J Optim Theory Appl (2011) 149: 528–539
DOI 10.1007/s10957-011-9802-x

On Approximate KKT Condition and its Extension to Continuous Variational Inequalities

Gabriel Haeser · María Laura Schuverdt

Published online: 21 January 2011

Abstract In this work, we introduce a necessary sequential Approximate-Karush-Kuhn-Tucker (AKKT) condition for a point to be a solution of a continuous variational inequality, and we prove its relation with the Approximate Gradient Projection condition (AGP) of Gárciga-Otero and Svaiter. We also prove that a slight variation of the AKKT condition is sufficient for a convex problem, either for variational inequalities or optimization. Sequential necessary conditions are more suitable to iterative methods than usual punctual conditions relying on constraint qualifications. The AKKT property holds at a solution independently of the fulfillment of a constraint qualification, but when a weak one holds, we can guarantee the validity of the KKT conditions.

Keywords Optimality conditions · Variational inequalities · Constraint qualifications · Practical algorithms

Communicated by G. Di Pillo.

We are indebted to two anonymous referees for insightful comments and suggestions.
G. Haeser was supported by CNPq Grant 503328/2009-0 and FAPESP Grant 09/09414-7.
M.L. Schuverdt was supported by PRONEX-CNPq/FAPERJ Grant E-26/171.164/2003.

G. Haeser
Institute of Mathematics and Statistics, University of São Paulo, São Paulo SP, Brazil
e-mail: gabriel.haeser@unifesp.br

Present address:
G. Haeser
Institute of Science and Technology, Federal University of São Paulo, São José dos Campos SP, Brazil

M.L. Schuverdt (✉)
CONICET, Department of Mathematics, FCE, University of La Plata, CP 172, 1900 La Plata, Argentina
e-mail: schuverd@mate.unlp.edu.ar

1 Introduction

In this paper we study necessary sequential conditions for optimization problems and continuous variational inequalities.

At least two kinds of necessary optimality conditions can be found in the literature for the general nonlinear programming problem (NLP). From the practical point of view, optimality conditions arise based on the sequence of iterands generated by a given algorithm. These so-called sequential optimality conditions do not require any constraint qualification, and they provide a natural stopping criterion for iterative methods. Sequential optimality conditions have been studied in [1–4] for optimization problems. As mentioned in [3], we say that a feasible point x^* verifies a sequential optimality condition if there exists a sequence $\{x^k\}$ that converges to x^* and verifies a mathematical proposition $\mathcal{P}$ that, in some sense, measures optimality of x^k. For example, the Approximate-Karush-Kuhn-Tucker sequential optimality condition for nonlinear programming problems (AKKT, [3]) requires the existence of a sequence $\{x^k\}$ converging to x^* with the property that x^k is "almost a Karush-Kuhn-Tucker (KKT) point" for every k, in the sense that there exists an appropriate sequence of Lagrange multipliers for which the gradient of the Lagrangian function at x^k converges to zero. Another sequential optimality condition was used to analyze the Inexact Restoration method in [5, 6]. The so-called Approximate Gradient Projection condition (AGP, [1]) is said to hold at a feasible point x^* if there exists a sequence $\{x^k\}$ converging to x^* such that the projected gradient on an approximated feasible set, at x^k, converges to zero. In [3], some stronger generalizations of AGP were presented when the sequence of iterates lies on a linear or convex set.

On the other hand, from the theoretical point of view, there exist the classical punctual optimality conditions that are associated with some constraint qualification. These optimality conditions can be of first order or second order. We know that when a local minimizer of a general smooth nonlinear problem verifies any first order constraint qualification then this point verifies the KKT condition. Thus, given any first order constraint qualification CQ1 we say that a feasible point x^* verifies a first order punctual optimality condition if x^* verifies the KKT condition or it does not verify CQ1, [7]. When second derivatives are present and a local minimizer verifies any second order constraint qualification, then it is possible to prove that the Hessian of the Lagrangian function is positive semi-definite on a tangent subspace [8]. Thus, given any second order constraint qualification CQ2 we say that a feasible point x^* verifies a second order punctual optimality condition, if x^* verifies KKT and PSD conditions or it does not verify CQ2, where PSD stands for positive semi-definiteness of the Hessian of the Lagrangian function on a tangent subspace. Weaker constraint qualifications are preferred since they provide stronger optimality conditions.

