

Approximate Karush–Kuhn–Tucker Condition in Multiobjective Optimization

Giorgio Giorgi¹ · Bienvenido Jiménez² · Vicente Novo²

Received: 12 October 2015 / Accepted: 15 July 2016 / Published online: 21 July 2016
© Springer Science+Business Media New York 2016

Abstract We extend the so-called approximate Karush–Kuhn–Tucker condition from a scalar optimization problem with equality and inequality constraints to a multiobjective optimization problem. We prove that this condition is necessary for a point to be a local weak efficient solution without any constraint qualification, and is also sufficient under convexity assumptions. We also state that an enhanced Fritz John-type condition is also necessary for local weak efficiency, and under the additional quasi-normality constraint qualification becomes an enhanced Karush–Kuhn–Tucker condition. Finally, we study some relations between these concepts and the notion of bounded approximate Karush–Kuhn–Tucker condition, which is introduced in this paper.

Keywords Approximate optimality conditions · Sequential optimality conditions · Enhanced Fritz John conditions · Enhanced Karush–Kuhn–Tucker conditions · Vector optimization problems

Mathematics Subject Classification 90C29 · 90C46

Communicated by Fabian Flores-Bazán.

✉ Vicente Novo
vnovo@ind.uned.es

Giorgio Giorgi
ggiorgi@eco.unipv.it

Bienvenido Jiménez
bjimenez@ind.uned.es

¹ University of Pavia, Pavia, Italy

² Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain

1 Introduction

Karush–Kuhn–Tucker (KKT) optimality conditions play an important role in optimization theory, both for scalar optimization and for multiobjective optimization. The KKT conditions are satisfied at a weak efficient point, provided a suitable constraint qualification holds.

Other kind of optimality conditions, named Fritz John-type conditions, do not require any constraint qualification. Such conditions are true in the scalar case and also for the multiobjective problem.

Another type of necessary optimality conditions in scalar problems, that do not require a constraint qualification, are the so-called “approximate optimality conditions” or also “asymptotic optimality conditions” or “sequential optimality conditions,” in which appropriate sequences of points and multipliers are considered. These conditions are investigated since several decades. We can quote the works of Kortanek and Evans [1], for pseudoconcave programming, the paper of Fiacco and McCormick [2], the paper of Zlobec [3], where this last author generalizes, in an asymptotic way, the classic optimality conditions given by Guignard and the book of Hestenes [4]. Asymptotic versions of the Karush–Kuhn–Tucker conditions are considered also by McShane [5], Craven [6] and Trudzik [7].

Recently, several works have been devoted to the study of such conditions for the interest they have in the design and analysis of algorithms to check this kind of approximate optimality conditions; see, for example, Haeser and Schuverdt [8], Andreani, Haeser and Martínez [9], Andreani, Martínez and Svaiter [10], Dutta et al. [11] and Haeser and Melo [12].

Despite of its great interest, this type of optimality conditions has not been studied in multiobjective optimization; in our knowledge this paper is the first dedicated to this topic.

In this paper, we deal with the study of the approximate KKT condition for a continuously differentiable multiobjective problem in finite dimensional spaces, whose feasible set is defined by inequality and equality constraints. In the main result (Sect. 3) we extend to this context the necessary optimality conditions obtained for scalar problems via approximate KKT conditions. We also prove that these conditions are also sufficient conditions under convexity assumptions. Illustrative examples are given.

In Sect. 4 we obtain a necessary optimality condition of enhanced Fritz John-type (in this case, in addition to the classic Fritz John conditions, the multipliers satisfy some additional sensitivity-like conditions). This class of multipliers was obtained by Bertsekas and Ozdaglar [13] for scalar optimization problems. Moreover, we study the relationships between these conditions and the above one, we introduce the notions of enhanced (respectively weak enhanced) KKT-point and state that the approximate KKT condition is sufficient to have an enhanced KKT-point under a weak constraint qualification, called quasi-normality constraint qualification. We also introduce the notion of bounded approximate KKT condition and study the relations with the above concepts.

2 Preliminaries

Recall that, for a scalar function, the so-called Clarke subdifferential is defined as follows [14]: The upper Clarke directional derivative of a locally Lipschitz function $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ at $x \in \mathbb{R}^n$ in the direction $d \in \mathbb{R}^n$ is

$$\varphi^\circ(x, d) := \limsup_{y \rightarrow x, t \downarrow 0} \frac{\varphi(y + td) - \varphi(y)}{t},$$

and the Clarke subdifferential of φ at x is given by

$$\partial_C \varphi(x) := \{\xi \in \mathbb{R}^n : \langle \xi, d \rangle \leq \varphi^\circ(x, d) \forall d \in \mathbb{R}^n\},$$

where $\langle \cdot, \cdot \rangle$ stands for the Euclidean scalar product. When φ is continuously differentiable, one has $\partial_C \varphi(x) = \{\nabla \varphi(x)\}$. Next, we collect some properties of this subdifferential that can be found in [14].

Proposition 2.1 (i) *Let $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ be locally Lipschitz and $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ continuously differentiable. Then,*

$$\partial_C(\varphi + \psi)(x) = \partial_C \varphi(x) + \{\nabla \psi(x)\}.$$

(ii) *If x^0 is a local minimum of φ , then $0 \in \partial_C \varphi(x^0)$.*

(iii) *If $\varphi(x) = \max\{\varphi_1(x), \dots, \varphi_p(x)\}$, where $\varphi_1, \dots, \varphi_p : \mathbb{R}^n \rightarrow \mathbb{R}$ are continuously differentiable, then*

$$\partial_C \varphi(x) = \text{conv}\{\nabla \varphi_j(x) : j = 1, \dots, p \text{ such that } \varphi_j(x) = \varphi(x)\},$$

(here conv denotes the convex hull).

We denote by $B(x^0, \delta)$ (respectively $\bar{B}(x^0, \delta)$) the open (respectively closed) ball centered at $x^0 \in \mathbb{R}^n$ and radius $\delta > 0$. In general, for a vector $x \in \mathbb{R}^n$, we denote its components by x_i ($i = 1, \dots, n$). We denote by \mathbb{R}_+^n the nonnegative orthant of \mathbb{R}^n .

For $a \in \mathbb{R}$, we denote $a_+ := \max\{0, a\}$, $a_+^2 := (a_+)^2$, $\|\cdot\|$ is the Euclidean norm of \mathbb{R}^n except otherwise is specified. For $y, z \in \mathbb{R}^n$, as usual we write $y \leq z$ if $y_i \leq z_i$ for $i = 1, \dots, n$; $y < z$ if $y_i < z_i$ for $i = 1, \dots, n$.

We consider the following multiobjective optimization problem:

$$(\text{MOP}) \min_{\mathbb{R}_+^p} f(x) \text{ subject to } x \in S,$$

where $S := \{x \in \mathbb{R}^n : g(x) \leq 0, h(x) = 0\}$, $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $h : \mathbb{R}^n \rightarrow \mathbb{R}^r$ are continuously differentiable functions. The components of these functions are f_l ($l = 1, \dots, p$), g_j ($j = 1, \dots, m$) and h_i ($i = 1, \dots, r$). The set of active indexes at a point $x \in S$ is given by $J(x) := \{j : g_j(x) = 0\}$.

A point $x^0 \in S$ is an efficient (respectively weak efficient) solution for (MOP) iff there exists no $x \in S$ such that $f(x) \leq f(x^0)$, $f(x) \neq f(x^0)$ (respectively

$f(x) < f(x^0)$). The set of all efficient (respectively weak efficient) solutions of (MOP) is denoted by $\text{Min}(f, S)$ (respectively $\text{WMin}(f, S)$). A point $x^0 \in S$ is a local weak efficient solution for (MOP) if there exists a neighborhood U of x^0 such that x^0 is a weak efficient solution on $S \cap U$.

3 Approximate KKT Condition for Multiobjective Optimization Problems

Next we introduce the concept of approximate Karush–Kuhn–Tucker condition for the multiobjective problem (MOP) inspired by the works by Andreani, Haeser and Martínez [9], Andreani, Martínez and Svaiter [10], Haeser and Schuverdt [8] and Dutta et al. [11]. We point out that all the mentioned authors are concerned with scalar optimization problems.

