\, \bar{x},\, \bar{\mu})$ for the constraints defining the branch $\mathcal{KKT}_{(I_0^1 \cup I_0^+,\, I_0^2 \cup I_0^N)}$ for some partition $(I_0^1,\, I_0^2)$ of I_0^0.*

Then $(d,\, \xi,\, \nu) \in T_{\mathcal{KKT}}(\bar{\sigma},\, \bar{x},\, \bar{\mu})$.

In particular, piecewise MFCQ for $\mathcal{KKT}$ at $(\bar{\sigma},\, \bar{x},\, \bar{\mu})$ implies

$$T_{\mathcal{KKT}}(\bar{\sigma},\, \bar{x},\, \bar{\mu}) = C_{\mathcal{KKT}}(\bar{\sigma},\, \bar{x},\, \bar{\mu}) = \mathcal{L}. \tag{3.11}$$

Of course, the result of Proposition 3.2 can be obtained by the standard argument combined with piecewise analysis. We include this proposition merely as one of the illustrations for the use of the 2-regularity concept.

If we *assume* (3.6), then it can be seen that 2-regularity of F at $\bar{z}$ with respect to $\bar{\zeta}$ is actually equivalent to (3.9). In particular, strict complementarity in (2.1) at $\bar{\zeta}$ is a necessary condition for 2-regularity of F at $\bar{z}$ with respect to $\bar{\zeta}$.

Consider again the case of canonical perturbations (3.7) and assume that MFCQ holds. In this case, from (3.8) it easily follows that piecewise MFCQ holds for $\mathcal{KKT}$ at $(\bar{\sigma},\, \bar{x},\, \bar{\mu})$, which implies (3.11). In particular, under these assumptions, the KKT multifunction is protodifferentiable at $\bar{\sigma}$ for $(\bar{x},\, \bar{\mu})$, and its contingent derivative at $\bar{\sigma}$ for $(\bar{x},\, \bar{\mu})$ is a multifunction from $\mathbf{R}^s$ to $2^{\mathbf{R}^n \times \mathbf{R}^m}$ whose graph coincides with $\mathcal{L}$. This fact was established in [16, Proposition 2.5.1].

We next discuss the assumptions of Corollary 2.1 which is our existence and uniqueness result. Recall that in this setting, we consider the existence of solutions for a given perturbation of the form $\bar{\sigma} + td + \rho(t)$, where the direction $d \in \mathbf{R}^s$ and the mapping $\rho : \mathbf{R}_+ \to \mathbf{R}^s$, $\rho(t) = o(t)$, are fixed. Therefore, the relevant object to study is

$$\tilde{F} : \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^n \times \mathbf{R}^m, \quad \tilde{F}(p) = F(z), \quad p = (t,\, x,\, \mu),\; z = (\bar{\sigma} + td + \rho(t),\, x,\, \mu).$$

By the analysis similar to the above (possibly after the rearrangement of the indices), for $\bar{p} = (0,\, \bar{x},\, \bar{\mu})$ it can be seen that

$$\tilde{F}'(\bar{p}) = \begin{pmatrix} \tilde{\Lambda} \\ 0 \end{pmatrix},$$

where the matrix $\tilde{\Lambda}$ of dimension $(n + |I_+| + |N|) \times (1 + n + m)$ is given by

$$\tilde{\Lambda} = \begin{pmatrix} \frac{\partial \Psi}{\partial \sigma}(\bar{\sigma},\, \bar{x},\, \bar{\mu})d & \frac{\partial \Psi}{\partial x}(\bar{\sigma},\, \bar{x},\, \bar{\mu}) & -\left(\frac{\partial G}{\partial x}(\bar{\sigma},\, \bar{x})\right)^{\mathrm{T}} \\ Ad & B & 0 \\ 0 & 0 & C \end{pmatrix}.$$

In particular,

$$\operatorname{im} \tilde{F}'(\bar{z}) = \operatorname{im} \tilde{\Lambda} \times \{0\}. \tag{3.12}$$

