Journal of Optimization Theory and Applications
https://doi.org/10.1007/s10957-019-01603-x

New Constraint Qualifications with Second-Order Properties in Nonlinear Optimization

G. Haeser[1] · **A. Ramos**[2]

Received: 8 January 2018 / Accepted: 12 October 2019

Abstract
In this paper, we present and discuss new constraint qualifications to ensure the validity of well-known second-order properties in nonlinear optimization. Here, we discuss conditions related to the so-called basic second-order condition, where a new notion of polar pairing is introduced in order to replace the polar operation, useful in the first-order case. We then proceed to define our second-order constraint qualifications, where we present an approach similar to the Guignard constraint qualification in the first-order case.

Keywords Nonlinear optimization · Constraint qualifications · Second-order optimality conditions

Mathematics Subject Classification 90C30 · 90C46

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

Communicated by Nikolai Pavlovich Osmolovskii.

✉ G. Haeser
ghaeser@ime.usp.br

A. Ramos
albertoramos@ufpr.br

1 Department of Applied Mathematics, University of São Paulo, São Paulo, SP, Brazil

2 Department of Mathematics, Federal University of Paraná, Curitiba, PR, Brazil

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 details, see [1, Framework 7.13, page 513], [2, Chapter 12] and [3].

Necessary optimality conditions are usually 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.

When seeking for weak constraint qualifications, which imply the validity of second-order conditions 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 start by seeking constraint qualifications ensuring the validity of standard second-order optimality conditions in nonlinear optimization theory (see Definition 2.1). In this paper, we consider only the so-called basic second-order optimality condition, whose formulation depends on the maximum of a quadratic form over all Lagrange multipliers, leaving the discussion about conditions that can be checked with a single Lagrange multiplier to a later study.

The paper is organized as follows: In Sect. 2, we present the basic concepts and definitions. first-order constraint qualifications with second-order properties, highlighting the central aspect of WSOC in second-order algorithms. In Sect. 3, we discuss second-order constraint qualifications and, in Sect. 4, we define new weak conditions with respect to the basic second-order optimality condition. Section 5 presents some conclusions and remarks.

2 Basic Definitions

A constraint qualification (CQ) is any property about the analytic description of the feasible set around a local minimizer that ensures the existence of Lagrange multipliers [see (2)]. We are particularly interested in CQs, which guarantee the existence of special Lagrange multipliers, namely the ones that can be used to formulate second-order necessary optimality conditions. For a very nice recent review of CQs in general, see [4]. A CQ is called a first-order one, if it uses only the gradients of the constraints in its formulation, while a second-order one is defined in terms of gradient and Hessian. Both types of conditions can yield second-order optimality conditions.

We continue with some notation. We denote by $\mathbb{R}^n$ the n-dimensional real Euclidean space, $n \in \mathbb{N}$, while $\mathbb{R}^n_+ \subseteq \mathbb{R}^n$ is the set of vectors, whose components are nonnegative. The set of symmetric matrices of order n is denoted by $\mathrm{Sym}(n)$. We use $\langle \cdot, \cdot \rangle$ to denote the Euclidean inner product on $\mathbb{R}^n$, with $\| \cdot \|$ the associated norm. For a cone $\mathcal{K} \subseteq \mathbb{R}^s$, its polar is $\mathcal{K}^\circ := \{v \in \mathbb{R}^s : \langle v, k \rangle \leq 0 \text{ for all } k \in \mathcal{K}\}$.

Consider the nonlinear constrained optimization problem

$$\begin{array}{lll} \text{minimize} & f(x), & \\ \text{subject to} & h_i(x) = 0 & \forall i \in \mathcal{E} := \{1, \ldots, m\}, \\ & g_j(x) \le 0 & \forall j \in \mathcal{I} := \{1, \ldots, p\}, \end{array} \tag{1}$$

where $f, h_i, g_j : \mathbb{R}^n \to \mathbb{R}$ are twice continuously differentiable functions. Denote by Ω the feasible set of (1). Given $x \in \Omega$, we define $A(x) := \{j \in \mathcal{I} : g_j(x) = 0\}$ as the set of indices of active inequalities.

The Lagrangian function is $L(x; \lambda, \mu) := f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{j=1}^p \mu_j g_j(x)$ where $(x, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p_+$. We use $\nabla_x L(x; \lambda, \mu)$ and $\nabla^2_{xx} L(x; \lambda, \mu)$ for the gradient and the Hessian of $L(x; \lambda, \mu)$ with respect to x, respectively.

Several second-order optimality conditions have been proposed in the literature, from both a theoretical and practical point of view; see [1,2,5–17] and references therein. In order to describe second-order conditions, we introduce some important sets.

For a feasible point $\bar{x}$, let us denote by $\Lambda(\bar{x})$ the set of Lagrange multipliers, that is, the set of vectors $(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p_+$ with

$$\nabla_x L(x; \lambda, \mu) = 0, \quad \text{and} \quad \mu_j g_j(\bar{x}) = 0, \ \forall j \in \mathcal{I}. \tag{2}$$

The cone of critical directions (critical cone) is defined as follows:

$$C(\bar{x}) := \left\{ d \in \mathbb{R}^n : \begin{array}{l} \langle \nabla f(\bar{x}), d \rangle \le 0; \ \langle \nabla h_i(\bar{x}), d \rangle = 0, i \in \mathcal{E}; \\ \langle \nabla g_j(\bar{x}), d \rangle \le 0, \ j \in A(\bar{x}) \end{array} \right\}. \tag{3}$$

Obviously, $C(\bar{x})$ is a non-empty, closed and convex cone. From the algorithmic point of view, an important set is the critical subspace, given by:

$$S(\bar{x}) := \left\{ d \in \mathbb{R}^n : \ \langle \nabla h_i(\bar{x}), d \rangle = 0, i \in \mathcal{E}; \ \langle \nabla g_j(\bar{x}), d \rangle = 0, \ j \in A(\bar{x}) \right\}. \tag{4}$$

In the case when $\Lambda(\bar{x}) \neq \emptyset$, a simple inspection shows that the critical subspace, $S(\bar{x})$, is the lineality space of the critical cone $C(\bar{x})$.

