

Constraint Qualifications for Karush-Kuhn-Tucker Conditions in Constrained Multiobjective Optimization

Gabriel Haeser ^{*} Alberto Ramos [†]

March 4, 2020

Abstract

The notion of a normal cone of a given set is paramount in optimization and variational analysis. In this work, we give a definition of a multiobjective normal cone which is suitable for studying optimality conditions and constraint qualifications for multiobjective optimization problems. A detailed study of the properties of the multiobjective normal cone is conducted. With this tool, we were able to characterize weak and strong Karush-Kuhn-Tucker conditions by means of a Guignard-type constraint qualification. Furthermore, the computation of the multiobjective normal under the error bound property is provided. The important statements are illustrated by examples.

Key words: Multiobjective optimization, optimality conditions, constraint qualifications.

1 Introduction

Multiobjective optimization problems (MOPs) is a class of optimization problems involving more than one objective function to be optimized. Many real-life problems can be formulated as MOPs which include engineering design, economics, financial investment, mechanics, etc. See [18, 17] and references therein. For several theoretical and numerical issues related to MOPs, see for instance, [29, 34, 19, 28].

^{*}Department of Applied Mathematics, Institute of Mathematics and Statistics, University of São Paulo, São Paulo SP, Brazil. Email: ghaeser@ime.usp.br

[†]Department of Mathematics, Federal University of Paraná, Curitiba, PR, Brazil. Email: albertoramos@ufpr.br

Due to the possible conflicting nature of the objective functions, it is not expected to find a solution that can optimize all of the objective functions simultaneously. Thus, different notions of optimality are developed for MOPs, where one aims at finding best trade-off options (strong/weak Pareto points, see Definition 2.1). As in nonlinear programming, optimality conditions play a crucial role in the development of efficient methods for solving unconstrained and constrained MOPs. Some methods are Newton's methods [20], SQP methods [21], trust-region methods [12, 37], and scalarization methods such as weighted sum methods [23, 22], ε -constraint methods [15], Charnes-Cooper scalarization [30, 11] and others.

In scalar optimization problems, the Karush-Kuhn-Tucker (KKT) conditions play a major role in theoretical and numerical optimization. To guarantee the fulfillment of KKT conditions at a minimizer one needs constraint qualifications (CQs), that is, properties depending only on the functions defining the constraints, which ensure the KKT conditions at local minimizers. For MOPs, the corresponding KKT conditions take into account simultaneously multipliers for the constraints and for the objective functions with the multipliers associated to the objective functions being non-negative. If there exists at least one positive multiplier corresponding to the objective functions, we say that weak Karush-Kuhn-Tucker (weak KKT) conditions hold, and when all the multipliers corresponding to the objective functions are positive, we say that strong Karush-Kuhn-Tucker (strong KKT) conditions hold. But differently from the scalar case, usual CQs, in general, do not ensure weak/strong KKT conditions. Thus, additional conditions which take into account the objective functions are introduced: the so-called *regularity conditions* (RC). Such conditions have been used for establishing first-order and second-order optimality conditions for smooth MOPs. For more information see [31, 32, 14, 9] and references therein.

In this work, we take a new look at CQs for MOPs. We are interested in establishing the weakest possible CQs that guarantee the validity of weak/strong KKT conditions at locally weak Pareto solutions. To develop our new CQs, we rely on a new notion of normal cone adapted to MOPs and some variational analysis tools. For details see Section 4.

The paper is organized as follows: In Section 2 we introduce the basic assumptions and concepts useful in our analysis. In Section 3 we introduce the multiobjective normal cone and we study its properties and calculus rules. In Section 4 we introduce new CQs for MOP and we study their properties. Our concluding remarks are given in Section 5.

2 Preliminaries and Basic Assumptions

Our notation is standard in optimization and variational analysis: \mathbb{R}^n is the n -dimensional real Euclidean space, $n \in \mathbb{N}$. \mathbb{R}_+ is the set of non-negative numbers and $a^+ := \max\{0, a\}$, $a \in \mathbb{R}$. We denote by $\|\cdot\|$ the Euclidean norm in \mathbb{R}^r and by $\|\cdot\|_1$ the ℓ_1 -norm, i.e., $\|\theta\|_1 := |\theta_1| + \dots + |\theta_r|$, for $\theta \in \mathbb{R}^r$. Given a differentiable mapping $F : \mathbb{R}^s \rightarrow \mathbb{R}^d$, we use $\nabla F(x)$ to denote the Jacobian matrix of F at x .

For a set-valued mapping $\Gamma : \mathbb{R}^s \rightrightarrows \mathbb{R}^d$, the sequential Painlevé-Kuratowski outer limit of $\Gamma(z)$ as $z \rightarrow z^*$ is denoted by

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

For a cone $\mathcal{K} \subset \mathbb{R}^s$, its polar is $\mathcal{K}^\circ := \{v \in \mathbb{R}^s : \langle v, k \rangle \leq 0 \text{ for all } k \in \mathcal{K}\}$. Always $\mathcal{K} \subset \mathcal{K}^{\circ\circ}$ and when \mathcal{K} is a closed convex set, equality holds $\mathcal{K}^{\circ\circ} = \mathcal{K}$.

Given $X \subset \mathbb{R}^n$ and $\bar{z} \in X$, we define the *tangent cone* to X at \bar{z} by

$$T(\bar{z}, X) := \{d \in \mathbb{R}^n : \exists t_k \downarrow 0, d^k \rightarrow d \text{ with } \bar{z} + t_k d^k \in X\}, \quad (1)$$

and the *regular normal cone* to X at $\bar{z} \in X$ by

$$\widehat{N}(\bar{z}, X) := \left\{ w \in \mathbb{R}^n : \limsup_{z \rightarrow \bar{z}, z \in X} \frac{\langle w, z - \bar{z} \rangle}{\|z - \bar{z}\|} \leq 0 \right\}. \quad (2)$$

The (Mordukhovich) *limiting normal cone* to X at $\bar{z} \in X$ is defined by

$$N(\bar{z}, X) := \limsup_{z \rightarrow \bar{z}, z \in X} \widehat{N}(z, X). \quad (3)$$

Here, we consider the smooth constrained multiobjective optimization problem (MOP) of the form:

$$\begin{aligned} & \text{minimize} && f(x) = (f_1(x), f_2(x), \dots, f_r(x)) \\ & \text{subject to} && h_i(x) = 0, \quad i = 1, \dots, m; \\ & && g_j(x) \leq 0, \quad j = 1, \dots, p, \end{aligned} \quad (4)$$

where $h(x) = (h_1(x), \dots, h_m(x))$ and $g(x) = (g_1(x), \dots, g_p(x))$ are continuously differentiable on \mathbb{R}^n . We denote by $\Omega = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\}$ the feasible set of (4) which is assumed to be a nonempty set.

Due to the possible conflicting nature of the objective functions, it is not expected to find a solution that can optimize all of the objective functions simultaneously. Therefore, different notions of optimality are developed for MOPs, where one aim at finding best trade-off options (Pareto points, see Definition 2.1). We give below some important definitions.

Definition 2.1. Let \bar{x} be a feasible point of (4). Then,

1. We say that \bar{x} is a local (*strong*) *Pareto* optimal point if there is a $\delta > 0$ such that there is no $x \in \Omega \cap B(\bar{x}, \delta)$ with $f_\ell(x) \leq f_\ell(\bar{x}), \forall \ell \in \{1, \dots, r\}$ and $f(x) \neq f(\bar{x})$;
2. We say that \bar{x} is a local *weak Pareto* optimal point if there is a $\delta > 0$ such that there is no $x \in \Omega \cap B(\bar{x}, \delta)$ with $f_\ell(x) < f_\ell(\bar{x}), \forall \ell \in \{1, \dots, r\}$.

There is a natural way to relate weak/strong Pareto optimal point with a scalar optimization problem. Indeed, given a set of not all zero scalars (*weights*) $\theta_1 \geq 0, \dots, \theta_r \geq 0$, all local minimizers of the scalarized problem of minimizing $\theta_1 f_1(x) + \dots + \theta_r f_r(x)$ subject to $x \in \Omega$, are weak Pareto points while the reverse implication holds under convexity. See [29]. By considering the scalarized problem with positive weights $\theta_1 > 0, \dots, \theta_r > 0$, all its local minimizers are Pareto solutions, and, under convexity and compactness assumptions, a dense subset of the Pareto solutions (so-called proper Pareto solutions) coincides with the local minimizers of the scalarized problem with positive weights. See [36].