Definition 3.1 We say that the approximate Karush–Kuhn–Tucker condition (AKKT) is satisfied for (MOP) at a feasible point $x^0 \in S$ iff there exist sequences $(x^k) \subset \mathbb{R}^n$ and $(\lambda^k, \mu^k, \tau^k) \subset \mathbb{R}_+^p \times \mathbb{R}_+^m \times \mathbb{R}^r$ such that

$$(A0) \quad x^k \rightarrow x^0,$$

$$(A1) \quad \sum_{l=1}^p \lambda_l^k \nabla f_l(x^k) + \sum_{j=1}^m \mu_j^k \nabla g_j(x^k) + \sum_{i=1}^r \tau_i^k \nabla h_i(x^k) \rightarrow 0,$$

$$(A2) \quad \sum_{l=1}^p \lambda_l^k = 1,$$

$$(A3) \quad g_j(x^0) < 0 \Rightarrow \mu_j^k = 0 \text{ for sufficiently large } k, \quad j = 1, \dots, m.$$

This condition is introduced here for the first time in multiobjective optimization problems, and it extends in a natural way the AKKT condition studied in scalar optimization problems (see, for example, [8]). Points satisfying the AKKT condition will be called AKKT points. Let us observe that the sequence of points (x^k) is not required to be feasible.

Remark 3.1 Assuming $\mu^k \in \mathbb{R}_+^m$, condition (A3) is clearly equivalent to

$$\mu_j^k g_j(x^k) \geq 0 \text{ for sufficiently large } k, \forall j \notin J(x^0). \tag{1}$$

Each of these conditions implies the condition

$$\mu_j^k g_j(x^k) \rightarrow 0 \quad \forall j \notin J(x^0). \tag{2}$$

In order to establish necessary optimality conditions for problem (MOP), we are going to scalarize it. To this aim, we consider the nonsmooth function $\phi : \mathbb{R}^p \rightarrow \mathbb{R}$ defined by

$$\phi(y) := \max_{1 \leq i \leq p} \{y_i\}.$$

Clearly, $\phi(y) \leq 0 \Leftrightarrow y \leq 0$, and $\phi(y) < 0 \Leftrightarrow y < 0$.

Furthermore the following result is well known. We report here for the reader's convenience the result and its proof.

Lemma 3.1 *If $x^0 \in \text{WMin}(f, S)$, then $x^0 \in \text{Min}(\phi(f(\cdot) - f(x^0)), S)$.*

Proof Suppose that $x^0 \notin \text{Min}(\phi(f(\cdot) - f(x^0)), S)$, then there exists $\hat{x} \in S$ such that $\phi(f(\hat{x}) - f(x^0)) < \phi(f(x^0) - f(x^0)) = 0$. It follows that $f(\hat{x}) - f(x^0) < 0$, which contradicts the hypothesis. \square

Theorem 3.1 *If $x^0 \in S$ is a local weak efficient solution of problem (MOP), then x^0 satisfies the AKKT condition with sequences (x^k) and $(\lambda^k, \mu^k, \tau^k)$. In addition, for these sequences we have that*

$$(E1) \quad \mu^k = b_k g(x^k)_+ \text{ and } \tau^k = c_k h(x_k), \text{ where } b_k, c_k > 0, \forall k,$$

$$(E2) \quad f_l(x^k) - f_l(x^0) + \sum_{j=1}^m \mu_j^k g_j(x^k) + \sum_{i=1}^r \tau_i^k h_i(x^k) \leq 0, \quad \forall k, l = 1, \dots, p.$$

Proof By assumption, there exists $\delta > 0$ such that $x^0 \in \text{WMin}(f, S \cap B(x^0, \delta))$, and by Lemma 3.1, we deduce that $x^0 \in \text{Min}(\phi(f(\cdot) - f(x^0)), S \cap B(x^0, \delta))$. In consequence, we may suppose that x^0 is the unique solution of problem

$$\text{Minimize } \phi(f(x) - f(x^0)) + \frac{1}{2} \|x - x^0\|^2, \text{ subject to } x \in S, \|x - x^0\| \leq \delta \quad (3)$$

(choosing a small δ if necessary).

We define, for each $k \in \mathbb{N}$,

$$\psi_k(x) := \phi(f(x) - f(x^0)) + \frac{1}{2} \|x - x^0\|^2 + \sum_{j=1}^m k g_j(x)_+^2 + \sum_{i=1}^r k h_i(x)^2,$$

and let x^k be a solution of problem

$$\text{Minimize } \psi_k(x), \text{ subject to } \|x - x^0\| \leq \delta. \quad (4)$$

Let us observe that x^k exists because ψ_k is continuous and $\bar{B}(x^0, \delta)$ is compact.

Let z be an accumulation point of (x^k) . We may suppose that $x^k \rightarrow z$ (choosing a subsequence if necessary).

On one hand, we have

$$\phi(f(x^k) - f(x^0)) \leq \psi_k(x^k)$$

because

$$\psi_k(x^k) - \phi(f(x^k) - f(x^0)) = \frac{1}{2} \|x^k - x^0\|^2 + \sum_{j=1}^m k g_j(x^k)_+^2 + \sum_{i=1}^r k h_i(x^k)^2 \geq 0.$$

On the other hand, as x^0 is a feasible point of problem (4) and x^k is a solution, one has

$$\psi_k(x^k) \leq \psi_k(x^0) = 0 \quad \forall k \in \mathbb{N} \tag{5}$$

since $x^0 \in S$ and $\phi(f(x^0) - f(x^0)) = \phi(0) = 0$.

Let us prove that z is a feasible point of problem (3).

Indeed, first as $\|x^k - x^0\| \leq \delta$ it follows that $\|z - x^0\| \leq \delta$.

Second, suppose that

$$\sum_{j=1}^m g_j(z)_+^2 + \sum_{i=1}^r h_i(z)^2 > 0.$$

Then, there exists $c > 0$ such that

$$\sum_{j=1}^m g_j(x^k)_+^2 + \sum_{i=1}^r h_i(x^k)^2 > c$$

for all k large enough, by continuity and because $x^k \rightarrow z$. Now, as

$$\begin{aligned} \psi_k(x^k) &= \phi(f(x^k) - f(x^0)) + \frac{1}{2}\|x^k - x^0\|^2 + k \left(\sum_{j=1}^m g_j(x^k)_+^2 + \sum_{i=1}^r h_i(x^k)^2 \right) \\ &> \phi(f(x^k) - f(x^0)) + kc, \end{aligned}$$

taking the limit we obtain $\psi_k(x^k) \rightarrow +\infty$, which contradicts (5). In consequence, $\sum_{j=1}^m g_j(z)_+^2 + \sum_{i=1}^r h_i(z)^2 = 0$, and this implies that $z \in S$.

From (5) one has

$$\psi_k(x^k) = \phi(f(x^k) - f(x^0)) + \frac{1}{2}\|x^k - x^0\|^2 + \sum_{j=1}^m k g_j(x^k)_+^2 + \sum_{i=1}^r k h_i(x^k)^2 \leq 0 \tag{6}$$

and, as $\sum_{j=1}^m k g_j(x^k)_+^2 + \sum_{i=1}^r k h_i(x^k)^2 \geq 0$, we get

$$\phi(f(x^k) - f(x^0)) + \frac{1}{2}\|x^k - x^0\|^2 \leq 0.$$

Taking the limit it results

$$\phi(f(z) - f(x^0)) + \frac{1}{2}\|z - x^0\|^2 \leq 0.$$

As x^0 is the unique solution of problem (3) with value 0, we conclude that $z = x^0$. Therefore, $x^k \rightarrow x^0$ and $\|x^k - x^0\| < \delta$ for all k sufficiently large.