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

Definition 2.1 Let $\bar{x}$ be a feasible point with $\Lambda(\bar{x}) \neq \emptyset$. We say that

1. the *strong*[1] *second-order optimality condition* (SSOC) holds at $\bar{x}$, if there exists $(\lambda, \mu) \in \Lambda(\bar{x})$ such that $\langle \nabla^2_{xx} L(\bar{x}; \lambda, \mu) d, d \rangle \ge 0$ for every $d \in C(\bar{x})$;
2. the *weak second-order optimality condition* (WSOC) holds at $\bar{x}$, if there exists $(\lambda, \mu) \in \Lambda(\bar{x})$ such that $\langle \nabla^2_{xx} L(\bar{x}; \lambda, \mu) d, d \rangle \ge 0$ for every $d \in S(\bar{x})$;
3. the *basic second-order optimality condition* (BSOC) holds at $\bar{x}$, if $\forall d \in C(\bar{x})$, there exists $(\lambda, \mu) \in \Lambda(\bar{x})$ such that $\langle \nabla^2_{xx} L(\bar{x}; \lambda, \mu) d, d \rangle \ge 0$.

[1] Note that in the literature of sufficient second-order conditions, the term *strong* is usually associated with a condition, where the critical cone is replaced by the smallest subspace containing it, which is not the use we consider here.

For a discussion about the different types of second-order conditions and their relevance, we refer the reader to the extended version of [18]. In this paper, we focus our attention on BSOC.

We end this section with the following technical lemma. For the proof, see [19, Lemma 1.29].

Lemma 2.1 *Let $\rho : [0, \infty[\to [0, \infty[$ be a function having right-hand derivative $\rho'_+(0)$, such that $\rho(0) = \rho'_+(0) = 0$ and $\rho(t) \leq \alpha + \beta t$, $\forall t \geq 0$ for some $\alpha, \beta > 0$. Then, there is a continuously differentiable function ϕ such that $\phi(0) = \phi'_+(0) = 0$ and $\phi(t) > \rho(t)$, $\forall t > 0$.*

3 Second-Order Constraint Qualifications

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

It is well known that Guignard's CQ (see [20]) is the weakest CQ, for differentiable data, which implies the existence of Lagrange multipliers at a local minimizer. The main goal of this section is to pursue a similar idea but taking into account Lagrange multipliers connected to second-order optimality conditions.

In order to formulate our results, let us review Guignard's CQ and its weakness property. The first proof of this result, due to Gould and Tolle [21], was quite involving, but one can more easily understand the result with tools of variational analysis, which we define below. Given a set-valued mapping $\Gamma : \mathbb{R}^s \rightrightarrows \mathbb{R}^d$, the *sequential Painlevé–Kuratowski outer limit* of $\Gamma(z)$ as $z \to z^*$ is defined by

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

and we say that Γ is *outer semicontinuous* at z^* whenever

$$\limsup_{z \to z^*} \Gamma(z) \subseteq \Gamma(z^*).$$

For a given $S \subseteq \mathbb{R}^n$ and $z^* \in S$, the *tangent cone* to S at z^* is defined by $T_S(z^*) := \{d \in \mathbb{R}^n : \operatorname{dist}(z^* + t_n d, S) = o(t_n) \text{ for some } t_n \downarrow 0\}$, and the *regular normal cone* to S at $z^* \in S$ as

$$N_S(z^*) := \left\{ w \in \mathbb{R}^n : \limsup_{z \in \Omega, z \to z^*} \|z - z^*\|^{-1} \langle w, z - z^* \rangle \leq 0 \right\}.$$

The regular normal cone has the following remarkable properties. See [6].

Theorem 3.1 *Given $S \subseteq \mathbb{R}^n$ and $z^* \in S$. Then $w \in N_S(z^*)$, if and only if there is a continuously differentiable function f, which achieves its global minimum relative to S at z^*, such that $-\nabla f(z^*) = w$. Furthermore, $N_S(z^*) = T_S(z^*)^{\circ}$.*

When the set S under consideration is the feasible set Ω of (1), for $\bar{x} \in \Omega$, we use the notation $N_1(\bar{x}) := N_\Omega(\bar{x})$ and $T_1(\bar{x}) := T_\Omega(\bar{x})$. In that case, we also define the *first-order linearized cone* to Ω at $\bar{x} \in \Omega$ as

$$L_1(\bar{x}) := \left\{ d \in \mathbb{R}^n : \langle \nabla h_i(\bar{x}), d \rangle = 0, i \in \mathcal{E}; \;\; \langle \nabla g_i(\bar{x}), d \rangle \leq 0, \;\; i \in A(\bar{x}) \right\}.$$

Clearly, $T_1(\bar{x}) \subseteq L_1(\bar{x})$ always holds, while the reciprocal inclusion may not hold. It is easy to see that the KKT conditions hold at $\bar{x}$, for the objective function f with respect to Ω, if and only if $-\nabla f(\bar{x}) \in L_1(\bar{x})^\circ$, where the polar cone of $L_1(\bar{x})$ is given by

$$L_1(\bar{x})^\circ = \left\{ w \in \mathbb{R}^n : w = \sum_{i \in \mathcal{E}} \lambda_i \nabla h_i(\bar{x}) + \sum_{j \in A(\bar{x})} \mu_j \nabla g_j(\bar{x}), \lambda_i \in \mathbb{R}, \mu_j \geq 0 \right\}.$$

