Marko M. Mäkelä | Ville-Pekka Eronen | Napsu Karmitsa

On Nonsmooth Optimality Conditions with Generalized Convexities

Turku Centre for Computer Science

TUCS Technical Report
No 1056, August 2012

On Nonsmooth Optimality Conditions with Generalized Convexities

Marko M. Mäkelä
University of Turku, Department of Mathematics and Statistics
FI-20014 Turku, Finland
makela@utu.fi
Ville-Pekka Eronen
University of Turku, Department of Mathematics and Statistics
FI-20014 Turku, Finland
vpoero@utu.fi
Napsu Karmitsa
University of Turku, Department of Mathematics and Statistics
FI-20014 Turku, Finland
napsu@karmitsa.fi

TUCS Technical Report
No 1056, August 2012

Abstract

Optimality conditions are an essential part of mathematical optimization theory, affecting heavily, for example to the optimization method development. Different types of generalized convexities have proved to be the main tool when constructing optimality conditions, particularly sufficient conditions for optimality. The purpose of this paper is to present some sufficient and necessary optimality conditions for locally Lipschitz continuous multiobjective problems. In order to prove sufficient optimality conditions some generalized convexity properties for functions are introduced. For necessary optimality conditions we will need some constraint qualifications.

Keywords: Generalized convexities; Clarke derivatives; nonsmooth analysis; nondifferentiable programming; optimality conditions; constraint qualifications

TUCS Laboratory
TOpGroup

1 Introduction

Optimality conditions are an essential part of mathematical optimization theory, affecting heavily, for example to the optimization method development. When constructing optimality conditions convexity has been the most important concept during the last decades. Recently there have been numerous attempts to generalize the concept of convexity in order to weaken the assumptions of the attained results (see e.g. [1, 4, 8, 14, 16, 25, 28, 30]).

Different kinds of generalized convexities have proved to be the main tool when constructing optimality conditions, particularly sufficient conditions. There exist a wide amount of papers published for smooth (continuously differentiable) single-objective case (see [25] and references therein). For nonsmooth (not continuously differentiable) problems there is an additional degree of freedom in choosing the way how to deal with the nonsmoothness. There are many different generalized directional derivatives to do this. For example, necessary and sufficient conditions for nonsmooth single-objective optimization by using the Dini directional derivatives were developed in [8]. These results were extended for nonsmooth multiobjective problems in [3].

Another degree of freedom is how to generalize convexity. In [21] sufficient conditions for nonsmooth multiobjective programs were derived by using the $(\mathcal{F}, \rho)$-convexity defined by Preda [26] and its extension for nonsmooth case defined by Bhatia and Jain [4]. Recently, the concept of invexity defined by Hanson [9] has become a very popular research concept. It was used to formulate necessary and sufficient conditions for differentiable multiobjective case in [24], for arcwise connected functions in [5] and for nonsmooth multiobjective programming in [6, 13, 22, 23].

In this paper, we present optimality conditions for nonsmooth multiobjective problems with locally Lipschitz continuous functions. Three types of constraint sets are considered. First, we discuss general set constraint, then, only inequality constraints and, finally, both inequality and equality constraints. To deal with the nonsmoothness we use the Clarke subdifferential as a generalization to gradient. For the necessary condition we require that certain constraint qualifications holds. For sufficient conditions we use f°-pseudo- and quasiconvexities [14] as a generalization to convexity. The necessary conditions with inequality constraints relies mainly on [15]. In [13] a sufficient condition was presented which differs from ours mainly by the formulation of object function. Moreover, f°-quasiconcave inequality constraints were not considered in [13].

Nonsmooth problems with locally Lipschitz continuous functions were considered also in [11, 23, 29]. Our presentation differs from [23] and [11] by constraint qualifications and the formulation of KKT conditions. Also, in [23] the necessary optimality condition relied on a theorem, which required the subdifferential of equality constraint functions to be a singleton. For the sufficient conditions we need generalized pseudo- and quasiconvexities. Contrary to [23], the invexity and its generalizations are not used here. In [29] general constraint set was used in the derivation of conditions for weak Pareto optimality. Our presentation has different, more spesific formulation for these conditions.

This article is organized as follows. In Section 2 we recall some basic tools from nonsmooth analysis. In Section 3 results concerning generalized pseudo- and quasiconvexity are presented. In Section 4 we present Karush-Kuhn-Tucker type necessary and sufficient conditions of weak Pareto optimality for nonsmooth multiobjective optimization problems with different constraint sets. Finally, some concluding remarks are given in Section 5.

2 Nonsmooth Analysis

In this section we collect some notions and results from nonsmooth analysis. Most of the proofs of this section are omitted, since they can be found, for example in [7, 17]. Nevertheless, we start by recalling the notion of convexity and Lipschitz continuity. The function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda \in [0, 1]$ we have

$$f\big(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}).$$

A function is *locally Lipschitz continuous at a point* $\boldsymbol{x} \in \mathbb{R}^n$ if there exist scalars $K > 0$ and $\delta > 0$ such that

$$|f(\boldsymbol{y}) - f(\boldsymbol{z})| \leq K\|\boldsymbol{y} - \boldsymbol{z}\| \quad \text{for all } \boldsymbol{y}, \boldsymbol{z} \in B(\boldsymbol{x}; \delta),$$

where $B(\boldsymbol{x}; \delta) \subset \mathbb{R}^n$ is an open ball with center $\boldsymbol{x}$ and radius δ. If a function is locally Lipschitz continuous at every point then it is called *locally Lipschitz continuous*. Note that both convex and smooth functions are always locally Lipschitz continuous (see, e.g. [7]). In what follows the considered functions are assumed to be locally Lipschitz continuous.

DEFINITION 2.1. [7] Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous at $\boldsymbol{x} \in S \subset \mathbb{R}^n$. The *Clarke generalized directional derivative* of f at $\boldsymbol{x}$ in the direction of $\boldsymbol{d} \in \mathbb{R}^n$ is defined by

$$f^\circ(\boldsymbol{x}; \boldsymbol{d}) = \limsup_{\substack{\boldsymbol{y} \to \boldsymbol{x} \\ t \downarrow 0}} \frac{f(\boldsymbol{y} + t\boldsymbol{d}) - f(\boldsymbol{y})}{t}$$

and the *Clarke subdifferential* of f at $\boldsymbol{x}$ by

$$\partial f(\boldsymbol{x}) = \{\boldsymbol{\xi} \in \mathbb{R}^n \mid f^\circ(\boldsymbol{x}; \boldsymbol{d}) \geq \boldsymbol{\xi}^T \boldsymbol{d} \text{ for all } \boldsymbol{d} \in \mathbb{R}^n\}.$$

Each element $\boldsymbol{\xi} \in \partial f(\boldsymbol{x})$ is called a *subgradient* of f at $\boldsymbol{x}$.

Note that the Clarke generalized directional derivative $f^\circ(\boldsymbol{x}; \boldsymbol{d})$ always exists for a locally Lipschitz continuous function f. Furthermore, if f is smooth $\partial f(\boldsymbol{x})$ reduces to $\partial f(\boldsymbol{x}) = \{\nabla f(\boldsymbol{x})\}$ and if f is convex $\partial f(\boldsymbol{x})$ coincides with the classical subdifferential of convex function (cf. [27]), in other words the set of $\boldsymbol{\xi} \in \mathbb{R}^n$ satisfying

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \boldsymbol{\xi}^T(\boldsymbol{y} - \boldsymbol{x}) \quad \text{for all } \boldsymbol{y} \in \mathbb{R}^n.$$

The following properties derived in [7] are characteristic to the generalized directional derivative and subdifferential.

THEOREM 2.2. *If* $f : \mathbb{R}^n \to \mathbb{R}$ *is locally Lipschitz continuous at* $\boldsymbol{x} \in \mathbb{R}^n$*, then*

(i) $\boldsymbol{d} \mapsto f^\circ(\boldsymbol{x};\boldsymbol{d})$ *is positively homogeneous, subadditive and Lipschitz continuous function such that* $f^\circ(\boldsymbol{x};-\boldsymbol{d}) = (-f)^\circ(\boldsymbol{x};\boldsymbol{d})$.

(ii) $\partial f(\boldsymbol{x})$ *is a nonempty, convex and compact set.*

(iii) $f^\circ(\boldsymbol{x};\boldsymbol{d}) = \max\{\boldsymbol{\xi}^T\boldsymbol{d} \mid \boldsymbol{\xi} \in \partial f(\boldsymbol{x})\}$ *for all* $\boldsymbol{d} \in \mathbb{R}^n$.

(iv) $f^\circ(\boldsymbol{x};\boldsymbol{d})$ *is upper semicontinuous as a function of* $(\boldsymbol{x},\boldsymbol{d})$.

From the last part of (i) in Theorem 2.2 we can easily deduce the following lemma.

LEMMA 2.3. *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be locally Lipschitz continuous and* $\boldsymbol{x} \in \mathbb{R}^n$*. Then*

$$\partial(-f)(\boldsymbol{x}) = -\partial f(\boldsymbol{x}).$$

PROOF. By Theorem 2.2 (i) we have

$$\begin{aligned} \partial(-f)(\boldsymbol{x}) &= \{\boldsymbol{\xi} \mid (-f)^\circ(\boldsymbol{x};\boldsymbol{d}) \geq \boldsymbol{\xi}^T\boldsymbol{d}, \text{ for all } \boldsymbol{d} \in \mathbb{R}^n\} \\ &= \{\boldsymbol{\xi} \mid f^\circ(\boldsymbol{x};-\boldsymbol{d}) \geq (-\boldsymbol{\xi})^T(-\boldsymbol{d}), \text{ for all } \boldsymbol{d} \in \mathbb{R}^n\} \\ &= \{-\boldsymbol{\xi} \mid f^\circ(\boldsymbol{x};-\boldsymbol{d}) \geq \boldsymbol{\xi}^T(-\boldsymbol{d}), \text{ for all } \boldsymbol{d} \in \mathbb{R}^n\}. \end{aligned}$$

Using the fact that $\boldsymbol{d} \in \mathbb{R}^n$ iff $-\boldsymbol{d} \in \mathbb{R}^n$ we obtain

$$\begin{aligned} &\{-\boldsymbol{\xi} \mid f^\circ(\boldsymbol{x};-\boldsymbol{d}) \geq \boldsymbol{\xi}^T(-\boldsymbol{d}), \text{ for all } \boldsymbol{d} \in \mathbb{R}^n\} \\ = \ &-\{\boldsymbol{\xi} \mid f^\circ(\boldsymbol{x};\boldsymbol{d}) \geq \boldsymbol{\xi}^T\boldsymbol{d}, \text{ for all } \boldsymbol{d} \in \mathbb{R}^n\} = -\partial f(\boldsymbol{x}). \end{aligned}$$

Hence, $\partial(-f)(\boldsymbol{x}) = -\partial f(\boldsymbol{x})$. □

In order to maintain equalities instead of inclusions in subderivation rules we need the following regularity property.

DEFINITION 2.4. The function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *subdifferentially regular* at $\boldsymbol{x} \in \mathbb{R}^n$ if it is locally Lipschitz continuous at $\boldsymbol{x}$ and for all $\boldsymbol{d} \in \mathbb{R}^n$ the classical directional derivative

$$f'(\boldsymbol{x};\boldsymbol{d}) = \lim_{t \downarrow 0} \frac{f(\boldsymbol{x}+t\boldsymbol{d}) - f(\boldsymbol{x})}{t}$$

exists and $f'(\boldsymbol{x};\boldsymbol{d}) = f^\circ(\boldsymbol{x};\boldsymbol{d})$.

Note, that the equality $f'(\boldsymbol{x};\boldsymbol{d}) = f^\circ(\boldsymbol{x};\boldsymbol{d})$ is not necessarily valid in general even if $f'(\boldsymbol{x};\boldsymbol{d})$ exists. This is the case, for instance, with concave nonsmooth functions. However, convexity, as well as smoothness implies subdifferential regularity [7]. Furthermore, it is easy to show that a necessary and sufficient condition for convexity is that for all $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$ we have

$$\begin{aligned} f(\boldsymbol{y}) - f(\boldsymbol{x}) &\geq f^\circ(\boldsymbol{x};\boldsymbol{y}-\boldsymbol{x}) \\ &= f'(\boldsymbol{x};\boldsymbol{y}-\boldsymbol{x}). \end{aligned} \tag{1}$$

Next we present two subderivation rules of composite functions, namely the finite maximum and positive linear combination of subdifferentially regular functions.

THEOREM 2.5. *Let* $f_i : \mathbb{R}^n \to \mathbb{R}$ *be locally Lipschitz continuous at* $\boldsymbol{x}$ *for all* $i = 1, \ldots, m$. *Then the function*

$$f(\boldsymbol{x}) = \max \{f_i(\boldsymbol{x}) \mid i = 1, \ldots, m\}$$

is locally Lipschitz continuous at $\boldsymbol{x}$ *and*

$$\partial f(\boldsymbol{x}) \subset \operatorname{conv} \{\partial f_i(\boldsymbol{x}) \mid f_i(\boldsymbol{x}) = f(\boldsymbol{x}),\ i = 1, \ldots, m\}, \tag{2}$$

where conv *denotes the convex hull of a set. In addition, if* f_i *is subdifferentially regular at* $\boldsymbol{x}$ *for all* $i = 1, \ldots, m$, *then* f *is also subdifferentially regular at* $\boldsymbol{x}$ *and equality holds in (2).*

THEOREM 2.6. *Let* $f_i : \mathbb{R}^n \to \mathbb{R}$ *be locally Lipschitz continuous at* $\boldsymbol{x}$ *and* $\lambda_i \in \mathbb{R}$ *for all* $i = 1, \ldots, m$. *Then the function*

$$f(\boldsymbol{x}) = \sum_{i=1}^{m} \lambda_i f_i(\boldsymbol{x})$$

is locally Lipschitz continuous at $\boldsymbol{x}$ *and*

$$\partial f(\boldsymbol{x}) \subset \sum_{i=1}^{m} \lambda_i \partial f_i(\boldsymbol{x}). \tag{3}$$

In addition, if f_i *is subdifferentially regular at* $\boldsymbol{x}$ *and* $\lambda_i \geq 0$ *for all* $i = 1, \ldots, m$, *then* f *is also subdifferentially regular at* $\boldsymbol{x}$ *and equality holds in (3).*

In the following, for a given set $S \subset \mathbb{R}^n$ we denote by d_S the *distance function* of S, that is,

$$d_S(\boldsymbol{x}) = \inf \{\|\boldsymbol{x} - \boldsymbol{s}\| \mid \boldsymbol{s} \in S\}. \tag{4}$$

If S is nonempty, then d_S is locally Lipschitz continuous with the constant one [7]. The closure of a set S is denoted $\operatorname{cl} S$. By the Weierstrass Theorem we may replace $\inf$ by $\min$ in (4) if $S \neq \emptyset$ is closed. Note also that $d_S(\boldsymbol{x}) = 0$ if $\boldsymbol{x} \in \operatorname{cl} S$.