In this work, we are interested on the first class of optimality conditions mentioned, the sequential ones. We will study the sufficiency of AKKT conditions in convex optimization problems and extend its definition to a more general framework of continuous variational inequalities (VI). The main tool for this extension is the geometric interpretation of the general variational inequality as a nonlinear programming problem. The extension of the AGP sequential optimality condition to variational inequalities has been recently done in [9].

Algorithms defined to obtain a KKT point are more suited for VI than NLP. The reason is that the KKT system for NLP problems may not distinguish a local maximum from a minimum. An algorithm developed to handle the VI directly requires the use of a merit function, as, for example, the gap function. The main drawback is the evaluation of such function. As it was mentioned in the book [10], methods based on merit functions are effective in practice in the cases where the feasible set is relatively simple.

This paper is organized as follows. In Sect. 2 we define the Strong AKKT (SAKKT) condition for optimization problems and prove sufficiency for convex problems under a mild assumption. We also prove that AKKT with a particular choice of sequences is a strong optimality condition, since it implies the KKT condition under the quasi-normality constraint qualification. In Sect. 3 we define AKKT conditions for continuous variational inequalities and prove sufficiency for convex problems. In Sect. 4 we prove relations between AKKT conditions and the AGP condition. Final remarks are given in Sect. 5.

2 AKKT for Optimization

Consider the problem

$$\text{Minimize } f(x), \text{ subject to } x \in \Omega, \tag{1}$$

where

$$\Omega = \{x \in \mathbb{R}^n | h(x) = 0, g(x) \leq 0\}, \tag{2}$$

$f : \mathbb{R}^n \to \mathbb{R}, h : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^n \to \mathbb{R}^p$ are continuously differentiable functions.

We will denote $\mathbb{R}_+ = \{t \in \mathbb{R} | t > 0\}$. We say that the Approximate-Karush-Kuhn-Tucker condition (AKKT, [3]) is satisfied at a feasible point $x^* \in \Omega$ if, and only if, there exist sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^k\} \subset \mathbb{R}^m$, $\{\mu^k\} \subset \mathbb{R}^p_+$, such that $x^k \to x^*$,

$$\nabla f(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{j=1}^{p} \mu_j^k \nabla g_j(x^k) \to 0 \tag{3}$$

and

$$g_j(x^*) < 0 \Rightarrow \mu_j^k = 0 \text{ for sufficiently large } k. \tag{4}$$

This AKKT condition corresponds to AKKT(∅) of [3]. Given $I \subset \{1, \dots, p\}$, AKKT($I$) says that the sequence $\{x_k\}$ must also lie in the interior of the feasible set associated to I (this means that $g_i(x_k) < 0, \ \forall i \in I$). In [3], the authors proved that under a sufficient interior hypothesis, if x^* is a local minimizer, then there is a sequence that verifies the AKKT(I) condition. Here, we are interested on the AKKT(I) condition when $I = \emptyset$.

The AKKT condition is used as a natural stopping criterion to the Augmented Lagrangian solver Algencan [11] (available at www.ime.usp.br/~egbirgin/tango).

AKKT is a strong optimality condition since it implies the KKT condition under the constant positive linear dependence condition (CPLD, [7, 12]), which is a constraint qualification weaker than the constant rank [13] and the Mangasarian-Fromovitz condition [14]. Note that, if x^* is a minimum point to (1) such that a constraint qualification holds, then AKKT holds for constant sequences $x^k = x^*$, $\lambda^k = \lambda^*$, $\mu^k = \mu^*$, where $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^p_+$ are the Lagrange multipliers associated to x^*. Interesting cases arise when x^* is not a KKT point. Recently, [15, 16] a necessary optimality condition and a constraint qualification were defined by means of the image space analysis [17] for general constraint problems.

In [3], it was proved that AKKT(I) is an optimality condition if I satisfies a sufficient interior hypothesis. When $I = \emptyset$, the sufficient interior hypothesis is void and the proof is much simpler. We include it here for completeness.

Theorem 2.1 *If x^* is a local minimum point of* (1)–(2), *then x^* satisfies the AKKT condition* (3)–(4).