Therefore, from Theorem 3.1, the assumption $N_1(\bar{x}) = L_1(\bar{x})^\circ$, which is known as the Guignard CQ, is equivalent to the property that $\bar{x}$ is a KKT point for every continuously differentiable objective function f, such that $\bar{x}$ is a local minimizer of f over Ω. Thus, the Guignard CQ is the weakest condition, independently of the objective function, ensuring the validity of KKT at a local minimizer. Due to Theorem 3.1, the Guignard CQ is usually stated as $T_1(\bar{x})^\circ = L_1(\bar{x})^\circ$, which gives rise to the more well-known (and stronger) Abadie's CQ, namely $T_1(\bar{x}) = L_1(\bar{x})$.

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 Ω at $x \in \Omega$ is defined by the function $d \in \mathbb{R}^n \mapsto T_2(x, d)$, given by:

$$T_2(x, d) := \left\{ z \in \mathbb{R}^n : \operatorname{dist}\left(x + t_n d + \frac{1}{2} t_n^2 z, \Omega \right) = o\left(t_n^2\right) \text{ for some } t_n \downarrow 0 \right\}.$$

Note that $T_2(x, d) = \emptyset$ if $d \notin T_1(x)$. The set $T_2(x, d)$ is known as the outer second-order tangent to Ω at x in the direction d; see, for instance, [12,22].

Now, we consider the second-order geometric optimality condition, given in [10] (see also [23]), which is relevant only when $T_2(\bar{x}, d) \neq \emptyset$.

Theorem 3.2 *Let $\bar{x}$ be a local minimizer of (1). Then, for every $d \in T_1(\bar{x}) \cap \{\nabla f(\bar{x})\}^\perp$, we have $\langle \nabla f(\bar{x}), z \rangle + \langle \nabla^2 f(\bar{x}) d, d \rangle \geq 0$, for all $z \in T_2(\bar{x}, d)$.*

Now, let us define the linearization of the second-order tangent set

$$L_2(x, d) := \left\{ z \in \mathbb{R}^n : \begin{array}{ll} \langle \nabla h_i(x), z \rangle + \langle \nabla^2 h_i(x) d, d \rangle = 0, & i \in \mathcal{E}, \\ \langle \nabla g_i(x), z \rangle + \langle \nabla^2 g_i(x) d, d \rangle \leq 0, & i \in A(x, d) \end{array} \right\},$$

where $A(x, d) := \{i \in A(x) : \langle \nabla g_i(x), d \rangle = 0\}$.

Clearly, for every $d \in T_1(x)$, we have $T_2(x, d) \subseteq L_2(x, d)$. If for every $d \in C(x) = L_1(x) \cap \{d : \langle \nabla f(x), d \rangle = 0\}$ we impose that $T_2(x, d) = L_2(x, d)$ (what is called the second-order Abadie CQ, see [24,25]), we have that BSOC holds when

x is a local minimizer. This assumption can be replaced by the equality of the polars, $T_2(x,d)^\circ = L_2(x,d)^\circ$, for every $d \in C(x)$, which was called the second-order Guignard CQ [24]. In [26], the following CQ, which can be seen as a second-order version of MFCQ, is introduced: $\{\nabla h_i(x)\}_{i=1}^m$ is linearly independent and there exist $d \in C(x)$ and $z \in \mathbb{R}^n$, such that $\langle \nabla h_i(x), z\rangle + \langle \nabla^2 h_i(x)d, d\rangle = 0, i \in \mathcal{E}$ and $\langle \nabla g_i(x), z\rangle + \langle \nabla^2 g_i(x)d, d\rangle < 0, i \in A(x,d)$. 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.

Since the second-order tangent set $T_2(x,d)$ can be an empty set for certain directions $d \in T_1(x)$, following [10], we consider the *projective second-order tangent set* of Ω at $\bar{x}$, defined by:

$$T_2^{proj}(\bar{x},d) := \limsup_{(t,r,t/r)\downarrow 0} \frac{\Omega - td - \bar{x}}{\frac{t^2}{2r}}.$$

Thus, $(z,r) \in T_2^{proj}(\bar{x},d)$ iff there is a sequence $(z^k, r_k, t_k) \in \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{R}_+$, such that $(z^k, r_k, t_k) \to (z, r, 0)$, $t_k/r_k \downarrow 0$, $r_k > 0$ and $\bar{x} + t_k d + \frac{t_k^2}{2r_k} z^k \in \Omega$.

This concept has been used in the context of composite optimization [27] and vector optimization [28,29]. The set $T_2^{proj}(\bar{x},d)$ is a non-empty closed cone by [10, Proposition 2.1.]. It is not difficult to see that if $(z,r) \in T_2^{proj}(\bar{x},d)$ for some $r > 0$, then $z \in T_2(\bar{x},d)$, and if $z \in T_2(\bar{x},d)$, then $(z,1) \in T_2^{proj}(\bar{x},d)$. Using the projective second-order tangent set, we have the following necessary optimality condition, for twice continuously differentiable data. See [10].