Now, as x^k is a solution of the nonsmooth problem (4) and it is an interior point of the feasible set, for k large enough, from Proposition 2.1(ii) it follows that $0 \in \partial_C \psi_k(x^k)$. By applying Proposition 2.1, parts (i) and (iii), we have

$$0 \in \text{conv} \left(\bigcup_{1 \leq l \leq p} \{\nabla f_l(x^k)\} \right) + x^k - x^0 + \sum_{j=1}^m k g_j(x^k)_+ \nabla g_j(x^k) + \sum_{i=1}^r k h_i(x^k) \nabla h_i(x^k).$$

Hence, there exist $\lambda_l^k \geq 0, l = 1, \dots, p$, such that $\sum_{l=1}^p \lambda_l^k = 1$ and

$$\sum_{l=1}^p \lambda_l^k \nabla f_l(x^k) + \sum_{j=1}^m k g_j(x^k)_+ \nabla g_j(x^k) + \sum_{i=1}^r k h_i(x^k) \nabla h_i(x^k) = x^0 - x^k \rightarrow 0.$$

Choosing for each k

$$\mu_j^k = k g_j(x^k)_+ \geq 0 \ (j = 1, \dots, m) \text{ and } \tau_i^k = k h_i(x^k) \ (i = 1, \dots, r) \quad (7)$$

we see that x^0 satisfies conditions (A1), (A2) and (E1). If $g_j(x^0) < 0$, then $g_j(x^k) < 0$ for all k large enough, and so $\mu_j^k = k g_j(x^k)_+ = 0$; therefore (A3) is also satisfied.

From (7), clearly for all $j = 1, \dots, m$ and for all $i = 1, \dots, r$,

$$\mu_j^k g_j(x^k) = \mu_j^k g_j(x^k)_+ = k g_j(x^k)_+^2 \text{ and } \tau_i^k h_i(x^k) = k h_i(x^k)^2 \quad \forall k.$$

In consequence, (6) can be rewritten as

$$\phi(f(x^k) - f(x^0)) + \frac{1}{2} \|x^k - x^0\|^2 + \sum_{j=1}^m \mu_j^k g_j(x^k) + \sum_{i=1}^r \tau_i^k h_i(x^k) \leq 0.$$

From here, condition (E2) follows and the proof is finished. □

We highlight that in Theorem 3.1, we do not require any constraint qualification, that is, these necessary optimality conditions are true for any local weak efficient solution without additional requirements. Let us observe also that condition (E1) says that the multipliers μ^k and τ^k are proportional, respectively, to $g(x^k)_+$ and $h(x^k)$. Condition (E1) is usually satisfied, for example, when the external penalty method is applied [8].

Next we illustrate Theorem 3.1 with an example.

Example 3.1 Consider the following multiobjective problem:

$$\text{Min } f(x_1, x_2) = (f_1(x_1, x_2), f_2(x_1, x_2)) \text{ subject to } g(x_1, x_2) = x_2^2 - x_1 \leq 0,$$

where

$$f_1(x_1, x_2) = 4x_1 - x_2^2, \quad f_2(x_1, x_2) = -2x_1 - x_2.$$

Let us note that it is a nonconvex problem. The point $x^0 = (1, 1)$ is a weak efficient solution as can be checked, and so Theorem 3.1 can be applied. In order to find sequences satisfying (A0)–(A3), (E1) and (E2), first let us solve the equation

$$\lambda_1 \nabla f_1(x_1, x_2) + \lambda_2 \nabla f_2(x_1, x_2) + \mu \nabla g(x_1, x_2) = (0, 0), \tag{8}$$

with $\lambda_1, \lambda_2, \mu \geq 0, \lambda_1 + \lambda_2 = 1$. We obtain for $x_2 < -1$ or $x_2 \geq 0$,

$$\lambda_1 = \frac{4x_2 + 1}{10x_2 + 1}, \quad \lambda_2 = \frac{6x_2}{10x_2 + 1}, \quad \mu = \frac{4x_2 + 4}{10x_2 + 1}. \tag{9}$$

Second, consider the nonfeasible points $x^\varepsilon = (1, 1 + 6\varepsilon), \varepsilon > 0$. Then, from (9) it results

$$\lambda_1^\varepsilon = \frac{5 + 24\varepsilon}{11 + 60\varepsilon}, \quad \lambda_2^\varepsilon = \frac{6 + 36\varepsilon}{11 + 60\varepsilon}, \quad \bar{\mu}^\varepsilon = \frac{8 + 24\varepsilon}{11 + 60\varepsilon}.$$

Let $\mu^\varepsilon = \frac{8}{11+60\varepsilon}$ and $\tilde{\mu}^\varepsilon = \frac{24\varepsilon}{11+60\varepsilon}$ so that $\bar{\mu}^\varepsilon = \mu^\varepsilon + \tilde{\mu}^\varepsilon$. In view of (8), one has $\lambda_1^\varepsilon \nabla f_1(x^\varepsilon) + \lambda_2^\varepsilon \nabla f_2(x^\varepsilon) + (\mu^\varepsilon + \tilde{\mu}^\varepsilon) \nabla g(x^\varepsilon) = 0, \forall \varepsilon > 0$. From here,

$$\lambda_1^\varepsilon \nabla f_1(x^\varepsilon) + \lambda_2^\varepsilon \nabla f_2(x^\varepsilon) + \mu^\varepsilon \nabla g(x^\varepsilon) = -\tilde{\mu}^\varepsilon \nabla g(x^\varepsilon) \rightarrow 0 \text{ when } \varepsilon \downarrow 0.$$

Clearly, (A0)–(A3) are satisfied with $(x^\varepsilon, \lambda_1^\varepsilon, \lambda_2^\varepsilon, \mu^\varepsilon)$ (if necessary we transform the points in sequences selecting $\varepsilon = \frac{1}{k}$). Moreover, (E1) is fulfilled selecting $b_\varepsilon = \frac{\mu^\varepsilon}{g(x^\varepsilon)_+} > 0$. Condition (E2) is also satisfied since (after some calculations)

$$f_1(x^\varepsilon) - f_1(x^0) + \mu^\varepsilon g(x^\varepsilon) = \frac{-(2160\varepsilon + 588)\varepsilon^2}{11 + 60\varepsilon} < 0 \quad \forall \varepsilon > 0,$$

$$f_2(x^\varepsilon) - f_2(x^0) + \mu^\varepsilon g(x^\varepsilon) = \frac{-192\varepsilon^2}{11 + 60\varepsilon} < 0 \quad \forall \varepsilon > 0.$$

Conditions (E1) and (E2) are good properties as it is showed in Remark 3.2 and in Theorem 4.1, but it may not be easy to find a sequence with such properties. However, sequences satisfying (A0)–(A3) are easily obtained. As instance, if x^0 is a KKT-point (see Definition 4.4) and $\sum_{l=1}^p \lambda_l = 1$, then every sequence $(x^k, \lambda^k, \mu^k, \tau^k) \subset \mathbb{R}^n \times \mathbb{R}_+^p \times \mathbb{R}_+^m \times \mathbb{R}^r$ converging to $(x^0, \lambda, \mu, \tau)$ satisfy (A0)–(A3) whenever $\mu_j^k = 0$ for sufficiently large k if $g_j(x^0) < 0$.

The reciprocal of Theorem 3.1 is not true, as the following example shows.

Example 3.2 Consider problem (MOP) with the following data:

$$f_1 = x_1 - 2x_2^2, \quad f_2 = -x_2, \quad g = -x_1 \leq 0,$$

$$x^0 = (0, 0), \quad x^k = \left(-\frac{1}{k}, \frac{1}{k}\right), \quad \lambda_1^k = 1, \quad \lambda_2^k = 0, \quad \mu^k = 1.$$

One has

$$\lambda_1^k \nabla f_1(x^k) + \lambda_2^k \nabla f_2(x^k) + \mu^k \nabla g(x^k) = \left(0, \frac{-4}{k}\right) \rightarrow (0, 0).$$

Therefore conditions (A0)–(A3) are fulfilled. Moreover, (E1) is satisfied with $b_k = \frac{\mu^k}{g(x^k)_+} = k$. Condition (E2) is also fulfilled since

$$\begin{aligned} f_1(x^k) - f_1(x^0) + \mu^k g(x^k) &= \frac{-2}{k^2} < 0 \quad \forall k, \\ f_2(x^k) - f_2(x^0) + \mu^k g(x^k) &= 0 \quad \forall k. \end{aligned}$$

However x^0 is not a local weak efficient solution, because for the feasible points $x(t) = (t^2, t)$ one has $f(x(t)) = (-t^2, -t) < f(x^0)$ for all $t > 0$.