A set $S \subset \mathbb{R}^n$ is a *cone* if $\lambda s \in S$ for all $\lambda \geq 0$ and $s \in S$. We also denote

$$\operatorname{ray} A = \{\lambda \boldsymbol{a} \mid \lambda \geq 0,\ \boldsymbol{a} \in A\} \quad \text{and} \quad \operatorname{cone} A = \operatorname{ray} \operatorname{conv} A.$$

In other words $\operatorname{ray} A$ is the smallest cone containing A and $\operatorname{cone} A$ is the smallest convex cone containing A.

DEFINITION 2.7. The *Clarke normal cone* of the set $S \subset \mathbb{R}^n$ at $\boldsymbol{x} \in S$ is given by the formula

$$N_S(\boldsymbol{x}) = \operatorname{cl} \operatorname{ray} \partial d_S(\boldsymbol{x}).$$

It is easy to derive that $N_S(\boldsymbol{x})$ is a closed convex cone (see, for example [7]). In convex case the normal cone can be expressed by the following simple inequality condition.

THEOREM 2.8. *If S is a convex set, then*

$$N_S(\boldsymbol{x}) = \{\boldsymbol{z} \in \mathbb{R}^n \mid \boldsymbol{z}^T(\boldsymbol{y} - \boldsymbol{x}) \le 0 \quad \textit{for all} \quad \boldsymbol{y} \in S\}.$$

The *contingent cone, polar cone* and *strict polar cone* of set $A \in \mathbb{R}^n$ at point $\boldsymbol{x}$ are defined respectively as

$$\begin{aligned} T_A(\boldsymbol{x}) &= \{\boldsymbol{d} \in \mathbb{R}^n \mid \text{there exist } t_i \downarrow 0 \text{ and } \boldsymbol{d}_i \to \boldsymbol{d} \text{ with } \boldsymbol{x} + t_i\boldsymbol{d}_i \in A\} \\ A^- &= \{\boldsymbol{d} \mid \boldsymbol{a}^T\boldsymbol{d} \le 0, \ \text{for all } \boldsymbol{a} \in A\} \\ A^s &= \{\boldsymbol{d} \mid \boldsymbol{a}^T\boldsymbol{d} < 0, \ \text{for all } \boldsymbol{a} \in A\}. \end{aligned}$$

Next we will present some basic results that are useful in section 4.

LEMMA 2.9. *Let $S_i \subset \mathbb{R}^n$, $i = 1, 2, \dots, I$ be convex sets and $C \subset \mathbb{R}^n$ be a convex cone. Assume that all the sets are nonempty. Then*

(i) $\operatorname{conv} \bigcup_{i=1}^I S_i = \{\sum_{i=1}^I \lambda_i \boldsymbol{s}_i \mid \boldsymbol{s}_i \in S_i, \ \lambda_i \ge 0, \sum_{i=1}^I \lambda_i = 1\}$

(ii) $\operatorname{cone} \bigcup_{i=1}^I S_i = \{\sum_{i=1}^I \mu_i \boldsymbol{s}_i \mid \boldsymbol{s}_i \in S_i, \ \mu_i \ge 0\} = \sum_{i=1}^I \operatorname{ray} S_i$

(iii) $\bigcup_{i=1}^I (S_i + C) = \bigcup_{i=1}^I S_i + C$

(iv) $\operatorname{conv} \bigcup_{i=1}^I (S_i + C) = \operatorname{conv} \bigcup_{i=1}^I S_i + C.$

PROOF. (i): Since $S_i \subset \bigcup_{i=1}^I S_i$ for all $i = 1, 2, \dots, I$, we have

$$\left\{ \sum_{i=1}^I \lambda_i \boldsymbol{s}_i \mid \boldsymbol{s}_i \in S_i, \ \lambda_i \ge 0, \ \text{for all } i = 1, 2, \dots, I, \ \sum_{i=1}^I \lambda_i = 1 \right\} \subset \operatorname{conv} \bigcup_{i=1}^I S_i.$$

Let $\boldsymbol{s} \in \operatorname{conv} \bigcup_{i=1}^I S_i$ be arbitrary. Then

$$\boldsymbol{s} = \sum_{j=1}^J \alpha_j \boldsymbol{s}_j, \ \alpha_j > 0, \sum_{j=1}^J \alpha_j = 1, \boldsymbol{s}_j \in \bigcup_{i=1}^I S_i, \ \text{for all } j = 1, 2, \dots, J.$$

Denote J_i the set of indices for which $\boldsymbol{s}_j \in S_i$, that is, $J_i = \{j \mid \boldsymbol{s}_j \in S_i\}$ and $\hat{I} \subset \{1, 2, \dots, I\}$ the set for which $J_i \neq \emptyset$. Denote also $\alpha_i = \sum_{j \in J_i} \alpha_j$. Then

$$\boldsymbol{s} = \sum_{i \in \hat{I}} \sum_{j \in J_i} \alpha_j \boldsymbol{s}_j = \sum_{i \in \hat{I}} \alpha_i \sum_{j \in J_i} \frac{\alpha_j}{\alpha_i} \boldsymbol{s}_j.$$

Since $\frac{\alpha_j}{\alpha_i} > 0$ and $\sum_{j \in J_i} \frac{\alpha_j}{\alpha_i} = 1$ for all $i \in \hat{I}$ we have $\sum_{j \in J_i} \frac{\alpha_j}{\alpha_i} \boldsymbol{s}_j = \hat{\boldsymbol{s}}_i \in S_i$. Noting that $\sum_{i \in \hat{I}} \alpha_i = \sum_{j=1}^J \alpha_j = 1$ we obtain

$$\boldsymbol{s} = \sum_{i \in \hat{I}} \alpha_i \hat{\boldsymbol{s}}_i \in \left\{ \sum_{i=1}^I \lambda_i \boldsymbol{s}_i \mid \boldsymbol{s}_i \in S_i, \ \lambda_i \ge 0, \sum_{i=1}^I \lambda_i = 1 \right\}.$$

(ii): Follows from (i) by taking ray from both sides.
(iii): The relation is clear from the following deduction

$$\begin{aligned}\bigcup_{i=1}^{I}(S_i + C) &= \{\boldsymbol{s} + \boldsymbol{c} \mid \boldsymbol{s} \in S_i \text{ for some } i = 1, 2 \ldots, I,\ \boldsymbol{c} \in C\} \\ &= \{\boldsymbol{s} + \boldsymbol{c} \mid \boldsymbol{s} \in \bigcup_{i=1}^{I} S_i, \boldsymbol{c} \in C\} = \bigcup_{i=1}^{I} S_i + C.\end{aligned}$$

(iv): By relation $\bigcup_{i=1}^{I} S_i \subset \operatorname{conv} \bigcup_{i=1}^{I} S_i$ and relation (iii) we have

$$\operatorname{conv} \bigcup_{i=1}^{I}(S_i + C) = \operatorname{conv}(\bigcup_{i=1}^{I} S_i + C) \subset \operatorname{conv}(\operatorname{conv} \bigcup_{i=1}^{I} S_i + C).$$

Furthermore, since $\operatorname{conv} \bigcup_{i=1}^{I} S_i + C$ is convex we have

$$\operatorname{conv}(\operatorname{conv} \bigcup_{i=1}^{I} S_i + C) = \operatorname{conv} \bigcup_{i=1}^{I} S_i + C.$$

For the other part suppose $\boldsymbol{s} \in \operatorname{conv} \bigcup_{i=1}^{I} S_i + C$. Then by (i) we have $\lambda_i \geq 0$ for all $i = 1, 2, \ldots, I$ such that $\sum_{i=1}^{I} \lambda_i = 1$ and

$$\boldsymbol{s} = \sum_{i=1}^{I} \lambda_i \boldsymbol{s}_i + \boldsymbol{c} = \sum_{i=1}^{I} \lambda_i (\boldsymbol{s}_i + \boldsymbol{c}) \in \operatorname{conv} \bigcup_{i=1}^{I}(S_i + C),$$

where in last relation part (i) can be applied since C, S_i, and thus, $S_i + C$ are convex for all $i = 1, 2, \ldots, I$. □

LEMMA 2.10. *Let $A, B \subset \mathbb{R}^n$ be convex compact sets. Then*

$$S = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} = \lambda \boldsymbol{a} + (1 - \lambda)\boldsymbol{b}, \boldsymbol{a} \in A, \boldsymbol{b} \in B, 0 \leq \lambda \leq 1\} = \operatorname{conv}(A \cup B)$$

and S is compact.

PROOF. Let $A, B \subset \mathbb{R}^n$ be convex compact sets. Relation $S = \operatorname{conv}(A \cup B)$ follows from Lemma 2.9 (i). Let $(\boldsymbol{x}_i) \subset \operatorname{conv}(A \cup B)$ be an arbitrary converging subsequence with $\lim_{i \to \infty} \boldsymbol{x}_i = \hat{\boldsymbol{x}}$. Then

$$\boldsymbol{x}_i = \lambda_i \boldsymbol{a}_i + (1 - \lambda_i)\boldsymbol{b}_i, \quad \boldsymbol{a}_i \in A,\ \boldsymbol{b}_i \in B,\ \lambda_i \in [0, 1] \text{ for all } i \in \mathbb{N}.$$

Consider the sequence $(\boldsymbol{z}_i) = (\boldsymbol{a}_i, \boldsymbol{b}_i, \lambda_i)$. Suppose that there is finitely many different points in sequence $(\boldsymbol{z}_i)$. Then the sequence is converging. Suppose then that there exist infinitely many different points. Since $A \times B \times [0, 1]$ is compact, the Bolzano-Weierstrass Theorem implies that the sequence has an accumulation point $\hat{\boldsymbol{z}}$. By the definition of accumulation point there exists convergent subsequence $(\boldsymbol{z}_{i_j})$ such that $i_j < i_{\hat{j}}$ for all $j < \hat{j}$. Since $(\boldsymbol{x}_i)$ is convergent we have

$$\lim_{i\to\infty} \boldsymbol{x}_i = \lim_{j\to\infty} \boldsymbol{x}_{i_j}.$$

Hence, without loss of generality we may assume that sequence $(\boldsymbol{z}_i)$ converges.

Since sets A, B and $[0,1]$ are closed, we have

$$\lim_{i\to\infty} \boldsymbol{a}_i = \hat{\boldsymbol{a}} \in A, \quad \lim_{i\to\infty} \boldsymbol{b}_i = \hat{\boldsymbol{b}} \in B, \quad \lim_{i\to\infty} \lambda_i = \hat{\lambda} \in A.$$

Thus,

$$\begin{aligned}\hat{\boldsymbol{x}} &= \lim_{i\to\infty}(\lambda_i \boldsymbol{a}_i + (1-\lambda_i)\boldsymbol{b}_i) \\ &= \lim_{i\to\infty}\lambda_i \lim_{i\to\infty}\boldsymbol{a}_i + (1-\lim_{i\to\infty}\lambda_i)\lim_{i\to\infty}\boldsymbol{b}_i \\ &= \hat{\lambda}\hat{\boldsymbol{a}} + (1-\hat{\lambda})\hat{\boldsymbol{b}} \in \operatorname{conv}(A\cup B)\end{aligned}$$

implying $\operatorname{conv}(A\cup B)$ is closed.

Since A and B are bounded there exists $r_A > 0$ and $r_B > 0$ such that $A \subset B(\mathbf{0}; r_A)$ and $B \subset B(\mathbf{0}; r_B)$. Denote $r = \max\{r_A, r_B\}$. Then $A\cup B \subset B(\mathbf{0}; r)$. Since $B(\mathbf{0}; r)$ is convex also $\operatorname{conv}(A\cup B) \subset B(\mathbf{0}; r)$ implying $\operatorname{conv}(A\cup B)$ is bounded. Hence $\operatorname{conv}(A\cup B)$ is compact. □

COROLLARY 2.11. *Let $A_1, A_2, \ldots, A_k \subset \mathbb{R}^n$ be convex compact sets. Then the set $\operatorname{conv}(\bigcup_{i=1}^k A_i)$ is a compact set.*

PROOF. The result follows from Lemma 2.10 by applying mathematical induction. □

To the end of this section we recall the classical necessary and sufficient nonsmooth unconstrained optimality condition.

THEOREM 2.12. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous at $\boldsymbol{x}^*$. If f attains its local minimum at $\boldsymbol{x}^*$, then*

$$\mathbf{0} \in \partial f(\boldsymbol{x}^*).$$

If, in addition, f is convex, then the above condition is sufficient for $\boldsymbol{x}^*$ to be a global minimum.

3 Generalized Convexities

In this section we present some generalizations of convexity, namely f°-pseudoconvexity, quasiconvexity and f°-quasiconvexity, that are used later. We also define f°-quasiconcavity. A famous generalization of convexity is pseudoconvexity introduced in [18]. For a pseudoconvex function f a point $\boldsymbol{x} \in \mathbb{R}$ is a global minimum if and only if $\nabla f(\boldsymbol{x}) = \mathbf{0}$. The classical pseudoconvexity requires the function to be smooth and, thus, it is not suitable for our purposes. However, with some modifications pseudoconvexity can be defined for nonsmooth functions as well. One such definition is presented in [10]. This definition requires the function to be merely locally Lipschitz continuous.

