A cone-continuity constraint qualification and algorithmic consequences *

Roberto Andreani[†] José Mário Martinez[‡] Alberto Ramos[§] Paulo J. S. Silva[¶]

February 12th, 2015

Abstract

Every local minimizer of a smooth constrained optimization problem satisfies the sequential Approximate Karush-Kuhn-Tucker (AKKT) condition. This optimality condition is used to define the stopping criteria of many practical nonlinear programming algorithms. It is natural to ask for conditions on the constraints under which AKKT implies KKT. These conditions will be called Strict Constraint Qualifications (SCQ). In this paper we define a Cone-Continuity Property (CCP) that will be showed to be the weakest possible SCQ. Its relation with other constraint qualifications will also be clarified. In particular, it will be proved that CCP is strictly weaker than the Constant Positive Generator (CPG) constraint qualification.

Keywords: Constrained optimization, Optimality conditions, Constraint qualifications, KKT conditions, Approximate KKT conditions.

1 Introduction

We will consider Constrained Optimization problems defined by

$$\text{Minimize } f(x) \text{ subject to } h(x) = 0, g(x) \leq 0, \tag{1.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}, h : \mathbb{R}^n \to \mathbb{R}^m$, and $g : \mathbb{R}^n \to \mathbb{R}^p$ admit continuous first derivatives onto $\mathbb{R}^n$.

Many scientific and technological problems require the solution of problems of this form. Finding global solutions of (1.1) is possible only when the problem is small or has some special structure. Even the verification that a given feasible point is a solution may be very hard. For this reason one relies on necessary optimality conditions. By this we mean computable conditions that must be verified by the minimizers of (1.1) and whose fulfillment indicate that, very likely, the point under consideration is an acceptable (perhaps approximate) solution of (1.1).

*This work was supported by PRONEX-Optimization (PRONEX-CNPq/FAPERJ E-26/111.449/2010-APQ1), CEPID-CeMEAI (FAPESP 2013/07375-0), CNPq (Grants 304618/2013-6, 482549/2013-0, and 303013/2013-3), FAPESP (Grants 2012/20339-0 and 2013/05475-7), CAPES, and National Project on Industrial Mathematics.

†Department of Applied Mathematics, Institute of Mathematics, Statistics and Scientific Computing, University of Campinas, Campinas, SP, Brazil. Email: andreani@ime.unicamp.br.

‡Department of Applied Mathematics, Institute of Mathematics, Statistics and Scientific Computing, University of Campinas, Campinas, SP, Brazil. Email: martinez@ime.unicamp.br.

§Institute of Mathematics and Statistics, University of São Paulo, São Paulo, SP, Brazil. Email: aramos27@gmail.com.

¶Department of Applied Mathematics, Institute of Mathematics, Statistics and Scientific Computing, University of Campinas, Campinas, SP, Brazil. Email: pjssilva@ime.unicamp.br.

A point $x \in \mathbb{R}^n$ is said to satisfy the KKT conditions related to (1.1) if there exist $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$ such that

$$\nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x) + \sum_{i=1}^{p} \mu_i \nabla g_i(x) = 0, \tag{1.2}$$

$$h_i(x) = 0 \text{ for all } i = 1, \dots, m, \tag{1.3}$$

and

$$\min\{\mu_i, -g_i(x)\} = 0 \text{ for all } i = 1, \dots, p. \tag{1.4}$$

The condition (1.4) implies that $g(x) \leq 0$, $\mu \geq 0$, and $\mu_i = 0$ for all i such that $g_i(x) < 0$.

Given $x \in \mathbb{R}^n$, it is easy to check the existence of $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$ satisfying (1.2), (1.3), and (1.4). Unfortunately, the KKT conditions are not necessarily satisfied by minimizers of (1.1). For example $x^* = 0$ is a global minimizer of x subject to $x^2 = 0$ but there are no multipliers λ, μ that fulfill the KKT conditions for $x = x^*$. The properties of the constraints that guarantee that minimizers of the Constrained Optimization problem satisfy the KKT conditions are called *Constraint Qualifications* (CQ): If x is a local minimizer of (1.1) and some constraint qualification is fulfilled at x, then the KKT conditions are satisfied for appropriate multipliers $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$. In other words, if CQ is a constraint qualification, the property

$$\text{KKT or Not-CQ} \tag{1.5}$$

is fulfilled at every local minimizer of (1.1).

Obviously, necessary optimality conditions should be as strong as possible. Moreover, strength of (1.5) is linked to the weakness of the constraint qualification. The most popular constraint qualification is the linear independence of the gradients of active constraints (LICQ). Its attractiveness is due to two independent properties: On the one hand, LICQ is easily verifiable, and, on the other hand, it can be associated with many practical optimization algorithms, for which it can be proved that convergence occurs to points that satisfy "KKT or Not-LICQ". The Mangasarian-Fromovitz constraint qualification (MFCQ), which states that the gradients of active constraints are "positively linearly independent" at the feasible point under consideration, is obviously weaker than LICQ [18, 22]. Qi and Wei [21] introduced the "Constant Positive Linear Dependence" (CPLD) condition, which says that if some gradients of active constraints are positively linearly dependent at a point x, then the same gradients are linearly dependent in a neighborhood of x. They also showed that a particular Sequential Quadratic Programming algorithm converges to points that satisfy "KKT or not-CPLD". Curiously, Qi and Wei did not prove that CPLD was a constraint qualification. This property of CPLD was proved by Andreani, Martínez, and Schuverdt [8], who also described the status of CPLD with respect to other constraint qualifications proving that the new condition implies quasinormality. CPLD is weaker than MFCQ and is necessarily satisfied if the constraints of the problem are linear (a property that is not shared by MFCQ). This motivated a sequence of papers in which weaker constraint qualifications were introduced, with proved association with practical algorithms. See [4, 5] and references therein . This effort seemed to come to an end with the introduction of the Constant Positive Generators (CPG) constraint qualification in [5]. The definition of the CPG constraint qualification is the following.

Definition 1.1. Assume that $h(x^*) = 0$ and $g(x^*) \leq 0$. Define $I = \{1, \dots, m\}$. Let $J(x^*) \subset \{1, \dots, p\}$ be the indices of the active inequality constraints at x^*. Let J_- be set of indices $\ell \in J(x^*)$ such that, for all $\ell \in J_-$, there exist $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ and $\mu_j \in \mathbb{R}_+$ for all $j \in J(x^*)$, such that

$$-\nabla g_\ell(x^*) = \sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) + \sum_{j \in J(x^*)} \mu_j \nabla g_j(x^*). \tag{1.6}$$

Define $J_+ = J(x^*) - J_-$. We say that the Constant Positive Generator (CPG) condition holds at x^* if there exist (maybe empty) sets $I' \subset I$ and $J' \subset J_-$ such that

(i) The gradients $\nabla h_i(x^*)$ and $\nabla g_j(x^*)$ indexed by $i \in I'$ and $j \in J'$ are linearly independent.

(ii) For all x in a neighborhood of x^*, if

$$z = \sum_{i=1}^{m} \lambda'_i \nabla h_i(x) + \sum_{j \in J(x^*)} \mu'_j \nabla g_j(x),$$

with $\mu'_j \geq 0$ for all $j \in J(x^*)$, then for all $i \in I'$, $\ell \in J'$, and $j \in J_+$, there exist $\lambda''_i \in \mathbb{R}$, $\lambda'''_\ell \in \mathbb{R}$, and $\mu''_j \in \mathbb{R}_+$ such that

$$z = \sum_{i \in I'} \lambda''_i \nabla h_i(x) + \sum_{\ell \in J'} \lambda'''_\ell \nabla g_\ell(x) + \sum_{j \in J_+} \mu''_j \nabla g_j(x).$$

Remark 1. The item (i) of Definition 1.1 above is equivalent to state that the gradients $\{\nabla h_i(x^*), i \in I\}$; $\{\nabla g_l(x^*), l \in J'\}$ and $\{\nabla g_j(x^*), j \in J_+\}$ form a positive linear independent set, that is, the only solution of

$$\sum_{i \in I'} \lambda_i \nabla h_i(x) + \sum_{\ell \in J'} \gamma_\ell \nabla g_\ell(x) + \sum_{j \in J_+} \mu_j \nabla g_j(x) = 0,$$

with $\lambda_i \in \mathbb{R}, i \in I'$, $\gamma_\ell \in \mathbb{R}, \ell \in J'$ and $\mu_j \geq 0, j \in J_+$ is the trivial one.

