Università di Pisa

Dipartimento di Informatica

Technical Report: TR-04-11

Componentwise versus Global Approaches to Nonsmooth Multiobjective Optimization

Giancarlo Bigi

May 4, 2004

ADDRESS: Largo B.Pontecorvo 1, 56127 Pisa, Italia. TEL: +39 050 2212700 FAX: +39 050 2212726

Componentwise versus Global Approaches to Nonsmooth Multiobjective Optimization

Giancarlo Bigi[a,1]

[a]*Dipartimento di Informatica, Università di Pisa*
Largo B.Pontecorvo, 1, 56127 Pisa, Italia.

Abstract.

The paper aims to analyse and compare two different approaches to optimality conditions for multiobjective optimization problems, which involve nonsmooth functions. Both necessary and sufficient first order conditions are presented for the case in which constraint is given just as a set. Finally, the optimality conditions based on these two approaches are compared.

Keywords: Dini–Hadamard componentwise and set–valued derivatives, generalized Jacobians, tangent cones, optimality conditions.

1 Introduction

Nonsmooth multiobjective optimization has received much attention in the past years; in particular, many first order optimality conditions have been achieved, relying on different approaches based on nonsmooth analysis tools.

Since these tools are quite simple and powerful for real–valued functions, a standard approach is to consider suitable weighted sums of the components of vector–valued functions and to exploit them to study optimality (see [2, 7, 8]). However, such an approach inherits some well–known drawbacks of scalarization techniques. Another way to rely on nonsmooth tools for real–valued functions is just some kind of componentwise approach, that is considering generalized

[1]Tel: +39 050 2212789; fax: +39 050 2212726. *e-mail address:* giancarlo.bigi@di.unipi.it.

derivatives (see [13, 14, 15, 16]) or some suitable subdifferentials (see [5, 11, 12]) of the components of the considered functions.

Relying on the concept of Kuratowski limit, a very different approach has been introduced in [17] and then applied to study multiobjective optimization in some recent papers (see [3, 13, 18]). It can be considered somehow a global one, since set–valued directional derivatives of vector–valued functions are introduced without relying on components.

The aim of this paper is twofold. First, we deepen the analysis of the latter approach, providing new optimality conditions; we also show that well–known tools for vector functions such as generalized Jacobians (see [6]) can be incorporated into this global approach, just relying on particular set–valued derivatives. Then, we discuss whether or not the global approach provides sharper results than the componentwise one.

For the sake of simplicity, we focus our attention only on Dini–Hadamard type derivatives and the Pareto minimization problem

$$\min_{\mathbb{R}^\ell_+} f(x) \qquad \text{subject to} \qquad x \in X, \tag{1}$$

where $f = (f_1, ..., f_\ell) : \mathbb{R}^n \longrightarrow \mathbb{R}^\ell$ (let $K := \{1, ..., \ell\}$ be the corresponding index set) and $X \subseteq \mathbb{R}^n$ is just any subset.

We recall that $\bar{x} \in X$ is a *local vector minimum point* of (1) if there exists a neighbourhood N of $\bar{x}$ such that no $x \in X \cap N$ satisfies $f(x) - f(\bar{x}) \in -\mathbb{R}^\ell_+ \setminus \{0\}$, that is $f_k(x) \leq f_k(\bar{x})$ for all $k \in K$ with $f_{\bar{k}}(x) < f_{\bar{k}}(\bar{x})$ for at least an index; it is just a *local vector weak minimum point* of (P) if no $x \in X \cap N$ satisfies $f(x) - f(\bar{x}) \in -\operatorname{int} \mathbb{R}^\ell_+$, that is $f_{\bar{k}}(x) < f_{\bar{k}}(\bar{x})$ for all indexes.

Optimality conditions for (1) based on generalized derivatives are tested along particular sets of directions, which are related to the structure of the constraint X. The following approximations provide the adequate directions (see [1]).

Definition 1.1. The *contingent cone* of X at $\bar{x} \in X$ is:

$$T(X, \bar{x}) := \{w \in \mathbb{R}^q \ : \ \exists \{w_n\} \to w, \ \exists \{t_n\} \downarrow 0 \text{ s.t. } \bar{y} + t_n w_n \in X\}.$$

The *cone of interior displacements* of X at $\bar{x} \in X$ is:

$$IT(X, \bar{x}) := \{w \in \mathbb{R}^q \ : \ \forall \{t_n\} \downarrow 0, \ \forall \{w_n\} \to w, \ \forall \ n \geq \bar{n} \ \ \bar{x} + t_n w_n \in X\}.$$

The *cone of feasible directions* for X at $\bar{x}$ is:

$$F(X, \bar{x}) := \{ \ w \in \mathbb{R}^n \ : \ \exists \ \tau > 0, \ \ \forall \ t \in]0, \tau[, \ \ \bar{x} + tw \in X \ \}.$$

2 Componentwise approach

The most natural and simple way to analyse differential properties of nonsmooth functions is to rely on the generalized directional derivatives of their components. Thus, we start recalling the following definitions.

