OPTIMALITY CONDITIONS AND CHARACTERIZATIONS OF THE OPTIMAL SOLUTION SETS FOR UNCERTAIN CONVEX OPTIMIZATION PROBLEMS

NITHIRAT SISARAT

A Thesis Submitted to the Graduate School of Naresuan University
in Partial Fulfillment of the Requirements
for the Doctor of Philosophy Degree in Mathematics
May 2019

This thesis entitled "Optimality conditions and characterizations of the optimal solution sets for uncertain convex optimization problems"

by Nithirat Sisarat

has been approved by the Graduate School as partial fulfillment of the requirements for the Doctor of Philosophy Degree in Mathematics of Naresuan University

Oral Defense Committee

S. Suantai .... Chair

(Professor Suthep Suantai, Ph.D.)

Rabian Wangkeeree .... Advisor

(Associate Professor Rabian Wangkeeree, Ph.D.)

Somyot Plubtieng .... Internal Examiner

(Professor Somyot Plubtieng, Ph.D.)

Narin Petrot .... Internal Examiner

(Associate Professor Narin Petrot, Ph.D.)

P. Cholamjiak .... External Examiner

(Associate Professor Prasit Cholamjiak, Ph.D.)

Approved

....

(Professor Paisarn Muneesawang, Ph.D.)

Dean of the Graduate School

29 MAY 2019

Finally, I gratefully appreciate my beloved family for their moral support and encouragement throughout my study and research period.

Nithirat Sisarat

Title	OPTIMALITY CONDITIONS AND CHARACTERIZATIONS OF THE OPTIMAL SOLUTION SETS FOR UNCERTAIN CONVEX OPTIMIZATION PROBLEMS
Author	Nithirat Sisarat
Advisor	Associate Professor Rabian Wangkeeree, Ph.D.
Academic Paper	Thesis Ph.D. in Mathematics, Naresuan University, 2018.
Keywords	Lagrange multipliers, Karush-Kuhn-Tucker conditions, constraint qualifications, solution sets, convex optimization problems, subdifferentials, nonconvex constraints, pseudo Lagrangian functions, tangentially convex functions, uncertain convex optimization, robust solutions, multi-objective optimization, cone-convex functions, level set.

ABSTRACT

The primary aim of this thesis is to investigate the characterization of robust solution sets for uncertain convex optimization problems without convexity of constraint data uncertainty. In order to do this we study first such problems in the absence of data uncertainty and establish the weakest constraint qualification for guaranteeing the Lagrange multiplier conditions to be necessary and sufficient for optimality. After introducing the so-called pseudo-Lagrangian function, we then establish the constant pseudo-Lagrange property and employ it to derive a characterization of the solution set, which is expressed in terms of convex subdifferentials, tangential subdifferentials and Lagrange multipliers. Afterwards, with slightly consideration, characterizations of the robust optimal solution set for uncertain convex optimization problems with a robust convex constraint set described by locally Lipschitz constraints are derived. Beside, by means of linear scalarization, characterizations of weakly robust efficient solution set and properly robust efficient solution set of uncertain convex multi-objective optimization problems are also shown. Finally, we conclude our investigations by making an analysis to weaken the differentiability and the convexity assumptions considered in non-convex optimization problems and multi-objective optimization problems with cone constraints, respectively.

LIST OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF FIGURES (CONT.)

CHAPTER I

INTRODUCTION

Constrained optimization problems concern the minimization or maximization of functions over some set of conditions called *constraints*, and play a vital role in many fields of science as diverse as economics, accounting, computer science, engineering and others to select an optimal solution in many decisions based on computational methods. The original treatment of constrained optimization problems was to deal only with equality constraints via the introduction of the *Lagrange multiplier method* which consists of transforming a constrained problem into an unconstrained one. Later on, it was realized that constraints in the form of inequalities play a predominant role in modeling real world, and, therefore; it leads to more challenging necessary conditions for global optimality of an optimization problem in terms of a system of inequalities. Mathematically, let us consider the following optimization problem:

$$\min_{x \in \mathbb{R}^n} \{f(x) : g_i(x) \leq 0, \ i = 1, 2, \ldots, m\}, \tag{P}$$

where $f, g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$, are given functions. Johns paper [1] first brought the Lagrange multiplier rule to the domain of inequality constraint. Another key contribution in this way is due to Kuhn and Tucker [2]. They amended the defect of the Lagrange multiplier rule for convex and other nonlinear optimization problems involving differentiable functions as laid down by John. That is in order to guarantee the existence of Lagrangian multipliers, one is in need of imposing certain conditions on the constraints, known as constraint qualifications. Later, it was discovered that similar results were presented by Karush [3] back in 1939. These lead to the now famous *Karush-Kuhn-Tucker (shortly, KKT) conditions* which play a very fundamental role both in the theory of optimization and the analysis of optimization algorithms for constrained problems.

Admittedly, convexity is a very important hypothesis and is a natural framework for studying in mathematical optimization problems due to the fact that when a convex function is minimized over a convex feasible set which is commonly represented by convex inequality constraints, Lagrange multiplier condition becomes sufficient for global optimality.

Taken from another viewpoint, characterizations and properties of the optimal solution sets of optimization problems are of great interest, as they play an important role for understanding the behavior of solution methods for mathematical programs that have multiple optimal solutions. Mangasarian [4] initially presented several simple and elegant characterizations of the optimal solution sets of differentiable convex programs provided that a fixed minimizer is known and applied them to study monotone linear complementarity problems in [5]. This study was further extended to nonsmooth cases by Burke and Ferris [6]. Subsequently, such results were applied to characterize the problems that have weak sharp minimum [7] and to analyze properties of proximal bundle methods for finding global solutions of convex optimization problems [8]. Since then, by employing Lagrange multipliers and its properties, many interesting results on characterizations of the optimal solution sets of various classes of optimization problems have been obtained, see; cone-constrained convex programming problems in [9, 10], pseudo-linear programs in [11–14], pseudoconvex minimization problems in [15–17], variational inequality problems in [18, 19], quasi-convex programming problems in [20–22], nonconvex optimization problems with pseudo-invex functions in [23–25], semi-infinite optimization problems in [26–28] and cone constrained convex vector optimization problems in [29].

As we know, the majority of many practical constrained optimization problems often involve input data that are noisy or uncertain due to modeling, estimation errors, prediction errors as well as measurement errors [30–33]. Therefore, it is imperative to study the optimization problems with data uncertainty. Precisely stated, the problem (P) in the face of constraint data uncertainty can be captured by the following optimization problem:

$$\min_{x\in\mathbb{R}^n}\{f(x,u) : g_i(x,v_i) \leq 0,\ i = 1, 2, \dots, m\}, \tag{UP}$$

where $f : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}$ and $g_i : \mathbb{R}^n \times \mathcal{V}_i \to \mathbb{R}$, $i = 1, 2, \dots, m$, are given functions, and u and v_i are the uncertain parameters that are not exactly known, but are only known to reside in certain uncertainty sets $\mathcal{U} \subseteq \mathbb{R}^{q_0}$ and $\mathcal{V}_i \subseteq \mathbb{R}^{q_i}$, respectively. In addition, in many situations often we need to make decisions now before we can know the true values or have better estimations of the parameters, for instance, optimization problems arising in industry or commerce might involve various costs, financial returns, and future demands that might be unknown at the time of the decision. If the uncertainties are ignored while solving the optimization problem, it may lead to solutions which are suboptimal or even infeasible. As an illustration, we now consider the following uncertain linear program:

$$\min_{(x_1,x_2)\in\mathbb{R}^2}\{2x_1+3x_2 : -x_1-2x_2+1\leq 0,\ -3x_1-2x_2+2\leq 0\}, \qquad \text{(ULP)}$$

when (assume that) the datas in constraints are only estimates and can be inaccurate, i.e.,

$$\begin{bmatrix}-1 & -2 & 1\\ -3 & -2 & 2\end{bmatrix}+\begin{bmatrix}\pm 0.05 & \pm 0.05 & \pm 0.05\\ \pm 0.05 & \pm 0.05 & \pm 0.05\end{bmatrix}.$$

The optimal solution of the problem that have constraints $-x_1-2x_2+1\leq 0$ and $-3x_1-2x_2+2\leq 0$, is $(x_1^*,x_2^*):=(0.5,0.25)$ with an objective function value of 1.75. Unfortunately, a matrix of actual values that realize is

$$\begin{bmatrix}-1.05 & -2.05 & 1.05\\ -3 & -2 & 2,\end{bmatrix}$$

and therefore, the optimal solution (x_1^*,x_2^*) is no longer feasible for this realization.

Consequently, how to explicate mathematical approaches that are capable of treating data uncertainty in constrained optimization has become a critical question in mathematical optimization. As we have seen the problematic situations where a decision based on a model has to be taken here and now, we need naturally to the additional requirement that any feasible vectors must satisfy all constraints including each set of constraints corresponding to a possible realization of the uncertain parameters from the set uncertainty set. In this way, we refer to the so-called *robust (worst case) counterpart* and therefore its robust counterpart of (ULP) reads as follows:

$$\min_{(x_1,x_2)\in\mathbb{R}^2}\{2x_1+3x_2 : a_{11}x_1+a_{12}x_2+b_1\leq 0,\ a_{21}x_1+a_{22}x_2+b_2\leq 0$$
$$\forall\begin{bmatrix}a_{11} & a_{12} & b_1\end{bmatrix}\in\mathcal{U}_1,\ \begin{bmatrix}a_{21} & a_{22} & b_2\end{bmatrix}\in\mathcal{U}_2\}, \qquad \text{(RLP)}$$

where

$$\mathcal{U}_1:=\left\{\begin{bmatrix}-0.95 & -1.95 & 0.95\end{bmatrix},\begin{bmatrix}-1 & -2 & 1\end{bmatrix},\begin{bmatrix}-1.05 & -2.05 & 1.05\end{bmatrix}\right\},$$

$$\mathcal{U}_2 := \left\{ \begin{bmatrix} -2.95 & -1.95 & 1.95 \end{bmatrix}, \begin{bmatrix} -3 & -2 & 2 \end{bmatrix}, \begin{bmatrix} -3.05 & -2.05 & 2.05 \end{bmatrix} \right\}.$$

Robust optimization, which is its robust counterpart of an uncertain optimization problem, has emerged as a powerful deterministic approach for studying optimization problems with data uncertainty in the sense that it minimizes the objective function value in the worst case of all scenarios and gets a solution that works well even in the worst-case scenario, but also is immunized against the data uncertainty. For instance, we can see that the optimal solution of (RLP) is $(0.5000, 0.2560)$ with a corresponding objective function value of 1.7683. So, the solution has the advantage of satisfying all of the constraints without increasing the objective function too much, that is, the solution is robust or immune to uncertainty. In general, the robust counterpart of the problem (UP) which, by parametric reformulation of (UP) (see, [30]), is given by

$$\min_{(x,\alpha)\in\mathbb{R}^{n+1}} \{\alpha : f(x,u) - \alpha \leq 0,\ g_i(x,v_i) \leq 0,\ \forall u \in \mathcal{U},\ v_i \in \mathcal{V}_i,\ i = 1, 2, \ldots, m\},$$

or equivalently,

$$\min_{x\in\mathbb{R}^n} \{\sup_{u\in\mathcal{U}} f(x,u) : g_i(x,v_i) \leq 0,\ v_i \in \mathcal{V}_i,\ i = 1, 2, \ldots, m\}, \tag{RP}$$

where the uncertain constraint are enforced for every possible value of the parameters within their prescribed uncertainty and the global minimizer of the problem (RP) is known as robust optimal solution of the problem (UP). Over the years, a great deal of attention has been attracted to treat uncertain optimization problems by using robust optimization methodology. For issues related to optimality conditions and duality properties, see [34–40] and other references therein. Here, we specially mention the works on characterizations of robust optimal solution sets for optimization problems with data uncertainty due to [41–45]. More precisely, by using a robust Slater-type condition, some properties and characterizations of robust optimal solution sets of an uncertain convex optimization problem are obtained in [41]. Then, the main results of [41] are investigated in [42] under a new robust FarkasMinkowski constraint qualification which is strictly weaker than the robust Slater-type condition, and generalized to uncertain general programming problems; pseudo-convex robust optimization problem with constraint involving data uncertainty in [43], cone-constrained convex optimization problem with data uncertainty in both the objective and constraint functions in [44] and an uncertain fractional optimization problem in [45]. Beside, in almost all existing literature on robust convex optimization, the convexity assumption on the functions $g_i(\cdot, v_i)$, $i = 1, 2, \dots, m$, for all $v_i \in \mathcal{V}_i$, is principle and restrictive. In fact, even if $g_i(\cdot, v_i)$, $i = 1, 2, \dots, m$, are not convex for all $v_i \in \mathcal{V}_i$, it may happen that the so-called *robust feasible set* $\{x \in \mathbb{R}^n : \; x \in C, \; g_i(x, v_i) \leq 0, \; \forall v_i \in \mathcal{V}_i, \; i = 1, 2, \dots, m\}$ is convex. For a simple example, let $\mathcal{V} := [0, 1]$ and $g(x, v) := x - vx^3$ for all $x \in \mathbb{R}$ and $v \in \mathcal{V}$. It can be observed that $\{x \in \mathbb{R} : g(x, v) \leq 0, \; \forall v \in \mathcal{V}\} = [-1, 0]$, which is convex, while $g(\cdot, v)$ are not convex for each $v \in \mathcal{V}$. So, the aforementioned papers may go awry. How we do in this case? In order to motivate our study, it will be meaningful to consider the convex constraint set without uncertainties does not admit a *convex representation* in the sense that the constraint functions to represent the convex constraint set are non-necessarily convex.

Convex optimization problems without convex representation was discussed by Lasserre [46] in 2010 where the involving functions are differentiable. It covers a broad class of nonlinear programming problems, including the classical convex programming problems as well as convex minimization problems with quasi-convex constraint functions due to the quasi-convexity of constraint functions ensures that the constraint set is a convex set. Further study has been done about optimality conditions for some classes of smooth/non-smooth constrained optimization problems [47–50]. As far as we know, characterizations of the optimal solution set for convex optimization problem without convexity of constraints in the absence of data uncertainty have not also been studied yet. So, at first, we will investigate the characterization of the optimal solution sets of the following convex optimization problem:

$$\min_{x \in \mathbb{R}^n} \{f(x) : x \in C\}, \tag{CP}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function, the constraint set C, defined by

$$C := \{x \in \mathbb{R}^n : g_i(x) \leq 0, \; i \in I\}, \tag{1.0.1}$$

is a nonempty convex subset of the Euclidean space $\mathbb{R}^n$ and $g_i : \mathbb{R}^n \to \mathbb{R}$, $i \in I := \{1, 2, \dots, m\}$, are continuous functions, but they are *not assumed to be convex functions*. It is remarkable that the characterization of solution sets of (CP) is done by applying [27, Corollary 3.10.] if the functions $g_i : \mathbb{R}^n \to \mathbb{R}$, $i \in I$, are restricted to be locally Lipschitz and regular in the sense of Clarke [51] and additionally the pseudoconvexity in the first argument of the Lagrange function,

$$\mathcal{L}(\cdot, \lambda) := f(\cdot) + \sum_{i \in I} \lambda_i g_i(\cdot),\ \forall \lambda := (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{R}^m_+,$$

is satisfied. However, the pseudoconvexity assumption of $\mathcal{L}(\cdot, \lambda)$ for every $\lambda \in \mathbb{R}^m_+$ often fails (see Remark 3.1.23). Further, regularity requirements of g_i's may fail even if g_i's are differentiable functions due to the fact that differentiable functions are not necessarily regular unless they are continuously differentiable.

All in all, the optimality conditions, which as we shall see, are at the root of the development for mathematical optimization problems in many aspects. It is also worth maintaining here that optimality conditions have a relationship with the representation of the constraint set. Recently, Ho [52] went further in the case of scalar differentiable problems but moreover without the convexity of the constraint set and of the functions that are involved, and showed that necessary and sufficient KKT optimality conditions are then considered in relation to the presence of convexity of the level sets of objective function. In view of [47] and [53], the main results in [52] also suggest a way and motivate us to investigate KKT optimality conditions in (1) non-smooth optimization problems with inequality constraints without the presence of convexity of objective function, of constraint functions and of feasible set, and (2) differentiable multi-objective optimization problems over cone constraints without the convexity of the feasible set, and the cone-convexity of objectives and constraint functions.

Motivated and inspired by all above contributions, in this thesis, our aim is to perform study of theoretical side of optimization problems related to optimality conditions and characterizations of the optimal solution sets for convex optimization problems as well as robust convex optimization problems without convexity of constraints. More than that, in the absence of data uncertainty and convexity of constraint sets, we are going to consider (scalar) non-smooth optimization problems and differentiable multi-objective optimization problems over cone constraints.

In the following, we give a description of how is this thesis organized.

Chapter II. We will include several notions and preliminary results in order to make this thesis as self-contained as possible.

Chapter III. We draw our attention to the investigation of optimality conditions and characterizations of the optimal solution sets for convex optimization without convex representation. The chapter contains two different parts, the first one devoted to give characterizations of the solution set of (CP) without the pseudoconvexity assumption of Lagrange function. In order to make use of the obtained results for both the differentiable setting and the regular locally Lipschitz setting, we deal with the problem (CP) with continuous tangentially convex constraint functions (see [48]). First, we give the weakest constraint qualification for guaranteeing the Lagrange multiplier conditions to be necessary and sufficient for optimality of (CP). After introducing the so-called *pseudo-Lagrange function*, we then establish the constant pseudo-Lagrange property and employ it to derive a characterization of the solution set of (CP). These are expressed in terms of convex subdifferentials, tangential subdifferentials and Lagrange multipliers. Moreover, Lagrange multiplier characterizations of the solution set for optimization problems with a pseudoconvex locally Lipschitz objective function, without convexity of the constraint functions are given. With a slight consideration of the first part, the second one devotes to examine a robust optimization framework for studying characterizations of the robust optimal solution set for uncertain convex optimization problems with a robust convex feasible set described by locally Lipschitz constraints. In addition, the results are then applied to derive characterizations of weakly robust efficient solution set and properly robust efficient solution set of uncertain convex multi-objective optimization problems without convexity assumption on constraint functions.

Chapter IV. Along the line of Ho [52], the first part of this chapter is to the study of non-smooth optimization problems with inequality constraints without the presence of convexity of objective function, of constraint functions and of feasible set. We present necessary and sufficient KKT optimality conditions for these problems in terms of tangential subdifferentials. Our results contain and improve some recent ones in the literature. Many examples are also given to explain the advantages of our main results. In the second part of the chapter we deal with a class of differentiable multi-objective optimization problems (MOP) over cone constraints without the convexity of the feasible set, and the cone-convexity of objectives and constraint functions. We present relationships among constraint qualifications of multi-objective optimization problem (MOP) over cone constraint and establish necessary and sufficient KKT optimality conditions for a feasible point under the question to be a weak Pareto minimum of (MOP). We finally give sufficient conditions for guaranteeing a weak Pareto minimum to be a Pareto minimum of the problem (MOP).

Chapter V. We give the concluding remarks.

CHAPTER II

PRELIMINARIES

In this chapter, we will review the certain notations, basic definitions, and preliminary results that are related to our research. Throughout this thesis, all spaces under consideration are the n-dimensional Euclidean space $\mathbb{R}^n$. All vectors are considered to be column vectors which can be transposed to a row vector by the superscript T. For vectors $x := (x_1, x_2, \ldots, x_n)$ and $y := (y_1, y_2, \ldots, y_n)$ in $\mathbb{R}^n$, the (usual) inner product of x and y is denoted by $\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$, while the norm of x is given by $\|x\| = \sqrt{\langle x, x \rangle}$. The closed (resp. open, left closed right open) interval between $\alpha, \beta \in \mathbb{R}$ with $\alpha < \beta$ is denoted by $[\alpha, \beta]$ (resp. $]\alpha, \beta[$, $[\alpha, \beta[$). The non-negative orthant of $\mathbb{R}^n$ is denoted by $\mathbb{R}^n_+$ and is defined by $\mathbb{R}^n_+ := \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_i \geq 0, \ i = 1, 2, \ldots, n\}$. For any two sets $A_1, A_2 \subseteq \mathbb{R}^n$, define $A_1 + A_2 := \{a_1 + a_2 \in \mathbb{R}^n : a_1 \in A_1, \ a_2 \in A_2\}$. For any set $A \subseteq \mathbb{R}^n$ and any scalar $\alpha \in \mathbb{R}$, $\alpha A := \{\alpha a \in \mathbb{R}^n : a \in A\}$.

2.1 Basic concepts

Definition 2.1.1. A sequence $\{x_l \in \mathbb{R} : l = 1, 2, \ldots\}$ or simply $\{x_l\} \subset \mathbb{R}$ is said to **converge** if there exists $x \in \mathbb{R}$ such that for every $\varepsilon > 0$, we have

$$|x_l - x| < \varepsilon, \ \forall l \geq l_\varepsilon,$$

for some integer l_ε (that depend on ε). The scalar x is said to be the **limit** of $\{x_l\}$, and the sequence $\{x_l\}$ is said to **converge to** x. Symbolically, it is expressed as

$$x_l \to x \text{ or } \lim_{l \to +\infty} x_l = x.$$

Definition 2.1.2. A scalar sequence $\{x_l\}$ is said to be **bounded above** (resp. **below**) if there exists some scalar α such that $x_l \leq \alpha$ (resp. $x_l \geq \alpha$) for all $l \in \mathbb{N}$. It is said to be **bounded** if it is bounded above and bounded below. The sequence $\{x_l\}$ is said to be **monotonically nonincreasing** (resp. **nondecreasing**) if $x_{l+1} \leq x_l$ (resp. $x_{l+1} \geq x_l$) for all $l \in \mathbb{N}$. If $x_l \to x$ and $\{x_l\}$ is monotonically nonincreasing (resp. nondecreasing), we also use the notation $x_l \downarrow x$ (resp. $x_l \uparrow x$).

Given a sequence $\{x_l\} \subset \mathbb{R}$, let $y_r := \sup\{x_l : l \geq r\}$, $z_r := \inf\{x_l : l \geq r\}$. Note that the sequences $\{y_r\}$ and $\{z_r\}$ are nonincreasing and nondecreasing, respectively, and therefore have a limit whenever $\{x_l\}$ is bounded above or is bounded below, respectively.

Definition 2.1.3. The limit of $\{y_r\}$ is denoted by $\limsup_{l\to+\infty} x_l$, and is referred to as the **upper limit** of $\{x_l\}$. The limit of $\{z_r\}$ is denoted by $\liminf_{l\to+\infty} x_l$, and is referred to as the **lower limit** of $\{x_l\}$.

Definition 2.1.4. A sequence $\{x_l\}$ of vectors in $\mathbb{R}^n$ is said to **converge** to some $x \in \mathbb{R}^n$ if the i-th component of x_l converges to the i-th component of x for every $i = 1, 2, \ldots, n$. We use the notations $x_l \to x$ or $\lim_{l\to+\infty} x_l = x$ to indicate convergence for vector sequences as well.

Definition 2.1.5. The sequence $\{x_l\} \subset \mathbb{R}^n$ is bounded if there exists $M > 0$ such that $\|x_l\| \leq M$ for every $l \in \mathbb{N}$. A **subsequence** of $\{x_l\} \subset \mathbb{R}^n$ is a sequence $\{x_{l_j}\}$, $j = 1, 2, \ldots$, where each x_{l_j} is a member of the original sequence and the order of the elements as in the original sequence is maintained.

By $\mathbb{B}(x, r)$ we means an **open ball** of radius $r > 0$ with center at x, i.e., $\mathbb{B}(x, r) := \{y \in \mathbb{R}^n : \|y - x\| < r\}$.

Definition 2.1.6. We say that x is a **closure point** of a subset A of $\mathbb{R}^n$ if there exists a sequence $\{x_l\} \subset A$ that converges to x. The **closure** of A, denoted $\mathrm{cl}(A)$, is the set of all closure points of A. We also say that x is an interior point of a subset A of $\mathbb{R}^n$ if there exists $r > 0$ such that $\mathbb{B}(x, r) \subseteq A$. The **interior** of A, denoted $\mathrm{int}(A)$, is the set of all interior points of A.

Definition 2.1.7. A subset A of $\mathbb{R}^n$ is called

(i) **closed** if $A = \mathrm{cl}(A)$.

(ii) **open** if its complement, $\mathbb{R}^n \backslash A$, is closed, or equivalently, $A = \mathrm{int}(A)$.

(iii) **bounded** if there exists a scalar M such that $\|x\| \leq M$ for all $x \in A$.

(iv) **compact** if it is closed and bounded.

Wholesale, we prefer to deal with functions that are real-valued and are defined over $\mathbb{R}^n$. However, in some situations, prominently arising in the context of optimization, we will encounter operations on real-valued functions that produce *extended real-valued functions*, that is, functions that take values in $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$. As an example, a function of the form

$$f(x) = \sup_{i \in I} f_i(x),$$

where I is an infinite index set, can take the value ∞ even if the functions f_i are real-valued. Most rules with infinity are intuitively clear except possibly $0 \times (+\infty)$ and $\infty - \infty$. Because we will be dealing mainly with minimization problems, we will follow the convention

$$0 \times (+\infty) = (+\infty) \times 0 = 0 \text{ and } \infty - \infty = \infty.$$

Definition 2.1.8. An extended real-valued function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be a **proper** function if $f(x) > -\infty$ for every $x \in \mathbb{R}^n$ and the **domain** of f,

$$\text{dom } f := \{x \in \mathbb{R}^n : f(x) < +\infty\},$$

is nonempty.