This suggests the following definitions of Karush-Kuhn-Tucker (KKT) notions associated with weak and strong Pareto solutions. See [9, 8].

Definition 2.2. Let x be a feasible point of (4). Suppose that there exists a vector $(\theta, \lambda, \mu) \neq 0 \in \mathbb{R}_+^r \times \mathbb{R}^m \times \mathbb{R}_+^p$ such that $\mu_j g_j(x) = 0$, for $j = 1, \dots, p$ and

$$\sum_{\ell=1}^r \theta_\ell \nabla f_\ell(x) + \sum_{i=1}^m \lambda_i \nabla h_i(x) + \sum_{j=1}^p \mu_j \nabla g_j(x) = 0, \quad (5)$$

Then, we say that:

- a) x is a *weak KKT point* if (5) hold with $\theta \neq 0$.
- b) x is a *strong KKT point* if (5) hold with $\theta_\ell > 0, \forall \ell = 1, \dots, r$.

When $r = 1$, both concepts reduce to the classical KKT conditions, which is known to hold at a local minimizer only under a CQ. Among the most known ones we can mention the *linear independence CQ* (LICQ), which state the linear independence of all the gradients of active constraints and the *Mangasarian-Fromovitz CQ* (MFCQ), which states the positive linear independence of such gradients, see [33].

We mention that necessary optimality conditions are not only useful for identifying possible candidate solutions to a problem, but also crucial in

designing numerical methods for solving such problems. Also, many CQs (but not all) can be used in the convergence analysis of numerical methods for scalar optimization problems. See [3, 5, 6] and references therein.

Finally, we present some basic results on MOPs. In this paper, we use $V = (v_\ell)_{\ell=1}^r$ to denote a $n \times r$ matrix whose columns are the vectors v_1, v_2, \dots, v_r in \mathbb{R}^n . Thus, $V\theta = \theta_1 v_1 + \dots + \theta_r v_r$ for $\theta = (\theta_1, \dots, \theta_r) \in \mathbb{R}^r$. From [8, 16] we have the following

Lemma 2.1. *Let \bar{x} be a local weak Pareto point of (4). Consider the scalar optimization problem*

$$\text{minimize } \Phi(x) := \max_{\ell=1, \dots, r} \{f_\ell(x) - f_\ell(\bar{x})\} \text{ subject to } x \in \Omega. \quad (6)$$

Then, \bar{x} is a local solution of (6). Furthermore, if f is a C^1 mapping then $\max_{\ell=1, \dots, r} \langle \nabla f_\ell(\bar{x}), d \rangle \geq 0, \forall d \in T(\bar{x}, \Omega)$.

From [7], we have the following theorem of alternative

Lemma 2.2. *Let A, B, C be given matrices with n columns and $A \neq 0$. Then, exactly one of the following statements is true:*

- (s1) $Az > 0, Bz \geq 0, Cz = 0$ has a solution z ;
- (s2) $A^T y_1 + B^T y_2 + C^T y_3 = 0, y_1 \geq 0, y_2 \geq 0, y_1 \neq 0$ has solution y_1, y_2, y_3 .

3 A Multiobjective Normal Cone

In this section, we will introduce a new notion of normal cone for multiobjective optimization problems, and derive its properties. We start by considering a closed subset Ω in \mathbb{R}^n and a point $\bar{x} \in \Omega$.

Definition 3.1. Given $\bar{x} \in \Omega$ and $r \in \mathbb{N}$, the *regular r -multiobjective normal cone* to Ω at \bar{x} is the cone defined as

$$\widehat{N}(\bar{x}, \Omega; r) := \left\{ V = (v_\ell)_{\ell=1}^r \in \mathbb{R}^{n \times r} : \limsup_{x \rightarrow \bar{x}, x \in \Omega} \min_{\ell=1, \dots, r} \frac{\langle v_\ell, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \right\}. \quad (7)$$

For a lower semicontinuous mapping f having lower semicontinuous extended-real-valued components $f_\ell : \mathbb{R}^n \rightarrow (-\infty, \infty]$, $\ell = 1, \dots, r$, we define the

r -regular subdifferential of f at \bar{x} as

$$\widehat{\partial}f(\bar{x}; r) := \left\{ V = (v_\ell)_{\ell=1}^r \in \mathbb{R}^{n \times r} : \liminf_{x \rightarrow \bar{x}} \left(\max_{\ell=1, \dots, r} \frac{f_\ell(x) - f_\ell(\bar{x}) - \langle v_\ell, x - \bar{x} \rangle}{\|x - \bar{x}\|} \right) \geq 0 \right\}. \quad (8)$$

Furthermore, we define the *limiting r -multiobjective normal cone* to Ω at \bar{x} by

$$N(\bar{x}, \Omega; r) := \limsup_{x \rightarrow \bar{x}} \widehat{N}(x, \Omega; r) \quad (9)$$

and the *limiting r -multiobjective subdifferential* of f at \bar{x}

$$\partial f(\bar{x}, \Omega; r) := \limsup_{x \rightarrow \bar{x}} \widehat{\partial}f(x; r). \quad (10)$$

Clearly, when $r = 1$ the normal cones reduce to the classical regular, and limiting normal cones. In the rest of this section, we will analyse the main properties of the multiobjective normal cone. Thus, we start by computing them for some simple cases.

Proposition 3.1. *Let Ω be a closed convex set, $\bar{x} \in \Omega$, and $r \in \mathbb{N}$. Then, we have that $N(\bar{x}, \Omega; r) = \widehat{N}(\bar{x}, \Omega; r)$. Furthermore,*

$$\widehat{N}(\bar{x}, \Omega; r) = \{V = (v_\ell)_{\ell=1}^r \in \mathbb{R}^{n \times r} : \min_{\ell=1, \dots, r} \langle v_\ell, x - \bar{x} \rangle \leq 0, \quad \forall x \in \Omega\}. \quad (11)$$

Proof. First, we will show the equality in (11). Take $V \in \widehat{N}(\bar{x}, \Omega; r)$ and $x \in \Omega$. For every $t \in [0, 1]$, set $x(t) := \bar{x} + t(x - \bar{x}) \in \Omega$. Thus, since $x(t) - \bar{x} = t(x - \bar{x})$, we obtain that

$$\min_{\ell=1, \dots, r} \frac{\langle v_\ell, x - \bar{x} \rangle}{\|x - \bar{x}\|} = \lim_{t \rightarrow 0} \min_{\ell=1, \dots, r} \frac{\langle v_\ell, x(t) - \bar{x} \rangle}{\|x(t) - \bar{x}\|} \leq \limsup_{z \rightarrow \bar{x}, z \in \Omega} \min_{\ell=1, \dots, r} \frac{\langle v_\ell, z - \bar{x} \rangle}{\|z - \bar{x}\|} \leq 0,$$

which implies that $\min_{\ell=1, \dots, r} \langle v_\ell, x - \bar{x} \rangle \leq 0$.

Clearly, if $V = (v_\ell)_{\ell=1}^r$ satisfies $\min_{\ell=1, \dots, r} \langle v_\ell, x - \bar{x} \rangle \leq 0 \quad \forall x \in \Omega$, we have that $V \in \widehat{N}(\bar{x}, \Omega; r)$. From (11), we obtain that $N(\bar{x}, \Omega; r) = \widehat{N}(\bar{x}, \Omega; r)$. \square

When Ω is a polyhedral we can obtain a nice characterization of the r -multiobjective normal cone.

Proposition 3.2. *Suppose that $\Omega = \{x \in \mathbb{R}^n : Ax \leq 0\}$ for some matrix $A \in \mathbb{R}^{m \times n}$. Then, for every $\bar{x} \in \Omega$, $\widehat{N}(\bar{x}, \Omega; r) = N(\bar{x}, \Omega; r)$ holds and*

$$N(\bar{x}, \Omega; r) = \left\{ V = (v_\ell)_{\ell=1}^r \in \mathbb{R}^{n \times r} : \begin{array}{l} V\theta = A^T \lambda, \text{ with } \lambda \in N(A\bar{x}, \mathbb{R}_-^m) \\ \text{and } 0 \neq \theta \in \mathbb{R}_+^r \end{array} \right\}. \quad (12)$$

The proposition is a direct consequence of Corollary 4.7.

It is well-known that the regular normal cone can be characterized by viscosity sub gradients [40, Theorem 6.11]. Here, we have a variational characterization of the regular r -multiobjective normal cone, which extends the viscosity characterization of the classical normal cone when $r = 1$.