In [5] it was proved that CPG is a constraint qualification, strictly weaker than CPLD and some of its relaxations, and that it is useful in the context of several algorithms, for which it can be proved that limit points that satisfy CPG are KKT points.

The question that arose so far is: Is CPG the weakest constraint qualification that satisfy the above properties? Before going to the answer of this question we need to formulate it more precisely.

The second "nice property" of constraint qualifications is the association with practical algorithms. This association consists of the possibility of proving that, under a given CQ, the algorithm converges to KKT points. Some constraint qualifications have this property and others do not. This leads us to discuss the notion of *Sequential Optimality Condition* [19, 3, 9]. To fix ideas we will consider the most popular of these conditions, called *Approximate KKT* (AKKT) [21, 3, 10].

Definition 1.2. Assume that $h(x^*) = 0$ and $g(x^*) \leq 0$. We say that x^* satisfies AKKT if there exist sequences $\{x^k\} \subset \mathbb{R}^n$ ($\{x^k\}$ is called AKKT sequence), $\{\lambda^k\} \subset \mathbb{R}^m$, and $\{\mu^k\} \subset \mathbb{R}^p$ such that $\lim_{k\to\infty} x^k = x^*$,

$$\lim_{k\to\infty} \nabla f(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{i=1}^{p} \mu_i^k \nabla g_i(x^k) = 0, \tag{1.7}$$

and

$$\lim_{k\to\infty} \min\{\mu_i^k, -g_i(x^k)\} = 0 \text{ for all } i = 1, \ldots, p. \tag{1.8}$$

AKKT, as other sequential optimality conditions, has two main properties. The first is that it is a genuine necessary optimality condition, independently of the fulfillment of constraint qualifications [3, 10]. The second is that many optimization algorithms (but not all, see [7]) generate primal-dual $\{x^k, \lambda^k, \mu^k\}$ sequences for which (1.7) and (1.8) are fulfilled. These properties motivate the definition of *Strict Constraint Qualifications* (SCQ). A Strict Constraint Qualification is a property of feasible points of constrained optimization problem that, when satisfied by an AKKT point, guarantee that such point satisfies the KKT conditions [10]. In other words, Strict Constraint Qualifications are characterized by the property

$$AKKT + SCQ \Rightarrow KKT. \tag{1.9}$$

Since all local minimizers satisfy AKKT, the property (1.9) implies that strict constraint qualifications are, in fact, constraint qualifications. The reciprocal is not true. For instance, Abadie's CQ [1] or quasinormality [11] are constraint qualifications that are not strict constraint qualifications.

Now we are able to define precisely what we mean by "constraint qualifications associated with algorithms". Essentially, those constraint qualifications are the Strict Constraint Qualifications. The question about the weakness of CPG can be formulated as: Is CPG the weakest strict constraint qualification?

An attentive reader could argue in the following way: Being AKKT a genuine necessary optimality condition obviously associated with usual stopping criteria of practical algorithms, why should one worry with the constraint qualifications under which AKKT implies KKT (which, in fact, is not a genuine optimality condition)? The reason is that, in some important families of problems, AKKT is not strong enough for providing reliable detection of optimality. In such situations, the aditional fulfillment of KKT provides a sort of confirmation that we are approaching true minimizers. A typical example comes from considering the minimization of x_2 subject to $x_1 \geq 0$ and $x_1 x_2 = 0$ and the feasible point $x^* = (0, 1)$. This point is neither a minimizer nor a KKT point, but it satisfies AKKT according to Definition 1.2. The point x^* does not satisfy other sequential optimality conditions [3].

In this paper we will prove that CPG is not the weakest Strict Constraint Qualification. The weakest SCQ will be completely characterized as being the continuity of the cone generated by the gradients of active constraints (Cone Continuity Property CCP) and we will prove that CPG is strictly stronger than CCP [1].

As a consequence of these results we are able to present an updated landscape of constraint qualifications, strict constraint qualifications and sequential optimality conditions. Open questions remain, that will be probably the subject of future research.

This paper is organized as follows. In Section 2, we review basic definitions of optimization and variational analysis and introduce our principal object of study (2.11). In Section 3, we introduce the Cone-Continuity Property (CCP) and we prove that CCP is the weakest Strict Constraint Qualification. Section 4 shows the relationship between the Cone Continuity Property and other constraint qualifications as Abadie's CQ and quasinormality. Concluding remarks are discussed in Section 5.

2 Basic Definitions and Preliminaries

Our notation is standard in optimization and variational analysis; cf. [23, 13, 20]. $\mathbb{N}$ denotes the set of natural numbers, $\mathbb{R}^n$ stands for the n-dimensional real Euclidean space, $n \in \mathbb{N}$. We denote by $\mathbb{B}$ the closed unit ball in $\mathbb{R}^n$, and $\mathbb{B}(x, \eta) := x + \eta\mathbb{B}$ the closed ball centered at x with radius $\eta > 0$. $\mathbb{R}_+$ the set of positive scalars, $a^+ = \max\{0, a\}$, the positive part of $a \in \mathbb{R}$. We use $\langle \cdot, \cdot \rangle$ to denote the Euclidean inner product, $\|\cdot\|$ the associated norm. Given a set-valued mapping (multifunction) $F : \mathbb{R}^s \rightrightarrows \mathbb{R}^d$, the *sequential Painlevé-Kuratowski outer/upper limit* of $F(z)$ as $z \to z^*$ is denoted by

$$\limsup_{z \to z^*} F(z) := \{w^* \in \mathbb{R}^d : \exists \ \ (z^k, w^k) \to (z^*, w^*) \text{ with } w^k \in F(z^k)\} \tag{2.1}$$

and the inner limit by

$$\liminf_{z \to z^*} F(z) := \{w^* \in \mathbb{R}^d : \forall z^k \to z^* \ \ \exists \ \ w^k \to w^* \text{ with } w^k \in F(z^k)\}. \tag{2.2}$$

We say that F is *outer semicontinuous* (osc) at z^* if

$$\limsup_{z \to z^*} F(z) \subset F(z^*). \tag{2.3}$$

[1] This result was announced in the Workshop of Continuous Optimization held in Florianópolis in February 2014 [6] and cited in the book [10]

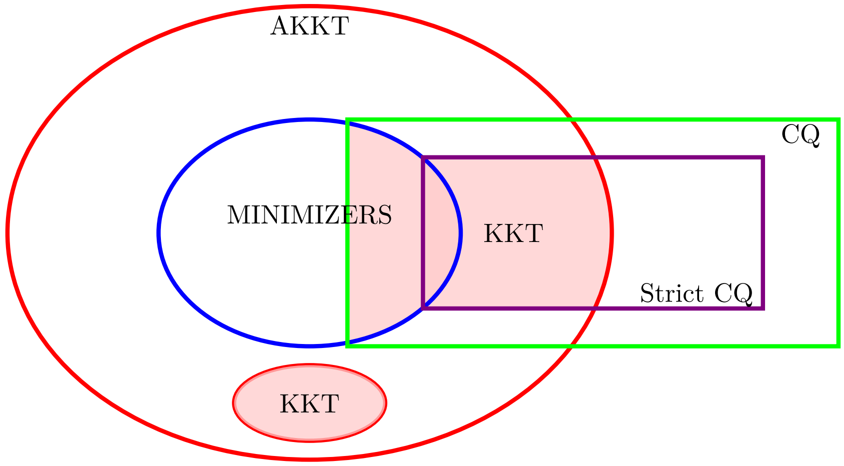


Figure 1: **Optimality Relations:** *Small ellipse:* Local minimizers. *Big ellipse:* Feasible points that satisfy the AKKT condition. *Small rectangle:* Feasible points that satisfy a Strict Constraint qualification. *Big rectangle:* Feasible points that satisfy a (not necessarily Strict) Constraint qualifications. *Shadowed area:* KKT points.