Theorem 3.3 *Let $\bar{x}$ be a local minimizer of* (1). *Then, for every $d \in T_1(\bar{x}) \cap \nabla f(\bar{x})^\perp$, we have that $\langle \nabla f(\bar{x}), z\rangle + r\langle \nabla^2 f(\bar{x})d, d\rangle \geq 0$, for all $(z,r) \in T_2^{proj}(\bar{x},d)$.*

Since $T_2^{proj}(\bar{x},d)$ can be a difficult object to deal with, we define the *projective linearized second-order tangent set* of Ω at $\bar{x}$ as

$$L_2^{proj}(\bar{x},d) := \left\{ (z,r) \in \mathbb{R}^n \times \mathbb{R}_+ : \begin{array}{l} \langle \nabla h_i(\bar{x}), z\rangle + r\langle \nabla^2 h_i(\bar{x})d, d\rangle = 0, i \in \mathcal{E}; \\ \langle \nabla g_i(\bar{x}), z\rangle + r\langle \nabla^2 g_i(\bar{x})d, d\rangle \leq 0, \\ i \in A(\bar{x},d). \end{array} \right\}. \tag{5}$$

Clearly, for every $d \in T_1(\bar{x})$, we see that $T_2^{proj}(\bar{x},d) \subseteq L_2^{proj}(\bar{x},d)$. As $L_2^{proj}(\bar{x},d)$ is defined by linear equalities and inequalities, its polar can be calculated, in fact: $(w,\eta) \in L_2^{proj}(\bar{x},d)^\circ$ if, and only if, there exist multipliers $\lambda_i \in \mathbb{R}$, $i \in \mathcal{E}$; $\mu_j \geq 0$, $j \in A(\bar{x},d)$, and $\beta \geq 0$, such that

$$\begin{pmatrix} w \\ \eta \end{pmatrix} = \sum_{i\in\mathcal{E}} \lambda_i \begin{pmatrix} \nabla h_i(\bar{x}) \\ \langle \nabla^2 h_i(\bar{x})d, d\rangle \end{pmatrix} + \sum_{j\in A(\bar{x},d)} \mu_j \begin{pmatrix} \nabla g_j(\bar{x}) \\ \langle \nabla^2 g_j(\bar{x})d, d\rangle \end{pmatrix} - \begin{pmatrix} 0 \\ \beta \end{pmatrix}. \tag{6}$$

Since our goal is to formulate a weak condition to ensure BSOC, we define the second-order normal cone $N_2(\bar{x}) \subseteq \mathbb{R}^n \times \mathrm{Sym}(n)$ to Ω at $\bar{x} \in \Omega$ as follows:

$$N_2(\bar{x}) := \left\{ (w, W) : \limsup_{x \in \Omega, x \to \bar{x}} \frac{\langle w, x - \bar{x}\rangle + (1/2) W(x - \bar{x})^2}{\|x - \bar{x}\|^2} \leq 0 \right\}. \tag{7}$$

In some sense, an element of the second-order normal cone stands for a vector-matrix pair, which plays the role of a gradient–Hessian pair in a vanishing, in Ω, Taylor-like expansion. The following result makes this formulation precise and is a second-order version of Theorem 3.1:

Theorem 3.4 *Let $\bar{x} \in \Omega$, $w \in \mathbb{R}^n$, and $W \in \mathrm{Sym}(n)$. Then, $(w, W) \in N_2(\bar{x})$ iff there exists a twice continuously differentiable function f, which attains its global minimum relative to Ω at $\bar{x}$, such that $-\nabla f(\bar{x}) = w$ and $-\nabla^2 f(\bar{x}) = W$.*

Proof First, we will prove the "if" implication. Let $(w, W) \in N_2(\bar{x})$ be an arbitrary element. Define the function $\eta_0 : [0, \infty[\to \mathbb{R}$ as

$$\eta_0(t) := \sup\{w^T(x - \bar{x}) + 1/2\langle W(x - \bar{x}), x - \bar{x}\rangle : \|x - \bar{x}\| \leq t, \;\; x \in \Omega\}.$$

Clearly, $\eta_0(t)$ is non-decreasing and $0 = \eta_0(0) \leq \eta_0(t)$. As $(w, W) \in N_2(\bar{x})$, we see that $\lim_{t\to 0} \eta_0(t)t^{-2} \leq 0$, and thus, there exist $M, T > 0$, such that $\eta_0(t) \leq Mt^2$, for all $t \leq T$. Modify $\eta_0(t)$ outside $[0, T[$ in such a way that $\eta_0(t) \leq Mt^2$ over $[0, \infty[$.

For a moment, let us suppose that there is a C^2 function $\phi : [0, \infty[\to [0, \infty[$ such that $\phi(0) = \phi'_+(0) = \phi''_+(0) = 0$ and $\phi(t) > \eta_0(t)$ for all $t > 0$. Thus, we can use ϕ to define a C^2 function, with the desired properties. Set $F(x) := \langle w, x - \bar{x}\rangle + 1/2\langle W(x - \bar{x}), (x - \bar{x})\rangle - \phi(\|x - \bar{x}\|)$. Observe that $F(x)$ is a C^2 function, due to the smoothness of ϕ. Moreover, $\nabla F(\bar{x}) = w$ and $\nabla^2 F(\bar{x}) = W$. We observe that, as a consequence of the inequality $\phi(\|x - \bar{x}\|) > \eta_0(\|x - \bar{x}\|)$ for every $x \neq \bar{x} \in \Omega$, we get that $F(x) < F(\bar{x})$ for all $x \neq \bar{x} \in \Omega$, and thus, $F(x)$ achieves its global maximum over Ω uniquely at $\bar{x}$.

