Encyclopedia of Operations Research and Management Science

CONSTRAINT QUALIFICATIONS

Mikhail V. Solodov

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

Abstract. We discuss assumptions on the constraint functions that allow constructive description of the geometry of a given set around a given point in terms of the constraints derivatives. Consequences for characterizing solutions of variational and optimization problems are discussed. In the optimization case, these include primal and primal-dual first- and second-order necessary optimality conditions.
Key words: constraint qualification, regularity, tangent cone, variational problem, constrained optimization, optimality conditions, Karush-Kuhn-Tucker conditions, Lagrange multipliers, error bound, metric regularity, second-order regularity.
See also: Optimality conditions for constrained optimization; Fritz John and KKT points; Duality and triality.

Roughly speaking, (first-order) constraint qualifications are properties of the analytical description of a set which ensure that the structure of the set around a given feasible point can be constructively captured by (first-order) approximations of the constraint functions defining the set. If it is so, the consequences are far reaching. Constraint qualifications are essential for deriving primal and primal-dual characterizations of solutions of optimization and variational problems, for duality relations, sensitivity and stability analysis, and for convergence and rate of convergence of computational methods for solving optimization and variational problems.

Let D be any set in $\mathbf{R}^n$. An appropriate object to describe the geometry of D around a feasible point $\bar{x} \in D$ is the *tangent (or contingent) cone*

$$\mathcal{T}_D(\bar{x}) = \left\{ d \in \mathbf{R}^n \;\middle|\; \begin{array}{c} \exists \{t_k\} \subset \mathbf{R}_+ \setminus \{0\},\ \{t_k\} \to 0,\ \exists \{d^k\} \subset \mathbf{R}^n,\ \{d^k\} \to d, \\ \text{such that } \bar{x} + t_k d^k \in D \text{ for all } k \end{array} \right\}.$$

The tangent cone includes all the feasible directions, if there are any, as well as "almost-feasible" ones in the stated sense. Suppose further that D is defined by a finite number of equality and inequality constraints, as is common in applications:

$$D = \{x \in \mathbf{R}^n \mid h(x) = 0,\ g(x) \le 0\}, \tag{1}$$

where $h : \mathbf{R}^n \to \mathbf{R}^l$ and $g : \mathbf{R}^n \to \mathbf{R}^m$ are given functions, which we assume to be continuously differentiable in the region of interest. Then constraint qualifications can be thought of as conditions imposed on the functions h, g and/or their derivatives at or around the point $\bar{x}$, that guarantee that the tangent cone $\mathcal{T}_D(\bar{x})$ has an explicit algebraic representation in terms of the derivatives of the constraint functions. This is crucial for developing optimality conditions in optimization, since whenever a point $\bar{x}$ is a local solution of the problem

$$\text{minimize} \quad f(x) \quad \text{subject to} \quad x \in D, \tag{2}$$

where $f : \mathbf{R}^n \to \mathbf{R}$ is differentiable at $\bar{x}$, it holds that

$$\langle f'(\bar{x}), x - \bar{x} \rangle \geq 0 \quad \forall\, x \in \bar{x} + \mathcal{T}_D(\bar{x}). \tag{3}$$

Or equivalently,

$$-f'(\bar{x}) \in (\mathcal{T}_D(\bar{x}))^\circ, \tag{4}$$

where $K^\circ = \{z \in \mathbf{R}^n \mid \langle z, y \rangle \leq 0\, \forall\, y \in K\}$ stands for the dual (negative polar) cone of a cone K in $\mathbf{R}^n$. Constraint qualifications allow explicit characterization of the tangent cone in (3) and of its dual in (4), which makes these abstract optimality conditions tractable.

For the same reasons, constraint qualifications are important for solving the more general variational problems of the form

$$\text{find } \bar{x} \in D \text{ such that } \langle F(\bar{x}), x - \bar{x} \rangle \geq 0 \quad \forall\, x \in \bar{x} + \mathcal{T}_D(\bar{x}), \tag{5}$$

or equivalently,

$$-F(\bar{x}) \in (\mathcal{T}_D(\bar{x}))^\circ, \tag{6}$$

where $F : \mathbf{R}^n \to \mathbf{R}^n$. When for some function $f : \mathbf{R}^n \to \mathbf{R}$ it holds that $F(x) = f'(x)$, $x \in \mathbf{R}^n$, then (5) represents the primal optimality conditions for the optimization problem (2), while (6) leads to the primal-dual optimality conditions. But in the variational setting F need not be integrable in general. When the feasible set D is convex, (5) becomes equivalent to the classical *variational inequality*

$$\text{find } \bar{x} \in D \text{ such that } \langle F(\bar{x}), x - \bar{x} \rangle \geq 0 \quad \forall\, x \in D.$$

Derivation of tractable first-order primal and primal-dual necessary optimality conditions via computing the tangent cone in (3) and its dual in (4) is perhaps the most important role of constraint qualifications. The term "constraint qualification" (CQ) was coined in [1]. Alternatively, the term *regularity* is also sometimes used in the literature to refer to (some of the) constraint qualifications. Constraint qualifications also appear in second-order necessary optimality conditions. And, as already mentioned, they play an important role in deriving duality relations, sensitivity/stability analysis, error bound estimates, and convergence and rate of convergence of computational methods.