It is *inner semicontinuous* (isc) at z^* if

$$F(z^*) \subset \liminf_{z \to z^*} F(z). \tag{2.4}$$

If F is both outer semicontinuous and inner semicontinuous at z^* we say that F is *continuous* at z^*.

Given the set S, the symbol $z \xrightarrow{S} z^*$ means that $z \to z^*$ with $z \in S$. For a cone $\mathcal{K} \subset \mathbb{R}^s$, its polar (negative dual) is $\mathcal{K}^\circ = \{v \in \mathbb{R}^s | \langle v, k\rangle \leq 0 \text{ for all } k \in \mathcal{K}\}$. We use the notation $\phi(t) \leq o(t)$ for any function $\phi : \mathbb{R}_+ \to \mathbb{R}^s$ such that $\limsup_{t \to 0_+} t^{-1}\phi(t) \leq 0$.

Given $S \subset \mathbb{R}^n$ and $z^* \in S$, define the (Bouligand-Severi) *tangent/contingent cone* to S at z^* by

$$T_S(z^*) := \limsup_{t \downarrow 0} \frac{S - z^*}{t} = \{d \in \mathbb{R}^n : \exists \;\; t_k \downarrow 0, d^k \to d \text{ with } z^* + t_k d^k \in S\}. \tag{2.5}$$

The (Fréchet) *regular normal cone* to S at $z^* \in S$

$$\widehat{N}_S(z^*) := \{w \in \mathbb{R}^n : \langle w, z - z^*\rangle \leq o(|z - z^*|) \text{ for } z \in S\} = T_S^\circ(z^*). \tag{2.6}$$

The (Mordukhovich) *limiting normal cone* to S at $x^* \in S$

$$N_S(z^*) := \limsup_{z \xrightarrow{S} z^*} \widehat{N}_S(z). \tag{2.7}$$

When S is a convex set, both regular and limiting normal cones reduce to the classical normal cone of convex analysis and then the common notation $N_S(z^*)$ is used. For general sets we have the inclusion $\widehat{N}_S(z^*) \subset N_S(z^*)$ for all $z^* \in S$.

Denote by Ω the feasible set associated with (1.1), $\Omega := \{x \in \mathbb{R}^n | \, h(x) = 0, g(x) \leq 0\}$. Let $J(x^*)$ be the set of indices of active inequality constraints, and let $r = |J(x^*)|$ be the number of elements of $J(x^*)$. Let $x^* \in \Omega$ be a local minimizer of (1.1). The geometrical first-order necessary optimality condition states that $\langle \nabla f(x^*), d\rangle \geq 0$ for all $d \in T_\Omega(x^*)$. In other words,

$$-\nabla f(x^*) \in T_\Omega(x^*)^\circ. \tag{2.8}$$

Let us denote $I = \{1, \ldots, n\}$. The *linearized cone* $L_\Omega(x^*)$ is defined as follows.

$$L_\Omega(x^*) := \{d \in \mathbb{R}^n \mid \langle \nabla h_i(x^*), d\rangle = 0,\ \forall i \in I, \langle \nabla g_j(x^*), d\rangle \leq 0,\ \forall j \in J(x^*)\}. \tag{2.9}$$

By the geometric first-order necessary optimality condition (2.8), if x^* satisfies

$$T_\Omega(x^*)^\circ = L_\Omega(x^*)^\circ, \tag{2.10}$$

then the KKT conditions hold at x^*. The condition (2.10) was introduced by Guignard [15]. Gould and Tolle [16] proved that Guignard's condition (2.10) is the weakest constraint qualification. Abadie's constraint qualification, which reads $L_\Omega(x^*) = T_\Omega(x^*)$, is stronger than Guignard's CQ but in many cases is preferred in pratice because it does not involve the polar operation.

Given $x^* \in \Omega$, we define:

$$K(x) = \left\{ \sum_{i=1}^{m} \lambda_i \nabla h_i(x) + \sum_{j \in J(x^*)} \mu_j \nabla g_j(x) : \quad \mu_j \in \mathbb{R}_+, \lambda_i \in \mathbb{R} \right\} \tag{2.11}$$

Clearly, $K(x)$ is a closed convex cone and coincides with $L_\Omega(x^*)^\circ$ at x^*. Using this cone, the KKT condition can now be written $-\nabla f(x^*) \in K(x^*)$. The cone (2.11) also allows us re-write the CPG condition in a geometric setting. For this purpose, define the cone:

$$K_{I',J'}(x) = \left\{ \sum_{i \in I'} \lambda_i \nabla h_i(x) + \sum_{\ell \in J'} \gamma_\ell \nabla g_\ell(x) + \sum_{j \in J_+} \mu_j \nabla g_j(x) \colon\ \mu_j \in \mathbb{R}_+, \lambda_i, \gamma_\ell \in \mathbb{R} \right\} \tag{2.12}$$

where $I' \subset I$, $J' \subset J_-$ and $J_+ = J(x^*) - J_-$.

The re-statement of the CPG condition is as follows. We say that CPG holds at x^* if there exists $I' \subset I$, $J' \subset J_-$ and a neighborhood V of x^* such that:

1. The gradients $\nabla h_i(x^*)$ and $\nabla g_j(x^*)$ indexed by $i \in I'$ and $j \in J'$ are linearly independent;
2. The inclusion:

$$K(x) \subset K_{I',J'}(x) \text{ holds for all } x \in V. \tag{2.13}$$

Clearly, from (2.13), we have that $K(x^*) = K_{I',J'}(x^*)$. The cone $K_{I',J'}(x)$ is outer semicontinous at x^* as the following technical lemma shows:

Lemma 2.1. *Let $x^* \in \Omega$, $I' \subset I, J' \subset J_-$ and $J_+ = J(x^*) - J_-$ such that*

- *The gradients $\nabla h_i(x^*)$ and $\nabla g_j(x^*)$ indexed by $i \in I'$ and $j \in J'$ are linearly independent.*

Then the set-valued mapping $x \in \mathbb{R}^n \rightrightarrows K_{I',J'}(x)$ is outer semicontinous at x^*.