Observe that if there exists $(\xi,\, \nu) \in \mathbf{R}^n \times \mathbf{R}^m$ satisfying $(d,\, \xi,\, \nu) \in \mathcal{L}$, then (2.1) implies that the first column in $\tilde{\Lambda}$ can be obtained as a linear combination of the other columns. Hence,

$$\operatorname{im} \tilde{\Lambda} = \operatorname{im} \begin{pmatrix} \frac{\partial \Psi}{\partial x}(\bar{\sigma},\, \bar{x},\, \bar{\mu}) & -\left(\frac{\partial G}{\partial x}(\bar{\sigma},\, \bar{x})\right)^{\mathrm{T}} \\ B & 0 \\ 0 & C \end{pmatrix}. \tag{3.13}$$

Evidently, (2.7) implies that the matrix in the right-hand side has full row rank. Therefore, $(\operatorname{im} \tilde{\Lambda})^{\perp} = \{0\}$, and by (3.12), we have that

$$(\operatorname{im} \tilde{F}'(\bar{z}))^{\perp} = \{0\} \times \mathbf{R}^{|I_0|}.$$

It is now easy to see that under the assumptions of Corollary 2.1, 2-regularity of $\tilde{F}$ at $\bar{p}$ with respect to $q = (1,\, \xi,\, \nu)$ is equivalent to saying that the matrix

$$\begin{pmatrix} \frac{\partial \Psi}{\partial \sigma}(\bar{\sigma},\, \bar{x},\, \bar{\mu})d & \frac{\partial \Psi}{\partial x}(\bar{\sigma},\, \bar{x},\, \bar{\mu}) & -\left(\frac{\partial G_{I_0^+ \cup I_+}}{\partial x}(\bar{\sigma},\, \bar{x})\right)^{\mathrm{T}} \\ \frac{\partial G_{I_0^+ \cup I_+}}{\partial \sigma}(\bar{\sigma},\, \bar{x})d & \frac{\partial G_{I_0^+ \cup I_+}}{\partial x}(\bar{\sigma},\, \bar{x}) & 0 \end{pmatrix}$$

has full row rank. Taking again into account that the first column above can be represented as a linear combination of the other columns (by (2.1) and (2.3)), the latter condition is equivalent to (2.7).

Moreover, if we *assume* that the matrix

$$\begin{pmatrix} \frac{\partial \Psi}{\partial x}(\bar{\sigma},\, \bar{x},\, \bar{\mu}) & -\left(\frac{\partial G_I}{\partial x}(\bar{\sigma},\, \bar{x})\right)^{\mathrm{T}} \\ \frac{\partial G_{I_+}}{\partial x}(\bar{\sigma},\, \bar{x}) & 0 \end{pmatrix}$$

has full row rank, which is equivalent to the assumption that the matrix in the right-hand side of (3.13) has full row rank, then 2-regularity of $\tilde{F}$ at $\bar{p}$ with respect to q is equivalent to the assumptions of Corollary 2.1, i.e., $I_0^0 = \emptyset$ and (2.7) (recall that according to the discussion above, the strict complementarity condition $I_0^0 = \emptyset$ is necessary for 2-regularity of $\tilde{F}$ at $\bar{p}$ with respect to q).

4 Some Examples

We start this section with the following result exhibiting some further properties of the set $\mathcal{L}$. Related pairs of dual linear programs are known to be very useful in sensitivity analysis. But Lemma 4.1 appears to be new. Conclusions which can be deduced by using Lemma 4.1 will be given after the proof and illustrated by the examples below.