Here, we will show that there exists such $\phi : [0, \infty[\to [0, \infty[$. Define ϕ_0 as $\phi_0(t) = \eta_0(t)/t$ (if $t > 0$) and $\phi_0(0) = 0$. Since $\eta_0(0) = \eta'_0(0) = 0$, we have $\lim_{t\to 0} \eta_0(t)t^{-2} \leq 0$ and $\eta_0(t) \leq Mt^2$, $t \geq 0$. The function $\phi_0(t)$ satisfies the hypotheses of Lemma 2.1. Thus, there is a C^1 function $\phi_1 : [0, \infty[\to [0, \infty[$, such that $\phi_1(0) = \phi'_{1+}(0) = 0$, and $t\phi_1(t) > \eta_0(t)$ for all $t > 0$. Finally, set $\phi(t) = \frac{1}{t}\int_t^{2t} s\phi_1(s)\mathrm{d}s$, (for $t > 0$) and $\phi(0) = 0$. Since $t\phi_1(t)$ is non-decreasing, $\phi(t)$ meets all the required properties.

The "only if" part is a simple consequence of the Taylor expansion. □

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

$$[A, B(\cdot)] := \left\{ (w, W) \in \mathbb{R}^n \times \mathrm{Sym}(n) : \begin{array}{l} \langle w, d\rangle \leq 0, \quad \text{for all } d \in A \\ \langle w, z\rangle + r\langle Wd, d\rangle \leq 0, \text{ for all} \\ (d, (z, r)) \in K_A(w) \times B(d) \end{array} \right\}, \tag{8}$$

where $K_A(v) := A \cap v^{\perp}$ for $v \in \mathbb{R}^n$.

Remark 3.1 If $A := L_1(x)$, then $K_{L_1(x)}(\nabla f(x))$ is the critical cone $C(x)$.

Using the second-order normal cone, the polar pairing, and Theorem 3.4 and Theorem 3.3 can be rewritten as:

$$N_2(\bar{x}) \subseteq \left[T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)\right]. \tag{9}$$

This can been seen as a second-order version of the inclusion $N_1(\bar{x}) \subseteq T_1(\bar{x})^{\circ}$. In the next section, we will use $N_2(\bar{x})$ to associate a weak CQ with BSOC.

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

The polar pairing operator unifies the geometric optimality condition, when it is applied to the tangent objects $(T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot))$, while when applied to the linearized objects $(L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot))$, the polar pairing 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 $-\nabla f(\bar{x}) \in T_1(\bar{x})^{\circ}$ gives the geometric optimality condition, while the KKT conditions are given by $-\nabla f(\bar{x}) \in L_1(\bar{x})^{\circ}$. 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 $N_2(\bar{x})$ is not known.

4 Weak Constraint Qualification for BSOC

Let us define several weak CQs that ensure the validity of BSOC. We start by showing the following inclusion:

Proposition 4.1 *It always holds that* $[L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)] \subseteq [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$.

Proof Take $(w, W) \in [L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$. By the definition of the polar pairing (8), w belongs to $L_1(\bar{x})^{\circ}$, and since $L_1(\bar{x})^{\circ} \subseteq T_1(\bar{x})^{\circ}$, we get that $w \in T_1(\bar{x})^{\circ}$. We see that the expression: $\langle w, z\rangle + r\langle Wd, d\rangle \leq 0$, for every $(d, (z, r)) \in K_{L_1(\bar{x})}(w) \times L_2^{proj}(\bar{x}, d)$, is equivalent to saying that $(w, r) \in L_2^{proj}(\bar{x}, d)^{\circ}$ for every $d \in K_{L_1(\bar{x})}(w)$. This equivalence implies that $(w, r) \in T_2^{proj}(\bar{x}, d)^{\circ}$ for every $d \in K_{T_1(\bar{x})}(w)$, since $T_2^{proj}(\bar{x}, d) \subseteq L_2^{proj}(\bar{x}, d)$ and $K_{T_1(\bar{x})}(w) \subseteq K_{L_1(\bar{x})}(w)$. This shows that $(w, W) \in [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$. □

We will use $[L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$ to characterize BSOC in the following sense:

Proposition 4.2 *BSOC holds at the feasible point* $\bar{x}$ *for the* C^2 *objective function* f *if, and only if* $(-\nabla f(\bar{x}), -\nabla^2 f(\bar{x})) \in [L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$.

Proof First, we wil see that if BSOC holds at $\bar{x} \in \Omega$, for some C^2 function f, then $(-\nabla f(\bar{x}), -\nabla^2 f(\bar{x})) \in [L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$. Indeed, from BSOC, we get for every $d \in C(\bar{x}) = L_1(\bar{x}) \cap \nabla f(\bar{x})^{\perp}$, there exists $(\lambda, \mu) \in \Lambda(x^*)$ (which may depend on d), such that $\langle \nabla^2_{xx} L(\bar{x}, \lambda, \mu)d, d\rangle \geq 0$. Thus, $\Lambda(\bar{x}) \neq \emptyset$. Now, using (6), we get that $(-\nabla f(\bar{x}), -\langle \nabla^2 f(\bar{x})d, d\rangle) \in L_2^{proj}(\bar{x}, d)^{\circ}$, $\forall d \in L_1(\bar{x}) \cap \nabla f(\bar{x})^{\perp}$, which implies that $(-\nabla f(\bar{x}), -\nabla^2 f(\bar{x})) \in [L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$.