In the following remark we study the relationships between conditions (E1) and (E2) and other conditions, some of them already considered in the literature.

Remark 3.2 The following implications are true:

- (i) Condition (E1) implies the following condition (sign condition, SGN in short): for every k one has

$$(SGN) \quad \mu_j^k g_j(x^k) \geq 0 \quad (\forall j = 1, \dots, m) \quad \text{and} \quad \tau_i^k h_i(x^k) \geq 0 \quad (\forall i = 1, \dots, r), \quad (10)$$

and, moreover, $\mu_j^k > 0 \Leftrightarrow g_j(x^k) > 0$, and $\tau_i^k \neq 0 \Leftrightarrow h_i(x^k) \neq 0$.

- (ii) Conditions (A0), (E2) and (SGN) imply the following convergence condition:

$$(CVG) \quad \mu_j^k g_j(x^k) \rightarrow 0 \quad (\forall j = 1, \dots, m) \quad \text{and} \quad \tau_i^k h_i(x^k) \rightarrow 0 \quad (\forall i = 1, \dots, r). \quad (11)$$

- (iii) (CVG) implies the following condition, that we call sum converging to zero condition:

$$(SCZ) \quad \sum_{j=1}^m \mu_j^k g_j(x^k) + \sum_{i=1}^r \tau_i^k h_i(x^k) \rightarrow 0.$$

- (iv) (SGN) and (SCZ) imply (CVG).

Indeed, part (i) is immediate since $\mu_j^k g_j(x^k) = b_k g_j(x^k)^2$ if $g_j(x^k) \geq 0$ and $\mu_j^k g_j(x^k) = 0$ if $g_j(x^k) < 0$. So $\mu_j^k g_j(x^k) \geq 0$. Similarly, $\tau_i^k h_i(x^k) = c_k h_i(x^k)^2 \geq 0$. The rest of part (i) is obvious.

For part (ii), let $s_k := \sum_{j=1}^m \mu_j^k g_j(x^k) + \sum_{i=1}^r \tau_i^k h_i(x^k)$. Using (SGN) and (E2), we have $0 \leq s_k \leq f_l(x^0) - f_l(x^k)$ for all k and l . From here, $s_k \rightarrow 0$ follows by the continuity of f_l and property (A0).

Parts (iii)–(iv) are obvious.

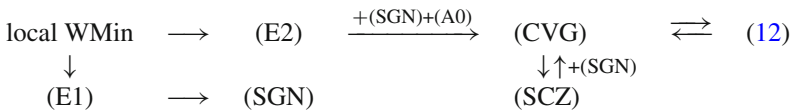
Taking into account Remark 3.2(i), from (E2) it follows that $f_l(x^k) \leq f_l(x^0)$, $l = 1, \dots, p$, i.e., the points x^k are better than x^0 . This condition is similar to the property given by (22), which is used further on to define a strong EFJ-point. As a consequence, if we wish a sequence (x^k) satisfying (E1) and (E2), then we have to look for it in the set defined by the system $f_l(x) \leq f_l(x^0)$, $l = 1, \dots, p$.

Remark 3.3 In some works, in order to define the AKKT condition for scalar optimization problems, several variants of condition (A3) or some of the above ones are used. For example, in [10] the authors use the CAKKT condition, in which (A3) is replaced with

$$\sum_{j=1}^m |\mu_j^k g_j(x^k)| + \sum_{i=1}^r |\tau_i^k h_i(x^k)| \rightarrow 0, \tag{12}$$

which is clearly equivalent to (CVG).

We summarize the above relations in the following diagram (where the arrow means “implies”):



In next remark, we show that the reciprocal implications of Remark 3.2 are not true and the invalidity if any assumption is not satisfied.

Remark 3.4 (i) (SGN) $\not\Rightarrow$ (E1). Consider Example 3.1, with $x^0 = (1, 1)$, $x^k = (1 + \frac{1}{k}, 1)$, $\lambda_1^k = \frac{5}{11}$, $\lambda_2^k = \frac{6}{11}$, $\mu^k = \frac{8}{11}$. With these data, (SGN) is satisfied but (E1) is not since $g(x^k)_+ = 0$, and so $b_k g(x^k)_+ = 0 \neq \mu^k$ for all $b_k > 0$.

(ii) (CVG) $\not\Rightarrow$ (E2) even if (SGN) holds. Consider Example 3.1, with $x^0 = (1, 1)$, $x^k = (1 - \frac{1}{k}, 1)$, $\lambda_1^k = \frac{5}{11}$, $\lambda_2^k = \frac{6}{11}$, $\mu^k = \frac{8}{11}$. One has $\mu^k g(x^k) \rightarrow 0$ and (SGN) is satisfied, but (E2) is not since

$$f_2(x^k) - f_2(x^0) + \mu^k g(x^k) = \frac{6k - 8}{11k^2} > 0 \quad \forall k > 1.$$

(iii) (E2) $\not\Rightarrow$ (CVG) if (SGN) is not satisfied. Consider the following data in problem (MOP):

$$f_1 = x_1 + x_2, \quad f_2 = -x_1 + x_2, \quad g_1 = x_1^2 - x_2 \leq 0, \quad g_2 = -2x_1^2 + x_2 \leq 0,$$

$x^0 = (0, 0)$, $x^k = (0, -\frac{1}{k})$, $\lambda_1^k = \frac{1}{2}$, $\lambda_2^k = \frac{1}{2}$, $\mu_1^k = 1 + k$, $\mu_2^k = k$. One has that (E2) is satisfied since

$$f_l(x^k) - f_l(x^0) + \mu_1^k g_1(x^k) + \mu_2^k g_2(x^k) = 0 \quad \forall k, \quad l = 1, 2,$$

however (CVG) is not satisfied because $\mu_1^k g_1(x^k) = (1+k) \frac{1}{k} = \frac{1}{k} + 1 \rightarrow 1$.
 (iv) (SCZ) $\not\Rightarrow$ (CVG) if (SGN) is not satisfied. The same data of part (iii) show this fact since

$$\mu_1^k g_1(x^k) + \mu_2^k g_2(x^k) = (1+k) \frac{1}{k} + k \frac{-1}{k} = \frac{1}{k} \rightarrow 0,$$

so (SCZ) holds but (CVG) does not.

Theorem 3.1 extends to multiobjective optimization (and improves in some cases) Theorem 3.3 in Andreani, Martínez and Svaiter [10], Theorem 2.1 in Haeser and Schuverdt [8] and Theorem 2.1 (with $I = \emptyset$) in Andreani, Haeser and Martínez [9].

Next we establish that the reciprocal of Theorem 3.1 is true for convex programs.