Lemma 4.1 *If for a given $(d,\, \xi) \in \mathbf{R}^s \times \mathbf{R}^n$, there exists $\nu \in \mathbf{R}^m$ such that $(d,\, \xi,\, \nu) \in \mathcal{L} = \mathcal{L}(\bar{\sigma},\, \bar{x},\, \bar{\mu})$, then ξ is a solution of the LP problem*

$$\begin{array}{ll} \text{minimize}_x & \langle \Phi(\bar{\sigma},\, \bar{x}),\, x\rangle \\ \text{subject to} & \langle G_i'(\bar{\sigma},\, \bar{x}),\, (d,\, x)\rangle \geq 0,\ i \in I, \end{array} \tag{4.1}$$

while $\bar{\mu}$ is a solution of the dual LP problem

$$\begin{array}{ll} \text{maximize}_\mu & -\langle \mu,\, \frac{\partial G}{\partial \sigma}(\bar{\sigma},\, \bar{x})d\rangle \\ \text{subject to} & \left(\frac{\partial G}{\partial x}(\bar{\sigma},\, \bar{x})\right)^{\mathrm{T}} \mu = \Phi(\bar{\sigma},\, \bar{x}), \\ & \mu_I \geq 0,\ \mu_N = 0. \end{array} \tag{4.2}$$

Proof. By the definition of $\mathcal{L}$ (see (2.1)), ξ is feasible in (4.1) since

$$\langle G_i'(\bar{\sigma},\, \bar{x}),\, (d,\, \xi)\rangle \geq 0 \quad \forall\, i \in I_0, \quad \langle G_i'(\bar{\sigma},\, \bar{x}),\, (d,\, \xi)\rangle = 0 \quad \forall\, i \in I_+.$$

At the same time, $\bar{\mu}$ is feasible in (4.2) since the constraints of (4.2) can be stated in the form $(\bar{\sigma},\, \bar{x},\, \mu) \in \mathcal{KKT}$. Furthermore, the duality relation holds:

$$\begin{aligned} \langle \Phi(\bar{\sigma},\, \bar{x}),\, \xi\rangle &= \left\langle \left(\frac{\partial G}{\partial x}(\bar{\sigma},\, \bar{x})\right)^{\mathrm{T}} \bar{\mu},\, \xi \right\rangle \\ &= \left\langle \bar{\mu},\, \frac{\partial G}{\partial x}(\bar{\sigma},\, \bar{x})\xi \right\rangle \\ &= \sum_{i \in I_+} \bar{\mu}_i \left\langle \frac{\partial G_i}{\partial x}(\bar{\sigma},\, \bar{x}),\, \xi \right\rangle \end{aligned}$$

$$
\begin{aligned}
&= -\sum_{i\in I_+} \bar\mu_i \left\langle \frac{\partial G_i}{\partial\sigma}(\bar\sigma,\, \bar x),\, d\right\rangle \\
&= -\left\langle \bar\mu,\, \frac{\partial G}{\partial\sigma}(\bar\sigma,\, \bar x)d\right\rangle,
\end{aligned}
$$

where the inclusion $(\bar\sigma,\, \bar x,\, \bar\mu) \in \mathcal{KKT}$ and the definition of $\mathcal{L}$ (see (2.1)) were taken into account. ■

Lemma 4.1 leads to the following conclusions. If $\frac{\partial G}{\partial\sigma}(\bar\sigma,\, \bar x)d \neq 0$ for a given $d \in \mathbf{R}^s$, and there exist $\xi \in \mathbf{R}^n$ and $\nu \in \mathbf{R}^m$ such that $(d,\, \xi,\, \nu) \in \mathcal{L}$, then the objective function of (4.2) is non-constant, and it is "quite likely" that $\bar\mu$ is the unique solution of (4.2). In that case, $\bar\mu$ is a vertex of the polyhedral set of multipliers by necessity. Suppose that SMFCQ is not satisfied, i.e., the set of multipliers (of the nonperturbed KKT system) associated with $\bar x$ is not a singleton. Then we can conclude that $I_0 \neq \emptyset$, i.e., the *strict complementarity* condition is violated at the solution $(\bar x,\, \bar\mu)$ of the nonperturbed KKT system. Indeed, if it were the case that $\bar\mu$ is a vertex and $I_0 = \emptyset$, then $\bar\mu$ would have been the unique solution of the equality-part of constraints in (4.2), which contradicts nonuniqueness of the multiplier.