Now we move on to define the semicontinuities of a real-valued function that involve the limit infimum and limit supremum of the function.

Definition 2.1.9. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be **lower semicontinuous (lsc)** at $\bar{x} \in \mathbb{R}^n$ if for every sequence $\{x_l\} \subset \mathbb{R}^n$ converging to $\bar{x}$,

$$f(\bar{x}) \leq \liminf_{l \to +\infty} f(x_l).$$

Equivalently,

$$f(\bar{x}) \leq \liminf_{x \to \bar{x}} f(x) := \lim_{\delta \downarrow 0} \inf_{x \in \mathbb{B}(\bar{x}, \delta)} f(x).$$

The function f is lsc over a set $A \subseteq \mathbb{R}^n$ if f is lsc at every $\bar{x} \in A$.

In tandem with the concept of lower semicontinuity and limit infimum, we next define the upper semicontinuity and the limit supremum of a function.

Definition 2.1.10. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be **upper semicontinuous (usc)** at $\bar{x} \in \mathbb{R}^n$ if for every sequence $\{x_l\} \subset \mathbb{R}^n$ converging to $\bar{x}$,

$$f(\bar{x}) \geq \limsup_{l\to+\infty} f(x_l).$$

Equivalently,

$$f(\bar{x}) \geq \limsup_{x\to\bar{x}} f(x) := \lim_{\delta\downarrow 0} \sup_{x\in\mathbb{B}(\bar{x},\delta)} f(x).$$

The function f is usc over a set $A \subseteq \mathbb{R}^n$ if f is usc at every $\bar{x} \in A$.

Definition 2.1.11. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be **continuous** at $\bar{x}$ if it is lsc as well as usc at $\bar{x}$, that is,

$$\lim_{x\to\bar{x}} f(x) = f(\bar{x}).$$

The next result, a generalization of the classical theorem of Weierstrass, suggests a way that whether an optimal solution exists.

Theorem 2.1.12. [54, p. 86] *Let A be a nonempty closed subset of $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be lsc over A. Assume that one of the following conditions holds:*

(i) *A is bounded.*

(ii) *Some level set $\{x \in \mathbb{R}^n : f(x) \leq \alpha\}$ is nonempty and bounded.*

(iii) *For every sequence $\{x_l\} \subset A$ such that $\|x_l\| \to +\infty$, we have $f(x_l) \to \infty$ as $l \to +\infty$.*

Then, the set of ***minima*** *(or* ***minimizer****) of f over A, i.e., $\{x \in A : f(x) \leq f(y),\ \forall y \in A\}$, is nonempty and compact.*

Next we go through the notions of differentiability for real-valued functions as well as vector-valued functions. We begin with by recalling the partial derivative.

Definition 2.1.13. For $f : \mathbb{R}^n \to \mathbb{R}$ be some function, fix some $x \in \mathbb{R}^n$, and consider the expression

$$\lim_{t\to 0} \frac{f(x+te_i) - f(x)}{t},$$

where e_i is the i-th unit vector (all components are 0 except for the i-th component which is 1). If the above limit exists, it is called the i-th **partial derivative** of f at the vector x and it is denoted by $\frac{\partial f(x)}{\partial x_i}$.

Definition 2.1.14. Let f be a function from $\mathbb{R}^n$ to $\overline{\mathbb{R}}$, and let x be a point where f is finite. We say that f is **(Fréchet) differentiable** at x if and only if there exists a vector ξ (necessarily unique) with the property that

$$\lim_{y \to x} \frac{f(y) - f(x) - \langle \xi, y - x \rangle}{\|y - x\|} = 0.$$

Definition 2.1.15. For a proper function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, we define the **one-sided directional derivative** of f at $\bar{x} \in \operatorname{dom} f$ in the direction $d \in \mathbb{R}^n$ to be

$$f'(\bar{x}; d) := \lim_{t \downarrow 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t},$$

provided that $+\infty$ and $-\infty$ are allowed as limits.

Remark 2.1.16. If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is differentiable at $\bar{x}$ with corresponding vector ξ and $f(\bar{x}) \in \mathbb{R}$, we have

$$f'(\bar{x}; d) = \langle \xi, d \rangle, \ \forall d \in \mathbb{R}^n.$$

Moreover,

$$\xi = \nabla f(x) := \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n} \right).$$

Proof. Suppose that f is differentiable at $\bar{x}$. It then follows from the definition that there exists $\xi \in \mathbb{R}^n$ such that for any $d \neq 0$,

$$\begin{aligned} 0 &= \lim_{t \downarrow 0} \frac{f(\bar{x} + td) - f(\bar{x}) - \langle \xi, td \rangle}{t\|d\|} \\ &= \frac{f'(\bar{x}; d) - \langle \xi, d \rangle}{\|d\|}. \end{aligned}$$

Therefore, $f'(\bar{x}, d)$ exists and is a linear function of d:

$$f'(\bar{x}; d) = \langle \xi, d \rangle, \ \forall d \in \mathbb{R}^n.$$

In particular, for $i = 1, 2, \dots, n$,

$$\langle \xi, e_i \rangle = \lim_{t \to 0} \frac{f(\bar{x} + te_i) - f(x)}{t} = \frac{\partial f(x)}{\partial x_i},$$

which in turn implies that $\xi = \nabla f(x)$. ☐

As can be seen, such a ξ, if it exists, without loss of generality, is called the **gradient** of f at x and is denoted by $\nabla f(x)$. In addition, if f is **continuously differentiable**, that is, ∇f is continuous over $\mathbb{R}^n$, then f is called a **smooth function**. If f is not smooth, it is referred to as being **nonsmooth**.

Example 2.1.17. [Differentiable but nonsmooth function] Consider the following real-valued function

$$f(x) = \begin{cases} 0, & \text{if } x = 0; \\ x^2 \sin(\frac{1}{x}), & \text{ohterwise.} \end{cases}$$

It is clear that the function f is differentiable at $x \neq 0$ and its derivative is

$$f'(x) = 2x \sin(1/x) - \cos(1/x), \ \forall x \neq 0.$$

Since $f(0 + x) - f(0) = x^2 \sin(\frac{1}{x})$ for any $x \neq 0$ and $\lim_{x \to 0} x \sin(1/x) = 0$, the function f is differentiable at $x = 0$ and $f'(0) = 0$. Therefore, f is differentiable on $\mathbb{R}$. However, f is not continuously differentiable because the limit $\lim_{x \to 0} f'(x)$ does not exist.

Definition 2.1.18. A vector-valued function $f : \mathbb{R}^n \to \mathbb{R}^p$ is called **differentiable** (or **smooth**) if each component f_j of f, $j = 1, 2, \ldots, p$, is differentiable (or smooth, respectivelty). The **Jacobian** of f, denoted $\nabla f(x)$, is the $p \times n$ matrix and can be expressed as

$$\nabla f(x) := (\nabla f_1(x), \nabla f_2(x), \ldots, \nabla f_p(x))^T.$$

Let $f : \mathbb{R}^n \to \mathbb{R}^p$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable vector-valued functions, and let h be their composition, i.e.,

$$h(x) := g(f(x)), \ \forall x \in \mathbb{R}^n.$$

Then, the **chain rule** for differentiation [54, p. 19] states that

$$\nabla h(x) = \nabla f(x)^T \nabla g(f(x)), \ \forall x \in \mathbb{R}^n. \tag{2.1.1}$$

Definition 2.1.19. A **set-valued map** Φ from $\mathbb{R}^n$ to $\mathbb{R}^m$ associates every $x \in \mathbb{R}^n$ to a set in $\mathbb{R}^m$; that is, for every $x \in \mathbb{R}^n$, $\Phi(x) \subseteq \mathbb{R}^m$. Symbolically, it is expressed as $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$.

A set-valued map $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to be **upper semicontinuous (usc)** at $x \in \mathbb{R}^n$ if for any sequences $\{\xi_l\}$ and $\{x_l\}$ tending to ξ and x respectively, and if $\xi_l \in \Phi(x_l)$ for each $l \in \mathbb{N}$, then $\xi \in \Phi(x)$.

2.2 Convex and nonsmooth analysis

We start this section by recalling the notions of a convex set and an affine set.

Definition 2.2.1. A set $A \subseteq \mathbb{R}^n$ is said to be

(i) **convex** if $\alpha a_1 + (1-\alpha)a_2 \in A,\ \forall a_1,\ a_2 \in A,\ \forall \alpha \in [0,1]$.

(ii) **affine** if $\alpha a_1 + (1-\alpha)a_2 \in A,\ \forall a_1,\ a_2 \in A,\ \forall \alpha \in \mathbb{R}$.

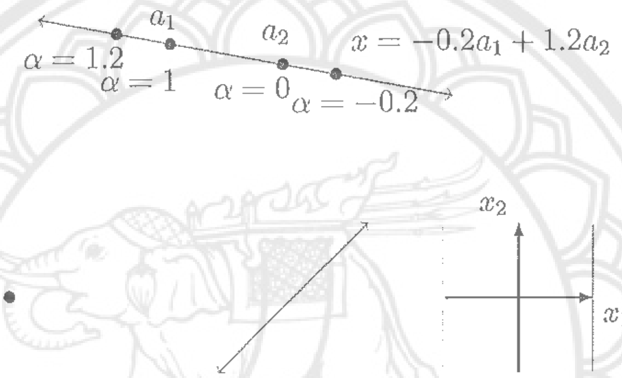


Figure 1: Example of affine sets on a two-dimensional space.

Next we state some important properties of convex sets.

Proposition 2.2.2 (Operations on convex sets).

(i) *The intersection of an arbitrary collection of convex sets is convex.*

(ii) *For a convex set $A \subseteq \mathbb{R}^n$ and scalar $\alpha_1 \geq 0$ and $\alpha_2 \geq 0$, $(\alpha_1 + \alpha_2)A = \alpha_1 A_1 + \alpha_2 A_2$ which is convex.*

Theorem 2.2.3. [55, Theorem 3.16] *Let A and B be nonempty convex subsets of $\mathbb{R}^n$ with $\operatorname{int} A \neq \emptyset$. Then $\operatorname{int} A \cap B = \emptyset$ if and only if there exist a vector $\xi \in \mathbb{R}^n \backslash \{0\}$ and a real number α with*

$$\langle \xi, a \rangle \leq \alpha \leq \langle \xi, b \rangle \text{ for all } a \in A \text{ and all } b \in B$$

and

$$\langle \xi, a \rangle < \alpha \text{ for all } a \in \operatorname{int} A.$$

Definition 2.2.4. The **convex hull** of a set $A \subseteq \mathbb{R}^n$, is denoted by co A, is the smallest convex set containing A and can be expressed as

$$\text{co } A = \left\{ x \in \mathbb{R}^n : x = \sum_{i=1}^{m} \alpha_i a_i,\ a_i \in A,\ \alpha_i \geq 0,\ i = 1, 2, \ldots, m,\ \sum_{i=1}^{m} \alpha_i = 1 \right\},$$

for some $m \in \mathbb{N}$.

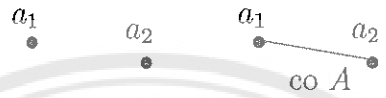


Figure 2: Illustration of a convex hull.

Another important property of a convex set is the *line segment principle.* To arrive there, we now need the notions of the affine hull and the relative interior of convex sets.

Definition 2.2.5. The **affine hull** of a set $A \subseteq \mathbb{R}^n$, is denoted by aff A, is the smallest affine set containing A and can be expressed as

$$\text{aff } A = \left\{ x \in \mathbb{R}^n : x = \sum_{i=1}^{m} \alpha_i a_i,\ a_i \in A,\ \alpha_i \in \mathbb{R},\ i = 1, 2, \ldots, m,\ \sum_{i=1}^{m} \alpha_i = 1 \right\},$$

for some $m \in \mathbb{N}$.

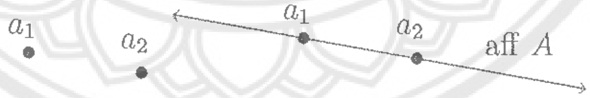


Figure 3: Illustration of an affine hull.

Definition 2.2.6. [56, p. 44] The **relative interior** of a convex set $A \subseteq \mathbb{R}^n$, ri A, is the interior of A relative to the affine hull of A, that is,

$$\text{ri } A := \{x \in A : \exists \varepsilon > 0 \text{ s.t. } \mathbb{B}(x, \varepsilon) \cap \text{aff } A \subseteq A\}.$$

For an n-dimensional convex set $A \subseteq \mathbb{R}^n$, i.e., the dimension of a subspace which parallel to aff A, aff $A = \mathbb{R}^n$ and thus ri A = int A.

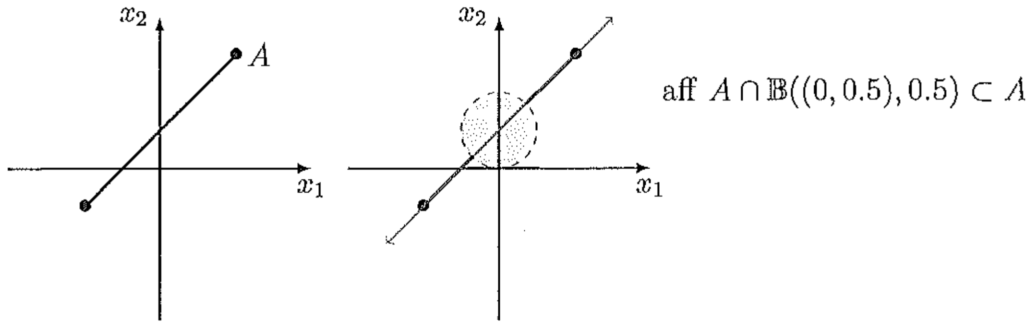


Figure 4: Illustration of a relative interior of a one-dimensional convex set on a two-dimensional space.

Proposition 2.2.7. [56, p. 45] *Consider a nonempty convex set $A \subseteq \mathbb{R}^n$. Then, the following assertions hold:*

(i) ri A *is nonempty.*

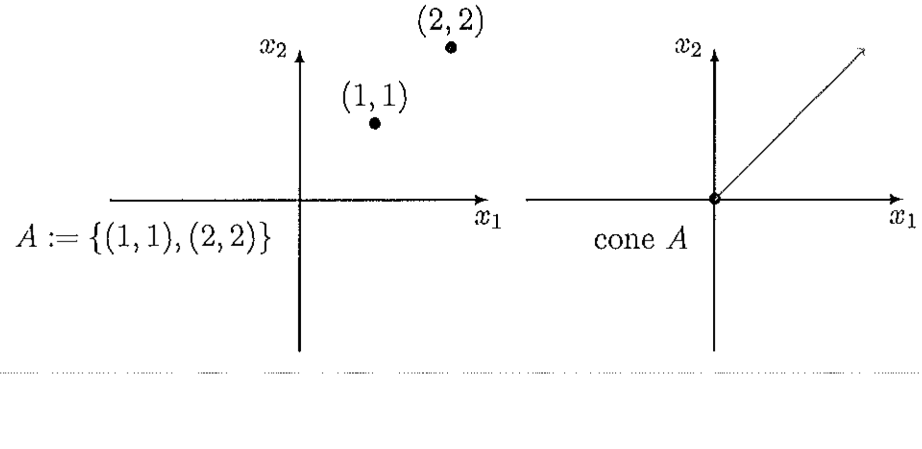
(ii) [56, Theorem 6.1, Line Segment Principle] *Let $x \in$ ri A and $y \in$ cl A. Then for $\alpha \in [0, 1[$,*

$$(1 - \alpha)x + \alpha y \in \text{ri } A.$$

On the one hand, a special class of convex sets is a convex cone. So, we shall begin by recalling the concept of cones.

Definition 2.2.8. A set $K \subseteq \mathbb{R}^n$ is said to be a **cone** if for every $x \in K$, $\alpha x \in K$ for every $\alpha \geq 0$. For any set $A \subseteq \mathbb{R}^n$, the **cone generated** by A is denoted by cone A and is defined as

$$\text{cone } A := \bigcup_{\alpha \geq 0} \alpha A = \{x \in \mathbb{R}^n : x = \alpha a,\ a \in A,\ \alpha \geq 0\}.$$



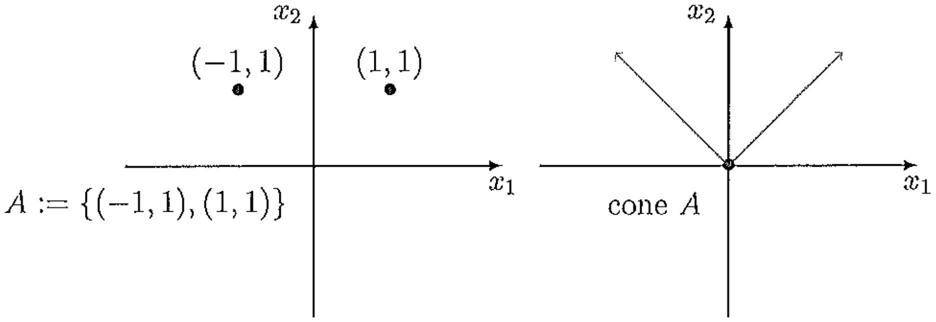


Figure 5: Illustration of the cone generated by the nonconvex set A.

From now on, we moving on to the notion of the convex cones.

Definition 2.2.9. The set $K \subseteq \mathbb{R}^n$ is said to be **convex cone** if it is convex as well as a cone. For any set $A \subseteq \mathbb{R}^n$, the **convex cone generated by** A is denoted by cone co A and is expressed as

$$\text{cone co } A = \left\{ x \in \mathbb{R}^n : x = \sum_{i=1}^{m} \alpha_i a_i,\ a_i \in A,\ \alpha_i \geq 0,\ i = 1, 2, \ldots, m,\ m \in \mathbb{N} \right\}.$$

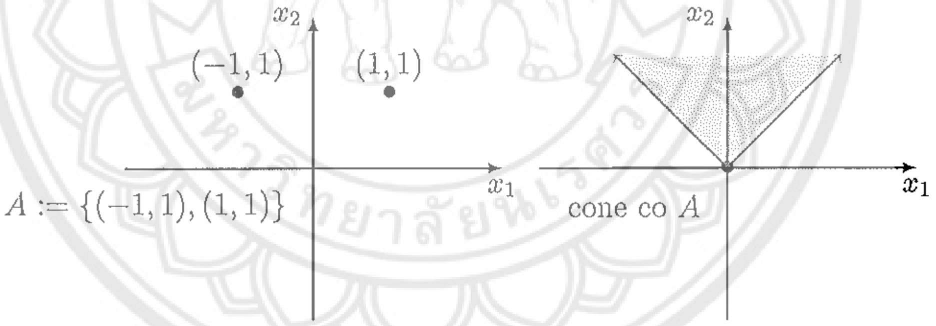


Figure 6: Illustration of a convex cone generated by a nonconvex set A.

In addition, for a collection of convex sets $A_i \subseteq \mathbb{R}^n$, $i = 1, 2, \ldots, m$, the convex cone generated by A_i, $i = 1, 2, \ldots, m$, can be shown to be expressed as

$$\text{cone co} \bigcup_{i=1}^{m} A_i = \bigcup_{\substack{\alpha_i \geq 0 \\ i=1,2,\ldots,m}} \sum_{i=1}^{m} \alpha_i A_i. \tag{2.2.1}$$

See, [57, p. 97], for more details.

The following theorem suggests a way that whether a given cone is convex.

Theorem 2.2.10. *A cone $K \subseteq \mathbb{R}^n$ is convex if and only if $K + K \subseteq K$.*

Definition 2.2.11. Consider a set $A \subseteq \mathbb{R}^n$. The cone defined as

$$A^\circ := \{\xi \in \mathbb{R}^n : \langle \xi, x \rangle \leq 0, \ \forall x \in A\}$$

is called the **polar cone** of A. Note that the polar cone of the set A is a closed convex cone.

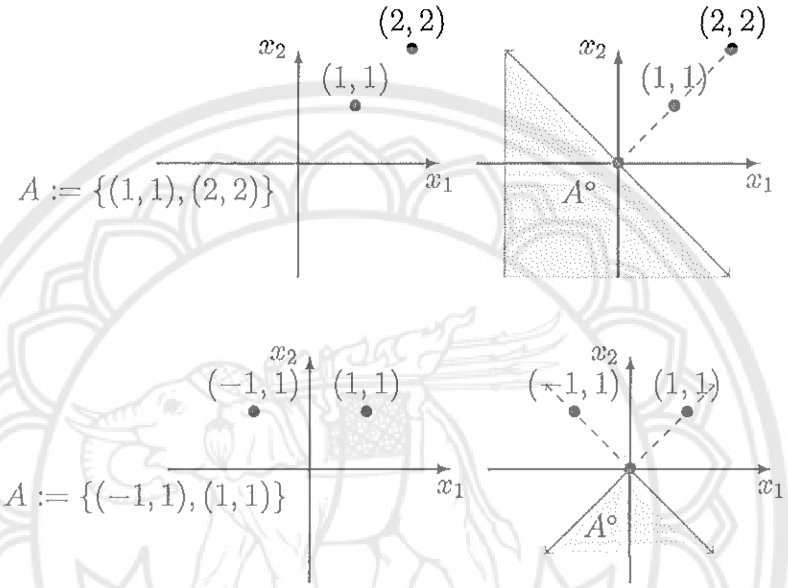


Figure 7: Polar cones of the sets.

Now we present some properties of polar and bipolar cones.

Proposition 2.2.12. [54, Proposition 3.2.1(a)] *Let $a_1, a_2, \ldots, a_m$ be vectors in $\mathbb{R}^n$. Then, the finitely generated cone*

$$A := \text{cone}\{a_1, a_2, \ldots, a_m\}$$

is closed and its polar cone is the polyhedral cone given by

$$A^\circ = \{d \in \mathbb{R}^n : \langle a_i, d \rangle \leq 0, \ i = 1, 2, \ldots, m\}.$$

Proposition 2.2.13.

(i) *Consider two sets $A_1, \ A_2 \subseteq \mathbb{R}^n$ such that $A_1 \subseteq A_2$. Then $A_2^\circ \subseteq A_1^\circ$.*

(ii) [The bipolar cone theorem] *Consider a nonempty set $A \subseteq \mathbb{R}^n$. Then*

$$A^{\circ\circ} = \text{cl cone co } A.$$

Definition 2.2.14. Consider a set $A \subseteq \mathbb{R}^n$. The **positive polar cone (dual cone)** to the set A is defined as

$$A^* := \{\xi \in \mathbb{R}^n : \langle \xi, x \rangle \geq 0, \ \forall x \in A\}.$$

Observe that $A^* = (-A)^\circ = -A^\circ$.

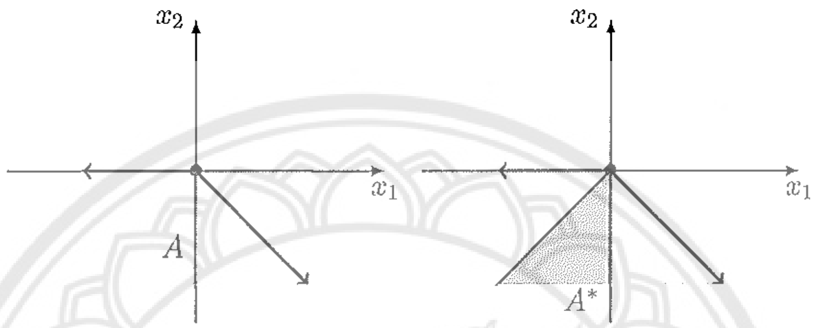


Figure 8: Dual cone of the set.

Lemma 2.2.15. [55, Lemma 3.21, p. 77] *Let K be a convex cone in $\mathbb{R}^p$.*

(i) *If K is closed, then $K = \{x \in \mathbb{R}^p : \langle \xi, x \rangle \geq 0$ for all $\xi \in K^*\}$.*

(ii) *If $\text{int}K \neq \emptyset$, then $\text{int}K = \{x \in \mathbb{R}^p : \langle \xi, x \rangle > 0$ for all $\xi \in K^* \backslash \{0\}\}$.*

Definition 2.2.16. Consider a set $A \subseteq \mathbb{R}^n$ and $\bar{x} \in A$. The (Bouligand) **tangent cone** to the set A at $\bar{x}$, $T(A, \bar{x})$, is defined by

$$T(A, \bar{x}) := \left\{ d \in \mathbb{R}^n : \exists \{x_l\} \subset A, \ x_l \to \bar{x}, \ t_l \downarrow 0 \text{ s.t. } \frac{1}{t_l}(x_l - \bar{x}) \to d \text{ as } l \to +\infty \right\}.$$

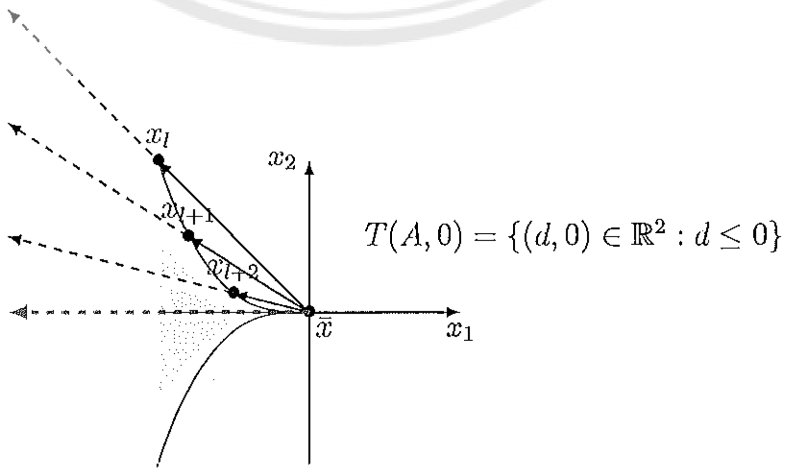


Figure 9: Illustration of the behavior of the vector in a tangent cone.