Proof Let $\delta > 0$ be such that $f(x^*) \le f(x)$ for every $x \in \Omega$, $\|x - x^*\|_2 \le \delta$. Then x^* is the unique global minimum point of the problem

$$\text{Minimize} \left[f(x) + \frac{1}{2}\|x - x^*\|_2^2 \right], \quad \text{subject to } x \in \Omega,\ \|x - x^*\|_2 \le \delta. \tag{5}$$

If $v \in \mathbb{R}^p$, we denote $v_+ = (\max\{v_1, 0\}, \dots, \max\{v_p, 0\})^T$. For every $k \in \mathbb{N}$, let x^k be a global minimum point of

$$\text{Minimize} \left[f(x) + \frac{1}{2}\|x - x^*\|_2^2 + \sum_{i=1}^m k h_i(x)^2 + \sum_{j=1}^p k g_j(x)_+^2 \right], \quad \text{subject to } \|x - x^*\|_2 \le \delta. \tag{6}$$

We have that $\{x^k\}$ is well defined since the objective function is continuous and the constraint set is non-empty and compact. Let us consider a limit point z of the sequence $\{x^k\}$. From the definition of x^k and the fact that $x^* \in \Omega$ is feasible to problem (6), we have that

$$f(x^k) \le f(x^k) + \frac{1}{2}\|x^k - x^*\|_2^2 + \sum_{i=1}^m k h_i(x^k)^2 + \sum_{j=1}^p k g_j(x^k)_+^2 \le f(x^*).$$

Taking limits for an appropriate subsequence yields $f(z) \le f(x^*)$. We will show that z is feasible for problem (5), which will enable us to conclude that $z = x^*$. From the definition of x^k, $\|x^k - x^*\|_2 \le \delta$, hence, we must have $\|z - x^*\|_2 \le \delta$. Ab absurdo, let us assume $\sum_{i=1}^m h_i(z)^2 + \sum_{j=1}^p g_j(z)_+^2 > 0$, then there exists $c > 0$ such that $\sum_{i=1}^m h_i(x^k)^2 + \sum_{j=1}^p g_j(x^k)_+^2 > c$ and $x^k \to z$ for an appropriate subsequence. Thus, since $x^* \in \Omega$, we may write

$$f(x^k)+\sum_{i=1}^{m}kh_i(x^k)+\sum_{j=1}^{p}kg_j(x^k)_+^2>f(x^*)+\sum_{i=1}^{m}kh_i(x^*)$$
$$+\sum_{j=1}^{p}kg_j(x^*)_+^2+kc+f(x^k)-f(x^*).$$

Since $f(x^k)\to f(z)$, we have that $kc+f(x^k)-f(x^*)>0$ for sufficiently large k, thus

$$f(x^k)+\sum_{i=1}^{m}kh_i(x^k)+\sum_{j=1}^{p}kg_j(x^k)_+^2>f(x^*)+\sum_{i=1}^{m}kh_i(x^*)+\sum_{j=1}^{p}kg_j(x^*)_+^2.$$

But x^* is feasible to problem (6), and this contradicts the definition of x^k. Therefore, z is feasible to (5) and $z=x^*$. Since $\{x^k\}$ is bounded, x^* is the unique limit point of $\{x^k\}$. We conclude that $x^k\to x^*$, thus $\|x^k-x^*\|_2<\delta$ for sufficiently large k, and from the definition of x^k we have that the gradient of the objective function of problem (6) at x^k must vanish:

$$\nabla f(x^k)+x^k-x^*+\sum_{i=1}^{m}kh_i(x^k)\nabla h_i(x^k)+\sum_{j=1}^{p}kg_j(x^k)_+\nabla g_j(x^k)=0.$$

Defining $\lambda_i^k=kh_i(x^k)$ and $\mu_j^k=kg_j(x^k)_+\geq 0$ we have (3). From the continuity of $g_j(x)$, (4) follows, which concludes the proof. □

It is interesting to notice that a stronger version of the AKKT condition is sufficient for optimality in convex problems. We will say that a feasible $x^*\in\Omega$ satisfies the Strong AKKT condition (SAKKT) if, and only if, there exist sequences $x^k\to x^*$, $\{\lambda^k\}\subset\mathbb{R}^m$, $\{\mu^k\}\subset\mathbb{R}_+^p$ such that (3) holds and

$$g_j(x^k)<0\Rightarrow \mu_j^k=0. \tag{7}$$

It is clear from the proof of Theorem 2.1 that SAKKT is also a sequential optimality condition, that is, every local minimizer satisfies SAKKT.