1 Tangent cone and first-order primal-dual optimality conditions

Consider a set D defined by a finite number of equality and inequality constraints, as in (1). Let $\bar{x} \in D$ be any feasible point. If $g_i(\bar{x}) < 0$ for some $i \in \{1, \dots, m\}$, by continuity $g_i(x) < 0$ for all $x \in \mathbf{R}^n$ close to $\bar{x}$, and it is clear that such constraints (*inactive* at $\bar{x}$) do not influence the geometry of the set D around the point $\bar{x}$. Let

$$A(\bar{x}) = \{i = 1, \dots, m \mid g_i(\bar{x}) = 0\}$$

be the set of inequality constraints *active* at $\bar{x}$. All the equality constraints are, of course, active at any feasible point. Consider now the cone of directions obtained by linearizing all the constraints active at $\bar{x}$:

$$\mathcal{L}_D(\bar{x}) = \{d \in \mathbf{R}^n \mid h'(\bar{x})d = 0,\ \langle g_i'(\bar{x}), d\rangle \leq 0\ \forall\, i \in A(\bar{x})\},$$

which is an intuitively natural candidate to represent directions tangent to D at the point $\bar{x}$. It is easy to see that $\mathcal{T}_D(\bar{x}) \subset \mathcal{L}_D(\bar{x})$ always. The fundamental question is when in fact it holds that

$$\mathcal{T}_D(\bar{x}) = \mathcal{L}_D(\bar{x}). \tag{7}$$

When the latter is the case, applying a theorem of the alternatives [2] to compute the dual of $\mathcal{L}_D(\bar{x})$, we have that

$$(\mathcal{T}_D(\bar{x}))^\circ = \{z \in \mathbf{R}^n \mid z = \sum_{i=1}^{l} \lambda_i h_i'(\bar{x}) + \sum_{i \in A(\bar{x})} \mu_i g_i'(\bar{x}),\ \lambda \in \mathbf{R}^l,\ \mu_i \geq 0\ \forall\, i \in A(\bar{x})\}. \tag{8}$$

Then the characterization (6) of solutions of the variational problem (5) immediately translates into the following: there exists $(\lambda, \mu) \in \mathbf{R}^l \times \mathbf{R}^m$ such that

$$\begin{gathered} -F(\bar{x}) = \sum_{i=1}^{l} \lambda_i h_i'(\bar{x}) + \sum_{i=1}^{m} \mu_i g_i'(\bar{x}), \\ h(\bar{x}) = 0, \\ g(\bar{x}) \leq 0, \quad \mu \geq 0, \quad \mu_i = 0\ \forall\, i \notin A(\bar{x}). \end{gathered} \tag{9}$$

In the case of $F(x) = f'(x), x \in \mathbf{R}^n$, corresponding to the optimization problem, (9) are the *Karush-Kuhn-Tucker optimality conditions*

$$\begin{gathered} \frac{\partial L}{\partial x}(\bar{x}, \lambda, \mu) = 0, \\ h(\bar{x}) = 0, \\ g(\bar{x}) \leq 0, \quad \mu \geq 0, \quad \mu_i = 0\ \forall\, i \notin A(\bar{x}), \end{gathered} \tag{10}$$

where

$$L : \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m \to \mathbf{R}, \quad L(x, \lambda, \mu) = f(x) + \langle \lambda, h(x)\rangle + \langle \mu, g(x)\rangle$$

is the Lagrangian of the problem.

The key question, therefore, is when (7) or more generally (8) are guaranteed to hold. Obviously, (7) is sufficient for (8) but not necessary. Condition (7) is called Abadie CQ [3], and condition (8) is called Guignard CQ [4]. Guignard CQ is in a sense the weakest CQ that ensures that KKT conditions (10) are necessary optimality conditions for the associated optimization problem [5]. It should be emphasized though that neither Abadie CQ nor Guignard CQ is verifiable directly in general, since they require the knowledge of the tangent cone or its dual. They are more akin the desired properties we would like to ensure than CQs as such.

Before proceeding with the presentation of CQs, we make one final observation. CQs and the desired equality (7) depend not only on the geometry of the set D but also on its analytic representation, i.e., on the choice of the constraint functions h and g in (1). For example, consider the set $D = \{0\}$ which has the unique feasible point $\bar{x} = 0$, so that $\mathcal{T}_D(\bar{x}) = \{0\}$. If this set is represented by the equality constraint with $h(x) = x$, then in (1) we get $\mathcal{L}_D(\bar{x}) = \{0\}$ which gives the correct tangent cone, i.e., (7) holds. If, on the other hand, the set is represented by the equality constraint with $h(x) = \|x\|^2$, then in (1) we get $\mathcal{L}_D(\bar{x}) = \mathbf{R}^n$ and (7) is no longer valid.