In view of the definition, to construct a tangent cone we consider all the sequences $\{x_l\}$ in A the converge to the given point $\bar{x} \in A$, and then calculate all the directions $d \in \mathbb{R}^n$ that are tangential to the sequences at $\bar{x}$. However, if A is a convex set, the tangent cone can be obtained by the following way.

Theorem 2.2.17. *Consider a set $A \subseteq \mathbb{R}^n$ and $\bar{x} \in A$. Then the following hold:*

(i) *$T(A, \bar{x})$ is closed.*

(ii) *If A is convex, $T(A, \bar{x}) = \text{cl cone}(A - \bar{x})$ and hence $T(A, \bar{x})$ is convex.*

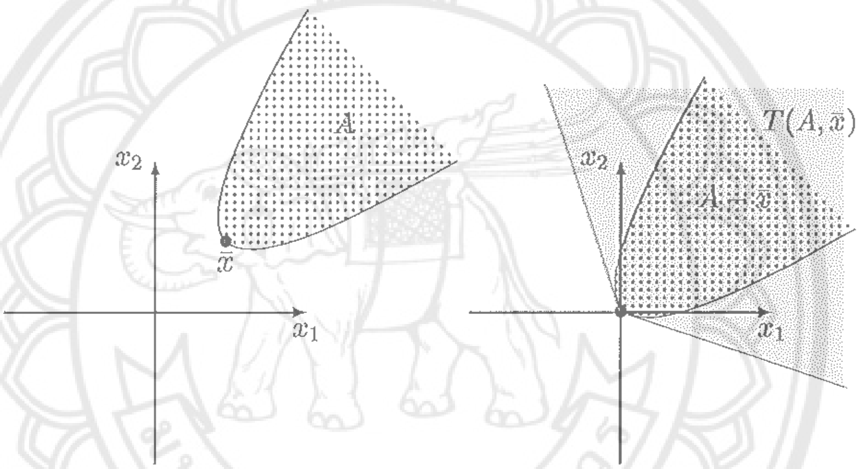


Figure 10: Illustration of a tangent cone of a convex set.

Definition 2.2.18. Consider a convex set $A \subseteq \mathbb{R}^n$ and $\bar{x} \in A$. The **normal cone** to the set A at $\bar{x}$, $N(A, \bar{x})$, is given by

$$N(A, \bar{x}) := \{d \in \mathbb{R}^n : \langle d, x - \bar{x} \rangle \leq 0, \ \forall x \in A\}.$$

By employing the polarity, we have the following connection between the contingent cone and the normal cone.

Proposition 2.2.19. *Consider a convex set $A \subseteq \mathbb{R}^n$. Then,*

$$N(A, \bar{x}) = (T(A, \bar{x}))^\circ \text{ and } T(A, \bar{x}) = (N(A, \bar{x}))^\circ.$$

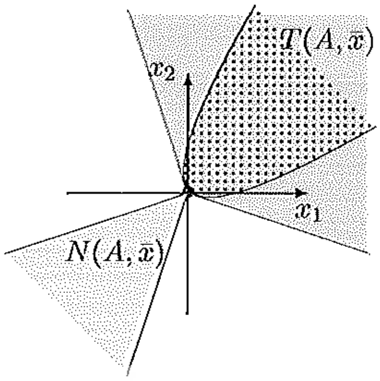


Figure 11: Tangent cone and normal cone of a convex set.

With all these background on convex sets, we now move on to consider the convexity of a function.

Definition 2.2.20. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be **convex** if for any $x, y \in \mathbb{R}^n$ and $\alpha \in [0, 1]$ we have

$$f((1-\alpha)x + \alpha y) \leq (1-\alpha)f(x) + \alpha f(y).$$

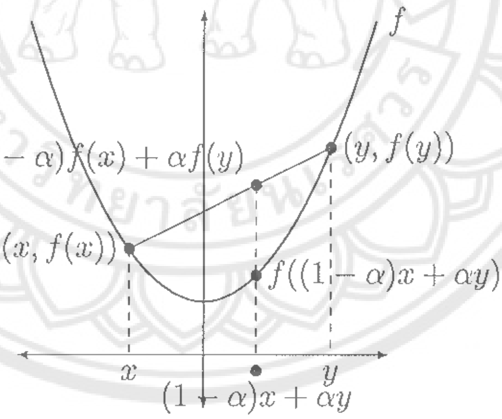


Figure 12: Geometric interpretation of convex functions.

Example 2.2.21. Consider a set $A \subseteq \mathbb{R}^n$. The **indicator function**, $\delta_A : \mathbb{R}^n \to \overline{\mathbb{R}}$, to the set A is defined as

$$\delta_A(x) := \begin{cases} 0, & \text{if } x \in A; \\ +\infty, & \text{ohterwise.} \end{cases}$$

It can be easily shown that δ_A is lsc and convex if and only if A is closed and convex, respectively.

Proposition 2.2.22 (Operations that preserve convexity).

(i) *Consider proper convex functions* $f_i : \mathbb{R}^n \to \overline{\mathbb{R}}$ *and* $\alpha_i \geq 0$, $i = 1, 2, \ldots, m$. *Then* $f := \sum_{i=1}^{m} \alpha_i f_i$ *is also a convex function.*

(ii) *Consider a family of proper convex functions* $f_i : \mathbb{R}^n \to \overline{\mathbb{R}}$, $i \in \mathcal{I}$, *where* $\mathcal{I}$ *is an arbitrary index set. Then* $f := \sup_{i \in \mathcal{I}} f_i$ *is a convex function.*

Theorem 2.2.23. [56, Theorem 10.1] *A proper convex function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *is continuous on* ri dom f.

A special class of convex functions is a sublinear function.

Definition 2.2.24. A proper function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be a **sublinear function** if f is subadditive and positively homogeneous, that is,

$$f(x_1 + x_2) \leq f(x_1) + f(x_2),\ \forall x_1, x_2 \in \mathbb{R}^n \text{ (subadditive property)}$$
$$f(\alpha x) = \alpha f(x),\ \forall x \in \mathbb{R}^n,\ \forall \alpha > 0 \text{ (positively homogeneous property).}$$

Lemma 2.2.25. [48, Lemma 8] *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be a sublinear function. If* f *vanishes on an open set, then it is nonnegative everywhere.*

We now consider some properties related to the directional derivative of convex functions.

Theorem 2.2.26. [56, Theorem 23.1] *Consider a proper convex function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *and* $\bar{x} \in$ dom f. *Then for every* $d \in \mathbb{R}^n$, *the directional derivative* $f'(\bar{x}; d)$ *exists. Moreover,* $f'(\bar{x}; d)$ *is a sublinear function in* d *for every* $d \in \mathbb{R}^n$.

Next we recall the subgradient and the subdifferential of a convex function.

Definition 2.2.27. [56, p. 214] Consider a proper convex function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\bar{x} \in$ dom f. Then $\xi \in \mathbb{R}^n$ is said to be the **subgradient** of the function f at $\bar{x}$ if

$$f(x) - f(\bar{x}) \geq \langle \xi, x - \bar{x} \rangle,\ \forall x \in \mathbb{R}^n.$$

The collection of all such vectors constitute the **subdifferential** of f at $\bar{x}$ and is denoted by $\partial f(\bar{x})$. For $\bar{x} \notin$ dom f, $\partial f(\bar{x})$ is empty.

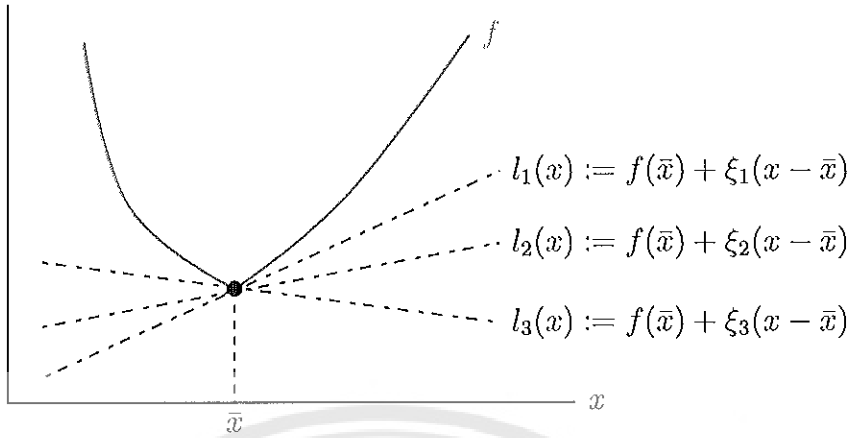


Figure 13: Geometric interpretation of subdifferentials.

Remark 2.2.28. [56, p. 215] For a convex set A, $\partial\delta_A(\cdot) = N(A, \cdot)$.

Several elementary facts about the subdifferential will now be listed.

Theorem 2.2.29. [56, Theorem 23.2] *Consider a proper convex function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *and* $\bar{x} \in \operatorname{dom} f$. *Then*

$$\partial f(\bar{x}) = \{\xi \in \mathbb{R}^n : f'(\bar{x}; d) \geq \langle \xi, d \rangle, \ \forall d \in \mathbb{R}^n\}.$$

Proposition 2.2.30. [56, Theorem 25.1] *Consider a convex function* $f : \mathbb{R}^n \to \mathbb{R}$ *differentiable at* $\bar{x}$ *with gradient* $\nabla f(\bar{x})$. *Then,* $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$.

Proposition 2.2.31. [56, Theorem 23.4] *Consider a proper convex function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *and* $\bar{x} \in \operatorname{dom} f$. *Then* $\partial f(\bar{x})$ *is closed and convex. For* $\bar{x} \in \operatorname{ri} \operatorname{dom} f$, $\partial f(\bar{x}) \neq \emptyset$. *Furthermore, if* $\bar{x} \in \operatorname{int} \operatorname{dom} f$, $\partial f(\bar{x})$ *is nonempty and compact.*

Theorem 2.2.32. [56, Theorem 23.8, Moreau-Rockafellar Sum Rule] *Consider two proper convex function* $f_1, f_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$. *Suppose that* $\operatorname{ri} \operatorname{dom} f_1 \cap \operatorname{ri} \operatorname{dom} f_2 \neq \emptyset$. *Then*

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x)$$

for every $x \in \operatorname{dom}(f_1 + f_2)$.

Lemma 2.2.33. [41, Lemma 2.1] *Let* $\mathcal{U} \subseteq \mathbb{R}^p$ *be a convex compact set, and let* $f : \mathbb{R}^n \times \mathbb{R}^{q_0} \to \mathbb{R}$ *be a function such that for each fixed* $u \in \mathcal{U}$, $f(\cdot, u)$ *is a convex* *function on* $\mathbb{R}^n$ *and for each fixed* $x \in \mathbb{R}^n$, $f(x, \cdot)$ *is a concave function on* $\mathbb{R}^{q_0}$. *Then,*

$$\partial \left(\max_{u \in \mathcal{U}} f(\cdot, u) \right) (\bar{x}) = \bigcup_{u \in \mathcal{U}(\bar{x})} \partial f(\cdot, u)(\bar{x}),$$

where $\mathcal{U}(\bar{x}) := \{\bar{u} \in \mathcal{U} : f(\bar{x}, \bar{u}) = \max_{u \in \mathcal{U}} f(\bar{x}, u)\}$.

As we also deal with a class of nonconvex functions instead of convex functions for constraint functions, we shall need generalized subdifferential to nonconvex function. In what follows we firstly refer to the notion of local Lipschitz continuity and the basic properties.

Definition 2.2.34. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be **locally Lipschitz** at $x \in \mathbb{R}^n$, if there exist an open neighborhood U and a constant L such that, for all y and z in U, one has

$$|f(y) - f(z)| \leq L\|y - z\|.$$

If the function f is locally Lipschitz at every point $x \in \mathbb{R}^n$, one says that f is a locally Lipschitz function on $\mathbb{R}^n$.

Proposition 2.2.35. [51, Corollary, p. 32] *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be continuously differentiable at* x, *then* f *is a locally Lipschitz at* x.

Theorem 2.2.36. *Consider a proper convex function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. *Then* f *is locally Lipschitz at* x *for every* $x \in \text{ri dom } f$.

We are now ready to state a generalization of the ordinary directional derivative which always exists for locally Lipschitz continuous functions.

Definition 2.2.37. [51, p. 25] Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz at a given point $x \in \mathbb{R}^n$. The **Clarke generalized directional derivative** of f at x in the direction $d \in \mathbb{R}^n$, denoted $f^o(x; d)$, is defined as

$$f^o(x; d) := \limsup_{\substack{y \to x \\ t \to 0^+}} \frac{f(y + td) - f(y)}{t},$$

Moreover, $f^o(x; d)$ is a sublinear function in d for every $d \in \mathbb{R}^n$ (see [51, Proposition 2.1.1(a)]).

Now, in analogous to the property in Theorem 2.2.29 for convex functions, we are in the position to define the subdifferential to nonconvex locally Lipschitz continuous functions.

Definition 2.2.38. [51, p. 27] Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz at a given point $x \in \mathbb{R}^n$. The **Clarke generalized subdifferential** of f at x, denoted by $\partial^o f(x)$, is defined as

$$\partial^o f(x) := \{\xi \in \mathbb{R}^n : f^o(x; d) \geq \langle \xi, d \rangle, \ \forall d \in \mathbb{R}^n\}.$$

If $f : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz at x, it is well known that the Clarke generalized subdifferential of f at x is a nonempty, convex and compact subset of $\mathbb{R}^n$ (see [51, Proposition 2.1.2(a)]), and

$$f^o(x; d) = \max_{\xi \in \partial^o f(x)} \langle \xi, d \rangle, \ \forall d \in \mathbb{R}^n.$$

(see [51, Propoition 2.1.2(b)]). Moreover, $\partial^o f$ is upper semicontinuous at x (see [51, Proposition 2.1.5(d)]).

Remember that max-functions are always encountered in the robust optimization. We shall need the results of differential calculus for max-functions. In what follows, we need the following regularity property.

Definition 2.2.39. [51, Definition 2.3.4] Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz at a given point $x \in \mathbb{R}^n$. The function f is said to be **regular** at $x \in \mathbb{R}^n$ if $f'(x; \cdot)$ and $f^o(x; \cdot)$ both exist and coincide.

We now note some sufficient conditions for a function to be regular.

Proposition 2.2.40. [51, Proposition 2.3.6 together with Corollary (p. 32)] *A function $f : \mathbb{R}^n \to \mathbb{R}$ is regular at a given point x if*

(i) *f is continuously differentiable at x, or*

(ii) *f is convex.*

The continuous differentiability is critical in the regularity as the next example shows.

Example 2.2.41. [51, Example 2.2.3] Remember that the function f in the Example 2.1.17 is differentiable but it is not continuously differentiable. In addition, $f^o(0; d) = |d|$ and $f'(0; d) = 0$ for all $d \in \mathbb{R}$. Therefore, f is not regular at $x = 0$.

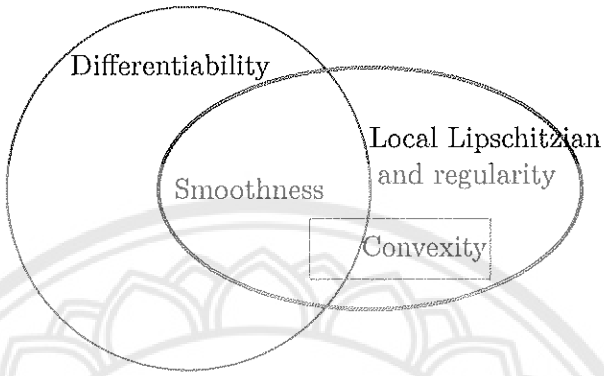


Figure 14: Relationship among differentiability, smoothness, convexity, and local Lipschitzian and regularity.

In order to formulate a differential calculus rule for max-functions, the following assumptions will be considered in the sequel.

Assumptions [58, Theorem 2.1] (see also [37, p. 2041]) Let $\mathcal{V}$ be a compact subset of $\mathbb{R}^q$. Suppose $g : \mathbb{R}^n \times \mathcal{V} \to \mathbb{R}$, is a function satisfying the following conditions:

(A1) $g(x, v)$ is upper semicontinuous in (x, v);

(A2) g is locally Lipschitz in the first argument uniformly in the second argument, i.e. for all $x \in \mathbb{R}^n$, there exist neighborhood U of x and a constant $L > 0$ such that for all y and z in U, and $v \in \mathcal{V}$, one has

$$|g(y, v) - g(z, v)| \leq L\|y - z\|;$$

(A3) g is regular with respect to x;

(A4) The generalized gradient $\partial_x^o g(x, v)$ with respect to the first component is upper semicontinuous in (x, v).

Remark 2.2.42. [58, Remark of Theorem 2.1] Note that, if one of the following conditions holds, then the conditions (A2), (A3), and (A4) hold:

(i) The function g is convex in x and continuous in v.

(ii) The derivative $\nabla_x g(x, v)$ with respect to x exists and is continuous in (x, v).

The following lemma will be useful in our later analysis especially a differential calculus rule for max-functions.

Lemma 2.2.43. [58, Theorem 2.1, Danskin theorem in nonsmooth setting](see also [59, Theorem 2]) *Let $\mathcal{V}$ be a nonempty compact subset of $\mathbb{R}^q$ and let $g : \mathbb{R}^n \times \mathcal{V} \to \mathbb{R}$ be such that the conditions (A1)-(A4) are fulfilled. Let $\psi(x) := \sup_{v \in \mathcal{V}} g(x, v)$. Denote $\mathcal{V}(x) := \{v \in \mathcal{V} : g(x, v) = \psi(x)\}$. Then the function ψ is locally **Lipschitz, directionally differentiable, regular** for each $x \in \mathbb{R}^n$ and*

$$\begin{aligned} \psi^o(x; d) &= \max\{g_x^o(x, v; d) : v \in \mathcal{V}(x)\} \\ &= \max\{\langle \xi, d \rangle : \xi \in \partial_x^o g(x, v), v \in \mathcal{V}(x)\}, \ \forall d \in \mathbb{R}^n. \end{aligned}$$

Lemma 2.2.44. [36] *Let $\mathcal{V}$ be a nonempty compact convex subset of $\mathbb{R}^q$ and let $g : \mathbb{R}^n \times \mathcal{V} \to \mathbb{R}$ be such that the conditions (A1)-(A4) are fulfilled. In addition, suppose that $g(x, \cdot)$ is concave on $\mathcal{V}$, for each $x \in U$. Then the following statements hold:*

(i) *The set $\mathcal{V}(x)$ is convex and compact.*

(ii) *The set $\partial_x^o g(x, \mathcal{V}(x)) := \{\xi \in \mathbb{R}^n : \exists v \in \mathcal{V}(x) \text{ s.t. } \xi \in \partial_x^o g(x, v)\}$ is convex and compact.*

(iii) *$\partial^o \psi(x) = \partial_x^o g(x, \mathcal{V}(x))$, where ψ is defined in Lemma 2.2.43.*

Observe that the notion of regularity plays a pivotal role and it does not generalize differentiable function. However, it is possible to unify the differentiability and the local Lipschitz regularity based on the following notion.

Definition 2.2.45. The function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be **tangentially convex** at $x \in \mathbb{R}^n$ (see [60] and [61]) if for every $d \in \mathbb{R}^n$ the one-sided directional derivative of f at x, $f'(x; d)$, exists, is finite, and is a convex function of d.

Note also that if f is tangentially convex at x, the function $f(x;\cdot)$ is a sublinear function.

Remark 2.2.46. [Classes of tangentially convex functions] The following points are taken from [48].

(i) Every convex function which has an open domain is tangentially convex at each point of its domain.

(ii) Every function which is *Gâteaux differentiable* at a point x, the directional derivative at x exists in all directions and is a linear function, is tangentially convex at x by the linearity of the directional derivative $f'(x;\cdot)$.

(iii) Every locally Lipschitz function which is regular in the sense of Clarke at a point x is tangentially convex at x, since in such a case the classical one-sided directional derivative is convex because it coincides with the Clarke directional derivative.

(iv) The class of tangentially convex functions at a given point is a real vector space, and hence some tangentially convex functions (not necessarily convex and differentiable) will follow from the sum of a convex function with a differentiable function as the next example shows.

Example 2.2.47. A function $f:\mathbb{R}\to\mathbb{R}$, defined as $f(x) := |x| - x^3$ for all $x \in \mathbb{R}$, is a tangentially convex function at 0.

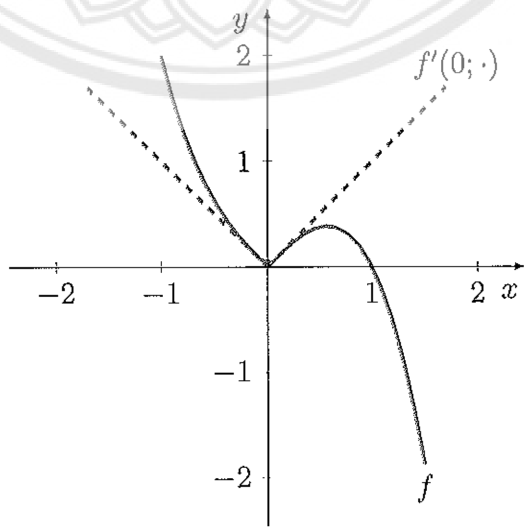


Figure 15: Example of a tangentially convex function.

The concept of subdifferential for tangentially convex function is implicitly given in [60].

Definition 2.2.48. The **tangential subdifferential** of (a tangentially convex function) f at $x \in \mathbb{R}^n$ is the set $\partial_T f(x)$ given as

$$\partial_T f(x) := \{\xi \in \mathbb{R}^n : f'(x;d) \geq \langle \xi, d \rangle, \ \forall d \in \mathbb{R}^n\},$$

which is a nonempty compact convex set.

It is important to note that if f is tangentially convex at $x \in \mathbb{R}^n$ such that $f(x) \in \mathbb{R}$, the function $f'(x;\cdot)$ is the support function of the tangential subdifferential, that is,

$$f'(x;d) = \max_{\xi \in \partial_T f(x)} \langle \xi, d \rangle, \ \text{for all } d \in \mathbb{R}^n. \tag{2.2.2}$$

Let $f : \mathbb{R}^n \to \mathbb{R}$ be tangentially convex at $x \in \mathbb{R}^n$. Surprisingly, the tangential subdifferentials enjoy nice calculus properties including the positive homogeneous rule and the sum rule, i.e.,

(i) for every $\theta \geq 0$, $\partial_T(\theta f)(x) = \theta \partial_T f(x)$;

(ii) if f and g is tangentially convex at the same point x, one has

$$\partial_T(f+g)(x) = \partial_T f(x) + \partial_T g(x).$$

Remark 2.2.49. For a given tangentially convex function $f : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$, it is easily to verify that the function $y \longmapsto f'(x; y-x)$ is convex and

$$\partial_T f(x) = \partial f'(x; \cdot - x)(x) = \partial f'(x;\cdot)(0).$$

Example 2.2.50. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined as $f(x_1, x_2) := \sqrt{x_1^2 + x_2^2} - x_1^3 - x_2^3$. Then, for every $(d_1, d_2) \in \mathbb{R}^2$, we can verify that

$$\frac{f((0,0) + t(d_1,d_2)) - f(0,0)}{t} = \sqrt{d_1^2 + d_2^2} - t^2 d_1^3 - t^2 d_2^3 \text{ for all } t > 0.$$

Letting $t \to 0^+$, we get $f'((0,0);(d_1,d_2)) = \sqrt{d_1^2 + d_2^2}$, from which it follows that $(d_1, d_2) \longmapsto f'((0,0);(d_1,d_2))$ is convex. So, f is tangentially convex at $(0,0)$ and its tangential subdifferential at $(0,0)$ is

$$\partial_T f(0,0) = \partial f'((0,0);\cdot)((0,0)) = [-1,1] \times [-1,1].$$

2.3 Lagrange multiplier theory revisited

Consider the problem (P) where $f, g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \dots, m$, are differentiable functions. The main draw back of the necessary optimality conditions for the problem (P) to attain its global minimizer laid down by John [1] as stated below.