It is known that SAKKT is too strong for practical applications [2], since reasonable algorithms might fail to verify it. Consider for example the problem

$$\text{Minimize } x,\quad \text{subject to } -x\leq 0,$$

at the minimum point $x^*=0$. A reasonable algorithm might generate a sequence $x^k\to x^*$ with $x^k>0$ for every k, hence the complementarity condition (7) would imply $\mu^k=0$. Then, since $\nabla f(x^k)=1$, this sequence fails to satisfy (3). However, it is interesting to notice that SAKKT is sufficient for optimality for convex problems, under some assumptions.

Theorem 2.2 *Let f, g be convex functions and h affine. Let $x^*\in\Omega$ be such that SAKKT holds. If the sequences $\{x^k\}$, $\{\lambda^k\}$ are such that $\lambda_i^k h_i(x^k)\geq 0$ for every $i=1,\dots,m$ and for every $k\in\mathbb{N}$, then x^* is a minimum point of* (1)–(2).

Proof Let $x \in \Omega$ be an arbitrary feasible point. From the convexity assumption we have

$$f(x) \geq f(x^k) + \nabla f(x^k)^T (x - x^k), h_i(x) \\ = h_i(x^k) + \nabla h_i(x^k)^T (x - x^k), g_j(x) \geq g_j(x^k) + \nabla g_j(x^k)^T (x - x^k).$$

Multiplying accordingly by λ_i^k, μ_j^k and adding, since $h_i(x) = 0$ and $g_j(x) \leq 0$, we get:

$$\begin{aligned} f(x) &\geq f(x) + \sum_{i=1}^m \lambda_i^k h_i(x) + \sum_{j=1}^p \mu_j^k g_j(x) \geq f(x^k) + \sum_{i=1}^m \lambda_i^k h_i(x^k) + \sum_{j=1}^p \mu_j^k g_j(x^k) \\ &\quad + \left(\nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j=1}^p \mu_j^k \nabla g_j(x^k)\right)^T (x - x^k) \\ &\geq f(x^k) + \left(\nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j=1}^p \mu_j^k \nabla g_j(x^k)\right)^T (x - x^k). \end{aligned}$$

From $x^k \to x^*$ and (3), it follows that $f(x^*) \leq f(x)$, thus x^* is a minimum point. □

Note that for many iterative methods, to assume that $h(x^k) = 0$ for every k, when h is an affine function, is not a serious drawback and this implies $\lambda_i^k h_i(x^k) \geq 0$ for every i and k. This condition is also satisfied if there are no equality constraints, or if equality constraints are transformed into two inequalities in the obvious way. Note also that this condition can be replaced in Theorem 2.2 by a weaker one, that is, $\sum_{i=1}^m \lambda_i^k h_i(x^k) \geq 0$.

From the proof of Theorem 2.1, it can be seen that SAKKT with $\lambda_i^k h_i(x^k) \geq 0$ is a necessary optimality condition. Thus SAKKT with this choice of $\{x^k\}, \{\lambda^k\}$ is a necessary and sufficient condition for optimality in a convex problem. A common condition usually satisfied by the multipliers λ^k, μ^k is that:

$$\lambda^k = A_k h(x^k) \quad \text{and} \quad \mu^k = B_k g(x^k)_+, \; A_k > 0, B_k > 0, \tag{8}$$

for some real numbers A_k, B_k. This is the case, for example, when the external penalty method is applied. Notice that AKKT with this choice of multipliers is also an optimality condition.

We will show that the AKKT condition with this common choice (8) of multipliers is a stronger optimality condition. For this we will define the quasi-normality constraint qualification [18, 19], which is weaker than the CPLD [7].

Definition 2.1 (quasi-normality) We say that $x^* \in \Omega$ satisfies the quasi-normality constraint qualification if, and only if, for any couple of multipliers $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}_+^p$, satisfying

$$\sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{p} \mu_j \nabla g_j(x^*) = 0, \qquad \mu_j g_j(x^*) = 0, \tag{9}$$

there is no sequence $x^k \to x^*$ with the property:

$$\lambda_i \neq 0 \Rightarrow \lambda_i h_i(x^k) > 0 \tag{10}$$

and

$$\mu_j > 0 \Rightarrow \mu_j g_j(x^k) > 0. \tag{11}$$

Quasi-normality is not the weakest first order constraint qualification. The Abadie constraint qualification [20], defined using the tangent and linearized tangent cones, is weaker than quasi-normality.

