A survey of constraint qualifications with second-order properties in nonlinear optimization

Gabriel Haeser*       Alberto Ramos†

January 7, 2018

Abstract

In this paper we discuss recent developments in first- and second-order constraint qualifications that imply second-order optimality conditions. In the first-order case, we are particularly interested in conditions that ensure the validity of second-order necessary optimality conditions that can be checked with a single Lagrange multiplier and defined in terms of positive semidefiniteness on a subspace, due to its connections with global convergence of second-order algorithms. In the second-order case, we discuss necessary and sufficient conditions to ensure the validity of second-order optimality conditions.

Keywords: Nonlinear optimization, Constraint qualifications, Global convergence, Second-order optimality conditions.
AMS Classification: 90C46, 90C30

1 Introduction

Numerical optimization deals with the design of algorithms with the aim of finding a point with the lowest possible value of a certain function over a constraint set. Useful tools for the design of algorithms are the necessary optimality conditions, i.e., conditions satisfied by every local minimizer. Not all necessary optimality conditions serve that purpose. Optimality conditions should be computable with the information provided by the algorithm, where its fulfillment indicates that the considered point is an acceptable solution. For constrained optimization problems, the Karush-Kuhn-Tucker (KKT) conditions are the basis for most optimality conditions. In fact, most algorithms for constrained optimization are iterative and in their implementation, the KKT conditions serve as a theoretical guide for developing suitable stopping criteria. For more details, see [55, Framework 7.13, page 513], [35, Chapter 12] and [9].

---

*Department of Applied Mathematics, University of São Paulo, São Paulo-SP, Brazil. Email: ghaeser@ime.usp.br
†Department of Mathematics, Federal University of Paraná, Curitiba, PR, Brazil. e-mail: albertoramos@ufpr.br.
When the second-order information is available, necessary optimality conditions can be of first- or second-order depending on whether the first- or second-order derivatives are used in the formulation. Second-order necessary optimality conditions are much stronger than first-order ones and hence are mostly desirable, since they allow ruling out possible non-minimizers accepted as solution when only first-order information is considered.

Global convergence proofs of second-order algorithms are based on second-order necessary optimality condition of the form: If a local minimizer satisfies some constraint qualification, then the WSOC condition holds, where WSOC stands for the Weak Second-order Optimality Condition, that states that the Hessian of the Lagrangian at a KKT point, for some Lagrange multiplier, is positive semidefinite on the subspace orthogonal to the gradients of active constraints, see Definition 2.1.

Thus, we are interested in assumptions guaranteeing that local minimizers satisfy WSOC, given its implications to numerical algorithms.

When seeking for the weakest possible constraint qualification that implies the validity of WSOC at local minimizers, it is more than natural to consider second-order constraint qualifications, that is, to take into account the second-order information of the constraints in its formulation. We will pursue this goal in this paper, and in this context, it is natural to consider also minimal constraint qualifications ensuring more standard second-order optimality conditions in nonlinear optimization theory (see Definitions 2.1 and 2.2).

In Section 2 we present the basic concepts and definitions while reviewing first-order constraint qualifications with second-order properties, highlighting the central aspect of WSOC in second-order algorithms. In Section 3 we discuss second-order constraint qualifications and we define the weakest possible ones with respect to several second-order optimality conditions. In Section 4 we present our concluding remarks.

## 2 First-order Constraint Qualifications

A constraint qualification is any property about the analytic description of the feasible set around a local minimizer that ensures the existence of Lagrange multipliers. We are particularly interested in constraint qualifications that guarantee the existence of special Lagrange multipliers, namely, the ones that can be used to formulate second-order necessary optimality conditions. A constraint qualifications is called a first-order one if it only uses the gradients of the constraints in its formulation, while a second-order one is defined in terms of gradients and Hessians. Both types of conditions can yield second-order optimality conditions and in this section we review the first-order ones.

First, we start with the basic notation. \( \mathbb{R}^n \) stands for the \( n \)-dimensional real Euclidean space, \( n \in \mathbb{N} \). \( \mathbb{R}^n_+ \subset \mathbb{R}^n \) is the set of vectors whose components are non-negative. The canonical basis of \( \mathbb{R}^n \) is denoted by \( e_1, \ldots, e_n \). A set \( \mathcal{R} \subset \mathbb{R}^n \) is a ray if \( \mathcal{R} := \{rd_0 : r \geq 0 \} \) for some \( d_0 \in \mathbb{R}^n \). Given a convex cone \( \mathcal{K} \subset \mathbb{R}^n \), we define the linearity set of \( \mathcal{K} \) as \( \mathcal{K} \cap -\mathcal{K} \), which is the largest subspace contained
in \( \mathcal{K} \). We say that \( \mathcal{K} \) is a first-order cone if \( \mathcal{K} \) is the direct sum of a subspace and a ray. The closed unit ball around the origin in \( \mathbb{R}^n \) is denote by \( B \), while \( B(x, \eta) := x + \eta B \) is the closed ball centered at \( x \) with radius \( \eta > 0 \). The set of symmetric matrix of order \( n \) is denoted by \( \text{Sym}(n) \). We use \( \langle \cdot, \cdot \rangle \) to denote the Euclidean inner product on \( \mathbb{R}^n \) and on \( \text{Sym}(n) \), with \( \| \cdot \| \) the associated norm. Given a symmetric \( n \times n \) matrix \( A \) and \( v \in \mathbb{R}^n \), we use \( Av^2 \) to denote \( \langle v, Av \rangle \).

Given the set \( S \), the symbol \( z \overset{S}{\to} z^\ast \) means that \( z \to z^\ast \) with \( z \in S \). For a cone \( \mathcal{K} \subset \mathbb{R}^n \), its polar (negative dual) is \( \mathcal{K}^\circ := \{ v \in \mathbb{R}^n | \langle v, k \rangle \leq 0 \text{ for all } k \in \mathcal{K} \} \).

Consider the non-linear constrained optimization problem

\[
\begin{array}{ll}
\text{minimize} & f(x), \\
\text{subject to} & h_i(x) = 0 \quad \forall i \in \mathcal{E} := \{1, \ldots, m\}, \\
& g_j(x) \leq 0 \quad \forall j \in \mathcal{I} := \{1, \ldots, p\},
\end{array}
\tag{2.1}
\]

where \( f, h_i, g_j : \mathbb{R}^n \to \mathbb{R} \) are assumed to be, at least, twice continuously differentiable functions.

Denote by \( \Omega \) the feasible set of (2.1). For a point \( x \in \Omega \), we define \( A(x) := \{ j \in \mathcal{I} : g_j(x) = 0 \} \) as the set of indices of active inequalities. A feasible point \( x^* \in \Omega \) satisfies the Mangasarian-Fromovitz Constraint Qualification (MFCQ) if \( \{ \nabla h_i(x^*) : i \in \mathcal{E} \} \) is a linearly independent set and there is a direction \( d \in \mathbb{R}^n \) such that \( \nabla h_i(x^*)^T d = 0, i \in \mathcal{E} \) and \( \nabla g_j(x^*)^T d < 0, j \in A(x^*) \). Given \( x^* \), define by \( J(x) \) the matrix whose first \( m \) rows are formed by \( \nabla h_i(x)^T, i \in \mathcal{E} \) and the remaining rows by \( \nabla g_j(x)^T, j \in A(x^*) \).

We denote the Lagrangian function by \( L(x, \lambda, \mu) := f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{p} \mu_j g_j(x) \) where \( (x, \lambda, \mu) \) is in \( \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p_+ \) and the generalized Lagrangian function as \( L^g(x, \lambda_0, \lambda, \mu) := \lambda_0 f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{p} \mu_j g_j(x) \) where \( (x, \lambda_0, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}^m \times \mathbb{R}^p_+ \). Clearly, \( L^g(x, 1, \lambda, \mu) = L(x, \lambda, \mu) \). The symbols \( \nabla_x L^g(x, \lambda_0, \lambda, \mu) \) and \( \nabla^2_{xx} L^g(x, \lambda_0, \lambda, \mu) \) stand for the gradient and the Hessian of \( L^g(x, \lambda_0, \lambda, \mu) \) with respect to \( x \), respectively. Similar notation holds for \( L(x, \lambda, \mu) \).

The generalized first-order optimality condition at the feasible point \( x^* \) is

\[
\nabla L^g_2(x^*, \lambda_0, \lambda, \mu) = 0 \quad \text{with} \quad \mu^T g(x^*) = 0, \quad \lambda_0 \geq 0, \quad \mu \geq 0, \quad (\lambda_0, \lambda, \mu) \neq (0, 0, 0).
\tag{2.2}
\]

The set of vectors \( (\lambda_0, \lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^m_+ \times \mathbb{R}^p_+ \) satisfying (2.2) is the set of generalized Lagrange multipliers (or Fritz John multipliers), denoted by \( \Lambda_0(x^*) \).

Note that (2.2) with \( \lambda_0 = 1 \) corresponds to the Karush-Kuhn-Tucker (KKT) conditions, the standard first-order condition in numerical optimization. We denote by \( \Lambda(x^*) := \{ (\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p_+ : (1, \lambda, \mu) \in \Lambda_0(x^*) \} \) the set of all Lagrange multipliers. At every minimizer, there are Fritz John multipliers such that (2.2) holds, that is, \( \Lambda_0(x^*) \neq \emptyset \). In order to get existence of Lagrange multipliers, additional assumptions have to be required. Assumptions on the analytic description of the feasible set that guarantee the validity of the KKT conditions at local minimizers are called constraint qualification (CQ). Thus, under any CQ, the KKT conditions are necessary for optimality.