Definition 2.1. The *upper Dini–Hadamard derivative* of $h : \mathbb{R}^n \to \mathbb{R}$ at $\bar{x}$ in the direction w is

$$d^+h(\bar{x}; w) := \limsup_{w' \to w, t \downarrow 0} \quad t^{-1}[h(\bar{x} + tw') - h(\bar{x})].$$

Definition 2.2. The *lower Dini–Hadamard derivative* of $h : \mathbb{R}^n \to \mathbb{R}$ at $\bar{x}$ in the direction w is

$$d^-h(\bar{x}; w) := \liminf_{w' \to w, t \downarrow 0} \quad t^{-1}[h(\bar{x} + tw') - h(\bar{x})].$$

Some componentwise optimality conditions have been shown in [13, 14, 16] for the unconstrained and inequality constrained cases. In this section we present similiar results for the set constrained one.

Theorem 2.1. *If $\bar{x}$ is a local vector weak minimum point of (1), then*

$$\max_{k \in K} \; d^+f_k(\bar{x}; w) \geq 0 \tag{2}$$

holds for any $w \in T(X, \bar{x})$.

Proof. Ab absurdo, suppose there exists some $w \in T(X, \bar{x})$ such that (2) does not hold. Consider any two sequences $t_n \downarrow 0$ and $w_n \to w$ such that $\bar{x} + t_n w_n \in X$; since $d^+f_k(\bar{x}; w) < 0$, by standard properties of the maximum limit we get that

$$t^{-1}[f_k(\bar{x} + t_n w_n) - f_k(\bar{x})] < 0 \tag{3}$$

holds for all $k \in K$ whenever n is large enough. This contradicts that $\bar{x}$ is a local vector minimum point of (1). □

Relying instead on Clarke derivatives, analogous optimality conditions have been obtained in [15].

It would be reasonable to try to replace the upper Dini–Hadamard derivatives with the lower ones for at least some components of f, since a stronger condition would be achieved; the following example shows that it is not possible to substitute even one of them with the corresponding lower derivative.

Example 2.1. Consider problem (1) with $n = \ell = 2$, the constraint given by the set $X = \{x \in \mathbb{R}^2_+ : x_1 \geq x_2^2 \geq x_1^4\}$ and $f(x_1, x_2) = (x_2 - x_1, x_1 - x_2)$ if $x \in X$ and $f(x_1, x_2) = (-x_1 - x_2, -x_1 - x_2)$ if $x \notin X$. Thus, $\bar{x} = (0, 0)$ is a vector minimum point and $d^+ f_1(\bar{x}; w) = w_2 - w_1 = -d^+ f_2(\bar{x}; w)$ so that (2) holds. Since $d^- f_i(\bar{x}; w) = -w_1$ whenever $w_2 = 0$ and $d^- f_i(\bar{x}; w) = -w_2$ whenever $w_1 = 0$, then we have

$$\max \{d^+ f_1(\bar{x}; w), d^- f_2(\bar{x}; w)\} = -w_1 < 0$$

for any nonzero $(w_1, 0) \in T(X, \bar{x}) = \mathbb{R}^2_+$ and

$$\max \{d^- f_1(\bar{x}; w), d^+ f_2(\bar{x}; w)\} = -w_2 < 0$$

for any nonzero $(0, w_2) \in T(X, \bar{x})$.

In the above example the directions which do not allow such substitutions are just those on the boundary of the contingent cone. This suggest that considering a smaller approximation of X as the set of "test" directions might allow to achieve this goal. A general scheme to build componentwise necessary optimality conditions based on different derivatives and tangent cones has been developed in [4].

In the differentiable case, that is $d^+ f_k(\bar{x}; w) = d^- f_k(\bar{x}; w) = \nabla f_k(\bar{x}) \cdot w$, if the necessary condition (2) holds as a strict inequality for all nonzero directions, then it turns out to be also sufficient. However, the standard procedure of replacing inequalities with strict inequalities to achieve sufficiency does not work in this nonsmooth framework, as the following example shows.

Example 2.2. Consider (1) with $n = \ell = 2$, the constraint $X = \mathbb{R}^2_+$ and $f(x_1, x_2) = (x_1, x_2 \cos(1/x_2))$ if $x_2 \neq 0$ and $f(x_1, 0) = (x_1, 0)$. Thus, $\bar{x} = (0, 0)$ is not a vector minimum point, since $f(0, 1/(2n+1)\pi) = (0, -1/(2n+1)\pi)$; on the contrary,

$$\max \{d^+ f_1(\bar{x}; w), d^+ f_2(\bar{x}; w)\} = \max \{w_1, w_2\} > 0$$

holds for any nonzero $w \in T(X, \bar{x}) = \mathbb{R}^2_+$.

It is easy to check that in this example the maximum of the lower Dini–Hadamard derivatives is not positive for all the test directions; in fact, the following sufficient criterion holds (notice that the corresponding conditions with inequalities instead of strict ones are not neccessary for optimality).

Theorem 2.2. *Let $\bar{x} \in X$. If*

$$\max_{k \in K} \; d^- f_k(\bar{x}; w) > 0 \tag{4}$$

holds for any nonzero $w \in T(X, \bar{x})$, then $\bar{x}$ is a local vector minimum point of (1).

Proof. Ab absurdo, suppose there exists a sequence $\{x_n\} \subseteq X$ such that $x_n \to \bar{x}$ and $f_k(x_n) - f_k(\bar{x}) < 0$ for all $k \in K$. Setting $t_n := ||x_n - \bar{x}||$ and $w_n := t_n^{-1}(x_n - \bar{x})$, we have $x_n = \bar{x} + t_n w_n$. Taking the suitable subsequence, we can suppose that $t_n \downarrow 0$ and $w_n \to w$ for some nonzero $w \in \mathbb{R}^n$. Thus, we have both $w \in T(X, \bar{x})$ and

$$d^- f_k(\bar{x}; w) \leq \liminf_{n \to +\infty} \;\; t_n^{-1}[f_k(\bar{x} + t_n w_n) - f_k(\bar{x})] \leq 0$$

for any index $k \in K$ in contradiction with the assumption. $\square$

3 Global approach

The componentwise approach has the weak feature of being suitable only to study Pareto optimization, since in general components are not strictly related to ordering cones other than the non–negative orthant $\mathbb{R}^\ell_+$.