Probably the most obvious CQ that guarantees (7) is *linearity* of constraints around the point in question. Specifically, if there exists a neighborhood U of $\bar{x}$ such that

$$h \text{ and } g_i \ \forall\, i \in A(\bar{x}) \text{ are affine in } U,$$

then (7) holds. Moreover, in that case the tangent cone is the set of feasible directions at $\bar{x}$.

The linear independence constraint qualification (LICQ) consists of saying that the gradients of equality and active inequality constraints are linearly independent at $\bar{x}$:

$$\text{the set } \{h_i'(\bar{x}),\ i = 1, \ldots, l\} \cup \{g_i'(x),\ i \in A(\bar{x})\} \text{ is linearly independent.} \tag{11}$$

Apart from the characterization (7) of the tangent cone, LICQ further implies that the multiplier (λ, μ) satisfying primal-dual conditions (9) is actually unique.

The *Mangasarian-Fromovitz constraint qualification* (MFCQ, [6]) assumes that

$$\begin{array}{c} \text{the set } \{h_i'(\bar{x}),\ i = 1, \ldots, l\} \text{ is linearly independent and} \\ \exists\, d \in \mathbf{R}^n \text{ such that } h'(\bar{x})d = 0,\ \langle g_i'(\bar{x}), d\rangle < 0 \ \forall\, i \in A(\bar{x}). \end{array} \tag{12}$$

Applying a theorem of the alternatives [2], the equivalent dual form of MFCQ states that

$$\begin{array}{c} \text{zero is the unique solution of the linear system} \\ \displaystyle\sum_{i=1}^{l} \lambda_i h_i'(\bar{x}) + \sum_{i \in A(\bar{x})} \mu_i g_i'(\bar{x}) = 0, \quad \mu_i \geq 0 \ \forall\, i \in A(\bar{x}). \end{array} \tag{13}$$

MFCQ (12) also implies the characterization (7) of the tangent cone, while being evidently weaker than LICQ (11). Also, using the dual form (13), it is not difficult to see that at solutions of the variational problem (5) MFCQ is equivalent to the property of the multiplier set of (λ, μ) satisfying the primal-dual conditions (9) being nonempty and bounded [7]. If $\bar{x}$ satisfies the primal-dual conditions (9) with some (λ, μ), then the stronger property of uniqueness of Lagrange multipliers is called the *strict Mangasarian-Fromovitz constraint qualification* (SMFCQ, [8]). MFCQ is stable in the following sense. If MFCQ holds at $\bar{x} \in D$ then there exists a neighborhood U of $\bar{x}$ such that MFCQ holds at each $x \in D \cap U$.

Slater constraint qualification consists of the following assumptions:

$$\begin{array}{c} h \text{ is an affine function, each } g_i \text{ is a convex function, and} \\ \exists\, \hat{x} \in \mathbf{R}^n \text{ such that } h(\hat{x}) = 0,\ g_i(\hat{x}) < 0 \ \forall\, i \in \{1, \ldots, m\}. \end{array}$$

If g is convex differentiable and no equations appear, then Slater CQ is equivalent to MFCQ (12) holding at every point $\bar{x} \in D$ [9, 10].

In the case when the description (1) of the set D does not contain inequality constraints LICQ, MFCQ and SMFCQ all reduce to the classical *regularity* condition

$$\operatorname{rank} h'(\bar{x}) = l.$$

The fact that under this assumption $\mathcal{T}_D(\bar{x})$ is the tangent subspace $\ker h'(\bar{x})$ is a consequence of the Lyusternik-Graves Theorem [11, 12]; see also [13]. Furthermore, KKT conditions (10) for optimization reduce in this case to the classical Lagrange optimality conditions.

The *constant rank constraint qualification* (CRCQ, [14]) holds at $\bar{x} \in D$ if there exists a neighborhood U of $\bar{x}$ such that

$$\begin{array}{c}\text{for every pair of index sets } I \subset \{1, \dots, l\} \text{ and } J \subset A(\bar{x}) \\ \text{the set } \{h_i'(x),\ i \in I\} \cup \{g_i'(x),\ i \in J\} \text{ has the same rank for all } x \in D \cap U.\end{array} \tag{14}$$

In (14) the rank in question depends on the choice of I and J but not on the point $x \in D \cap U$. Clearly, LICQ (11) implies CRCQ (14). Linearity of constraints also implies CRCQ. Under CRCQ it holds that the tangent cone $\mathcal{T}_D(\bar{x})$ has the form (7) [14]. CRCQ is neither weaker nor stronger than MFCQ (12). Thus nothing can be said about the multiplier set in KKT conditions (9), except that it is nonempty. Note also that unlike MFCQ, if CRCQ holds at $\bar{x} \in D$, it will continue to hold if any of the equality constraints $h_i(x) = 0$ were to be replaced by the two inequalities $h_i(x) \leq 0$ and $-h_i(x) \leq 0$. MFCQ and CRCQ are related, however, in the following sense: it can be shown that under CRCQ there exists an alternative representation of the feasible set for which MFCQ holds [15].