When the second-order information is available, we can consider second-order conditions. Several second-order optimality conditions have been proposed.
in the literature, both from a theoretical and practical point of view, see [15, 59, 55, 35, 34, 17, 22, 58, 24, 21, 12, 14, 28, 36, 20] and references therein. In order to describe second-order conditions, we introduce some important sets. We start with the cone of critical directions (critical cone), defined as follows:

\[ C(x^*) := \{ d \in \mathbb{R}^n \mid \nabla f(x^*)^T d = 0; \nabla h_i(x^*)^T d = 0, i \in E; \nabla g_j(x^*)^T d \leq 0, j \in A(x^*) \}. \quad (2.3) \]

Obviously, \( C(x^*) \) is a non-empty closed convex cone. When \( \Lambda(x^*) \neq \emptyset \), the critical cone \( C(x^*) \) can be written as

\[
\left\{ d \in \mathbb{R}^n : \begin{array}{l}
\nabla h_i(x^*)^T d = 0, \text{ for } i \in E, \\
\nabla g_j(x^*)^T d \leq 0, \text{ if } \mu_j > 0 \\
\n\text{if } \mu_j = 0, j \in A(x^*)
\end{array} \right\}, \quad (2.4)
\]

for every \((\lambda, \mu) \in \Lambda(x^*)\). From the algorithmic point of view, an important set is the critical subspace (or weak critical cone), given by:

\[ S(x^*) := \{ d \in \mathbb{R}^n \mid \nabla h_i(x^*)^T d = 0, i \in E; \nabla g_j(x^*)^T d = 0, j \in A(x^*) \}. \quad (2.5) \]

In the case when \( \Lambda(x^*) \neq \emptyset \), a simple inspection shows that the critical subspace \( S(x^*) \) is the linearity space of the critical cone \( C(x^*) \). Under strict complementarity, \( S(x^*) \) coincides with \( C(x^*) \).

Now, we are able to define the classical second-order conditions.

**Definition 2.1.** Let \( x^* \) be a feasible point with \( \Lambda(x^*) \neq \emptyset \). We have the following definitions

1. We say that the **strong second-order optimality condition** (SSOC) holds at \( x^* \) if there is a \((\lambda, \mu) \in \Lambda(x^*)\) such that \( d^T \nabla^2 L(x^*, \lambda, \mu) d \geq 0 \) for every \( d \in C(x^*) \).

2. We say that the **weak second-order optimality condition** (WSOC) holds at \( x^* \) if there is a \((\lambda, \mu) \in \Lambda(x^*)\) such that \( d^T \nabla^2 L(x^*, \lambda, \mu) d \geq 0 \) for every \( d \in S(x^*) \).

The classical second-order condition SSOC is particularly important from the point of view of passing from necessary to sufficient optimality conditions. In this case, strengthening the sign of the inequality from \( \geq 0 \) to \( > 0 \) in the definition of SSOC, that is, instead of positive semi-definiteness of the Hessian of the Lagrangian on the critical cone, we require its positive definiteness on the same cone (minus the origin), we get a sufficient optimality condition, see [15, 22]. Furthermore, this sufficient condition also ensures that the local minimizer \( x^* \) is isolated. Besides these nice properties, from the practical point of view, SSOC has some disadvantages. In fact, to verify the validity of SSOC at a given point, is in general, an NP-hard problem, [54, 56]. Also, it is well known that very simple second-order methods fail to generate sequences in which SSOC holds at its accumulation points, see [39].

As we have mentioned, WSOC has two important features that makes it relevant in practical algorithms, which is our main motivation for describing it.
The first one is that it does not rely on the whole set of Lagrange multipliers, in contrast with other second-order conditions in the literature, and the second one is that positive semi-definiteness of the Hessian of the Lagrangian must be verified in a subspace (a more tractable task) rather than at a pointed cone.

This is compatible with the implementation of an algorithm that globally converges to a point $x^*$ fulfilling WSOC. At each iteration, one has available an approximation $x^k$ to a solution and a single Lagrange multiplier approximation $(\lambda^k, \mu^k)$ if one wishes to declare convergence to a second-order stationary point (see details in [4] and references therein). Of course, this is still a non-trivial computational task, so this only makes sense when most of the effort for checking WSOC has already been done as part of the computation of the iterate. This is the case of algorithms that try to compute a descent direction and a negative curvature direction [2, 51]. Near a KKT point, once the procedure for computing the negative curvature direction fails, WSOC is approximately satisfied.

This is an important difference with respect to other second-order conditions. In order to verify an optimality condition that relies on the whole set of Lagrange multipliers, one needs an algorithm that generates all multipliers, which may be difficult. Even more, in classical second-order conditions, one must check if a matrix is positive semi-definite on a pointed cone, which is a far more difficult problem than checking it on a subspace (see [54]). Finally, we are not aware of any reasonable iterative algorithm that generates subsequences that converge to a point that satisfies a classical, more accurate, second-order optimality condition based on a pointed cone. The discussion in [39] indicates that such algorithms probably do not exist.

From this point of view, WSOC seems to be the most adequate second-order condition when dealing with global convergence of second-order methods. In fact, all second-order algorithms known by the authors only guarantee convergence to points satisfying WSOC, see [2, 23, 25, 26, 29, 31, 33, 37, 52] and references therein.

This situation in which a most desirable theoretical property is not suitable in an algorithmic framework is not particular only to the second-order case. Even in the first-order case, it is known, for example, that the Guignard constraint qualification is the weakest possible assumption to yield KKT conditions at a local minimizer [38]. In other words, a good first-order necessary optimality condition is of the form “KKT or not-Guignard”. But this is too strong for practical purposes, since no algorithm is known to fulfill such condition at limit points of sequences generated by it, in fact, the convergence assumptions of algorithms require stronger constraint qualifications [5, 6, 8, 9]. For second-order algorithms, the situation is quite similar, with the peculiarity that the difficulty is not only on the required constraint qualification, but also in the verification of the optimality condition, since, numerically, we can only guarantee a partial second-order property, that is, for directions in the critical subspace, which is a subset of the desirable critical cone of directions.

As the KKT conditions, SSOC and WSOC hold at minimizers only if some
additional condition is valid. As we will explore in the next section, only MFCQ is not enough to ensure the existence of some Lagrange multiplier where SSOC holds. Even WSOC can not be assured to hold under MFCQ alone. Under MFCQ alone, we have the following result, [22, 16]:

**Definition 2.2.** Let $x^*$ be a feasible point. We say that the basic second-order optimality condition (BSOC) holds at $x^*$ if (i) the set of multipliers $\Lambda(x^*) \neq \emptyset$ and (ii)

$$\forall d \in C(x^*), \exists (\lambda, \mu) \in \Lambda(x^*) \text{ such that } d^T \nabla^2_{xx} L(x^*, \lambda, \mu)d \geq 0.$$  (2.6)

**Theorem 2.1.** Let $x^*$ be a local minimizer of (2.1) that satisfies MFCQ, Then $x^*$ satisfies BSOC.

Note that for each critical direction, we have an associated Lagrange multiplier $(\lambda, \mu) \in \Lambda(x^*)$, in opposition to SSOC or WSOC, where we require the same Lagrange multiplier for all critical directions. Observe that (2.6) does not imply WSOC (and neither SSOC).

From the proof of Theorem 2.1 (see, for instance, [22]), we can see that for each $d \in C(x^*)$, the Lagrange multiplier $(\lambda, \mu)$ constructed is such that $\mu_j = 0$ for all $j \in A(x^*)$ such that $\nabla g_j(x^*)^T d = 0$. This specification turns out to be relevant in our analysis, so we single out the following definition:

**Definition 2.3.** Given a feasible point $x^* \in \Omega$ and $d \in C(x^*)$, we say that the basic second-order optimality condition in the direction $d$ (BSOC($d$)) holds at $x^*$ if there is a Lagrange multiplier $(\lambda, \mu) \in \Lambda(x^*)$ such that $d^T \nabla^2_{xx} L(x^*, \lambda, \mu)d \geq 0$, where $\mu_j = 0$ for all $j \in A(x^*)$ with $\nabla g_j(x^*)^T d = 0$.

Observe that when $\Lambda(x^*)$ is a compact set (namely, when MFCQ holds), (2.6) can be written in the more compact form below:

$$\forall d \in C(x^*), \max\{d^T \nabla^2_{xx} L(x^*, \lambda, \mu)d : (\lambda, \mu) \in \Lambda(x^*)\} \geq 0.$$  

Even when no constraint qualification is assumed, a second-order optimality condition can be formulated, relying on Fritz John multipliers (2.2) (see [22]):

**Theorem 2.2.** Let $x^*$ be a local minimizer of (2.1). Then, for every $d$ in the critical cone $C(x^*)$, there is a Fritz John multiplier $(\lambda_0, \lambda, \mu) \in \Lambda_0(x^*)$ such that

$$d^T \nabla^2_{xx} L^0(x^*, \lambda_0, \lambda, \mu)d \geq 0.$$  (2.7)

Theorem 2.1 can be seen as a consequence of Theorem 2.2, as MFCQ is equivalent to the non existence of a Fritz-John multiplier of the form $(0, \lambda, \mu) \in \Lambda_0(x^*)$. However, we will see in the next section a standalone proof of Theorem 2.1 under a constraint qualification weaker than MFCQ.

The optimality condition of Theorem 2.2 has been studied a lot over the years, [32, 16, 46, 22, 11]. An important property is that it can be transformed into a sufficient optimality condition by simply replacing the non-negative sign “$\geq 0$” by “$> 0$” (except at the origin), without any additional assumption. For
this reason, this condition is said to be a “no-gap” optimality condition. Note that this is different from the case of SSOC, since an additional assumption must be made for the necessary condition to hold.

We emphasize that even though optimality conditions given by Theorems 2.1 and 2.2 have nice theoretical properties, they require the knowledge of the whole set of (generalized) Lagrange multipliers at the basis point, whereas in practice, we only have access to (an approximation of) a single Lagrange multiplier. In the case of the optimality condition given by Theorem 2.2, one could argue that the possibility of verifying it with \( \lambda_0 = 0 \), and hence independently of the objective function, is not useful at all as an optimality condition. This is arguably the case for the first-order Fritz John optimality condition, but since Theorem 2.2 gives a “no-gap” optimality condition, this argument is not convincing in the second-order case. In fact, one could show that if the sufficient optimality condition associated with Theorem 2.2 is fulfilled with \( \lambda_0 = 0 \) for all critical directions, then the basis point is an isolated feasible point, and hence a local solution independently of the objective function. We take the point of view that algorithms naturally treat differently the objective function and the constraint functions, in a way that a multiplier associated with the objective function is not present, hence our focus on Lagrange multipliers, rather than on Fritz John multipliers.