Proof. Let ω^* be an element of $\limsup_{x \to x^*} K_{I',J'}(x)$, so there are sequences x^k, ω^k such that $x^k \to x^*$, $\omega^k \to \omega^*$ and $\omega^k \in K_{I',J'}(x^k)$ with

$$\omega^k = \sum_{i \in I'} \lambda_i^k \nabla h_i(x^k) + \sum_{l \in J'} \gamma_\ell^k \nabla g_\ell(x^k) + \sum_{j \in J_+} \mu_j^k \nabla g_j(x^k) \tag{2.14}$$

for some sequence $\{\lambda_i^k \in \mathbb{R}, i \in I'; \gamma_\ell^k \in \mathbb{R}, \ell \in J'; \mu_j^k \in \mathbb{R}_+, j \in J_+\}$. Define $M_k = \max\{|\lambda_i^k|, i \in I'; |\gamma_\ell^k|, \ell \in J'; \mu_j^k, j \in J_+\}$. We have two possibilities:

- If $\{M_k\}$ has a bounded subsequence. So we can assume, by possibly extracting an adequate subsequence, that for all $i \in I', \ell \in J'$ and $j \in J_+$ the subsequences of $\lambda_i^k, \gamma_\ell^k, \mu_j^k$ have limits $\lambda_i^*, \gamma_\ell^*, \mu_j^*$ respectively. Now, taking the limit at (2.14) we get

$$\omega^* = \sum_{i \in I'} \lambda_i^* \nabla h_i(x^*) + \sum_{l \in J'} \gamma_\ell^* \nabla g_\ell(x^*) + \sum_{j \in J_+} \mu_j^* \nabla g_j(x^*) \in K_{I',J'}(x^*).$$

- Otherwise, we have $M_k \to \infty$. Dividing (2.14) by M_k, we arrive at

$$\frac{\omega^k}{M_k} = \sum_{i \in I'} \frac{\lambda_i^k}{M_k} \nabla h_i(x^k) + \sum_{l \in J'} \frac{\gamma_\ell^k}{M_k} \nabla g_\ell(x^k) + \sum_{j \in J_+} \frac{\mu_j^k}{M_k} \nabla g_j(x^k). \tag{2.15}$$

  Since $\max\{|\lambda_i^k/M_k|, i \in I'; |\gamma_\ell^k/M_k|, \ell \in J'; \mu_j^k/M_k, j \in J_+\} = 1$ for all $k \in \mathbb{N}$, we can extract a convergent subsequence. Thus, taking limits in (2.15), we get a contradiction to the fact that $\{\nabla h_i(x^*), i \in I\}$, $\{\nabla g_\ell(x^*), \ell \in J'\}$ and $\{\nabla g_j(x^*), j \in J_+\}$ form a positive linear independent set, see Remark 1.

□

3 Cone-Continuity Constraint Qualification

Definition 3.1. We say that $x^* \in \Omega$ satisfies the Cone-Continuity Property CCP if the set-valued mapping (multifunction) $\mathbb{R}^n \ni x \rightrightarrows K(x)$, defined in (2.11), is outer semicontinuous at x^*, that is

$$\limsup_{x \to x^*} K(x) \subset K(x^*). \tag{3.1}$$

The AKKT condition is naturally associated with the CCP condition. The best way to see this is to write it a more compact but equivalent form [3]. The AKKT condition holds at $x^* \in \Omega$ if, and only if, there exist sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^k\} \subset \mathbb{R}^m$, and $\{\mu^k\} \subset \mathbb{R}^p_+$ with $\mu_j^k = 0$ for $j \notin J(x^*)$, such that $\lim_{k \to \infty} x^k = x^*$ and

$$\lim_{k \to \infty} \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j \in J(x^*)} \mu_j^k \nabla g_j(x^k) = 0. \tag{3.2}$$

The expression (3.2) says that if the AKKT condition holds the vector $-\nabla f(x^k)$ gets arbitrarily close to the cone $K(x^k)$ as k goes to ∞.

Note that the multifunction $x \in \mathbb{R}^n \rightrightarrows K(x)$ is always inner semicontinuous due to the continuity of the gradients and the definition of $K(x)$. For this reason, outer semicontinuity is sufficient for the continuity of $K(x)$ at x^*.

In the following theorem we show that CCP plays, with respect to AKKT, the same role as the Guignard's CQ plays with respect to local optimality. Guignard's CQ is the weakest constraint qualification that guarantees that local minimality implies KKT [16], in the same sense that CCP is the weakest condition on the constraints that guaranteeing that AKKT implies KKT.

Theorem 3.1. *The CCP condition is the weakest property under which AKKT implies KKT, independently of the objective function. (In other words, CCP is the weakest strict constraint qualification.)*

Proof. Let us show first that, if CCP holds, the sequential AKKT condition implies the KKT condition independently of the objective function. Let f be an objective function such that the sequential AKKT condition holds at x^*, then there are sequences $\{x^k\} \to x^*$, $\{\lambda^k\} \in \mathbb{R}^n$, $\{\mu^k\} \in \mathbb{R}^p_+$ and $\{\zeta^k\} \in \mathbb{R}^m$ such that $\mu_j^k = 0$ for $j \notin J(x^*)$ and

$$\zeta^k = \nabla f(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{j \in J(x^*)} \mu_j^k \nabla g_j(x^k) \to 0. \tag{3.3}$$

Define $\omega^k = \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j\in J(x^*)} \mu_j^k \nabla g_j(x^k)$, we see that

$$\omega^k \in K(x^k) \ \text{ and } \ \omega^k = \zeta^k - \nabla f(x^k). \tag{3.4}$$

Taking limits in (3.4) when k goes to infinity, using the continuity of the gradient of f and $\zeta^k \to 0$, we get

$$-\nabla f(x^*) = \lim \omega^k \in \limsup_{k\to\infty} K(x^k) \subset \limsup_{x\to x^*} K(x) \subset K(x^*), \tag{3.5}$$

where the last inclusion follows from the Cone-Continuity Property. Therefore, $-\nabla f(x^*) \in K(x^*)$, which is equivalent to say that x^* satisfies the KKT condition.

Now, let us prove that, if the sequential AKKT condition implies the KKT condition for every objective function, then CCP holds. Let $\omega^* \in \limsup_{x\to x^*} K(x)$, by definition of outer limit, there are sequences x^k, ω^k tais que $x^k \to x^*$, $\omega^k \to \omega^*$ and $\omega^k \in K(x^k)$. Define $f(x) = -\langle w^*, x\rangle$ for all $x \in \mathbb{R}^n$. Then the AKKT condition holds at x^* for this function with $\{x^k\}$ as AKKT sequence since $\nabla f(x^k) + \omega^k = -\omega^* + \omega^k \to 0$. So by hypothesis the KKT condition holds at x^*, that is, $-\nabla f(x^*) = \omega^* \in K(x^*)$. □

Since the AKKT condition is a necessary optimality condition, we have the next corollary, cf. [3].

Corollary 3.2. *The Cone Continuity Property CCP is a Constraint Qualification.*

Remark 2. The book [10] introduces a small variation of the CCP called the U-condition. It was based on CCP [6], but tried to avoid the nomenclature of variational analysis. In particular, Theorems 3.3 and 3.4 from [10] shows that the U-condition fulfills the thesis of Theorem 3.1 above. This is an indirect proof that the U-condition is equivalent to the Cone-Continuity Property, even though a direct proof of such equivalence can be easily derived. In several practical constrained optimization algorithms that generate AKKT sequences (for example, the Sequential Quadratic Programming algorithm of Qi and Wei [21], the interior-point method of Chen and Goldfarb [14], and Augmented Lagrangian algorithms [2, 10]) convergence to KKT points has been proved assuming different constraint qualifications. By Theorem 3.1, in all these cases the constraint qualification employed may be replaced with the weaker Cone-Continuity Property and cannot be improved using another constraint qualifications.

4 Relations with other constraint qualifications.

In this section, we study the relations between the CCP condition with other constraint qualifications.

4.1 Cone-Continuity Property and CPG condition

In this subsection we will show that Cone-Continuity Property is strictly weaker that CPG.

Theorem 4.1. *The CPG condition implies the CCP condition.*

Proof. From the definition of CPG, there are set of indexes I', J', J_+ such that the gradients $\nabla h_i(x^*)$ and $\nabla g_j(x^*)$, $(i, j) \in (I', J')$ are linearly independent and a neighborhood V of x^* such that

$$K(x) \subset K_{I',J'}(x) \ \ \forall x \in V. \tag{4.1}$$

Now taking limits in (4.1) and using the outer semicontinuity of $K_{I',J'}(x)$ at x^*, Lemma (2.1), we get

$$\limsup_{x\to x^*} K(x) \subset \limsup_{x\to x^*} K_{I',J'}(x) \subset K_{I',J'}(x^*) = K(x^*), \tag{4.2}$$

which implies the outer semicontinuity of $K(x)$ at x^*. □

Now the next example shows in fact that CCP condition is strictly weaker that CPG.