The *relaxed constant rank constraint qualification* (rCRCQ, [16]) holds at $\bar{x} \in D$ if there exists a neighborhood U of $\bar{x}$ such that

$$\begin{array}{c}\text{for every index set } J \subset A(\bar{x}) \\ \text{the set } \{h_i'(x),\ i = 1, \dots, l\} \cup \{g_i'(x),\ i \in J\} \text{ has the same rank for all } x \in U.\end{array} \tag{15}$$

It is clear that CRCQ (14) implies rCRCQ (15). It can be seen from the following example [16] that the reverse implication is not valid: $D = \{x \in \mathbf{R}^2 \mid x_1 - x_2 = 0,\ -x_1 \leq 0,\ -x_1 - x_2^2 \leq 0\}$, $\bar{x} = 0$. It holds that rCRCQ is still sufficient for the tangent cone $\mathcal{T}_D(\bar{x})$ to have the desired form (7) [16]. When there are no inequality constraints, rCRCQ reduces to the *weak constant rank condition* introduced in [17].

The point $\bar{x} \in D$ satisfies the *constant positive linear dependence* condition (CPLD, [19]) if there exists a neighborhood U of $\bar{x}$ such that

$$\begin{array}{c}\text{whenever for some index sets } I \subset \{1, \dots, l\} \text{ and } J \subset A(\bar{x}) \text{ the system} \\ \displaystyle\sum_{i \in I} \lambda_i h_i'(\bar{x}) + \sum_{i \in J} \mu_i g_i'(\bar{x}) = 0, \quad \mu_i \geq 0\ \forall\, i \in J \\ \text{has a nonzero solution,} \\ \text{the set } \{h_i'(x),\ i \in I\} \cup \{g_i'(x),\ i \in J\} \text{ is linearly dependent for all } x \in U.\end{array} \tag{16}$$

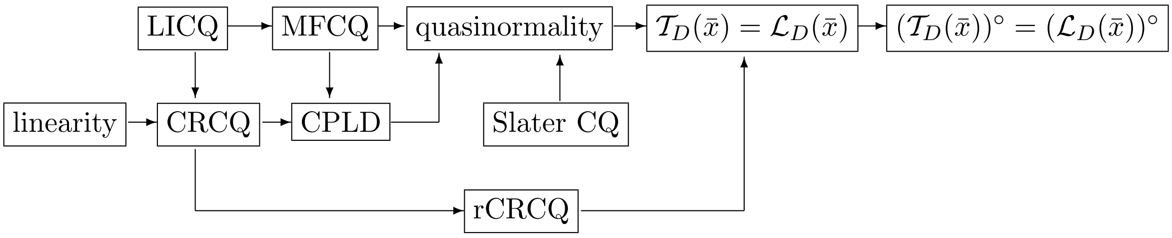
Comparing the dual form (13) of MFCQ with (16), it is immediate that MFCQ implies CPLD but not vice versa. It can be seen that CPLD is also weaker than CRCQ [20]. Nevertheless, CPLD still guarantees that the tangent cone has the desired representation (7); this follows from the results in [20, 21]; and also from the error bound in [16], see Section 3. Hence, KKT conditions (10) are necessary optimality conditions, which had also been shown in [20]. Finally, CPLD is neither weaker nor stronger than rCRCQ, as can be seen from the following example [16]: $D = \{x \in \mathbf{R}^2 \mid x_2 = 0,\ x_1 - x_2^2 \leq 0,\ -x_1 - x_2^2 \leq 0\}$, $\bar{x} = 0$.

The point $\bar{x} \in D$ is said to be *quasinormal* [22] (see also [21, 23]) if

$$\begin{array}{c}\text{there exist no nonzero } (\lambda, \mu) \in \mathbf{R}^l \times \mathbf{R}^m_+ \text{ and no sequence } \{x^k\} \to \bar{x} \text{ such that} \\ \displaystyle\sum_{i=1}^{l} \lambda_i h_i'(\bar{x}) + \sum_{i=1}^{m} \mu_i g_i'(\bar{x}) = 0, \\ \text{and for all } k, \lambda_i h_i(x^k) > 0 \text{ for all } i \text{ with } \lambda_i \neq 0,\ \mu_i g_i(x^k) > 0 \text{ for all } i \text{ with } \mu_i \neq 0.\end{array} \tag{17}$$

Quasinormality implies that the tangent cone has the desired representation (7) [22]. Quasinormality is implied by CPLD [20].