DEFINITION 3.1. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *f°-pseudoconvex*, if it is locally Lipschitz continuous and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$

$$f(\boldsymbol{y}) < f(\boldsymbol{x}) \qquad \text{implies} \qquad f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) < 0.$$

Note that due to (1) a convex function is always f°-pseudoconvex. Sometimes the reasoning chain in the definition of f°-pseudoconvexity needs to be converted.

LEMMA 3.2. *A locally Lipschitz continuous function f is f°-pseudoconvex, if and only if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$*

$$f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) \geq 0 \qquad \textit{implies} \qquad f(\boldsymbol{y}) \geq f(\boldsymbol{x}).$$

PROOF. Follows directly from the definition of f°-pseudoconvexity. □

The important sufficient extremum property of pseudoconvexity remains also for f°-pseudoconvexity.

THEOREM 3.3. *An f°-pseudoconvex function f attains its global minimum at $\boldsymbol{x}^*$, if and only if*

$$\mathbf{0} \in \partial f(\boldsymbol{x}^*).$$

PROOF. If f attains its global minimum at $\boldsymbol{x}^*$, then by Theorem 2.12 we have $\mathbf{0} \in \partial f(\boldsymbol{x}^*)$. On the other hand, if $\mathbf{0} \in \partial f(\boldsymbol{x}^*)$ and $\boldsymbol{y} \in \mathbb{R}^n$, then by Definition 2.1 we have

$$f^\circ(\boldsymbol{x}^*; \boldsymbol{y} - \boldsymbol{x}^*) \geq \mathbf{0}^T(\boldsymbol{y} - \boldsymbol{x}^*) = 0$$

and, thus by Lemma 3.2 we have

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}^*).$$

□

Note that it follows from Theorem 3.3 that pseudoconvexity implies f°-pseudoconvexity.

The notion of quasiconvexity is the most widely used generalization of convexity and, thus, there exist various equivalent definitions and characterizations. Next we recall the most commonly used definition of quasiconvexity (see [1]).

DEFINITION 3.4. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *quasiconvex*, if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda \in [0, 1]$

$$f(\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y}) \leq \max\{f(\boldsymbol{x}), f(\boldsymbol{y})\}.$$

Note that, unlike pseudoconvexity, the previous definition of quasiconvexity does not require differentiability nor continuity. We give also a useful result concerning a finite maximum of quasiconvex functions.

THEOREM 3.5. *Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be quasiconvex at $\boldsymbol{x}$ for all $i = 1, \ldots, m$. Then the function*

$$f(\boldsymbol{x}) = \max\{f_i(\boldsymbol{x}) \mid i = 1, \ldots, m\}$$

is also quasiconvex.

PROOF. Follows directly from the definition of quasiconvexity. □

Analogously to the Definition 3.1 we can define the corresponding generalized concept, which is a special case of h-quasiconvexity defined by Komlósi [14] when h is the Clarke generalized directional derivative.

DEFINITION 3.6. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *f°-quasiconvex*, if it is locally Lipschitz continuous and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$

$$f(\boldsymbol{y}) \leq f(\boldsymbol{x}) \qquad \text{implies} \qquad f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) \leq 0.$$

With f°-quasiconvexity we can define f°-quasiconcavity

DEFINITION 3.7. A function $f : \mathbb{R}^n \to \mathbb{R}$ is f°-quasiconcave if $-f$ is f°-quasiconvex.

THEOREM 3.8. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is f°-quasiconcave if it is locally Lipschitz continuous and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$*

$$f(\boldsymbol{y}) \leq f(\boldsymbol{x}) \qquad \textit{implies} \qquad f^\circ(\boldsymbol{y}; \boldsymbol{y} - \boldsymbol{x}) \leq 0.$$

PROOF. By Definitions 3.6 and 3.7 we have

$$-f(\boldsymbol{x}) \leq -f(\boldsymbol{y}) \qquad \text{implies} \qquad (-f)^\circ(\boldsymbol{y}; \boldsymbol{x} - \boldsymbol{y}) \leq 0.$$

Using Theorem 2.2 (i) we obtain

$$f(\boldsymbol{y}) \leq f(\boldsymbol{x}) \qquad \text{implies} \qquad f^\circ(\boldsymbol{y}; \boldsymbol{y} - \boldsymbol{x}) \leq 0$$

which proves the theorem. □

Next, we give few results concerning relations between the previously presented generalized convexities. The proofs for these results can be found in [16].

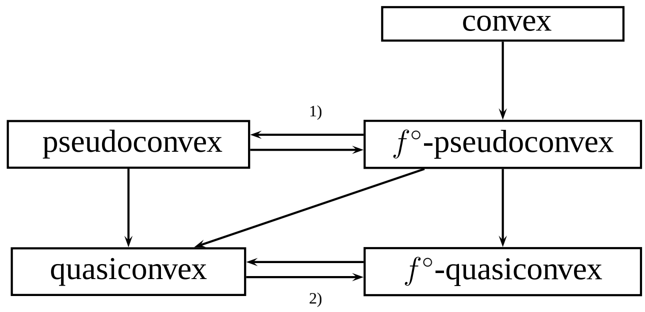
THEOREM 3.9. *If $f : \mathbb{R}^n \to \mathbb{R}$ is f°-pseudoconvex, then f is f°-quasiconvex and quasiconvex.*

THEOREM 3.10. *If $f : \mathbb{R}^n \to \mathbb{R}$ is f°-quasiconvex, then f is quasiconvex.*

THEOREM 3.11. *If $f : \mathbb{R}^n \to \mathbb{R}$ is subdifferentially regular and quasiconvex then f is f°-quasiconvex.*

The following figure illustrates the relations between different convexities.

Figure 1: Relations between different convexity types



1) demands continuous differentiability,

2) demands subdifferential regularity.

4 Optimality Conditions for Nonsmooth Multiobjective Problem

In this section we present some necessary and sufficient optimality conditions for multiobjective optimization.

Consider first a general multiobjective optimization problem

$$\begin{cases} \text{minimize} & \{f_1(\boldsymbol{x}), \dots, f_q(\boldsymbol{x})\} \\ \text{subject to} & \boldsymbol{x} \in S, \end{cases} \tag{5}$$

where $f_k : \mathbb{R}^n \to \mathbb{R}$ for $k = 1, 2, \dots, q$ are locally Lipschitz continuous functions and $S \subset \mathbb{R}^n$ is an arbitrary nonempty set. Denote

$$F(\boldsymbol{x}) = \bigcup_{k \in Q} \partial f_k(\boldsymbol{x}) \quad \text{and} \quad Q = \{1, 2, \dots, q\}.$$

We start the consideration by defining the notion of optimality for the multiobjective problem (5).

DEFINITION 4.1. A vector $\boldsymbol{x}^*$ is said to be a *global Pareto optimum* of (5), if there does not exist $\boldsymbol{x} \in S$ such, that $f_k(\boldsymbol{x}) \leq f_k(\boldsymbol{x}^*)$ for all $k = 1, \dots, q$ and $f_l(\boldsymbol{x}) < f_l(\boldsymbol{x}^*)$ for some l. Vector $\boldsymbol{x}^*$ is said to be a *global weak Pareto optimum* of (5), if there does not exist $\boldsymbol{x} \in S$ such, that $f_k(\boldsymbol{x}) < f_k(\boldsymbol{x}^*)$ for all $k = 1, \dots, q$. Vector $\boldsymbol{x}^*$ is a *local (weak) Pareto optimum* of (5), if there exists $\delta > 0$ such, that $\boldsymbol{x}^*$ is a global (weak) Pareto optimum on $B(\boldsymbol{x}^*; \delta) \cap S$.

Next we will present some optimality conditions of problem (5) in terms of cones. We also consider the unconstrained case, that is, when $S = \mathbb{R}^n$. We begin the considerations with the following lemma which can be found in [15] (Lemma 4.2).

LEMMA 4.2. *If $\boldsymbol{x}^*$ is a local weak Pareto optimum of problem* (5), *then $F^s(\boldsymbol{x}^*) \cap T_S(\boldsymbol{x}^*) = \emptyset$.*

PROOF. Let $\boldsymbol{x}^*$ be a local weak Pareto optimum. Then, there exists $\varepsilon > 0$ such that for every $\boldsymbol{y} \in S \cap B(\boldsymbol{x}^*, \varepsilon)$ there exists $k \in Q$ such that inequality $f_k(\boldsymbol{y}) \geq f_k(\boldsymbol{x}^*)$ holds. Let $\boldsymbol{d} \in T_S(\boldsymbol{x}^*)$ be arbitrary. Then, there exist sequences $(\boldsymbol{d}_i)$ and (t_i) such that $\boldsymbol{d}_i \to \boldsymbol{d}$, $t_i \downarrow 0$ and $\boldsymbol{x}^* + t_i\boldsymbol{d}_i \in S$ for all $i \in \mathbb{N}$. Also, there exists an index I_1 such that $\boldsymbol{x}^* + t_i\boldsymbol{d}_i \in S \cap B(\boldsymbol{x}^*, \varepsilon)$ for all $i > I_1$. Then for every $i > I_1$ there exists k_i such that $f_{k_i}(\boldsymbol{x}^* + t_i\boldsymbol{d}_i) \geq f_{k_i}(\boldsymbol{x}^*)$. Since the set Q is finite, there exists $\bar{k} \in Q$ and subsequences $(\boldsymbol{d}_{i_j}) \subset (\boldsymbol{d}_i)$ and $(t_{i_j}) \subset (t_i)$ such that

$$f_{\bar{k}}(\boldsymbol{x}^* + t_{i_j}\boldsymbol{d}_{i_j}) \geq f_{\bar{k}}(\boldsymbol{x}^*) \tag{6}$$

for all i_j with $j \in \mathbb{N}$ large enough. Denote $I_2 = \{i_j \mid i_j > I_1, j \in \mathbb{N}\}$. The Mean-Value Theorem (see e.g. [7]) implies that for all $\bar{i} \in I_2$ there exists $\tilde{t}_{\bar{i}} \in (0, t_{\bar{i}})$ such that

$$f_{\bar{k}}(\boldsymbol{x}^* + t_{\bar{i}}\boldsymbol{d}_{\bar{i}}) - f_{\bar{k}}(\boldsymbol{x}^*) \in \partial f_{\bar{k}}(\boldsymbol{x}^* + \tilde{t}_{\bar{i}}\boldsymbol{d}_{\bar{i}})^T t_{\bar{i}}\boldsymbol{d}_{\bar{i}}. \tag{7}$$

From the definition of generalized directional derivative (Definition 2.1), (6) and (7) we obtain

$$f^\circ_{\bar{k}}(\boldsymbol{x}^* + \tilde{t}_{\bar{i}}\boldsymbol{d}_{\bar{i}}; \boldsymbol{d}_{\bar{i}}) = \max_{\boldsymbol{\xi} \in \partial f_{\bar{k}}(\boldsymbol{x}^* + \tilde{t}_{\bar{i}}\boldsymbol{d}_{\bar{i}})} \boldsymbol{\xi}^T\boldsymbol{d}_{\bar{i}} \geq \frac{1}{t_{\bar{i}}}(f_{\bar{k}}(\boldsymbol{x}^* + t_{\bar{i}}\boldsymbol{d}_{\bar{i}}) - f_{\bar{k}}(\boldsymbol{x}^*)) \geq 0.$$

Thus, for all $\bar{i} \in I_2$ we have $f^\circ_{\bar{k}}(\boldsymbol{x}^* + \tilde{t}_{\bar{i}}\boldsymbol{d}_{\bar{i}}; \boldsymbol{d}_{\bar{i}}) \geq 0$. Since $\boldsymbol{d}_{\bar{i}} \to \boldsymbol{d}$ and $\boldsymbol{x}^* + \tilde{t}_{\bar{i}}\boldsymbol{d}_{\bar{i}} \to \boldsymbol{x}^*$ the upper semicontinuity of function $f^\circ_{\bar{k}}$ (Theorem 2.2, (iv)) implies

$$f^\circ_{\bar{k}}(\boldsymbol{x}^*, \boldsymbol{d}) \geq \lim_{\bar{i} \to \infty} f^\circ_{\bar{k}}(\boldsymbol{x}^* + \tilde{t}_{\bar{i}}\boldsymbol{d}_{\bar{i}}; \boldsymbol{d}_{\bar{i}}) \geq 0.$$

Thus, there exists $\boldsymbol{\xi} \in \partial f_{\bar{k}}(x^*) \subset F(\boldsymbol{x}^*)$ such that $\boldsymbol{\xi}^T\boldsymbol{d} \geq 0$ implying $\boldsymbol{d} \notin F^s(\boldsymbol{x}^*)$. □

Next, we will present a result for the unconstrained case. The result is analogous to Theorem 2.12.

THEOREM 4.3. *Let f_k be locally Lipschitz continuous for all $k \in Q$ and $S = \mathbb{R}^n$. If $\boldsymbol{x}^*$ is a local weak Pareto optimum of problem* (5), *then*

$$\boldsymbol{0} \in \operatorname{conv} F(\boldsymbol{x}^*)$$

PROOF. Since $S = \mathbb{R}^n$ we have $T_S(\boldsymbol{x}^*) = \mathbb{R}^n$ as well. Then by Lemma 4.2 we have $F^s(\boldsymbol{x}^*) = \emptyset$. Hence, for any $\boldsymbol{d} \in \mathbb{R}^n$ there exists $\boldsymbol{\xi} \in F(\boldsymbol{x}^*) \subset \operatorname{conv} F(\boldsymbol{x}^*)$ such that

$$\boldsymbol{d}^T\boldsymbol{\xi} \geq 0. \tag{8}$$

Suppose that $\mathbf{0} \notin \operatorname{conv} F(\boldsymbol{x}^*)$. Since the sets $\operatorname{conv} F(\boldsymbol{x}^*)$ and $\{\mathbf{0}\}$ are closed convex sets, there exists $\boldsymbol{d} \in \mathbb{R}^n$ and $a \in \mathbb{R}$ such that

$$0 = \boldsymbol{d}^T \mathbf{0} \geq a \quad \text{and} \quad \boldsymbol{d}^T \boldsymbol{\xi} < a \text{ for all } \boldsymbol{\xi} \in \operatorname{conv} F(\boldsymbol{x}^*)$$

according to the Separation Theorem (see e.g. [2]). From the first inequality we see that $a \leq 0$. Then the second inequality contradicts with inequality (8). Hence, $\mathbf{0} \in \operatorname{conv} F(\boldsymbol{x}^*)$. □

In the following we shall present the necessary optimality condition of problem (5) in terms of Clarke normal cone. The proof is quite similar to the proof for single objective case in [17, p. 72–73]. Before the condition we will present a useful lemma.