The examples presented in this section highlight the situation where SMFCQ is violated, i.e., the multiplier associated with $\bar x$ at the base value $\bar\sigma$ of the parameter is not unique. In particular, we demonstrate that the branches of solutions of the perturbed KKT system may depend drastically on the specific choice of the multiplier (which should be already clear from Lemma 4.1).

Example 4.1 Let $s = 1$, $n = 2$, $m = 3$, $f(\sigma,\, x) = x_1 + x_2^2/2$, $G(\sigma,\, x) = (x_1+\sigma,\, x_1+x_2,\, x_1 - x_2)$. Consider the parametric optimization problem (1.3) with the feasible set defined in (1.4).

When $\sigma = \bar\sigma = 0$, this problem has the unique solution $\bar x = 0$, and

$$
\frac{\partial G}{\partial\sigma}(\bar\sigma,\, \bar x) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{\partial G}{\partial x}(\bar\sigma,\, \bar x) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}.
$$

The KKT system for this problem takes the form (1.1) with Φ defined in (1.2): $\Phi(\sigma,\, x) = \frac{\partial f}{\partial x}(\sigma,\, x) = (1,\, x_2)$, and the set of multipliers associated with $\bar x$ is $\{\mu \in \mathbf{R}^3 \mid \mu = (1 - 2\theta,\, \theta,\, \theta),\, \theta \in [0,\, 1/2]\}$.

For $d < 0$, problem (4.2) has the unique solution $\bar\mu = (1,\, 0,\, 0)$. With this choice of the multiplier, $I_0 = \{2,\, 3\}$, $I_+ = \{1\}$, $N = \emptyset$,

$$
\frac{\partial\Psi}{\partial\sigma}(\bar\sigma,\, \bar x,\, \bar\mu) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \frac{\partial\Psi}{\partial x}(\bar\sigma,\, \bar x,\, \bar\mu) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},
$$

and $\mathcal{L}$ is the solution set of the following system of equations:

$$
\begin{aligned}
\nu_1 + \nu_2 + \nu_3 &= 0, \\
\xi_2 - \nu_2 + \nu_3 &= 0, \\
\min\{\xi_1 + \xi_2,\, \nu_2\} &= 0, \\
\min\{\xi_1 - \xi_2,\, \nu_3\} &= 0, \\
d + \xi_1 &= 0.
\end{aligned}
$$

Consider next the branches of $\mathcal{L}$.

If $I_1 = \emptyset$, $I_2 = I_0$, then the corresponding branch is the ray spanned in $\mathbf{R}^s \times \mathbf{R}^n \times \mathbf{R}^m$ by $(-1, (1, 0), 0)$, and for every non-zero element of this ray, $I_0^0 = \emptyset$ and (2.7) holds.

If $I_1 = I_0$ and $I_2 = \emptyset$, the corresponding branch is the ray spanned by $(0, 0, (-2, 1, 1))$, and for every non-zero element of this ray, $I_0^0 = \emptyset$ and (2.7) holds as well.

As is not difficult to check, the other two branches (corresponding to $I_1 = \{2\}$ and $I_2 = \{3\}$, $I_1 = \{3\}$ and $I_2 = \{2\}$, respectively) are trivial. Hence, for the given choice of $\bar{\mu}$, the tangent cone $T_{\mathcal{KKT}}(\bar{\sigma}, \bar{x}, \bar{\mu})$ consists of two rays (the first two branches above); this follows from (2.2) and Proposition 3.2.

Moreover, according to Corollary 2.1, for every $t > 0$ small enough, the perturbed KKT system corresponding to the parameter values $\sigma = -t + o(t)$ has the unique solution of the form $((t, 0), 0) + o(t)$. Similarly, for the parameter values $\sigma = o(t)$, there is a unique solution of the form $(0, (-2t, t, t)) + o(t)$.