Theorem 3.2 *Assume that f_l ($l = 1, \dots, p$) and g_j ($j = 1, \dots, m$) are convex and h_i ($i = 1, \dots, r$) are affine. If $x^0 \in S$ satisfies the AKKT condition and (SCZ) is fulfilled, then x^0 is a (global) weak efficient solution of (MOP).*

Proof Suppose that x^0 is not a weak efficient solution. Then, there exists $\hat{x} \in S$ such that

$$f_l(\hat{x}) < f_l(x^0) \quad l = 1, \dots, p. \tag{13}$$

Let (x^k) , and $(\lambda^k, \mu^k, \tau^k)$ be the sequences that satisfy (A0)–(A3). Without any loss of generality we may assume that $\lambda^k \rightarrow \lambda^0$, with $\lambda^0 \geq 0$ and $\sum_{l=1}^p \lambda_l^0 = 1$. As f_l, g_j are convex and h_i are affine, for all k one has

$$f_l(\hat{x}) \geq f_l(x^k) + \nabla f_l(x^k)(\hat{x} - x^k) \quad \forall l = 1, \dots, p, \tag{14}$$

$$g_j(\hat{x}) \geq g_j(x^k) + \nabla g_j(x^k)(\hat{x} - x^k) \quad \forall j = 1, \dots, m, \tag{15}$$

$$h_i(\hat{x}) = h_i(x^k) + \nabla h_i(x^k)(\hat{x} - x^k) \quad \forall i = 1, \dots, r. \tag{16}$$

Multiplying (14) by λ_l^k , (15) by μ_j^k and (16) by τ_i^k and adding up, it results (the first inequality is valid because $\hat{x} \in S$)

$$\begin{aligned} \sum_{l=1}^p \lambda_l^k f_l(\hat{x}) &\geq \sum_{l=1}^p \lambda_l^k f_l(x^k) + \sum_{j=1}^m \mu_j^k g_j(x^k) + \sum_{i=1}^r \tau_i^k h_i(x^k) + \gamma_k \\ &\geq \sum_{l=1}^p \lambda_l^k f_l(x^k) + \sum_{j=1}^m \mu_j^k g_j(x^k) + \sum_{i=1}^r \tau_i^k h_i(x^k) + \gamma_k, \end{aligned} \tag{17}$$

where $\gamma_k = (\sum_{l=1}^p \lambda_l^k \nabla f_l(x^k) + \sum_{j=1}^m \mu_j^k \nabla g_j(x^k) + \sum_{i=1}^r \tau_i^k \nabla h_i(x^k))(\hat{x} - x^k)$. As $x^k \rightarrow x^0, \gamma_k \rightarrow 0$ by (A1) and $\sum_{j=1}^m \mu_j^k g_j(x^k) + \sum_{i=1}^r \tau_i^k h_i(x^k) \rightarrow 0$ by (SCZ), taking the limit in (17) we obtain

$$\sum_{l=1}^p \lambda_l^0 f_l(\hat{x}) \geq \sum_{l=1}^p \lambda_l^0 f_l(x^0). \tag{18}$$

As $\lambda^0 \geq 0$ and $\lambda^0 \neq 0$, from (13) it follows that $\sum_{i=1}^p \lambda_i^0 f_i(\hat{x}) < \sum_{i=1}^p \lambda_i^0 f_i(x^0)$, which contradicts (18). \square

This theorem extends to multiobjective optimization Theorem 4.2 in Andreani, Martínez and Svaiter [10] and Theorem 2.2 in Haeser and Schuverdt [8].

We illustrate Theorem 3.2 with the following example.

Example 3.3 Consider problem (MOP) with the following data:

$$f_1 = x_1^2 + x_2^2, \quad f_2 = -x_1, \quad g = x_1 - 2 \leq 0,$$

$$x^0 = (2, 1), \quad x^k = (2, 1), \quad \lambda_1^k = 0, \quad \lambda_2^k = 1, \quad \mu^k = 1.$$

With these data, x^0 satisfies AKKT and moreover (SCZ) holds, so by applying Theorem 3.2 we conclude that x^0 is a global weak efficient solution.

The conclusion of Theorem 3.2 cannot be that x^0 is a global efficient solution as this example shows, because x^0 is not efficient since $f(2, 0) \leq f(x^0)$ and $f(2, 0) \neq f(x^0)$.

4 Relations of the AKKT Condition with Other Optimality Conditions

Next we state necessary optimality conditions, which are alternative to the conditions of Theorem 3.1 and we study the relations with AKKT conditions. These conditions, in scalar optimization, are called by Bertsekas and Ozdaglar [13] “enhanced Fritz John” conditions because the associated multipliers have some additional properties (see also Hestenes [4, Theorem5.7.1]).

Definition 4.1 We say that a point $x^0 \in S$ satisfies the enhanced Fritz John conditions (EFJ) (or it is an EFJ-point) iff there exist $(\lambda, \mu, \tau) \in \mathbb{R}_+^p \times \mathbb{R}_+^m \times \mathbb{R}^r$ such that

(EFJ1) $\sum_{i=1}^p \lambda_i \nabla f_i(x^0) + \sum_{j=1}^m \mu_j \nabla g_j(x^0) + \sum_{i=1}^r \tau_i \nabla h_i(x^0) = 0,$

(EFJ2) $(\lambda, \mu, \tau) \neq 0,$

(EFJ3) if the index set $J \cup I$ is nonempty, where

$$J = \{j : \mu_j \neq 0\}, \quad I = \{i : \tau_i \neq 0\},$$

then there exists a sequence $(x^k) \subset \mathbb{R}^p$ that converges to x^0 and is such that, for all k ,

$$\mu_j g_j(x^k) > 0, \quad \forall j \in J, \quad \tau_i h_i(x^k) > 0, \quad \forall i \in I, \tag{19}$$

$$g_j(x^k)_+ = o(w(x^k)), \quad \forall j \notin J, \quad |h_i(x^k)| = o(w(x^k)), \quad \forall i \notin I, \tag{20}$$

where

$$w(x) = \min \left\{ \min_{j \in J} \{g_j(x)_+\}, \min_{i \in I} \{|h_i(x)|\} \right\}. \tag{21}$$

We say that x^0 satisfies the strong EFJ-condition (or it is a strong EFJ-point) iff it satisfies (EFJ1)–(EFJ3) and, in addition, the sequence (x^k) in (EFJ3) satisfies

$$f(x^k) < f(x^0). \quad (22)$$

We say that x^0 satisfies the weak EFJ-condition (or it is a weak EFJ-point) iff it satisfies (EFJ1)–(EFJ3), but conditions (20)–(21) are removed.

Remark 4.1 Condition (EFJ3) implies $\mu_j = 0$ if $g_j(x^0) < 0$, i.e., the usual complementary condition, since if $\mu_j > 0$, then in view of (19) one has that $g_j(x^k) > 0$ for x^k near x^0 , and this is a contradiction.

Extending the above definition, we say that $x^0 \in S$ is an EKKT-point iff it satisfies the conditions (EFJ1), (EFJ3) and, moreover, $\lambda \neq 0$ instead of (EFJ2). Similarly it is defined a strong EKKT-point and a weak EKKT-point.

In scalar optimization, multipliers satisfying (EFJ1), (EFJ2) and (EFJ3) with (22) are called by Bertsekas and Ozdaglar [13] informative multipliers. See [13] for interesting properties of this class of multipliers. Let us observe that (EFJ3) with (22) says that, whenever $(\mu, \tau) \neq 0$, the constraints whose indexes are in the set $J \cup I$ can be violated by a sequence of (infeasible) points x^k converging to x^0 , that improve all the objectives f_j ; the remaining constraints, whose indices do not belong to $J \cup I$, may also be violated, but the degree of their violation is arbitrarily small relative to the other constraints according to (20)–(21). Another consequence of (19) is the following: if $g_j(x) \leq 0$ on some neighborhood of x^0 , then $\mu_j = 0$.

In order to obtain a multiplier rule with $\lambda \neq 0$, usually a constraint qualification is utilized. The following constraint qualification is one of the weakest and was introduced by Hestenes [4].

Definition 4.2 We say that $x^0 \in S$ satisfies the quasi-normality constraint qualification (QNCQ) iff there is not any multiplier $(\mu, \tau) \in \mathbb{R}_+^m \times \mathbb{R}^r$ such that

- (i) $(\mu, \tau) \neq 0$,
- (ii) $\sum_{j=1}^m \mu_j \nabla g_j(x^0) + \sum_{i=1}^r \tau_i \nabla h_i(x^0) = 0$,
- (iii) in every neighborhood of x^0 there is a point $x \in \mathbb{R}^n$ such that $\mu_j g_j(x) > 0$ for all j having $\mu_j \neq 0$, and $\tau_i h_i(x) > 0$ for all i having $\tau_i \neq 0$.

It is well known that quasi-normality implies regularity (see Hestenes [4, Theorem 5.8.1]), i.e., the Abadie constraint qualification is satisfied, and this property implies the existence of KKT multipliers at a local minimum.