Theorem 2.3.1. *Consider the problem* (P) *where* $f, g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \dots, m$, *are differentiable functions and let* a *be a point of global minimizer of* (P). *Then there exist* $\lambda_i \geq 0$, $i = 0, 1, 2, \dots, m$, *not all simultaneously zero, such that*

$$\lambda_0 \nabla f(a) + \sum_{i=1}^{m} \lambda_i \nabla g_i(a) = 0 \text{ and } \lambda_i g_i(a) = 0, \ i = 1, 2, \dots, m.$$

In view of Theorem 2.3.1, a multiplier that associated with the gradient of the objective function can also become zero. Consequently, making the objective function play no role in the optimization process. In addition, however, even if a feasible point a satisfies two conditions above and f, g_i, $i = 1, 2, \dots, m$, are convex, we can not conclude that a is a point of minimizer of (P). To see this, for every $x \in \mathbb{R}^n$ such that $g_i(x) \leq 0$ for all $i = 1, 2, \dots, m$,

$$\begin{aligned} 0 &= \lambda_0 \langle \nabla f(a), x - a \rangle + \sum_{i=1}^{m} \lambda_i \langle \nabla g_i(a), x - a \rangle \\ &\leq \lambda_0 (f(x) - f(a)) + \sum_{i=1}^{m} \lambda_i (g_i(x) - g_i(a)) \text{ (by convexity of } f \text{ and } g_i) \\ &\leq \lambda_0 (f(x) - f(a)). \end{aligned}$$

So, one needs a positive multiplier associated with the gradient of the objective function. Kuhn and Tucker [2] realized that in order to achieve this one needs to impose certain conditions on the constraints, known as constraint qualifications, and then the Fritz-John optimality conditions become in the form of Lagrange multiplier rule or KKT conditions. One of such constraint qualification, due to Slater [62], was introduced in 1950 and it still plays a very central role especially in convex optimization theory. So, it will be useful to briefly state the *Slater's constraint qualification.*

Definition 2.3.2. The **Slater's constraint qualification** is said to hold for the system $x \in \mathbb{R}^n$, $g_i(x) \leq 0$, $i = 1, 2, \ldots, m$, if there exists $x_0 \in \mathbb{R}^n$ such that $g_i(x_0) < 0$ for all $i = 1, 2, \ldots, m$.

In the recent past, there has been a renewed interest in the study of KKT optimality conditions for a convex optimization problem (CP), whose constraint set C is described as in (1.0.1) and every constraint functions are not necessarily convex. As a breakthrough to this, Lassere [46] considered in differentiable problem fulfilling Slater's constraint qualification and *non-degeneracy condition*: For all $i = 1, 2, \ldots, m$,

$$\nabla g_i(x) \neq 0, \text{ whenever } x \in C \text{ and } g_i(x) = 0.$$

He showed that as far as KKT optimality conditions are concerned, the convexity of constraint functions can be replaced by the convexity of the constraint set. This result has been obtained by Dutta and Lalitha [47] for the non-differentiable case involving Lipschitzian data of g_i which meet a regular condition in the sense of Clarke. In the discussion in [47], g_i, $i = 1, 2, \ldots, m$, are only need to be continuous [48]. We now give a description of some results in the aforementioned papers which will be useful in our later analysis.

Proposition 2.3.3. [47, Proposition 2.2] *Let $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$, be locally Lipschitz and regular in the sense of Clarke. Let $C := \{x \in \mathbb{R}^n : g_i(x) \leq 0,\ i = 1, 2, \ldots, m\}$ be nonempty. If C is convex, then, for each $i = 1, 2, \ldots, m$,*

$$g_i^o(x, y - x) \leq 0, \ \forall x, y \in C \text{ with } g_i(x) = 0. \tag{2.3.1}$$

Moreover, if Slater's constraint qualification holds, and $0 \notin \partial^o g_i(x)$ whenever $x \in C$ and $g_i(x) = 0$ (non-degeneracy condition), then, (2.3.1) *implies that C is convex.*

Theorem 2.3.4. [47, Theorem 2.4] *Let us consider the problem* (CP). *Let us assume that each g_i is locally Lipschitz and regular in the sense of Clarke. Further assume that the Slater's constraint qualification holds and non-degeneracy condition is satisfied. Then $\bar{x} \in C$ is a global minimizer of f over C and only if there exist $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$, such that*

i) $0 \in \partial f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \partial^o g_i(\bar{x})$,

ii) $\lambda_i g_i(\bar{x}) = 0, \ \forall i = 1, 2, \ldots, m.$

We next recall the notion of pseudoconvexity in tangentially convex setting.

Definition 2.3.5. [48, Definition 7] A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ which is tangentially convex at $x \in \operatorname{dom} f$ is said to be **pseudoconvex** at x if $f(y) \geq f(x)$ for every $y \in \mathbb{R}^n$ such that $f'(x; y - x) \geq 0$.

Theorem 2.3.6. [48, Theorem 9] *Suppose that functions $g_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $i = 1, 2, \ldots, m$, are continuous, the system $x \in \mathbb{R}^n$, $g_i(x) \leq 0$, $i = 1, 2, \ldots, m$, satisfies the Slater's constraint qualification, and for every $x \in C := \{x \in \mathbb{R}^n : g_i(x) \leq 0, \ i = 1, 2, \ldots, m\}$ and $i \in \{1, 2, \ldots, m\}$ such that $g_i(x) = 0$ the function g_i is tangentially convex at x and $\partial_T g_i(x) \neq \{0\}$. Assume further that the set C is convex, and the function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is tangentially convex at $\bar{x} \in C$. If either $g_i(\bar{x}) < 0$ for every $i \in \{1, 2, \ldots, m\}$ or for every $i \in \{1, 2, \ldots, m\}$ such that $g_i(\bar{x}) = 0$, $0 \notin \partial_T g_i(\bar{x})$, then a necessary condition for $\bar{x}$ to be a global minimizer of f over C is the existence of real numbers $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$, such that*

i) $0 \in \partial_T f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \partial_T g_i(\bar{x}),$

ii) $\lambda_i g_i(\bar{x}) = 0, \ \forall i = 1, 2, \ldots, m.$

We have already seen that Fritz-John optimality conditions are one of the key results to obtain the KKT conditions. To the end of this subsection, we now present a generalization of Fritz-John optimality conditions due to [60, p. 88, Corollary].

Theorem 2.3.7. *Let φ_i, $i = 0, 1, 2, \ldots, m$, are real-valued functions on $\mathbb{R}^n$. Let $\bar{x}$ be a solution to the following problem:*

$$\min_{x \in \mathbb{R}^n} \{\varphi_0(x) : \varphi_i(x) \leq 0, \ i = 1, 2, \ldots, m\}.$$

Suppose that there exist convex functions h_i, $i = 0, 1, 2, \ldots, m$, such that

$$\lim_{t \downarrow 0} \frac{\varphi_i(\bar{x} + td) - \varphi_i(\bar{x})}{t} \leq h_i(d), \ \forall d \in \mathbb{R}^n.$$

Then, there exist $\lambda_i \geq 0$, $i = 0, 1, 2, \ldots, m$, not all zero, such that

$$\sum_{i=0}^{m} \lambda_i h_i(d) \geq 0, \ \forall d \in \mathbb{R}^n.$$

2.4 Multi-objective optimization problems

Multi-objective (vector-valued) optimization is a subject of mathematical programming that extensively studied and applied in various decision-making contexts like economics, human decision making, control engineering, transportation and many others. We refer the reader to [63–66]. For comprehensive treatment of theoretical issues concerning multi-objective optimization can be found in [55, 67–69]. A multi-objective optimization problem can be formulated in mathematical terms as:

$$\mathbb{R}^p_+ - \min_{x \in \mathbb{R}^n} \{f(x) := (f_1(x), f_2(x), \ldots, f_p(x)) : x \in \Omega\}, \tag{MOP}$$

where the integer $p \geq 2$ is the number of objectives, each $f_j : \mathbb{R}^n \to \mathbb{R}$, $j = 1, 2, \ldots, p$, is a scalar function, the set Ω is the constraint set of decision vectors. An element $x^* \in \Omega$ is called a **feasible solution** or a **feasible decision**. A vector $z^* = f(x^*) \in \mathbb{R}^p$ for a feasible solution x^* is called an **objective vector** or an **outcome**.

In the multi-objective setting, the scalar concept of optimality does not apply directly due to the fact that all the objectives can not be simultaneously optimized with a single solution. To this effect, some of the objective vectors can be extracted for examination in a way that such vectors are those where none of the components can be improved without deteriorating their performance in at least one of the rest. In this way, we refer to a Pareto minimum [70] which usually uses coordinate-wise ordering (induced by the positive orthant as ordering cone) to examine the objective vectors. A more formal definition of Pareto optimality is the following:

Definition 2.4.1. A feasible point $x^* \in \Omega$ is said to be a **Pareto minimum point** (or an **efficient solution** or a **nondominated point**) of problem (MOP), if there is no feasible point $x \in \Omega$ such that

$$f_j(x) \leq f_j(x^*) \text{ for all } j = 1, 2, \ldots, p,$$

and

$$f_k(x) < f_k(x^*) \text{ for some } k \in \{1, 2, \ldots, p\},$$

or equivalently, $(\{f(x^*)\} - \mathbb{R}^p_+ \backslash \{0\}) \cap f(\Omega) = \emptyset$.

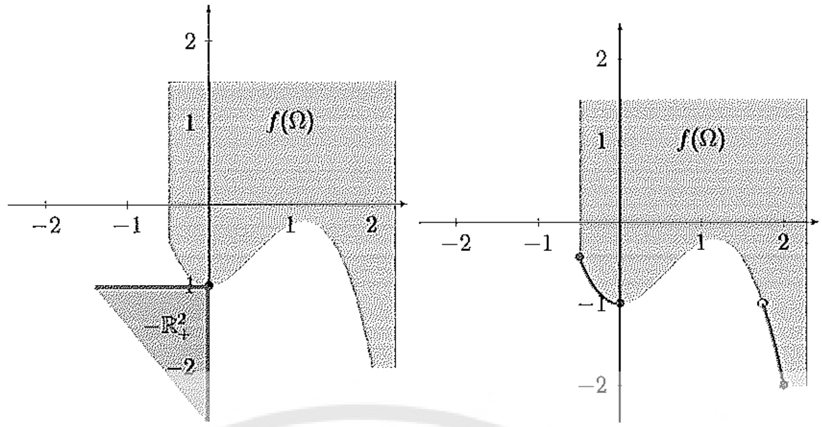


Figure 16: A geometric view of Pareto minimality.

In addition to Pareto optimality, the necessary optimality conditions for the problem (MOP) to attain its Pareto minimum rely on the way to establish necessary optimality conditions of the following weakly Pareto optimal solutions, which are often relevant from a technical point of view because they are sometimes easier to generate than Pareto optimal points. However, they are not always useful in practice, because they have no meaning on the economic scene.

Definition 2.4.2. A feasible point $x^* \in \Omega$ is called a **weak Pareto minimum point** (or a **weakly efficient solution** of (MOP) if there does not exist another feasible point $x \in \Omega$ such that

$$f_j(x) < f_j(x^*) \text{ for all } j = 1, 2 \ldots, p,$$

or equivalently, $(\{f(x^*)\} - \text{int } \mathbb{R}^p_+) \cap f(\Omega) = \emptyset$.

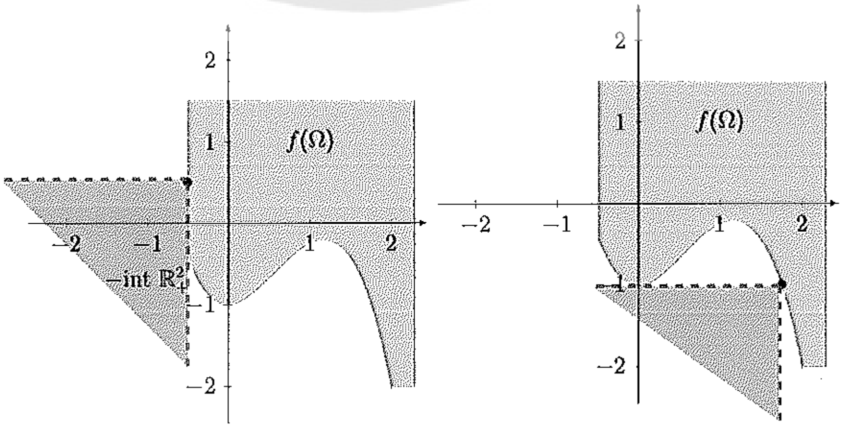


Figure 17: Geometric interpretation of a weak Pareto minimality.

Interestingly, when comparing two Pareto minimum points, they either obtain the same performance (all objectives equal), or, each beats the other in at least one objective. The latter case leads to the study of how much worse we must do in one or more objectives in order to do better in some other objectives. As an illustration, consider a bi-criterion problem. We might ask how much we must pay in the second objective to obtain an improvement in the first objective. This notion refers to the proper Pareto optimality due to Geoffrion [71].

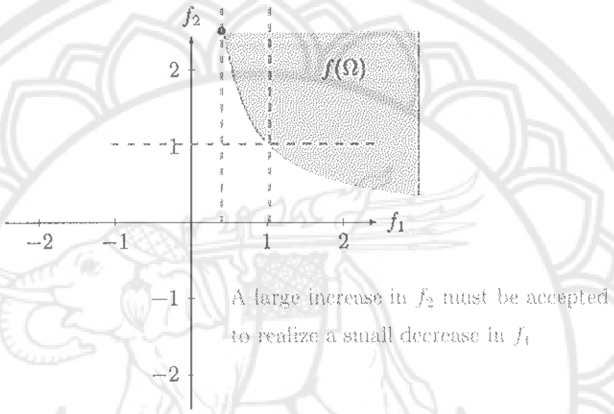


Figure 18: Illustration of the behavior of non-proper Pareto minimality.

Definition 2.4.3. A feasible point $x^* \in \Omega$ is said to be a **proper Pareto minimum point** (or a **properly efficient solution**) of problem of (MOP) if it is a Pareto minimum point and there exists $M > 0$ such that, for each $x \in \Omega$ and each $j \in \{1, 2, \ldots, p\}$ satisfying $f_j(x) < f_j(x^*)$, there exists $k \in \{1, 2, \ldots, p\}$ such that $f_k(x) > f_k(x^*)$ and

$$\frac{f_j(x^*) - f_j(x)}{f_k(x) - f_k(x^*)} \leq M,$$

or equivalently, $\mathrm{cl\ cone}(f(\Omega) + \mathbb{R}^p_+ - f(x^*)) \cap (-\mathbb{R}^p_+) = \{0\}$ (see [67, Theorem 3.1.4, p. 40]).

In other words, a solution is properly Pareto optimal if there is at least one pair of objectives for which a finite decrement in one objective is possible only at the expense of some reasonable increment in the other objective.

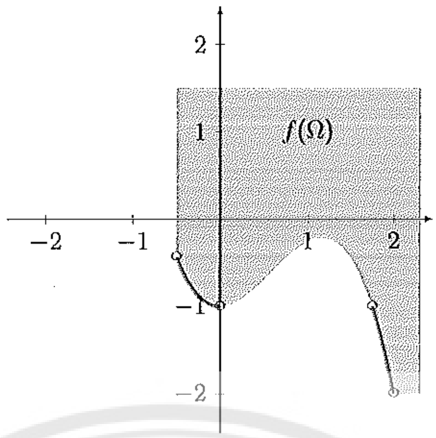


Figure 19: Illustration of proper Pareto minimality.

As far as we know, the search for weak Pareto minimum (resp. Pareto minimum) to (MOP) has been carried out through via scalarization. Scalarization means that the problem is converted into a single (scalar) or a family of single objective optimization problems, possibly depending on some parameters. After the multi-objective optimization problems has been scalarized, the widely developed theory and methods for single objective optimization can be used. Among the contributions on this way we remind the papers of [72], where the authors employ the linear scalarization, which is almost the most well-known scalarization technique in multi-objective optimization problems, to characterize the set of weak Pareto minimum and proper Pareto minimum under convexity assumptions of objective function and of constraint set. Before presenting let us consider the following scalar parameterized convex problem of (MOP) depending on a parameter $\theta := (\theta_1, \theta_2, \ldots, \theta_p) \in \mathbb{R}^p_+$:

$$\min_{x \in \mathbb{R}^n} \left\{ \sum_{j=1}^{p} \theta_j f_j(x) : x \in \Omega \right\}. \tag{P_θ}$$

Assume that the solution set of problem (P_θ), denoted by S_θ, is nonempty. So, the set of weakly efficient solutions $(W(\Omega))$ and the set of properly efficient solutions $(P(\Omega))$ for (MOP) can be characterized as follows:

(i) $x^* \in W(\Omega)$ if and only if there exists $\theta \in \mathbb{R}^p_+ \backslash \{0\}$ such that $x^* \in S_\theta$.

(ii) $x^* \in P(\Omega)$ if and only if there exists $\theta \in \text{int}\, \mathbb{R}^p_+$ such that $x^* \in S_\theta$.

CHAPTER III

CHARACTERIZATIONS OF THE SOLUTION SETS

3.1 Convex optimization problems without convexity of constraints

In this section, we aim to give characterizations of the solution set of (CP) without the pseudoconvexity assumption of Lagrange function. We begin by the following lemma.

Lemma 3.1.1. *Let C be defined as in (1.0.1), $x \in C$ and $I(x) := \{i \in I : g_i(x) = 0\}$. Assume that for every $i \in I(x)$ the function g_i is tangentially convex at x. If the set C is convex, then for any $y \in C$, one has*

$$g_i'(x; y - x) \leq 0, \ \forall i \in I(x). \tag{3.1.1}$$

Moreover, for each $i \in I(x)$, $\partial_T g_i(x) \subseteq N(C, x)$.

Proof. Suppose that the set C is convex. For any $y \in C$, the convexity of C implies that $x + t(y - x) \in C$ for any $t \in [0, 1]$. Consequently, for each $i \in I$,

$$g_i(x + t(y - x)) \leq 0 \text{ for all } t \in [0, 1].$$

Therefore, for every $i \in I(x)$, one has

$$\frac{g_i(x + t(y - x)) - g_i(x)}{t} = \frac{g_i(x + t(y - x))}{t} \leq 0 \text{ for all } t \in]0, 1].$$

By passing the last inequalities to the limit as $t \to 0^+$, as each g_i being continuous, we get (3.1.1).

Furthermore, (2.2.2) and (3.1.1) yield, for any $\xi \in \partial_T g_i(x)$, $\langle \xi, y - x \rangle \leq 0, \forall y \in C$, which gives that $\xi \in N(C, x)$, thereby establishing the desired result. □

Remark 3.1.2. In view of Lemma 3.1.1, we can see that for every $\lambda_i \geq 0$, $i \in I(\bar{x})$, $\sum_{i \in I(\bar{x})} \lambda_i \partial_T g_i(x) \subseteq \sum_{i \in I(\bar{x})} \lambda_i N(C, x) \subseteq N(C, x)$ due to $N(C, x)$ is a convex cone (see, Theorem 2.2.10 also).

It should be noted that, in convex programs, Slaters constraint qualification is usually used to obtain the Lagrange multiplier conditions which characterize optimality (see [9, 12, 14, 24, 26, 27] and other references therein). However, the Lagrange multiplier conditions for the convex optimization problems without convexity of the constraint functions may fail under the Slaters constraint qualification. Recently, Lagrange multiplier conditions have been obtained under the Slaters constraint qualification together with an additional condition on the constraints. Some constraint qualifications, which are also necessary for the existence of Lagrange multipliers for convex optimization problems without convexity of the constraint functions, has been introduced in [49, 73]. In an analogous manner as [49], we introduce the following constraint qualification in terms of tangential subdifferentials and show that it is the weakest constraint qualification for guaranteeing the Lagrange multiplier conditions to be necessary and sufficient for optimality of (CP).

Definition 3.1.3. [74] Let $x \in C$ and g_i, $i \in I(x)$, be tangentially convex at x. The *normal cone condition* is satisfied at x if

$$N(C, x) = \text{cone co} \bigcup_{i \in I(x)} \partial_T g_i(x).$$

Theorem 3.1.4. [74, **Weakest CQ for Lagrange multiplier conditions**] *Let $\bar{x} \in C$ be given, and for every $i \in I(\bar{x})$ the functions g_i be tangentially convex at $\bar{x}$. Then, the following assertions are equivalent:*

(i) *The normal cone condition is satisfied at $\bar{x}$;*

(ii) *For each convex function $f : \mathbb{R}^n \to \mathbb{R}$ attaining its global minimizer over C at $\bar{x}$, there exist $\lambda_i \geq 0$, $i \in I$, such that*

$$0 \in \partial f(\bar{x}) + \sum_{i \in I} \lambda_i \partial_T g_i(\bar{x}), \tag{3.1.2}$$

and

$$\lambda_i g_i(\bar{x}) = 0, \ \forall i \in I. \tag{3.1.3}$$

Proof. [(i) $\Rightarrow$ (ii)]. Suppose that (i) holds. Let $f : \mathbb{R}^n \to \mathbb{R}$ be any convex function such that $\bar{x} \in C$ is a global minimizer of (CP). It then follows from the convexity of C that, for each $t \in [0, 1]$,

$$f(\bar{x}) \leq f(\bar{x} + t(x - \bar{x})), \ \forall x \in C,$$

which gives

$$f'(\bar{x}; \bar{x} - \bar{x}) = 0 \leq f'(\bar{x}; x - \bar{x}), \ \forall x \in C. \tag{3.1.4}$$

It means that $\bar{x}$ is a minimizer of the convex function $f'(\bar{x}; \cdot - \bar{x})$ over C and it can be equivalently expressed as $0 \in \partial(f'(\bar{x}; \cdot - \bar{x}) + \delta_C)(\bar{x})$. In view of Theorem 2.2.26, we have $\text{dom} f'(\bar{x}; \cdot - \bar{x}) = \mathbb{R}^n$ and so, $\text{ri dom} f'(\bar{x}; \cdot - \bar{x}) \cap \text{ri dom } \delta_C = \text{ri } C \neq \emptyset$ (see, Proposition 2.2.7(i)). By Theorem 2.2.32, Remark 2.2.28 and Remark 2.2.49, we arrive at the assertion that

$$0 \in \partial f'(\bar{x}; \cdot - \bar{x})(\bar{x}) + N(C, \bar{x}) = \partial f(\bar{x}) + N(C, \bar{x}). \tag{3.1.5}$$

The condition (i) yields that there exists $\lambda_i \geq 0$, $i \in I(\bar{x})$, such that

$$0 \in \partial f(\bar{x}) + \sum_{i \in I(\bar{x})} \lambda_i \partial_T g_i(\bar{x}).$$

Setting $\lambda_i = 0$ for $i \notin I(\bar{x})$, the above expression can be rewritten as

$$0 \in \partial f(\bar{x}) + \sum_{i \in I} \lambda_i \partial_T g_i(\bar{x}) \text{ and } \lambda_i g_i(\bar{x}) = 0, \ \forall i \in I,$$

and hence (ii) has been justified.

[(ii) $\Rightarrow$ (i)]. Suppose that (ii) holds. By the virtue of Remark 3.1.2, we only need to prove that

$$N(C, \bar{x}) \subseteq \text{cone co} \bigcup_{i \in I(\bar{x})} \partial_T g_i(\bar{x}).$$

In fact, let $\xi \in N(C, \bar{x})$ be given. The definition of $N(C, \bar{x})$ yields that $\langle -\xi, \bar{x} \rangle \leq \langle -\xi, x \rangle$ for all $x \in C$. It can be seen that $f(x) := \langle -\xi, x \rangle$, $x \in \mathbb{R}^n$, is a convex function attaining its global minimizer over C at $\bar{x}$. So, from (ii) and $\partial f(\bar{x}) = \{-\xi\}$, there exist $\lambda_i \geq 0$, $i \in I$, such that

$$0 \in \{-\xi\} + \sum_{i \in I} \lambda_i \partial_T g_i(\bar{x}) \text{ and } \lambda g_i(\bar{x}) = 0, \ \forall i \in I.$$

This together with (2.2.1) in turn implies that

$$\xi \in \sum_{i \in I(\bar{x})} \lambda_i \partial_T g_i(\bar{x}) \subseteq \text{cone co} \bigcup_{i \in I(\bar{x})} \partial_T g_i(\bar{x}),$$

thereby leading to the desired result. $\square$

Remark 3.1.5. [Sufficient condition for the normal cone condition] As seen before, for each $\xi \in N(C,\bar{x})$, $f(x) := \langle -\xi, x\rangle$, $x \in \mathbb{R}^n$, is a convex function attaining its global minimizer over C at $\bar{x}$. Thus, if the system $g_i(x) \leq 0$, $i \in I$, satisfies the Slater's constraint qualification and the *non-degeneracy condition* at $\bar{x}$, i.e., for every $i \in I(\bar{x})$,

$$0 \notin \partial_T g_i(\bar{x}),$$

then Theorem 2.3.6 guarantees the existence of multipliers $\lambda_i \geq 0$, $i \in I$, such that

$$0 \in \{-\xi\} + \sum_{i\in I} \lambda_i \partial_T g_i(\bar{x}) \text{ and } \lambda g_i(\bar{x}) = 0,\ \forall i \in I,$$

consequently, $\xi \in \text{cone co} \bigcup_{i\in I(\bar{x})} \partial_T g_i(\bar{x})$. Therefore, the normal cone condition holds at $\bar{x}$.

Remark 3.1.6. In view of the proof of (3.1.4) and (3.1.5) in the proof of Theorem 3.1.4, one can notice that if For each convex function $f : \mathbb{R}^n \to \mathbb{R}$ attaining its global minimizer over C at $\bar{x}$ is replaced by For each tangentially convex and pseudoconvex function $f : \mathbb{R}^n \to \mathbb{R}$ at $\bar{x}$ attaining its global minimizer over C at $\bar{x}$, then its conclusions hold also true when the convex subdifferential $\partial f(\bar{x})$ is replaced by the tangential subdifferential $\partial_T f(\bar{x})$.