Therefore, it is reasonable to develop also some kind of global approach, which deals with vector–valued functions as a whole, that is without relying on components. Though the vector *incremental ratio*

$$h_f(x, t, w) = t^{-1}[f(x + tw) - f(x)]$$

is unlike to converge to some unique point unless f is differentiable, we can adapt the Kuratowski convergence of sets to this framework; thus, we can define set–valued directional derivatives as the set of all the cluster points of suitable sequences based on the above incremental ratio.

Definition 3.1. The *Dini–Hadamard (global) set–valued derivative* of f at $\bar{x}$ in the direction w is

$$Df(\bar{x}; w) := \{z \in \mathbb{R}^\ell \; : \; \exists t_n \downarrow 0, \; \exists w_n \to w \text{ s.t. } h_f(\bar{x}, t_n, w) \to z\}.$$

Since in the real–valued case (i.e. $\ell = 1$) we just have

$$d^+ f(\bar{x}; w) = \sup Df(\bar{x}; w), \quad d^- f(\bar{x}; w) = \inf Df(\bar{x}; w),$$

it would be reasonable to consider just the maximal and minimal points of $Df(\bar{x}; w)$ as done in some papers [3, 13, 17, 18]. It would allow to get both lower and upper set–valued derivatives but some assumptions to guarantee the existence of the derivatives would be needed; furthermore, such definitions would be related to a precise ordering cone, while the above one avoids these drawbacks.

Some results require that f satisfies a *Lipschitz condition*, that is

$$||f(y) - f(x)|| \leq L_f ||y - x|| \tag{5}$$

holds for some suitable $L_f > 0$. We say that f is *Lipschitz near* $\bar{x}$ if (5) holds for any y and x in a neighbourhood of $\bar{x}$ while we say that f is *locally Lipschitz at* $\bar{x}$ if if (5) holds for $x = \bar{x}$ and any y in a neighbourhood of $\bar{x}$.

Proposition 3.1. *Suppose f is continuous at $\bar{x} \in \mathbb{R}^n$.*

(i) The set–valued mapping $w \rightrightarrows Df(\bar{x}; w)$ is closed, i.e.

$$w_n \to w, \ z_n \in Df(\bar{x}; w_n), \ z_n \to z \implies z \in Df(\bar{x}; w).$$

(ii) If f is locally Lipschitz at $\bar{x}$, then $Df(\bar{x}; w)$ is a nonempty compact set.

Proof. (i) For any $z_n \in Df(\bar{x}; w_n)$ there exist sequences $\tau_m \downarrow 0$ and $v_m \to w_n$ such that $\tau_m^{-1}[f(\bar{x} + \tau_m w_m) - f(\bar{x})]$ converges to z_n. Let's just choose an index $\bar{m}$ such that $||w_n - v_{\bar{m}}|| \leq 1/n$ and $||z_n - \tau_{\bar{m}}^{-1}[f(\bar{x} + \tau_{\bar{m}} w_{\bar{m}}) - f(\bar{x})]|| \leq 1/n$. Setting $t_n = \tau_{\bar{m}}$ and $\hat{w}_n = w_{\bar{m}}$ we get both $\hat{w}_n \to w$ and

$$||z - t_n^{-1}[f(\bar{x} + t_n \hat{w}_n) - f(\bar{x})]|| \leq ||z - z_n|| + ||z_n - \tau_{\bar{m}}^{-1}[f(\bar{x} + \tau_{\bar{m}} w_{\bar{m}}) - f(\bar{x})]|| \to 0.$$

Therefore, we have $z \in Df(\bar{x}; w_n)$.

(ii) Given any $\varepsilon > 0$ and any sequences $t_n \downarrow 0$ and $w_n \to w$ we have

$$||t_n^{-1}[f(\bar{x} + t_n w_n) - f(\bar{x})]|| \leq L_f ||w_n|| \leq L_f(||w|| + \varepsilon)$$

whenever n is large enough. Therefore, we have $Df(\bar{x}; w) \subseteq B(0, L_f ||w||)$. Since (i) implies that $Df(\bar{x}; w)$ is a closed set, then compactness follows. Moreover, the boundedness of the sequence $\{t_n^{-1}[f(\bar{x}+t_n w_n)-f(\bar{x})]\}$ implies that it admits a converging subsequence and therefore $Df(\bar{x};w)$ is nonempty. $\square$

It is worth noting that $Df(\bar{x};w)$ may be nonconvex, as Example 3.1 below shows. The next proposition investigates the relationships between the set–valued derivative of f and those of its components.

Proposition 3.2. *Suppose $f=(\varphi_1,\varphi_2)$ with $\varphi_i:\mathbb{R}^n\to\mathbb{R}^{\ell_i}$, $\ell_1+\ell_2=\ell$.*

(i) $Df(\bar{x};w)\subseteq D\varphi_1(\bar{x};w)\times D\varphi_2(\bar{x};w)$

(ii) If φ_2 is differentiable at $\bar{x}$, then $Df(\bar{x};w)=D\varphi_1(\bar{x};w)\times\{\nabla\varphi_2(\bar{x})\cdot w\}$.