LEMMA 4.4. *If $\boldsymbol{x}^*$ is a local weak Pareto optimum of problem* (5)*, then it is local weak Pareto optimum of unconstrained problem*

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} \{f_1(\boldsymbol{x}) + K d_S(\boldsymbol{x}), f_2(\boldsymbol{x}) + K d_S(\boldsymbol{x}), \ldots, f_q(\boldsymbol{x}) + K d_S(\boldsymbol{x})\}, \tag{9}$$

where $K = \max\{K_1, K_2, \ldots, K_q\}$ and K_k is the Lipschitz constant of function f_k at point $\boldsymbol{x}^*$.

PROOF. From the definition of K and local weak Pareto optimality we see that there exists $\varepsilon > 0$ such that the Lipschitz condition holds for all f_k at $B(\boldsymbol{x}^*; \varepsilon)$ and $\boldsymbol{x}^*$ is weak Pareto optimum at $B(\boldsymbol{x}^*; \varepsilon) \cap S$. Suppose on the contrary that $\boldsymbol{x}^*$ is not a local weak Pareto optimum of problem (9). Then there exists $\boldsymbol{y} \in B(\boldsymbol{x}^*; \frac{\varepsilon}{2})$ such that

$$f_k(\boldsymbol{y}) + K d_S(\boldsymbol{y}) < f_k(\boldsymbol{x}^*) + K d_S(\boldsymbol{x}^*) = f_k(\boldsymbol{x}^*) \quad \text{for all } k \in Q. \tag{10}$$

Suppose $\boldsymbol{y} \in \operatorname{cl} S$. Then $K d_S(\boldsymbol{y}) = 0$ and by the continuity of f_k there exists $\delta > 0$ such that $f_k(\boldsymbol{z}) < f_k(\boldsymbol{x}^*)$ for all $k \in Q$ and $\boldsymbol{z} \in B(\boldsymbol{y}; \delta) \subset B(\boldsymbol{x}^*; \frac{\varepsilon}{2})$. Since $\boldsymbol{y} \in \operatorname{cl} S$ we have $S \cap B(\boldsymbol{y}; \delta) \cap B(\boldsymbol{x}^*; \frac{\varepsilon}{2}) \neq \emptyset$ and, thus, $\boldsymbol{x}^*$ is not a weak Pareto optimum of (5) in $S \cap B(\boldsymbol{x}^*; \varepsilon)$ contradicting the assumption. Hence, $\boldsymbol{y} \notin \operatorname{cl} S$ and $d_S(\boldsymbol{y}) > 0$.

By the definition of $d_S(\boldsymbol{y})$ there exists $\boldsymbol{c} \in \operatorname{cl} S$ such that $d_S(\boldsymbol{y}) = \|\boldsymbol{y} - \boldsymbol{c}\|$. Furthermore,

$$\|\boldsymbol{c} - \boldsymbol{y}\| \leq \|\boldsymbol{x}^* - \boldsymbol{y}\| < \frac{\varepsilon}{2}.$$

Thus,

$$\|\boldsymbol{c} - \boldsymbol{x}^*\| \leq \|\boldsymbol{c} - \boldsymbol{y}\| + \|\boldsymbol{y} - \boldsymbol{x}^*\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

implying $\boldsymbol{c} \in B(\boldsymbol{x}^*; \varepsilon)$. By inequality (10) and local weak Pareto optimality of $\boldsymbol{x}^*$ there exists $k_1 \in Q$ such that

$$f_{k_1}(\boldsymbol{y}) < f_{k_1}(\boldsymbol{x}^*) \leq f_{k_1}(\boldsymbol{c}).$$

Hence,

$$|f_{k_1}(\boldsymbol{x}^*) - f_{k_1}(\boldsymbol{y})| \leq |f_{k_1}(\boldsymbol{c}) - f_{k_1}(\boldsymbol{y})| \leq K \|\boldsymbol{y} - \boldsymbol{c}\| = K d_S(\boldsymbol{y})$$

implying $f_{k_1}(\boldsymbol{x}^*) \leq f_{k_1}(\boldsymbol{y}) + K d_S(\boldsymbol{y})$. This contradicts with inequality (10). Thus, $\boldsymbol{x}^*$ is a local weak Pareto optimum of problem (9). □

Finally, we can state the necessary optimality condition of problem (5) with arbitrary nonempty feasible set $S \subset \mathbb{R}^n$.

THEOREM 4.5. *If $\boldsymbol{x}^*$ is a local weak Pareto minimum of* (5)*, then*

$$\mathbf{0} \in \operatorname{conv} F(\boldsymbol{x}^*) + N_S(\boldsymbol{x}^*). \tag{11}$$

PROOF. By Lemma 4.4 $\boldsymbol{x}^*$ is a local weak Pareto optimum of unconstrained problem (9). Consider kth objective function of the unconstrained problem. By Theorem 2.6 we have

$$\partial(f_k(\boldsymbol{x}) + K d_S(\boldsymbol{x})) \subset \partial f_k(\boldsymbol{x}) + K\partial d_S(\boldsymbol{x}).$$

The Definition 2.7 of normal cone implies $K\partial d_S(\boldsymbol{x}) \subset N_S(\boldsymbol{x})$. Since $\boldsymbol{x}^*$ is a local weak Pareto optimum of problem (9), Lemma 4.3 implies

$$\mathbf{0} \in \operatorname{conv} \bigcup_{k\in Q} \partial(f_k(\boldsymbol{x}^*) + K d_S(\boldsymbol{x}^*)) \subset \operatorname{conv} \bigcup_{k\in Q} (\partial f_k(\boldsymbol{x}^*) + N_S(\boldsymbol{x}^*)).$$

By Lemma 2.9 (iv) we have

$$\operatorname{conv} \bigcup_{k\in Q} (\partial f_k(\boldsymbol{x}^*) + N_S(\boldsymbol{x}^*)) = \operatorname{conv} F(\boldsymbol{x}^*) + N_S(\boldsymbol{x}^*),$$

as desired. □

Since Pareto optimality implies weak Pareto optimality we get immediately the following consequence.

COROLLARY 4.6. *Condition* (11) *is also necessary for $\boldsymbol{x}^*$ to be a local Pareto optimum of* (5).

To prove a sufficient condition for global optimality we need the assumptions that S is convex and f_k are f°-pseudoconvex for all $k \in Q$.

THEOREM 4.7. *Let f_k be f°-pseudoconvex for all $k \in Q$ and S convex. Then $\boldsymbol{x}^* \in S$ is a global weak Pareto minimum of* (5)*, if and only if*

$$\mathbf{0} \in \operatorname{conv} F(\boldsymbol{x}^*) + N_S(\boldsymbol{x}^*).$$

PROOF. The necessity follows directly from Theorem 4.5. For sufficiency let $\mathbf{0} \in \operatorname{conv} F(\boldsymbol{x}^*) + N_S(\boldsymbol{x}^*)$. Then there exist $\boldsymbol{\xi}_* \in \operatorname{conv} F(\boldsymbol{x}^*)$ and $\boldsymbol{z}_* \in N_S(\boldsymbol{x}^*)$ such that $\boldsymbol{\xi}_* = -\boldsymbol{z}_*$. Then by Theorem 2.8 we have for all $\boldsymbol{x} \in S$ that

$$0 \leq -\boldsymbol{z}_*^T(\boldsymbol{x} - \boldsymbol{x}^*) = \boldsymbol{\xi}_*^T(\boldsymbol{x} - \boldsymbol{x}^*) = \sum_{k=1}^{q} \lambda_k \boldsymbol{\xi}_k^T(\boldsymbol{x} - \boldsymbol{x}^*),$$

where $\lambda_k \geq 0$, $\boldsymbol{\xi}_k \in \partial f_k(\boldsymbol{x}^*)$ for all $k \in Q$ and $\sum_{k=1}^{q} \lambda_k = 1$. Thus, there exists k_1 such that $f_{k_1}^\circ(\boldsymbol{x}^*, \boldsymbol{x} - \boldsymbol{x}^*) \geq \boldsymbol{\xi}_{k_1}^T(\boldsymbol{x} - \boldsymbol{x}^*) \geq 0$. Then by Lemma 3.2 the f°-pseudoconvexity of f_{k_1} implies $f_{k_1}(\boldsymbol{x}) \geq f_{k_1}(\boldsymbol{x}^*)$. Thus, there exists no feasible point $\boldsymbol{x} \in S$ with $f_k(\boldsymbol{x}) < f_k(\boldsymbol{x}^*)$ for all $k \in Q$ implying $\boldsymbol{x}^*$ is a global weak Pareto optimum. □

4.1 Inequality constraints

Now we shall consider problem (5) with inequality constraints:

$$\begin{cases} \text{minimize} & \{f_1(\boldsymbol{x}), \ldots, f_q(\boldsymbol{x})\} \\ \text{subject to} & g_i(\boldsymbol{x}) \leq 0 \quad \text{for all } i = 1, \ldots, m, \end{cases} \tag{12}$$

where also $g_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$ are locally Lipschitz continuous functions. Denote $M = \{1, 2, \ldots, m\}$ and the *total constraint function* by

$$g(\boldsymbol{x}) = \max\{g_i(\boldsymbol{x}) \mid i = 1, \ldots, m\}.$$

Problem (12) can be seen as a special case of (5), where

$$S = \{\boldsymbol{x} \in \mathbb{R}^n \mid g(\boldsymbol{x}) \leq 0\}.$$

Denote also

$$G(\boldsymbol{x}) = \bigcup_{i \in I(\boldsymbol{x})} \partial g_i(\boldsymbol{x}), \text{ where } I(\boldsymbol{x}) = \{i \mid g_i(\boldsymbol{x}) = 0\}.$$

For necessary conditions we need some constraint qualifications. We restrict ourselves to constraint qualifications that give conditions in terms of feasible set or constraint functions. This makes the constraint qualifications easily applicable to both single and multiobjective problems. There are many constraint qualifications involving the objective functions too (see e.g. [15]), but they are not considered here.

In order to formulate Karush-Kuhn-Tucker (KKT) type optimality conditions we need one of the following constraint qualifications

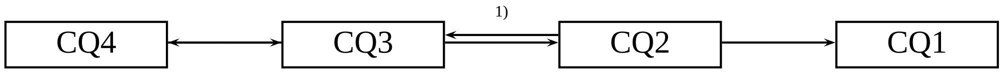
$$\begin{array}{ll} \text{(CQ1)} & G^-(\boldsymbol{x}) \subset T_S(\boldsymbol{x}) \\ \text{(CQ2)} & \mathbf{0} \notin \partial g(\boldsymbol{x}) \\ \text{(CQ3)} & G^s(\boldsymbol{x}) \neq \emptyset \\ \text{(CQ4)} & \mathbf{0} \notin \operatorname{conv} G(\boldsymbol{x}), \end{array}$$

where we assume $I(\boldsymbol{x}) \neq \emptyset$ for all the constraint qualifications. Due to Theorem 2.2 (ii) the assumption $I(\boldsymbol{x}) \neq \emptyset$ guarantees that $G(\boldsymbol{x}) \neq \emptyset$. Note that the sets $G^-(\boldsymbol{x})$ and $G^s(\boldsymbol{x})$ can be defined also in terms of generalized directional derivatives. For example

$$\begin{aligned} G^-(\boldsymbol{x}) &= \{\boldsymbol{d} \mid \boldsymbol{\xi}^T \boldsymbol{d} \leq 0, \text{ for all } \boldsymbol{\xi} \in \bigcup_{i \in I(\boldsymbol{x})} \partial g_i(\boldsymbol{x})\} \\ &= \{\boldsymbol{d} \mid g_i^\circ(\boldsymbol{x}; \boldsymbol{d}) \leq 0, \text{ for all } i \in I(\boldsymbol{x})\}. \end{aligned}$$

In [15] CQ1 and CQ3 were called nonsmooth analogs of Abadie qualification and Cottle qualification respectively, while both CQ4 and CQ2 were called Cottle constraint qualifications in [19] and [17] respectively. In [15] it was shown that CQ1 follows from CQ3. In the appendix we will show that the following relations hold between the given constraint qualifications.

Figure 2: Relations between different constraint qualifications



[1)] If all constraint functions are subdifferentially regular or f°-pseudoconvex.

Next, we will prove a KKT Theorem in the case where the constraint qualification is CQ1. As seen in Figure 2, CQ1 is the weakest condition of the above qualifications. Thus, CQ1 can be replaced by any of CQ2, CQ3 or CQ4. The proof of the KKT Theorem is in practice the same as in [15]. The idea is quite similar to the proof in [2, p. 165] for differentiable single objective case. The outline of the proof goes as follows. First we characterize a necessary conditon for (weak Pareto) optimality in terms of contingent cone and objective function(s). Then, by some constraint qualification we replace the contingent cone by another cone, related to constraint functions and, finally, by some alternative theorem we may express the optimality in the form of KKT conditions. The main difference between the differentiable and nondifferentiable case is that the cones are defined with generalized directional derivatives (or subdifferentials) instead of classical gradients.

The weak Pareto optimality was expressed in terms of contingent cone and objective functions in Lemma 4.2. Let us then prove the theorem of alternatives needed in the proof of the KKT Theorem.

LEMMA 4.8. *Let $A \subset \mathbb{R}^n$ be a nonempty closed convex set and let $C \subset \mathbb{R}^n$ be a nonempty closed convex cone. Then one and only one of the following relations hold*

1. $A \cap C \neq \emptyset$
2. $A^s \cap -C^- \neq \emptyset$.