As we have mentioned, known practical methods are only guaranteed to converge to points satisfying WSOC, and hence, we focus our attention on conditions ensuring it at local minimizers.

We start with [14], where the authors investigate the issue of verifying (2.6) for the same Lagrange multiplier:

**Theorem 2.3** ([14]). Let \( x^* \) be a local minimizer of (2.1). Assume that MFCQ holds at \( x^* \) and that \( \Lambda(x^*) \) is a (bounded) line segment. Then, for every first-order cone \( K \subset C(x^*) \), there is a \( (\lambda^K, \mu^K) \in \Lambda(x^*) \) such that

\[
\forall d \in K, \quad d^T \nabla^2_{xx} L(x^*, \lambda^K, \mu^K) d \geq 0. 
\]  
(2.8)

We are mainly interested in the special case \( K := S(x^*) \). Thus, (2.8) holds at a local minimizer \( x^* \) when \( \Lambda(x^*) \) is a line segment and MFCQ holds at \( x^* \) (or, equivalently, \( \Lambda(x^*) \) is a bounded line segment). Note that in this case, (2.8) is equivalent to WSOC.

In order to prove Theorem 2.3 a crucial result is Yuan’s Lemma [61], which was generalized for first-order cones in [14]. For further applications of Yuan’s Lemma, see [47, 27].

**Lemma 2.4** (Yuan [61, 14]). Let \( P, Q \in \mathbb{R}^{n \times n} \) be two symmetric matrices and \( K \subset \mathbb{R}^n \) a first-order cone. Then the following conditions are equivalent:

- \( \max \{ d^T P d, d^T Q d \} \geq 0, \quad \forall d \in K; \)

- There exist \( \alpha \geq 0 \) and \( \beta \geq 0 \) with \( \alpha + \beta = 1 \) such that \( d^T (\alpha P + \beta Q) d \geq 0, \quad \forall d \in K. \)
A sufficient condition to guarantee that $\Lambda(x^*)$ is a line segment, is to require that the rank of the Jacobian matrix $J(x^*)$ is row-deficient by at most one, that is, the rank is one less than the number of rows. The fact that the rank assumption yields the one-dimensionality of $\Lambda(x^*)$ is a simple consequence of the rank-nullity theorem. Thus, we have the following result:

**Theorem 2.5** (Baccari and Trad [14]). Let $x^*$ be a local minimizer of (2.1) such that MFCQ holds and the rank of the Jacobian matrix $J(x^*) \in \mathbb{R}^{(m+q)\times n}$ is $m + q - 1$, where $q$ is the number of active inequality constraints at $x^*$. Then, there exists a Lagrange multiplier $(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p_+$ such that WSOC holds.

By a generalization of Yuan’s Lemma 2.4, Theorem 2.3 has been recently generalized in [42] by replacing the assumption that $\Lambda(x^*)$ is a line segment by the assumption that the set of matrices $\{\nabla^2_{xx}L(x^*, \lambda, \mu)\}$, where $(\lambda, \mu)$ ranges over the vertices of the bounded polyhedral set $\Lambda(x^*)$, is a set of rank at most 2 when viewed as a subset of the vector space $\text{Sym}(n)$.

Another line of reasoning in order to arrive at second-order optimality conditions is to use Janin’s version of the classical Constant Rank theorem ([60], Theorem 2.9). See [44, 3, 50].

**Theorem 2.6** (Constant Rank). Let $x^* \in \Omega$ and $d \in C(x^*)$. Let $E \subset \{1, \ldots, p\}$ be the set of indices $j$ such that $\nabla g_j(x^*)^T d = 0, j \in A(x^*)$. If $\{\nabla h_i(x), i \in E; \nabla g_j(x), j \in E\}$ has constant rank in a neighbourhood of $x^*$, then, there are $\varepsilon > 0$ and a twice continuously differentiable function $\xi : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ such that $\xi(0) = x^*, \xi'(0) = d, h_i(\xi(t)) = 0, i \in E; g_j(\xi(t)) = 0, j \in E$ for $t \in (-\varepsilon, \varepsilon)$ and $g(\xi(t)) \leq 0$ for $t \in [0, \varepsilon]$.

The proof that the function $\xi$ is twice continuously differentiable was done in [50]. Using a variation of the constant rank theorem jointly with MFCQ, in [7], Andreani, Martínez and Schuverdt have proved the existence of multipliers satisfying WSOC at a local minimizer. This joint condition was also used in the convergence analysis of a second-order augmented Lagrangian method.

**Theorem 2.7** (Andreani, Martínez and Schuverdt [7]). Let $x^*$ be a local minimizer of (2.1) with MFCQ holding at $x^*$. Assume the Weak Constant Rank (WCR) condition holds, namely, that the rank of the Jacobian matrix $J(x) \in \mathbb{R}^{(m+q)\times n}$ is constant around $x^*$, where $q$ is the number of active inequality constraints at $x^*$. Then, WSOC holds at $x^*$.

The proof is done using Theorem 2.6 for $d \in S(x^*)$ and $E = \{1, \ldots, p\}$, using the fact that $t = 0$ is a local minimizer of $f(\xi(t)), t \geq 0$.

This result was further improved in [3], where they noticed that MFCQ can be replaced by the non-emptiness of $\Lambda(x^*)$. This was also done independently in [40]. In fact, WSOC can be proved to hold for all Lagrange multipliers:

**Theorem 2.8** (Andreani, Echagüe and Schuverdt [3]). Let $x^*$ be a local minimizer of (2.1) such that the rank of the Jacobian matrix $J(x) \in \mathbb{R}^{(m+q)\times n}$ is constant around $x^*$, where $q$ is the number of active inequality constraints at $x^*$. Then, every Lagrange multiplier $(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p_+$ (if any exists) is such that WSOC holds.
This same technique can be employed under the Relaxed Constant Rank CQ (RCRCQ, [49]), that is,
\[ \nabla h_i(x), i \in E; \nabla g_j(x), j \in E \]
has constant rank around \( x^* \) for every \( E \subset A(x^*) \), to prove the stronger result that all Lagrange multipliers satisfy SSOC. See [3, 50]. These results can be strengthened by replacing the use of the Constant Rank theorem by the assumption that the critical cone is a subset of the Tangent cone of a modified feasible set (Abadie-type assumptions). See details in [1, 20].

In \( \mathbb{R}^3 \), the remarkable example by Arutyunov [11]/Anitescu [10] shows that if the rank increases by more than two around \( x^* \), WSOC may fail for all Lagrange multipliers (of course, SSOC also fails).

We describe below a modification of the example, given in [13].

**Example 2.1.**

Minimize \( x_3 \),
\[
\begin{align*}
x_3 &\geq 2\sqrt{3}x_1x_2 - 2x_2^2, \\
x_3 &\geq x_2^2 - 3x_1^2, \\
x_3 &\geq -2\sqrt{3}x_1x_2 - 2x_2^2.
\end{align*}
\]

Here \( x^* := (0, 0, 0) \) is a global minimizer. The critical subspace is the whole plane \( x_3 = 0 \) and \( \Lambda(x^*) \) is the simplex defined by \( \mu_1 + \mu_2 + \mu_3 = 1, \mu_j \geq 0, j = 1, 2, 3 \). Figure 1 shows the graph of the right-hand side of each constraint, where the feasible set is the set of points above all surfaces. Note that along every direction in the critical cone, there is a convex combination of the constraints that moves upwards and (2.6) holds, but for any convex combinations of the constraints, there exists a direction in the critical cone that moves downwards. This means that WSOC fails for all Lagrange multipliers (note that WSOC is equivalent to SSOC in this example, as \( C(x^*) = S(x^*) \)). Note also in Figure 1 that around \( x^* \) there is no feasible curve such that all constraints are active along this curve, which is the main property allowing the proof of WSOC under constant rank assumptions (Theorem 2.6).

![Figure 1: MFCQ alone is not enough to ensure the validity of SSOC or WSOC.](image-url)
To conclude this section we note that when the rank is constant, the Hessian of the Lagrangian does not depend on the Lagrange multiplier. This explains why results under constant rank conditions hold for all Lagrange multipliers.

**Theorem 2.9.** Suppose that $\Lambda(x^*) \neq \emptyset$. If the rank of $J(x)$ is constant around a point $x^*$, then the quadratic form $d^2 \nabla^2_x L(x^*, \lambda, \mu)d$ for $d \in S(x^*)$ does not depend on $(\lambda, \mu) \in \Lambda(x^*)$.

**Proof.** By simplicity, assume $m = 0$ and $A(x^*) = \{1, \ldots, p\}$. By Theorem 2.6, for each $d \in S(x^*)$, there exists a smooth curve $\xi(t)$, $t \in (-\varepsilon, \varepsilon)$ with $g(\xi(t)) = 0$ for all $t$, with $\xi(0) = x^*$ and $\xi'(0) = d$. Take $\bar{\mu} \in \text{Ker}(J(x^*)^T)$ and let us define the function $R(t) := \sum_{i=1}^{p} \bar{\mu}_i g_i(\xi(t))$, which is constantly zero for small $t$. Straightforward calculations show that $R''(0) = d^T \sum_{i=1}^{p} \bar{\mu}_i \nabla^2 g_i(x^*)d + \xi''(0)^T J(x^*)^T \bar{\mu} = 0$. Hence, $d^T \sum_{i=1}^{p} \mu_i \nabla^2 g_i(x^*)d = 0$.