Theorem 3.3. *Let \bar{x} be a feasible point of (4). Then, the following statements are equivalent*

- (a) $V = (v_\ell)_{\ell=1}^r \in \mathbb{R}^{n \times r}$ belongs to $\widehat{N}(\bar{x}, \Omega; r)$;
- (b) There is a C^1 mapping $f(x) = (f_\ell(x))_{\ell=1}^r$ such that \bar{x} is a local Pareto optimal point relative to Ω , and $v_\ell = -\nabla f_\ell(\bar{x})$, for every $\ell = 1, \dots, r$;
- (c) For every $d \in T(\bar{x}, \Omega)$, we have that $\min_{\ell=1, \dots, r} \langle v_\ell, d \rangle \leq 0$.

As a consequence

$$\widehat{N}(\bar{x}, \Omega; r) = \{V = (v_\ell)_{\ell=1}^r \in \mathbb{R}^{n \times r} : \min_{\ell=1, \dots, r} \langle v_\ell, d \rangle \leq 0, \forall d \in T(\bar{x}, \Omega)\}. \quad (13)$$

Proof. First, we will show that (a) implies (b). Take $V = (v_\ell)_{\ell=1}^r \in \widehat{N}(\bar{x}, \Omega; r)$. We will use an argument similar to [40, Theorem 6.11]. We define

$$\eta_0(r) := \text{Sup} \left\{ \min_{\ell=1, \dots, r} \langle v_\ell, x - \bar{x} \rangle : x \in \Omega, \|x - \bar{x}\| \leq r \right\}. \quad (14)$$

Clearly, $\eta_0(r) \leq r \min_{\ell=1, \dots, r} \|v_\ell\|$ and $0 \leq \eta_0(0) \leq \eta_0(r) \leq o(r)$. By using [40, Theorem 6.11], there is $\eta : \mathbb{R}_+ \rightarrow \mathbb{R}$ continuously differentiable on \mathbb{R}_+ with $\eta_0(r) \leq \eta(r)$, $\eta'(r) \rightarrow 0$ and $\eta(r)/r \rightarrow 0$ as $r \rightarrow 0$. Thus, for every $\ell \in \{1, \dots, r\}$ define $f_\ell(x) := -\langle v_\ell, x - \bar{x} \rangle + \eta(\|x - \bar{x}\|)$ and $f(x) = (f_\ell(x))_{\ell=1}^r$. Clearly, f is a C^1 mapping with $f(\bar{x}) = 0$. To show that \bar{x} is a local weak Pareto point for f relative to Ω , suppose that this is not the case. Then, there is a sequence $x^k \rightarrow \bar{x}$ with $x^k \in \Omega$, $\forall k \in \mathbb{N}$ such that $f_\ell(x^k) < f_\ell(\bar{x}) = 0$, $\forall \ell = 1, \dots, r$. Since $\eta_0(r) \leq \eta(r)$ and by definition of $f_\ell(x)$, we see that $-\langle v_\ell, x^k - \bar{x} \rangle + \eta_0(\|x^k - \bar{x}\|) < 0$, $\forall \ell$ and hence $\eta_0(\|x^k - \bar{x}\|) < \min_{\ell=1, \dots, r} \langle v_\ell, x^k - \bar{x} \rangle$ for k large enough, which is a contradiction with (14).

To show that (b) implies (c), suppose that there exists a C^1 mapping $f(x) = (f_\ell(x))_{\ell=1}^r$ such that \bar{x} is a local weak Pareto optimal point relative to Ω . From Lemma 2.1, we have that $\max_{\ell=1, \dots, r} \langle \nabla f_\ell(\bar{x}), d \rangle \geq 0$, $\forall d \in T(\bar{x}, \Omega)$. Thus, if $V := (v_\ell)_{\ell=1}^r$ with $v_\ell := -\nabla f_\ell(\bar{x})$, $\forall \ell$, we get that $\min_{\ell=1, \dots, r} \langle v_\ell, d \rangle \leq 0$.

Finally, to see that (c) implies (a), assume that V does not belong to $\widehat{N}(\bar{x}, \Omega; r)$. Thus, there exist a scalar $\alpha > 0$ and a sequence $\{x^k\} \subset \Omega$ converging to \bar{x} such that

$$\lim_{k \rightarrow \infty} \min_{\ell=1, \dots, r} \frac{\langle v_\ell, x^k - \bar{x} \rangle}{\|x^k - \bar{x}\|} > 2\alpha > 0. \quad (15)$$

Taking a further subsequence of $\{x^k\}$, we may assume that $(x^k - \bar{x})/\|x^k - \bar{x}\|$ converges to some vector $d \in \mathbb{R}^n$. Clearly, $d \in T(\bar{x}, \Omega)$. From (15), we get that $\min_{\ell=1, \dots, r} \langle v_\ell, d \rangle \geq \alpha > 0$, a contradiction. Thus, $V \in \widehat{N}(\bar{x}, \Omega; r)$. \square

Remark. When $r = 1$, the right-hand side of (13) coincides with $T(\bar{x}, \Omega)^\circ$. Thus, $\widehat{N}(\bar{x}, \Omega) = \widehat{N}(\bar{x}, \Omega; 1) = T(\bar{x}, \Omega)^\circ$.

We have also the following

Lemma 3.4. *Consider f and g two lower semicontinuous mappings having lower semicontinuous extended-real-valued components, and Ω a closed subset of \mathbb{R}^n . Then, we always have that*

1. $\widehat{\partial}\bar{\iota}_\Omega(x; r) = \widehat{N}(x, \Omega; r)$ where $\bar{\iota}_\Omega$ is the indicator mapping, that is, a mapping defined by $\bar{\iota}_\Omega(x) = 0$ if $x \in \Omega$ and $\bar{\iota}_\Omega(x) = \infty$ otherwise.
2. $V = (v_\ell)_{\ell=1}^r \in \widehat{\partial}f(x; r)$ iff there is a C^1 mapping h such that x is a local Pareto solution of $f - h$ and $v_\ell = \nabla h_\ell(x)$, $\forall \ell = 1, \dots, r$.
3. If x is a local Pareto solution of f , then $0 \in \widehat{\partial}f(x; r)$.
4. If f is a C^1 mapping, then

$$\widehat{\partial}f(x; r) = \{U \in \mathbb{R}^{n \times r} : U\theta = \nabla f(x)^T \theta, \quad \|\theta\|_1 = 1, \theta \geq 0\}.$$

5. If f is a C^1 mapping, then $\widehat{\partial}(f + g)(x; r) = \nabla f(x) + \widehat{\partial}g(x; r)$.

Proof. First, it is straightforward to see that items 1 and 5 hold. Also, observe that item 2 is just a modification of Theorem 3.3.

Item 3: Since \bar{x} is a local Pareto solution of the mapping f , we have that $\max_{\ell=1, \dots, r} \{f_\ell(x) - f_\ell(\bar{x})\} \geq 0$ for every x near \bar{x} , and thus, $0 \in \widehat{\partial}f(\bar{x}; r)$.

Item 4: Take $V = (v_\ell)_{\ell=1}^r \in \widehat{\partial}f(\bar{x}; r)$. By item 2, there exists a C^1 mapping h such that \bar{x} is a local Pareto solution of $f - h$ and $v_\ell = \nabla h_\ell(\bar{x})$, $\forall \ell = 1, \dots, r$. Then, $\max_{\ell=1, \dots, r} \{f_\ell(x) - h_\ell(x) - (f_\ell(\bar{x}) - h_\ell(\bar{x}))\} \geq 0$, for every x near \bar{x} . Thus, by Fermat's rule and computing Clarke's subdifferential,

there exists $\theta \geq 0$ with $\|\theta\|_1 = 1$ such that $\sum_{\ell=1}^r \theta_\ell (\nabla f_\ell(\bar{x}) - \nabla h_\ell(\bar{x})) = 0$, i.e.,
 $V\theta = \nabla f(\bar{x})^T \theta$.