The main relationships discussed above can be summarized as follows:



When comparing different constraint qualifications, it should also be kept in mind that conditions like CPLD (16) and rCRCQ (15) depend not only on the properties of the problem data at the point $\bar{x}$ but also on their properties in some neighborhood of this point. They require more information and, unless some stronger sufficient conditions hold, they are usually much more difficult to verify directly than, say, the classical MFCQ (12).

The more general format of constraints is given by

$$D = \{x \in \mathbf{R}^n \mid S(x) \in Q\}, \tag{18}$$

where Q is a subset of $\mathbf{R}^s$ and $S : \mathbf{R}^n \to \mathbf{R}^s$. The set (1) defined by a finite number of equality and inequality constraints is clearly a special case of (18) given by $s = l + m$, $S(x) = (h(x), g(x))$ and $Q = \{0\} \times (-\mathbf{R}^m_+)$. When Q is a closed convex set, the *Robinson CQ* [10] holds at $\bar{x} \in D$ if

$$0 \in \operatorname{int} \{S(\bar{x}) + \operatorname{im} S'(\bar{x}) - Q\}. \tag{19}$$

Robinson CQ ensures that

$$\mathcal{T}_D(\bar{x}) = \{d \in \mathbf{R}^n \mid S'(\bar{x})d \in \mathcal{T}_Q(S(\bar{x}))\}. \tag{20}$$

Furthermore, if Q is a closed convex cone and $\bar{x}$ is a local solution of the optimization problem (2) then there exists $\nu \in \mathbf{R}^s$ such that

$$f'(\bar{x}) + (S'(\bar{x}))^\top \nu = 0, \quad S(\bar{x}) \in Q,\ \nu \in Q^\circ,\ \langle \nu, S(\bar{x}) \rangle = 0. \tag{21}$$

In the case of equality and inequality constraints (1), Robinson CQ (19) reduces to MFCQ (12), characterization of the tangent cone (20) reduces to (7), and optimality conditions (21) reduce to the KKT conditions (10).

2 Second-order necessary optimality conditions

Constraint qualifications are also important for deriving second-order necessary optimality conditions. But not all CQs discussed in Section 1 are suitable for this purpose.

In this section the problem data is assumed to be twice continuously differentiable. Let $\bar{x} \in D$ be a local minimizer of f in D. We shall denote by $\mathcal{M}(\bar{x})$ the set of Lagrange multipliers at $\bar{x}$, i.e., all $(\lambda, \mu) \in \mathbf{R}^l \times \mathbf{R}^m$ satisfying the KKT conditions (10). Denote by

$$\mathcal{C}(\bar{x}) = \{d \in \mathbf{R}^n \mid \langle f'(\bar{x}), d\rangle \le 0,\ h'(\bar{x})d = 0,\ \langle g_i'(\bar{x}), d\rangle \le 0\ \forall\, i \in A(\bar{x})\}$$

the *critical cone* of the optimization problem at $\bar{x} \in D$. Then under MFCQ (12) it holds that

$$\forall\, d \in \mathcal{C}(\bar{x}) \quad \exists\, (\lambda, \mu) \in \mathcal{M}(\bar{x}) \quad \text{such that} \quad \left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \lambda, \mu)d, d \right\rangle \geq 0. \tag{22}$$

It is known [24] that when SMFCQ does not hold (i.e., $\mathcal{M}(\bar{x})$ is not a singleton) the stronger version

$$\exists\, (\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x}) \quad \text{such that} \quad \left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}, \bar{\mu})d, d \right\rangle \geq 0 \quad \forall\, d \in \mathcal{C}(\bar{x}) \tag{23}$$

is not a necessary optimality condition in general. At the same time, under CRCQ (14) even a stronger property than (23) is valid [25]:

$$\forall\, (\lambda, \mu) \in \mathcal{M}(\bar{x}) \quad \left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \lambda, \mu)d, d \right\rangle \geq 0 \quad \forall\, d \in \mathcal{C}(\bar{x}).$$

Weaker forms of second-order conditions make use of the subspace

$$\mathcal{C}^+(\bar{x}) = \{d \in \mathbf{R}^n \mid h'(\bar{x})d = 0,\ \langle g_i'(\bar{x}), d\rangle = 0\ \forall\, i \in A(\bar{x})\},$$

which is in general smaller than the critical cone $\mathcal{C}(\bar{x})$. The two cones coincide when the *strict complementarity* condition $\bar{\mu}_i > 0\ \forall\, i \in A(\bar{x})$ holds. The following result had been established in [25]. If $\bar{x}$ is a local minimizer satisfying KKT conditions (i.e., $\mathcal{M}(\bar{x}) \neq \emptyset$) and the weak constant rank condition holds, i.e., there is a neighborhood U of $\bar{x}$ such that

$$\text{the set } \{h_i'(x),\ i = 1, \ldots, l\} \cup \{g_i'(x),\ i \in A(\bar{x})\} \text{ has the same rank for all } x \in U,$$

then

$$\forall\, (\lambda, \mu) \in \mathcal{M}(\bar{x}) \quad \left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \lambda, \mu)d, d \right\rangle \geq 0 \quad \forall\, d \in \mathcal{C}^+(\bar{x}). \tag{24}$$

In particular, any CQ that ensures that $\mathcal{M}(\bar{x}) \neq \emptyset$ (see Section 1) in combination with the weak constant rank condition guarantees that (24) holds at a local minimizer $\bar{x}$. For example, (24) holds under rCRCQ (15), as the latter implies both $\mathcal{M}(\bar{x}) \neq \emptyset$ and the weak constant rank condition.