Next, we establish that a point satisfying conditions AKKT and (E1)–(E2) is an EFJ-point, and if additionally quasi-normality holds, then it is an EKKT-point.

Theorem 4.1 Assume that $x^0 \in S$ satisfies the AKKT condition with the sequences (x^k) and $(\lambda^k, \mu^k, \tau^k)$ and condition (E1), with $c_k = b_k$ for all k . Then,

- (i) x^0 is an EFJ-point.
- (ii) If, in addition, condition (E2) is satisfied, then x^0 is a strong EFJ-point.

Proof We simultaneously prove parts (i) and (ii). Let $t_k := \|(\lambda^k, \mu^k, \tau^k)\|$, where $\| \cdot \|$ is the ℓ_1 -norm. Then, $t_k \geq 1$ since $\sum_{l=1}^p \lambda_l^k = 1$ and $\lambda^k \geq 0$. As $\|\frac{1}{t_k}(\lambda^k, \mu^k, \tau^k)\| = 1 \forall k$, without any loss of generality we may suppose that $\frac{1}{t_k}(\lambda^k, \mu^k, \tau^k) \rightarrow (\lambda^0, \mu^0, \tau^0)$ with $(\lambda^0, \mu^0, \tau^0) \neq 0$. Clearly $\lambda^0 \geq 0$ and $\mu^0 \geq 0$. Dividing in (A1) by t_k and taking the limit it results (EFJ1) (with $(\lambda, \mu, \tau) = (\lambda^0, \mu^0, \tau^0)$).

For checking condition (EFJ3), assume that $(\mu^0, \tau^0) \neq 0$ and set $s_k := \sum_{j=1}^m \mu_j^k g_j(x^k) + \sum_{i=1}^r \tau_i^k h_i(x^k)$. By Remark 3.2(i), one has that (SGN) holds.

If $\mu_j^0 > 0$, as $\mu_j^0 = \lim \frac{\mu_j^k}{t_k}$, then it follows that $\mu_j^k > 0$ for all k large enough. By Remark 3.2(i), we derive that $g_j(x^k) > 0$ [the first part of (19) holds] and so $\mu_j^k g_j(x^k) > 0$ and $\mu_j^k g_j(x^k) > 0$ for all k large enough. Therefore, taking into account (10), one has $s_k > 0$.

If $\tau_i^0 \neq 0$, as $\tau_i^0 = \lim \frac{\tau_i^k}{t_k} = \lim \frac{c_k h_i(x^k)}{t_k}$ one has $h_i(x^k) \neq 0$ and the second part of (19) follows. To prove (22), we analyze two possibilities:

- (a) If $\tau_i^0 > 0$, then $\tau_i^k > 0$ and $h_i(x^k) > 0$, so $\tau_i^k h_i(x^k) > 0$ and $\tau_i^0 h_i(x^k) > 0$ for all k large enough. In consequence, according with (10), one has $s_k > 0$.
- (b) If $\tau_i^0 < 0$, then $\tau_i^k < 0$ and $h_i(x^k) < 0$, so $\tau_i^k h_i(x^k) > 0$ and $\tau_i^0 h_i(x^k) > 0$ for all k large enough. In consequence, in view of (10), one has $s_k > 0$.

Therefore, if $(\mu^0, \tau^0) \neq 0$, then one has $s_k > 0$, and consequently, making use of (E2), $f_l(x^0) - f_l(x^k) < 0$ for all $l = 1, \dots, p$ and condition (22) is proved.

In order to prove (20), assume that in (E1) $c_k = b_k$ for all k . Then, $g_j(x^k)_+ = \frac{\mu_j^k}{b_k}$ for $j = 1, \dots, m$ and $|h_i(x^k)| = \frac{|\tau_i^k|}{b_k}$ for $i = 1, \dots, r$, and so

$$w(x^k) = \frac{1}{b_k} \min \left\{ \min_{j \in J} \{\mu_j\}, \min_{i \in I} \{|\tau_i|\} \right\}.$$

Therefore

$$\frac{g_{j_0}(x^k)_+}{w(x^k)} = \frac{g_{j_0}(x^k)_+/t_k}{w(x^k)/t_k} = \frac{\mu_{j_0}/t_k}{\min \{ \min_{j \in J} \{\mu_j/t_k\}, \min_{i \in I} \{|\tau_i|/t_k\} \}}.$$

Now, clearly this quotient tends to zero for $j_0 \notin J$ because $\mu_{j_0}/t_k \rightarrow \mu_{j_0}^0 = 0$ and the denominator is equal to $\varepsilon_k \geq \alpha > 0$ for some α and for all k large enough, since $\mu_j^k/t_k \rightarrow \mu_j^0 > 0 \forall j \in J$ and $|\tau_i^k|/t_k \rightarrow |\tau_i^0| > 0 \forall i \in I$. □

Remark 4.2 Note that the multipliers given in Theorem 3.1, which are defined by (7), satisfy condition (E1) with $b_k = c_k$.

Proposition 4.1 *If $x^0 \in S$ is an EFJ-point (respectively strong EFJ-point or weak EFJ-point) and (QNCQ) holds, then x^0 an EKKT-point (respectively strong EKKT-point or weak EKKT-point).*

Proof If $\lambda^0 = 0$, it follows that (QNCQ) does not hold since in every neighborhood of x^0 there is a point x^k that does not satisfy condition (iii) of the definition of (QNCQ). □

Part (ii) of Theorem 4.1, taking into account Proposition 4.1, extends to multiobjective optimization and improves Theorem 2.3 in Haeser and Schuverdt [8], which is true for scalar problems and where it is not required that the multipliers satisfy (20)–(22).

The following result is an immediate consequence of Theorems 3.1 and 4.1 taking into account Remark 4.2 and Proposition 4.1.

Theorem 4.2 *If $x^0 \in S$ is a local weak solution of (MOP), then*

- (i) x^0 is a strong EFJ-point.
- (ii) If, in addition, (QNCQ) holds at x^0 , then x^0 is a strong EKKT-point.

Part (i) of this theorem extends to multiobjective optimization and improves Theorem 5.7.1 in Hestenes [4] and Proposition 3.3.5 in Bertsekas [15] (in both, only it is proved that x^0 is a weak EFJ-point). It also extends (partially) Proposition 2.1 in Bertsekas and Ozdaglar [13], where a scalar problem is considered with equality and inequality constraints, but also with a set constraint.

Next we introduce a new notion of approximate KKT type, requiring boundedness of the multipliers, and then we will prove that this condition is stronger than AKKT, and we will study its relations with weak EKKT-points and KKT-points.

Definition 4.3 We say that the bounded approximate Karush–Kuhn–Tucker condition (BAKKT, in short) is satisfied at a feasible point $x^0 \in S$ iff there exist sequences $(x^k) \subset \mathbb{R}^n$ and $(\lambda^k, \mu^k, \tau^k) \subset \mathbb{R}_+^p \times \mathbb{R}_+^m \times \mathbb{R}^r$ such that conditions (A0)–(A3) hold and

$$\text{the sequence } (\lambda^k, \mu^k, \tau^k) \text{ is bounded.} \tag{23}$$

Clearly condition (23) can be replaced with the convergence condition of the multipliers:

$$\text{the sequence } (\lambda^k, \mu^k, \tau^k) \text{ converges to } (\lambda^0, \mu^0, \tau^0). \tag{24}$$

Proposition 4.2 *If $x^0 \in S$ satisfies (BAKKT), then x^0 satisfies also the AKKT condition and, in addition, (CVG) holds.*

Proof We only have to prove condition (CVG), but this follows immediately from (24), the continuity of g_j and h_i and the fact that $x^k \rightarrow x^0$. □

The converse implication is not true, as the next example shows.

Example 4.1 Consider problem (MOP) with the data $f(x_1, x_2) = x_1$ and $g(x_1, x_2) = -x_1^3 + x_2^2$. Clearly $x^0 = (0, 0)$ is a minimum of f subject to $g \leq 0$. So by Theorem 3.1, the AKKT condition is satisfied.