Proof. (i) It is enough to notice that whenever a sequence $h_f(\bar{x},t_n,w_n)$ converges to some $z\in Df(\bar{x};w)$, then we have $h_{\varphi_i}(\bar{x},t_n w_{,n})\to z^i$ with $z=(z^1,z^2)$ and therefore $z^i\in D\varphi_i(\bar{x};w)$.

(ii) It follows from (i) just noting that every sequence $h_{\varphi_2}(\bar{x},t_n,w_n)$ converges to $\nabla\varphi_2(\bar{x})\cdot w$. $\square$

Obviously, the inclusion in (i) may be strict, as the following example shows; moreover, notice that otherwise the global approach would coincide with the componentwise one.

Example 3.1. Consider the function $f:\mathbb{R}\to\mathbb{R}^2$ given by

$$f(x)=(x\sin(1/x),x\cos(1/x)),\quad x\neq 0$$

with $f(0)=(0,0)$ and consider the point $\bar{x}=0$. Given any $\theta\in[0,2\pi]$ and any $w\in\mathbb{R}$, let $w_n=w$ and $t_n=1/w(\theta+2n\pi)$; then,

$$t_n^{-1}[f(\bar{x}+t_n w_n)-f(\bar{x})]=(w\sin\theta,w\cos\theta).$$

Therefore, we have both

$$Df(\bar{x};w)=\{z\in\mathbb{R}^2\ :\ z_1^2+z_2^2=w^2\}$$

and

$$Df_1(\bar{x};w)=Df_2(\bar{x};w)=[-w,w].$$

The fact that the considered derivatives are now set–valued leads to at least two different types of necessary optimality conditions: all the elements of the derivative do not have all negative components (empty intersection) or there is at least one element with such a property (not contained).

Theorem 3.1. *Suppose $\bar{x} \in X$ is a local vector weak minimum point of (1). Then*

$$Df(\bar{x}; w) \cap -\operatorname{int} \mathbb{R}^{\ell}_{+} = \emptyset \tag{6}$$

holds for any $w \in IT(X, \bar{x})$. Furthermore, if f is locally Lipschitz near $\bar{x}$, then (6) holds for any $w \in F(X, \bar{x})$ while if f is just locally Lipschitz at $\bar{x}$, then

$$Df(\bar{x}; w) \nsubseteq -\operatorname{int} \mathbb{R}^{\ell}_{+} \tag{7}$$

holds for any $w \in T(X, \bar{x})$.

Proof. Ab absurdo, suppose (6) does not hold for some $w \in IT(X, \bar{x})$. Therefore, let

$$z \in Df(\bar{x}; w) \cap -\operatorname{int} \mathbb{R}^{\ell}_{+}.$$

Since $z = \lim\ t_n^{-1}[f(\bar{x} + t_n w_n) - f(\bar{x})]$ for suitable sequences $t_n \downarrow 0$ and $w_n \to w$, we have both $\bar{x} + t_n w_n \in X$ and $f(\bar{x}) - f(\bar{x} + t_n w_n) \in \operatorname{int} \mathbb{R}^{\ell}_{+}$, contradicting the optimality of $\bar{x}$. When f is locally Lipschitz near $\bar{x}$, then

$$\lim\ t_n^{-1}[f(\bar{x} + t_n w_n) - f(\bar{x})] = \lim\ t_n^{-1}[f(\bar{x} + t_n w) - f(\bar{x})];$$

hence, the same argument shows that (6) holds also for any $w \in F(X, \bar{x})$.

Analogously, suppose (7) does not hold for some $w \in T(X, \bar{x})$. Let $t_n \downarrow 0$ and $w_n \to w$ be such that $\bar{x} + t_n w_n \in X$. Taking the suitable subsequences, we can suppose that $t_n^{-1}[f(\bar{x} + t_n w_n) - f(\bar{x})] \to z$ for some $z \in \mathbb{R}^{\ell}$. Then, we have $z \in Df(\bar{x}; w) \subseteq -\operatorname{int} \mathbb{R}^{\ell}_{+}$ and we achieve a contradiction just as in the previous case. □

Clearly, (6) implies (7); on the contrary, the latter is satisfied by a larger cone of directions. The following example shows that the former may not hold for some of these directions.

Example 3.2. Consider the same problem of Example 2.1. Given any nonzero $w = (w_1, 0) \in T(X, \bar{x}) \setminus IT(X, \bar{x})$, calculations show that

$$Df(\bar{x}; w) = \{d^+ f(\bar{x}; w)\} \cup \{d^- f(\bar{x}; w)\} = \{(-w_1, w_1), (-w_1, -w_1)\}.$$

Therefore, (6) does not hold for all $w \in T(X, \bar{x})$.

It is easy to check that condition (6) holds if and only if the disjunction is achieved just considering the minimal points of the derivative; in the same way, condition (7) holds if and only if it is the set of maximal points to be not contained into the cone $-\operatorname{int} \mathbb{R}^\ell_+$.

We now investigate how necessary conditions have to be strengthened to become also sufficient. The following result adjusts a well–known differential property of convex functions to our framework.

Proposition 3.3. *Suppose f is convex. Then, the inclusion*

$$f(y) - f(x) - Df(\bar{x}; y - x) \subseteq \mathbb{R}^\ell_+ \tag{8}$$

holds for any $x, y \in \mathbb{R}^n$.