Example 4.1 (CCP does not imply CPG)**.** In $\mathbb{R}^2$, consider $x^* = (0,0)$ and the inequality constraints defined by

$$\begin{array}{rcl} g_1(x_1,x_2) & = & x_1; \\ g_2(x_1,x_2) & = & (x_1^+)^2 \exp((x_2^+)^2). \end{array}$$

Clearly, $x^* = (0,0)$ is a feasible point and both constraints are active at x^*. By direct calculations:

$$\nabla g_1(x_1,x_2) = (1,0) \text{ and } \nabla g_2(x_1,x_2) = (2x_1^+ \exp((x_2^+)^2), 2x_2^+(x_1^+)^2 \exp((x_2^+)^2)), \forall x = (x_1,x_2) \in \mathbb{R}^2.$$

Since $\nabla g_1(x^*) = (1,0)$ and $\nabla g_2(x^*) = (0,0)$, we have that $K(x^*) = \mathbb{R}_+ \times \{0\}$ and the unique choice for a positive basis of $K(x^*)$ is $\{I^{'} = \emptyset, J^{'} = \emptyset, J_+ = \{1\}\}$. Thus we get $K_{I',J'}(x) = \mathbb{R}_+ \times \{0\}$ for every $x = (x_1,x_2) \in \mathbb{R}^2$. On the other hand

$$K(x) = \{(\mu_1 + 2\mu_2 x_1^+ \exp((x_2^+)^2), 2\mu_2 x_2^+(x_1^+)^2 \exp((x_2^+)^2)) \mid \quad \mu_1, \mu_2 \geq 0\}$$

Thus, $K(x)$ cannot be a subset of $K_{I',J'}(x)$ in any neighborhood of x^* and CPG is not fulfilled.

Now let us prove that $K(x)$ is continuous at x^*. Let $\omega^* \in \limsup_{x\to x^*} K(x)$. Therefore, there are sequences x^k and ω^k, such that $x^k = (x_1^k, x_2^k) \to x^*$, $\omega^k = (\omega_1^k, \omega_2^k) \to \omega^*$, and

$$\omega^k = \mu_1^k(1,0) + \mu_2^k(2(x_1^k)^+ \exp(((x_2^k)^+)^2), 2(x_2^k)^+((x_1^k)^+)^2 \exp(((x_2^k)^+)^2)) \in K(x^k), \tag{4.3}$$

where μ_1^k, μ_2^k are non-negative scalars. Suppose, by contradiction, that $\omega^* = (\omega_1^*, \omega_2^*)$ does not belong to $K(x^*) = \mathbb{R}_+ \times \{0\}$. So, ω_2^* must be nonzero. From (4.3) we have that there exists $\rho > 0$ such that

$$|\omega_2^k| = 2\mu_2^k |(x_2^k)^+((x_1^k)^+)^2 \exp(((x_2^k)^+)^2)| > \rho > 0 \tag{4.4}$$

for k large enough. In particular, both x_1 and x_2 are strictly positive. Moreover, since $\mu_1^k \geq 0$ and using (4.4), we get

$$\omega_1^k = \mu_1^k + 2\mu_2^k (x_1^k)^+ \exp((x_2^k)^+)^2 \geq \frac{\omega_2^k}{(x_1^k)^+(x_2^k)^+} > \frac{\rho}{(x_1^k)^+(x_2^k)^+} > 0. \tag{4.5}$$

Taking limits in (4.5), we obtain $\omega_1^k \to \infty$. This is a contradiction with the fact that $\omega^k \to \omega^*$. Hence, ω^* must be in $K(x^*)$. Note that this example, $K(x)$ is the positive cone generated by $(1,0)$ and $(1, x_1^+ x_2^+)$.

4.2 Cone-Continuity Property and Abadie's constraint qualification

Next we show that CCP is stronger than Abadie's Constraint Qualification and independent of quasinormality and pseudonormality conditions. The implications are based on the following two lemmas. The first one can be found in [23, Theorem 6.11].

Lemma 4.2. *Let $\bar{x}$ in Ω. For every $v \in T_\Omega^\circ(\bar{x})$, there exists a smooth function F such that $-\nabla F(\bar{x}) = v$ and attains its global minimum relative to Ω uniquely at $\bar{x}$.*

Lemma 4.3. *For all $\bar{x} \in \Omega$ and $v \in T_\Omega^\circ(\bar{x})$, there exist sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^k\} \subset \mathbb{R}^m$, and $\{\mu^k\} \subset \mathbb{R}_+^p$ with $x^k \to \bar{x}$ such that:*

(i) $\omega^k := \sum_{i=1}^m \lambda_j^k \nabla h_j(x^k) + \sum_{j=1}^p \mu_j^k \nabla g_j(x^k)$ *converges to* v*;*

(ii) $\lambda_j^k = kh_j(x^k)$ *for all* $i = 1, \ldots, m$ *and* $\mu_j^k = kg_j(x^k)^+$ *for all* $j = 1, \ldots, p$*.*

Proof. Let v be an element of $T_\Omega^\circ(\bar{x})$. By Lemma (4.2), there exists a smooth function F such that $-\nabla F(\bar{x}) = v$ and F attains its global minimum relative to Ω uniquely at $\bar{x}$. Consider, for each $k \in \mathbb{N}$, the following optimization problem:

$$\begin{array}{ll} \text{Minimize} & F_k(x) = F(x) + \frac{k}{2}\left(\sum_{j=1}^m h_j(x)^2 + \sum_{j=1}^p (g_j(x)^+)^2\right) \\ \text{subject to} & x \in \mathbb{B}(\bar{x}, \eta). \end{array} \tag{4.6}$$

Since $\mathbb{B}(\bar{x}, \eta)$ is a compact set and $F_k(x)$ is continuous, by Weierstrass' theorem, there is at least one solution for (4.6), namely x^k. Since x^k is a solution of (4.6) we have:

$$F(x^k) \leq F(x^k) + \frac{k}{2}\left(\sum_{j=1}^m h_j(x^k)^2 + \sum_{j=1}^p (g_j(x^k)^+)^2\right) = F_k(x^k) \leq F_k(\bar{x}) = F(\bar{x}). \tag{4.7}$$

By (4.7) and using the fact that F is bounded in the compact set $\mathbb{B}(\bar{x}, \eta)$, we get

$$\lim \left|g_j(x^k)^+\right| = 0 \;\; \forall \;\; j = 1, \ldots, p \;\; \text{and} \;\; \lim \left|h_j(x^k)\right| = 0 \;\; \forall \;\; j = 1, \ldots, m.$$

By the continuity of the constraints we have that every limit point of $\{x^k\}$ is feasible. Moreover, the sequence x^k converges to $\bar{x}$. In fact, if x^∞ is a limit point of the sequence, by (4.7) we have that $F(x^k) \leq F(\bar{x})$ and, taking limits in this inequality, $F(x^\infty) \leq F(\bar{x})$. Therefore, x^∞ is also a global miinimizer of F. Since $\bar{x}$ is the unique global minimizer, $x^\infty = \bar{x}$. Thus, for k large enough, $x^k \in \text{int}\mathbb{B}(\bar{x}, \eta)$. Since x^k is a solution of (4.6) and $x^k \in \text{int}\,\mathbb{B}(\bar{x}, \eta)$, we get:

$$\nabla F_k(x^k) = \nabla F(x^k) + \sum_{j=1}^m kh_j(x^k)\nabla h_i(x^k) + \sum_{j=1}^p kg_j(x^k)^+ \nabla g_j(x^k) = 0. \tag{4.8}$$

Define $\lambda_j^k := kh_i(x^k)$, $\mu_j^k := kg_j(x^k)^+$, and $\omega^k := \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j=1}^p \mu_j^k \nabla g_j(x^k)$. From (4.8) and the continuity of $\nabla F(x)$ we have $\omega^k = -\nabla F(x^k) \to -\nabla F(x^*) = v$. Therefore, the items (i) and (ii) hold. □

Theorem 4.4. *CCP implies Abadie's Constraint Qualification.*

Proof. Our aim is to prove $T_\Omega(x^*) = L_\Omega(x^*)$. The inclusion $T_\Omega(x^*) \subset L_\Omega(x^*)$ is well known. To show that $L_\Omega(x^*) \subset T_\Omega(x^*)$, we will first prove that $N_\Omega(x^*) \subset L_\Omega(x^*)^\circ$, which is equivalent to $N_\Omega(x^*) \subset K(x^*)$ (by Farkas' lemma $L_\Omega(x^*)^\circ = K(x^*)$).