But $\Lambda(x^*) = (\bar{\mu} + \text{Ker}(J(x^*)^T)) \cap \mathbb{R}^p_+$ for a fixed Lagrange multiplier $\bar{\mu} \in \Lambda(x^*)$. Hence $\mu \in \Lambda(x^*)$ if, and only if, $\mu = \bar{\mu} + \hat{\mu}$, for some $\hat{\mu} \in \text{Ker}(J(x^*)^T)$, with $\bar{\mu} + \hat{\mu} \geq 0$. It follows that $d^T \nabla^2_{xx} L(x^*, \mu)d = d^T \nabla^2_{xx} L(x^*, \bar{\mu})d$, as we wanted to show. Observe that $x^*$ is not necessarily a local minimizer, we only require $\Lambda(x^*) \neq \emptyset$. \qed

### 3 Second-order Constraint Qualifications

Let us recall that a second-order constraint qualification may involve the second-order derivatives on the formulation of the condition.

It is well known that the Guignard constraint qualification is the weakest condition for differentiable data, independent of the objective function, that implies the existence of Lagrange multipliers at a local minimizer. The main goal of this section is to define the weakest possible constraint qualification that implies not only the existence of Lagrange multipliers, but also Lagrange multipliers connected to second-order optimality conditions.

It is natural to consider second-order constraint qualifications for this task, namely, one should involve the Hessians of the constraints on the formulation of the constraint qualification. Second-order constraint qualifications have been considered in [4, 41] to prove global convergence of second-order algorithms to a point satisfying WSOC under conditions weaker than RCRCQ or the joint condition MFCQ plus the rank assumption of Theorem 2.7.

In order to formulate our results, let us review the minimality property of the Guignard constraint qualification. The first proof of this result, due to Gould and Tolle [38], was quite involving, but one can more easily understand the result with tools of variational analysis, which we define below. First, let us recall some definitions. Given a set-valued mapping (multifunction) $\Gamma: \mathbb{R}^s \rightrightarrows \mathbb{R}^d$, the **sequential Painlevé-Kuratowski outer limit** of $\Gamma(z)$ as $z \to z^*$ is defined by

$$\limsup_{z \to z^*} \Gamma(z) := \{w^* \in \mathbb{R}^d : \exists (z^k, w^k) \to (z^*, w^*) \text{ with } w^k \in \Gamma(z^k)\}. \quad (3.1)$$

We say that $\Gamma$ is **outer semicontinuous** (osc) at $z^*$ if $\limsup_{z \to z^*} \Gamma(z) \subseteq \Gamma(z^*)$.  


Given $S \subseteq \mathbb{R}^n$ and $z^* \in S$, the tangent cone to $S$ at $z^*$ is
\[
T_S(z^*) := \{ d \in \mathbb{R}^n \mid \text{dist}(x + t_n d, S) = o(t_n) \text{ for some } t_n \to 0^+ \} \tag{3.2}
\]
and the regular normal cone to $S$ at $z^* \in S$ is
\[
\hat{N}_S(z^*):= \{ w \in \mathbb{R}^n : \limsup_{z \in \Omega, z \to z^*} \frac{\langle w, z - z^* \rangle}{\|z - z^*\|} \leq 0 \}. \tag{3.3}
\]
The regular normal cone has the following remarkable properties. See [59].

**Theorem 3.1.** Given a non-empty set $S \subseteq \mathbb{R}^n$ and $z^* \in S$, the following conditions are equivalent:

i) $w \in \hat{N}_S(z^*)$;

ii) There is a continuously differentiable function $f$ that achieves its global minimum relative to $S$ at $z^*$ such that $-\nabla f(z^*) = w$;

iii) $w \in T^*_S(z^*)$.

When the set $S$ under consideration is the feasible set $\Omega$ of (2.1), for $x^* \in \Omega$, we use the notation $\hat{N}_1(x^*) := \hat{N}_\Omega(x^*)$ and $T_1(x^*) := T_{\Omega}(x^*)$. In that case, we also define the first-order linearized cone to $\Omega$ at $x^* \in \Omega$ as
\[
L_1(x^*) := \{ d \in \mathbb{R}^n : \nabla h_i(x^*)^T d, i \in E; \nabla g_i(x^*)^T d \leq 0, \quad i \in A(x^*) \}.
\]
Clearly, $T_1(x^*) \subseteq L_1(x^*)$ always holds, while the reciprocal inclusion may not hold. It is easy to see that the KKT conditions hold at $x^*$ for the objective function $f$ with respect to the constraint set $\Omega$ if, and only if, $-\nabla f(x^*) \in L_1(x^*)^\circ$, where the polar cone of $L_1(x^*)$ is given by
\[
L_1^\circ(x^*) = \{ w \in \mathbb{R}^n : w = \sum_{i \in E} \lambda_i \nabla h_i(x^*) + \sum_{i \in A(x^*)} \mu_i \nabla g_i(x^*), \lambda_i \in \mathbb{R}, \mu_i \geq 0 \}.
\]

Therefore, from Theorem 3.1, the assumption $\hat{N}_1(x^*) = L_1^\circ(x^*)$, which is known as the Guignard constraint qualification, is equivalent to the property that $x^*$ is a KKT point for every continuously differentiable objective function $f$ such that $x^*$ is a local minimizer of $f$ over $\Omega$. That is, the Guignard constraint qualification is the weakest condition independent of the objective function ensuring the validity of KKT at a local minimizer. Due to Theorem 3.1 again, Guignard constraint qualification can be stated more usually as $T^*_1(x^*) = L^\circ_1(x^*)$, which gives rise to the more well known (and stronger) Abadie constraint qualification, namely, $T_1(x^*) = L_1(x^*)$.

We will define the second-order analogues of these objects in order to propose a similar minimal CQ with second-order properties. The second-order tangent point-to-set mapping $T_2(x, \cdot) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ of $\Omega$ at $x \in \Omega$ is defined by the function $d \in \mathbb{R}^n \mapsto T_2(x, d)$ given by:
\[
T_2(x, d) := \{ \bar{z} \in \mathbb{R}^n \mid \text{dist}(x + t_n d + \frac{1}{2}t_n^2 \bar{z}, \Omega) = o(t_n^2) \text{ for some } t_n \to 0^+ \}.
\]
Note that $T_2(x,d) = \emptyset$ if $d \notin T_1(x)$. The second-order geometric optimality condition is given by [58], which we include a proof due to its simplicity:

**Theorem 3.2.** Let $x^* \in \Omega$ be a local minimizer of $f$ over $\Omega$. Then, for every $d \in T_1(x^*)$ such that $\nabla f(x^*)^T d = 0$ we have $\nabla f(x^*)^T z + \nabla^2 f(x^*) d^2 \geq 0$, for all $z \in T_2(x^*,d)$.

**Proof.** Let $d \in T_1(x^*)$ with $\nabla f(x^*)^T d = 0$ and $z \in T_2(x^*,d)$. By the definition of $T_2(x^*,d)$, there are sequences $\{x^k\} \subset \Omega$ and $t_k \to 0^+$ such that $x^k = x^* + t_k d + \frac{1}{2} t_k^2 z + o(t_k^2)$, hence, the Taylor expansion yields:

$$f(x^k) = f(x^*) + t_k \nabla f(x^*)^T d + \frac{1}{2} t_k^2 (\nabla f(x^*)^T z + \nabla^2 f(x^*) d^2) + o(t_k^2).$$

The result follows dividing by $t_k^2$ and taking the limit in $k$, as $x^k \to x^*$, $o(t_k^2)/t_k^2 \to 0$, $\nabla f(x^*)^T d = 0$ and $f(x^*) \leq f(x^k)$ for sufficiently large $k$. □

Theorem 3.2 is a second-order analogue of the first-order geometric optimality condition (namely, the implication $ii) \Rightarrow iii)$ in Theorem 3.1). In fact, since $T_2(x^*,0) = T_1(x^*)$, the particular case $d = 0$ in Theorem 3.2 yields the first-order geometric optimality condition. Now, let us define the linearization of the second-order tangent point-to-set mapping as

$$L_2(x,d) := \left\{ z \in \mathbb{R}^n : \nabla h_i(x)^T z + \nabla^2 h_i(x) d^2 = 0, \quad i \in \mathcal{E}, \right. \nabla g_i(x)^T z + \nabla^2 g_i(x) d^2 \leq 0, \quad i \in A(x,d) \right\},$$

where $A(x,d) := \{ i \in A(x) : \nabla g_i(x)^T d = 0 \}$. Clearly, for $d \in T_1(x)$, we have $T_2(x,d) \subset L_2(x,d)$. If for every $d \in C(x) = L_1(x) \cap \{ d : \nabla f(x)^T d = 0 \}$ we impose that $T_2(x,d) = L_2(x,d)$ (what is called the second-order Abadie constraint qualification [45, 43]) we have that BSOC holds whenever $x$ is a local minimizer. In fact, this assumption can be replaced by the weaker equality of the polars $T_2^e(x,d) = L_2^e(x,d)$ for every $d \in C(x)$, which was called the second-order Guignard constraint qualification [45]. In [16], the following constraint qualification, which can be seen as a second-order version of MFCQ, is introduced: $\{ \nabla h_i(x) \}_{i=1}^m$ is linearly independent and for some $d \in C(x)$, there exists $z \in \mathbb{R}^n$ such that $\nabla h_i(x)^T z + \nabla^2 h_i(x) d^2 = 0, i \in \mathcal{E}$ and $\nabla g_i(x)^T z + \nabla^2 g_i(x) d^2 < 0, i \in A(x,d)$. We note that the second-order versions of Abadie and Guignard are in fact stronger than their first order counterparts, while the second-order MFCQ is weaker than MFCQ, and still implies BSOC for $d$ in the critical cone.