Let us check that (BAKKT) does not hold. Suppose that there exist sequences $x^k = (x_1^k, x_2^k) \in \mathbb{R}^2$ and $\mu^k \in \mathbb{R}_+$ such that $x^k \rightarrow x^0$ and

$$\nabla f(x^k) + \mu^k \nabla g(x^k) =: (w_1^k, w_2^k) \rightarrow (0, 0).$$

From here, $(1, 0) + \mu^k (-3(x_1^k)^2, 2x_2^k) = (w_1^k, w_2^k)$. Therefore there is solution $\mu^k = \frac{1-w_1^k}{3(x_1^k)^2} \rightarrow +\infty$ whenever $x_1^k \neq 0$ and $\frac{x_2^k}{(x_1^k)^2} \rightarrow 0$. So every solution (μ^k) is not bounded, and consequently (BAKKT) is not satisfied at x^0 . Note that this can also be deduced from Theorem 4.4 since x^0 is not a KKT-point.

Let us note that by Theorem 4.2(i), x^0 is an EFJ-point, and also that any constraint qualification does not hold because x^0 is not a KKT-point.

Next theorem proves that the converse of Proposition 4.2 is true if a constraint qualification holds at the point x^0 .

Recall that the Mangasarian–Fromovitz constraint qualification (MFCQ) holds at $x^0 \in S$ if the gradients $\nabla h_1(x^0), \dots, \nabla h_r(x^0)$ are linearly independent and there exists a vector $d \in \mathbb{R}^n$ such that $\nabla g_j(x^0)d < 0$ for all $j \in J(x^0)$ and $\nabla h_i(x^0)d = 0$ for all $i = 1, \dots, r$. It is well known that (MFCQ) is equivalent to the positive linearly independent constraint qualification [4], i.e., the following implication is true: $\mu_j \geq 0, j \in J(x^0), \tau \in \mathbb{R}^r,$

$$\sum_{j \in J(x^0)} \mu_j \nabla g_j(x^0) + \sum_{i=1}^r \tau_i \nabla h_i(x^0) = 0 \Rightarrow \mu_j = 0, \forall j \in J(x^0), \tau = 0.$$

It is also known that (MFCQ) is stronger than the so-called constant positive linear dependence condition, which implies (QNCQ) [16].

Theorem 4.3 *Assume that (MFCQ) holds at x^0 . Then, for every pair of sequences $(x^k) \subset \mathbb{R}^n$ and $(\lambda^k, \mu^k, \tau^k) \subset \mathbb{R}_+^p \times \mathbb{R}_+^m \times \mathbb{R}^r$ that satisfy (A0)–(A3), one has that the second sequence is bounded.*

Proof Let (s_k) be the sequence defined by

$$s_k := \left\| \sum_{l=1}^p \lambda_l^k \nabla f_l(x^k) + \sum_{j=1}^m \mu_j^k \nabla g_j(x^k) + \sum_{i=1}^r \tau_i^k \nabla h_i(x^k) \right\| \rightarrow 0. \tag{25}$$

If the sequence (μ^k, τ^k) is not bounded, then, without any loss of generality we can suppose that $t_k := \|(\mu^k, \tau^k)\| \rightarrow +\infty$ and that $\frac{1}{t_k}(\mu^k, \tau^k)$ converges to (μ^0, τ^0) with $\|(\mu^0, \tau^0)\| = 1$. Dividing (25) by t_k , one has

$$\left\| \sum_{l=1}^p \frac{\lambda_l^k}{t_k} \nabla f_l(x^k) + \sum_{j=1}^m \frac{\mu_j^k}{t_k} \nabla g_j(x^k) + \sum_{i=1}^r \frac{\tau_i^k}{t_k} \nabla h_i(x^k) \right\| = \frac{s_k}{t_k}.$$

The sequence (λ^k) is bounded since $\lambda \geq 0$ and (A2) holds. So $\lambda^k/t_k \rightarrow 0$, and as $s_k/t_k \rightarrow 0$, taking the limit it results

$$\sum_{j=1}^m \mu_j^0 \nabla g_j(x^0) + \sum_{i=1}^r \tau_i^0 \nabla h_i(x^0) = 0. \tag{26}$$

By (A3), $\mu_j^0 = 0$ for all $j \notin J(x^0)$. With this observation, (26) contradicts (MFCQ) because $(\mu^0, \tau^0) \neq 0$ and $\mu^0 \geq 0$. Therefore, the sequence (μ^k, τ^k) is bounded and so condition (23) holds. \square

Definition 4.4 We say that a feasible point $x^0 \in S$ is a KKT-point iff there exist $(\lambda, \mu, \tau) \in \mathbb{R}_+^p \times \mathbb{R}_+^m \times \mathbb{R}^r$ such that

$$(K1) \quad \sum_{l=1}^p \lambda_l \nabla f_l(x^0) + \sum_{j=1}^m \mu_j \nabla g_j(x^0) + \sum_{i=1}^r \tau_i \nabla h_i(x^0) = 0,$$

$$(K2) \quad \lambda \neq 0,$$

$$(K3) \quad \mu_j = 0 \text{ if } g_j(x^0) < 0, \quad j = 1, \dots, m.$$

Theorem 4.4 If $x^0 \in S$ satisfies (BAKKT), then x^0 is a KKT-point.

Proof Let $(x^k), (\lambda^k, \mu^k, \tau^k)$ be the sequences that satisfy (A0)–(A3) and (24). Then, taking the limit in (A1) and (A2), we obtain

$$\sum_{l=1}^p \lambda_l^0 \nabla f_l(x^0) + \sum_{j=1}^m \mu_j^0 \nabla g_j(x^0) + \sum_{i=1}^r \tau_i^0 \nabla h_i(x^0) = 0 \quad (27)$$

and $\sum_{l=1}^p \lambda_l^0 = 1$, with $\lambda^0 \geq 0, \mu^0 \geq 0$. The condition $\mu_j^0 g_j(x^0) = 0$ for all $j \notin J(x^0)$ follows from (A3). \square

The following result is a direct consequence of Theorems 4.3 and 4.4

Theorem 4.5 If $x^0 \in S$ satisfies the AKKT condition and (MFCQ) holds at x^0 , then x^0 is a KKT-point.

Following a notion due to Dutta et al. [11], given a sequence of positive numbers $\varepsilon_k \downarrow 0$ and points $x^k \rightarrow x^0$, we say that (x^k) is a sequence of ε_k -KKT-points if there exist $(\lambda^k, \mu^k, \tau^k) \in \mathbb{R}_+^p \times \mathbb{R}_+^m \times \mathbb{R}^r$ such that

$$\left\| \sum_{l=1}^p \lambda_l^k \nabla f_l(x^k) + \sum_{j=1}^m \mu_j^k \nabla g_j(x^k) + \sum_{i=1}^r \tau_i^k \nabla h_i(x^k) \right\| \leq \varepsilon_k,$$

$\mu_j^k g_j(x^k) = 0$ for $j = 1, \dots, m$ and (A2) holds.

Note that we do not require x^k to be feasible. It is obvious that, if (x^k) is a sequence of ε_k -KKT-points, then x^0 satisfies the AKKT condition, and so the following result is an immediate consequence of Theorem 4.5.

Proposition 4.3 If (x^k) is a sequence of ε_k -KKT-points and (MFCQ) holds at x^0 , then x^0 is a KKT-point.

Theorem 3.2 in Dutta et al. [11] is a particular case of Proposition 4.3. These authors only consider a scalar function f and inequality constraints.