PROOF. Assume that $A \cap C \neq \emptyset$. If $A^s = \emptyset$ then trivially $A^s \cap -C^- = \emptyset$. If $\boldsymbol{d} \in A^s \neq \emptyset$, we have $\boldsymbol{a}^T\boldsymbol{d} < 0$ for all $\boldsymbol{a} \in A \cap C$. Thus, $\boldsymbol{d} \notin -C^- = \{\boldsymbol{x} \mid \boldsymbol{x}^T\boldsymbol{c} \geq 0, \forall\, \boldsymbol{c} \in C\}$ and $A^s \cap -C^- = \emptyset$.

Assume next that $A \cap C = \emptyset$. Since A and C are closed convex sets the Separation Theorem (see e.g. [2]) implies there exist $\boldsymbol{d} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ such that

$$\boldsymbol{d}^T\boldsymbol{a} < \alpha \quad \forall\, \boldsymbol{a} \in A \tag{13}$$

$$\boldsymbol{d}^T\boldsymbol{c} \geq \alpha \quad \forall\, \boldsymbol{c} \in C. \tag{14}$$

Since C is a cone, $\mathbf{0} \in C$ and C is unbounded, we can choose $\alpha = 0$. Then, equation (13) means that $\boldsymbol{d} \in A^s$ and equation (14) means that $\boldsymbol{d} \in -C^-$. Thus, $\boldsymbol{d} \in A^s \cap -C^- \neq \emptyset$. □

The following results are useful in the proof of necessary conditions.

LEMMA 4.9. *Let f_k, $k \in Q$ and g_i, $i \in M$ be locally Lipschitz continuous and $A \subset \mathbb{R}^n$ an arbitrary set. Then*

$$A^- = (\operatorname{cl} A)^-, \quad F^s(\boldsymbol{x}) = (\operatorname{conv} F(\boldsymbol{x}))^s \quad \textit{and} \quad G^-(\boldsymbol{x}) = (\operatorname{cone} G(\boldsymbol{x}))^-.$$

PROOF. Since

$$A \subset \operatorname{cl} A, \quad F(\boldsymbol{x}) \subset \operatorname{conv} F(\boldsymbol{x}) \quad \text{and} \quad G(\boldsymbol{x}) \subset \operatorname{cone} G(\boldsymbol{x})$$

clearly

$$(\operatorname{cl} A)^- \subset A^-, \quad (\operatorname{conv} F(\boldsymbol{x}))^s \subset F^s(\boldsymbol{x}) \quad \text{and} \quad (\operatorname{cone} G(\boldsymbol{x}))^- \subset G^-(\boldsymbol{x}).$$

Suppose that $\boldsymbol{d} \in A^-$. If $\boldsymbol{d} \notin (\operatorname{cl} A)^-$ then $\boldsymbol{d}^T\boldsymbol{a} > 0$ for some $\boldsymbol{a} \in \operatorname{cl} A$. By the continuity of function $\boldsymbol{d}^T\boldsymbol{a}$ there exists $\varepsilon > 0$ such that $\boldsymbol{d}^T\boldsymbol{b} > 0$ for all $\boldsymbol{b} \in B(\boldsymbol{a};\varepsilon)$. This contradicts with assumption $\boldsymbol{d} \in A^-$ as $B(\boldsymbol{a};\varepsilon) \cap A \neq \emptyset$.

Suppose that $\boldsymbol{d} \in F^s(\boldsymbol{x})$. Then for every $\boldsymbol{\xi} \in \bigcup_{k\in Q} \partial f_k(\boldsymbol{x})$ we have $\boldsymbol{d}^T\boldsymbol{\xi} < 0$. Then

$$\boldsymbol{d}^T(\sum_{k=1}^{q} \lambda_k \boldsymbol{\xi}_k) = \sum_{k=1}^{q} \lambda_k \boldsymbol{d}^T \boldsymbol{\xi}_k < 0,$$

for all $\boldsymbol{\xi}_k \in \partial f_k(\boldsymbol{x})$ and $\lambda_k \geq 0$, $\sum_{k=1}^q \lambda_k = 1$. Hence, $\boldsymbol{d} \in (\operatorname{conv} F(\boldsymbol{x}))^s$.

Suppose that $\boldsymbol{d} \in G^-(\boldsymbol{x})$. Likewise to the previous case we can show that $\boldsymbol{d} \in (\operatorname{conv} G(\boldsymbol{x}))^-$. Then

$$\boldsymbol{d}^T\boldsymbol{\xi} \leq 0 \quad \text{implying} \quad \boldsymbol{d}^T\lambda\boldsymbol{\xi} \leq 0$$

for all $\lambda \geq 0$ and $\boldsymbol{\xi} \in \operatorname{conv} G(\boldsymbol{x})$. Hence, $\boldsymbol{d} \in (\operatorname{cone} G(\boldsymbol{x}))^-$. □

Now, we are ready to formulate the necessary condition for local weak Pareto optimality.

THEOREM 4.10. *If $\boldsymbol{x}^*$ is a local weak Pareto optimum and CQ1 holds then*

$$\mathbf{0} \in \operatorname{conv} F(\boldsymbol{x}^*) + \operatorname{cl} \operatorname{cone} G(\boldsymbol{x}^*). \tag{15}$$

PROOF. By Lemma 4.2 $F^s(\boldsymbol{x}^*) \cap T_S(\boldsymbol{x}^*) = \emptyset$. Since the CQ1 holds we have

$$F^s(\boldsymbol{x}^*) \cap G^-(\boldsymbol{x}^*) \subset F^s(\boldsymbol{x}^*) \cap T_S(\boldsymbol{x}^*) = \emptyset.$$

By Lemma 4.9 we have

$$\begin{aligned} F^s(\boldsymbol{x}^*) \cap G^-(\boldsymbol{x}^*) &= (\operatorname{conv} F(\boldsymbol{x}^*))^s \cap (\operatorname{cone} G(\boldsymbol{x}^*))^- \\ &= (\operatorname{conv} F(\boldsymbol{x}^*))^s \cap (\operatorname{cl} \operatorname{cone} G(\boldsymbol{x}^*))^- = \emptyset. \end{aligned}$$

Since $F(\boldsymbol{x}^*)$ and $G(\boldsymbol{x}^*)$ are nonempty $(I(\boldsymbol{x}^*) \neq \emptyset)$, $\operatorname{conv} F(\boldsymbol{x}^*)$ is a closed convex set (Corollary 2.11) and $\operatorname{cl} \operatorname{cone} G(\boldsymbol{x}^*)$ is a closed convex cone. Then Lemma 4.8 implies

$$\operatorname{conv} F(\boldsymbol{x}^*) \cap -\operatorname{cl} \operatorname{cone} G(\boldsymbol{x}^*) \neq \emptyset.$$

This is equivalent with $\mathbf{0} \in \operatorname{conv} F(\boldsymbol{x}^*) + \operatorname{cl} \operatorname{cone} G(\boldsymbol{x}^*)$. □

Since Pareto optimality implies weak Pareto optimality we get immediately the following consequence.

COROLLARY 4.11. *Condition* (15) *is also necessary for* $\boldsymbol{x}^*$ *to be a local Pareto optimum of* (12).

In Theorem 4.10 it was assumed that $I(\boldsymbol{x}) \neq \emptyset$. If this is not the case, then we have $g(\boldsymbol{x}) < 0$. By continuity of g there exists $\varepsilon > 0$ such that $B(\boldsymbol{x}; \varepsilon)$ belongs to the feasible set. Then $N_S(\boldsymbol{x}) = \{0\}$ and with Theorem 4.5 we may deduce that condition in Theorem 4.3 holds. From that we may deduce that assumption $I(\boldsymbol{x}) \neq \emptyset$ could be omitted if in (15) $\operatorname{cl} \operatorname{cone} G(\boldsymbol{x}^*)$ is replaced by $\{0\} \cup \operatorname{cl} \operatorname{cone} G(\boldsymbol{x}^*)$.

A condition stronger than (15) was developed for CQ3 in [15] and [19]. Next we shall study the stronger condition. For that we need the following lemma.

LEMMA 4.12. *If CQ4 (or equivalently CQ3) holds at* $\boldsymbol{x} \in \mathbb{R}^n$, *then* $\operatorname{cone} G(\boldsymbol{x})$ *is closed.*

PROOF. Let $(\boldsymbol{d}_j) \subset \operatorname{cone} G(\boldsymbol{x})$ be an arbitrary converging sequence such that $\lim_{j\to\infty} \boldsymbol{d}_j = \hat{\boldsymbol{d}}$. For every j there exists $\lambda_j \geq 0$ and $\boldsymbol{\xi}_j \in \operatorname{conv} G(\boldsymbol{x})$ such that $\boldsymbol{d}_j = \lambda_j \boldsymbol{\xi}_j$. By Corollary 2.11 $\operatorname{conv} G(\boldsymbol{x})$ is a compact set. Then there exists a converging subsequence $(\boldsymbol{\xi}_{j_i})$ such that $\lim_{i\to\infty} \boldsymbol{\xi}_{j_i} = \hat{\boldsymbol{\xi}}$. By closedness of $\operatorname{conv} G(\boldsymbol{x})$ we have $\hat{\boldsymbol{\xi}} \in \operatorname{conv} G(\boldsymbol{x})$. Since $\mathbf{0} \notin \operatorname{conv} G(\boldsymbol{x})$ sequence

$$\lambda_{j_i} = \frac{\|\boldsymbol{d}_{j_i}\|}{\|\boldsymbol{\xi}_{j_i}\|}$$

is converging too. Denote $\lim_{i\to\infty} \lambda_{j_i} = \hat{\lambda}$. Then

$$\hat{\boldsymbol{d}} = \hat{\lambda}\hat{\boldsymbol{\xi}} \in \operatorname{cone} G(\boldsymbol{x})$$

implying that $\operatorname{cone} G(\boldsymbol{x})$ is closed. □

THEOREM 4.13. *If* $\boldsymbol{x}^*$ *is a local weak Pareto optimum and CQ3 holds, then*

$$\mathbf{0} \in \operatorname{conv} F(\boldsymbol{x}^*) + \operatorname{cone} G(\boldsymbol{x}^*).$$

PROOF. From Lemma 4.12 it follows that if CQ3 holds then $\operatorname{cl} \operatorname{cone} G(\boldsymbol{x}^*) = \operatorname{cone} G(\boldsymbol{x}^*)$. Then the result follows directly from Theorem 4.10. □

Consider then the sufficient conditions of problem (12). It is well-known that the convexity of the functions f_k, $k \in Q$, and g_i, $i \in M$, guarantees the sufficiency of the KKT optimality condition for global weak Pareto optimality in Theorem 4.13 (see [19, p. 51]). We will present the sufficient conditions in more detail later. Namely, they can be obtained as a special case of sufficient conditions for problems with both inequality and equality constraints.

4.2 Equality constraints

Consider problem (5) with both inequality and equality constraints.

$$\begin{cases} \text{minimize} & \{f_1(\boldsymbol{x}), \dots, f_q(\boldsymbol{x})\} \\ \text{subject to} & g_i(\boldsymbol{x}) \le 0 \quad \text{for all } i = 1, \dots, m, \\ & h_j(\boldsymbol{x}) = 0 \quad \text{for all } j = 1, \dots, p, \end{cases} \tag{16}$$

where all functions are supposed to be locally Lipschitz continuous. Denote $H(\boldsymbol{x}) = \bigcup_{j=1}^{p} \partial h_j(\boldsymbol{x})$ and $J = \{1, 2, \dots, p\}$. By Lemma 2.3 we see that

$$-H(\boldsymbol{x}) = -\bigcup_{j \in J} \partial h_j(\boldsymbol{x}) = \bigcup_{j \in J} \partial(-h_j)(\boldsymbol{x}).$$

A straightforward way to deal with an equality constraint $h_j(\boldsymbol{x}) = 0$ is to replace it with two inequality constraints

$$h_j(\boldsymbol{x}) \le 0 \quad \text{and} \quad -h_j(\boldsymbol{x}) \le 0. \tag{17}$$

Then, we may use the results obtained for problem (12) to derive results for problem (16). However, some constraint qualifications are not satisfied if this kind of operation is done as we will see soon.

Consider first the CQ1. Denote

$$\begin{aligned} G_*^-(\boldsymbol{x}) &= \{\boldsymbol{d} \mid g_i^\circ(\boldsymbol{x}; \boldsymbol{d}) \le 0,\ i \in I(\boldsymbol{x}),\ h_j^\circ(\boldsymbol{x}; \boldsymbol{d}) \le 0,\ (-h_j)^\circ(\boldsymbol{x}; \boldsymbol{d}) \le 0,\ j \in J\} \\ &= G^-(\boldsymbol{x}) \cap H^-(\boldsymbol{x}) \cap (-H)^-(\boldsymbol{x}). \end{aligned}$$

It is good to note that we can replace $(-h_j)^\circ(\boldsymbol{x}; \boldsymbol{d}) \le 0$ by $h_j^\circ(\boldsymbol{x}; -\boldsymbol{d}) \le 0$ in the definition of $G_*^-(\boldsymbol{x})$ according to Theorem 2.2 (i). We can use a new cone instead of the cone $H^-(\boldsymbol{x}) \cap (-H)^-(\boldsymbol{x})$ as the next lemma shows.