For $d > 0$, problem (4.2) has the unique solution $\bar{\mu} = (0, 1/2, 1/2)$. With this choice, $I_0 = \{1\}$, $I_+ = \{2, 3\}$, $N = \emptyset$, and it is easy to see that $\mathcal{L}$ is the ray spanned in $\mathbf{R}^s \times \mathbf{R}^n \times \mathbf{R}^m$ by $(1, 0, 0)$, and for every non-zero element of this ray, $I_0^0 = \emptyset$, and (2.7) holds. We conclude that $T_{\mathcal{KKT}}(\bar{\sigma}, \bar{x}, \bar{\mu})$ coincides with this ray, and for every $t > 0$ small enough, the perturbed KKT system corresponding to the parameter values $\sigma = t + o(t)$ has the unique solution of the form $o(t)$.

Note that according to Lemma 4.1, any other choice of the multiplier $\bar{\mu}$ will result in the cone $\mathcal{L} \subset \{0\} \times \mathbf{R}^n \times \mathbf{R}^m$.

Example 4.2 *([1, Example 4.99])* Let $s = n = m = 2$, $f(\sigma, x) = ((x_1 - 1)^2 + x_2^2)/2$, $G(\sigma, x) = (-x_1, -x_1 - \sigma_1 x_2 - \sigma_2)$.

When $\sigma = \bar{\sigma} = 0$, optimization problem (1.3) with the feasible set defined in (1.4) has the unique solution $\bar{x} = 0$, and

$$\frac{\partial G}{\partial \sigma}(\bar{\sigma}, \bar{x}) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \frac{\partial G}{\partial x}(\bar{\sigma}, \bar{x}) = \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix}.$$

The KKT system for this problem takes the form (1.1) with Φ defined in (1.2): $\Phi(\sigma, x) = \frac{\partial f}{\partial x}(\sigma, x) = (x_1 - 1, x_2)$, and the set of multipliers associated with $\bar{x}$ is $\{\mu \in \mathbf{R}^2 \mid \mu = (1 - \theta, \theta), \theta \in [0, 1]\}$.

For $d = (d_1, d_2)$ with $d_2 < 0$, problem (4.2) has the unique solution $\bar{\mu} = (1, 0)$. With this choice of the multiplier, $I_0 = \{2\}$, $I_+ = \{1\}$, $N = \emptyset$,

$$\frac{\partial \Psi}{\partial \sigma}(\bar{\sigma}, \bar{x}, \bar{\mu}) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \frac{\partial \Psi}{\partial x}(\bar{\sigma}, \bar{x}, \bar{\mu}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

(note that $\frac{\partial \Psi}{\partial x}(\bar{\sigma}, \bar{x}, \bar{\mu})$ does not actually depend on the choice of $\bar{\mu}$), and $\mathcal{L}$ is the solution set of the following system of equations:

$$\begin{aligned} \xi_1 + \nu_1 + \nu_2 &= 0, \\ \xi_2 &= 0, \\ \min\{-d_2 - \xi_1, \nu_2\} &= 0, \\ -\xi_1 &= 0. \end{aligned}$$

Obviously, $\mathcal{L}$ consists of the following two branches in $\mathbf{R}^s \times \mathbf{R}^n \times \mathbf{R}^m$:

$$\{((\alpha,\,-\beta),\,0,\,0) \mid \alpha \in \mathbf{R},\,\beta \in \mathbf{R}_+\}$$

and

$$\{((\alpha,\,0),\,0,\,(-\gamma,\,\gamma)) \mid \alpha \in \mathbf{R},\,\gamma \in \mathbf{R}_+\};$$

their intersection is the straight line spanned by $((1,\,0),\,0,\,0)$. For every element of the two branches with $\beta > 0$ and $\gamma > 0$, respectively, we have that $I_0^0 = \emptyset$ and (2.7) holds. It follows that $T_{\mathcal{KKT}}(\bar{\sigma},\,\bar{x},\,\bar{\mu})$ consists of these two branches.