Since the second-order tangent cone $T_2(x,d)$ can be an empty set for certain directions $d \in T_1(x)$, following [57], we consider the projective second-order tangent cone of $\Omega$ at $x^*$ defined by:

$$T_{2,p}(x^*,d) := \left\{ (z,r) \in \mathbb{R}^n \times \mathbb{R}_+ : \exists (z^k,r_k,t_k) \in \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{R}_+ \text{ such that } \begin{align*} (z^k,r_k,t_k) &\to (z,r,0), \quad t_k/r_k \to 0, \quad r_k > 0 \\
&\text{and } x^* + t_k d + \frac{t_k^2}{2r_k} z^k \in \Omega, \text{ for all } k \in \mathbb{N} \end{align*} \right\}.$$
From the definition, \( T_{2,p}(x^*, d) \) is a closed cone which is always a non-empty set. Furthermore, it is not difficult to see that if \( (z, r) \in T_{2,p}(x^*, d) \) for some \( r > 0 \), then \( z \in T_2(x^*, d) \), and if \( z \in T_2(x^*, d) \) then \( (z, 1) \in T_{2,p}(x^*, d) \). Since \( T_{2,p}(x^*, d) \) can be a difficult object to deal with, we define the \textit{projective linearized second-order tangent cone} of \( \Omega \) at \( x^* \) as

\[
L_{2,p}(x^*, d) := \left\{ (z, r) \in \mathbb{R}^n \times \mathbb{R}_+ : \begin{array}{l}
\nabla h_i(x^*)^T z + r \nabla^2 h_i(x^*) d^2 \leq 0, \quad i \in \mathcal{E}, \\
\nabla g_i(x^*)^T z + r \nabla^2 g_i(x^*) d^2 \leq 0, \quad i \in A(x^*, d) \end{array} \right\}.
\]

(3.5)

Clearly, for every \( d \in T_1(x^*) \) we get \( T_{2,p}(x^*, d) \subset L_{2,p}(x^*, d) \). Since \( L_{2,p}(x^*, d) \) is defined by linear equalities and inequalities, the polar of \( L_{2,p}(x^*, d) \) can be explicitly calculated in the following way:

\[
L_{2,p}^*(x^*, d) = \left\{ \sum_{i \in \mathcal{E}} \lambda_i \left( \frac{\nabla h_i(x^*)}{\nabla^2 h_i(x^*) d} \right) + \sum_{j \in A(x^*, d)} \mu_j \left( \frac{\nabla g_j(x^*)}{\nabla^2 g_j(x^*) d} \right) - \frac{\beta}{2} : \begin{array}{l}
\lambda \in \mathbb{R}^m; \\
\mu_j \geq 0 \quad \forall j; \\
\beta \geq 0
\end{array} \right\}.
\]

(3.6)

Using the projective second-order tangent cone we have the following necessary optimality condition for twice continuously differentiable data.

**Theorem 3.3** (Penot [57]). Let \( x^* \in \Omega \) be a local minimizer of \( f \) over \( \Omega \). Then, for every \( d \in T_1(x^*) \) with \( \nabla f(x^*)^T d = 0 \) we have that \( \nabla f(x^*)^T z + r \nabla^2 f(x^*) d^2 \geq 0 \), for all \( (z, r) \in T_{2,p}(x^*, d) \).

Since our goal is to formulate the weakest possible condition to ensure BSOC (also WSOC and SSOC), we define the second-order normal cone to \( \Omega \) at \( x^* \in \Omega \) as follows:

\[
\hat{N}_2(x^*) := \{(w, W) \in \mathbb{R}^n \times \text{Sym}(n) : \limsup_{x \in \Omega, x \to x^*} \frac{\langle w, x - x^* \rangle + (1/2)W(x - x^*)^2}{\|x - x^*\|^2} \leq 0\}.
\]

(3.7)

In some sense, an element of the second-order normal cone stands for a vector-matrix pair that plays the role of a gradient-hessian pair in a vanishing Taylor-like expansion in \( \Omega \). The following result makes this formulation precise, and it is a second-order version of Theorem 3.1:

**Theorem 4.** Let \( x^* \in \Omega \), \( w \in \mathbb{R}^n \) and \( W \in \text{Sym}(n) \). Then, the following statements are equivalent:

(i) \( (w, W) \) belongs to \( \hat{N}_2(x^*) \);

(ii) There is a twice continuously differentiable function \( f \) that attains its global minimum relative to \( \Omega \) at the point \( x^* \) such that \( -\nabla f(x^*) = w \) and \( -\nabla^2 f(x^*) = W \).

**Proof.** Let us prove the implication \((i) \Rightarrow (ii)\). Take \((w, W) \in \hat{N}_2(x^*)\) an arbitrary element. Denote

\[
\eta_0(t) := \sup \{ \langle w, x - x^* \rangle + (1/2)W(x - x^*)^2 : \|x - x^*\| \leq t, \; x \in \Omega \} \text{ for } t \geq 0.
\]

(3.8)
Clearly, \( \eta_0(t) \) is non-decreasing on \([0, \infty)\) and \( 0 = \eta_0(0) \leq \eta_0(t) \). From the definition of \( \tilde{N}_2(x^*) \), \( |\eta_0(t)| \leq Mt^2 \) for all \( t \leq T \), where \( M \) and \( T \) are positive scalars. Modify \( \eta_0(t) \) outside \([0, T)\) in such a way that \( |\eta_0(t)| \leq Mt^2 \) over \([0, \infty)\).

Since \( \eta_0(t) \) satisfies all the hypotheses of Corollary A.2 from the Appendix, there is a twice continuously differentiable convex non-decreasing function \( \phi : [0, \infty) \to [0, \infty) \) such that \( \phi(0) = \phi_\epsilon'(0) = \phi_\epsilon''(0) = 0 \) and \( \phi(t) > \eta_0(t) \) for all \( t > 0 \). Define \( F(x) := \langle w, x-x^* \rangle + (1/2)W(x-x^*)^2 - \phi(\|x-x^*\|) \). Observe that \( F(x) \) is a twice differentiable function at any \( x \neq x^* \) due to the smoothness of \( \phi \) and the norm \( \| \cdot \| \) at non-zero points. Now, by the classical chain rule and the smoothness of \( \phi \), we have \( \nabla F(x^*) = w \) and \( \nabla^2 F(x^*) = W \). Since \( F(x) < F(x^*) \) for all \( x \in \Omega \), we have that \( F(x) \) achieves its global maximum over \( \Omega \) uniquely at \( x^* \). The other implication is a simple consequence of the Taylor formulae.

Now, we will rewrite the geometric second-order optimality condition using new geometric objects. Given a set \( A \subset \mathbb{R}^n \) and a point-to-set mapping \( B(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}_+, d \rightarrow B(d) \), we define the bracket of \( A \) and \( B(\cdot) \) (denoted \([A, B(\cdot)]\)) as

\[
[A, B(\cdot)] := \left\{ (w, W) \in \mathbb{R}^n \times \text{Sym}(n) : \begin{array}{ll}
\langle w, d \rangle \leq 0, & \text{for all } d \in A \\
\langle w, z \rangle + \langle W, rdd^T \rangle \leq 0, & \text{for all } (d, (z, r)) \in K_A(w) \times B(d)
\end{array} \right\},
\]

where \( K_A(v) := A \cap v^\perp \) for \( v \in \mathbb{R}^n \). Note that in the case \( A := L_1(x) \), the set \( K_{L_1(x)}(\nabla f(x)) \) coincides with the critical cone \( C(x) \). Using the projective second-order normal cone and the bracket, the second-order geometric optimality condition obtained in Theorem 3.3 can be rewritten as the inclusion

\[
\tilde{N}_2(x^*) \subset [T_1(x^*), T_{2, p}(x^*, \cdot)].
\]

In the next subsections, we will use \( \tilde{N}_2(x^*) \) to associate minimal CQs associated with the conditions WSOC, SSOC and BSOC.

Note that the bracket operator generalizes the polar operator in the sense that, given \( A \subset \mathbb{R}^n \), \( \{0\}, A(\cdot) = A^\circ \times \text{Sym}(n) \), where \( A(\cdot) \equiv (A, 0) \). Note also that \([A, B(\cdot)]\) is a closed convex cone regardless of \( A \) and \( B(\cdot) \).

The bracket operator unifies the geometric optimality condition, when it is applied to the tangent objects \((T_1(x^*), T_{2, p}(x^*, \cdot))\), while when applied to the linearized objects \((L_1(x^*), L_{2, p}(x^*, \cdot))\), the bracket operator gives rise to BSOC, as we discuss in the next subsection. In the first-order case, this role is played by the polar operator, as \(-\nabla f(x^*) \in T_1^c(x^*)\) gives the geometric optimality condition, while the KKT conditions are represented by \(-\nabla f(x^*) \in L_1^c(x^*)\). In this case, the (first-order) normal cone coincides with the polar of the tangent cone, which gives a nice geometric interpretation for the Guignard CQ. The situation is slightly less favorable in the second-order case, as a characterization of the second-order normal cone \( \tilde{N}_2(x^*) \) is not known.
3.1 Minimal CQ for BSOC

Here we introduce minimal CQs for the different second-order conditions which can be defined in terms of the bracket operator. From the obvious inclusions $T_1(x^*) \subset L_1(x^*)$ and $T_{2,p}(x^*, d) \subset L_{2,p}(x^*, d)$ for all $d \in T_1(x^*)$, we have

$$[L_1(x^*), L_{2,p}(x^*, \cdot)] \subset [T_1(x^*), T_{2,p}(x^*, \cdot)].$$

Now, take a $C^2$ function $f$ such that $(-\nabla f(x^*), -\nabla^2 f(x^*)) \in [L_1(x^*), L_{2,p}(x^*, \cdot)]$. Then, from the definition we have that (i) the set of Lagrange multipliers $\Lambda(x^*)$ associated with $f$ is non-empty and (ii) for every $d \in L_1(x^*) \cap \{\nabla f(x^*)\}^\perp$, there is a Lagrange multiplier $(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p_+$ with $\mu_i = 0$, for $j \notin A(x, d)$ such that $\nabla f(x^*) + \sum_{i \in E} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^p \mu_j \nabla g_j(x^*) = 0$ and

$$\nabla^2 f(x^*)d^2 + \sum_{i \in E} \lambda_i \nabla^2 h_i(x^*)d^2 + \sum_{j=1}^p \mu_j \nabla^2 g_j(x^*)d^2 \geq 0.$$