Proof. Given any $x, y \in \mathbb{R}^n$ and any $z \in Df(x; y - x)$, consider the sequences $t_n \downarrow 0$ and $w_n \to (y - x)$ such that $t_n^{-1}[f(x + t_n w_n) - f(x_n)] \to z$. The convexity of f implies that

$$f(x + w_n) - f(x) - t_n^{-1}[f(x + t_n w_n) - f(x)] \in \mathbb{R}^\ell_+$$

whenever $t_n \in [0, 1]$. Since convexity implies continuity, taking the limit we get

$$f(y) - f(x) - z \in \mathbb{R}^\ell_+$$

and the thesis follows. □

Thus, we can achieve the following complete characterization of weak optimality for the convex case.

Theorem 3.2. *Suppose f and X are both convex. Then, $\bar{x} \in X$ is a vector weak minimum point of (1) if and only if (7) holds for any $w \in F(X, \bar{x})$.*

Proof. The "only if" part follows immediately from Theorem 3.1, just recalling that $F(X, \bar{x}) \subseteq T(X, \bar{x})$. Vice versa, suppose $\bar{x}$ is not a vector weak minimum point of (1). Therefore, there exists $x \in X$ such that $f(x) - f(\bar{x}) \in -\operatorname{int} \mathbb{R}^\ell_+$. Proposition 3.3 implies that

$$Df(\bar{x}; x - \bar{x}) \subseteq f(x) - f(\bar{x}) - \mathbb{R}^\ell_+ \subseteq -\operatorname{int} \mathbb{R}^\ell_+.$$

Since X is convex, then $(x - \bar{x}) \in F(X, \bar{x})$ contradicting (7). □

An equivalent characterization has been achieved in [18], considering just the maximal points of the set–valued derivative.

When problem (1) does not satisfy convexity assumptions, we can obtain a sufficient criterion, just strengthening the necessary condition (6) in the following way.

Theorem 3.3. *Suppose f is locally Lipschitz at $\bar{x} \in X$. If*

$$Df(\bar{x}; w) \cap -\mathbb{R}^\ell_+ = \emptyset \tag{9}$$

holds for any nonzero $w \in T(X, \bar{x})$, then $\bar{x}$ is a local vector minimum point of (1).

Proof. Ab absurdo, suppose there exists a sequence $x_n \to \bar{x}$ with $x_n \in X$ and $f(x_n) - f(\bar{x}) \in -\mathbb{R}^\ell_+ \setminus \{0\}$. Arguing as in Theorem 2.2, we can suppose $x_n = \bar{x} + t_n w_n$ with $t_n \downarrow 0$ and $w_n \to w$ for some nonzero $w \in \mathbb{R}^n$. Since f is locally Lipschitz at $\bar{x}$, then the sequence

$$\{t_n^{-1}[f(\bar{x} + t_n w_n) - f(\bar{x})]\} \subseteq -\mathbb{R}^\ell_+ \setminus \{0\}$$

admits a subsequence converging to some $z \in -\mathbb{R}^\ell_+$. Furthermore, we also have $z \in Df(\bar{x}; w)$ contradicting (9). □

Let us stress that we can not achieve sufficiency replacing (9) with the weaker condition $Df(\bar{x}; w) \nsubseteq -\mathbb{R}^\ell_+$, as shown by the following example.

Example 3.3. Consider (1) with $n = 1$, $\ell = 2$, $X = \mathbb{R}_+$ and

$$f(x) = (x \sin(\log |x|), x \cos(\log |x|)), \quad x \neq 0$$

with $f(0) = (0,0)$. Arguing in the same fashion as in Example 3.1, we can easily check that

$$Df(\bar{x};w) = \{z \in \mathbb{R}^2 \ : \ z_1^2 + z_2^2 = w^2\} \nsubseteq -\mathbb{R}^2_+$$

but $\bar{x} = 0$ is not a local vector minimum point of (1).

It is worth stressing that all the optimality criteria of this section can be generalized to the case in which $\mathbb{R}^\ell$ is ordered by any pointed convex cone C: it is enough to just replace $\mathbb{R}^\ell_+$ with C in all the conditions.

3.1 Comparison with the Generalized Jacobian

As a matter of fact, there is another kind of global approach, which is very different from the one considered in the previous section. In fact, the famous Rademacher's theorem states that a vector–valued function f, which is locally Lipschitz near $\bar{x}$, is actually differentiable almost everywhere on a suitable neighbourhood of $\bar{x}$. Therefore, the idea is to consider the convex hull of all the cluster points of sequences of Jacobians as the differentiability points approach $\bar{x}$. Obviously, throughout this subsection we suppose that f is locally Lipschitz near $\bar{x}$.

Definition 3.2. The *generalized Jacobian* of f at $\bar{x}$ is

$$\partial f(\bar{x}) := \operatorname{conv}\{A \in \mathbb{R}^{\ell\times n} \ : \ \exists\ x_n \to x \text{ s.t. } \exists\ \nabla f(x_n) \text{ and } \nabla f(x_n) \to A\}.$$

It is always a nonempty convex compact subset of $\mathbb{R}^{\ell\times n}$ and the set–valued mapping $x \rightrightarrows \partial f(x)$ is closed and upper semicontinuous. For a detailed account on all its properties, see [6].