3 Error bounds and metric regularity

Describing local structure of a set D defined by a finite number of equality and inequality constraints (1) is also closely related to the so-called *error bounds*, which are estimates of the distance from a given point to the set D in terms of computable quantities measuring violation of its constraints. Specifically, one would like to know when there exist a neighborhood U of $\bar{x} \in D$ and a constant $c > 0$ such that

$$\operatorname{dist}(x, D) \leq c(\|h(x)\| + \|\max\{0, g(x)\}\|) \quad \forall x \in U. \tag{25}$$

Indeed, if the estimate (25) is valid then the consequences for the characterization of the tangent cone as in (7) are immediate. It is enough to observe that for $d \in \mathcal{L}_D(\bar{x})$ and $t > 0$ small enough it holds that $h(\bar{x} + td) = h(\bar{x}) + th'(\bar{x})d + o(t) = o(t)$, $g_i(\bar{x} + td) < 0$ for all $i \notin A(\bar{x})$ and

$g_i(\bar{x}+td) = g_i(\bar{x}) + t\langle g_i'(\bar{x}), d\rangle + o(t) \le o(t)$ for all $i \in A(\bar{x})$. Then (25) immediately implies that $\mathrm{dist}(x+td, D) = o(t)$ so that $d \in \mathcal{T}_D(\bar{x})$.

In the case of linear constraints the error bound (25) (with $U = \mathbf{R}^n$) is the classical Hoffman's Lemma [26]. More generally, (25) is valid under rCRCQ (15) [16, 18]. The error bound is also valid assuming MFCQ (12) [10] or, more generally, CPLD (16) [16].

A stronger property than the error bound (25) is that of *metric regularity* [9, 27]. Consider the right-hand side perturbation of the set D, i.e.,

$$D(p, q) = \{x \in \mathbf{R}^n \mid h(x) = p,\ g(x) \le q\}, \quad p \in \mathbf{R}^l,\ q \in \mathbf{R}^m. \tag{26}$$

Let $\bar{x} \in D(0, 0)$. Then the system in (26) is metrically regular at $(\bar{x}, 0, 0)$ if there exist a neighborhood V of $(\bar{x}, 0, 0)$ and a constant $C > 0$ such that

$$\mathrm{dist}(x, D(p, q)) \le C(\|h(x) - p\| + \|\max\{0, g(x) - q\}\|) \quad \forall\, (x, p, q) \in V.$$

For smooth constraint systems, metric regularity holds if, and only if, MFCQ (12) holds for the unperturbed set D defined in (1) [10, 28]. This is an important stability property that highlights the special role of MFCQ among all the other CQs.

Robinson CQ (19) for the more general smooth constraints (18) is also equivalent to metric regularity in the following sense. Defining the perturbed set

$$D(p) = \{x \in \mathbf{R}^n \mid S(x) + p \in Q\}, \quad p \in \mathbf{R}^s,$$

with $\bar{x} \in D(0)$, metric regularity holds at $(\bar{x}, 0)$ if there exist a neighborhood V of $(\bar{x}, 0)$ and a constant $C > 0$ such that

$$\mathrm{dist}(x, D(p)) \le C\,\mathrm{dist}(S(x) + p, Q) \quad \forall\, (x, p) \in V.$$

4 Second-order regularity

Constraint qualifications discussed until now were based on at most first-order information about the constraint functions. Sometimes, in particular when classical CQs are violated, second derivatives need to be employed.

One line of analysis is concerned with deriving second-order necessary optimality conditions of the type (22) under assumptions weaker than MFCQ (12). For a feasible set defined by (1), the *second-order regularity condition* holds at $\bar{x} \in D$ *in a direction* $d \in \mathbf{R}^n$ if

$$\begin{array}{c}\text{the set } \{h_i'(\bar{x}),\ i = 1,\, \ldots,\, l\} \text{ is linearly independent and} \\ \exists\, \xi \in \mathbf{R}^n \text{ such that } h'(\bar{x})\xi + h''(\bar{x})[d,\, d] = 0,\ \langle g_i'(\bar{x}),\, \xi\rangle + \langle g_i''(\bar{x})d,\, d\rangle < 0\ \forall\, i \in A(\bar{x}).\end{array} \tag{27}$$