Theorem 2.3 *Let $x^* \in \Omega$ be such that AKKT holds with sequences $\{x^k\}, \{\lambda^k\}$ and $\{\mu^k\}$ satisfying (8). If x^* satisfies quasi-normality then x^* is a KKT point of problem (1)–(2).*

Proof Suppose there exist sequences $x^k \to x^*$, $\{\lambda^k\} \subset \mathbb{R}^m$, $\{\mu^k\} \subset \mathbb{R}^p_+$, such that (3), (4) and (8) hold. Let

$$\delta_k = \sqrt{1 + \sum_{i=1}^{m} (\lambda_i^k)^2 + \sum_{j=1}^{p} (\mu_j^k)^2}.$$

Since $\|(\frac{1}{\delta_k}, \frac{\lambda^k}{\delta_k}, \frac{\mu^k}{\delta_k})\|_2 = 1$ for every k, we may take a subsequence such that $(\frac{1}{\delta_k}, \frac{\lambda^k}{\delta_k}, \frac{\mu^k}{\delta_k}) \to (\mu_0^*, \lambda^*, \mu^*) \neq 0$ and $\mu_j^* \geq 0$ for every $j = 0, 1, \ldots, p$. Dividing the expression in (3) by δ_k and taking limits we get:

$$\mu_0^* \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^{p} \mu_j^* \nabla g_j(x^*) = 0.$$

Since $\delta_k > 0$, we have:

$$\lambda_i^* \neq 0 \Rightarrow \lambda_i^* \lambda_i^k > 0$$

and

$$\mu_j^* > 0 \Rightarrow \mu_j^* \mu_j^k > 0.$$

Thus, if $\mu_0^* = 0$, from (8), quasi-normality would fail. We conclude that $\mu_0^* > 0$. Complementarity follows from (4), hence x^* is a KKT point. □

Note that AKKT with the choice (8) of multipliers implies the SAKKT condition and $\lambda_i^k h_i(x^k) \geq 0$, thus AKKT and (8) is also necessary and sufficient for optimality on a convex problem.

3 AKKT for Continuous Variational Inequalities

In this section we will extend the AKKT condition to continuous variational inequalities, in the same fashion as the AGP condition [1] has been extended in [9].

Let $F:\mathbb{R}^n\to\mathbb{R}^n$ be a continuous function, $\Omega\subset\mathbb{R}^n$ defined as in (2) and $\langle\cdot,\cdot\rangle$ the Euclidean inner product. The continuous variational inequality problem VI(F,Ω) is to find $x^*\in\Omega$, such that

$$\langle F(x^*), x-x^*\rangle\geq 0,\ \forall x\in\Omega. \tag{12}$$

If a solution x^* satisfies any constraint qualification (such as Mangasarian-Fromovitz [14], CPLD [7, 12], quasi-normality [18, 19], etc.), it can be shown that the Karush-Kuhn-Tucker condition for variational inequalities holds, that is, there exist $\lambda\in\mathbb{R}^m$, $\mu\in\mathbb{R}^p_+$ such that

$$F(x^*)+\sum_{i=1}^m\lambda_i\nabla h_i(x^*)+\sum_{j=1}^p\mu_j\nabla g_j(x^*)=0 \tag{13}$$

and

$$g_j(x^*)<0\Rightarrow\mu_j=0. \tag{14}$$

This is a consequence of the usual KKT condition for optimization, since x^* is a solution to VI(F,Ω) if, and only if x^* is a minimum point to the nonlinear programming problem

$$\text{Minimize } \langle F(x^*),x\rangle,\quad \text{subject to } x\in\Omega. \tag{15}$$

We define the AKKT condition to VI(F,Ω) in a natural way: we say that $x^*\in\Omega$ satisfies AKKT if, and only if, there exist sequences $x^k\to x^*$, $\{\lambda^k\}\subset\mathbb{R}^m$, $\{\mu^k\}\subset\mathbb{R}^p_+$ such that

$$F(x^k)+\sum_{i=1}^m\lambda_i^k\nabla h_i(x^k)+\sum_{j=1}^p\mu_j^k\nabla g_j(x^k)\to 0, \tag{16}$$

and

$$g_j(x^*)<0\Rightarrow\mu_j^k=0. \tag{17}$$

Theorem 3.1 *If x^* is a solution to VI(F,Ω), then x^* satisfies the AKKT condition* (16)–(17).