It is worth stressing that the above definition does not suggest how to calculate the generalized Jacobian in practice, since differentiability points are not known explicitly very often. On the contrary, the global approach of the previous section seems more practical; however, they are strictly related, just considering the *Clarke (global) set–valued derivative*, that is

$$D_C f(\bar{x};w) := \{z \in \mathbb{R}^\ell \ : \ \exists t_n \downarrow 0,\ \exists x_n \to x \text{ s.t. } h_f(x_n,t_n,w) \to z\}.$$

In fact, the convex hull of this directional derivative is related to the generalized Jacobian in the following way.

Lemma 3.1. *The equality*

$$\partial f(\bar{x}) \cdot w = \operatorname{conv} D_C f(\bar{x}; w)$$

holds for any $w \in \mathbb{R}^n$.

Proof. Consider a sequence $x_n \to \bar{x}$ such that $\nabla f(x_n)$ exists and $\nabla f(x_n) \to A$ for some $A \in \mathbb{R}^{\ell \times n}$. Since f is differentiable at x_n, we can choose t_n such that

$$||t_n^{-1}[f(x_n + t_n w) - f(x_n) - t_n \nabla f(x_n) \cdot w]|| \leq 1/n$$

and $t_n \downarrow 0$. Thus, we have

$$||t_n^{-1}[f(x_n + t_n w) - f(x_n)] - A \cdot w|| \leq 1/n + ||(A - \nabla f(x_n)) \cdot w|| \to 0.$$

Therefore, $A \cdot w \in D_C f(\bar{x}; w)$ and hence $\partial f(\bar{x}) \cdot w \subseteq \operatorname{conv} D_C f(\bar{x}; w)$.

Vice versa, let $z \in D_C f(\bar{x}; w)$ so that $t_n^{-1}[f(x_n + t_n w) - f(x_n)] \to z$ for some $x_n \to \bar{x}$ and $t_n \downarrow 0$. The vector mean value theorem (see, for instance, Proposition 2.6.5 in [6]) implies that

$$z_n = t_n^{-1}[f(x_n + t_n w) - f(x_n)] \in \operatorname{conv} \{\partial f(x_n + \tau w) \cdot w \ : \ \tau \in [0, t_n]\}.$$

Therefore,

$$z_n = \sum_{i=1}^{\ell+1} \lambda_i^{(n)} d_i^{(n)} \tag{10}$$

for some $d_i^{(n)} \in \partial f(x_n + \tau_n w) \cdot w$ and $\lambda_i^{(n)} \geq 0$ such that $\sum_{i=1}^{\ell+1} \lambda_i^{(n)} = 1$. Since f is locally Lipschitz near $\bar{x}$, then the sequences $\{d_i^{(n)}\}$ are bounded. Hence, taking the suitable subsequences, we can suppose $\lambda_i^{(n)} \to \lambda_i$ and $d_i^{(n)} \to d_i$ for some $\lambda_i \geq 0$ and $d_i \in \mathbb{R}^\ell$. Since the set–valued mapping $x \rightrightarrows \partial f(\bar{x})$ is closed, we have $d_i \in \partial f(\bar{x}) \cdot w$. Taking the limit in (10), we get

$$z = \sum_{i=1}^{\ell+1} \lambda_i d_i \in \partial f(\bar{x}) \cdot w.$$

Then, the inclusion $\operatorname{conv} D_C f(\bar{x}; w) \subseteq \partial f(\bar{x}) \cdot w$ follows immediately from the convexity of the generalized Jacobian. $\square$

This equality clarifies also the relation between the generalized Jacobian and the Dini–Hadamard set–valued derivative.

Corollary 3.1. *$Df(\bar{x};w) \subseteq \partial f(\bar{x}) \cdot w$.*

Proof. Since f is locally Lipschitz near $\bar{x}$, then $Df(\bar{x};w) \subseteq D_C f(\bar{x};w)$. $\square$

The following example shows that the above inclusion may be strict.

Example 3.4. Consider the function

$$f(x) = (x^2 \cos(1/x), -x^2 \sin(1/x)), \quad x \neq 0$$

with $f(0) = (0,0)$. Then, f is locally Lispchitz near $\bar{x} = 0$ and furthermore it is differentiable at any $x \neq \bar{x}$ with

$$\nabla f(x) = (2x\cos(1/x) + \sin(1/x), -2x\sin(1/x) + \cos(1/x)).$$

Therefore, we have

$$\partial f(\bar{x}) = \text{conv } \{z \in \mathbb{R}^2 \ : \ z_1^2 + z_2^2 = 1\} = B(0,1)$$

which implies that $\partial f(\bar{x}) \cdot w = B(0,w)$ for any $w \in \mathbb{R}$. On the contrary, it is easy to show that $Df(\bar{x};w) = \{(0,0)\}$.

As a consequence of Corollary 3.1, both the necessary optimality conditions (6) and (7) imply $\partial f(\bar{x}) \cdot w \nsubseteq -\text{int}\, \mathbb{R}^\ell_+$; moreover, $\partial f(\bar{x}) \cdot w \cap -\mathbb{R}^\ell_+ = \emptyset$ implies the sufficient optimality condition (9) in the same way. Therefore, the below criteria based on the generalized Jacobian follow easily from Theorems 3.1 and 3.3.

Theorem 3.4. *If $\bar{x} \in X$ is a local vector weak minimum point of (1), then*

$$\partial f(\bar{x}) \cdot w \nsubseteq -\text{int}\, \mathbb{R}^\ell_+ \tag{11}$$

holds for any $w \in T(X,\bar{x})$. Vice versa, any $\bar{x} \in X$ such that

$$\partial f(\bar{x}) \cdot w \cap -\mathbb{R}^\ell_+ = \emptyset \tag{12}$$

holds for any nonzero $w \in T(X,\bar{x})$ is a local vector minimum point of (1).