That is, BSOC holds. Furthermore, we can see that BSOC holds at the feasible point $x^*$ for the objective function $f$ if, and only if $(-\nabla f(x^*), -\nabla^2 f(x^*)) \in [L_1(x^*), L_{2,p}(x^*, \cdot)]$. Thus, the minimal CQ that ensures that BSOC holds at $x^*$, for every $C^2$ objective function $f$ which has $x^*$ as a local minimizer constrained to $\Omega$, in view of Theorem 3.4, is

$$\tilde{N}_2(x^*) \subset [L_1(x^*), L_{2,p}(x^*, \cdot)].$$

Using the bracket and the inclusion $\tilde{N}_2(x^*) \subset [T_1(x^*), T_{2,p}(x^*, \cdot)]$ given by Theorem 3.3, a weak CQ that ensures BSOC at local minimizers is

$$[T_1(x^*), T_{2,p}(x^*, \cdot)] = [L_1(x^*), L_{2,p}(x^*, \cdot)],$$

which we call second-order Guignard CQ for BSOC. This gives rise to two new weak CQs to ensure BSOC, namely, another Guignard-type CQ, given by

$$T_1(x^*) = L_1(x^*) \text{ and } T_{2,p}(x^*, d) = L_{2,p}(x^*, d) \text{ for all } d \in \mathbb{R}^n,$$

and an Abadie-type one given by

$$T_1(x^*) = L_1(x^*) \text{ and } T_{2,p}(x^*, d) = L_{2,p}(x^*, d) \text{ for every } d \in \mathbb{R}^n.$$

Certainly, the last two conditions imply $[T_1(x^*), T_{2,p}(x^*, \cdot)] = [L_1(x^*), L_{2,p}(x^*, \cdot)]$. A well-known CQ for BSOC is the second-order Abadie condition introduced by Kawasaki [45], which states the equality $T_2(x^*, d) = L_2(x^*, d)$ for every direction $d \in T_1(x^*)$. This assumption implicitly assumes that $T_2(x^*, d) \neq \emptyset$ for every direction, which is not true in general. Under this assumption, we have that $T_2(x^*, d) = L_2(x^*, d)$ implies equality between the sets $[T_1(x^*), T_{2,p}(x^*, \cdot)]$ and $[L_1(x^*), L_{2,p}(x^*, \cdot)]$, as the next proposition will show.

**Proposition 3.5.** Assume that $T_2(x^*, d) \neq \emptyset$ and $T_2(x^*, d) = L_2(x^*, d)$, for all $d \in T_1(x^*)$. Then, $[T_1(x^*), T_{2,p}(x^*, \cdot)] = [L_1(x^*), L_{2,p}(x^*, \cdot)]$. 

15
Proof. Take \((w,W) \in [T_1(x^*), T_{2,p}(x^*,\cdot)]\). Since \(T_2(x^*,d) = L_2(x^*,d)\) holds for \(d = 0\), we have \(T_1(x^*) = L_1(x^*)\). Thus, \(w \in L_1(x^*)^\circ\) and \(w \in T_1(x^*)^\circ\) by the definition of the bracket. Take \(d \in K_A(w)\). By \(T_2(x^*,d) = L_2(x^*,d)\) and using the definition of the bracket, we have that \(\langle w, z \rangle + \langle W, dd^T \rangle \leq 0\), for all \(z \in L_2(x^*,d)\). From strong duality for linear programming (since \(L_2(x^*,d)\) is defined by affine constraints), there exist multipliers \((\hat{\lambda}, \hat{\mu}) \in \mathbb{R}^m \times \mathbb{R}_+^p\) with \(\mu_j = 0\), for \(j \notin A(x^*,d)\) such that

\[-w + \sum_{i=1}^m \hat{\lambda}_i \nabla h_i(x^*) + \sum_{j=1}^p \hat{\mu}_j \nabla g_j(x^*) = 0\]

and

\[-W d^2 + \sum_{i=1}^m \hat{\lambda}_i \nabla^2 h_i(x^*) d^2 + \sum_{j=1}^p \hat{\mu}_j \nabla^2 g_j(x^*) d^2 \geq 0\].

Then, using (3.6), we see that \((w,Wd^2) \in L_{2,p}(x^*,d)\), that is, \(\langle w, z \rangle + \langle W, rdd^T \rangle \leq 0\), for all \((z,r) \in L_{2,p}(x^*,d)\) and thus \((w,W) \in [L_1(x^*), L_{2,p}(x^*,\cdot)]\).

Now, instead of considering conditions that jointly take into account all directions \(d\) in the critical cone, let us fix a direction \(d \in T_1(x^*)\) and let us develop weak conditions to ensure the validity of BSOC\((d)\). The second-order regular normal cone of \(\Omega\) in the direction \(d\) at \(x^*\) is defined as:

\[\tilde{N}_2(x^*,d) := \{ (w, d^T W d) \in \mathbb{R}^n \times \mathbb{R} : (w,W) \in \tilde{N}_2(x^*) \text{ and } w \perp d \}. \tag{3.10}\]

Loosely speaking, \(\tilde{N}_2(x^*,d)\) represents the set of all \(C^2\) functions such that \(x^*\) is a local minimizer relative to \(\Omega\) that have the direction \(d\) as a critical direction. It is not difficult to see that \(\tilde{N}_2(x^*,d)\) is a non-empty convex cone in \(\mathbb{R}^n \times \mathbb{R}\).

Using the directional second-order normal cone, the second implication in the geometric optimality condition given by Theorem 3.3 can be rewritten as the inclusion

\[\tilde{N}_2(x^*,d) \subset T_{2,p}(x^*,d), \text{ for each } d \in T_1(x^*).\]

Using the polar cone of \(L_{2,p}(x^*,d)\), we see that \((-\nabla f(x^*), -\nabla^2 f(x^*) d^2) \in L_{2,p}(x^*,d)\) if, and only if, there are multipliers \((\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}_+^p\) with \(\mu_j = 0\), for \(j \notin A(x,d)\) such that

\[\nabla L(x^*, \lambda, \mu) = 0 \text{ and } d^T \nabla^2_{xx} L(x^*, \lambda, \mu) d \geq 0.\]

Hence, \((-\nabla f(x^*), -\nabla^2 f(x^*) d^2) \in L_{2,p}^*(x^*,d)\) if, and only if, BSOC\((d)\) holds at \(x^*\). With this tool, the minimal condition to guarantee that BSOC\((d)\) holds for every \(C^2\) function which has \(x^*\) as a local minimizer relative to \(\Omega\) and \(d\) as critical direction, is \(\tilde{N}_2(x^*,d) \subset L_{2,p}^*(x^*,d)\). Using the geometric second-order optimality conditions, we see that weak CQs that guarantee that BSOC\((d)\) is an optimality condition when \(d\) is a critical direction, are \(T_{2,p}^*(x^*,d) = L_{2,p}^*(x^*,d)\) (Guignard-type CQ) and \(T_{2,p}(x^*,d) = L_{2,p}(x^*,d)\) (Abadie-type CQ).

### 3.2 Minimal CQ for WSOC

Here we define a minimal CQ to guarantee that WSOC holds at any local minimizer. For this purpose, we consider the following linearized cone related
with WSOC.

\[ L_{2,W}(x^*) := \left\{ (z, Z) \in \mathbb{R}^n \times \text{Sym}_+(n) : \begin{array}{l}
\nabla h_i(x^*)^T z + (\nabla^2 h_i(x^*), Z) = 0, i = 1 \in \mathcal{E}, \\
\n\nabla g_j(x^*)^T z + (\nabla^2 g_j(x^*), Z) \leq 0, i \in A(x^*), \\
Z \in \text{conv}\{dd^T : d \in S(x^*)\}. 
\end{array} \right\}, \tag{3.11} \]

Computing the polar cone associated with \( L_{2,W}(x^*) \) we obtain

\[ L^0_{2,W}(x^*) = \left\{ \begin{array}{l}
\sum_{i \in \mathcal{E}} \lambda_i \left( \nabla h_i(x^*) + \nabla^2 h_i(x^*) \right) + \sum_{j=1}^q \mu_j \left( \nabla g_j(x^*) \right) - \left( 0, D \right) : \\
\lambda_i \in \mathbb{R}, i \in \mathcal{E}, \\
\mu_j \in \mathbb{R}^+, j = 1, \ldots, q, \\
\mu_j = 0, \forall j \notin A(x^*), \\
d^T D d \geq 0 \quad \forall d \in S(x^*). 
\end{array} \right\}, \tag{3.12} \]

Using this cone, WSOC can be equivalently stated as \((-\nabla f(x^*), -\nabla^2 f(x^*)) \in L^0_{2,W}(x^*)\). Using Theorem 3.4, we see that the weakest condition which ensures that WSOC holds at a local minimum is \( \hat{N}_2(x^*) \subset L^0_{2,W}(x^*) \). To introduce a Guigard-type condition for WSOC consider the set

\[ T_{2,W}(x^*) := \text{conv}\{ (z, rdd^T) : (z, r) \in T_{2,p}(x^*), d \in S(x^*) \}. \tag{3.13} \]

Clearly, \( T_{2,W}(x^*) \) is a non-empty closed convex cone and from the inclusions \( T_{2,p}(x^*, d) \subset L_{2,p}(x^*, d) \) for every \( d \in S(x^*) \), we get \( T_{2,W}(x^*) \subset L_{2,W}(x^*) \). Thus, \( L_{2,W}(x^*) \) can be considered as a linear approximation of \( T_{2,W}(x^*) \). We can infer a simple relationship between the polar of \( T^0_{2,W}(x^*) \) and the bracket \([T_1(x^*), T_{2,p}(x^*, \cdot)] \). In fact, the following proposition follows direct from the definition:

**Proposition 3.6.** It always holds that \( \hat{N}_2(x^*) \subset [T_1(x^*), T_{2,p}(x^*, \cdot)] \subset T^0_{2,W}(x^*) \).