Let $v \in N_\Omega(x^*)$, so by definition of normal cone (2.7) there are sequences $\{x^k\} \in \Omega$, $\{v^k\}$ such that

$$x^k \to x^* \;\; , \;\; v^k \to v \quad \text{and} \;\; v^k \in T_\Omega^\circ(x^k).$$

By Lemma (4.3), for each $v^k \in T_\Omega^\circ(x^k)$ there exist sequences $\{x^{k,\ell}\}$ and $\{v^{k,\ell}\}$ satisfying the items (i) and (ii) of the Lemma (4.3). This means that, for all $k \in \mathbb{N}$, we have:

$$\lim_{\ell\to\infty} \omega^{k,\ell} := \lim_{\ell\to\infty} \sum_{j=1}^m \lambda_j^{k,\ell} \nabla h_j(x^{k,\ell}) + \sum_{j=1}^p \mu_j^{k,\ell} \nabla g_j(x^{k,\ell}) = v^k,$$

where

$$\mu_j^{k,\ell} = \ell g_j(x^{k,\ell})^+ \text{ for all } j = 1, \ldots, p \text{ and } \lambda_j^{k,\ell} = \ell h_j(x^{k,\ell}) \;\text{ for all } j = 1, \ldots, m.$$

Thus, for all $k \in \mathbb{N}$, there exists $\ell(k)$ such that:

- $\|x^k - x^{k,\ell(k)}\| < 1/2^k$;
- $\omega^{k,\ell(k)} = \sum_{j=1}^m \lambda_j^{k,\ell(k)} \nabla h_j(x^{k,\ell(k)}) + \sum_{j=1}^p \mu_j^{k,\ell(k)} \nabla g_j(x^{k,\ell(k)})$;
- $\|v^k - w^{k,\ell(k)}\| < 1/2^k$;
- $\mu_j^{k,\ell(k))} = \ell(k) g_j(x^{k,\ell(k)})^+$ for all $j = 1, .., p$ and $\lambda_j^{k,\ell(k)} = \ell(k) h_j(x^{k,\ell(k)})$ for all $j = 1, \ldots, m$.

Clearly,

$$\lim_{k\to\infty} x^{k,\ell(k)} = x^* \quad \text{and} \quad \lim_{k\to\infty} \omega^{k,\ell(k)} = v.$$

Furthermore, for k large enough, $\mu_j^{k,\ell(k)} = \ell(k) g_j(x^{k,\ell(k)})^+ = 0$ for $j \notin J(x^*)$. Therefore, $\omega^{k,\ell(k)}$ belongs to $K(x^{k,\ell(k)})$ if k is large enough. So, we have found sequences such that $x^{k,\ell(k)} \to x^*$ and $\omega^{k,\ell(k)} \to v$, with $\omega^{k,\ell(k)} \in K(x^{k,\ell(k)})$. By the Cone-Continuity property and the definition of outer limit we have that $v \in \limsup_{x\to x^*} K(x) \subset K(x^*)$. Therefore, $N_\Omega(x^*) \subset K(x^*) = L_\Omega(x^*)^\circ$. Taking polar conjugation we deduce:

$$L_\Omega(x^*) = K(x^*)^\circ \subset N_\Omega(x^*)^\circ. \tag{4.9}$$

Finally, by [23, Theorems 6.26 and 6.28(b)], we have that $N_\Omega(x^*)^\circ \subset T_\Omega(x^*)$. Therefore, by (4.9), $L_\Omega(x^*) \subset T_\Omega(x^*)$, as we wanted to prove. □

Now we wish to show that Abadie's CQ is strictly weaker than CCP. Note that there is no contradiction with the fact that CCP is the weakest Strict Constraint Qualification, because Abadie's CQ is not a Strict Constraint Qualification. We are going to show an example in which Abadie's CQ holds but CCP does not. In fact, in the forthcoming example the Quasinormality constraint qualification is fulfilled. The fulfillment of Quasinormality is enough to prove that Abadie's CQ also holds [11]. The example will be given in the following subsection, where we will also show that CCP does not imply Quasinormality.

4.3 Cone-Continuity Property and Quasinormality

Let us recall the definition of Quasinormality [17, 11]. We say that the Quasinormality Constraint Qualification holds at $x^* \in \Omega$ if whenever $\sum_{j=1}^m \lambda_j \nabla h_j(x^*) + \sum_{j\in J(x^*)} \mu_j \nabla g_j(x^*) = 0$ for some $\lambda \in \mathbb{R}^m$ and $\mu_j \in \mathbb{R}_+$ for every $j \in J(x^*)$, there is no sequence $x^k \to x^*$ such that for every $k \in \mathbb{N}$, $\lambda_i h_i(x^k) > 0$ when λ_i is nonzero and $g_j(x^k) > 0$ when $\mu_j > 0$.

Example 4.2 (Quasinormality does not imply CCP)**.** Consider $x^* = (0, 1)$ and the constraints

$$\begin{aligned} h(x_1, x_2) &= x_1 x_2 \\ g(x_1, x_2) &= -x_1. \end{aligned}$$

The point x^* is feasible, both constraints are active, $\nabla g(x_1, x_2) = (-1, 0)$ and $\nabla h(x_1, x_2) = (x_2, x_1)$. Let us show that Quasinormality holds at $x^* = (0, 1)$.

We have that $\nabla g(x^*) = -\nabla h(x^*) = (-1, 0)$, so if $\mu \nabla g(x^*) + \lambda \nabla h(x^*) = (0, 0)$ with non-null coefficients, we have $\mu = \lambda > 0$. Assume by contradiction that there is a sequence $(x_1^k, x_2^k) \to (0, 1)$, such that $\lambda h(x_1^k, x_2^k) > 0$ and $\mu g(x_1^k, x_2^k) > 0$ for all $k \in \mathbb{N}$. From $\mu g(x_1^k, x_2^k) = -\mu x_1^k > 0$ we get $x_1^k < 0$ and from $\lambda h(x_1^k, x_2^k) = \lambda x_1^k x_2^k > 0$, we get $x_2^k < 0$ for all $k \in \mathbb{N}$. This is impossible since $x_2^k \to 1$. Therefore, Quasinormality holds at x^*.

Now, let us prove that CCP does not hold at x^*.

We observe that $K(x^*) = \{\lambda \nabla h(x^*) + \mu \nabla g(x^*)\} = \mathbb{R} \times \{0\}$. To show that CCP does not hold at x^*, we must find $\omega^* \in \limsup_{x\to x^*} K(x)$ such that $\omega^* \notin K(x^*)$. Define $x^k = (x_1^k, x_2^k)$ as $x_1^k = 1/k$, $x_2^k = 1$ and $\lambda^k = \mu^k = k$. With this choice define $\omega^k := \lambda^k \nabla h(x^k) + \mu^k \nabla g(x^k) \in K(x^k)$. Replacing λ^k, μ^k into ω^k, we have $\omega^k = \lambda^k(1, 1/k) + \mu^k(-1, 0) = (0, 1)$, thus we get $w^* = \lim\ \omega^k = (0, 1) \in \limsup_{x\to x^*} K(x)$. However, $(0, 1)$ does not belong to the cone $K(x^*) = \mathbb{R} \times \{0\}$.