Now, take $U = (u_\ell)_{\ell=1}^r$ such that $U\theta = \nabla f(x)^T \theta$ for $\theta \geq 0$, $\|\theta\|_1 = 1$. Observe that

$$\begin{aligned} \max_{\ell=1, \dots, r} \frac{f_\ell(x) - f_\ell(\bar{x}) - \langle u_\ell, x - \bar{x} \rangle}{\|x - \bar{x}\|} &\geq \sum_{\ell=1}^r \frac{\theta_\ell f_\ell(x) - \theta_\ell f_\ell(\bar{x}) - \theta_\ell \langle u_\ell, x - \bar{x} \rangle}{\|x - \bar{x}\|} \\ &\geq \sum_{\ell=1}^r \frac{\theta_\ell f_\ell(x) - \theta_\ell f_\ell(\bar{x}) - \theta_\ell \langle \nabla f_\ell(\bar{x}), x - \bar{x} \rangle}{\|x - \bar{x}\|} \\ &\geq \sum_{\ell=1}^r \theta_\ell \left(\frac{f_\ell(x) - f_\ell(\bar{x}) - \langle \nabla f_\ell(\bar{x}), x - \bar{x} \rangle}{\|x - \bar{x}\|} \right). \end{aligned}$$

Taking limit inferior when $x \rightarrow \bar{x}$ in the last expression, we get $U \in \widehat{\partial} f(x; r)$. \square

We end this section with the following observation: If \bar{x} is a local Pareto solution of minimizing $f(x) = (f_1(x), \dots, f_r(x))$ subject to $x \in \Omega$, then, \bar{x} is also a local Pareto solution of the unconstrained MOP of minimizing $f(x) + \bar{v}_\Omega(x)$, where $\bar{v}_\Omega(x)$ is the indicator mapping defined in Lemma 2.1, Item 1. Thus, by Lemma 2.1, Item 3, we obtain that $0 \in \widehat{\partial}(f + \bar{v}_\Omega)(\bar{x}; r)$. If we assume that $f(x)$ is a C^1 mapping, by Lemma 2.1, Item 5, we obtain that $0 \in (\nabla f_1(x), \dots, \nabla f_r(x)) + \widehat{N}(\bar{x}, \Omega; r)$.

Corollary 3.5. *If \bar{x} is a local Pareto optimal point of (4), then*

$$0 \in (\nabla f_1(x), \dots, \nabla f_r(x)) + \widehat{N}(x, \Omega; r). \quad (16)$$

Thus, (16) can be interpreted as Fermat's rule using the r -multiobjective normal cone.

4 On Constraint Qualifications for Multiobjective Optimization

In this section, we will focus on the use of the r -multiobjective normal cone to define CQs to characterize the weak/strong KKT conditions. Later, we will analyse the relations of the r -multiobjective cones with the tangent cone and normal cone of the feasible set.

Most CQs for classical KKT conditions relate some geometric objects (as the tangent cone) with some analytic objects (as the linearized tangent cone). We recall that the *tangent cone to Ω* at \bar{x} is defined as

$$T(\bar{x}, \Omega) = \{d \in \mathbb{R}^n : \exists t_k \downarrow 0, d^k \rightarrow d \text{ with } \bar{z} + t_k d^k \in \Omega\}.$$

Despite the importance of the tangent cone in optimization and variational analysis, $T(\bar{x}, \Omega)$ can be a difficult geometric object to compute. Thus, it is common to consider a first-order approximation of $T(\bar{x}, \Omega)$, which is called *linearized tangent cone to Ω* at \bar{x} , and is defined by

$$L(\bar{x}, \Omega) := \left\{ d \in \mathbb{R}^n : \begin{array}{l} \langle \nabla h_i(\bar{x}), d \rangle = 0, \quad i = 1, \dots, m \\ \langle \nabla g_j(\bar{x}), d \rangle \leq 0, \quad j \in A(\bar{x}) \end{array} \right\}, \quad (17)$$

where $A(\bar{x}) := \{j = 1 \dots, p : g_j(\bar{x}) = 0\}$. With such objects, important CQs were introduced in the literature. For instance, given a feasible point $\bar{x} \in \Omega$, *Guignard's CQ* is stated as $T(\bar{x}, \Omega)^\circ = L(\bar{x}, \Omega)^\circ$, while *Abadie's CQ* states the stronger equality $T(\bar{x}, \Omega) = L(\bar{x}, \Omega)$.

In this section, we define new CQs for weak/strong KKT conditions based on the multiobjective normal cone, and we analyze their properties.

4.1 Constraint Qualifications for Weak KKT Points

Here, we start by giving the weakest CQ to guarantees that a local weak Pareto minimizer is, in fact, a weak KKT point. This is characterized by the fact that the fulfillment of the CQ is equivalent to the fact that independently of the objective function, local Pareto implies weak KKT.

Theorem 4.1. *Let \bar{x} be a feasible point of (4). Then, the weakest property under which every local weak Pareto point satisfies the weak KKT conditions for every continuously differentiable mapping $f(x) = (f_1(x), \dots, f_r(x))$ is*

$$L(\bar{x}, \Omega) \subset \{d \in \mathbb{R}^n : \min_{\ell=1, \dots, r} \langle v_\ell, d \rangle \leq 0, \text{ for all } V = (v_i)_{i=1}^r \in \widehat{N}(\bar{x}, \Omega; r)\}. \quad (18)$$

Proof. Assume that every local weak Pareto point is actually a weak KKT point. We will show that inclusion (18) holds. Indeed, take $d \in L_\Omega(\bar{x})$ and $V = (v_\ell)_{\ell=1}^r \in \widehat{N}(\bar{x}, \Omega; r)$. By Theorem 3.3, there is a C^1 mapping f such that \bar{x} is a local Pareto minimizer of f over Ω and $v_\ell = -\nabla f_\ell(\bar{x})$, for every $\ell = 1, \dots, r$. By hypotheses, we get that \bar{x} is a weak KKT point and thus, there exist $\theta \in \mathbb{R}_+^m$, $\mu \in \mathbb{R}_+^p$ and $\lambda \in \mathbb{R}^m$ such that

$$\sum_{\ell=1}^r \theta_\ell v_\ell = \sum_{i=1}^m \lambda_i \nabla h_i(\bar{x}) + \sum_{j=1}^p \mu_j \nabla g_j(\bar{x}), \quad \text{and} \quad \mu_j = 0, \quad \forall j \notin A(\bar{x}).$$

Furthermore, we can assume that $\|\theta\|_1 = 1$. Since $d \in L(\bar{x}, \Omega)$, we see that

$$\min_{\ell=1, \dots, r} \langle v_\ell, d \rangle \leq \sum_{\ell=1}^r \theta_\ell \langle v_\ell, d \rangle = \sum_{i=1}^m \lambda_i \langle \nabla h_i(\bar{x}), d \rangle + \sum_{j=1}^p \mu_j \langle \nabla g_j(\bar{x}), d \rangle \leq 0.$$

Now, assume that (18) holds. Thus, take $f(x) = (f_1(x), \dots, f_r(x))$ a C^1 mapping such that \bar{x} is a local weak Pareto optimal point over Ω . Define $v_\ell := -\nabla f_\ell(\bar{x})$, $\forall \ell = 1, \dots, r$. By Motzkin's Theorem of the Alternative theorem [33], we see that \bar{x} is weak KKT point if the following system

$$\begin{aligned} \nabla h_i(\bar{x})^T d &= 0, \quad i = 1, \dots, m; & \nabla g_j(\bar{x})^T d &\leq 0, \quad j \in A(\bar{x}); \\ \nabla f_\ell(\bar{x})^T d &< 0, \quad \ell = 1, \dots, r \end{aligned} \quad (19)$$

has no solution d . Now, if we suppose that the system (19) admits a solution d , we get that $d \in L_\Omega(\bar{x})$, and by (18), $\min_{\ell=1, \dots, r} \langle v_\ell, d \rangle \leq 0$ which is equivalent to $\max_{\ell=1, \dots, r} \langle \nabla f_\ell(\bar{x}), d \rangle \geq 0$, a contradiction. Thus, (19) has no solution. \square

Recall that expression (18) is a true CQ, since it only depends on the feasible set, and no information of the objective functions are required. To continue our discussion, define

$$\mathfrak{L}(\bar{x}, \Omega; r) = \left\{ d \in \mathbb{R}^n : \min_{\ell=1, \dots, r} \langle v_\ell, d \rangle \leq 0, \quad \forall V = (v_\ell)_{\ell=1}^r \in \widehat{N}(\bar{x}, \Omega; r) \right\}. \quad (20)$$

Using (20), the weakest CQ guaranteeing that a local weak Pareto minimizer is a weak KKT point is the inclusion $L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r)$. Note that in the scalar case ($r = 1$), the inclusion (18) reduces to Guignard's CQ. Indeed, when $r = 1$, we see that $\mathfrak{L}(\bar{x}, \Omega; 1) = (\widehat{N}(\bar{x}, \Omega))^\circ = T(\bar{x}, \Omega)^\circ$. Thus, (18) is equivalent to $L(\bar{x}, \Omega) \subset T(\bar{x}, \Omega)^\circ$, which is the classical Guignard's CQ, usually stated as $T(\bar{x}, \Omega)^\circ = L(\bar{x}, \Omega)^\circ$.