Proof Since x^* is a solution to (15), then by Theorem 2.1 there exist sequences $x^k\to x^*$, $\{\lambda^k\}\subset\mathbb{R}^m$, $\{\mu^k\}\subset\mathbb{R}^p_+$ such that (17) holds and

$$F(x^*)+\sum_{i=1}^m\lambda_i^k\nabla h_i(x^k)+\sum_{j=1}^p\mu_j^k\nabla g_j(x^k)\to 0, \tag{18}$$

since $F(x^k)\to F(x^*)$, we conclude that (16) holds. □

In the same fashion, we will say that SAKKT is satisfied for VI(F, Ω) when there exist sequences such that (17) and (7) hold. Now we will prove that SAKKT is sufficient for $x^* \in \Omega$ to be a solution to VI(F, Ω), under the same mild assumption of Theorem 2.2.

Theorem 3.2 *Let g be a convex function and h affine. If $x^* \in \Omega$ satisfies SAKKT to VI(F, Ω) and the sequences $\{x^k\}, \{\lambda^k\}$ are such that $\lambda_i^k h_i(x^k) \geq 0$ for every $i = 1, \ldots, m$ and for every $k \in \mathbb{N}$, then x^* is a solution to VI(F, Ω).*

Proof Suppose there exist sequences $x^k \to x^*$, $\{\lambda^k\} \subset \mathbb{R}^m$, $\{\mu^k\} \subset \mathbb{R}_+^p$ such that $\lambda_i^k h_i(x^k) \geq 0$, (16) and (7) hold. Clearly, x^* satisfies SAKKT for the optimization problem (15), then from Theorem 2.2, x^* is a minimum point to (15), hence a solution to VI(F, Ω). □

Observe that, differently from the optimization case, for the convex VI(F, Ω) it is not necessary to impose convexity of F. The main reason is that we are considering VI(F, Ω) in its reformulation as the nonlinear convex problem (15).

Some known NLP methods have been extended for the finite dimensional variational inequality with feasible set defined by convex inequalities and where F is a maximal monotone operator; see, for example [21, 22] and references therein. This shows the importance of the convex case.

4 The Relation with the AGP Condition

In [1], Martínez and Svaiter introduced the Approximate Gradient Projection optimality condition (AGP) for optimization problems, which gives a natural stopping criterion for the inexact restoration method [5, 6, 23]. In [9] AGP has been extended to the context of variational inequalities. We state the AGP condition according to [9] for the VI. If $v \in \mathbb{R}^p$, we denote $v_- = (\min\{v_1, 0\}, \ldots, \min\{v_p, 0\})^T$.

Definition 4.1 (AGP) Given $\gamma \in [-\infty, 0]$, a feasible point $x^* \in \Omega$ satisfies the AGP(γ) condition for VI(F, Ω) if, and only if, there exists a sequence $x^k \to x^*$ such that

$$P_{\Omega_\gamma(x^k)}(x^k - F(x^k)) - x^k \to 0,$$

where $P_X(\cdot)$ is the Euclidean projection on the closed convex set X and

$$\Omega_\gamma(x^k) = \left\{ z \in \mathbb{R}^n \,\middle|\, \begin{array}{ll} g_j(x^k)_- + \nabla g_j(x^k)^{\mathrm{T}}(z - x^k) \leq 0 & \text{if } \gamma < g_j(x^k) \\ \nabla h_i(x^k)^{\mathrm{T}}(z - x^k) = 0 & \end{array} \right\}.$$

For the convex VI(F, Ω), a similar result to our Theorem 3.2 was obtained using the AGP condition, see Corollary 3.2 of [9]. There, the authors used the original definition of the problem and a different hypothesis for the equality constraints. Thus, in the convex VI(F, Ω) with only inequality constraints, both theorems establish the same thesis: if x^* satisfies SAKKT (respectively AGP) then x^* is a solution of the

VI(F, Ω). This motivates us to find the relation between SAKKT and AGP. We will prove that SAKKT ⇒ AGP ⇒ AKKT for continuous variational inequalities, which generalize the results in [2], where these were proved for optimization problems.