(see [29] for a different but equivalent statement of this property). It can be easily seen that (27) holds automatically in any direction $d \in \mathbf{R}^n$ provided MFCQ (12) holds at $\bar{x}$. But (27) may hold for a given $d \in \mathbf{R}^n$ when MFCQ is violated. When second-order regularity condition holds at $\bar{x}$ in a direction $d \in \mathcal{L}_D(\bar{x})$, one can constructively characterize the so-called *second-order tangent set in a direction* d, defined by

$$\mathcal{T}_2(\bar{x},\, d) = \left\{ \xi \in \mathbf{R}^n \;\middle|\; \begin{array}{c} \exists\, \{t_k\} \subset \mathbf{R}_+ \setminus \{0\},\ \{t_k\} \to 0, \exists\, \{\xi^k\} \subset \mathbf{R}^n,\ \{\xi^k\} \to \xi, \\ \text{such that } \bar{x} + t_k d + \frac{1}{2} t_k^2 \xi^k \in D \text{ for all } k \end{array} \right\}.$$

Furthermore, for $d \in \mathcal{C}(\bar{x})$, this characterization allows to establish the second-order necessary optimality condition of the form

$$\exists\,(\lambda,\,\mu) \in \mathcal{M}(\bar{x}) \quad \text{such that} \quad \left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x},\,\lambda,\,\mu)d,\,d \right\rangle \geq 0. \tag{28}$$

Note that the latter subsumes that $\mathcal{M}(\bar{x})$ is nonempty, which means that the KKT conditions (10) hold as well.

Another important concept of second-order regularity was developed in [30, 31] for the case when there are no inequality constraints, under the name of 2-*regularity*. In somewhat different terms, we say that this condition holds at $\bar{x} \in D$ *in a direction* $d \in \mathbf{R}^n$ if

$$\operatorname{im} h'(\bar{x}) + h''(\bar{x})[d,\,\ker h'(\bar{x})] = \mathbf{R}^l.$$

The counterpart of this concept for problems with no equality constraints and the related theory were developed in [32]. An extension of these works to the case of equality and inequality constraints, as in (1), was derived in [33]. Specifically, according to [33], 2-regularity holds at $\bar{x} \in D$ in a direction $d \in \mathbf{R}^n$ if

$$\begin{gathered} \operatorname{im} h'(\bar{x}) + h''(\bar{x})[d,\,\mathcal{L}_D(\bar{x})] = \mathbf{R}^l \text{ and} \\ \exists\,\xi^1 \in \mathbf{R}^n,\ \exists\,\xi^2 \in \mathcal{L}_D(\bar{x}) \text{ such that} \\ h'(\bar{x})\xi^1 + h''(\bar{x})[d,\,\xi^2] = 0,\ \langle g_i'(\bar{x}),\,\xi^1\rangle + \langle g_i''(\bar{x})d,\,\xi^2\rangle < 0\ \forall\, i \in A(\bar{x}). \end{gathered} \tag{29}$$

As in the case of second-order regularity (27), 2-regularity (29) holds automatically in any direction $d \in \mathbf{R}^n$ provided MFCQ (12) holds at $\bar{x}$. But (29) may hold for a given $d \in \mathbf{R}^n$ when MFCQ is violated. The role of 2-regularity is to characterize those directions $d \in \mathcal{L}_D(\bar{x})$ that actually belong to $\mathcal{T}(\bar{x})$ in the cases when (7) does not hold. See [33] for the detailed exposition of the related theory of first- and second-order necessary optimality conditions, and in particular, for combinations of 2-regularity with second-order regularity and its relevant extensions. It is important to emphasize that, unlike second-order regularity, 2-regularity does not imply that (10) is necessary for optimality: $\mathcal{M}(\bar{x})$ can be empty, and the first- and second-order necessary optimality conditions that can be established under 2-regularity are generally weaker than (10) and (28), respectively.

Finally, in [34] the 2-regularity theory was extended to the general constraints, as in (18). The corresponding condition of 2-regularity at $\bar{x} \in D$ in a direction $d \in \mathbf{R}^n$ has the form

$$0 \in \operatorname{int}\{S(\bar{x}) + \operatorname{im} S'(\bar{x}) + S''(\bar{x})[d,\,(S'(\bar{x}))^{-1}(Q - S(\bar{x}))] - Q\}.$$

It is clear that this condition is implied by Robinson CQ (19).

References

[1] Kuhn, HW, Tucker, AW. Nonlinear programming. Pp 481–492 in Neyman, J., ed. Proceedings of Second Berkeley Symposium of Mathematical Statistics and Probability. 1950. University of California Press, Berkeley, CA.

[2] Mangasarian, OL. Nonlinear Programming. McGraw Hill, New York. 1969.

[3] Abadie, J. On the Kuhn-Tucker Theorem. In Abadie, J., ed. Nonlinear Programming. North-Holland, Amsterdam. 1967.

[4] Guignard, M. Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space. SIAM J. Control 1969; 7:232–241.