We proceed by analyzing the relation of (18) with Abadie and Guignard CQs. It is known that Abadie's CQ is enough to ensure that every weak Pareto point satisfies the weak KKT conditions, see for instance [41], but Guignard's CQ does not ensure the validity of the weak KKT conditions at weak Pareto points, [1]. This is the so called "gap" in scalar and multiobjective optimization which has attracted the attention of many researchers. For more details see also [13, 42, 26]. Here, we will try to elucidate this "gap" by using the cone $\mathfrak{L}(\bar{x}, \Omega; r)$. First, consider Proposition 4.2 where some properties about $\mathfrak{L}(\bar{x}, \Omega; r)$ are stated and also provides another proof of the fact that Abadie's CQ ensures the existence of weak KKT points at local weak Pareto minimizers (see (23)).

Proposition 4.2. *Given $\bar{x} \in \Omega$ and $r \in \mathbb{N}$, we always have*

$$\mathfrak{L}(\bar{x}, \Omega; r) \subset \widehat{N}(\bar{x}, \Omega)^\circ, \quad (21)$$

$$T(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r), \quad (22)$$

$$\text{Abadie's CQ at } \bar{x} \text{ implies } L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r), \quad (23)$$

$$L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r) \text{ implies Guignard's CQ.} \quad (24)$$

Proof. To see that (21) holds, take $d \in \mathfrak{L}(\bar{x}, \Omega; r)$ and $w \in \widehat{N}(\bar{x}, \Omega)$. By Theorem 3.3 in the case that $r = 1$, we see that there exists a C^1 function $h(x)$ such that \bar{x} is a local minimizer of $h(x)$ over Ω and $w = -\nabla h(\bar{x})$. Now, define the C^1 mapping $f(x) = (f_1(x), \dots, f_r(x))$ where each component $f_\ell(x) = h(x)$, $\forall x, \forall \ell = 1, \dots, r$. It is not difficult to see that \bar{x} is a local weak Pareto minimizer for $f(x)$ over Ω . Since $\nabla f_\ell(\bar{x}) = \nabla h(\bar{x})$ for every ℓ , we see that $\langle w, d \rangle = \langle -\nabla h(\bar{x}), d \rangle = \min_{\ell=1, \dots, r} \langle -\nabla f_\ell(\bar{x}), d \rangle \leq 0$. Thus, we

obtain that $d \in \widehat{N}(\bar{x}, \Omega)^\circ$ and hence (21) holds. Inclusion (22) is a simple consequence of (13). The implication (23) follows directly from the inclusion (22), when Abadie's CQ holds at \bar{x} (i.e., $T(\bar{x}, \Omega) = L(\bar{x}, \Omega)$). To show (24), from (21) and the equality $\widehat{N}(\bar{x}, \Omega) = T(\bar{x}, \Omega)^\circ$, we see that

$$T(\bar{x}, \Omega) \subset L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r) \subset \widehat{N}(\bar{x}, \Omega)^\circ = T(\bar{x}, \Omega)^{\circ\circ}. \quad (25)$$

Taking polar in the last sequence of inclusions, we get $L(\bar{x}, \Omega)^\circ = T(\bar{x}, \Omega)^\circ$, which is Guignard's CQ at \bar{x} . \square

Finally, we show that Abadie's CQ, Guignard's CQ and inclusion (18) are independent CQs.

Example 4.1. *(Guignard's CQ is strictly weaker than $L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r)$) In \mathbb{R}^2 , take $\bar{x} = (0, 0)$ and the multiobjective problem*

$$\text{minimize } (f_1(x_1, x_2) := x_1, \quad f_2(x_1, x_2) := x_2) \text{ s.t. } h(x_1, x_2) := x_1 x_2 = 0.$$

It is not difficult to see that \bar{x} is a local weak Pareto point which is not a weak KKT point. Thus, Theorem 4.1 says that (18) fails. Now, we also observe that $L(\bar{x}, \Omega) = \mathbb{R}^2$ and $T(\bar{x}, \Omega) = \{(d_1, d_2) \in \mathbb{R}^2 : d_1 d_2 = 0\}$, and thus $L(\bar{x}, \Omega)^\circ = T(\bar{x}, \Omega)^\circ$ holds, that is, Guignard's CQ is valid at \bar{x} .

Clearly, Abadie's CQ is strictly stronger than inclusion (18), since in the scalar case, Guignard's CQ is equivalent to (18).

Finally, observe that if $\mathfrak{L}(\bar{x}, \Omega; r)$ is a closed convex cone and Guignard's CQ holds at \bar{x} , then the inclusion $L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r)$ holds.

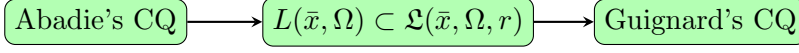


Figure 1: Gap between Abadie's and Guignard's CQ for weak KKT points.

Proposition 4.3. *Let \bar{x} be a feasible point such that $\mathfrak{L}(\bar{x}, \Omega; r)$ is a closed convex cone. Then, Guignard's CQ holds at \bar{x} iff $L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r)$ holds.*

Proof. By Proposition 4.2, inclusion $L(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r)$ implies Guignard's CQ holds. For the other implication, from (21), (22) and Guignard's CQ, we get that $T(\bar{x}, \Omega) \subset \mathfrak{L}(\bar{x}, \Omega; r) \subset T(\bar{x}, \Omega)^{\circ\circ} = L(\bar{x}, \Omega)$. Taking polarity and using again Guignard's CQ, we see that $L(\bar{x}, \Omega)^{\circ} \subset \mathfrak{L}(\bar{x}, \Omega; r)^{\circ} \subset T(\bar{x}, \Omega)^{\circ} = L(\bar{x}, \Omega)^{\circ}$ implies that $L(\bar{x}, \Omega) = \mathfrak{L}(\bar{x}, \Omega; r)^{\circ\circ} = \mathfrak{L}(\bar{x}, \Omega; r)$. \square

4.2 Constraint Qualifications for Strong KKT Points

Now, we focus on CQs and the fulfillment of the strong KKT conditions. It is well-known that Abadie's CQ is not sufficient to ensure the validity of strong KKT at local weak Pareto minimizers. Using the multiobjective normal cone, similarly to Theorem 4.1, we obtain the weakest CQ needed to guarantee that a local weak Pareto minimizer is a strong KKT point.

Theorem 4.4. *Let \bar{x} be a feasible point of (4). The weakest property under which every local weak Pareto point is a strong KKT point for every continuously differentiable mapping $f(x) = (f_1(x), \dots, f_r(x))$ is*

$$L(\bar{x}, \Omega) \subset \left\{ d \in \mathbb{R}^n : \begin{array}{l} \text{for every } V = (v_i)_{i=1}^r \in \widehat{N}(\bar{x}, \Omega; r), \text{ we have that} \\ \min_{\ell=1, \dots, r} \langle v_{\ell}, d \rangle < 0 \text{ or } \min_{\ell=1, \dots, r} \langle v_{\ell}, d \rangle = \min_{\ell=1, \dots, r} \langle v_{\ell}, d \rangle = 0 \end{array} \right\}. \quad (26)$$

Proof. To prove the inclusion (26), take $V = (v_{\ell})_{\ell=1}^r \in \widehat{N}(\bar{x}, \Omega; r)$ and a direction $d \in L(\bar{x}, \Omega)$. By Theorem 3.3, there is a C^1 mapping $f(x)$ having \bar{x} as a local Pareto point over Ω such that $-\nabla f_{\ell}(\bar{x}) = v_{\ell}$ for every $\ell = 1, \dots, r$. By hypotheses, \bar{x} must be a strong KKT point. So, there exist $\theta \in \mathbb{R}_+^r$, $\mu \in \mathbb{R}_+^p$ and $\lambda \in \mathbb{R}^m$ with $\theta_{\ell} > 0$, $\forall \ell = 1, \dots, r$, such that $\|\theta\|_1 = 1$ and

$$\sum_{\ell=1}^r \theta_{\ell} v_{\ell} = \sum_{i=1}^m \lambda_i \nabla h_i(\bar{x}) + \sum_{j=1}^p \mu_j \nabla g_j(\bar{x}), \quad \mu_j = 0, \quad \forall j \notin A(\bar{x}). \quad (27)$$

Now, from (27) and since $d \in L_\Omega(\bar{x})$, we see that

$$\sum_{\ell=1}^r \theta_\ell \langle v_\ell, d \rangle = \sum_{i=1}^m \lambda_i \langle \nabla h_i(\bar{x}), d \rangle + \sum_{j=1}^p \mu_j \langle \nabla g_j(\bar{x}), d \rangle \leq 0. \quad (28)$$

From (28), we get that $\min_{\ell=1, \dots, r} \langle v_\ell, d \rangle < 0$ or $\min_{\ell=1, \dots, r} \langle v_\ell, d \rangle = 0$. In the last case, since $\theta_\ell > 0, \forall \ell$, we get $\langle v_\ell, d \rangle = 0, \forall \ell$ and hence $\max_{\ell=1, \dots, r} \langle v_\ell, d \rangle = 0$.