LEMMA 4.14. *Let $h : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz continuous function. Then*

$$\begin{aligned} \partial h(\boldsymbol{x})^- \cap (-\partial h(\boldsymbol{x}))^- &= \{\boldsymbol{d} \mid h^\circ(\boldsymbol{x}; \boldsymbol{d}) \le 0,\ h^\circ(\boldsymbol{x}; -\boldsymbol{d}) \le 0\} \\ &\subset \{\boldsymbol{d} \mid h^\circ(\boldsymbol{x}; \boldsymbol{d}) = 0\} \end{aligned}$$

PROOF. Suppose $\boldsymbol{d} \in \partial h(\boldsymbol{x})^- \cap (-\partial h(\boldsymbol{x}))^-$. By the subadditivity of h° (Theorem 2.2 (i)) we have

$$0 = h^\circ(\boldsymbol{x}; \boldsymbol{0}) \le h^\circ(\boldsymbol{x}; -\boldsymbol{d}) + h^\circ(\boldsymbol{x}; \boldsymbol{d}) \le 0, \tag{18}$$

which is possible only if $h^\circ(\boldsymbol{x}; -\boldsymbol{d}) = h^\circ(\boldsymbol{x}; \boldsymbol{d}) = 0$. Namely, if one would be strictly negative the other should be strictly positive in order to satisfy inequality (18). This is impossible as $\boldsymbol{d} \in \partial h(\boldsymbol{x})^- \cap (-\partial h(\boldsymbol{x}))^-$. □

Denote

$$H^0(\boldsymbol{x}) = \{\boldsymbol{d} \mid h_j^\circ(\boldsymbol{x}; \boldsymbol{d}) = 0 \ \text{for all } j \in J\}.$$

From Lemma 4.14 we can easily deduce that $H^-(\boldsymbol{x}) \cap (-H)^-(\boldsymbol{x}) \subset H^0(\boldsymbol{x})$. However, in general $H^0(\boldsymbol{x}) \not\subset H^-(\boldsymbol{x}) \cap (-H)^-(\boldsymbol{x})$. To see this, consider a function

$$h(x) = \begin{cases} -x & \text{, if } x \leq 0 \\ 0 & \text{, otherwise.} \end{cases}$$

Then $h^\circ(0,1) = 0$ and $h^\circ(0,-1) = 1$. Thus, $1 \in H^0(0)$ but $1 \notin H^-(0) \cap (-H)^-(0)$.

Now we can present two constraint qualifications for problem (16):

$$\begin{array}{ll} \text{(CQ5)} & G^-(\boldsymbol{x}) \cap H^-(\boldsymbol{x}) \cap (-H)^-(\boldsymbol{x}) \subset T_S(\boldsymbol{x}) \\ \text{(CQ6)} & G^-(\boldsymbol{x}) \cap H^0(\boldsymbol{x}) \subset T_S(\boldsymbol{x}), \end{array}$$

where again $I(\boldsymbol{x}) \neq \emptyset$. From Lemma 4.14 we see that CQ6 implies CQ5. Thus, we can derive KKT conditions with CQ6 if we can do so for CQ5.

Consider next the constraint qualification CQ2. Assume our problem has an equality constraint $h_1(\boldsymbol{x}) = 0$. Then, at the feasible points the total constraint function will be

$$g(\boldsymbol{x}) = \max\{h_1(\boldsymbol{x}), -h_1(\boldsymbol{x}), l(\boldsymbol{x})\} = \max\{\max\{h_1(\boldsymbol{x}), -h_1(\boldsymbol{x})\}, l(\boldsymbol{x})\},$$

where $l(\boldsymbol{x})$ contains the other terms. It is clear that function $\max\{h_1(\boldsymbol{x}), -h_1(\boldsymbol{x})\}$ is non-negative. Consequently, g is non-negative too. Then, 0 is minimum value for g and it is attained at every feasible point of problem (16). Thus, for any feasible $\boldsymbol{x}$ we have $\mathbf{0} \in \partial g(\boldsymbol{x})$ according to Theorem 2.12 and, thus, CQ2 does not hold. Hence, CQ2 is not suitable for problems with equality constraints.

Next, we shall consider CQ3. Denote

$$\begin{aligned} G^s_*(\boldsymbol{x}) &= \{\boldsymbol{d} \mid g_i^\circ(\boldsymbol{x};\boldsymbol{d}) < 0,\, i \in I(\boldsymbol{x}),\, h_j^\circ(\boldsymbol{x};\boldsymbol{d}) < 0,\, (-h_j)^\circ(\boldsymbol{x};\boldsymbol{d}) < 0,\, j \in J\} \\ &= G^s(\boldsymbol{x}) \cap \{\boldsymbol{d} \mid h_j^\circ(\boldsymbol{x};\boldsymbol{d}) < 0,\, h_j^\circ(\boldsymbol{x};-\boldsymbol{d}) < 0,\, j \in J\}. \end{aligned}$$

Let $\boldsymbol{x}, \boldsymbol{d} \in \mathbb{R}^n$ and $j \in J$ be arbitrary. By the subadditivity of h_j° we have

$$0 = h_j^\circ(\boldsymbol{x}, \mathbf{0}) \leq h_j^\circ(\boldsymbol{x}, \boldsymbol{d}) + h_j^\circ(\boldsymbol{x}, -\boldsymbol{d}). \tag{19}$$

From inequality (19) it is easy to see that $\{\boldsymbol{d} \mid h_j^\circ(\boldsymbol{x};\boldsymbol{d}) < 0,\, h_j^\circ(\boldsymbol{x};-\boldsymbol{d}) < 0\} = \emptyset$. Hence, CQ3 does not hold implying that the constraint qualification CQ3 (or CQ4) is not suitable for equality constraints.

Before the proof of the KKT Theorem of problem (16) we need the following lemma.

LEMMA 4.15. *If A and B are nonempty cones then* $\operatorname{cl}(A+B) \subset \operatorname{cl} A + \operatorname{cl} B$.

PROOF. Since $A \subset \operatorname{cl} A$ and $B \subset \operatorname{cl} B$ we have $A + B \subset \operatorname{cl} A + \operatorname{cl} B$. By Lemma 2 in [20] $\operatorname{cl} A + \operatorname{cl} B$ is closed. Thus, $\operatorname{cl}(A+B) \subset \operatorname{cl} A + \operatorname{cl} B$. □

Finally, we can state the theorem corresponding to Theorem 4.10 with constraint qualification CQ5.

THEOREM 4.16. *If $\boldsymbol{x}^*$ is a local weak Pareto optimum of* (16) *and CQ5 holds at $\boldsymbol{x}^*$, then*

$$\mathbf{0} \in \operatorname{conv} F(\boldsymbol{x}^*) + \operatorname{cl} \operatorname{cone} G(\boldsymbol{x}^*) + \operatorname{cl} \operatorname{cone} H(\boldsymbol{x}^*) - \operatorname{cl} \operatorname{cone} H(\boldsymbol{x}^*). \tag{20}$$

PROOF. From Theorem 4.10 and previous considerations we see that

$$\mathbf{0} \in \operatorname{conv} F(\boldsymbol{x}^*) + \operatorname{cl} \operatorname{cone}(G(\boldsymbol{x}^*) \cup H(\boldsymbol{x}^*) \cup -H(\boldsymbol{x}^*)). \tag{21}$$

By using Lemma 2.9 (ii) twice and Lemma 4.15 we obtain

$$\begin{aligned}
& \operatorname{cl} \operatorname{cone}(G(\boldsymbol{x}^*) \cup H(\boldsymbol{x}^*) \cup -H(\boldsymbol{x}^*)) \\
= \; & \operatorname{cl} \left(\sum_{i \in I(\boldsymbol{x}^*)} \operatorname{ray} \partial g_i(\boldsymbol{x}^*) + \sum_{j \in J} \operatorname{ray} \partial h_j(\boldsymbol{x}^*) + \sum_{j \in J} \operatorname{ray} \partial(-h_j(\boldsymbol{x}^*)) \right) \\
= \; & \operatorname{cl}(\operatorname{cone} G(\boldsymbol{x}^*) + \operatorname{cone} H(\boldsymbol{x}^*) - \operatorname{cone} H(\boldsymbol{x}^*)) \\
\subset \; & \operatorname{cl} \operatorname{cone} G(\boldsymbol{x}^*) + \operatorname{cl} \operatorname{cone} H(\boldsymbol{x}^*) - \operatorname{cl} \operatorname{cone} H(\boldsymbol{x}^*).
\end{aligned}$$

Combining this with relation (21) proves the theorem. □

There are papers dealing with equality constraints in nonsmooth problems without turning them into inequality constraints (see e.g. [12]). However, the conditions are expressed in terms of generalized Jacobian of multivalued mapping $h : \mathbb{R}^m \to \mathbb{R}^n$. We shall not consider generalized Jacobians here and, thus, will not discuss these type of conditions further.

There are also papers where closures are not needed in conditions in Theorem 4.16 (see e.g [11]). But there they used constraint qualifications including objective functions which we shall not consider either.

After the necessary conditions we shall now study sufficient conditions. For that we do not need the constraint qualifications but we have to make some assumptions on objective and constraint functions. More accurately, we assume that objective functions are f°-pseudoconvex and inequality constraint functions are f°-quasiconvex. The equality constraints may be f°-quasiconvex or f°-quasiconcave. Denote

$$H_+(\boldsymbol{x}) = \bigcup_{j \in J_+} \partial h_j(\boldsymbol{x}) \quad \text{and} \quad H_-(\boldsymbol{x}) = \bigcup_{j \in J_-} \partial h_j(\boldsymbol{x}),$$

where $J_- \cup J_+ = J$ and h_j is f°-quasiconvex if $j \in J_+$ and h_j is f°-quasiconcave if $j \in J_-$.

THEOREM 4.17. *Let $\boldsymbol{x}^*$ be a feasible point of problem* (16). *Suppose f_k are f°-pseudoconvex for all $k \in Q$, g_i are f°-quasiconvex for all $i \in M$, h_j are f°-quasiconvex for all $j \in J_+$ and f°-quasiconcave for all $j \in J_-$. If*

$$\mathbf{0} \in \operatorname{conv} F(\boldsymbol{x}^*) + \operatorname{cone} G(\boldsymbol{x}^*) + \operatorname{cone} H_+(\boldsymbol{x}^*) - \operatorname{cone} H_-(\boldsymbol{x}^*), \tag{22}$$

then $\boldsymbol{x}^*$ is a global weak Pareto optimum of (16).

PROOF. Note that if (22) is satisfied then $I(\boldsymbol{x}^*) \neq \emptyset$. Let $\boldsymbol{x} \in \mathbb{R}^n$ be an arbitrary feasible point. Then $g_i(\boldsymbol{x}) \leq g_i(\boldsymbol{x}^*)$ if $i \in I(\boldsymbol{x}^*)$, $h_j(\boldsymbol{x}) = h_j(\boldsymbol{x}^*)$ for all $j \in J_+ \cup J_-$ and f°-quasiconvexity implies that

$$g_i^\circ(\boldsymbol{x}^*; \boldsymbol{x} - \boldsymbol{x}^*) \leq 0 \text{ for all } i \in I(\boldsymbol{x}^*) \tag{23}$$

$$h_j^\circ(\boldsymbol{x}^*; \boldsymbol{x} - \boldsymbol{x}^*) \leq 0 \text{ for all } j \in J_+. \tag{24}$$

The f°-quasiconcavity implies that

$$h_j^\circ(\boldsymbol{x}^*; \boldsymbol{x}^* - \boldsymbol{x}) \leq 0 \text{ for all } j \in J_-. \tag{25}$$

According to (22) there exist $\boldsymbol{\xi}_k \in \partial f_k(\boldsymbol{x}^*)$, $\boldsymbol{\zeta}_i \in \partial g_i(\boldsymbol{x}^*)$, $\boldsymbol{\eta}_j \in \partial h_j(\boldsymbol{x}^*)$ and coefficients $\lambda_k, \mu_i, \nu_j \geq 0$, for all $k \in Q$, $i \in I(\boldsymbol{x}^*)$ and $j \in J$ such that $\sum_{k=1}^q \lambda_k = 1$ and

$$\mathbf{0} = \sum_{k \in Q} \lambda_k \boldsymbol{\xi}_k + \sum_{i \in I(\boldsymbol{x}^*)} \mu_i \boldsymbol{\zeta}_i + \sum_{j \in J_+} \nu_j \boldsymbol{\eta}_j - \sum_{j \in J_-} \nu_j \boldsymbol{\eta}_j. \tag{26}$$

Multiplying equation (26) by $\boldsymbol{x} - \boldsymbol{x}^*$, using Definition 2.1 and equations (23), (24) and (25) we obtain

$$\begin{aligned}
& - \sum_{k \in Q} \lambda_k \boldsymbol{\xi}_k^T (\boldsymbol{x} - \boldsymbol{x}^*) \\
= \quad & \sum_{i \in I(\boldsymbol{x}^*)} \mu_i \boldsymbol{\zeta}_i^T (\boldsymbol{x} - \boldsymbol{x}^*) + \sum_{j \in J_+} \nu_j \boldsymbol{\eta}_j^T (\boldsymbol{x} - \boldsymbol{x}^*) + \sum_{j \in J_-} \nu_j \boldsymbol{\eta}_j^T (\boldsymbol{x}^* - \boldsymbol{x}) \\
\leq \quad & \sum_{i \in I(\boldsymbol{x}^*)} \mu_i g_i^\circ(\boldsymbol{x}^*; \boldsymbol{x} - \boldsymbol{x}^*) + \sum_{j \in J_+} \nu_j h_j^\circ(\boldsymbol{x}^*; \boldsymbol{x} - \boldsymbol{x}^*) + \sum_{j \in J_-} \nu_j h_j^\circ(\boldsymbol{x}^*; \boldsymbol{x}^* - \boldsymbol{x}) \\
\leq \quad & \sum_{i \in I(\boldsymbol{x}^*)} \mu_i \cdot 0 + \sum_{j \in J_+} \nu_j^+ \cdot 0 + \sum_{j \in J_-} \nu_j \cdot 0 = 0.
\end{aligned}$$

Thus,

$$0 \leq \sum_{k \in Q} \lambda_k \boldsymbol{\xi}_k^T (\boldsymbol{x} - \boldsymbol{x}^*) \leq \sum_{k \in Q} \lambda_k f_k^\circ(\boldsymbol{x}^*; \boldsymbol{x} - \boldsymbol{x}^*).$$

Since $\lambda_k \geq 0$ for all $k \in Q$ and $\sum_{k \in Q} \lambda_k = 1 > 0$ there exists $k_1 \in Q$ such that

$$0 \leq f_{k_1}^\circ(\boldsymbol{x}^*; \boldsymbol{x} - \boldsymbol{x}^*).$$

Then, f°-pseudoconvexity of f_{k_1} implies that $f_{k_1}(\boldsymbol{x}^*) \leq f_{k_1}(\boldsymbol{x})$. Since $\boldsymbol{x}$ is an arbitrary feasible point there exists no feasible point $\boldsymbol{y} \in \mathbb{R}^n$ such that $f_k(\boldsymbol{y}) < f_k(\boldsymbol{x}^*)$ for all $k \in Q$. Thus, $\boldsymbol{x}^*$ is a global weak Pareto optimum of problem (16). □

Note, that due to Theorems 3.9 and 3.11 the previous result is valid also for f°-pseudoconvex and subdifferentially regular quasiconvex inequality constraint functions. Also, the implicit assumption $I(\boldsymbol{x}^*) \neq \emptyset$ could be omitted by replacing $\operatorname{cone} G(x^*)$ by $\{0\} \cup \operatorname{cone} G(x^*)$.