[5] Gould, FJ, Tolle, JW. A necessary and sufficient qualification for constrained optimization. SIAM J. Appl. Math. 1971; 20:164–172.

[6] Mangasarian, OL, Fromovitz, S. The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. J. Math. Anal. Appl. 1967; 17:37–47.

[7] Gauvin, J. A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming. Math. Program. 1977; 12:136–138.

[8] Kyparisis, J. On uniqueness of Kuhn–Tucker multipliers in nonlinear programming. Math. Program. 1985; 32:242–246.

[9] Robinson, SM. Regularity and stability for convex multivalued functions. Math. Oper. Res. 1976; 1:130–143.

[10] Robinson, SM. Stability theorems for systems of inequalities. Part II: Differentiable nonlinear systems. SIAM J. Numer. Anal. 1976; 13:497–513.

[11] Lyusternik, LA. On the conditional extrema of functionals. Math. Sbornik. 1934; 41:390-401.

[12] Graves, LM. Some mapping theorems. Duke Math. J. 1950; 17:111–114.

[13] Dontchev, AL. The Graves theorem revisited. J. Convex Anal. 1996; 3:45–53.

[14] Janin, R. Direction derivative of the marginal function in nonlinear programming. Math. Program. Stud. 1984; 21:127–138

[15] Lu, S. Implications of the constant rank constraint qualification. Math. Program 2009. Published online at DOI 10.1007/s10107-009-0288-3.

[16] Minchenko, L, Stakhovski, S. On relaxed constant rank regularity condition in mathematical programming. Optimization, to appear in 2010.

[17] Penot, JP. A new constraint qualification condition. J. Optim. Theory Appl. 1986; 48:459–468.

[18] Lu, S. Relations between the constant rank and the relaxed constant rank constraint qualifications. Submitted to Optimization, 2009. Available at http://www.unc.edu/~shulu/

[19] Qi, L, Wei, Z. On the constant positive linear independence condition and its application to SQP methods. SIAM J. Optim. 2000; 10:963–981.

[20] Andreani, R, Martínez, JM, Schuverdt, ML. On the relation between the constant positive linear dependence condition and quasinormality constraint qualification. J. Optim. Theory Appl. 2005; 125:473–485.

[21] Bertsekas, DP, Ozdaglar, AE. Pseudonormality and a Lagrange multiplier theory for constrained optimization. J. Optim. Theory Appl. 2002; 114:287–343.

[22] Hestenes, MR. Optimization Theory: The Finite Dimensional Case. Wiley, New York. 1975.

[23] Bertsekas, DP, Ozdaglar, AE. The relation between pseudonormality and quasiregularity in constrained optimization. Optim. Methods Soft. 2004; 19:493–506.

[24] Arutyunov, AV. Perturbations of extremum problems with constraints and necessary optimality conditions. J. Sov. Math. 1991; 54:1342–1400.

[25] Andreani, R, Echagüe, CE, Schuverdt, ML. Constant rank condition and second-order constraint qualification. Optimization 2010.

[26] Hoffman, AJ. On approximate solutions of systems of linear inequalities. J. Research of the National Bureau of Standards. 1952; 49:263–265.

[27] Borwein, JM. Stability and regular points of inequality systems. J. Optim. Theory Appl. 1986; 48:9–52.

[28] Cominetti, R. Metric regularity, tangent sets and second-order optimality conditions. Appl. Math. Optim. 1990; 21:265–287.

[29] Ben-Tal, JF. Second-order and related extremality conditions in nonlinear programming. J. Optim. Theory Appl. 1980; 31:143–165.

[30] Tretyakov, AA. Necessary and sufficient conditions for optimality of p-th order. USSR Comput. Math. Math. Phys. 1984; 24:123–127.

[31] Avakov, ER. Extremum conditions for smooth problems with equality-type constraints. USSR Comput. Math. Math. Phys. 1985; 25:24–32.

[32] Izmailov, AF, Solodov, MV. Optimality conditions for irregular inequality-constrained problems. SIAM J. Control. Optim. 2001; 40:1280–1295.

[33] Arutyunov, AV, Avakov ER, Izmailov, AF. Necessary conditions for an extremum in a mathematical programming problem. Proceedings of the Steklov Institute of Mathematics. 2007; 256:2–25.

[34] Arutyunov, AV, Avakov ER, Izmailov, AF. Necessary optimality conditions for constrained optimization problems under relaxed constrained qualifications. Math. Program. 2008; 114:37–68.

Further reading list

For extensions to infinite-dimensional spaces and nonsmooth data, applications in duality and sensitivity, see:

Rockafellar, RT, Wets, JB. Variational Analysis. Springer–Verlag, New York. 1997.

Bonnans, JF, Shapiro, A. Perturbation Analysis of Optimization Problems. Springer–Verlag, New York. 2000.

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

Facchinei, F, Pang, JS. Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer–Verlag, New York. 2003.

Mordukhovich, BS. Variational Analysis and Differentiation. Springer–Verlag, New York. 2006.