Now, let $f(x) = (f_1(x), \dots, f_r(x))$ be a C^1 mapping such that \bar{x} is a local Pareto optimal point over Ω . Set $v_\ell := -\nabla f_\ell(\bar{x}), \forall \ell = 1, \dots, r$. Using Motzkin's Theorem of the Alternative theorem [33], we see that \bar{x} is strong KKT point if the following system

$$\begin{aligned} \nabla h_i(\bar{x})^T d &= 0, \quad i = 1, \dots, m & \nabla g_j(\bar{x})^T d &\leq 0, \quad j \in A(\bar{x}); \\ \nabla f_\ell(\bar{x})^T d &\leq 0, \quad \ell = 1, \dots, r; & \nabla f_\ell(\bar{x})^T d &< 0, \quad \text{for some } \ell \in \{1, \dots, r\} \end{aligned} \quad (29)$$

has no solution d . Indeed, if (29) admits a solution d , we obtain that $d \neq 0$ belongs to $L(\bar{x}, \Omega)$. From inclusion (26), we obtain a contradiction. Thus, \bar{x} is a strong KKT point. \square

When $r = 1$, the right-hand side of (26) collapses to $\widehat{N}(\bar{x}, \Omega)^\circ$. Thus, (26) coincides with $L(\bar{x}, \Omega) \subset T(\bar{x}, \Omega)^{\circ\circ}$ which is the Guignard's CQ.

The inclusion (26) is a true CQ, since it does not require any information of the objective functions for the formulation, moreover, it characterizes the fulfillment of the strong KKT conditions at local weak Pareto minimizers which coincides with Guignard's CQ when $r = 1$. Unfortunately, for multi-objective problems ($r > 1$), the inclusion (26) is too strong in the sense that even in well-behaved constrained system where LICQ holds, it may fail.

Example 4.2. (Inclusion (26) does not imply LICQ) Consider [2, Example 4.18]: In \mathbb{R} , take $\bar{x} = 0$ and the constraints $g_1(x) := x$ and $g_2(x) := -\exp(x) + 1$. Note that $\nabla g_1(x) = 1$ and $\nabla g_2(x) = -\exp(x)$, and thus LICQ fails. Now, from $L(\bar{x}, \Omega) = \{0\}$ the inclusion (26) holds.

Example 4.3. (LICQ does not imply the inclusion (26)) Indeed, in \mathbb{R}^2 , take $\bar{x} = (0, 0)$, $a \in \mathbb{R}$ and the multiobjective problem

$$\begin{aligned} &\text{minimize } (f_1(x_1, x_2) := x_1^2, \quad f_2(x_1, x_2) := (x_1 - a)^2) \\ &\text{subject to } g(x_1, x_2) := x_1^2 - x_2 \leq 0. \end{aligned}$$

It is not difficult to see that \bar{x} is a local weak Pareto point. Straightforward calculations show that \bar{x} is not a strong KKT point. Thus, Theorem 4.4 says that (26) fails. Now, since $\nabla g(x_1, x_2) := (2x_1, -1)$, LICQ holds at \bar{x} .

From the examples above, we see that usual CQs are not sufficient to obtain strong KKT points at local weak Pareto solutions. Besides this observation, it is still possible to use classical CQs for obtaining strong KKT, but for a stronger notion of optimality. We focus on the concept of efficient given in [24].

Definition 4.1. We say that a feasible point \bar{x} is a local *Geoffrion-properly efficient* if it is a local strong Pareto solution in a neighbourhood \mathcal{N} and there exists a scalar $M > 0$ such that, for each ℓ ,

$$\frac{f_\ell(x) - f_\ell(\bar{x})}{f_j(\bar{x}) - f_j(x)} \leq M, \quad (30)$$

for some j such that $f_j(\bar{x}) < f_j(x)$ whenever $x \in \Omega \cap \mathcal{N}$ and $f_\ell(\bar{x}) > f_\ell(x)$.

Proposition 4.5. *Let \bar{x} be a feasible point. If \bar{x} is Geoffrion-properly efficient point with Abadie's CQ holding at \bar{x} . Then, \bar{x} is a strong KKT point.*

Proof. By Motzkin's Theorem of the Alternative [33], \bar{x} is a strong KKT point of (4) if the following system

$$\begin{aligned} \nabla f_\ell(\bar{x})^T d &\leq 0, \quad \ell = 1, \dots, r; & \nabla f_\ell(\bar{x})^T d &< 0, \quad \text{for some } \ell \in \{1, \dots, r\} \\ \nabla h_i(\bar{x})^T d &= 0, \quad i = 1, \dots, m & \nabla g_j(\bar{x})^T d &\leq 0, \quad j \in A(\bar{x}) \end{aligned} \quad (31)$$

has no solution d . Thus, suppose that (31) admits a non trivial solution d . From Abadie's CQ, there is a sequence $x^k := \bar{x} + t_k d^k \in \Omega$, with $t_k \rightarrow 0$, $d^k \rightarrow d$ and $x^k \rightarrow \bar{x}$.

Set $\mathcal{I}_- := \{\ell : \nabla f_\ell(\bar{x})^T d < 0\}$ and $\mathcal{I}_0 := \{\ell : \nabla f_\ell(\bar{x})^T d = 0\}$. Now, consider the set $\mathcal{K} := \{k \in \mathbb{N} : f_\ell(x^k) > f_\ell(\bar{x}) \text{ for some } \ell \in \mathcal{I}_0\}$. The set \mathcal{K} is infinite, otherwise we get a contradiction with the fact that \bar{x} is a local strong Pareto solution. Moreover, taking a further subsequence if necessary, there exists $\ell_0 \in \mathcal{I}_0$ such that $f_{\ell_0}(x^k) > f_{\ell_0}(\bar{x})$, $\forall k \in \mathcal{K}$. Take $\ell_1 \in \mathcal{I}_-$. Using the Taylor expansion, we get

$$\begin{aligned} f_{\ell_1}(x^k) &= f_{\ell_1}(\bar{x}) + t_k \nabla f_{\ell_1}(\bar{x})^T d^k + o(\|t_k d^k\|), \\ f_{\ell_0}(x^k) &= f_{\ell_0}(\bar{x}) + t_k \nabla f_{\ell_0}(\bar{x})^T d^k + o(\|t_k d^k\|). \end{aligned}$$

From $f_{\ell_0}(x^k) > f_{\ell_0}(\bar{x})$, $\forall k \in \mathcal{K}$, we see that $\nabla f_{\ell_0}(\bar{x})^T d^k + o(\|t_k d^k\|) t_k^{-1} > 0$ and converges to $\nabla f_{\ell_0}(\bar{x})^T d = 0$ from the right. Thus,

$$\frac{f_{\ell_1}(x^k) - f_{\ell_1}(\bar{x})}{f_{\ell_0}(\bar{x}) - f_{\ell_0}(x^k)} = \frac{-\nabla f_{\ell_1}(\bar{x})^T d^k - \frac{o(\|t_k d^k\|)}{t_k}}{\nabla f_{\ell_0}(\bar{x})^T d^k + \frac{o(\|t_k d^k\|)}{t_k}} \rightarrow \infty \text{ as } k \rightarrow \infty,$$

which contradict the definition of Geoffrion-properly efficient point. \square

4.3 Computation of the Multiobjective Normal Cone under the Error Bound Condition

Here, we characterize the multiobjective normal cone under the error bound property, see Definition 4.2. The error bound property has a large range of applications, which include convergence analysis of iterative algorithms, optimality condition, stability of inequality systems, to cite few of them. For more informations see [35, 10] and references therein.