For $d = (d_1,\,d_2)$ with $d_2 > 0$, problem (4.2) has the unique solution $\bar{\mu} = (0,\,1)$. With this choice, $I_0 = \{1\}$, $I_+ = \{2\}$, $N = \emptyset$,

$$\frac{\partial \Psi}{\partial \sigma}(\bar{\sigma},\,\bar{x},\,\bar{\mu}) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

and $\mathcal{L}$ is the solution set of the following system of equations:

$$\begin{aligned} \xi_1 + \nu_1 + \nu_2 &= 0, \\ d_1 + \xi_2 &= 0, \\ \min\{-\xi_1,\,\nu_1\} &= 0, \\ -d_2 - \xi_1 &= 0. \end{aligned}$$

The two branches of $\mathcal{L}$ are

$$\{((\alpha,\,\beta),\,(-\beta,\,-\alpha),\,(0,\,\beta)) \mid \alpha \in \mathbf{R},\,\beta \in \mathbf{R}_+\}$$

and

$$\{((\alpha,\,0),\,(0,\,-\alpha),\,(\gamma,\,-\gamma)) \mid \alpha \in \mathbf{R},\,\gamma \in \mathbf{R}_+\};$$

their intersection is the straight line spanned by $((1,\,0),\,(0,\,-1),\,0))$. For every element of the two branches with $\beta > 0$ and $\gamma > 0$, respectively, we again have that $I_0^0 = \emptyset$ and (2.7) holds. Hence, $T_{\mathcal{KKT}}(\bar{\sigma},\,\bar{x},\,\bar{\mu})$ consists of these two branches.

Finally, let $d = (d_1,\,0)$. In this case, every $\bar{\mu} = (1-\theta,\,\theta)$, $\theta \in [0,\,1]$, provides a solution to (4.2). Take $\theta \in (0,\,1)$ (the values $\theta = 0$ and $\theta = 1$ were already considered above). With this choice, $I_0 = N = \emptyset$, $I_+ = \{1,\,2\}$,

$$\frac{\partial \Psi}{\partial \sigma}(\bar{\sigma},\,\bar{x},\,\bar{\mu}) = \begin{pmatrix} 0 & 0 \\ \theta & 0 \end{pmatrix},$$

and $\mathcal{L}$ is the solution set of the following system of linear equations:

$$\begin{aligned} \xi_1 + \nu_1 + \nu_2 &= 0, \\ \theta d_1 + \xi_2 &= 0, \\ -\xi_1 &= 0, \\ -d_2 - \xi 1 &= 0. \end{aligned}$$

Hence, $\mathcal{L}$ is the subspace

$$\{((\alpha,\, 0),\, (0,\, -\theta\nu\alpha),\, (\gamma,\, -\gamma)) \mid \alpha,\, \gamma \in \mathbf{R}\}.$$

Summarizing the above analysis, we see that for every direction $d = (d_1,\, d_2)$ with $d_2 \neq 0$, the corresponding "primal" part of the tangent vector to $\mathcal{KKT}$ at $(\bar{\sigma},\, \bar{x},\, \bar{\mu})$ is uniquely defined: $\xi = 0$ if $d_2 < 0$ and $\xi = (-d_2,\, -d_1)$ if $d_2 > 0$. For $d = (d_1,\, 0)$, situation is more complicated, as for every $\theta \in [0,\, 1]$ there exists the tangent vector with the "primal" part $\xi = (0,\, -\theta d_1)$. This can be explained (in some sense) by the following observation: the behavior of the solution of the original optimization problem under a perturbation of the parameter σ along such directions depends drastically on the higher-order terms of such perturbation. For instance, if for $t \geq 0$ we take $\sigma = (t,\, 0) + \rho(t)$ with some $\rho : \mathbf{R}_+ \to \mathbf{R}^s$ such that $\rho(t) = o(t)$, then the branches of the solutions corresponding to different choices of ρ are not necessarily tangent to each other.