To prove the other implication, take a C^2 objective function f, such that $(-\nabla f(\bar{x}), -\nabla^2 f(\bar{x})) \in [L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$. From the definition of the polar pairing $[L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$, we see that

(i) $-\nabla f(\bar{x}) \in L_1(\bar{x})^{\circ}$, and thus, $\Lambda(\bar{x})$ is non-empty and;
(ii) for every direction $d \in C(\bar{x}) = L_1(\bar{x}) \cap \{\nabla f(\bar{x})\}^{\perp}$, we obtain that $(-\nabla f(\bar{x}), -\langle \nabla^2 f(\bar{x})d, d\rangle) \in L_2^{proj}(\bar{x}, d)^{\circ}$. Hence, by (6), there exist multipliers $(\lambda, \mu) \in \Lambda(\bar{x})$, such that

$$\langle \nabla^2 f(\bar{x})d, d\rangle + \sum_{i \in \mathcal{E}} \lambda_i \langle \nabla^2 h_i(\bar{x})d, d\rangle + \sum_{j=1}^{p} \mu_j \langle \nabla^2 g_j(\bar{x})d, d\rangle \geq 0,$$

that is, BSOC holds at $\bar{x}$. □

Thus, in view of Theorem 3.4 and Proposition 4.2, we have the following result:

Theorem 4.1 *BSOC holds at $\bar{x}$ for every C^2 objective function f having $\bar{x}$ as a local minimizer constrained to Ω if, and only if,*

$$N_2(\bar{x}) \subseteq \left[L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)\right]. \tag{10}$$

Using the polar pairing and the inclusion $N_2(\bar{x}) \subseteq [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$, given by Theorem 3.3, a weak CQ that ensures BSOC at local minimizers is

$$\left[T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)\right] = \left[L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)\right], \tag{11}$$

which we call *weak second-order Guignard CQ for BSOC*. This gives rise to two new weak CQs to ensure BSOC. The first, we call *second-order Guignard CQ for BSOC*,

$$T_1(\bar{x}) = L_1(\bar{x}) \text{ and } T_2^{proj}(\bar{x}, d)^{\circ} = L_2^{proj}(\bar{x}, d)^{\circ} \text{ for all } d \in \mathbb{R}^n, \tag{12}$$

and the second, we call *second-order Abadie CQ for BSOC*

$$T_1(\bar{x}) = L_1(\bar{x}) \text{ and } T_2^{proj}(\bar{x}, d) = L_2^{proj}(\bar{x}, d) \text{ for every } d \in \mathbb{R}^n. \tag{13}$$

We see that (12) and (13) are indeed CQs for BSOC, because they imply (11), which in turn implies (10), the characterization of BSOC.

Table 1 New constraint qualifications for BSOC

Characterization of BSOC	$N_2(\bar{x}) \subseteq [L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$
Weak Guignard CQ for BSOC	$[T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)] = [L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$
Guignard CQ for BSOC	$T_1(\bar{x}) = L_1(\bar{x})$ and $T_2^{proj}(\bar{x}, d)^\circ = L_2^{proj}(\bar{x}, d)^\circ, \forall d \in \mathbb{R}^n$
Abadie CQ for BSOC	$T_1(\bar{x}) = L_1(\bar{x})$ and $T_2^{proj}(\bar{x}, d) = L_2^{proj}(\bar{x}, d), \forall d \in \mathbb{R}^n$

A well-known CQ for BSOC is the second-order Abadie condition introduced by Kawasaki [24], which states the equality $T_2(\bar{x}, d) = L_2(\bar{x}, d)$, for every $d \in T_1(\bar{x})$. This assumes that $T_2(\bar{x}, d) \neq \emptyset$ for every direction, which is not true in general. Under this assumption, we have that $T_2(\bar{x}, d) = L_2(\bar{x}, d)$ implies equality between the sets $[T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$ and $[L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$, as the next proposition will show:

Proposition 4.3 *Assume that $T_2(\bar{x}, d) \neq \emptyset$ and $T_2(\bar{x}, d) = L_2(\bar{x}, d)$, for all $d \in T_1(\bar{x})$. Then, $[T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)] = [L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$.*

Proof Take $(w, W) \in [T_1(\bar{x}), T_2^{proj}(\bar{x}, \cdot)]$. First, we will show that w is in $L_1(\bar{x})^\circ$. Indeed, since the equality $T_2(\bar{x}, d) = L_2(\bar{x}, d)$ holds for $d = 0$, we have $T_1(\bar{x}) = L_1(\bar{x})$. Thus, $w \in T_1(\bar{x})^\circ = L_1(\bar{x})^\circ$.

Now, take $d \in K_{T_1(\bar{x})}(w)$. By the definition of the polar pairing, using the equality $T_2(\bar{x}, d) = L_2(\bar{x}, d)$, and since $(z, 1) \in T_2^{proj}(\bar{x}, d)$ implies that $z \in T_2(\bar{x}, d)$, we get that $\langle w, z\rangle + \langle Wd, d\rangle \leq 0$, for all $z \in L_2(\bar{x}, d)$. Thus, consider the maximization problem below, whose optimal value is not positive:

$$\text{Maximize } \{\langle w, z\rangle + \langle Wd, d\rangle\}, \text{ subject to } z \in L_2(\bar{x}, d).$$

From strong duality for linear programming, (since $L_2(\bar{x}, d)$ is defined by affine constraints and it is non-empty), there exist multipliers $(\hat{\lambda}, \hat{\mu}) \in \mathbb{R}^m \times \mathbb{R}^p_+$, with $\mu_j = 0$ for $j \notin A(\bar{x}, d)$, such that $-w + \sum_{i=1}^m \hat{\lambda}_i \nabla h_i(\bar{x}) + \sum_{j=1}^p \hat{\mu}_j \nabla g_j(\bar{x}) = 0$ and $-\langle Wd, d\rangle + \sum_{i=1}^m \hat{\lambda}_i \langle \nabla^2 h_i(\bar{x})d, d\rangle + \sum_{j=1}^p \hat{\mu}_j \langle \nabla^2 g_j(\bar{x})d, d\rangle \geq 0$. Then, by using (6), we see that $(w, \langle Wd, d\rangle) \in L_2^{proj}(\bar{x}, d)^\circ$, that is, $\langle w, z\rangle + r\langle Wd, d\rangle \leq 0$, for every $(z, r) \in L_2^{proj}(\bar{x}, d)$. Thus, $(w, W) \in [L_1(\bar{x}), L_2^{proj}(\bar{x}, \cdot)]$.