Definition 4.2. We say that the *Error Bound Property* (EBP) holds for Ω at the feasible point \bar{x} if there is a neighbourhood \mathcal{U} of \bar{x} such that

$$\text{dist}(x, \Omega) \leq L \max\{\|h(x)\|, \max\{0, g(x)\}\}, \quad \text{for every } x \in \mathcal{U}. \quad (32)$$

Under EBP, it is possible to get a nice characterization of the multiobjective normal cone. Indeed, we have that

Theorem 4.6. *Suppose that EBP holds at \bar{x} . Then, there is a neighbourhood \mathcal{U} of \bar{x} such that for every feasible point $x \in \mathcal{U}$, we have that:*

If $V = (v_\ell)_{\ell=1}^r \in N(x, \Omega; r)$, then, there exist $\theta \in \mathbb{R}_+^r$ with $\|\theta\|_1 = 1$ and multipliers $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}_+^p$ with $\mu_j g_j(x) = 0$, $\forall j$ such that

$$V\theta = \nabla h(x)^T \lambda + \nabla g(x)^T \mu, \quad \text{and} \quad \|(\lambda, \mu)\| \leq L\|V\theta\|. \quad (33)$$

Proof. First, consider the case when $V \in \widehat{N}(x, \Omega; r)$. By Theorem 3.3, there exist a C^1 mapping f having x as a local weak Pareto solution and $v_\ell = -\nabla f_\ell(x)$, $\ell = 1, \dots, r$. Consider $\delta > 0$ such that x is the unique solution of

$$\text{Minimize} \max_{\ell=1, \dots, r} \{f_\ell(z) - f_\ell(x)\} + \frac{1}{2}\|z - x\|^2 \quad \text{s.t.} \quad \|z - x\| \leq \delta, \quad z \in \Omega. \quad (34)$$

If in (34), we apply Fermat's rule at x by using Clarke's subdifferentials, we obtain an optimality condition based on Clarke's normal cone, which is usually larger than the limiting normal cone. To avoid the use of Clarke's subdifferentials, we proceed as follow:

For each $\eta > 0$, consider the smoothing approximation of the max function, $g_\eta(x)$, defined as

$$g_\eta(x) := \eta \ln \left\{ \sum_{\ell=1}^r \exp \left(\frac{f_\ell(x) - f_\ell(\bar{x})}{\eta} \right) \right\} - \eta \ln r, \quad \text{for every } x \in \mathbb{R}^n. \quad (35)$$

Note for each x , $g_\mu(x) \rightarrow \max_{\ell=1,\dots,r} \{f_\ell(z) - f_\ell(x)\}$ as $\mu \rightarrow 0$. Derivating $g_\eta(z)$, we get that

$$\nabla g_\eta(z) = \sum_{\ell=1}^r \theta_\ell(z) \nabla f_\ell(z) \text{ with } \theta_\ell(z) := \frac{\exp((f_\ell(z) - f_\ell(x))/\eta)}{\sum_{\ell=1}^r \exp((f_\ell(z) - f_\ell(x))/\eta)}. \quad (36)$$

Note that $\|\theta_\ell(z)\|_1 = 1$. Furthermore, we see that

$$\max_{\ell=1,\dots,r} \{f_\ell(z) - f_\ell(x)\} \leq g_\eta(z) + \eta \ln r \leq \max_{\ell=1,\dots,r} \{f_\ell(z) - f_\ell(x)\} + \eta \ln r, \quad \forall z \in \mathbb{R}^n. \quad (37)$$

Now, consider the smooth optimization problem

$$\text{Minimize } g_\eta(z) + \frac{1}{2} \|z - x\|^2 \text{ subject to } \|z - x\| \leq \delta, \quad z \in \Omega. \quad (38)$$

Let $\{\eta^k\}$ be a sequence of positive parameters converging to 0 and denote by $x^k \in \Omega$ the global minimizer of (38) with $\eta = \eta^k$. We continue by showing that $x^k \rightarrow x$. In fact, using the optimality of x^k and (37), we get that

$$g_{\eta^k}(x^k) + \frac{1}{2} \|x^k - x\|^2 \leq g_{\eta^k}(x) \leq \max_{\ell=1,\dots,r} \{f_\ell(x) - f_\ell(x)\} = 0. \quad (39)$$

But, from (37) and since x is a weak Pareto point for f , we see that

$$\begin{aligned} \frac{1}{2} \|x^k - x\|^2 - \eta^k \ln r &\leq \max_{\ell=1,\dots,r} \{f_\ell(x^k) - f_\ell(x)\} + \frac{1}{2} \|x^k - x\|^2 - \eta^k \ln r \\ &\leq g_{\eta^k}(x^k) + \frac{1}{2} \|x^k - x\|^2 \leq 0. \end{aligned}$$

Taking limit in the last expression and since $\eta^k \rightarrow 0$, we get that $x^k \rightarrow x$.

Consider k large enough such that $\|x^k - x\| < \delta$. By optimality of x^k and applying Fermat's rule for limiting normal cones, we obtain that

$$-\nabla g_{\eta^k}(x^k) - (x^k - x) = -\sum_{\ell=1}^r \theta_\ell(x^k) \nabla f_\ell(x^k) - (x^k - x) \in N(x^k, \Omega). \quad (40)$$

Thus, by [25, Theorem 3] applied to $N(x^k, \Omega)$, there are multipliers $\lambda^k \in \mathbb{R}^m$, $\mu^k \in \mathbb{R}_+^p$ with $\mu_j^k g_j(x^k) = 0$, $\forall j$ such that

$$V^k \theta^k - (x^k - x) = \nabla h(x^k)^T \lambda^k + \nabla g(x^k)^T \mu^k, \quad (41)$$

$$\|(\lambda^k, \mu^k)\| \leq L \|V^k \theta^k - (x^k - x)\| \text{ and } \|\theta^k\|_1 = 1, \quad (42)$$

where $V^k := (-\nabla f_\ell(x^k))_{\ell=1}^r$. Clearly, $V^k \rightarrow V$. From (41), there is no loss of generality (after possibly taking a further subsequence) if we assume that λ^k and θ^k converge to λ and θ respectively. Note that $\|\theta\|_1 = 1$ and $\mu_j g_j(x) = 0, \forall j$. Thus, taking limits in (41), (42), we obtain the results when V is in $\widehat{N}(x, \Omega; r)$.

For the general case, take $V \in N(x, \Omega; r)$. Thus, there are sequences $\{V^k\} \subset \widehat{N}(x^k, \Omega; r)$, $\{x^k\} \subset \Omega$ such that $V^k \rightarrow V$ and $x^k \rightarrow x$. Using the above results for each V^k and after taking a further subsequence, we obtain the desired result for V . \square

Corollary 4.7. *If EBP holds at \bar{x} . Then, $N(\bar{x}, \Omega; r) = \widehat{N}(\bar{x}, \Omega; r)$ and*

$$N(\bar{x}, \Omega; r) = \left\{ V \in (\mathbb{R}^n)^r : \begin{array}{l} V\theta = \nabla h(x)^T \lambda + \nabla g(x)^T \mu, \text{ with} \\ \lambda \in \mathbb{R}^m, \mu \in \mathbb{R}_+^p \text{ and } \theta \neq 0 \in \Theta \end{array} \right\}. \quad (43)$$

Proof. By Theorem 4.6, $N(\bar{x}, \Omega; r)$ is included in the right-side set of (43).

Now, take $V = (v_\ell)_{\ell=1}^r \in (\mathbb{R}^n)^r$ such that $V\theta = \nabla h(x)^T \lambda + \nabla g(x)^T \mu$ with $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}_+^p$ and $\theta \neq 0 \in \Theta$. Without loss of generality, we assume that $\|\theta\|_1 = 1$. Consider $d \in T(\bar{x}, \Omega)$ which implies that $\langle \nabla h_i(\bar{x}), d \rangle = 0, i = 1, \dots, m, \langle \nabla g_j(x), d \rangle \leq 0, j \in A(\bar{x})$. Thus,

$$\begin{aligned} \min_{\ell=1, \dots, r} \langle v_\ell, d \rangle &\leq \sum_{\ell=1}^r \theta_\ell \langle v_\ell, d \rangle = \left\langle \sum_{\ell=1}^r \theta_\ell v_\ell, d \right\rangle = \langle V\theta, d \rangle \\ &= \langle \nabla h(x)^T \lambda + \nabla g(x)^T \mu, d \rangle \\ &= \sum_{i=1}^m \lambda_i \langle \nabla h_i(\bar{x}), d \rangle + \sum_{j=1}^p \mu_j \langle \nabla g_j(\bar{x}), d \rangle \leq 0, \end{aligned}$$

Thus, $V \in \widehat{N}(\bar{x}, \Omega; r) \subset N(\bar{x}, \Omega; r)$, which complete the proof. \square

5 Conclusions

Constraint Qualifications have an important role in nonlinear programming (NLP). These conditions are not only useful for ensuring the validity of the KKT conditions, as they are employed also in stability analysis and convergence of methods for solving NLPs, but their role for MOPs is not so clear, since different notions of solutions and KKT conditions are presented. In this work we addressed the role of CQs for MOPs. Our main contribution is the use of the concept of multiobjective normal cone to obtain new CQs for MOPs. With this tool, we were able to characterize the weakest

CQ possible to guarantee the fulfilment of weak/strong KKT conditions at weak local Pareto minimizers. To analyze the strength of such CQs, we compared with others CQs stated in the literature. Furthermore, we provided a characterization of the multiobjective normal cone under the error bound condition.