Acknowledgment. We would like to thank A. Shapiro for encouragement and fruitful discussion on the subject. We also thank the anonymous referees whose comments helped us to improve the presentation.

References

[1] J.F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems.* Springer–Verlag, New York, 2000.

[2] B. Gollan. On the marginal function in nonlinear programming. *Mathematics of Operations Research*, 9:208–221, 1984.

[3] F. Facchinei and J.-S. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems.* Springer–Verlag, New York, 2003.

[4] A.F. Izmailov and M.V. Solodov. Error bounds for 2-regular mappings with Lipschitzian derivatives and their applications. *Mathematical Programming*, 89:413–435, 2001.

[5] A.F. Izmailov and M.V. Solodov. Karush-Kuhn-Tucker systems: regularity conditions, error bounds and a class of Newton-type methods. *Mathematical Programming*, 95:631–650, 2003.

[6] A.F. Izmailov and M.V. Solodov. The theory of 2-regularity for mappings with Lipschitzian derivatives and its applications to optimality conditions. *Mathematics of Operations Research*, 27:614–635, 2002.

[7] C. Kanzow. Some equation-based methods for the nonlinear complementarity problem. *Optimization Methods and Software*, 3:327–340, 1994.

[8] D. Klatte. Upper Lipschitz behaviour of solutions to perturbed $C^{1,1}$ programs. *Mathematical Programming*, 88:285–311, 2000.

[9] D. Klatte and B. Kummer. Contingent derivatives of implicit (multi-) functions. *Annals of Operations Research*, 101:313–331, 2001.

[10] D. Klatte and B. Kummer. Generalized Kojima functions and Lipschitz stability of critical points. *Computational Optimization and Applications*, 13:61–85, 1999.

[11] D. Klatte and B. Kummer. *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications*. Kluwer Academic Publishers, Dordrecht, 2002.

[12] M. Kojima. Strongly stable stationary solutions in nonlinear programs. In S.M. Robinson, ed., *Analysis and Computation of Fixed Points*, 93–138. Academic Press, New York, 1980.

[13] B. Kummer. Generalized equations: Solvability and regulaity. *Mathematical Programming Studies*, 21:199–212, 1984.

[14] J. Kyparisis. Parametric variational inequalities with multivalued solution sets. *Mathematics of Operations Research*, 17:341–364, 1992.

[15] J. Kyparisis. Sensitivity analysis for nonlinear programs and variational inequalities. *Mathematics of Operations Research*, 15:286–298, 1990.

[16] A. Levy. Solution sensitivity from general principles. *SIAM Journal on Control and Optimization*, 40:1–38, 2001.

[17] A. Levy and R.T. Rockafellar. Sensitivity analysis of solutions in nonlinear programming problems with nonunique multipliers. In D. Du, L. Qi, and R. Womersley, eds., *Recent Advances in Optimization*, 1–38. World Scientific Publishers, Singapore, 1995.

[18] J. Liu. Strong stability in variational inequalities. *SIAM Journal on Control and Optimization*, 33:725–749, 1995.

[19] Y. Qiu and T.L. Magnanti. Sensitivity analysis for variational inequalities. *Mathematics of Operations Research*, 17:61–76, 1992.

[20] S.M. Robinson. Generalized equations and their solutions. Part I: Basic theory. *Mathematical Programming Study*, 10:128–141, 1979.

[21] S.M. Robinson. Generalized equations and their solutions. Part II: Applications to nonlinear programming. *Mathematical Programming Study*, 19:200–221, 1982.

[22] S.M. Robinson. Local epi-continuity and local optimization. *Mathematical Programming*, 37:208–223, 1987.

[23] S.M. Robinson. Some continuity properties of polyhedral multifunctions. *Mathematical Programming Study*, 14:206–214, 1981.

[24] A. Shapiro. Sensitivity analysis of generalized equations. *Journal of Mathematical Sciences*, 115:2554–2565, 2003.