As discussed in the previous subsection, since Quasinormality implies the Abadie's CQ, the example above shows that the Abadie's CQ does not imply CCP. Therefore, the Abadie's CQ is strictly weaker than CCP. The next example shows that CCP condition does not imply Quasinormality. As a consequence, Quasinormality and CCP are independent constraint qualifications.

Example 4.3 (CCP does not imply quasinormality)**.** Consider $x^* = (0, 0)$ and the inequality constraints defined by

$$\begin{aligned} g_1(x_1, x_2) &= x_1^3; \\ g_2(x_1, x_2) &= x_1 \exp x_2. \end{aligned}$$

Clearly, x^* is feasible and both constraints are active at x^*. The gradients are $\nabla g_1(x_1, x_2) = (3x_1^2, 0)$ and $\nabla g_2(x_1, x_2) = (\exp x_2, x_1 \exp x_2)$ for all (x_1, x_2) in $\mathbb{R}^2$.

Let us show that CCP holds at $x^* = (0, 0)$. First, $K(x^*) = \{\mu_1 \nabla g_1(x^*) + \mu_2 \nabla g_2(x^*) : \mu_1, \mu_2 \geq 0\} = \{\mu_1(0, 0) + \mu_2(1, 0)\} = \mathbb{R}_+ \times \{0\}$. Take $\omega^* \in \limsup K(x)$. From the definition of outer limit there are sequences x^k and ω^k such that $x^k = (x_1^k, x_2^k) \to x^* = (0, 0)$, $\omega^k = (\omega_1^k, \omega_2^k) \to \omega^*$ and

$$\omega^k = \mu_1^k(3(x_1^k)^2, 0) + \mu_2^k(\exp(x_2^k), x_1^k \exp x_2^k) \in K(x^k) \tag{4.10}$$

where μ_1^k, μ_2^k are non-negative scalars. Suppose, by contradiction, that $\omega^* = (\omega_1^*, \omega_2^*)$ does not belong to $K(x^*) = \mathbb{R}_+ \times \{0\}$, so ω_2^* must be nonzero. By (4.10), we have that for k large enough

$$|\omega_2^k| = \mu_2^k |x_1^k \exp x_2^k| > \rho > 0 \tag{4.11}$$

where $\rho = |\omega_2^*|/2 > 0$. In particular $x_1^k \neq 0$. Since $\mu_1^k \geq 0$, using (4.11) we get

$$\omega_1^k = 3\mu_1^k(x_1^k)^2 + \mu_2^k \exp x_2^k \geq \mu_2^k \exp x_2^k \geq \frac{|\omega_2^k|}{|x_1^k|} > \frac{\rho}{|x_1^k|} > 0 \tag{4.12}$$

Taking limits in (4.12) we obtain $\omega_2^k \to \infty$, a contradiction with its convergence. Then, ω^* must be in $K(x^*)$.

Now, let us show that Quasinormality does not hold at $x^* = (0, 0)$. Define $x_1^k = x_2^k = 1/k$, $\mu_1 = 1$ and $\mu_2 = 0$. With this choice $\mu_1 \nabla g_1(x^*) + \mu_2 \nabla g_2(x^*) = 1.(0, 0) + 0.(1, 0) = (0, 0)$ and $\mu_1 g_1(x_1^k, x_2^k) = (x_1^k)^3 > 0$ for all $k \in \mathbb{N}$. So Quasinormality does not hold.

4.4 Cone-Continuity Property and Pseudonormality

In this subsection we will prove that CCP and the pseudonormality are independent of each other. We say that the Pseudonormality Constraint Qualification holds at $x^* \in \Omega$, [11, 12], if whenever $\sum_{j=1}^m \lambda_j \nabla h_j(x^*) + \sum_{j \in J(x^*)} \mu_j \nabla g_j(x^*) = 0$ for some $\lambda \in \mathbb{R}^m$ and $\mu_j \in \mathbb{R}_+$ for every $j \in J(x^*)$, there is no sequence $x^k \to x^*$ such that $\sum_{i=1}^m \lambda_j h_j(x^k) + \sum_{j \in J(x^*)} \mu_j \nabla g_j(x^k) > 0$ for all $k \in \mathbb{N}$. Trivially, pseudonormality implies quasinormality. Since CCP does not imply quasinormality, it turns out that CCP does not imply pseudonormality either. In order to show that pseudonormality does not imply CCP, consider the following example:

Example 4.4 (Pseudonormality does not imply CCP)**.** Consider $x^* = (0, 0)$ and the inequality constraints defined by

$$\begin{aligned} g_1(x_1, x_2) &= -x_1; \\ g_2(x_1, x_2) &= x_1 - x_1^2 x_2^2. \end{aligned}$$

The point x^* is clearly feasible and both constraints are active at x^*. The gradients are given by $\nabla g_1(x_1, x_2) = (-1, 0)$ and $\nabla g_2(x_1, x_2) = (1 - 2x_1 x_2^2, -2x_1^2 x_2)$ for all (x_1, x_2) in $\mathbb{R}^2$.

Let us show that Pseudonormality holds at $x^* = (0,0)$. Suppose that $\mu_1 \nabla g_1(x^*) + \mu_2 \nabla g_2(x^*) = (0,0)$ for some positive scalars μ_1 and μ_2, then $\mu_1 = \mu_2 = \mu$, but $\mu_1 g_1(x_1,x_2) + \mu_2 g_2(x_1,x_2) = -\mu x_1^2 x_2^2 \le 0$ for every $(x_1,x_2) \in \mathbb{R}^2$. So there no sequence that contradicts the pseudonormality condition.
Now, we will show that CCP does not hold at x^*. In x^* we have $K(x^*) = \mathbb{R} \times \{0\}$. Take $x_1^k = x_2^k = 1/k$ and define $\mu_2^k := (2(x_1^k)^2 x_2^k)^{-1}$, $\mu_1^k := \mu_2^k(1 - 2x_1^k(x_2^k)^2)$ and $\omega^k := \mu_1^k \nabla g_1(x_1^k, x_2^k) + \mu_2^k \nabla g_2(x_1^k, x_2^k)$. Clearly, $\omega^k \in K(x_1^k, x_2^k)$ for all $k \in \mathbb{N}$ and $\omega^k = \mu_1^k(-1,0) + \mu_2^k(1 - 2x_1^k(x_2^k)^2, -2(x_1^k)^2 x_2^k) = (0,-1)$. So $\lim_{k\to\infty} \omega^k = (0,-1) \notin K(x^*)$. Thus, CCP does not hold at $x^* = (0,0)$.

We end this section taking a closer look at the example of the introduction.