Finally, by modifying somewhat the proof we get the sufficient KKT optimality condition for global Pareto optimality with an extra assumption for the multipliers.

COROLLARY 4.18. *The condition of Theorem 4.17 is also sufficient for $\boldsymbol{x}^*$ to be a global Pareto optimum of* (16)*, if in addition $\lambda_j > 0$ for all $k \in Q$.*

PROOF. By the proof of Theorem 4.17 we know that inequality

$$0 \leq \sum_{k \in Q} \lambda_k \boldsymbol{\xi}_k^T (\boldsymbol{x} - \boldsymbol{x}^*) \leq \sum_{k \in Q} \lambda_k f_k^\circ(\boldsymbol{x}^*; \boldsymbol{x} - \boldsymbol{x}^*) \tag{27}$$

holds for arbitrary feasible $\boldsymbol{x}$. Suppose there exists $k_1 \in Q$ such that $f_{k_1}^\circ(\boldsymbol{x}^*; \boldsymbol{x} - \boldsymbol{x}^*) < 0$. Because $\lambda_k > 0$ for all $k \in Q$, by inequality (27) there must be also $k_2 \in Q$ such that $f_{k_2}^\circ(\boldsymbol{x}^*; \boldsymbol{x} - \boldsymbol{x}^*) > 0$. By Theorem 3.9 f_{k_2} is f°-quasiconvex and by Definition 3.4 we have $f_{k_2}(\boldsymbol{x}) > f_{k_2}(\boldsymbol{x}^*)$. Since $\boldsymbol{x}$ were arbitrary, $\boldsymbol{x}^*$ is Pareto optimal.

Suppose then that $f_k^\circ(\boldsymbol{x}^*; \boldsymbol{x} - \boldsymbol{x}^*) \geq 0$ for all $k \in Q$. Then the f°-pseudoconvexity implies that $f_k(\boldsymbol{x}^*) \leq f_k(\boldsymbol{x})$ and, thus, $\boldsymbol{x}^*$ is Pareto optimal. □

As the next example shows a global minimum $\boldsymbol{x}^*$ does not necessarily satisfy the conditions in Theorem 4.17.

EXAMPLE 4.1. Consider the problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) = -x_1 \\ \text{subject to} & g(\boldsymbol{x}) = (x_1 - 2)^2 + (x_2 + 2)^2 - 2 \leq 0 \\ & h(\boldsymbol{x}) = (x_1 - 4)^2 + x_2^2 - 10 = 0. \end{array}$$

All the functions are convex and, thus, the assumptions of Theorem 4.17 are satisfied.

The global minimum to this problem is $\boldsymbol{x}^* = (3, -3)^T$. The gradients at this point are

$$\nabla f(\boldsymbol{x}^*) = (-1, 0)^T,\ \nabla g(\boldsymbol{x}^*) = (2, -2)^T \text{ and } \nabla h(\boldsymbol{x}^*) = (-2, -6)^T.$$

The gradients are illustrated in Figure 3. The lengths of the gradients in figure are scaled for clarity. The bolded curve represents the feasible set.

In Figure 4 the cone in condition (22) is illustrated by shaded region. From Figure 4 we see that $\mathbf{0} \notin \nabla f(\boldsymbol{x}^*) + \operatorname{cone} \nabla g(\boldsymbol{x}^*) + \operatorname{cone} \nabla h(\boldsymbol{x}^*)$. Thus we have a global optimum but the sufficient condition is not satisfied.

Let us then apply necessary conditions (Theorem 4.16) to the given example. It is easy to see that qualifications CQ5 and CQ6 are equivalent if functions h_j are differentiable for all $j \in J$. Clearly,

$$\begin{array}{rcl} T_S(\boldsymbol{x}^*) & = & \{\lambda(-3, 1) \mid \lambda \geq 0\}, \\ H^0(\boldsymbol{x}^*) & = & \{\lambda(-3, 1) \mid \lambda \in \mathbb{R}\} \text{ and} \\ G^-(\boldsymbol{x}^*) & = & \{(d_1, d_2) \mid d_1, d_2 \in \mathbb{R},\ d_1 \leq d_2\}. \end{array}$$

Thus, $G^-(\boldsymbol{x}^*) \cap H^0(\boldsymbol{x}^*) = T_S(\boldsymbol{x}^*)$ implying that CQ6 is satisfied. According to Theorem 4.16, relation (20) should hold at global minimum $\boldsymbol{x}^*$. Indeed,

$$\begin{array}{rcl} \mathbf{0} & = & \nabla f(\boldsymbol{x}^*) + \frac{3}{8} \nabla g(\boldsymbol{x}^*) + 0 \nabla h(\boldsymbol{x}^*) - \frac{1}{8} \nabla h(\boldsymbol{x}^*) \\ & \subset & \operatorname{conv} F(\boldsymbol{x}^*) + \operatorname{cl} \operatorname{cone} G(\boldsymbol{x}^*) + \operatorname{cl} \operatorname{cone} H(\boldsymbol{x}^*) - \operatorname{cl} \operatorname{cone} H(\boldsymbol{x}^*). \end{array}$$

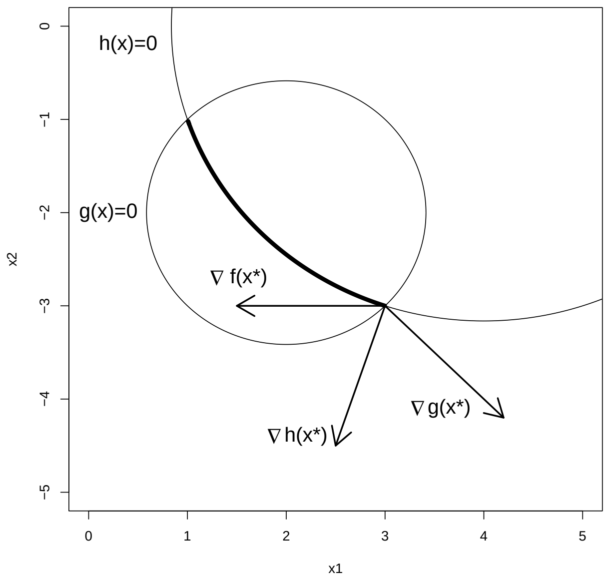


Figure 3: Gradients at the global minimum.

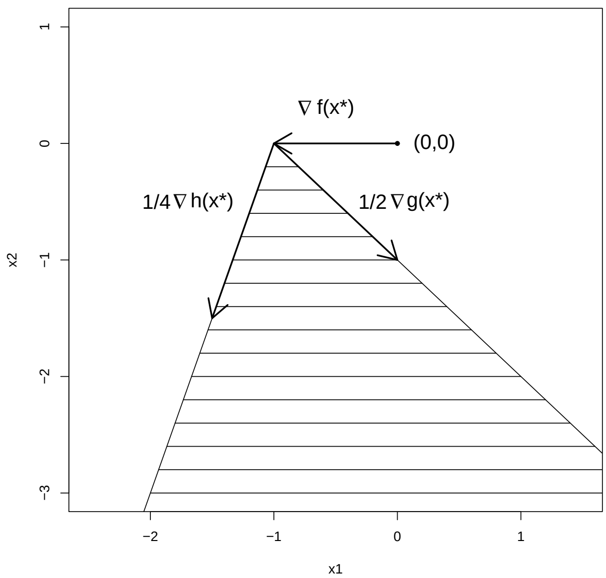


Figure 4: The set of sufficient KKT condition.

The relations in the necessary conditions are illustrated in Figure 5.

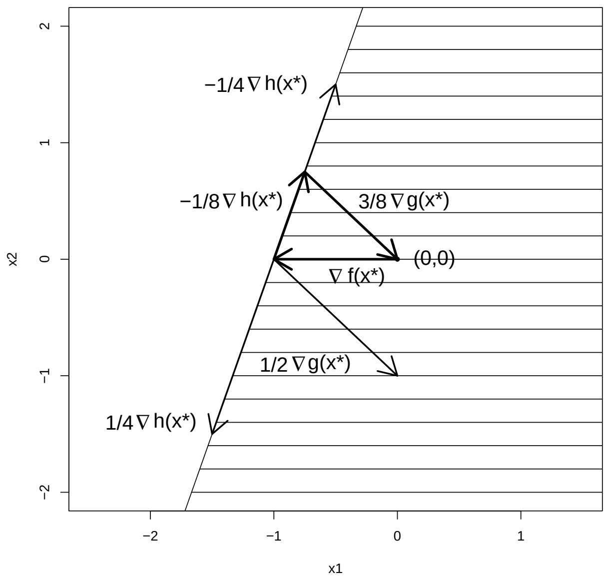


Figure 5: The gradients in the necessary KKT condition.

5 Concluding Remarks

We have considered KKT type necessary and sufficient conditions for nonsmooth multiobjective optimization problems. Both inequality and equality constraints were considered. The optimality were characterized as weak Pareto optimality. In necessary conditions CQ1–CQ6 constraint qualifications were needed. In sufficient conditions the main tools used were the generalized pseudo- and quasiconvexities based on the Clarke generalized directional derivative. It was assumed that the objective functions are f°-pseudoconvex and the constraint functions are f°-quasiconvex. Due to relations between different generalized convexities the results are valid also for f°-pseudoconvex and subdifferentially regular quasiconvex constraint functions.

References

[1] Avriel, M., Diewert, W. E., Schaible, S., and Zang, I. *Generalized Concavity*. Plenum Press, New York, 1988.

[2] Bazaraa, M., Sherali, H. D., and Shetty, C. M. *Nonlinear Programming Theory and Algorithms*. John Wiley and Sons, Inc., New York, 1979.

[3] BHATIA, D., AND AGGARWAL, S. Optimality and duality for multiobjective nonsmooth programming. *European Journal of Operation Research 57* (1992), 360–367.

[4] BHATIA, D., AND JAIN, P. Generalized (f,ρ)-convexity and duality for non smooth multi-objective programs. *Optimization 31* (1994), 153–164.

[5] BHATIA, D., AND MEHRA, A. Optimality conditions and duality involving arcwise connected and generalized arcwise connected functions. *Journal of Optimization Theory and Applications 100* (1999), 181–194.

[6] BRANDĀO, A. J. V., ROJAS-MEDAR, M. A., AND SILVA, G. N. Invex nonsmooth alternative theorem and applications. *Optimization 48* (2000), 239–253.

[7] CLARKE, F. H. *Optimization and Nonsmooth Analysis*. Wiley-Interscience, New York, 1983.

[8] DIEWERT, W. E. *Alternative Characterizations of Six Kinds of Quasiconcavity in the Nondifferentiable Case with Applications to Nonsmooth Programming*. "Generalized Concavity in Optimization and Economics" (Eds. Schaible, S. and Ziemba, W. T.), Academic Press, New York, pp. 51–95.

[9] HANSON, M. A. On sufficiency of the Kuhn-Tucker conditions. *Journal of Mathematical Analysis and Applications 80* (1981), 545–550.

[10] HIRIART-URRUTY, J. B. New concepts in nondifferentiable programming. *Bulletin de la Sociète Mathèmatiques de France, Mémoires 60* (1979), 57–85.

[11] HUU, S. P., MYUNG, L. G., AND SANG, K. D. Efficiency and generalized convexity in vector optimisation problems. *ANZIAM Journal 45* (2004), 523–546.

[12] JOURANI, A. Constraint qualifications and lagrange multipliers in nondifferentiable programming problems. *Journal of Optimization Theory and Applications 81*, 3 (1994), 533–548.

[13] KIM, D., AND LEE, H. Optimality conditions and duality in nonsmooth multiobjective programs. *Journal of Inequalities and Applications 2010* (2010). http://dx.doi.org/10.1155/2010/939537.

[14] KOMLÓSI, S. Generalized monotonicity and generalized convexity. *Journal of Optimization Theory and Applications 84* (1995), 361–376.

[15] LI, X. Constraint qualifications in nonsmooth multiobjective optimization. *Journal of Optimization Theory and Applications 106*, 2 (2000), 373–398.

[16] MÄKELÄ, M. M., KARMITSA, N., AND ERONEN, V.-P. On generalized pseudo- and quasiconvexities for nonsmooth functions. Tech. Rep. 989, TUCS Technical Report, Turku Centre for Computer Science, Turku, 2010.

[17] MÄKELÄ, M. M., AND NEITTAANMÄKI, P. *Nonsmooth Optimization: Analysis and Algorithms with Applications to Optimal Control*. World Scientific Publishing Co., Singapore, 1992.

[18] MANGASARIAN, O. L. Pseudoconvex functions. *SIAM Journal on Control 3* (1965), 281–290.

[19] MIETTINEN, K. *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, Boston, 1999.

[20] MIETTINEN, K., AND MÄKELÄ, M. M. On cone characterizations of weak, proper and pareto optimality in multiobjective optimization. *Mathematical Methods of Operations Research 53* (2001), 233–245.

[21] MISHRA, S. K. On sufficiency and duality for generalized quasiconvex nonsmooth programs. *Optimization 38* (1996), 223–235.