In the unconstrained case such results have been presented very recently in [9] while in the constrained case they seem fairly new. Furthermore, in the unconstrained case condition (11) implies the existence of a nonzero $\theta \in \mathbb{R}^\ell_+$ such that

$$0 \in \sum_{k=1}^{\ell} \theta_k \partial f_k(\bar{x}),$$

which is a well–known componentwise necessary optimality condition for locally Lipschitz problems (see [10], [12]).

4 Componentwise versus Global Approaches

Since the two approaches are actually different, it is interesting to know how they are related. Therefore, let us consider the vectors whose components are just the Dini–Hadamard derivatives of the components of f, that is

$$d^+ f(\bar{x}; w) := (d^+ f_1(\bar{x}; w), ..., d^+ f_\ell(\bar{x}; w)),$$

$$d^- f(\bar{x}; w) := (d^- f_1(\bar{x}; w), ..., d^- f_\ell(\bar{x}; w)).$$

Proposition 4.1. *Given any $\bar{x} \in \mathbb{R}^n$ and $w \in \mathbb{R}^n$, the following relations*

$$d^+ f(\bar{x}; w) - z \in \mathbb{R}^\ell_+, \qquad z - d^- f(\bar{x}; w) \in \mathbb{R}^\ell_+$$

hold for any $z \in Df(\bar{x}; w)$.

Proof. Given any $z \in Df(\bar{x}; w)$ and any sequences $t_n \downarrow 0$ and $w_n \to w$ such that $t_n^{-1}[f(\bar{x} + t_n w_n) - f(\bar{x})] \to z$, standard properties of the maximum and mimimum limits imply that

$$d^- f_k(\bar{x}; w) \leq \lim_{n \to +\infty} t_n^{-1}[f_k(\bar{x} + t_n w_n) - f_k(\bar{x})] = z_k \leq d^+ f_k(\bar{x}; w)$$

hold for any $k \in K$. Therefore, the thesis follows. □

Therefore, $d^+ f(\bar{x}; w)$ and $d^- f(\bar{x}; w)$ can be considered as some kind of upper and lower bound for $Df(\bar{x}; w)$; however, it is worth stressing that they are generally not the ideal maximal and minimal point of this set, since it is very unlikely that they belong to the set–valued derivative.

Example 4.1. Consider the same function of the Example 3.1. Then, given any $w \geq 0$, we have

$$d^+ f(\bar{x}; w) = (w, w) \notin Df(\bar{x}; w), \quad d^- f(\bar{x}; w) = (-w, -w) \notin Df(\bar{x}; w).$$

Now, we can compare the "global" necessary criterions in Theorem 3.1 with the corrresponding "componentwise" one in Theorem 2.1.

Proposition 4.2. *Suppose f is locally Lipschitz at $\bar{x}$. Given any $w \in \mathbb{R}^n$, then (7) holds if and only if (2) holds.*

Proof. Suppose (2) does not hold, that is $d^+f(\bar{x}; w) \in -\operatorname{int}\mathbb{R}^\ell_+$. Therefore, Proposition 4.1 (i) implies that

$$Df(\bar{x}; w) \subseteq d^+f_k(\bar{x}; w) - \mathbb{R}^\ell_+ \subseteq -\operatorname{int}\mathbb{R}^\ell_+ - \mathbb{R}^\ell_+ = \operatorname{int}\mathbb{R}^\ell_+.$$

Vice versa, suppose (2) holds. Thus, $d^+f_k(\bar{x}; w) \geq 0$ for some $k \in K$ and furthermore

$$d^+f_k(\bar{x}; w) = \lim\ t_n^{-1}[f(\bar{x} + t_n w_n) - f(\bar{x})]$$

for suitable sequences $t_n \downarrow 0$ and $w_n \to w$. Since f is locally Lipschitz, the sequence $\{t_n^{-1}[f(\bar{x}+t_n w_n) - f(\bar{x})]\}$ is bounded. Eventually taking a subsequence, we can suppose that it converges to some $z \in \mathbb{R}^\ell$. Therefore, we have both $z_k = d^+f_k(\bar{x}; w)$ and $z \in Df(\bar{x}; w)$, which imply (7). $\square$

According to these results, the two approaches seem to be equivalent but it is not so. In fact, Proposition 4.1 guarantees that (6) implies (2) but the vice versa does not hold, as shown by the following example.

Example 4.2. Consider (1) with $n = 1$, $\ell = 2$, the same function f as in Example 3.1 and $X = \mathbb{R}$. Obviously, $\bar{x} = 0$ is not a local vector weak minimum point but the necessary conditions of Theorem 2.1 are satisfied, since

$$\max\ \{d^+f_1(\bar{x}; w), d^+f_2(\bar{x}; w)\} = |w| \geq 0$$

holds for any $w \in T(X, \bar{x}) = \mathbb{R}$; on the contrary, (6) does not hold since, for instance, $(-4|w|/5, -3|w|/5) \in Df(\bar{x}; w) \cap -\operatorname{int}\mathbb{R}^2_+$ for any $w \neq 0$.

This example shows that the global approach provides sharper conditions at least in the unconstrained case.