Clearly, \( T^0_{2,W}(x^*) = L^0_{2,W}(x^*) \) is a Guigard-type weak CQ to guarantee WSOC at a local minimizer \( x^* \), while \( T_{2,W}(x^*) = L_{2,W}(x^*) \) is an Abadie-type CQ for WSOC. It is known that the condition MFCQ+WCR implies that WSOC holds at local minimizers (see Theorem 2.7). In fact, MFCQ+WCR implies a stronger conclusion.

**Theorem 3.7.** Let \( x^* \) be a feasible point. If MFCQ and WCR hold at \( x^* \), then \( T_{2,W}(x^*) = L_{2,W}(x^*) \).

**Proof.** Take \((z, Z) \in L_{2,W}(x^*)\). From the definition, there is a sequence \( \{Z^k\} \subset \text{conv}\{dd^T : d \in S(x^*)\} \) such that \( Z^k \to Z \). Without loss of generality we can assume \( Z^k = \sum_{\ell \in I(k)} d_{\ell k} d_{\ell k}^T \), with \( \{d_{\ell k} : \ell \in I(k)\} \subset S(x^*) \) and the cardinality of \( I(k) \) uniformly bounded, i.e., \(|I(k)| \leq M \) for some \( M \geq 0 \). Since \( \{Z^k\} \) is bounded, we have that \( \{d_{\ell k} : \ell \in I(k), k \in \mathbb{N}\} \) is also a bounded sequence.

For each \( k \in \mathbb{N} \), consider the optimization problem

\[
\text{Minimize } \frac{1}{2} \|z - \sum_{\ell \in I(k)} z_{\ell k}\|^2, \\
\text{s.t. } \nabla h_i(x^*)^T z_{\ell k} + d_{\ell k}^T \nabla^2 h_i(x^*) d_{\ell k} = 0, \quad i \in \mathcal{E}, \\
\nabla g_j(x^*)^T z_{\ell k} + d_{\ell k}^T \nabla^2 g_j(x^*) d_{\ell k} \leq 0, \quad j \in A(x^*), \quad \ell \in I(k). \tag{3.14} \]

\[ 17 \]
Denote by \( \{ z^k_\ell : \ell \in I(k) \} \) the solution of the problem above. By feasibility, \( z^k_\ell \in L_2(x^*, d_{ik}) \) for all \( \ell \in I(k) \). Since MFCQ holds at \( x^* \), we get that \( L_2(x^*, d_{ik}) = T_2(x^*, d_{ik}) \) and \( \sum_{\ell \in I(k)} (z^k_\ell, d_{ik}^k ) = (\sum_{\ell \in I(k)} z^k_\ell, Z^k) \in T_{2,W}(x^*) \). Furthermore, the sequence \( \{ \sum_{\ell \in I(k)} z^k_\ell : k \in \mathbb{N} \} \) is bounded. In fact, by WCR and following the proof of the Theorem 2.6 and [48], for each \( d_{ik} \), there is a \( C^2 \) curve \( \alpha_{ik}(t) \) such that \( \alpha_{ik}(0) = x^* \), \( \alpha'_{ik}(0) = d_{ik} \) and \( \hat{z}_{ik} := \alpha''_{ik}(0) \) with \( \hat{z}_{ik} \in L_2(x^*, d_{ik}) \). Furthermore, \( \hat{z}_{ik} \) is bounded by a constant that depends on the data and \( d_{ik} \). Since \( \{ d_{ik} : \ell \in I(k), k \in \mathbb{N} \} \) is bounded and \( I(k) \) is finite, the vector \( (\sum_{\ell \in I(k)} z^k_\ell : k \in \mathbb{N} \) is a feasible bounded sequence. From optimality \( \| z - \sum_{\ell \in I(k)} z^k_\ell \| \leq \| z - \sum_{\ell \in I(k)} \hat{z}_{ik} \| \) and thus we have that the sequence \( \{ \sum_{\ell \in I(k)} z^k_\ell : k \in \mathbb{N} \} \) is bounded as we wanted to show.

Thus, we can assume that \( \sum_{\ell \in I(k)} z^k_\ell \) converges to some vector \( \bar{z} \). By the KKT conditions, we have that there are multipliers \( (\lambda^k, \mu^k) \in \mathbb{R}^m \times \mathbb{R}^p \) with \( \mu_j^k = 0 \), for \( j \notin A(x^*) \) such that

\[
- \left( z - \sum_{\ell \in I(k)} z^k_\ell \right) + \sum_{i \in \mathcal{E}} \lambda^k_i \nabla h_i(x^*) + \sum_{j=1}^p \mu_j^k \nabla g_j(x^*) = 0.
\]

Furthermore, since MFCQ holds, the multipliers are bounded and we can assume that \( (\lambda^k, \mu^k) \to (\lambda, \mu) \). Thus, we have that

\[
\| z - \bar{z} \|^2 = \sum_{i \in \mathcal{E}} \lambda^*_i \nabla h_i(x^*)^T (z - \bar{z}) + \sum_{j=1}^p \mu^*_j \nabla g_j(x^*)^T (z - \bar{z}). \tag{3.15}
\]

Now, take \( j \in A(x^*) \). Since \( (z, Z) \in L_{2,W}(x^*) \), we have that \( \mu_j (\nabla g_j(x^*)^T z + \langle \nabla^2 g_j(x^*), Z \rangle) \leq 0 \), since \( \mu_j \geq 0 \). By the KKT conditions, we have that \( \mu_j^k (\nabla g_j(x^*)^T (\sum_{\ell \in I(k)} z^k_\ell + \langle \nabla^2 g_j(x^*), Z^k \rangle) = 0 \) which implies, taking limit, that \( \mu_j^k (\nabla g_j(x^*)^T (z + \langle \nabla^2 g_j(x^*), Z \rangle) = 0 \) and thus \( \mu_j^k \nabla g_j(x^*)^T (z - \bar{z}) \leq 0 \). Similarly, for \( i \in \mathcal{E} \), we get \( \lambda_i \nabla h_i(x^*)^T (z - \bar{z}) = 0 \). Using this expressions into (3.15), we get \( \bar{z} = z \). Thus, \( (z, Z) \in T_{2,W}(x^*) \). \( \square \)

### 3.3 Minimal CQ for SSOC

Now, let us define the minimal CQ to guarantee that SSOC holds at any local minimizer. For this, we first introduce the cone \( K^S_2(x^*) \), which is associated with SSOC:

\[
K^S_2(x^*) := \bigcup_{(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p, \mu_j = 0 \text{ for } j \notin A(x^*)} \left\{ (\sum_{i \in \mathcal{E}} \lambda_i \nabla h_i(x^*) + \sum_{j \in A(x^*)} \mu_j \nabla g_j(x^*), H) \text{ s.t. } H \leq \sum_{i \in \mathcal{E}} \lambda_i \nabla^2 h_i(x^*) + \sum_{j \in A(x^*)} \mu_j \nabla^2 g_j(x^*) \text{ on } C(x^*, \mu) \right\}, \tag{3.16}
\]

where \( C(x^*, \mu) \) is the critical cone associated with the vector \( \mu \) given by

\[
C(x^*, \mu) := \left\{ d \in \mathbb{R}^n : \begin{array}{l}
\langle \nabla h_i(x^*), d \rangle = 0, i \in \mathcal{E}, \\
\langle \nabla g_j(x^*), d \rangle = 0, \text{ if } \mu_j > 0, j \in A(x^*), \\
\langle \nabla g_j(x^*), d \rangle \leq 0, \text{ if } \mu_j^* = 0, j \in A(x^*).
\end{array} \tag{3.17}\right.
\]
When \( \mu \) is such that \( (\lambda, \mu) \) is a Lagrange multiplier associated with some objective function \( f \), the cone \( C(x^*, \mu) \) coincides with the critical cone \( C(x^*) = K_{L_1(x^*)}(\nabla f(x^*)) = L_1(x^*) \cap \nabla f(x^*)^\perp \).

Using \( K_2^S(x^*) \), the second-order condition SSOC holds at \( x^* \) if, and only if \((-\nabla f(x^*), -\nabla^2 f(x^*)) \) belongs to \( K_2^S(x^*) \). Using Theorem 3.4, we see that the weakest condition which ensures that SSOC holds at a local minimizer \( x^* \) is the inclusion \( \tilde{N}_2(x^*) \subset K_2^S(x^*) \). Let us introduce a Guignard-type condition for SSOC, for this, consider the following proposition, whose proof is a simple verification.

**Proposition 3.8.** Let \( x^* \) be a feasible point. We always have the inclusions: \( K_2^S(x^*) \subset [L_1(x^*), L_{2,p}(x^*, \cdot)] \subset [T_1(x^*), T_{2,p}(x^*, \cdot)] \).

Since \( \tilde{N}_2(x^*) \subset [T_1(x^*), T_{2,p}(x^*, \cdot)] \) always holds at every feasible point \( x^* \), the equality \( K_2^S(x^*) = [T_1(x^*), T_{2,p}(x^*, \cdot)] \) is a Guignard-type CQ for SSOC. From the proposition above we clearly have that \( K_2^S(x^*) = [T_1(x^*), T_{2,p}(x^*, \cdot)] \) implies \( [T_1(x^*), T_{2,p}(x^*, \cdot)] \). Thus, as expected, the Guignard-type condition for BSOC is weaker than the Guignard-type condition for SSOC. Now, it is well-known that under LICQ, every local minimizer satisfies SSOC [55, 17, 15]. In fact, let us prove the stronger result:

**Proposition 3.9.** Assume that \( x^* \) is a feasible point satisfying LICQ. Then, \( K_2^S(x^*) = [T_1(x^*), T_{2,p}(x^*, \cdot)] \). In particular, under LICQ, SSOC holds at any local minimizer.