[22] NOBAKHTIAN, S. Infine functions and nonsmooth multiobjective optimization problems. *Computers and Mathematics with Applications 51* (2006), 1385–1394.

[23] NOBAKHTIAN, S. Multiobjective problems with nonsmooth equality constraints. *Numerical Functional Analysis and Optimization 30* (2009), 337–351.

[24] OSUNA-GÓMEZ, R., BEATO-MORENO, A., AND RUFIAN-LIZANA, A. Generalized convexity in multiobjective programming. *Journal of Mathematical Analysis and Applications 233* (1999), 205–220.

[25] PINI, R., AND SINGH, C. A survey of recent [1985–1995] advances in generalized convexity with applications to duality theory and optimality conditions. *Optimization 39* (1997), 311–360.

[26] PREDA, V. On efficiency and duality for multiobjective programs. *Journal of Mathematical Analysis and Applications 166* (1992), 365–377.

[27] ROCKAFELLAR, R. T. *Convex Analysis*. Princeton University Press, Princeton, New Jersey, 1970.

[28] SCHAIBLE, S. *Generalized Monotone Maps*. ”Nonsmooth Optimization: Methods and Applications” (Ed. Giannessi, F.), Gordon and Breach Science Publishers, Amsterdam, pp. 392–408.

[29] STAIB, T. Necessary optimality conditions for nonsmooth multicriteria optimization problem. *SIAM Journal on Optimization 2* (1992), 153–171.

[30] YANG, X. M., AND LIU, S. Y. Three kinds of generalized convexity. *Journal of Optimization Theory and Applications 86* (1995), 501–513.

A Relations between the CQ constraint qualifications

Consider problem (12), that is, problem

$$\begin{cases} \text{minimize} & \{f_1(\boldsymbol{x}), \ldots, f_q(\boldsymbol{x})\} \\ \text{subject to} & g_i(\boldsymbol{x}) \leq 0 \quad \text{for all } i \in M = \{1, \ldots, m\}. \end{cases} \tag{28}$$

Next, we will study some relations between the constraint qualifications. From now on, we assume that $I(\boldsymbol{x}) \neq \emptyset$.

In [15] it was shown that CQ1 follows from CQ3. Next we will prove that CQ1 follows also from CQ2.

THEOREM A.1. *Let $\boldsymbol{x} \in \mathbb{R}^n$ be a feasible point of problem* (28) *such that $I(\boldsymbol{x}) \neq \emptyset$. If $\mathbf{0} \notin \partial g(\boldsymbol{x})$ then $G^-(\boldsymbol{x}) \subset T_S(\boldsymbol{x})$.*

PROOF. Assume that there exists $\boldsymbol{d}^* \in G^-(\boldsymbol{x})$ such that $\boldsymbol{d}^* \notin T_S(\boldsymbol{x})$. Since a contingent cone is a closed set there exists $\varepsilon > 0$ such that $\operatorname{cl} B(\boldsymbol{d}^*; \varepsilon) \cap T_S(\boldsymbol{x}) = \emptyset$. Since $\boldsymbol{d} \notin T_S(x)$, for every $\boldsymbol{d} \in \operatorname{cl} B(\boldsymbol{d}^*; \varepsilon)$ there exists $t(\boldsymbol{d}) > 0$ such that $g(\boldsymbol{x} + t_1 \boldsymbol{d}) > g(\boldsymbol{x})$ when $0 < t_1 < t(\boldsymbol{d})$. Thus,

$$g^\circ(\boldsymbol{x}; \boldsymbol{d}) \geq 0, \text{ for all } \boldsymbol{d} \in \operatorname{cl} B(\boldsymbol{d}^*; \varepsilon). \tag{29}$$

Since $\boldsymbol{d}^* \in G^-(x)$ we have

$$\begin{aligned} g^\circ(\boldsymbol{x}; \boldsymbol{d}^*) &= \max\left\{\boldsymbol{\zeta}^T \boldsymbol{d}^* \mid \boldsymbol{\zeta} \in \partial g(\boldsymbol{x})\right\} \\ &\leq \max\left\{\boldsymbol{\zeta}^T \boldsymbol{d}^* \mid \boldsymbol{\zeta} \in \operatorname{conv}\{\partial g_i(\boldsymbol{x}) \mid i \in I(\boldsymbol{x})\}\right\} \\ &= \max\{g_i^\circ(\boldsymbol{x}, \boldsymbol{d}^*) \mid i \in I(\boldsymbol{x})\} \leq 0. \end{aligned} \tag{30}$$

Then for all $\boldsymbol{\zeta} \in \partial g(\boldsymbol{x})$ we have $\boldsymbol{\zeta}^T \boldsymbol{d}^* \leq 0$. Since we have $\mathbf{0} \notin \partial g(\boldsymbol{x})$ the Separation Theorem (see e.g. [2]) implies that there exist $\alpha \in \mathbb{R}$ and $\boldsymbol{z}$, $\|\boldsymbol{z}\| = 1$ such that

$$\boldsymbol{z}^T \mathbf{0} > \alpha \quad \text{and} \quad \boldsymbol{z}^T \boldsymbol{\zeta} \leq \alpha$$

for all $\boldsymbol{\zeta} \in \partial g(\boldsymbol{x})$. Since $\boldsymbol{z}^T \mathbf{0} = 0$ we see that $\boldsymbol{z}^T \boldsymbol{\zeta} < 0$ for all $\boldsymbol{\zeta} \in \partial g(\boldsymbol{x})$. If $\bar{\boldsymbol{d}} = \boldsymbol{d}^* + \varepsilon \boldsymbol{z}$, then $\bar{\boldsymbol{d}} \in \operatorname{cl} B(\boldsymbol{d}^*; \varepsilon)$ and

$$\boldsymbol{\zeta}^T \bar{\boldsymbol{d}} = \boldsymbol{\zeta}^T \boldsymbol{d}^* + \varepsilon \boldsymbol{\zeta}^T \boldsymbol{z} < 0$$

for all $\boldsymbol{\zeta} \in \partial g(\boldsymbol{x})$. Then

$$g^\circ(\boldsymbol{x}; \bar{\boldsymbol{d}}) = \max\left\{\boldsymbol{\zeta}^T \bar{\boldsymbol{d}} \mid \boldsymbol{\zeta} \in \partial g(\boldsymbol{x})\right\} < 0$$

contradicting inequality (29). Thus, $G^-(\boldsymbol{x}) \subset T_S(\boldsymbol{x})$. □

There exist problems that satisfy the CQ1 constraint qualification, but does not satisfy the CQ2.

EXAMPLE A.1. Consider the problem (28) with $g(x) = |x|$. Then we have $G^{-}(0) = \{0\}$ and $T_S(0) = \{0\}$. Thus, $G^{-}(0) \subset T_S(0)$ and CQ1 holds at $x = 0$. However, $0 \in \partial g(0)$ and CQ2 does not hold.

Next we will consider the relations between CQ2 and CQ3. First we will show that CQ2 follows from CQ3.

THEOREM A.2. *If $I(\boldsymbol{x}) \neq \emptyset$ and $G^s(\boldsymbol{x}) \neq \emptyset$, then $\mathbf{0} \notin \partial g(\boldsymbol{x})$.*

PROOF. It follows from the condition $G^s(\boldsymbol{x}) \neq \emptyset$ that there exists $\boldsymbol{d}$, such that $g_i^\circ(\boldsymbol{x}; \boldsymbol{d}) < 0$ for all $i \in I(\boldsymbol{x})$. In other words, $\boldsymbol{d}^T \boldsymbol{\xi}_i < 0$ for all $\boldsymbol{\xi}_i \in \partial g_i(\boldsymbol{x})$ and $i \in I(\boldsymbol{x})$. Let $\lambda_i \geq 0$, $i \in I(\boldsymbol{x})$ be scalars such that $\sum_{i \in I(\boldsymbol{x})} \lambda_i = 1$. Then

$$\boldsymbol{d}^T \sum_{i \in I(\boldsymbol{x})} \lambda_i \boldsymbol{\xi}_i = \sum_{i \in I(\boldsymbol{x})} \lambda_i \boldsymbol{d}^T \boldsymbol{\xi}_i < 0.$$

Thus, $\boldsymbol{d}^T \boldsymbol{\xi} < 0$ for all $\boldsymbol{\xi} \in \operatorname{conv} \bigcup_{i \in I(\boldsymbol{x})} \partial g_i(\boldsymbol{x})$. Since $\partial g(\boldsymbol{x}) \subset \operatorname{conv} \bigcup_{i \in I(\boldsymbol{x})} \partial g_i(\boldsymbol{x})$, we have $g^\circ(\boldsymbol{x}; \boldsymbol{d}) < 0$ implying that $\mathbf{0} \notin \partial g(\boldsymbol{x})$. □

There exist problems for which CQ2 holds but CQ3 does not as the following example shows.

EXAMPLE A.2. Consider constraint functions

$$g_1(x) = x \quad \text{and} \quad g_2(x) = \begin{cases} x & , \text{ if } x < 0 \\ 0 & , \text{ if } x \geq 0. \end{cases}$$

Then $g(x) = \max\{g_1(x), g_2(x)\} = g_1(x)$ and $0 \notin \partial g(0)$. However, $0 \in \partial g_2(0)$ which implies $G^s(0) = \emptyset$.

Despite Example A.2 we can establish some conditions on constraint functions which guarantees that CQ2 implies CQ3. Namely, if all the constraint functions are subdifferentially regular or f°-pseudoconvex the CQ3 follows from CQ2.

THEOREM A.3. *Let $\boldsymbol{x} \in \mathbb{R}^n$ and $I(\boldsymbol{x}) \neq \emptyset$. If the functions g_i are subdifferentially regular for all $i \in M$ and $\mathbf{0} \notin \partial g(\boldsymbol{x})$, then $G^s(\boldsymbol{x}) \neq \emptyset$.*

PROOF. If $\mathbf{0} \notin \partial g(\boldsymbol{x})$, then there exists $\boldsymbol{d}$, such that $g^\circ(\boldsymbol{x}; \boldsymbol{d}) < 0$. Due to regularity we have $\partial g(\boldsymbol{x}) = \operatorname{conv} \bigcup_{i \in I(\boldsymbol{x})} \partial g_i(\boldsymbol{x})$. Hence,

$$\boldsymbol{d}^T \sum_{i \in I(\boldsymbol{x})} \lambda_i \boldsymbol{\xi}_i < 0, \text{ for all } \boldsymbol{\xi}_i \in \partial g_i(\boldsymbol{x}),\ \lambda_i \geq 0,\ \sum_{i \in I(\boldsymbol{x})} \lambda_i = 1,$$

implying $\boldsymbol{d}^T \boldsymbol{\xi}_i < 0$ for all $\boldsymbol{\xi}_i \in \partial g_i(\boldsymbol{x})$. In other words $g_i^\circ(\boldsymbol{x}; \boldsymbol{d}) < 0$ for all $i \in I(\boldsymbol{x})$. Thus, we have $\boldsymbol{d} \in G^s \neq \emptyset$. □

THEOREM A.4. *Let $\boldsymbol{x} \in \mathbb{R}^n$ and $I(\boldsymbol{x}) \neq \emptyset$. If the functions g_i are f°-pseudoconvex for all $i \in M$ and $\mathbf{0} \notin \partial g(\boldsymbol{x})$, then $G^s(\boldsymbol{x}) \neq \emptyset$.*

PROOF. On contrary, assume that $G^s = \emptyset$. Then for all $\boldsymbol{d} \in \mathbb{R}^n$ there exists $i \in I(\boldsymbol{x})$, for which $g_i^\circ(\boldsymbol{x}; \boldsymbol{d}) \geq 0$. Due to f°-pseudoconvexity we have $g_i(\boldsymbol{x} + t\boldsymbol{d}) \geq g_i(\boldsymbol{x})$ for all $t \geq 0$. Since $g(\boldsymbol{x}) \geq g_i(\boldsymbol{x})$ for all $i \in M$ we have $g(\boldsymbol{x} + t\boldsymbol{d}) \geq g(\boldsymbol{x})$ for all $\boldsymbol{d} \in \mathbb{R}^n$. Thus, $\boldsymbol{x}$ is a global minimum and $\mathbf{0} \in g(\boldsymbol{x})$ by Theorem 2.12. In other words, if $\mathbf{0} \notin g(\boldsymbol{x})$ we will have $G^s \neq \emptyset$. □

Finally, we will show that constraint qualification CQ3 is equivalent to CQ4.

THEOREM A.5. *Suppose* $I(\boldsymbol{x}) \neq \emptyset$. *Then* $\mathbf{0} \notin \operatorname{conv} G(\boldsymbol{x})$ *iff* $G^s(\boldsymbol{x}) \neq \emptyset$.

PROOF. The condition $\mathbf{0} \notin \operatorname{conv} G(\boldsymbol{x})$ is equivalent to condition $\operatorname{conv} G(\boldsymbol{x}) \cap \{\mathbf{0}\} = \emptyset$. By Corollary 2.11 $\operatorname{conv} G(\boldsymbol{x})$ is a closed convex set and trivially $\{\mathbf{0}\}$ is a closed convex cone. Also, $\{\mathbf{0}\}^- = \mathbb{R}^n = -\{\mathbf{0}\}^-$. By Lemma 4.8 $\operatorname{conv} G(\boldsymbol{x}) \cap \{\mathbf{0}\} = \emptyset$ is equivalent to

$$(\operatorname{conv} G(\boldsymbol{x}))^s \cap \mathbb{R}^n = (\operatorname{conv} G(\boldsymbol{x}))^s \neq \emptyset.$$

Furthermore, $(\operatorname{conv} G(\boldsymbol{x}))^s = G^s(\boldsymbol{x})$ according to Lemma 4.9. □

Turku Centre *for* Computer Science

Lemminkäisenkatu 14 A, 20520 Turku, Finland | www.tucs.fi

University of Turku

- Department of Information Technology
- Department of Mathematics

Åbo Akademi University

- Department of Computer Science
- Institute for Advanced Management Systems Research

Turku School of Economics and Business Administration

- Institute of Information Systems Sciences

ISBN 978-952-12-2770-7
ISSN 1239-1891