The following example illustrates that if the normal cone condition, condition (i) in Theorem 3.1.4, does not hold, then the optimality condition in Theorem 3.1.4 is not derived for a convex objective function.

Example 3.1.7. [Failure of Multiplier Characterization] Let us denote $x := (x_1, x_2) \in \mathbb{R}^2$, $g_1(x) := \sqrt{x_1^2 + x_2^2} - x_1^3 - 2$, $g_2(x) := -x_1^3 + \max\{-x_2, -x_2^3\}$, $g_3(x) := x_1 + \frac{1}{2}(x_2-1)^2 - \frac{1}{2}$, $C := \{x \in \mathbb{R}^2 : g_i(x) \leq 0,\ i \in I := \{1,2,3\}\}$ and $\bar{x} := (0,0)$. It is easy to verify that $C = \{x \in \mathbb{R}^2 : \sqrt{x_1^2 + x_2^2} - x_1^3 - 2 \leq 0,\ -x_1 - x_2 \leq 0,\ x_1 + \frac{1}{2}(x_2-1)^2 - \frac{1}{2} \leq 0\}$, $I(\bar{x}) = \{2,3\}$, $\partial_T g_2(\bar{x}) = \{0\} \times [-1,0]$ and $\partial_T g_3(\bar{x}) = \{(1,-1)\}$. It can be observed that

$$N(C,\bar{x}) = \text{cone co}\{(-1,-1),(1,-1)\}$$

and

$$\text{cone co}(\partial_T g_2(\bar{x}) \cup \partial_T g_3(\bar{x})) = \text{cone co}\{(0,-1),(1,1)\}.$$

Hence, we have that condition (i) of Theorem 3.1.4 does not hold. Thus for some convex function $f : \mathbb{R}^2 \to \mathbb{R}$, it may happen that the KKT optimality conditions may go awry at $\bar{x}$ even if it is a global minimizer. To see this, let $f(x) := e^{x_1} + x_2$. We see that

$$f(x) = e^{x_1} + x_2 \geq e^{x_1} - x_1 \geq 1 = f(\bar{x}) \text{ for all } x \in C.$$

So, $\bar{x}$ is a global minimizer of the convex function f over C. However, in this circumstance, we cannot find out $\lambda_1, \lambda_2, \lambda_3 \geq 0$ such that (3.1.2) and (3.1.3) hold. In fact, for $\lambda_1, \lambda_2, \lambda_3 \geq 0$ such that $\lambda_i g_i(\bar{x}) = 0$, $i \in I$, and

$$\begin{aligned} (0,0) &\in \partial f(\bar{x}) + \sum_{i \in I} \lambda_i \partial_T g_i(\bar{x}) \\ &= \{(1,1)\} + \lambda_1([-1,1] \times [-1,1]) + \lambda_2(\{0\} \times [-1,0]) + \lambda_3\{(1,-1)\}, \end{aligned}$$

it then follows that $\lambda_1 = 0$, and so, $1 + \lambda_3 = 0$ which contradicts to the validity of $\lambda_3 \geq 0$.

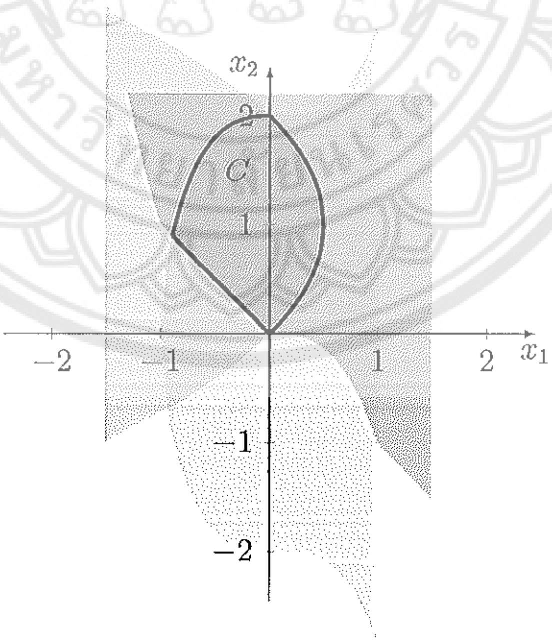


Figure 20: Illustration of a convex set without convex representation in Example 3.1.7.

Remark 3.1.8. In the case that for any $x \in C$ and $i \in I(x)$, g_i are locally Lipschitz and regular in the sense of Clarke, Theorem 3.2 in [49] can be obtained immediately by Theorem 3.1.4.

We will present some characterizations of the solution sets in terms of a given solution point of the convex minimization problem (CP).

Let $a \in S$ be a given solution point fulfilling the normal cone condition and for every $i \in I(a)$, the functions g_i be tangentially convex at a. Let $\lambda^a := (\lambda_1^a, \lambda_2^a \dots, \lambda_m^a) \in \mathbb{R}_+^m$ be a Lagrange multiplier vector corresponding to a such that (3.1.2) and (3.1.3) hold.

It is important to note that the constant Lagrangian-type property for the solution sets are commonly used to establish characterizations of solution sets for constrained optimization problems involving convex/pseudoconvex functions (see [9, 12, 24, 26, 27] and other references therein). However, the constant Lagrangian-type property for the solution sets may fail when some g_i are not convex even if the objective function is convex, for instance, let us define $f(x) := \max\{-x-1, 0\}$, $g_1(x) := \max\{x, x^3\}$ for $x \in \mathbb{R}$. We can see that f is a convex function while g_1 is not a convex function. Moreover, $a := 0$ is a minimizer of f on a convex set $C :=]-\infty, 0]$ with Lagrange multiplier $\lambda_1^a := 1$, and the solution set is $S = [-1, 0]$. However, the standard Lagrangian-type function $\mathcal{L}(x, \lambda_1^a) := f(x) + \lambda_1^a g_1(x)$ is not constant on the solution set S. In fact, $\mathcal{L}(a, \lambda_1^a) = 0 \neq -1 = \mathcal{L}(-1, \lambda_1^a)$. This situation motivates us to consider the so-called *pseudo Lagrangian-type function* $\mathcal{L}^P(\cdot, a, \lambda^a)$ [74], defined by

$$\mathcal{L}^P(x, a, \lambda^a) := f(x) + \sum_{i \in I(a)} \lambda_i^a g_i'(a; x-a), \text{ for all } x \in \mathbb{R}^n,$$

instead of the standard Lagrangian-type function. It can be seen that $\mathcal{L}^P(\cdot, a, \lambda^a)$ is constant on S, since $\mathcal{L}^P(x, a, \lambda^a) = \max\{-x-1, 0\} + \max\{x, 0\}$ for any $x \in \mathbb{R}$ and $\mathcal{L}^P(x, a, \lambda^a) = 0$ for any $x \in S$.

Next, we are in a position to prove that the pseudo Lagrangian function associated with a Lagrange multiplier corresponding to a solution is constant on a solution set S.

Proposition 3.1.9. [74, Constant pseudo Lagrangian-type property] For the problem (CP), assume that $a \in S$ satisfies the normal cone condition and the optimality conditions (3.1.2) and (3.1.3) hold with a Lagrange multiplier vector $\lambda^a := (\lambda_1^a, \lambda_2^a \dots, \lambda_m^a) \in \mathbb{R}_+^m$. Then for any $x \in S$,

$$\lambda_i^a g_i'(a; x-a) = 0, \ i \in I(a)$$

and $\mathcal{L}^P(\cdot, a, \lambda^a)$ is constant on S.

Proof. By Remark 2.2.49, $\partial_T g_i(a) = \partial g_i'(a; \cdot - a)(a)$ for all $i \in I(a)$. It follows from (3.1.2) and (3.1.3) that

$$0 \in \partial f(a) + \sum_{i \in I(a)} \lambda_i^a \partial g_i'(a; \cdot - a)(a) \subseteq \partial \mathcal{L}^P(\cdot, a, \lambda^a)(a),$$

and so,

$$f(x) + \sum_{i=1}^{m} \lambda_i^a g_i'(a; x-a) \geq f(a) \text{ for all } x \in \mathbb{R}^n.$$

By (3.1.3), it is easy to observe that $\lambda_i^a = 0$ for all $i \notin I(a)$. Therefore, by the fact that $f(x) = f(a)$, $\forall x \in S$, the above expression can be rewritten as

$$\sum_{i \in I(a)} \lambda_i^a g_i'(a; x-a) \geq 0 \text{ for all } x \in S.$$

Applying Lemma 3.1.1, we obtain $\lambda_i^a g_i'(a; x-a) = 0$, $\forall i \in I(a)$. Therefore, for any $x \in S$,

$$\mathcal{L}^P(x, a, \lambda^a) = f(x) + \sum_{i \in I(a)} \lambda_i^a g_i'(a; \cdot - a)(a) = f(x) = f(a),$$

thus yielding the desired results. □

Remark 3.1.10. [Pseudo Lagrangian-type function coincides with Lagrangian-type function] It is worth noting that if g_i, $i \in I(a)$, are pseudoconvex functions at a then, by Proposition 3.1.9, for any $x \in S$,

$$(\lambda_i^a g_i)'(a; x-a) = \lambda_i^a g_i'(a; x-a) = 0 \Longrightarrow \lambda_i^a g_i(x) \geq \lambda_i^a g_i(a) = 0.$$

This together with $x \in C$ yields, $\lambda_i^a g_i(x) = 0$, $i \in I(a)$. Furthermore,

$$\mathcal{L}^P(x, a, \lambda^a) = f(x) + \sum_{i \in I(a)} \lambda_i^a g_i'(a; x-a) = f(x) + \sum_{i=1}^{m} \lambda_i^a g_i(x), \ \forall x \in S.$$

It means that the pseudo Lagrangian-type function is the standard Lagrangian-type function on the solution set S.

Definition 3.1.11. A locally function Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be **pseudoconvex** if $f(y) \geq f(x)$ for every $x, y \in \mathbb{R}^n$ such that $f^o(x; y - x) \geq 0$.

Remark 3.1.12. In Proposition 3.1.9, if f is Locally Lipschitz, regular in the sense of Clarke and pseudoconvex, it is proved in [75, Lemma 3] that $\forall x, y \in \mathbb{R}^n$, one has

$$f(y) \leq f(x) \Longrightarrow f^o(x; y - x) \leq 0. \tag{3.1.6}$$

Applying (3.1.6), we can show that the conclusions given in Proposition 3.1.9 are still valid. Indeed, as $a \in S$, Theorem 3.1.4 asserts that there exists a Lagrange multiplier vector $\lambda^a := (\lambda_1^a, \lambda_2^a \ldots, \lambda_m^a) \in \mathbb{R}_+^m$ such that (3.1.2) and (3.1.3) hold. The fact that $\partial^o f(a) = \partial_T f(a)$ along with tangential subdifferential calculus rules at a imply that

$$f'(a; x - a) + \sum_{i \in I(a)} \lambda_i^a g_i'(a; x - a) \geq 0,\ \forall x \in \mathbb{R}^n.$$

Using (3.1.6) with the fact that $f(x) = f(a)$, $\forall x \in S$ and regularity of f, we deduce $\sum_{i \in I(a)} \lambda_i^a g_i'(a; x - a) \geq 0$, $\forall x \in S$, and hence, by Lemma 3.1.1, we obtain $\lambda_i^a g_i'(a; x - a) = 0$, $\forall i \in I(a)$, and for any $x \in S$, $\mathcal{L}^P(x, a, \lambda^a) = f(a)$.

In the sequel, we present characterizations of the solution set for problem (CP) in terms of convex subdifferentials, tangential subdifferentials and Lagrange multipliers. Denote by $\tilde{I}(a, \lambda^a)$ the following index set

$$\tilde{I}(a, \lambda^a) := \{i \in I(a) : \lambda_i^a > 0\}$$

and the set

$$X_1(\lambda^a) := \{x \in C : \forall i \in \tilde{I}(a, \lambda^a),\ \exists \eta_i \in \partial_T g_i(a),\ \langle \eta_i, x - a \rangle = 0\}.$$

Theorem 3.1.13. [74, Characterization of the solution set] *For the problem* (CP), *assume all conditions of Proposition 3.1.9. Then, the solution set S is characterized by*

$$S = \{x \in X_1(\lambda^a) : \exists \xi \in \partial f(x) \cap \partial f(a),\ \langle \xi, x - a \rangle = 0\}.$$

Proof. [$\subseteq$]. Let $x \in S$ be arbitrarily given. Then, x belongs to C. Furthermore, by (2.2.2) and Proposition 3.1.9, we have that for each $i \in \tilde{I}(a, \lambda^a)$,

$$\max_{\eta_i \in \partial_T g_i(a)} \langle \eta_i, x - a \rangle = g_i'(a; x - a) = 0.$$

Therefore, for each $i \in \tilde{I}(a, \lambda^a)$, there exists $\eta_i \in \partial_T g_i(a)$ such that

$$\langle \eta_i, x - a \rangle = 0.$$

On the other hand, it follows from (3.1.2) and (3.1.3) that there exist $\xi \in \partial f(a)$ such that $-\xi \in \sum_{i \in I} \lambda_i^a \partial_T g_i(a) = \partial_T(\sum_{i \in I} \lambda_i^a g_i)(a)$. That is, for any $d \in \mathbb{R}^n$,

$$\sum_{i \in I(a)} \lambda_i^a g_i'(a; d) = \left(\sum_{i \in I(a)} \lambda_i^a g_i \right)'(a; d) = \left(\sum_{i \in I} \lambda_i^a g_i \right)'(a; d) \geq \langle -\xi, d \rangle, \quad (3.1.7)$$

where the second equality follows from (3.1.3). Note from $x, a \in S$ that $f(x) = f(a)$. Letting $d := x - a$ in (3.1.7), one has $\sum_{i \in I(a)} \lambda_i^a g_i'(a; x - a) \geq \langle -\xi, x - a \rangle$, which together with Proposition 3.1.9 and $\xi \in \partial f(a)$ implies that

$$0 = \sum_{i \in I(a)} \lambda_i^a g_i'(a; x - a) \geq \langle -\xi, x - a \rangle \geq f(a) - f(x) = 0.$$

So, $\langle \xi, x - a \rangle = 0$. It remains to prove that $\xi \in \partial f(x)$. Now, for any $y \in \mathbb{R}^n$, we have

$$f(y) - f(x) = f(y) - f(a) \geq \langle \xi, y - a \rangle = \langle \xi, y - x \rangle + \langle \xi, x - a \rangle = \langle \xi, y - x \rangle.$$

Therefore, $\xi \in \partial f(x)$.

[$\supseteq$]. Conversely, let x be an arbitrary point of $\{x \in X_1(\lambda^a) : \exists \xi \in \partial f(x) \cap \partial f(a),\ \langle \xi, x - a \rangle = 0\}$. Then, $x \in C$ and there exists $\xi \in \partial f(x) \cap \partial f(a)$ such that $\langle \xi, a - x \rangle = 0$. So

$$f(a) - f(x) \geq \langle \xi, a - x \rangle = 0,$$

which together with the fact that $a \in S$ yields $f(x) = f(a)$, and so $x \in S$. □

As tangential convexity collapses to regularly locally Lipschitz setting and differentiability, the following corollaries are immediately direct consequences as special cases of Theorem 3.1.13.

Corollary 3.1.14. *For the problem* (CP), *let for any $x \in C$ and $i \in I(x)$ the functions g_i be locally Lipschitz and regular in the sense of Clarke, $a \in S$ be an optimal solution fulfilling the condition:*

$$N(C, a) = \text{cone co} \bigcup_{i \in I(a)} \partial^o g_i(a),$$

where $\partial^o g_i(a)$ *denotes the Clarke subdifferential of* g_i *at* a. *Assume that the optimality conditions* (3.1.2) *and* (3.1.3) *hold with a Lagrange multiplier vector* $\lambda^a := (\lambda_1^a, \lambda_2^a \ldots, \lambda_m^a) \in \mathbb{R}_+^m$. *Then, the solution set is characterized by*

$$S = \{x \in X_2(\lambda^a) : \exists \xi \in \partial f(x) \cap \partial f(a),\ \langle \xi, x - a\rangle = 0\},$$

where $X_2(\lambda^a) := \{x \in C : \forall i \in \tilde{I}(a, \lambda^a),\ \exists \eta_i \in \partial^o g_i(a),\ \langle \eta_i, x - a\rangle = 0\}$.

Proof. The desired results can be obtained immediately by Theorem 3.1.13, since every locally Lipschitz regular function g_i is tangentially convex at every point x, with $\partial_T g_i(a) = \partial^o g_i(a),\ \forall i \in \tilde{I}(a, \lambda^a)$. □

Corollary 3.1.15. *For the problem* (CP), *let* $f : \mathbb{R}^n \to \mathbb{R}$ *be a convex differentiable function and* $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$, *be differentiable functions,* $a \in S$ *be an optimal solution of* (CP) *fulfilling the condition:*

$$N(C, a) = \text{cone co} \bigcup_{i \in I(a)} \{\nabla g_i(a)\}.$$

Assume that the optimality conditions (3.1.2) *and* (3.1.3) *hold with a Lagrange multiplier vector* $\lambda^a := (\lambda_1^a, \lambda_2^a \ldots, \lambda_m^a) \in \mathbb{R}_+^m$. *Then,*

$$S = \{x \in C : \langle \nabla g_i(a), x - a\rangle = 0,\ \forall i \in \tilde{I}(a, \lambda^a),\ \nabla f(x) = \nabla f(a)\}.$$

Proof. It is clear that every differentiable functions g_i are tangentially convex at every point x, with $\partial_T g_i(x) = \{\nabla g_i(x)\},\ \forall i \in \tilde{I}(a, \lambda^a)$. It follows from Theorem 3.1.13 with $\partial f(x) = \{\nabla f(x)\}$ for every point $x \in \mathbb{R}^n$ that

$$S = \{x \in C : \langle \nabla g_i(a), x - a\rangle = 0,\ \forall i \in \tilde{I}(a, \lambda^a),\ \nabla f(x) = \nabla f(a), \\ \langle \nabla f(x), x - a\rangle = 0\}.$$

Further, since a satisfies optimality condition (3.1.2), we have

$$0 = \left\langle \sum_{i=1}^m \lambda_i^a \nabla g_i(a), x - a \right\rangle = \langle \nabla f(a), x - a\rangle = \langle -\nabla f(x), x - a\rangle.$$

Thus, the condition $\langle \nabla f(x), x - a\rangle = 0$ is superfluous. Therefore, $S = \{x \in C : \langle \nabla g_i(a), x - a\rangle = 0,\ \forall i \in \tilde{I}(a, \lambda^a),\ \nabla f(x) = \nabla f(a)\}$. □

When one solution of the considered problem is known, by using Theorem 3.1.13, we can find all of solutions of the convex optimization problem that have multiple solutions, and moreover at least one of the constraint functions g_i is not convex while the constraint set is convex. So, Theorem 2.2 and Corollary 2.1 in [9] cannot be applied in the following example.

Example 3.1.16. [Verifying solution set] Let us denote $x := (x_1, x_2) \in \mathbb{R}^2$. Consider the following constrained optimization problem:

$$\min_{x\in\mathbb{R}^n}\{f(x) : x \in C := \{x \in \mathbb{R}^2 : g_1(x) \leq 0,\ g_2(x) \leq 0,\ g_3(x) \leq 0\}\},$$

where $f(x) := \sqrt{x_1^2+x_2^2}-x_1-x_2$, $g_1(x) := \sqrt{x_1^2+x_2^2}-x_1^3-2$, $g_2(x) := \max\{-x_1, -x_1^3\}-x_2$, $g_3(x) := x_1$. Evidently, the function f is a convex function. Let us notice that

$$f(x) = \sqrt{x_1^2+x_2^2} - x_1 - x_2 \geq |x_2| - x_2 = 0, \text{ for all } x \in C.$$

Thus $a := (a_1, a_2) = (0,0) \in S$, $I(a) = \{2,3\}$, $\partial_T g_2(a) = \{(r,-1) : r \in [-1,0]\}$ and $\partial_T g_3(a) = \{(1,0)\}$. It is easy to verify that this problem satisfies the Slater's condition and non-degeneracy at a. Also, the convex subdifferential of f at any point x is given by

$$\partial f(x) = \begin{cases} \{(-1,-1)\} + \mathbb{B}(0,1) & \text{if } x = (0,0), \\ \left\{\left[\frac{1}{\sqrt{x_1^2+x_2^2}}\right](x_1,x_2) + (-1,-1)\right\} & \text{if } x \neq (0,0). \end{cases}$$

Let us select $\lambda^a := (\lambda_1^a, \lambda_2^a, \lambda_3^a) = (0,0,1)$. Therefore, by using Theorem 3.1.13, the solution set can be described simply as

$$\begin{aligned} S &= \{x \in C : \langle(\eta_{3,1}, \eta_{3,2}), (x_1,x_2)\rangle = 0 \text{ for some } (\eta_{3,1}, \eta_{3,2}) \in \partial_T g_3(a), \\ &\qquad \langle(\xi_1,\xi_2),(x_1,x_2)\rangle = 0 \text{ for some } (\xi_1,\xi_2) \in \partial f(x) \cap \partial f(a)\} \\ &= \Big\{x \in \mathbb{R}^2 : \sqrt{x_1^2+x_2^2} - x_1^3 - 2 \leq 0,\ \max\{-x_1,-x_1^3\} - x_2 \leq 0,\ x_1 \leq 0, \\ &\qquad \langle(1,0),(x_1,x_2)\rangle = 0,\ \left\langle\left(\frac{x_1}{\sqrt{x_1^2+x_2^2}} - 1, \frac{x_2}{\sqrt{x_1^2+x_2^2}} - 1\right), (x_1,x_2)\right\rangle = 0\Big\} \\ &= \{x \in \mathbb{R}^2 : x_1 = 0,\ \sqrt{x_2^2} - 2 \leq 0,\ -x_2 \leq 0\} \\ &= \{x \in \mathbb{R}^2 : x_1 = 0,\ 0 \leq x_2 \leq 2\}. \end{aligned}$$

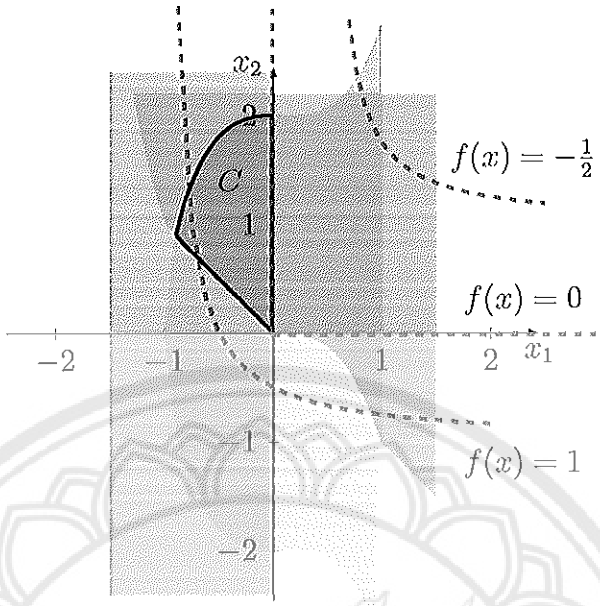


Figure 21: The behavior of the values of the objective function in Example 3.1.16.

Next, we give a characterization of S using subdifferentials of the pseudo Lagrangian-type function. To this aim, we need the following lemma.

Lemma 3.1.17. *For the problem* (CP), *assume all conditions of Proposition 3.1.9. Then for each* $x \in S$,

$$\partial \mathcal{L}^P(\cdot, a, \lambda^a)(x) = \partial \mathcal{L}^P(\cdot, a, \lambda^a)(a).$$

Proof. Now take any $\xi \in \partial \mathcal{L}^P(\cdot, a, \lambda^a)(x)$. Then, by the definition of convex subdifferential,

$$\mathcal{L}^P(y, a, \lambda^a) - \mathcal{L}^P(x, a, \lambda^a) \geq \langle \xi, y - x \rangle, \ \forall y \in \mathbb{R}^n. \tag{3.1.8}$$

Since $\mathcal{L}^P(\cdot, a, \lambda^a)$ is constant on S (Proposition 3.1.9) and $a \in S$, it follows from (3.1.8) that $\langle \xi, a - x \rangle = 0$ and so, $\langle \xi, y - x \rangle = \langle \xi, y - a \rangle + \langle \xi, a - x \rangle = \langle \xi, y - a \rangle$ for all $y \in \mathbb{R}^n$. This together with (3.1.8) entails

$$\mathcal{L}^P(y, a, \lambda^a) - \mathcal{L}^P(a, a, \lambda^a) \geq \langle \xi, y - a \rangle, \ \forall y \in \mathbb{R}^n,$$

which shows that $\xi \in \partial \mathcal{L}^P(\cdot, a, \lambda^a)(a)$. So, $\partial \mathcal{L}^P(\cdot, a, \lambda^a)(x) \subseteq \partial \mathcal{L}^P(\cdot, a, \lambda^a)(a)$. The proof of the converse inclusion is quite a similar argument and will be omitted. □

With the help of Proposition 3.1.9 and Lemma 3.1.17, we see now how the solution set can be characterized in terms of the pseudo Lagrangian-type function.