The other inclusion is a consequence of Proposition 4.1. □

Now, instead of considering conditions that jointly take into account all directions d in the critical cone, we will fix one direction in the tangent cone (Table 1). Under MFCQ, condition BSOC is well known to hold [9,26]. In fact, for each direction in the critical cone $d \in C(\bar{x})$, the Lagrange multipliers (λ, μ) constructed are such that, $\mu_j = 0$ for all $j \notin A(\bar{x}, d) := A(\bar{x}) \cap \{j\colon \langle \nabla g_j(\bar{x}), d\rangle = 0\}$. This specification turns out to be relevant in our analysis, so we single out the following definition:

Definition 4.1 Given $\bar{x} \in \Omega$ and $d \in C(\bar{x})$, we say that the *basic second-order optimality condition in the direction d* (BSOC(d)) holds at $\bar{x}$, if there exists $(\lambda, \mu) \in \Lambda(\bar{x})$ such that $\langle \nabla^2_{xx} L(\bar{x}, \lambda, \mu)d, d\rangle \geq 0$, with $\mu_j = 0, \forall j \notin A(\bar{x}, d)$.

Table 2 New constraint qualifications for BSOC(d)

Characterization of BSOC(d)	$N_2(\bar{x}, d) \subseteq L_2^{proj}(\bar{x}, d)^\circ$
Guignard CQ for BSOC(d)	$T_2^{proj}(\bar{x}, d)^\circ = L_2^{proj}(\bar{x}, d)^\circ$
Abadie CQ for BSOC(d)	$T_2^{proj}(\bar{x}, d) = L_2^{proj}(\bar{x}, d)$

Now, we proceed to develop weak conditions to ensure the validity of BSOC(d) at local minimizers. The main reason for doing this is that collecting all these conditions gives a new condition for the validity of BSOC, that is based only on the usual polar cone, rather than the polar pairing. Note however that the analysis of second-order optimality conditions along a specific direction d is not new, and it has been developed, for instance, in [30] and references therein.

Thus, consider the *second-order regular normal cone* of Ω *in the direction* d at $\bar{x}$, defined as:

$$N_2(\bar{x}, d) := \{(w, \langle Wd, d\rangle) \in \mathbb{R}^n \times \mathbb{R} : (w, W) \in N_2(\bar{x}) \text{ and } w \perp d\}. \tag{14}$$

Loosely speaking, in the light of Theorem 3.4, $N_2(\bar{x}, d)$ represents the set of all C^2 functions such that $\bar{x}$ is a local minimizer relative to Ω, that have the direction d as a critical direction. It is not difficult to see that $N_2(\bar{x}, d)$ is a non-empty and convex cone in $\mathbb{R}^n \times \mathbb{R}$ (Table 2).

Using the directional second-order normal cone and Theorem 3.4, the second implication in the geometric optimality condition, given by Theorem 3.3, can be rewritten as the inclusion

$$N_2(\bar{x}, d) \subseteq T_2^{proj}(\bar{x}, d)^\circ, \text{ for each } d \in T_1(\bar{x}). \tag{15}$$

By the polar cone of $L_2^{proj}(\bar{x}, d)$ [see (6)], $(-\nabla f(\bar{x}), -\langle\nabla^2 f(\bar{x})d, d\rangle)$ belongs to $L_2^{proj}(\bar{x}, d)^\circ$, if and only if there exists $(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p_+$, with $\mu_j = 0$ for $j \notin A(x, d)$, such that $\nabla_x L(\bar{x}; \lambda, \mu) = 0$ and $\langle\nabla^2_{xx} L(\bar{x}; \lambda, \mu)d, d\rangle \geq 0$. As a consequence, $(-\nabla f(\bar{x}), -\langle\nabla^2 f(\bar{x})d, d\rangle) \in L_2^{proj}(\bar{x}, d)^\circ$, if and only if BSOC($d$) holds at $\bar{x}$. Thus, we get the following result:

Theorem 4.2 *Let d be a vector in $\mathbb{R}^n$ and $\bar{x} \in \Omega$. BSOC(d) holds at $\bar{x}$ for every C^2 function, which has $\bar{x}$ as a local minimizer relative to Ω and has d as critical direction, iff $N_2(\bar{x}, d) \subseteq L_2^{proj}(\bar{x}, d)^\circ$.*

Using (15), we see that weak CQs that guarantee that BSOC(d) is an optimality condition when d is a critical direction are $T_2^{proj}(\bar{x}, d)^\circ = L_2^{proj}(\bar{x}, d)^\circ$ (which we call *second-order Guignard CQ for BSOC(d)*) and $T_2^{proj}(\bar{x}, d) = L_2^{proj}(\bar{x}, d)$ (which we refer as *second-order Abadie CQ for BSOC(d)*).