Theorem 4.1 *Given $\gamma \in [-\infty, 0]$, if $x^* \in \Omega$ satisfies AGP(γ), then x^* satisfies AKKT.*

Proof Let $x^k \to x^*$ be such that $y^k = P_{\Omega_\gamma(x^k)}(x^k - F(x^k))$ satisfies $y^k - x^k \to 0$. From the definition of the Euclidean projection, y^k solves the problem:

$$\text{Minimize } \frac{1}{2}\|y - (x^k - F(x^k))\|_2^2, \quad \text{subject to } y \in \Omega_\gamma(x^k). \tag{19}$$

From the linearity of the constraints, y^k satisfies the KKT condition for (19), that is, there exist $\{\lambda^k\} \subset \mathbb{R}^m$, $\{\mu^k\} \subset \mathbb{R}^p_+$ such that

$$y^k - (x^k - F(x^k)) + \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j=1}^p \mu_j^k \nabla g_j(x^k) = 0 \tag{20}$$

$$\mu_j^k (g_j(x^k)_- + \nabla g_j(x^k)^T (y^k - x^k)) = 0 \quad \text{if } \gamma < g_j(x^k) \tag{21}$$

$$\mu_j^k = 0 \text{ otherwise.} \tag{22}$$

Since $y^k - x^k \to 0$, (20) implies (16). To prove (17), suppose $g_j(x^*) < 0$, then, by Taylor expansion, $g_j(x^k)_- + \nabla g_j(x^k)^T (y^k - x^k) < 0$ for sufficiently large k, thus, from (21) and (22) we conclude $\mu_j^k = 0$. □

Remark Since AGP(γ) is satisfied at a solution of VI(F,Ω) (see [9]), this alternatively proves Theorem 3.1.

Theorem 4.2 *If $x^* \in \Omega$ satisfies SAKKT, then x^* satisfies AGP(γ) for every $\gamma \in [-\infty, 0]$.*

Proof From (16) and (7) we can write

$$x^k - F(x^k) - \left(x^k + \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j: g_j(x^k) \geq 0} \mu_j^k \nabla g_j(x^k) \right) \to 0,$$

thus, from the non-expansiveness of the Euclidean projection:

$$P_{\Omega_\gamma(x^k)}(x^k - F(x^k)) - P_{\Omega_\gamma(x^k)} \left(x^k + \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j: g_j(x^k) \geq 0} \mu_j^k \nabla g_j(x^k) \right) \to 0.$$

Writing the projection problem for the second projection, we conclude that it is equal to x^k, since x^k is a KKT point and that is sufficient for optimality given the convexity of the problem (see the proof of Theorem 1 in [1]). This concludes the proof. □

Now we will show that a reciprocal of Theorem 4.2 is true for the case $\gamma = 0$.

Theorem 4.3 *AGP*(0) *is equivalent to SAKKT.*

Proof Let $x^* \in \Omega$ be such that there exists a sequence $x^k \to x^*$ with $y^k = P_{\Omega_\gamma(x^k)}(x^k - F(x^k))$ satisfying $y^k - x^k \to 0$. Then y^k is a solution to (19). As in the proof of Theorem 4.1, from the linearity of the constraints, y^k satisfies the KKT condition (20)–(22) with $\gamma = 0$. This clearly implies (16) and (7). □

5 Final Remarks

In this work, we presented necessary and sufficient sequential conditions for optimality in convex optimization problems. In the general non-convex case, we presented strong practical optimality conditions with an appealing theoretical result: they guarantee the KKT condition at a solution imposing only the weak quasi-normality constraint qualification. We also showed that these sequential conditions can be generalized to continuous variational inequalities, and we proved the relationship between these conditions and the Approximate Gradient Projection condition for VI of Gárciga-Otero and Svaiter [9].

We realize that, using the reformulated problem, we are not able to exploit all the good properties that the original problem might possess. In spite of this limitation, the simplicity of the algorithms based on the KKT condition makes them very attractive and, if the constraints of the VI are not particularly simple, solving the KKT condition can be the preferred approach for solving the VI in practice.