*Proof.* From Proposition 3.8 we only need to show \([T_1(x^*), T_{2,p}(x^*, \cdot)] \subset K_2^S(x^*) \).

Take \((w, W) \in [T_1(x^*), T_{2,p}(x^*, \cdot)] \). Since LICQ holds, there are unique multipliers \((\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p_+ \) with \( \mu_j = 0, j \notin A(x^*) \) such that \( w = \sum_{i \in E} \lambda_i \nabla h_i(x^*) + \sum_{j \in A(x^*)} \mu_j \nabla g_j(x^*) \).

Now, to show that \((w, W) \in K_2^S(x^*) \), we only need to prove that \( H(\lambda, \mu) := \sum_{i \in E} \lambda_i \nabla^2 h_i(x^*) + \sum_{j \in A(x^*)} \mu_j \nabla^2 g_j(x^*) - W \) is a symmetric positive semi-definite matrix on \( C(x^*, \mu) \). Take \( d \in C(x^*, \mu) \). From the definition of \( C(x^*, \mu) \) and \( w \), we have that \( d \in K_{T_1(x^*)}(w) \). Furthermore, from LICQ, there is a \( C^2 \) curve \( \alpha(t), t \in (-\epsilon, \epsilon) \) such that \( \alpha(0) = x^*, \alpha'(0) = d \) and for every \( t \in (-\epsilon, \epsilon) \), for every \( i \in E \) and for all \( j \in A(x^*, d) \) we have \( h_i(\alpha(t)) = 0 \) and \( g_j(\alpha(t)) = 0 \). Set \( z := \alpha''(0) \). Since \( d \in K_{T_1(x^*)}(w) \), every \( j \) with \( \mu_j > 0 \) is in \( A(x^*, d) \). Now, taking derivative, we have that \( \nabla h_i(x^*)^T z + \nabla^2 h_i(x^*) d^2 = 0, \) (\( \forall i \)) and \( \nabla g_j(x^*)^T z + \nabla^2 g_j(x^*) d^2 = 0, \) (\( \forall j \) with \( \mu_j > 0 \)). Thus,

\[
H(\lambda, \mu) d^2 = \sum_{i \in E} \lambda_i \nabla^2 h_i(x^*) d^2 + \sum_{j \in A(x^*)} \mu_j \nabla^2 g_j(x^*) d^2 - W d^2 = - \sum_{i \in E} \lambda_i \nabla h_i(x^*)^T z - \sum_{j \in A(x^*)} \mu_j \nabla g_j(x^*)^T z - W d^2 = - (\sum_{i \in E} \lambda_i \nabla h_i(x^*) + \sum_{j \in A(x^*)} \mu_j \nabla g_j(x^*)^T z - W d^2
\]

where the last inequality holds since \((w, W) \in [T_1(x^*), T_{2,p}(x^*, \cdot)] \), \( T_{2,p}(x^*, d) = L_{2,p}(x^*, d) \) and \((z, 1) \in L_{2,p}(x^*, d) \). This completes the proof.

\( \square \)
3.4 Second-order cone and second-order sequential optimality conditions

In the recent years, several algorithms with second-order global convergence properties have been developed for constrained optimization. The convergence analysis of those methods has been improved by the study of the sequences generated by the method and the relationship with second-order sequential optimality conditions, see for instance [4, 41, 19]. Several similar applications of sequential optimality conditions for first-order methods have also been developed. See, for instance, [18] and references therein.

Given our previous discussion, the optimality condition relevant for global convergence of second-order algorithm is WSOC. Hence, each second-order sequential optimality condition has associated with it a companion CQ for WSOC. For instance, consider the sequential optimality condition called AKKT2 (second-order approximate KKT condition), [4]. Here, there is a companion CQ for WSOC, called second-order cone continuity property (CCP2) which states the outer semi-continuity at \( x^* \) of the set-valued mapping \( x \in \mathbb{R}^n \Rightarrow K^{W_2}(x) \), that is, \( \limsup_{x \to x^*} K^{W_2}(x) \subset K^{W_2}(x^*) \) where \( K^{W_2}(x) := \bigcup_{(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p, \mu_j = 0 \text{ for } j \notin A(x^*)} \left\{ (\sum_{i \in E} \lambda_i \nabla h_i(x) + \sum_{j \in A(x^*)} \mu_j \nabla g_j(x), H), \text{ such that } H \preceq \sum_{i \in E} \lambda_i \nabla^2 h_i(x) + \sum_{j \in A(x^*)} \mu_j \nabla^2 g_j(x) \text{ on } S(x, x^*) \right\} \). (3.19)

where \( S(x, x^*) := \{ d : \langle \nabla h_i(x), d \rangle = 0, i \in E; \langle \nabla g_j(x), d \rangle = 0, j \in A(x^*) \} \) is a perturbation of the critical subspace \( S(x^*) \) around the point \( x^* \). Clearly, \( K^{W_2}(x^*) = L^{2,w}(x^*) \). Now, there is a nice relationship between \( \tilde{N}_2(x^*) \) and the outer semi-continuity of \( K^{W_2}(x) \). In fact, following [4, Theorem 3.2] we obtain the inclusion \( \tilde{N}_2(x^*) \subset \limsup_{x \to x^*} K^{W_2}(x) \). (3.20)

From (3.20), it is easy to see that the outer semi-continuity of \( K^{W_2}(x) \) (CCP2) implies \( \tilde{N}_2(x^*) \subset K^{W_2}(x^*) \) and thus, we have that CCP2 is a CQ for WSOC. Similar results can be obtained for the Second-Order Complementarity-Approximate-KKT (CAKKT2) optimality condition, [41].

4 Final remarks

It is well known that Fritz-John type second-order optimality conditions have been thoroughly studied in the last years. They are present mostly in the most theoretical venues, given that they provide no-gap necessary and sufficient optimality conditions without the need of a constraint qualification. However, these conditions are difficult to be checked due to the necessity of knowing the full set of (generalized) Lagrange multipliers, and also, the involvement of the true critical cone, where checking positive semidefiniteness within it is a hard computational problem. Hence, it is an interesting question to know which problems
the Fritz-John multipliers can be replaced by Lagrange multipliers, besides the trivial case where all Fritz-John multipliers are Lagrange multipliers (MFCQ). Even though many research has appeared in the early days of optimization in second-order optimality conditions, minimal constraint qualifications ensuring the validity of the conditions have not been investigated thoroughly. In this paper, we reviewed the subject of second-order optimality conditions with the view of defining minimal CQs with respect to the three most well-known second-order optimality conditions under a CQ, namely, the basic, weak and strong second-order optimality conditions (BSOC, WSOC and SSOC, respectively). The refined analysis of each condition is relevant due to their use in different contexts, namely, WSOC is mostly relevant in global convergence analysis, while SSOC is a refinement of BSOC where only a single Lagrange multiplier is needed to check the condition. When only first-order derivatives of the constraints are considered in the formulation of the condition, most useful conditions are of constant rank of the set of gradients of active constraints type, where we showed that in this case, the second-order condition is expected to hold at any Lagrange multiplier. Considering gradients and Hessians of the constraints, we mainly investigated new Guignard-type constraint qualifications, where we mostly explore the connections of geometric optimality conditions with a conic description of second-order optimality conditions.

Acknowledgements

This research was funded by FAPESP grant 2013/05475-7 and CNPq.

References


A Appendix

We provide a proof for the technical Corollary needed in the proof of Theorem 3.4. First let us recall the lemma:

**Lemma A.1.** [53, Lemma 1.29] Let \( \rho : [0, \infty) \to [0, \infty) \) be a function having the right-hand derivative \( \rho'_+(0) \) and satisfying the conditions:

\[
\rho(0) = \rho'_+(0) = 0 \quad \text{and} \quad \rho(t) \leq \alpha + \beta t \quad \text{for all} \quad t \geq 0
\]

with positive constants \( \alpha \) and \( \beta \). Then, there is a non-decreasing, convex, continuously differentiable function \( \phi : [0, \infty) \to [0, \infty) \) such that

\[
\phi(0) = \phi'_+(0) = 0 \quad \text{and} \quad \phi(t) > \rho(t) \quad \text{for all} \quad t > 0.
\]

As a corollary, we have:

**Corollary A.2.** Let \( \rho : [0, \infty) \to [0, \infty) \) be a function having the right-hand derivative \( \rho'_+(0) \) with \( \lim_{t \to 0^+} \frac{\rho(t)}{t} = 0 \) and satisfying the conditions:

\[
\rho(0) = \rho'_+(0) = 0 \quad \text{and} \quad \rho(t) \leq \beta t^2 \quad \text{for all} \quad t \geq 0
\]

with a positive constant \( \beta \). Then, there is a non-decreasing, convex, twice-continuously differentiable function \( \phi : [0, \infty) \to [0, \infty) \) such that

\[
\phi(0) = \phi'_+(0) = \phi''_+(0) = 0 \quad \text{and} \quad \phi(t) > \rho(t) \quad \text{for all} \quad t > 0.
\]

**Proof.** Let \( \rho : [0, \infty) \to [0, \infty) \) a function satisfying the hypotheses. Define \( \phi_0 \) as \( \phi_0(t) = \rho(t)/t \) (if \( t > 0 \)) and \( \phi_0(0) = 0 \). The function \( \phi_0(t) \) is a function having the right-hand derivative and satisfying the hypotheses of the lemma above. Thus, there is a function \( \phi_1 \) there is a non-decreasing, convex, twice-continuously differentiable function \( \phi_1 : [0, \infty) \to [0, \infty) \) such that

\[
\phi_1(0) = \phi'_+ (0) = 0 \quad \text{and} \quad t\phi_1(t) > \rho(t) \quad \text{for all} \quad t > 0.
\]

Now, define \( \phi \) as \( \phi(t) = \frac{1}{t} \int_t^{2t} s \phi_1(s) ds \), (for \( t > 0 \)) and \( \phi(0) = 0 \). Since \( t\phi_1(t) \) is non-decreasing, we have that \( \phi(t) \) meets all the properties required. \( \square \)