5 Conclusions

In this paper, we investigated weak CQs ensuring the validity of the so-called basic second-order necessary optimality condition (BSOC). These are based on a characterization of a second-order normal cone at $\bar{x} \in \Omega$ in terms of a gradient–Hessian pair of an objective function attaining a global minimizer in Ω at $\bar{x}$. By defining an extension of the polar operation considering first- and second-order objects (polar pairing), we were able to characterize the second-order geometric optimality condition as the polar pairing of the first- and second-order tangent sets, while BSOC is given by the polar pairing of their linearizations. This allowed us to define new Abadie-type and Guignard-type CQs ensuring BSOC at a local minimizer. We also considered weak CQs that ensure BSOC at a given direction d, and in this cases, one relies on the usual polar operation.

Acknowledgements The first author was supported by FAPESP (Grants 2013/05475-7, 2017/18308-2 and 2018/24293-0) and CNPq. The second author was funded by CNPq, Grants 454798/2015-6 and 438185/2018-8.

References

1. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, Berlin (2006)
2. Fletcher, R.: Practical Methods of Optimization: Constrained Optimization. Wiley, New York (1981)
3. Andreani, R., Martínez, J.M., Ramos, A., Silva, P.J.S.: A cone-continuity constraint qualification and algorithmic consequences. SIAM J. Optim. **26**(1), 96–110 (2016)
4. Giorgi, G.: A guided tour in constraint qualifications for nonlinear programming under differentiability assumptions. http://economia.unipv.it/docs/dipeco/quad/ps/RePEc/pav/demwpp/DEMWP0160.pdf (2018). Accessed 1 Dec 2018
5. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Practical Methods of Optimization: Theory and Algorithms. Wiley, New York (2006)
6. Rockafellar, R.T., Wets, R.: Variational Analysis, Grundlehren der mathematischen Wissenschaften, vol. 317. Springer, Berlin (2009)
7. Fiacco, A.V., McCormick, G.P.: Nonlinear Programming Sequential Unconstrained Minimization Techniques. Wiley, New York (1968)
8. Bertsekas, D.P.: Nonlinear Programming. Athenas Scientific, Belmont (1999)
9. Bonnans, J.F., Shapiro, A.: Pertubation Analysis of Optimization Problems. Springer, Berlin (2000)
10. Penot, J.-P.: Second-order conditions for optimization problems with constraints. SIAM J. Control Optim. **37**, 303–318 (1998)
11. Casas, E., Troltzsch, F.: Second-order necessary and sufficient optimality conditions for optimization problems and applications to control theory. SIAM J. Optim. **13**(2), 406–431 (2002)
12. Bonnans, J.F., Cominetti, R., Shapiro, A.: Second-order optimality conditions based on parabolic second order tangent sets. SIAM J. Optim. **9**, 466–492 (1999)
13. Arutyunov, A.V., Pereira, F.L.: Second-order necessary optimality conditions for problems without a priori normality assumptions. Math. Oper. Res. **31**(1), 1–12 (2006)
14. Baccari, A., Trad, A.: On the classical necessary second-order optimality conditions in the presence of equality and inequality constraints. SIAM J. Optim. **15**(2), 394–408 (2005)
15. Daldoul, M., Baccari, A.: An application of matrix computations to classical second-order optimality conditions. Optim. Lett. **3**, 547–557 (2009)
16. Gfrerer, H.: Second-order necessary conditions for nonlinear optimization with abstract constraints: the degenerate case. SIAM J. Optim. **18**(2), 589–612 (2007)
17. Bomze, I.: Copositivity for second-order optimality conditions in general smooth optimization problems. Optimization **65**(4), 779–795 (2016)

18. Behling, R., Haeser, G., Ramos, A., and Viana, D.S. On a conjecture in second-order optimality conditions. J. Optim. Theory Appl. 176(3), 625–633 (2018). Extended version: arXiv:1706.07833
19. Murdokhovich, B.S.: Variational Analysis and Generalized Differentiation I: Basic Theory. Grundlehren der mathematischen Wissenschaften, vol. 330. Springer, Berlin (2005)
20. Guignard, M.: Generalized Kunh–Tucker conditions for mathematical programming in a Banach space. SIAM J. Control **7**, 232–241 (1969)
21. Gould, F.J., Tolle, J.W.: A necessary and sufficient qualification for constrained optimization. SIAM J. Appl. Math. **20**, 164–172 (1971)
22. Bonnans, J.F., Hermant, A., Arutyunov, A.V., Pereira, F.L.: No-gap second-order optimality conditions for optimal control problems with a single state constraint and control. Math. Program. Serie B **117**, 21–50 (2009)
23. Penot, J.-P.: Recent Advances on Second-Order Optimality Conditions. Springer, Berlin (2000)
24. Kawasaki, H.: Second-order necessary conditions of the Kuhn–Tucker type under new constraint qualifications. J. Optim. Theory Appl. **57**(2), 253–264 (1988)
25. Huy, N.Q., Tuyen, N.V.: New second-order optimality conditions for a class of differentiable optimization problems. J. Optim. Theory Appl. **171**(1), 27–44 (2016)
26. Ben-Tal, A.: Second-order and related extremality conditions in nonlinear programming. J. Optim. Theory Appl. **31**(2), 143–165 (1980)
27. Penot, J.-P.: Optimality conditions in mathematical programming and composite optimization. Math. Program. **67**, 225–245 (1994)
28. Giorgi, G., Jiménez, B., Novo, V.: An overview of second order tangent sets and their application to vector optimization. Bol. Soc. Esp. Mat. Apl. **52**, 73–96 (2010)
29. Jiménez, B., Novo, V.: Optimality conditions in differentiable vector optimization via second-order tangent sets. Appl. Math. Optim. **49**, 123–144 (2004)
30. Solodov, M.V.: Constraint qualifications. In: Cochran, J.J., et al. (eds.) Wiley Encyclopedia of Operations Research and Management Science. Wiley, New York (2010)