Actually, adding the constraint $X = \mathbb{R}_+$, the above example gives the same outcome. Since $IT(X, \bar{x}) \subset T(X, \bar{x})$, it is possible to build also examples such that (6) holds for any $w \in IT(X, \bar{x})$ while (2) does not hold for some direction $w \in T(X, \bar{x}) \setminus IT(X, \bar{x})$ at a point $\bar{x}$, which is not optimal. Therefore, the two optimality criterion are generally unrelated.

Anyway, if $T(X, \bar{x})$ is the closure of $IT(X, \bar{x})$, the upper semicontinuity of $d^+f_k(\bar{x}; \cdot)$ ensures that the "global" criterion implies the "componentwise" one; for instance, the above assumption holds if X is a convex epi-Lipschitzian set.

Therefore, we can state that the global approach provides sharper conditions also in the constrained case under mild assumptions.

The relationships between the two sufficient criteria are easier to investigate, since they are both based on "test" directions belonging to the contingent cone.

According to Proposition 4.1, condition (4) implies condition (9); on the contrary, the vice versa does not hold: the global approach provides a better sufficient condition than the componentwise one, as the following example shows.

Example 4.3. Consider (1) with $n = 1$, $\ell = 2$, $X = \mathbb{R}_+$ and

$$f(x) = (x \sin(\log |x|) + x, -x \sin(\log |x|)), \quad x \neq 0$$

with $f(0) = (0, 0)$. Then, $\bar{x} = 0$ is a vector minimum point but we have

$$\max \{d^- f_1(\bar{x}; w), d^- f_2(\bar{x}; w)\} = -w < 0$$

for any nonzero $w \in T(X, \bar{x}) = \mathbb{R}_+$; on the contrary, easy calculations show

$$Df(\bar{x}; w) = \{z \in \mathbb{R}^2 \ : \ z_2 = z_1 + w, \ z_1 \in [-w, w]\}$$

so that (9) holds.

References

[1] Bazaraa, M.S., and Shetty, C.M. (1976), *Foundations of Optimization*, Springer, Berlin.

[2] Bhatia, D. and Aggarwal, S. (1992), "Optimality and Duality for Multiobjective Nonsmooth Programming", *European Journal of Operational Research*, Vol. 57, pp. 360-367.

[3] Bigi, G. and Pappalardo, M. (2000), "Generalized Lagrange Multipliers: Regularity and Boundedness", in Di Pillo, G. and Giannessi, F. (eds.), *Nonlinear Optimization and Related Topics*, Kluwer, Dordrecht, pp. 1-14.

[4] Bigi, G. (2003), *Optimality and Lagrangian Regularity in Vector Optimization*, Ph.D. Thesis, S.E.U., Pisa.

[5] Censor, Y. (1977), "Pareto Optimality in Multiobjective Problems", *Applied Mathematics and Optimization*, Vol. 4, pp. 41-59.

[6] Clarke, F.H. (1983), *Optimization and Nonsmooth Analysis*, Wiley, New York.

[7] Coladas, L., Li, Z. and Wang, S. (1994), "Optimality Conditions for Multiobjective and Nonsmooth Minimization in Abstract Spaces", *Bulletin of the Australian Mathematical Society*, Vol. 50, pp. 205-218.

[8] Craven, B.D. (1989), "Nonsmooth Multiobjective Programming", *Numerical Functional Analysis and Optimization*, Vol. 10, pp. 49-64.

[9] Guerraggio, A. and Luc, D.T. (2001), "Optimality Conditions for $C^{1,1}$ Vector Optimization Problems", *Journal of Optimization Theory and Applications*, Vol. 109, pp. 615-629

[10] Kanniappan, P. (1983), "Necessary Conditions for Optimality of Nondifferentiable Convex Multiobjective Programming", *Journal of Optimization Theory and Applications*, Vol. 40, pp. 167-174.

[11] Minami, M. (1981), "Weak Pareto Optimality of Multiobjective Problems in a Locally Convex Linear Topological Space", *Journal of Optimization Theory and Applications*, Vol. 34, pp. 469-484.

[12] Minami, M. (1983), "Weak Pareto-Optimal Necessary Optimality Conditions in a Nondifferentiable Multiobjective Program on a Banach Space", *Journal of Optimization Theory and Applications*, Vol. 41, pp. 451-461.

[13] Pappalardo, M. and Stöcklin, W. (2001), "Necessary Optimality Conditions in Nondifferentiable Vector Optimization", *Optimization*, Vol. 50, pp. 233-251.

[14] Preda, V. (1996), "Optimality Conditions and Duality in Multiple Objective Programming Involving Semilocally Convex and Related Functions", *Optimization*, Vol. 36, pp. 219-230.

[15] Staib, T. (1991), "On Necessary and Sufficient Optimality Conditions for Multicriterial Optimization Problems" *Z.O.R. - Methods and Models of Operations Research*, Vol. 35, pp. 231-248.

[16] Taa, A. (1999), "Optimality Conditions for Vector Mathematical Programming via a Theorem of the Alternative", *Journal of Mathematical Analysis and Applications*, Vol. 233, pp. 233-245.

[17] Thibault, L. (1980), "Subdifferentials of Compactly Lipschitzian Vector-Valued Functions" *Annali di Matematica Pura ed Applicata*, Vol. 15, pp. 157-192.

[18] Yang, X.Q. (1998), "A Generalized Upper Dini-Directional Derivative in Vector Optimization", *Optimization*, Vol. 43, pp. 339-351.