Example 4.5 (AKKT methods can fail)**.** In $\mathbb{R}^2$, consider the following optimization problem

$$\text{Minimize } f(x_1,x_2) = x_2 \text{ subject to } h(x_1,x_2) = x_1x_2 = 0, g(x_1,x_2) = -x_1 \le 0. \tag{4.13}$$

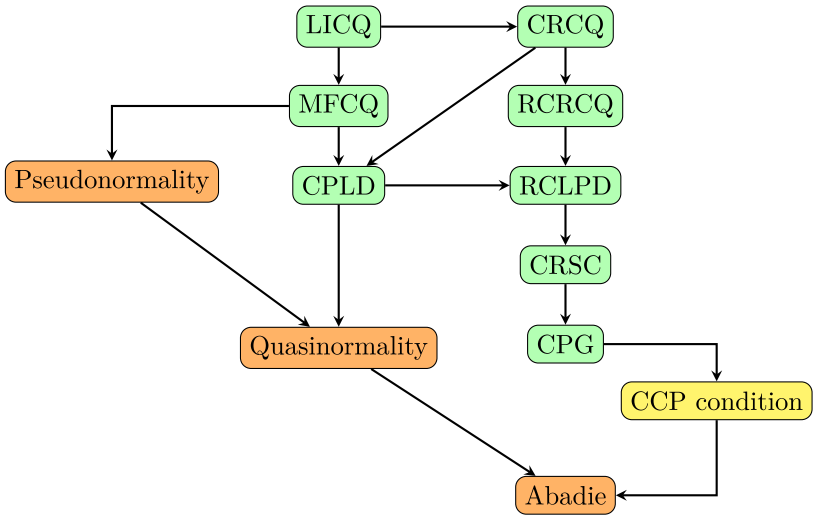
The constraints are the same as the Example 4.2 and as consequence Quasinormality and Abadie's CQ hold at $x^* = (0,1)$ but not the CCP condition. Furthermore, $x^* = (0,1)$ is an AKKT point for (4.13): To see this, take $x_1^k = 1/k$, $x_2^k = 1$, $\lambda^k = k$ and $\mu^k = k$. Calculating we obtain that $\nabla f(x^k) + \mu^k \nabla g(x^k) + \lambda^k \nabla h(x^k) = (0,1) + k(-1,0) + k(1,1/k) = (0,0)$. In spite of $x^* = (0,1)$ being an AKKT point, it means nothing for the optimization problem (4.13). The point x^* is not an optimal solution point neither a stationary point. But it can be attained by an algorithm that generates AKKT points (as an augmented lagrangian method, for instance). This means that the point $(0,1)$ fulfills any sensible practical KKT test and the algorithm will accept a point which has no relation with the optimization problem (4.13), certainly this cannot happen if instead of the quasinormality condition the point satisfies any constraint qualification which implies the CCP condition as LICQ, MFCQ, CRSC, CPG or even CCP itself.

5 Concluding remarks

Many Constrained Optimization algorithms (but not all, see [7]) generate points that satisfy the AKKT condition. This optimality condition is known to be strong because it implies "KKT or not-CQ" for many popular constraint qualifications as LICQ, MFCQ, CPLD, CRCQ, RCRCQ, RCPLD, CRSC, and the pleasantly weak Constant-Positive-Generator (CPG) constraint qualification [3, 4, 5, 8]. In this paper we asked for the weakest constraint qualification for which AKKT implies "KKT or not-CQ". We found that the CCP condition is this weakest constraint qualification, being even weaker than CPG. Therefore, the fact that AKKT implies "KKT or not-CCP" gives us the most accurate measure of the strength of AKKT as a stopping criterion for practical algorithms. In addition, we estabished the relations of CCP with other constraint qualifications that do not enjoy the "strictness property" (Pseudonormality, Quasinormality, and Abadie's CQ).

The updated landscape of Constraint Qualifications and Strict Constraint Qualifications is given in Figure 2, where arrows indicate implications. The reader may find useful to complete Figure 1 identifying the small rectangle with the points that satisfy CCP and the big rectangle with the points that satisfy Guignard's CQ.

Future work include the analysis of strict constraint qualifications associated with different sequential optimality conditions and, so, with different stopping criteria for constrained optimization algorithms. Among these, we can mention the Approximate Gradient Projection (AGP) condition [19], the L-AGP condition [3], and the Complementary Approximate Gradient Projection (CAKKT) condition [9].



References

[1] J. Abadie. On the Kuhn-Tucker Theorem. In: Abadie, J. (ed.) *Nonlinear Programming* , Noth-Holland, Amsterdam, 1967.

[2] R. Andreani, E. G. Birgin, J. M. Martínez, and M. L. Schuverdt. On Augmented Lagrangian Methods with General Lower-Level Constraints, *SIAM Journal on Optimization*, 18(4):1286, 2007.

[3] R. Andreani, G. Haeser, and J. M. Martínez. On sequential optimality conditions for smooth constrained optimization,*Optimization* 60, pp. 627-641, 2011.

[4] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva. A relaxed constant positive linear dependence constraint qualification and applications, *Mathematical Programming* 135, pp. 255-273, 2012.

[5] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva. Two new weak constraint qualifications and applications, *SIAM Journal on Optimization* 22, pp. 1109-1135, 2012.

[6] R. Andreani, J. M. Martínez, A. Ramos, and P. J. S. Silva. Constraint qualifications and approximate KKT-sequences, Contributed presentation at X Brazilian Workshop on Continuous Optimization, Florianópolis, Santa Catarina, Brazil, 2014.

[7] R. Andreani, J. M. Martínez, L. T. Santos, and B. F. Svaiter. On the behaviour of constrained optimization methods when Lagrange multipliers do not exist, *Optimization Methods and Software* 29, pp. 646–657, 2014.

[8] R. Andreani, J. M. Martínez and M. L. Schuverdt. On the relation between Constant Positive Linear Dependence Condition and Quasinormality Constraint Qualification, *Journal of Optimization Theory and Applications* 125, pp. 473-485, 2005.

[9] R. Andreani, J. M. Martínez, and B. F. Svaiter. A new sequential optimality condition for constrained optimization and algorithmic consequences, *SIAM Journal on Optimization* 20, pp. 3533-3554, 2010.

[10] E. G. Birgin and J. M. Martínez. *Practical Augmented Lagrangian Methods for Constrained Optimization*, SIAM Publications, 2014.

[11] D. P. Bertsekas, *Nonlinear Programming*, 2nd edition, Athena Scientific, Belmont, Mass., 1999.

[12] D. P. Bertsekas and A. Ozdaglar. Pseudonormality and a Lagrange Multiplier Theory for Constrained Optimization, *Journal of Optimization Theory and Applications* , 114 (2), pp. 287-343, 2002.

[13] J. M. Borwein and A. S. Lewis. *Convex Analysis and Nonlinear Optimization: Theory and Examples*,CMS Books in Mathematics,Springer,2006.

[14] L. Chen and D. Goldfarb. Interior-point 2 -penalty methods for nonlinear programming with strong global convergence properties, *Mathematical Programming*, 108, pp. 1?26, 2006.

[15] M. Guignard. Generalized Kuhn-Tucker conditions for mathematical programming problems in Banach space, *SIAM Journal on Control* 7, pp. 232-241, 1969.

[16] F. J. Gould and J. W. Tolle. A necessary and sufficient qualification for constrained optimization, *SIAM Journal on Applied Mathematics* 20, pp. 164-172, 1971.

[17] M. R. Hestenes. *Optimization Theory: The Finite Dimensional Case*, Wiley, New York, 1975.

[18] O. L. Mangasarian and S. Fromovitz, The Fritz-John necessary optimality conditions in presence of equality and inequality constraints, *Journal of Mathematical Analysis and Applications* 17 pp. 37–47, 1967.

[19] J. M. Martínez and B. F. Svaiter. A practical optimality condition without constraint qualifications for nonlinear programming. *Journal of Optimization Theory and Applications* 118, pp. 117-133, 2003.

[20] B. S. Mordukhovich. *Variational Analysis and Generalized Differentiation, I. Basis Theory*, Springer, Berlin, 1998.

[21] L. Qi and Z. Wei. On the constant positive linear dependence condition and its application to SQP methods, *SIAM Journal on Optimization* 10, pp. 963–981, 2000.

[22] R. T. Rockafellar. Lagrange multipliers and optimality, *SIAM Review* 35, pp. 183-238, 1993.

[23] R. T. Rockafellar and R. Wets. *Variational Analysis*, Springer, Berlin, 1998.