Proposition 3.1.18. [74] *For the problem* (CP), *assume all conditions of Proposition 3.1.9. Then,*

$$S = \{x \in X_2(\lambda^a) : 0 \in \partial \mathcal{L}^P(\cdot, a, \lambda^a)(x)\}.$$

Proof. Denote

$$S^* := \{x \in X_2(\lambda^a) : 0 \in \partial \mathcal{L}^P(\cdot, a, \lambda^a)(x)\}.$$

By Lemma 3.1.17 and the optimality condition (3.1.2), we get $0 \in \partial \mathcal{L}^P(\cdot, a, \lambda^a)(x)$ for all $x \in S$. This together with Proposition 3.1.9 implies easily that $S \subseteq S^*$. To establish the converse inclusion, let $x \in S^*$ be given. Then, by the definition of S^*, $x \in C$ and for each $i \in \tilde{I}(a, \lambda^a)$ there exist $\eta_i \in \partial_T g_i(a)$ such that

$$\langle \eta_i, x - a \rangle = 0,$$

which implies, for every $y \in \mathbb{R}^n$, that

$$\begin{aligned}
& f(y) + \sum_{i \in I(a)} \lambda_i^a g_i'(a; y - a) \\
\geq\ & f(x) + \sum_{i \in I(a)} \lambda_i^a g_i'(a; x - a) \\
=\ & f(x) + \sum_{i \in I(a) \setminus \tilde{I}(a, \lambda^a)} \lambda_i^a g_i'(a; x - a) + \sum_{i \in \tilde{I}(a, \lambda^a)} \lambda_i^a g_i'(a; x - a) \\
=\ & f(x) + \sum_{i \in \tilde{I}(a, \lambda^a)} \lambda_i^a g_i'(a; x - a) \\
\geq\ & f(x) + \sum_{i \in \tilde{I}(a, \lambda^a)} \lambda_i^a \langle \eta_i; x - a \rangle = f(x).
\end{aligned}$$

Taking $y := a$ in the last inequality, we get that $f(a) \geq f(x)$, and hence, for all $y \in C$,

$$f(y) \geq f(y) + \sum_{i \in I(a)} \lambda_i^a g_i'(a; y - a) \geq f(a) \geq f(x),$$

where the first inequality follows from Lemma 3.1.1. This proves that $x \in S$. $\square$

It turns out that Theorem 3.1.13 and Proposition 3.1.18 immediately yield the characterizations of the solution set for convex programs that was proposed in [9, Corollary 2.1 and Corollary 2.6].

Corollary 3.1.19. *For the problem* (CP), *let the functions* $g_i : \mathbb{R}^n \to \mathbb{R}$, $i \in I$ *be convex, and* $a \in S$ *an optimal solution fulfilling the normal cone condition and the optimality conditions* (3.1.2) *and* (3.1.3) *hold with a Lagrange multiplier vector* $\lambda^a := (\lambda_1^a, \lambda_2^a \ldots, \lambda_m^a) \in \mathbb{R}_+^m$. *Then, the solution set* S *of* (CP) *is characterized by*

$$S = \{x \in X(\lambda^a) : \exists \xi \in \partial f(x) \cap \partial f(a),\ \langle \xi, x - a \rangle = 0\},$$

where $X(\lambda^a) := \{x \in \mathbb{R}^n : g_i(x) = 0,\ \forall i \in \tilde{I}(a, \lambda^a),\ g_i(x) \leq 0,\ \forall i \in I \backslash \tilde{I}(a, \lambda^a)\}$.

Proof. By Theorem 3.1.13, we have that

$$S = \{x \in X_1(\lambda^a) : \exists \xi \in \partial f(x) \cap \partial f(a),\ \langle \xi, x - a \rangle = 0\}.$$

Let $x \in X_1(\lambda^a)$. As $\langle \eta_i, x - a \rangle = 0$ for some $\eta_i \in \partial g_i(a)$, $\forall i \in \tilde{I}(a, \lambda^a)$, we have $g_i'(a; x - a) \geq \langle \eta_i, x - a \rangle = 0$ for each $i \in \tilde{I}(a, \lambda^a)$. This together with $x \in C$, by Lemma 3.1.1, yields $g_i'(a; x - a) = 0$, $\forall i \in \tilde{I}(a, \lambda^a)$. Moreover, by Remark 3.1.10, we get that

$$\begin{aligned} &[x \in C,\ \forall i \in \tilde{I}(a, \lambda^a),\ \exists \eta_i \in \partial g_i(a),\ \langle \eta_i, x - a \rangle = 0] \\ \Longrightarrow\ &[x \in \mathbb{R}^n,\ g_i(x) = 0,\ \forall i \in \tilde{I}(a, \lambda^a),\ g_i(x) \leq 0,\ \forall i \in I \backslash \tilde{I}(a, \lambda^a)], \end{aligned}$$

consequently, $X_1(\lambda^a) \subseteq X(\lambda^a)$, and so, $S \subseteq \{x \in X_1(\lambda^a) : \exists \xi \in \partial f(x) \cap \partial f(a),\ \langle \xi, x - a \rangle = 0\}$. For the reverse containment, due to the fact $X(\lambda^a) \subseteq C$, the proof is similar to the one in Theorem 3.1.13, and so is omitted. □

Corollary 3.1.20. *For the problem* (CP), *let the functions* $g_i : \mathbb{R}^n \to \mathbb{R}$, $i \in I$ *be convex, and* $a \in S$ *an optimal solution fulfilling the normal cone condition and the optimality conditions* (3.1.2) *and* (3.1.3) *hold with a Lagrange multiplier vector* $\lambda^a := (\lambda_1^a, \lambda_2^a \ldots, \lambda_m^a) \in \mathbb{R}_+^m$. *Then,*

$$S = \{x \in X(\lambda^a) : 0 \in \partial \mathcal{L}(\cdot, \lambda^a)(x)\}.$$

Next, we will derive characterizations of the solution set of the following pseudoconvex minimization problem over a convex set (CP$'$):

$$\min_{x\in\mathbb{R}^n}\{f(x) : x \in C\}, \tag{CP$'$}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz, regular in the sense of Clarke and pseudoconvex, and $g_i : \mathbb{R}^n \to \mathbb{R}$, $i \in I$, are continuous functions and the constraint set C, defined as in (1.0.1), is a nonempty convex subset of $\mathbb{R}^n$. In view of Remark 3.1.12, we can obtain the following results.

Theorem 3.1.21. [74] *For the problem* (CP$'$), *let S' be the optimal solution set of* (CP$'$), *$\bar{x} \in S'$ an optimal solution fulfilling the normal cone condition and the optimality conditions* (3.1.2) *and* (3.1.3) *hold with a Lagrange multiplier vector $\lambda^a := (\lambda_1^a, \lambda_2^a \ldots, \lambda_m^a) \in \mathbb{R}_+^m$, and the functions g_i, $i \in I(a)$, be tangentially convex at a. Then,*

$$S' = \{x \in X_1(\lambda^a) : \exists p^x > 0,\ \exists \xi \in \partial^o f(x) \cap p^x \partial^o f(a),\ \langle \xi, x - a\rangle = 0\}.$$

Proof. [$\subseteq$]. Let us assume that $x \in S'$. By the same arguments given in the proof of Theorem 3.1.13, we can obtain that, for each $i \in \tilde{I}(a, \lambda^a)$, there exists $\eta_i \in \partial_T g_i(a)$ such that $\langle \eta_i, x - a\rangle = 0$. Furthermore, by [43, Lemma 3.4], there exists a real number $p^x > 0$ (depending on x) such that

$$\partial^o f(x) \cap p^x \partial^o f(a) \neq \emptyset.$$

It follows that there exists $\xi \in \mathbb{R}^n$ such that $\xi \in \partial^o f(x)$ and $\frac{1}{p^x}\xi \in \partial^o f(a)$. As $x, a \in S'$, $f(x) = f(a)$. It follows from [75, Lemma 3] that $f^o(x; x - a) \leq 0$ and $f^o(a; a - x) \leq 0$. So,

$$\langle \xi, x - a\rangle \leq 0 \text{ and } \left\langle \frac{1}{p^x}\xi, a - x \right\rangle \leq 0,$$

and hence $\langle \xi, x - a\rangle = 0$.

[$\supseteq$]. For every $x \in \{x \in X_1(\lambda^a) : \exists p^x > 0,\ \exists \xi \in \partial^o f(x) \cap p^x \partial^o f(a),\ \langle \xi, x - a\rangle = 0\}$, we get that $x \in C$, $\frac{1}{p^x}\xi \in \partial^o f(a)$ and $\langle \frac{1}{p^x}\xi, x - a\rangle = 0$ for some $p^x > 0$ and $\xi \in \partial^o f(x)$. In addition, for any $d \in \mathbb{R}^n$ such that $\langle \frac{1}{p^x}\xi, d\rangle \geq 0$, one has $f^o(x; d) \geq \langle \xi, d\rangle \geq 0$. Therefore,

$$x \ \in \ \{z \in C : \exists \xi \in \partial^o f(a),\ \langle \xi, z - a\rangle = 0;$$

$$\forall d \in \mathbb{R}^n, \ \langle \xi, d \rangle \geq 0 \Longrightarrow f^o(z; d) \geq 0\},$$

and hence, [75, Theorem 9] leads to $x \in S'$, thus yielding the desired results. $\square$

Before we end this section, let us illustrate the usefulness of Theorem 3.1.21 via an example.

Example 3.1.22. Consider the constrained optimization problem (CP′) where

$$f(x) := \max\{0, \frac{1}{2}x^2 - \frac{1}{2}, (x-1)^3 + 1\},$$
$$g_1(x) := \max\{x, x^3\},$$
$$g_2(x) := 4x - x^3,$$
$$g_3(x) := |x - 1| - 3, \text{ for every } x \in \mathbb{R}.$$

Evidently, the function f is a locally Lipschitz pseudoconvex function. Let us notice that

$$f(x) \geq 0 = \max\{0, -\frac{1}{2}, 0\} = f(0), \text{ for all } x \in C.$$

Then $a := 0 \in S'$, $I(a) = \{1, 2\}$ and each g_i, $i \in I(a)$, is tangentially convex at a. We can verify that

$$C = [-2, 0], \ \partial_T g_1(\bar{x}) = [0, 1], \ \partial_T g_2(\bar{x}) = \{4\}, \ \partial_T g_3(\bar{x}) = \{-1\}.$$

Also, the Clarke subdifferential of f at any point x is given by

$$\partial^o f(x) = \begin{cases} \{x\} & \text{if } x \in]-\infty, -1[, \\ [-1, 0] & \text{if } x = -1, \\ \{0\} & \text{if } x \in]-1, 0[, \\ [0, 3] & \text{if } x = 0, \\ \{3(x-1)^2\} & \text{if } x \in]0, +\infty[. \end{cases}$$

We can see that this problem does not satisfy non-degeneracy at a, the normal cone condition is fulfilled. Let us select $\lambda^a := (\lambda_1^a, \lambda_2^a, \lambda_3^a) = (1, 0, 0)$. Observe that for any $x \in [-1, 0[$, we can find $p^x > 0$ and $\xi \in \partial^o f(x) \cap p^x \partial^o f(a)$ such that $\langle \xi, x - a \rangle = 0$. So, by Theorem 3.1.21, we can obtain that the solution set can be described as $S' = [-1, 0]$.

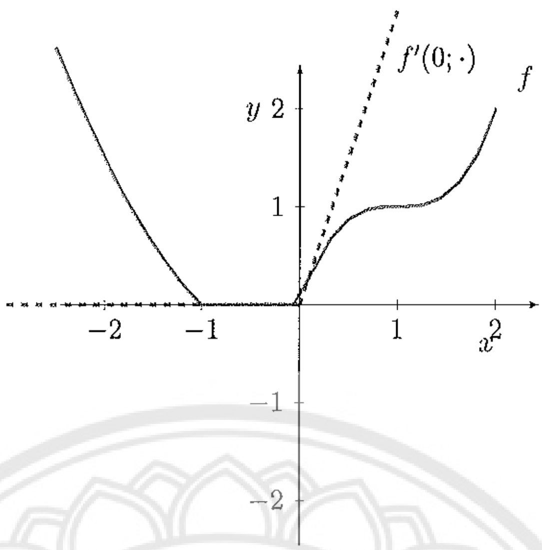


Figure 22: Plots of function f and its directional derivative in Example 3.1.22.

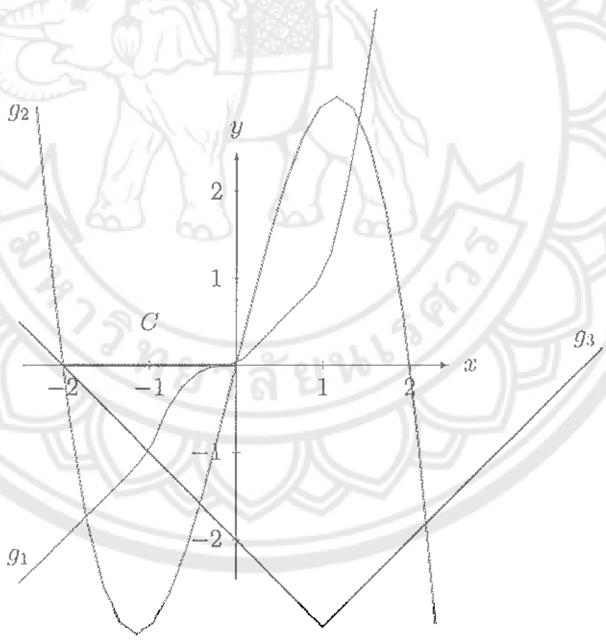


Figure 23: Illustration of constraint functions and the constraint set in Example 3.1.22.

Remark 3.1.23. In Example 3.1.22,

(i) g_1 is not pseudoconvex at $a := 0$, i.e., taking $y := -1$ we have $g'(a; y-a) = \max\{0, y\} = 0$, but $g_1(y) = -1 < 0 = g_1(a)$. Then Theorem 4.1 and 4.2 in [24] may not be relevant to this example.

(ii) The standard Lagrangian-type function with Lagrange multiplier $\lambda^a := (1,0,0)$, $\mathcal{L}(x,\lambda^a) = f(x) + \sum_{i=1}^{3} \lambda_i^a g_i(x)$, is not pseudoconvex at a, i.e., by taking $y := -1$ we get $\mathcal{L}(\cdot,\lambda^a)'(a; y-a) = \max\{4y, 0\} = 0$ while $\mathcal{L}(y,\lambda^a) = -1 < 0 = \mathcal{L}(a,\lambda^a)$. So, Theorem 3.3 in [27] cannot be applied.

(iii) Theorem 3.2 and Corollary 3.1 in [12] may not actually be relevant to this example because the constraint functions are not linear, and moreover, f is not pseudoconcave, i.e. $-f$ is not pseudoconvex, by considering $y := 2$ and $a = 0$ we have $(-f)'(a; y-a) = 0$, but $(-f)(y) = -2 < 0 = (-f)(a)$.

3.2 Uncertain convex optimization without convexity of constraint data uncertainty

In this section, we establish the characterizations of the robust optimal solution sets of (UP). In what follows, let us recall the following robust (worst case) counterpart optimization problem of (UP):

$$\min_{x\in\mathbb{R}^n} \left\{ \max_{u\in\mathcal{U}} f(x,u) : g_i(x,v_i) \leq 0,\ \forall v_i \in \mathcal{V}_i,\ i = 1,2,\ldots,m \right\}, \tag{RP}$$

where $f : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}$, and $g_i : \mathbb{R}^n \times \mathcal{V}_i \to \mathbb{R}$, $i = 1,2,\ldots,m$, are given functions and for each $i = 1,2,\ldots,m$, $(u,v_i) \in \mathcal{U} \times \mathcal{V}_i \subseteq \mathbb{R}^{q_0} \times \mathbb{R}^{q_i}$, where $\mathcal{U}$ and $\mathcal{V}_i$ are the specified nonempty convex and compact uncertainty sets. The robust feasible set of (UP) is given by

$$C_R := \{x \in \mathbb{R}^n : g_i(x,v_i) \leq 0,\ \forall v_i \in \mathcal{V}_i,\ i = 1,2,\ldots,m\}. \tag{3.2.1}$$

Throughout this section, we alway suppose that each $g_i : \mathbb{R}^n \times \mathcal{V}_i \to \mathbb{R}$, $i = 1,2,\ldots,m$, is a function satisfying the given assumptions (A1)-(A4), $g_i(x,\cdot)$ is a concave function on $\mathcal{V}_i$ for each $x \in \mathbb{R}^n$, and $C_R \neq \emptyset$. In addition, for each $i = 1,2,\ldots,m$, let us define a function ψ_i by

$$\psi_i(x) := \max_{v_i\in\mathcal{V}_i} g_i(x,v_i),\ x \in \mathbb{R}^n.$$

It is worth noting here, in view of Lemma 2.2.43, that for each $i = 1,2,\ldots,m$, ψ_i is a locally Lipschitz function on $\mathbb{R}^n$, and so it is a continuous function as well. It then follows that C_R is a closed set and for each $x \in \mathbb{R}^n$, by Lemma 2.2.44,

$$\mathcal{V}_i(x) := \{v_i \in \mathcal{V}_i : g_i(x, v_i) = \psi_i(x)\}$$

is a nonempty compact subset of $\mathbb{R}^{q_i}$.

Corresponding to any $x \in C_R$, for notational simplicity, we denote $I := \{1, 2, \ldots, m\}$ and decompose it into two index sets $I = I_1(x) \cup I_2(x)$, where

$$I_1(x) := \{i \in I : \exists v_i \in \mathcal{V}_i \text{ s.t. } g_i(x, v_i) = 0\}$$

is the active index set at $x \in C_R$ and $I_2 = I \backslash I_1(x)$. For $i \in I_1(x)$, let

$$\mathcal{V}_i^0(x) := \{v_i \in \mathcal{V}_i : g_i(x, v_i) = 0\}.$$

In order to derive characterizations of the robust optimal solution sets of (UP), one needs to investigate the multiplier characterization for the robust optimal solution of (UP) that suggests a way to obtain the Lagrange multipliers. In this way, we need to impose certain conditions on the constraints. Let us now introduce these conditions in the following definition.

Definition 3.2.1.

(i) **(Robust Slater-type constraint qualification [39]).** The set C_R is said to satisfy the *robust Slater-type constraint qualification* (RSCQ for short) if there exists $x_0 \in \mathbb{R}^n$ such that for each $i \in I$, it holds

$$g_i(x_0, v_i) < 0, \ \forall v_i \in \mathcal{V}_i.$$

(ii) **(Robust non-degeneracy condition [76]).** One says that C_R satisfies the *robust non-degeneracy condition* at $x \in C_R$ if for each $i \in I_1(x)$, it holds

$$0 \notin \partial_x^o g_i(x, v_i), \ \forall v_i \in \mathcal{V}_i^0(x).$$

One says that C_R satisfies the robust non-degeneracy condition whenever the robust non-degeneracy condition holds at every point $x \in C_R$.

(iii) **(Robust basic constraint qualification).** We say that the *robust basic constraint qualification* is satisfied at $x \in C_R$ if

$$N(C_R, x) = \bigcup_{\substack{\lambda_i \geq 0, \ v_i \in \mathcal{V}_i \\ \lambda_i g_i(x, v_i) = 0, \ i \in I}} \sum_{i \in I} \lambda_i \partial_x^o g_i(x, v_i).$$

Note that the robust basic constraint qualification which was introduced in [34], where the constraint data uncertainty $g_i(\cdot, v_i)$, $i = 1, \ldots, m$, are assumed to be convex for each $v_i \in \mathcal{V}_i$.

We begin by the following lemma which plays a key role in establishing the multiplier characterization as well as deriving characterizations of the robust optimal solution sets of (UP).

Lemma 3.2.2 (Characterizing convexity of robust feasible set). *Let C_R be defined as in* (3.2.1). *If C_R is convex, then, for each $z \in C_R$,*

$$g_{ix}^o(z, v_i; x - z) \leq 0, \ \forall x \in C_R, \ \forall i \in I_1(z), \ \forall v_i \in \mathcal{V}_i^0(z), \tag{3.2.2}$$

equivalently,

$$\psi_i^o(z; x - z) \leq 0, \ \forall x \in C_R, \ \forall i \in I_1(z).$$

Furthermore, if (RSCQ) holds and the robust non-degeneracy condition is satisfied, then (3.2.2) *implies that C_R is convex.*

Proof. In view of Proposition 2.3.3, the conclusion will follow if we show that ψ_i, $i \in I$, are regular in the sense of Clarke, for any $z \in C_R$ the nondegeneracy condition is satisfied, and the Slater's constraint qualification holds. The first and the second requirements will follow from Lemma 2.2.43 together with Lemma 2.2.44 that for each $z \in C_R$ and $i \in I_1(z)$, one has $\mathcal{V}_i^0(z) = \mathcal{V}_i(z)$,

$$\psi_i'(z; d) = \psi_i^o(z; d) = \max\{g_{ix}^o(z, v_i; d) : v_i \in \mathcal{V}_i(z)\}, \ \forall d \in \mathbb{R}^n,$$

and

$$0 \in \bigcap_{v_i \in \mathcal{V}_i^0} \mathbb{R}^n \backslash \Big(\partial_x^o g_i(z, v_i) \Big) = \mathbb{R}^n \backslash \Big(\bigcup_{v_i \in \mathcal{V}_i(z)} \partial_x^o g_i(z, v_i) \Big) = \mathbb{R}^n \backslash \partial^o \psi_i(z).$$

Finally, the robust Slater-type constraint qualification (RSCQ) leads us to the following strict inequality

$$\psi_i(x_0) = \max\{g_i(x_0, v_i) : v_i \in \mathcal{V}_i\} < 0, \ \forall i \in I \text{ for some } x_0 \in \mathbb{R}^n,$$

which means that the system $x \in \mathbb{R}^n$, $\psi_i(x) \leq 0$ $(i \in I)$ satisfies the Slater's constraint qualification. Taking into account Proposition 2.3.3, the proof is complete. □

Remark 3.2.3. In view of (3.2.2), we easily obtain that if C_R is convex then

$$C_R \subseteq \{z \in \mathbb{R}^n : g^o_{ix}(z, v_i; x - z) \leq 0,\ \forall x \in C_R,\ \forall i \in I_1(z),\ \forall v_i \in \mathcal{V}^0_i(z)\}.$$

Furthermore, for every $x \in C_R$ one has

$$\partial^o_x g_i(x, v_i) \subseteq N(C_R, x) \text{ whenever } i \in I_1(x) \text{ and } v_i \in \mathcal{V}^0_i(x).$$

Now the following theorem declares a result that the robust basic constraint qualification defined in Definition 3.2.1(iii) is a necessary and sufficient constraint qualification of a robust optimal solution for the given problem, that is, the robust basic constraint qualification holds if and only if the Lagrange multiplier conditions are satisfied for a robust optimal solution.

Theorem 3.2.4. [76, **Characterizing the robust basic constraint qualification**] *The following statements are equivalent:*

(i) *the robust basic constraint qualification holds at $\bar{x} \in C_R$;*

(ii) *for each real-valued convex-concave function f on $\mathbb{R}^n \times \mathcal{U}$, the following statements are equivalent:*

  (a) $\max_{u\in\mathcal{U}} f(x, u) \geq \max_{u\in\mathcal{U}} f(\bar{x}, u)$ *for all* $x \in C_R$,

  (b) *there exist* $\bar{u} \in \mathcal{U}$, $\bar{\lambda}_i \geq 0$, *and* $\bar{v}_i \in \mathcal{V}_i$, $i \in I$ *such that*

$$0 \in \partial_x f(\bar{x}, \bar{u}) + \sum_{i\in I} \bar{\lambda}_i \partial^o_x g_i(\bar{x}, \bar{v}_i),\ \bar{\lambda}_i g_i(\bar{x}, \bar{v}_i) = 0,\ \forall i \in I, \tag{3.2.3}$$

  and

$$f(\bar{x}, \bar{u}) = \max_{u\in\mathcal{U}} f(\bar{x}, u). \tag{3.2.4}$$

Proof. [(i) $\Rightarrow$ (ii)] Suppose that (i) holds. Let f be a real-valued convex-concave function on $\mathbb{R}^n \times \mathcal{U}$. Firstly, we assume that (a) holds. Then, $\bar{x}$ is a solution of the following constrained convex optimization problem:

$$\min_{x\in\mathbb{R}^n} \left\{ \max_{u\in\mathcal{U}} f(x, u) : x \in C_R \right\},$$

which can be equivalently expressed as,

$$0 \in \partial(\max_{u\in\mathcal{U}} f(\cdot, u))(\bar{x}) + N(C_R, \bar{x}).$$

By (i), there are $\bar{\lambda}_i \geq 0$, and $\bar{v}_i \in \mathcal{V}_i$, $i \in I$ such that

$$0 \in \partial \left(\max_{u \in \mathcal{U}} f(\cdot, u) \right)(\bar{x}) + \sum_{i \in I} \bar{\lambda}_i \partial_x^o g_i(\bar{x}, \bar{v}_i) \text{ and } \bar{\lambda}_i g_i(\bar{x}, \bar{v}_i) = 0, \ \forall i \in I.$$

Then, it follows from Lemma 2.2.33 that there exists $\bar{u} \in \mathcal{U}$ such that (3.2.3) and (3.2.4) hold.