Theorem 4.6 *If $x^0 \in S$ is a KKT-point, then x^0 is a weak EKKT-point. In addition, the sequence (x^k) satisfying (19) also satisfies*

$$\sum_{l=1}^p \lambda_l f_l(x^k) < \sum_{l=1}^p \lambda_l f_l(x^0).$$

Proof To give a unified treatment to the multipliers we transform the multipliers associated with equality constraints as follows: if $\tau_i \geq 0$, then we define $g_{m+i}(x) := h_i(x)$ and $\mu_{m+i} := \tau_i$; if $\tau_i < 0$, then we define $g_{m+i}(x) := -h_i(x)$ and $\mu_{m+i} := -\tau_i > 0$. In this way, (K1) can be written

$$\sum_{l=1}^p \lambda_l \nabla f_l(x^0) + \sum_{j=1}^s \mu_j \nabla g_j(x^0) = 0, \tag{28}$$

where $s = m + r$. Rearranging, we may assume that μ_1, \dots, μ_q are positive and μ_{q+1}, \dots, μ_s are zero. Let us observe that the indices j satisfying (K3) are in the second group. Through the following reduction process, we can assume that the gradients $\nabla g_1(x^0), \dots, \nabla g_q(x^0)$ are positive linearly independent.

Indeed, assume that there exist nonnegative numbers $\alpha_1, \dots, \alpha_q$, not all zero, such that $\sum_{j=1}^q \alpha_j \nabla g_j(x^0) = 0$. Then, the multipliers $\hat{\mu}_j = \mu_j - t\alpha_j$ ($j = 1, \dots, q$), $\hat{\mu}_j = \mu_j$ ($\forall j > q$) satisfy (28) for all $t \in \mathbb{R}$. So we can choose $t_0 > 0$ so that $\hat{\mu}_j \geq 0$ $\forall j = 1, \dots, q$ and $\hat{\mu}_{j_0} = 0$ for some index $j_0 \leq q$. So, rearranging we obtain a proper subset of subgradients $\nabla g_1(x^0), \dots, \nabla g_{q'}(x^0)$ and positive multipliers $\hat{\mu}_1, \dots, \hat{\mu}_{q'}$ with $q' < q$ (the rest of multipliers is zero).

If $q' = 0$ (i.e., if for t_0 , $\hat{\mu}_j = 0$ for all $j = 1, \dots, q$), then (28) is satisfied with $\mu_j = 0$, $j = 1, \dots, s$, and there is nothing to do in order to prove (EFJ3).

If $q' > 0$, repeating the process if necessary, we can assume that the gradients $\nabla g_1(x^0), \dots, \nabla g_{q'}(x^0)$ are positive linearly independent. So, by a theorem of the alternative, there exists $u \in \mathbb{R}^n$ such that

$$\nabla g_j(x^0)u > 0 \quad \forall j = 1, \dots, q.$$

As $g_j(x^0) = 0$, there exists $\delta > 0$ such that $g_j(x^0 + tu) > 0 \forall t \in]0, \delta[$ and for all $j = 1, \dots, q$. Therefore (22) is satisfied for this set of multipliers since $\mu_j g_j(x(t)) > 0 \forall t \in]0, \delta[$ and for all $j = 1, \dots, q$, where $x(t) = x^0 + tu$. Let us observe that, if $g_j = -h_i$ (and $\mu_j = -\tau_i > 0$), then one has $\tau_i h_i(x(t)) > 0$. This proves that x^0 is a weak EKKT-point.

Finally, if we apply (28) to the vector u , then it results that

$$\sum_{l=1}^p \lambda_l \nabla f_l(x^0)u = - \sum_{j=1}^s \mu_j \nabla g_j(x^0)u < 0,$$

and so for a suitable $\delta > 0$ we have that $\sum_{l=1}^p \lambda_l (f_l(x(t)) - f_l(x^0)) < 0$. □

Corollary 4.1 *The following statements are equivalent for $x^0 \in S$:*

- (a) x^0 is a KKT-point.
- (b) x^0 is a weak EKKT-point.
- (c) x^0 satisfies (BAKKT).

Proof (a) \Rightarrow (b) and (c) \Rightarrow (a) follow from Theorems 4.6 and 4.4, respectively. Let us prove the implication (b) \Rightarrow (c). Assume that x^0 is a weak EKKT-point. We may suppose that (EFJ1) is satisfied with $\sum_{l=1}^p \lambda_l = 1$. Then, we can choose $x^k = x^0$, $\lambda^k = \lambda$, $\mu^k = \mu$, $\tau^k = \tau$, and it is obvious that (BAKKT) is satisfied. \square

5 Conclusions

We have presented two type of results: sequential optimality conditions and enhanced FJ and KKT conditions for differentiable multiobjective optimization problems with equality and inequality constraints.

In scalar optimization, Bertsekas and Ozdaglar [13] considered a problem with a set constraint, but in [8–10, 16] the set constraint is not considered. It would be very interesting to extend the results of the present paper to a differentiable multiobjective problem with a set constraint, in such a way that in particular the results of [13] will be also generalized.

In the proof of Theorem 3.1, mainly properties of Clarke subdifferential are used. For this reason, we think that another promising research line is to extend our results to a multiobjective problem involving locally Lipschitz functions.

Another interesting investigation line is to study a necessary sequential AKKT condition for a point to be a solution of a vector variational inequality extending the results in [8, Sect. 3]. We believe that the sequential optimality conditions here exposed may be useful in order to generate or improve algorithms. Thus, in our opinion, further investigation is needed to provide suitable conditions that can be utilized in numerical structures for solving multiobjective problems and therefore for practical applications. See Haeser and Melo [12] where an AKKT-type condition is discussed as a stop criterion for an interactive algorithm.

Acknowledgments This research was partially supported (for the second and third author) by the Ministerio de Economía y Competitividad (Spain) under project MTM2012-30942. The authors are grateful to the associate editor and the anonymous referees for their helpful comments and suggestions.

References

1. Kortanek, K.O., Evans, J.P.: Asymptotic Lagrange regularity for pseudoconcave programming with weak constraint qualification. *Oper. Res.* **16**, 849–857 (1968)
2. Fiacco, A.V., McCormick, G.P.: The slacked unconstrained minimization technique for convex programming. *SIAM J. Appl. Math.* **15**, 505–515 (1967)
3. Zlobec, S.: Extensions of asymptotic Kuhn–Tucker conditions in mathematical programming. *SIAM J. Appl. Math.* **21**, 448–460 (1971)
4. Hestenes, M.R.: *Optimization Theory: The Finite Dimensional Case*. Wiley, New York (1975)
5. McShane, E.J.: The Lagrange multiplier rule. *Am. Math. Mon.* **80**, 922–925 (1973)

6. Craven, B.D.: Modified Kuhn–Tucker conditions when a minimum is not attained. *Oper. Res. Lett.* **3**, 47–52 (1984)
7. Trudzik, L.I.: Asymptotic Kuhn–Tucker conditions in abstract spaces. *Numer. Funct. Anal. Optim.* **4**, 355–369 (1982)
8. Haeser, G., Schuverdt, M.L.: On approximate KKT condition and its extension to continuous variational inequalities. *J. Optim. Theory Appl.* **149**, 528–539 (2011)
9. Andreani, R., Haeser, G., Martínez, J.M.: On sequential optimality conditions for smooth constrained optimization. *Optimization* **60**, 627–641 (2011)
10. Andreani, R., Martínez, J.M., Svaiter, B.F.: A new sequential optimality condition for constrained optimization and algorithmic consequences. *SIAM J. Optim.* **20**, 3533–3554 (2010)
11. Dutta, J., Deb, K., Tulshyan, R., Arora, R.: Approximate KKT points and a proximity measure for termination. *J. Glob. Optim.* **56**, 1463–1499 (2013)
12. Haeser, G., Melo, V.V.: Convergence detection for optimization algorithms: approximate-KKT stopping criterion when Lagrange multipliers are not available. *Oper. Res. Lett.* **43**, 484–488 (2015)
13. Bertsekas, D.P., Ozdaglar, A.E.: Pseudonormality and a Lagrange multiplier theory for constrained optimization. *J. Optim. Theory Appl.* **114**, 287–343 (2002)
14. Clarke, F.H.: *Optimization and Nonsmooth Analysis*. Wiley, New York (1983)
15. Bertsekas, D.P.: *Nonlinear Programming*. Athena Scientific, Belmont (1999)
16. Andreani, R., Martínez, J.M., Schuverdt, M.L.: On the relation between constant positive linear dependence condition and quasinormality constraint qualification. *J. Optim. Theory Appl.* **125**, 473–483 (2005)