We believe that sequential conditions can generate new algorithms for variational inequalities, as well as for optimization problems. Thus, in our opinion, further investigation on this topic is needed to provide promising alternative numerical schemes for solving variational inequalities.

References

1. Martínez, J.M., Svaiter, B.F.: A practical optimality condition without constraint qualifications for nonlinear programming. J. Optim. Theory Appl. **118**, 117–133 (2003)
2. Schuverdt, M.L.: Métodos de Lagrangiano aumentado com convergência utilizando a condição de dependência linear positiva constante. Ph.D. Thesis, UNICAMP (2006)
3. Andreani, R., Haeser, G., Martínez, J.M.: On sequential optimality conditions for smooth constrained optimization. Optimization (2010). doi:10.1080/02331930903578700
4. Haeser, G.: Condições sequenciais de otimalidade. Ph.D. Thesis, UNICAMP (2009)
5. Martínez, J.M., Pilotta, E.A.: Inexact restoration algorithms for constrained optimization. J. Optim. Theory Appl. **104**, 135–163 (2000)
6. Martínez, J.M., Pilotta, E.A.: Inexact restoration methods for nonlinear programming: advances and perspectives. In: Qi, L.Q., Teo, K.L., Yang, X.Q. (eds.) Optimization and Control with Applications, pp. 271–292. Springer, Berlin (2005)
7. Andreani, R., Martínez, J.M., Schuverdt, M.L.: On the relation between the Constant Positive Linear Dependence condition and quasinormality constraint qualification. J. Optim. Theory Appl. **125**, 473–485 (2005)
8. Andreani, R., Martínez, J.M., Schuverdt, M.L.: On second-order optimality conditions for nonlinear programming. Optimization **56**, 529–542 (2007)

9. Gárciga-Otero, R., Svaiter, B.F.: A new condition characterizing solutions of variational inequality problems. J. Optim. Theory Appl. **137**, 89–98 (2008)
10. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer, New York (2003)
11. Andreani, R., Birgin, E.G., Martínez, J.M., Schuverdt, M.L.: On augmented Lagrangian methods with general lower–level constraints. SIAM J. Control Optim. **18**, 1286–1309 (2007)
12. Qi, L., Wei, Z.: On the constant positive linear dependence condition and its application to SQP methods. SIAM J. Control Optim. **10**, 963–981 (2000)
13. Janin, R.: Direction derivative of the marginal function in nonlinear programming. Math. Program. Stud. **21**, 127–138 (1984)
14. Mangasarian, O.L., Fromovitz, S.: The Fritz-John necessary optimality conditions in presence of equality and inequality constraints. J. Math. Anal. Appl. **17**, 37–47 (1967)
15. Moldovan, A., Pellegrini, L.: On regularity for constrained extremum problem. Part 1: Sufficient optimality conditions. J. Optim. Theory Appl. **142**, 147–163 (2009)
16. Moldovan, A., Pellegrini, L.: On regularity for constrained extremum problem. Part 2: Necessary optimality conditions. J. Optim. Theory Appl. **142**, 165–183 (2009)
17. Giannessi, F.: Constrained Optimization and Image Space Analysis, Separation of Sets and Optimality Conditions, vol. 1. Springer, New York (2005)
18. Hestenes, M.R.: Optimization Theory—The Finite-Dimensional Case. Wiley, New York (1975)
19. Bertsekas, D.P.: Nonlinear Programming, 2nd edn. Athena Scientific, Belmont (1999)
20. Abadie, J.: On the Kuhn-Tucker theorem. In: Abadie, J. (ed.) Nonlinear Programming, pp. 21–36. North-Holland, Amsterdam (1967)
21. Iusem, A.N., Nasri, M.: Augmented Lagrangian methods for variational inequality problems. RAIRO. Rech. Opér. **44**, 5–25 (2010)
22. Auslender, A., Teboulle, M.: Lagrangian duality and related multiplier methods for variational inequality problems. SIAM J. Control Optim. **10**, 1097–1115 (1999)
23. Martínez, J.M.: Inexact restoration method with Lagrangian tangent decrease and new merit function for nonlinear programming. J. Optim. Theory Appl. **111**, 39–58 (2001)