As further subject of investigation we point out the relation of the multiobjective normal cone with the sequential optimality conditions stated in [27], since sequential optimality conditions for NLPs have been widely used for improving and unifying the global convergence analysis of several algorithms. Another interesting sequel to this study is to extend these ideas for multiobjective optimization problems with complementarity constraints, considering similar developments for single objective problems [43, 4, 38, 39].

References

- [1] B. Aghezzaf and M. Hachimi. On a gap between multiobjective optimization and scalar optimization. *Journal of Optimization Theory and Applications*, 109:431–435, 2001.
- [2] R. Andreani, G. Haeser, A. Ramos, and P.J.S. Silva. A second-order sequential optimality condition associated to the convergence of optimization algorithms. *IMA Journal of Numerical Analysis*, 47:53–63, 2017.
- [3] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva. Two new weak constraint qualifications and applications. *SIAM Journal on Optimization*, 22:1109–1135, 2012.
- [4] R. Andreani, G. Haeser, L. D. Secchin, and P. J. S. Silva. New sequential optimality conditions for mathematical programs with complementarity constraints and algorithmic consequences. *SIAM Journal on Optimization*, 29(4):3201–3230, 2019.
- [5] R. Andreani, J.M. Martínez, A. Ramos, and P.J.S. Silva. A cone-continuity constraint qualification and algorithmic consequences. *SIAM Journal on Optimization*, 26(1):96–110, 2016.
- [6] R. Andreani, J.M. Martínez, A. Ramos, and P.J.S. Silva. Strict constraint qualifications and sequential optimality conditions for constrained optimization. *Mathematics of Operations Research*, C2:693–1050, 2018.

- [7] A. Ben-Israel. Motzkin's transposition theorem, and the related theorems of Farkas, Gordan and Stiemke. In *Encyclopaedia of Mathematics, Supplement III*, Kluwer Academic Publishers, Dordrecht 2001.
- [8] G. Bigi. Optimality and lagrangian regularity in vector optimization. *Ph. D. Thesis*, University of Pisa, Pisa, 1999.
- [9] G. Bigi and M. Pappalardo. Regularity conditions in vector optimization. *Journal of Optimization Theory and Applications*, 102:83–96, 1999.
- [10] J. Bolte, T. P. Nguyen, J. Peypouquet, and B. W. Suter. From Error Bounds to the complexity of first-order descent methods for convex functions. *Mathematical Programming*, 165(2):471–507, 2017.
- [11] R. S. Burachik, C. Y. Kaya, and M. M. Rizvi. A new scalarization technique and new algorithms to generate Pareto fronts. *SIAM Journal on Optimization*, 27:2:1010–1034, 2017.
- [12] G. A. Carrizo, P. A. Lotito, and M. C. Maciel. Trust region globalization strategy for the nonconvex unconstrained multiobjective optimization problem. *Mathematical Programming*, 159:339–369, 2016.
- [13] M. Castellani and M. Pappalardo. About a gap between multiobjective optimization and scalar optimization. *Journal of Optimization Theory and Applications*, 109:437–439, 2001.
- [14] S. Chandra, J. Dutta, and C. S. Lalitha. Regularity conditions and optimality in vector optimization. *Numerical Functional Analysis and Optimization*, 25:479–501, 2004.
- [15] V. Chankong and Y. Y. Haimes. *Multiobjective Decision Making: Theory and Methodology*. North-Holland, 1983.
- [16] H. W. Corley. On optimality conditions for maximizations with respect to cones. *Journal of Optimization Theory and Applications*, 46 (1):67–78, 1985.
- [17] K. Deb. *Multi-objective optimization using evolutionary algorithms*. John Wiley and Sons, 2001.
- [18] K. Deb and R. Datta. Hybrid evolutionary multiobjective optimization and analysis of machining operations. *Engineering Optimization*, 44 (6):685–706.

- [19] M. Ehrgott. *Multicriteria optimization*. Springer, 2005.
- [20] J. Fliege, L. M. Grana Drummond, and B. F. Svaiter. Newton’s method for multiobjective optimization. *SIAM Journal on Optimization*, 20:602–626, 2009.
- [21] J. Fliege and A. I. F. Vaz. A method for constrained multiobjective optimization based on SQP techniques. *SIAM Journal on Optimization*, 26(4):2091–2119, 2016.
- [22] J. Fliege, A. I. F. Vaz, and L. N. Vicente. A new scalarization and numerical method for constructing weak Pareto front of multi-objective optimization problems. *Optimization*, 60:8:1091–1104, 2011.
- [23] S. Gass and T. Saaty. The computational algorithm for the parametric objective function. *Naval research logistics quarterly*, 2:39–45, 1955.
- [24] A. M. Geoffrion. Proper efficiency and the theory of vector optimization. *Journal of Mathematical Analysis*, 22:618–630, 1968.
- [25] H. Gfrerer and J. J. Ye. New constraint qualifications for mathematical programs with equilibrium constraints via variational analysis. *SIAM Journal on Optimization*, 27:842–865, 2017.
- [26] G. Giorgi. A note of the Guignard constraint qualification and the Guignard regularity condition in vector optimization. *Applied Mathematics*, 4: 4:734–740, 2012.
- [27] G. Giorgi, B. Jimnez, and V. Novo. Approximate Karush Kuhn Tucker condition in multiobjective optimization. *Journal of Optimization Theory and Applications*, 171:70–89, 2016.
- [28] A. Gopfert, C. Tammer, H. Riahi, and C. Zalinescu. *Variational Methods in Partially Ordered Spaces*. CMS books in mathematics, 2003.
- [29] J. Jahn. *Vector Optimization: Theory, Applications, and Extensions*. Springer-Verlag, 2011.
- [30] P. Kesarwani and J. Dutta. Charnes-Cooper scalarization and convex vector optimization. *To appear in Optimization Letters*, 2019.
- [31] M. C. Maciel, S. A. Santos, and G. N. Sottosantos. Regularity conditions in differentiable vector optimization revisited. *Journal of Optimization Theory and Applications*, 142:385–398, 2009.

- [32] T. Maeda. Constraint qualifications in multiobjective optimization problems: Differentiable case. *Journal of Optimization Theory and Applications*, 80:483–500, 1994.
- [33] O. L. Mangasarian. *Nonlinear Programming*. SIAM, 1994.
- [34] K. Miettinen. *Nonlinear Multiobjective Optimization*. Springer-Verlag, 1999.
- [35] J. S. Pang. Error bounds in mathematical programming. *Mathematical Programming*, 79:299–332, 1997.
- [36] M. Petschke. On a theorem of Arrow, Barankin, and Blackwell. *SIAM Journal on Control and Optimization*, 28(2):395–401, 1990.
- [37] S. Qu, M. Goh, and B. Liang. Trust region methods for solving multiobjective optimization. *Optimization Methods and Software*, 28 (4):796–811, 2013.
- [38] A. Ramos. Mathematical programs with equilibrium constraints: A sequential optimality condition, new constraint qualifications and algorithmic consequences. *To appear in Optimization Methods and Software*, 2016.
- [39] A. Ramos. Two new weak constraint qualifications for mathematical programs with equilibrium constraints and applications. *Journal on Optimization Theory and Applications*, pages 1–26, 2019.
- [40] R. T. Rockafellar and R. Wets. *Variational Analysis*. Series: Grundlehren der mathematischen Wissenschaften, Vol. 317, 2009.
- [41] C. Singh. Optimality conditions in multiobjective differentiable programming. *Journal of Optimization Theory and Applications*, 53 (1):115–123, 1987.
- [42] S. Y. Wang and F. M. Yang. A gap between multiobjective optimization and scalar optimization. *Journal of Optimization Theory and Applications*, 68:389–391, 1991.
- [43] P. Zhang, J. Zhang, G. H. Lin, and X. Ying. Constraint qualifications and proper pareto optimality conditions for multiobjective problems with equilibrium constraints. *Journal of Optimization Theory and Applications*, 176:763–782, 2018.