To prove sufficiency, assume that there exist $\bar{u} \in \mathcal{U}$, $\bar{\lambda}_i \geq 0$, and $\bar{v}_i \in \mathcal{V}_i$, $i \in I$ such that (3.2.3) and (3.2.4) hold. According to (3.2.3), we can find $\xi \in \partial f_x(\bar{x}, \bar{u})$ and $\eta_i \in \partial_x^o g_i(\bar{x}, \bar{v}_i)$, $i \in I$, such that

$$\xi + \sum_{i \in I} \bar{\lambda}_i \eta_i = 0. \tag{3.2.5}$$

It stems from $\xi \in \partial_x f(\bar{x}, \bar{u})$ and $\eta_i \in \partial_x^o g_i(\bar{x}, \bar{v}_i)$, $i \in I$, we get

$$f(x, \bar{u}) - f(\bar{x}, \bar{u}) \geq \langle \xi, x - \bar{x} \rangle \tag{3.2.6}$$

and

$$g_{ix}^o(\bar{x}, \bar{v}_i; x - \bar{x}) \geq \langle \eta_i, x - \bar{x} \rangle, \ \forall i \in I, \tag{3.2.7}$$

for any $x \in \mathbb{R}^n$. Multiplying each of inequalities in (3.2.7) by $\bar{\lambda}_i$ and summing up the obtained inequalities with (3.2.6), we obtain that, for all $x \in \mathbb{R}^n$,

$$f(x, \bar{u}) - f(\bar{x}, \bar{u}) + \sum_{i \in I} \bar{\lambda}_i g_{ix}^o(\bar{x}, \bar{v}_i; x - \bar{x}) \geq \left\langle \xi + \sum_{i \in I} \bar{\lambda}_i \eta_i, x - \bar{x} \right\rangle.$$

Taking (3.2.5) into account together with the condition $\bar{\lambda}_i g_i(\bar{x}, \bar{v}_i) = 0, \ i \in I$, we deduce

$$f(x, \bar{u}) - f(\bar{x}, \bar{u}) + \sum_{i \in I} \bar{\lambda}_i g_{ix}^o(\bar{x}, \bar{v}_i; x - \bar{x}) \geq 0, \ \forall x \in \mathbb{R}^n. \tag{3.2.8}$$

If $\bar{\lambda}_i = 0$ for all $i \in I$, the inequality (3.2.8), in particular, becomes

$$f(x, \bar{u}) - f(\bar{x}, \bar{u}) \geq 0, \ \forall x \in C_R. \tag{3.2.9}$$

Thus, together with $\max_{u \in \mathcal{U}} f(x, u) \geq f(x, \bar{u})$ for all $x \in \mathbb{R}^n$ and (3.2.4), we obtain

$$\max_{u \in \mathcal{U}} f(x, u) - \max_{u \in \mathcal{U}} f(\bar{x}, u) \geq 0, \ \forall x \in C_R.$$

It means that $\bar{x}$ is a robust optimal solution of problem (UP).

Otherwise, suppose that $\tilde{I} := \{i \in I : \bar{\lambda}_i > 0\} \neq \emptyset$. Let $x \in C_R$ be arbitrary. As $\bar{\lambda}_i g_i(\bar{x}, \bar{v}_i) = 0,\ i \in I$, we have $\tilde{I} \subseteq I_1(\bar{x})$ and so, $\bar{v}_i \in \mathcal{V}_i^0(\bar{x})$ for all $i \in \tilde{I}$. Thus, by Lemma 3.2.2, $g_{ix}^o(\bar{x}, \bar{v}_i; x - \bar{x}) \leq 0,\ \forall i \in \tilde{I}$. It then follows from (3.2.8) that

$$\begin{aligned} f(x,\bar{u}) - f(\bar{x},\bar{u}) &\geq f(x,\bar{u}) - f(\bar{x},\bar{u}) + \sum_{i\in\tilde{I}} \bar{\lambda}_i g_{ix}^o(\bar{x},\bar{v}_i; x-\bar{x}) \\ &= f(x,\bar{u}) - f(\bar{x},\bar{u}) + \sum_{i\in I} \bar{\lambda}_i g_{ix}^o(\bar{x},\bar{v}_i; x-\bar{x}) \\ &\geq 0, \end{aligned}$$

showing that (3.2.9) holds, and consequently, $\bar{x}$ is a robust optimal solution of problem (UP) as well.

[(ii) $\Rightarrow$ (i)] The proof is similar to the one in [44, Theorem 3.1], and so is omitted. □

In the uncertainty free case, we can easily obtain the following result, which was obtained by Yamamoto and Kuroiwa in [49].

Corollary 3.2.5. [49, Theorem 3.2] *Let $\bar{x} \in C' := \{x \in \mathbb{R}^n : g_i(x) \leq 0,\ \forall i \in I\}$ be a feasible solution, $g_i : \mathbb{R}^n \to \mathbb{R},\ i \in I$, be locally Lipschitz on $\mathbb{R}^n$. Assume further that for any $x \in C'$ and any $i \in I$ such that $g_i(x) = 0$, the function g_i is regular, and C' is convex. Then the following statement are equivalent:*

(i) $N(C', \bar{x}) = \bigcup_{\substack{\lambda_i \geq 0 \\ \lambda_i g_i(\bar{x})=0,\ i\in I}} \sum_{i\in I} \lambda_i \partial^o g_i(\bar{x})$;

(ii) *for each real-valued convex function f on $\mathbb{R}^n$, the following statements are equivalent:*

(a) $f(x) \geq f(\bar{x})$ *for all* $x \in C'$;

(b) *there exist* $\bar{\lambda}_i \geq 0$, $i \in I$ *such that*

$$0 \in \partial f(\bar{x}) + \sum_{i\in I} \bar{\lambda}_i \partial^o g_i(\bar{x}) \text{ and } \bar{\lambda}_i g_i(\bar{x}) = 0,\ \forall i \in I.$$

Remark 3.2.6. Both the robust Slater-type constraint qualification condition and robust non-degeneracy condition at $\bar{x}$ is a sufficient condition for guaranteeing the robust basic constraint qualification holds at $\bar{x}$. Indeed, according to Remark 3.2.3, we only have to show that

$$N(C_R, \bar{x}) \subseteq \bigcup_{\substack{\lambda_i \geq 0,\ v_i \in \mathcal{V}_i \\ \lambda_i g_i(\bar{x}, v_i) = 0,\ i \in I}} \sum_{i \in I} \lambda_i \partial_x^o g_i(\bar{x}, v_i).$$

Let $\eta \in N(C_R, \bar{x})$ be arbitrarily. Since the robust Slater-type constraint qualification condition and robust non-degeneracy are satisfied at $\bar{x}$, by Theorem 2.3.4 with $f := \langle -\eta, \cdot \rangle$, and $g_i := \psi_i$, $i \in I$, there exist $\bar{\lambda}_i \geq 0$, $i \in I$, such that

$$0 \in \{-\eta\} + \sum_{i \in I} \bar{\lambda}_i \partial^o \psi_i(\bar{x}) \text{ and } \bar{\lambda}_i \psi_i(\bar{x}) = 0,\ \forall i \in I.$$

It then follows from Lemma 2.2.44 that there exist $\bar{v}_i \in \mathcal{V}_i(\bar{x}) \subseteq \mathcal{V}_i$, $i \in I$, such that

$$\eta \in \sum_{i \in I} \bar{\lambda}_i \partial_x^o g_i(\bar{x}, \bar{v}_i) \text{ and } \bar{\lambda}_i g_i(\bar{x}, \bar{v}_i) = \bar{\lambda}_i \psi_i(\bar{x}) = 0,\ \forall i \in I.$$

This shows that

$$\eta \in \bigcup_{\substack{\lambda_i \geq 0,\ v_i \in \mathcal{V}_i \\ \lambda_i g_i(\bar{x}, v_i) = 0,\ i \in I}} \sum_{i \in I} \lambda_i \partial_x^o g_i(\bar{x}, v_i),$$

the result as require.

The following example is given to illustrate the condition (i) of Theorem 3.2.4 is essential.

Example 3.2.7. Let $x := (x_1, x_2) \in \mathbb{R}^2$, $v_1 := (v_{1,1}, v_{1,2})$, $v_2 := (v_{2,1}, v_{2,2})$, $v_3 := (v_{3,1}, v_{3,2})$, $\mathcal{V}_1 := \{(v_1, v_2) \in \mathbb{R}^2 : v_1^2 + v_2^2 \leq 1\}$, $\mathcal{V}_2 := [0,1] \times [1,2]$, $\mathcal{V}_3 := [0,1] \times [0,1]$,

$$\begin{aligned}
g_1(x, v_1) &:= v_{1,1} x_1 + v_{1,2} x_2 - x_1^3 - 2, \\
g_2(x, v_2) &:= -v_{2,1} x_1^3 + v_{2,2} \max\{-x_2, -x_2^3\}, \\
g_3(x, v_3) &:= v_{3,1} x_1 - v_{3,2} x_2^2, \\
C_R &:= \{x \in \mathbb{R}^2 : g_i(x, v_i) \leq 0,\ \forall v_i \in \mathcal{V}_i,\ i \in I := \{1,2,3\}\},
\end{aligned}$$

and $\bar{x} := (0,0)$. Then $C_R = \{x \in \mathbb{R}^2 : \sqrt{x_1^2 + x_2^2} - x_1^3 - 2 \leq 0,\ -x_1 - x_2 \leq 0,\ x_1 \leq 0\}$, $I(\bar{x}) = \{2,3\}$, $\partial_x^o g_2(\bar{x}, v_2) = \{0\} \times [-v_{2,2}, 0]$ and $\partial_x^o g_3(\bar{x}, v_3) = \{(v_{3,1}, 0)\}$. It can be observed that

$$N(C_R, \bar{x}) = \text{cone co}\{(-1,-1), (1,0)\}$$

and

$$\bigcup_{\substack{\lambda_i \geq 0,\ v_i \in \mathcal{V}_i \\ \lambda_i g_i(\bar{x}, v_i) = 0,\ i \in I}} \sum_{i \in I} \lambda_i \partial_x^\circ g_i(\bar{x}, v_i) = \text{cone co}\{(0,-1),(1,0)\}.$$

Hence, we have the condition (i) of Theorem 3.2.4 does not hold. Thus for some convex-concave function $f : \mathbb{R}^2 \times \mathcal{U} \to \mathbb{R}$, it may happen that the KKT optimality conditions may go awry at $\bar{x}$ even if it is a robust optimal solution for an uncertain problem,

$$\min_{x \in \mathbb{R}^2} \{f(x,u) : g_i(x, v_i) \leq 0,\ i \in I\},$$

where $u \in \mathcal{U}$ and $v_i \in \mathcal{V}_i$, $i \in I$. Actually, let $u := (u_1, u_2)$ be an uncertain parameter belong to uncertainty set $\mathcal{U} := \{(u_1, u_2) \in \mathbb{R}^2 : u_1^2 + u_2^2 \leq 1\}$, and $f(x,u) := e^{x_1} - u_1 x_1 - u_2 x_2 + x_1 + x_2$. We see that $\max_{u \in \mathcal{U}} f(x,u) = e^{x_1} + \sqrt{x_1^2 + x_1^2} + x_1 + x_2$ and

$$\max_{u \in \mathcal{U}} f(x,u) \geq e^{x_1} - x_1 \geq 1 = \max_{u \in \mathcal{U}} f(\bar{x}, u) \text{ for all } x \in C_R.$$

However, in this case, we cannot find out $u \in \mathcal{U}$, $v_i \in \mathcal{V}_i$ and $\lambda_i \geq 0$, $i \in I$, such that (3.2.3) and (3.2.4) hold. In fact, for $u \in \mathcal{U}$, $v_i \in \mathcal{V}_i$, $\lambda_i \geq 0$, such that $\lambda_i g_i(\bar{x}, v_i) = 0$, $i \in I$, $\max_{u \in \mathcal{U}} f(\bar{x}, u) = f(\bar{x}, u)$, and

$$\begin{aligned}(0,0) &\in \partial_x f(\bar{x}, u) + \sum_{i \in I} \lambda_i \partial_x^\circ g_i(\bar{x}, v_i) \\ &= \{(2 - u_1, 1 - u_2)\} + \lambda_1\{(v_{1,1}, v_{1,2})\} + \lambda_2(\{0\} \times [-v_{2,2}, 0]) + \lambda_3\{(v_{3,1}, 0)\},\end{aligned}$$

it then follows that $\lambda_1 = 0$, and so, $2 - u_1 + \lambda_3 v_{3,1} = 0$, a contradiction.

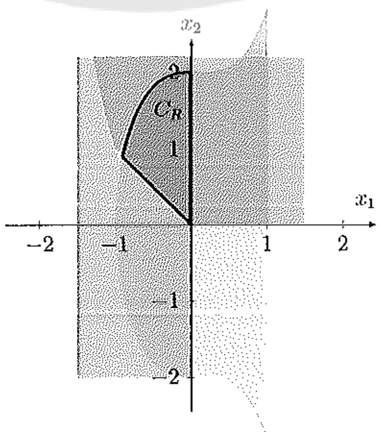


Figure 24: Illustration of the robust feasible set C_R in Example 3.2.7.

Remark 3.2.8. According to Remark 3.2.6, Example 3.2.7 demonstrates that only robust Slater-type constraint qualification condition is not sufficient to ensure the robust basic constraint qualification holds at consideration point. The reason is that the robust non-degeneracy condition at such a point is destroyed.

Next, we will establish some characterizations of robust optimal solution set in terms of a given robust solution point of the given problem. We begin by recalling the following constrained convex optimization problem in the face of data uncertainty (UP):

$$\min_{x\in\mathbb{R}^n}\{f(x,u) : g_i(x,v_i) \leq 0,\ i \in I\}, \tag{UP}$$

where $f : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}$ is a convex-concave function, the functions g_i, $i \in I$, satisfy the condition (A1)-(A4), $g_i(x,\cdot) : \mathcal{V}_i \to \mathbb{R}$, $i \in I$, are concave functions for any $x \in \mathbb{R}^n$, and the robust feasible set C_R is convex. Assume that the robust solution set of the problem (UP), denoted by

$$S_R := \left\{a \in C_R : \max_{u\in\mathcal{U}} f(a,u) \leq \max_{u\in\mathcal{U}} f(x,u),\ \forall x \in C_R\right\},$$

is nonempty. In what follows, for any given $y \in \mathbb{R}^n$, $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{R}^m_+$, $u \in \mathcal{U}$, $v_i \in \mathcal{V}_i$, $i \in I$ and $v := (v_1, v_2, \ldots, v_m)$, we introduce the so-called *pseudo Lagrangian-type function* $\mathcal{L}^P(\cdot, y, \lambda, u, v)$ [76] by, for all $x \in \mathbb{R}^n$,

$$\mathcal{L}^P(x,y,\lambda,u,v) := f(x,u) + \sum_{i\in I_1(y)} \lambda_i g^o_{ix}(y, v_i; x-y).$$

Now, we show that the pseudo Lagrangian-type function associated with a Lagrange multiplier vector and uncertainty parameters according to a solution is constant on S.

Proposition 3.2.9. [76] *Let $a \in S_R$ be a robust optimal solution fulfilling the robust basic constraint qualification. Then there exist a Lagrange multiplier vector $\lambda^a := (\lambda^a_1, \lambda^a_2, \ldots, \lambda^a_m) \in \mathbb{R}^m_+$, and uncertainty parameters $u^a \in \mathcal{U}$, $v^a_i \in \mathcal{V}_i$, $i \in I$, such that for any $x \in S_R$,*

$$\lambda^a_i g^o_{ix}(a, v^a_i; x-a) = 0,\ \forall i \in I_1(a),$$

$f(x,u^a) = \max_{u\in\mathcal{U}} f(x,u)$, and $\mathcal{L}^P(\cdot, a, \lambda^a, u^a, v^a)$ is constant on S_R.

Proof. It follows from $a \in S_R$ and Theorem 3.2.4 that there exist a Lagrange multiplier vector $\lambda^a := (\lambda_1^a, \lambda_2^a, \ldots, \lambda_m^a) \in \mathbb{R}_+^m$, and uncertainty parameters $u^a \in \mathcal{U}$, $v_i^a \in \mathcal{V}_i$, $i \in I$, satisfying the conditions (3.2.3) and (3.2.4). Then, it stems from the fact that $\partial_x^o g_i(a, v_i^a) = \partial g_{ix}^o(a, v_i^a; \cdot - a)(a)$ for all $i \in I_1(a)$ and (3.2.3), we get

$$0 \in \partial_x f(a, u^a) + \sum_{i \in I_1(a)} \lambda_i^a \partial g_{ix}^o(a, v_i^a; \cdot - a)(a) \subseteq \partial \mathcal{L}^P(\cdot, a, \lambda^a, u^a, v^a)(a),$$

which is noting else than

$$\begin{aligned} \max_{u \in \mathcal{U}} f(x, u) + \sum_{i \in I_1(a)} \lambda_i^a g_{ix}^o(a, v_i^a; x - a) &\geq f(x, u^a) + \sum_{i \in I_1(a)} \lambda_i^a g_{ix}^o(a, v_i^a; x - a) \\ &\geq f(a, u^a) \\ &= \max_{u \in \mathcal{U}} f(a, u), \ \forall x \in \mathbb{R}^n. \qquad (3.2.10) \end{aligned}$$

Notice that

$$\max_{u \in \mathcal{U}} f(x, u) = \max_{u \in \mathcal{U}} f(a, u), \text{ for any } a \in S_R \text{ and } x \in S_R, \qquad (3.2.11)$$

and taking this into account, (3.2.10) deduces

$$\sum_{i \in I_1(a)} \lambda_i^a g_{ix}^o(a, v_i^a; x - a) \geq 0, \text{ for any } x \in S_R.$$

Let us notice that for indices $i \in I_1(a)$ such that $\lambda_i^a > 0$, we have $g_i(a, v_i^a) = 0$, and consequently, $v_i^a \in \mathcal{V}_i^0(a)$. This in turn, by Remark 3.2.3, implies that for every $x \in S_R \subseteq C_R$,

$$\lambda_i^a g_{ix}^o(a, v_i^a; x - a) = 0, \ \forall i \in I_1(a). \qquad (3.2.12)$$

Now, we prove that

$$f(x, u^a) = \max_{u \in \mathcal{U}} f(x, u), \text{ for any } x \in S_R. \qquad (3.2.13)$$

In fact, by (3.2.10) and (3.2.12), we get the assertion

$$\max_{u \in \mathcal{U}} f(x, u) \geq f(x, u^a) \geq \max_{u \in \mathcal{U}} f(a, u).$$

This together with (3.2.11), (3.2.13) holds. Therefore, for any $x \in S_R$, (3.2.4), (3.2.11), (3.2.12) and (3.2.13) entail

$$\mathcal{L}^P(x, a, \lambda^a, u^a, v^a) = f(x, u^a) + \sum_{i \in I_1(a)} \lambda_i^a g_{ix}^o(a, v_i^a; x - a)$$

$$= \quad f(x,a) = \max_{u\in\mathcal{U}} f(x,u) = \max_{u\in\mathcal{U}} f(a,u) = f(a,u^a),$$

showing that $\mathcal{L}^P(\cdot, a, \lambda^a, u^a, v^a)$ is constant on S_R, and this completes the proof. □

Remark 3.2.10. It is worth noting that if $g_i(\cdot, v_i)$, $i \in I$, are convex functions for any $v_i \in \mathcal{V}_i$ then, for each $i \in I$, Proposition 3.2.9 gives

$$\lambda_i^a g_i(x, v_i^a) - \lambda_i^a g_i(a, v_i^a) \geq \lambda_i^a g_i'(a, v_i^a; x-a) = \lambda_i^a g_i^o(a, v_i^a; x-a) = 0 \text{ for any } x \in S_R.$$

This together with $x \in C_R$ and $\lambda_i^a g_i(a, v_i^a) = 0$, $i \in I$, arrives $\lambda_i^a g_i(x, v_i^a) = 0$, $i \in I$. Furthermore, it yields

$$\begin{aligned}\mathcal{L}^P(x, a, \lambda^a, u^a, v^a) &= f(x, u^a) + \sum_{i\in I_1(a)} \lambda_i^a g_{ix}^o(a, v_i^a; x-a) \\ &= f(x, u^a) \\ &= f(x, u^a) + \sum_{i=1}^{m} \lambda_i^a g_i(x, v_i^a), \ \forall x \in S_R.\end{aligned}$$

This shows that pseudo Lagrangian-type function collapses to the well-known Lagrangian-type function on the robust solution set S_R.

In the sequel, we are now in a position to establish the characterizations of the robust solution set for problem (UP) in terms of convex subdifferentials, Clarke subdifferentials and Lagrange multipliers. But before doing so it will thus be convenient to denote the following:

$$\begin{aligned}&\tilde{I}_1(a, \lambda^a) := \{i \in I_1(a) : \lambda_i^a > 0\},\\ &X_1(\lambda^a, v^a) := \{x \in C_R : \forall i \in \tilde{I}_1(a, \lambda^a),\ \exists \eta_i \in \partial_x^o g_i(a, v_i^a) \text{ s.t. } \langle \eta_i, x-a\rangle = 0\}\\ &X_1(\lambda^a, u^a, v^a) := \{x \in X_1(\lambda^a, v^a) : f(x, u^a) = \max_{u\in\mathcal{U}} f(x,u)\}\\ &\Gamma(x, u^a) := \{\xi \in \partial_x f(a, u^a) : \langle \xi, x-a\rangle \geq 0\} \text{ for any given } x \in C_R.\end{aligned}$$

Theorem 3.2.11. [76, **Characterizing the robust solution set**] *Let $a \in S_R$ be a robust optimal solution fulfilling the robust basic constraint qualification. Then there exist a Lagrange multiplier vector $\lambda^a := (\lambda_1^a, \ldots, \lambda_m^a) \in \mathbb{R}_+^m$, and uncertainty parameters $u^a \in \mathcal{U}$, $v_i^a \in \mathcal{V}_i$, $i \in I$, such that the robust solution set for the problem* (UP) *is characterized by*

$$S_R = S_1 = S_2 = S_3 = S_4 = S_5 = S_6 = S_7,$$

where

$$S_1 := \{x \in X_1(\lambda^a, u^a, v^a) : \exists \zeta \in \partial_x f(x, u^a) \cap \partial_x f(a, u^a) \ s.t. \ \langle \zeta, a - x \rangle = 0\},$$

$$S_2 := \{x \in X_1(\lambda^a, u^a, v^a) : \exists \zeta \in \partial_x f(x, u^a) \cap \partial_x f(a, u^a) \ s.t. \ \langle \zeta, a - x \rangle \geq 0\},$$

$$S_3 := \{x \in X_1(\lambda^a, u^a, v^a) : \exists \zeta \in \partial_x f(x, u^a), \ \exists \xi \in \Gamma(x, u^a) \ s.t. \ \langle \xi, x - a \rangle = \langle \zeta, a - x \rangle = 0\},$$

$$S_4 := \{x \in X_1(\lambda^a, u^a, v^a) : \exists \zeta \in \partial_x f(x, u^a), \ \exists \xi \in \Gamma(x, u^a) \ s.t. \ \langle \xi, x - a \rangle = \langle \zeta, a - x \rangle\},$$

$$S_5 := \{x \in X_1(\lambda^a, u^a, v^a) : \exists \zeta \in \partial_x f(x, u^a), \ \exists \xi \in \Gamma(x, u^a) \ s.t. \ \langle \xi, x - a \rangle \leq \langle \zeta, a - x \rangle\},$$

$$S_6 := \{x \in X_1(\lambda^a, u^a, v^a) : \exists \zeta \in \partial_x f(x, u^a) \ s.t. \ \langle \zeta, a - x \rangle = 0\},$$

$$S_7 := \{x \in X_1(\lambda^a, u^a, v^a) : \exists \zeta \in \partial_x f(x, u^a) \ s.t. \ \langle \zeta, a - x \rangle \geq 0\}.$$

Proof. Evidently, the following containments hold:

$$S_1 \subseteq S_2 \subseteq S_7$$

$$S_1 \subseteq S_6 \subseteq S_7,$$

$$S_1 \subseteq S_3 \subseteq S_4 \subseteq S_5 \subseteq S_7.$$

Hence, we only have to show that $S_R \subseteq S_1$ and $S_7 \subseteq S_R$. In order to establish $S_R \subseteq S_1$, let $x \in S_R$ be arbitrarily given. It follows from (3.2.3), we therefore obtain vectors $\zeta \in \partial_x f(a, u^a)$ and $\xi_i \in \partial_x^o g_i(a, v_i^a)$, $i \in I_1(a)$, such that

$$\zeta + \sum_{i \in I_1(a)} \lambda_i^a \xi_i = 0 \tag{3.2.14}$$

(since $\lambda_i^a = 0$ for $i \notin I_1(a)$). According to $\zeta \in \partial_x f(a, u^a)$, $\xi_i \in \partial_x^o g_i(a, v_i^a)$, $i \in I_1(a)$, and $x, a \in S_R$, one has

$$f(x, u^a) - f(a, u^a) \geq \langle \zeta, x - a \rangle \tag{3.2.15}$$

and

$$g_{ix}^o(a, v_i^a; x - a) \geq \langle \xi_i, x - a \rangle, \ \forall i \in I_1(a). \tag{3.2.16}$$

Once we have shown, in Proposition 3.2.9, that $\lambda_i^a g_{ix}^o(a, v_i^a; x - a) = 0$, $\forall i \in I_1(a)$, after multiplying both sides of (3.2.16) by λ_i^a, $i \in I_1(a)$ we get

$$0 \geq \langle \lambda_i^a \xi_i, x - a \rangle, \ \forall i \in I_1(a).$$

Summing up these inequalities and using (3.2.14) we obtain that

$$0 \geq \left\langle \sum_{i \in I(a)} \lambda_i^a \xi_i, x - a \right\rangle = \langle -\zeta, x - a \rangle. \tag{3.2.17}$$

Again, it follows from Proposition 3.2.9 that

$$f(x, u^a) = \max_{u \in \mathcal{U}} f(x, u), \tag{3.2.18}$$

and for each $i \in \tilde{I}_1(a)$, $\max_{\eta_i \in \partial_x^o g_i(a, v_i^a)} \langle \eta_i, x - a \rangle = g_{ix}^o(a, v_i^a; x - a) = 0$, the latter which in turn leads to there exists $\eta_i \in \partial_x^o g_i(a, v_i^a)$ such that

$$\langle \eta_i, x - a \rangle = 0.$$

On the one hand, taking (3.2.4) and (3.2.18) into account (3.2.15) we obtain

$$\langle \zeta, x - a \rangle \leq f(x, u^a) - f(a, u^a) = \max_{u \in \mathcal{U}} f(x, u) - \max_{u \in \mathcal{U}} f(a, u) = 0.$$

This together with (3.2.17) arrives at

$$\langle \zeta, x - a \rangle = 0.$$

Now, we only need to prove that $\zeta \in \partial_x f(x, u^a)$. In fact, for any $y \in \mathbb{R}^n$,

$$\begin{aligned} f(y, u^a) - f(x, u^a) &= f(y, u^a) - f(a, u^a) \\ &\geq \langle \zeta, y - a \rangle \\ &= \langle \zeta, y - x \rangle + \langle \zeta, x - a \rangle = \langle \zeta, y - x \rangle, \end{aligned}$$

which means $\zeta \in \partial_x f(x, u^a)$ and so, $x \in S_1$. This proves $S_R \subseteq S_1$.

To obtain $S_7 \subseteq S_R$, we now let x be arbitrary point of S_7. It follows that $x \in C_R$, and it is easy to see that

$$\max_{u \in \mathcal{U}} f(a, u) - \max_{u \in \mathcal{U}} f(x, u) = f(a, u^a) - f(x, u^a) \geq \langle \zeta, a - x \rangle \geq 0.$$

The last inequality together with the fact that $a \in S_R$ gives $x \in S_R$, and the proof is complete. □

Now, we give the following example to illustrate the significance of Theorem 3.2.11 that at least one of the constraint functions $g_i(\cdot, v_i)$ for some $v_i \in \mathcal{V}_i$, is not convex while the robust feasible set is convex. Then the results in [41, 42, 44, 45] may not be relevant to this example.

Example 3.2.12. Let us denote $x := (x_1, x_2) \in \mathbb{R}^2$, $u := (u_1, u_2)$, $v_1 := (v_{1,1}, v_{1,2})$, $v_2 := (v_{2,1}, v_{2,2})$, $v_3 := (v_{3,1}, v_{3,2})$, $\mathcal{U} := \{(u_1, u_2) \in \mathbb{R}^2 : u_1^2 + u_2^2 \leq 1\}$, $\mathcal{V}_1 := \{(v_1, v_2) \in \mathbb{R}^2 : v_1^2 + v_2^2 \leq 1\}$, $\mathcal{V}_2 = \mathcal{V}_3 := [0,1] \times [0,1]$. Consider the following constrained optimization problem with uncertainty data (UP):

$$\min_{x \in \mathbb{R}^2} \{f(x,u) : g_i(x, v_i) \leq 0,\ i \in I := \{1,2,3\}\}, \tag{UP}$$

where $u \in \mathcal{U}$, $v_i \in \mathcal{V}_i$, $i \in I$,

$$\begin{aligned}
f(x,u) &:= u_1 x_1 + u_2 x_2 - x_1 - x_2,\\
g_1(x, v_1) &:= v_{1,1} x_1 + v_{1,2} x_2 - x_1^3 - 2,\\
g_2(x, v_2) &:= v_{2,1} \max\{-x_1, -x_1^3\} - v_{2,2} x_2,\\
g_3(x, v_3) &:= v_{3,1} x_1 - v_{3,2} x_2^2.
\end{aligned}$$

A robust solution of (UP) is obtained by solving its robust (worst-case) counterpart (RP)

$$\min_{x \in \mathbb{R}^2} \left\{ \max_{u \in \mathcal{U}} f(x,u) : x \in C_R \right\}. \tag{RP}$$

Then $C_R = \{x \in \mathbb{R}^2 : \sqrt{x_1^2 + x_2^2} - x_1^3 - 2 \leq 0\} \cap \mathbb{R}^2_+ \cap \{x \in \mathbb{R}^2 : x_1 \leq 2\} = \{x \in \mathbb{R}^2 : x_1 = 0,\ 0 \leq x_2 \leq 2\}$. Evidently, the function $f : \mathbb{R}^2 \times \mathcal{U} \to \mathbb{R}$ is a convex-concave function. Let us notice that

$$\max_{u \in \mathcal{U}} f(x,u) = \sqrt{x_1^2 + x_2^2} - x_1 - x_2, \text{ for all } x \in \mathbb{R}^2,$$

and

$$\max_{u \in \mathcal{U}} f(x,u) \geq |x_2| - x_2 = 0 = \max_{u \in \mathcal{U}} f((0,0), u), \text{ for all } x \in C_R.$$

Thus $a := (a_1, a_2) = (0,0) \in S_R$, $I_1(a) = \{2,3\}$, $\partial_x^o g_2(a, v_2) = \{(r, -v_{2,2}) : r \in [-v_{2,1}, 0]\}$ for each $v_2 \in \mathcal{V}_2$ and $\partial_x^o g_3(a, v_3) = \{(v_{3,1}, 0)\}$ for each $v_3 \in \mathcal{V}_3$. So,

$$N(C_R, a) = \text{cone co}\{(-1,0), (0,-1), (1,0)\} = \bigcup_{\substack{\lambda_i \geq 0,\ v_i \in \mathcal{V}_i \\ \lambda_i g_i(a, v_i) = 0,\ i \in I}} \sum_{i \in I} \lambda_i \partial_x^o g_i(a, v_i),$$

which means that the robust basic constraint qualification holds at a. Also, for each $u \in \mathcal{U}$, the convex subdifferential of $f(\cdot, u)$ at any point x is given by

$$\partial_x f(x,u) = \{(u_1 - 1, u_2 - 1)\}.$$

Let us select $\lambda^a := (\lambda_1^a, \lambda_2^a, \lambda_3^a) = (0,0,1)$, $u^a := (0,1)$, $v_2^a := (1,1)$ and $v_3^a := (1,0)$. Therefore, $\tilde{I}_1(a) = \{3\}$ and by solving the following system, for $x \in \mathbb{R}^2$

$$\begin{cases} x_1 = \langle (1,0), (x_1,x_2) \rangle = 0, \\ \sqrt{x_2^2} - 2 \leq 0, \\ -x_2 \leq 0, \\ \partial_x f(x,u^a) \cap \partial_x f(a,u^a) = \{(-1,0)\}, \\ \langle (-1,0), (0,x_2) \rangle = 0, \\ -x_1 = \max_{u \in \mathcal{U}} f(x,u), \end{cases}$$

the robust solution set can be described simply as

$$S_R = S_1 = \{x \in \mathbb{R}^2 : x_1 = 0,\ 0 \leq x_2 \leq 2\}.$$

With the help of Proposition 3.2.9, we see now how the robust solution set can be characterized in terms of pseudo Lagrangian-type function.

Proposition 3.2.13. [76] *Assume all conditions of Theorem 3.2.4 hold. Let $a \in S_R$ be a robust optimal solution fulfilling the robust basic constraint qualification. Then there exist a Lagrange multiplier vector $\lambda^a := (\lambda_1^a, \lambda_2^a, \ldots, \lambda_m^a) \in \mathbb{R}_+^m$, and uncertainty parameters $u^a \in \mathcal{U}$, $v_i^a \in \mathcal{V}_i$, $i \in I$, such that*

$$S_R = \{x \in X_1(\lambda^a, u^a, v^a) : 0 \in \partial \mathcal{L}^P(\cdot, a, \lambda^a, u^a, v^a)(x)\}.$$

Proof. It will thus be convenient to denote

$$S^{**} := \{x \in X_1(\lambda^a, u^a, v^a) : 0 \in \partial \mathcal{L}^P(\cdot, a, \lambda^a, u^a, v^a)(x)\}.$$

By Proposition 3.2.9, we have that for each $x \in S_R$, $\lambda_i^a g_{ix}^o(a, v_i^a; x-a) = 0$, $\forall i \in I_1(a)$, $f(x,u^a) = \max_{u \in \mathcal{U}} f(x,u)$, and $\mathcal{L}^P(\cdot, a, \lambda^a, u^a, v^a)$ is constant on S_R. It then follows from optimality condition (3.2.3) that

$$\begin{aligned} & \mathcal{L}^P(y,a,\lambda^a,u^a,v^a) - \mathcal{L}^P(x,a,\lambda^a,u^a,v^a) \\ = \ & \mathcal{L}^P(y,a,\lambda^a,u^a,v^a) - \mathcal{L}^P(a,a,\lambda^a,u^a,v^a) \\ \geq \ & 0, \ \forall y \in \mathbb{R}^n, \end{aligned}$$

and so, $S_R \subseteq S^{**}$. To obtain the converse inclusion, let $x \in S^{**}$ be given. Then, by the definition of S^{**}, $x \in C_R$, there exist $\eta_i \in \partial_x^o g_i(a, v_i^a)$, $\forall i \in \tilde{I}({}_1a)$, such that

$\langle \eta_i, x - a \rangle = 0, \ \forall i \in \tilde{I}_1(a),$

$$f(x, u^a) = \max_{u \in \mathcal{U}} f(x, u) \tag{3.2.19}$$

and

$$\begin{aligned} f(y, u^a) + \sum_{i \in I_1(a)} \lambda_i^a g_{ix}^o(a, v_i^a; y - a) &= \mathcal{L}^P(y, a, \lambda^a, u^a, v^a) \\ &\leq \mathcal{L}^P(x, a, \lambda^a, u^a, v^a) \\ &= f(x, u^a) + \sum_{i \in I_1(a)} \lambda_i^a g_{ix}^o(a, v_i^a; x - a) \\ &= f(x, u^a) + \sum_{i \in \tilde{I}_1(a)} \lambda_i^a g_{ix}^o(a, v_i^a; x - a) \\ &= f(x, u^a) \text{ for all } y \in \mathbb{R}^n. \end{aligned}$$

Using (3.2.19) and taking $y := a$ in the last inequality, we get that

$$\max_{u \in \mathcal{U}} f(x, u) \geq \max_{u \in \mathcal{U}} f(a, u) \geq f(a, u^a) \geq f(x, u^a) = \max_{u \in \mathcal{U}} f(x, u).$$

Hence, $\max_{u \in \mathcal{U}} f(a, u) = \max_{u \in \mathcal{U}} f(x, u)$, which is noting else than $x \in S_R$. □

3.3 Uncertain multi-objective programs

In this section, as an application of the general results of the previous section, we examine the class of multiple-objective programs in the face of data uncertainty both in the objective and constraints that can be written by the following multi-objective optimization problem:

$$\min_{x \in \mathbb{R}^n} \{(f_1(x, u_1), f_2(x, u_2), \ldots, f_p(x, u_p)) : g_i(x, v_i) \leq 0, \ i \in I\}, \tag{UMP}$$

where $f_j : \mathbb{R}^n \times \mathcal{U}_j \to \mathbb{R}$, $j \in J := \{1, 2, \ldots, p\}$, are convex-concave functions, $g_i : \mathbb{R}^n \times \mathcal{V}_i \to \mathbb{R}$, $i \in I := \{1, 2, \ldots, m\}$, are functions satisfying the condition (A1)-(A4), $g_i(x, \cdot)$ are concave functions for any $x \in \mathbb{R}^n$, and u_j and v_i are uncertain parameters and they belong to nonempty convex compact sets $\mathcal{U}_j \subseteq \mathbb{R}^{q_j}$ and $\mathcal{V}_i \subseteq \mathbb{R}^{q_i}$, respectively.

We associate with (UMP) its robust counterpart, which is the worst case of (UMP),

$$\min_{x\in\mathbb{R}^n}\left\{\left(\max_{u_1\in\mathcal{U}_1} f_1(x,u_1), \max_{u_2\in\mathcal{U}_2} f_2(x,u_2),\ldots,\max_{u_p\in\mathcal{U}_p} f_p(x,u_p)\right) : x\in C_R\right\}, \quad \text{(RMP)}$$

where C_R stands for the robust feasible set of (UMP), defined by

$$C_R := \{x\in\mathbb{R}^n : g_i(x,v_i)\le 0,\ \forall v_i\in\mathcal{V}_i,\ i\in I\}.$$

In the same way, we will give three kind robust solutions for the problems (UMP) which has been introduced in [77].

Definition 3.3.1. $\bar{x}\in C_R$ is said to be

(i) a **robust efficient solution** of (UMP) if there does not exist a robust feasible solution x of (UMP) such that

$$\max_{u_j\in\mathcal{U}_j} f_j(x,u_j)\le \max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j) \text{ for all } j\in J,$$

and

$$\max_{u_k\in\mathcal{U}_k} f_k(x,u_k) < \max_{u_k\in\mathcal{U}_k} f_k(\bar{x},u_k) \text{ for some } k\in J.$$

(ii) a **weakly robust efficient solution** of (UMP) if there does not exist a robust feasible solution x of (UMP) such that

$$\max_{u_j\in\mathcal{U}_j} f_j(x,u_j) < \max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j) \text{ for all } j\in J.$$

(iii) a **properly robust efficient solution** of (UMP) if it is a robust efficient solution of (UMP) and there is a number $M>0$ such that for all $j\in J$ and $x\in C_R$ satisfying $\max_{u_j\in\mathcal{U}_j} f_j(x,u_j) < \max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j)$, there exists an index $k\in J$ such that $\max_{u_k\in\mathcal{U}_k} f_k(\bar{x},u_k) < \max_{u_k\in\mathcal{U}_k} f_k(x,u_k)$ and

$$\frac{\max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j) - \max_{u_j\in\mathcal{U}_j} f_j(x,u_j)}{\max_{u_k\in\mathcal{U}_k} f_k(x,u_k) - \max_{u_l\in\mathcal{U}_k} f_k(\bar{x},u_k)} \le M.$$

In view of these definitions, it is evidently that $\bar{x}\in C_R$ is a robust efficient solution (resp. weakly, properly robust efficient solution) of (UMP) if and only if

$\bar{x} \in C_R$ is a efficient solution (resp. weakly, properly efficient solution) of (RMP). The search for an efficient solution (resp. weakly, properly efficient solution) to multi-objective optimization problem has been carried out through solving a single (scalar) or a family of single objective optimization problems, possibly depending on some appropriate parameters. We refer the reader to [55, 67, 68, 72] and other references therein for necessary and sufficient conditions for (weakly, properly) efficient solutions to a multiobjective optimization by parameterization and linear scalarization (weighted sum approach).

In this section, we present characterizations of weakly robust efficient solution set ($WR(C_R)$) and properly robust efficient solution set ($PR(C_R)$) of the problem (UMP) by using linear scalarization approach. Before presenting, in the cases of study, let us consider the following scalar convex problem of (RMP) depending on a parameter $\theta := (\theta_1, \theta_2, \ldots, \theta_p) \in \mathbb{R}^p_+$:

$$\min_{x \in \mathbb{R}^n} \left\{ \sum_{j \in J} \theta_j \max_{u_j \in \mathcal{U}_j} f_j(x, u_j) : x \in C_R \right\}. \qquad (RP_\theta)$$

Suppose that the solution set of problem (RP_θ), denoted by S^θ_R is nonempty. It is well-known, in the literature, that weakly efficient solutions and properly efficient solutions of (RMP) can be characterized by solving some scalar parameterized convex problems (RP_θ). More precisely,

(i) $\bar{x} \in WR(C_R)$ if and only if there exists $\theta \in \mathbb{R}^p_+ \backslash \{0\}$ such that $\bar{x} \in S^\theta_R$.

(ii) $\bar{x} \in PR(C_R)$ if and only if there exists $\theta \in \operatorname{int} \mathbb{R}^p_+$ such that $\bar{x} \in S^\theta_R$.

Thus, by using Theorem 3.2.4, we can obtain immediately the following necessary and sufficient optimality conditions for weakly robust efficient solution as well as properly robust efficient solution of (UMP).

Theorem 3.3.2. [76] *For the problem* (UMP), *suppose all conditions of Theorem 3.2.4 hold and* $\bar{x} \in C_R := \{x \in \mathbb{R}^n : g_i(x, v_i) \leq 0, \ \forall v_i \in \mathcal{V}_i, \ i \in I\}$ *fulfilling the robust basic constraint qualification. Assume further that the set* C_R *is convex. Then,*

(i) $\bar{x} \in C_R$ *is a weakly robust efficient solution of* (UMP) *if and only if there exist* $\theta_j \geq 0$, $j \in J$, *not all zero*, $\lambda_i \geq 0$, $i \in I$, $\bar{u}_j \in \mathcal{U}_j$, $j \in J$, *and* $\bar{v}_i \in \mathcal{V}_i$, $i \in I$, *such that*

$$\begin{cases} 0 \in \sum_{j\in J} \theta_j \partial_x f_j(\bar{x}, \bar{u}_j) + \sum_{i\in I} \lambda_i \partial_x^o g_i(\bar{x}, \bar{v}_i), \\ \lambda_i g_i(\bar{x}, \bar{v}_i) = 0, \ \forall i \in I, \ \textit{and} \\ f_j(\bar{x}, \bar{u}_j) = \max_{u_j\in\mathcal{U}_j} f_j(\bar{x}, u_j), \ \forall j \in J. \end{cases} \tag{3.3.1}$$

(ii) $\bar{x} \in C_R$ *is a properly robust efficient solution of* (UMP) *if and only if there exist* $\theta_j > 0$, $j \in J$, $\lambda_i \geq 0$, $i \in I$, $\bar{u}_j \in \mathcal{U}_j$, $j \in J$, *and* $\bar{v}_i \in \mathcal{V}_i$, $i \in I$, *such that* (3.3.1) *holds.*

Proof. (i) As $\bar{x} \in C_R$ is a weakly robust efficient solution of (UMP) if and only if $\bar{x} \in C_R$ is a weakly efficient solution of (RMP), there exist $\theta_j \geq 0$, $j \in J$, not all zero, such that $\bar{x} \in C_R$ is a solution of (RP_θ). In the other word, $\bar{x} \in C_R$ is a robust solution of the following uncertain (only in the constraints) convex optimization problem:

$$\min_{x\in\mathbb{R}^n} \left\{ \sum_{j\in J} \theta_j \max_{u_j\in\mathcal{U}_j} f_j(x, u_j) : g_i(x, v_i) \leq 0, \ i \in I \right\}.$$

Applying Theorem 3.2.4, we get that there exist $\bar{\lambda}_i \geq 0$, and $\bar{v}_i \in \mathcal{V}_i$, $i \in I$, such that

$$0 \in \partial \left(\sum_{j\in J} \theta_j \max_{u_j\in\mathcal{U}_j} f_j(\cdot, u_j) \right)(\bar{x}) + \sum_{i\in I} \bar{\lambda}_i \partial^o g_i(\cdot, \bar{v}_i)(\bar{x}), \ \bar{\lambda}_i g_i(\bar{x}, \bar{v}_i) = 0, \ \forall i \in I.$$

By employing the summation, positively homogeneous and max-function of convex subdifferential rule, the result as required.

(ii) The proof of (ii) is quite similar to that of (i) and so is omitted. □

In the following proposition, we give a sufficient condition that a robust efficient solution of (UMP) can be a properly robust efficient solution of (UMP).

Proposition 3.3.3. [76] *For the problem* (UMP), *let* $\bar{x} \in C_R$ *be a robust feasible solution for* (UMP). *Assume all conditions of Theorem 3.2.4 hold. Assume further that the set* C_R *is convex*, $C_R \cap C_R(\bar{x}) \neq \emptyset$ *and*

$$N(C_R \cap C_R(\bar{x}), \bar{x})$$

$$= \quad cone\left\{\left(\bigcup_{\substack{u_j\in\mathcal{U}_j(\bar{x})\\ j\in J}}\partial_x f_j(\bar{x},u_j)\right)\bigcup\left(\bigcup_{\substack{v_i\in\mathcal{V}_i^0(\bar{x})\\ i\in I_1(\bar{x})}}\partial_x^o g_i(\bar{x},v_i)\right)\right\}, \tag{3.3.2}$$

where

$$C_R(\bar{x}) := \{x\in\mathbb{R}^n : \max_{u_j\in\mathcal{U}_j} f_j(x,u_j) \le \max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j),\ \forall j\in J\},\ \textit{and}$$

$$\mathcal{U}_j(\bar{x}) := \{\bar{u}_j\in\mathcal{U}_j : f_j(\bar{x},\bar{u}_j) = \max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j)\},\ j\in J,$$

$$\mathcal{V}_i^0(\bar{x}) := \{\bar{v}_i\in\mathcal{V}_i : g_i(\bar{x},\bar{v}_i)=0\},\ i\in I_1(\bar{x}).$$

If $\bar{x}$ is a robust efficient solution of (UMP), *then $\bar{x}$ is a properly robust efficient solution of* (UMP).

Proof. Let $\bar{x}$ be a robust efficient solution of (UMP). Then $\bar{x}$ is a minimizer of the following scalar convex problem:

$$\min_{x\in\mathbb{R}^n}\left\{\sum_{j\in J}\max_{u_j\in\mathcal{U}_j} f_j(x,u_j) : x\in C_R\cap C_R(\bar{x})\right\},$$

or equivalently, $0\in\sum_{j\in J}\partial_x\left(\max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j)\right)+N(C_R\cap C_R(\bar{x}),\bar{x})$. It follows that there exists $\eta\in N(C_R\cap C_R(\bar{x}),\bar{x})$ such that

$$-\eta\in\sum_{j\in J}\partial_x\left(\max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j)\right).$$

Then, by the condition (3.3.2), there exist $\theta_j\ge 0$, $\bar{u}_j\in\mathcal{U}_j$, $\xi_j\in\partial_x f_j(\bar{x},\bar{u}_j)$, $j\in J$, $\lambda_i\ge 0$, $\bar{v}_i\in\mathcal{V}_i^0(\bar{x})$ and $\zeta_i\in\partial_x^o g_i(\bar{x},\bar{v}_i)$, $i\in I_1(\bar{x})$, such that

$$\eta=\sum_{j\in J}\theta_j\xi_j+\sum_{i\in I(\bar{x})}\lambda_i\zeta_i \text{ and } f_j(\bar{x},\bar{u}_j)=\max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j),\ \forall j\in J.$$

It then follows that

$$\begin{aligned}
0 &= -\eta+\eta\\
&\in \sum_{j\in J}\partial_x\left(\max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j)\right)+\sum_{j\in J}\theta_j\partial_x f_j(\bar{x},\bar{u}_j)+\sum_{i\in I_1(\bar{x})}\lambda_i\partial_x^o g_i(\bar{x},\bar{v}_i)\\
&\subseteq \sum_{j\in J}\partial_x\left(\max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j)\right)+\sum_{j\in J}\theta_j\left(\bigcup_{u_j\in\mathcal{U}_j(\bar{x})}\partial_x f_j(\bar{x},u_j)\right)+N(C_R,\bar{x})
\end{aligned}$$

$$= \sum_{j\in J}\partial_x\left(\max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j)\right)+\sum_{j\in J}\theta_j\partial_x\left(\max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j)\right)+N(C_R,\bar{x})$$
$$= \sum_{j\in J}(1+\theta_j)\partial_x\left(\max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j)\right)+N(C_R,\bar{x}).$$

Therefore,

$$\sum_{j\in J}(1+\theta_j)\max_{u_j\in\mathcal{U}_j} f_j(x,u_j)\geq \sum_{j\in J}(1+\theta_j)\max_{u_j\in\mathcal{U}_j} f_j(\bar{x},u_j),\ \forall x\in C_R,$$

which gives that $\bar{x}\in S_R^{\tilde{\theta}}$ with $\tilde{\theta}:=(1+\theta_1,1+\theta_2,\ldots,1+\theta_p)\in \operatorname{int}\mathbb{R}^p_+$, and so $\bar{x}$ is a properly robust efficient solution of (UMP). □

Let $\theta:=(\theta_1,\theta_2,\ldots,\theta_p)\in\mathbb{R}^p_+\backslash\{0\}$ (resp. $\operatorname{int}\mathbb{R}^p_+$), $\lambda^\theta:=(\lambda^\theta_1,\lambda^\theta_2,\ldots,\lambda^\theta_m)$, $u^\theta:=(u^\theta_1,u^\theta_2,\ldots,u^\theta_p)$, $v^\theta:=(v^\theta_1,v^\theta_2,\ldots,v^\theta_m)$, and $a^\theta\in S^\theta_R$. We have seen already that if the robust basic constraint qualification holds at a^θ, the set of Lagrange multiplier and uncertain parameters $M(a^\theta)$ for (RP_θ) corresponding to a^θ, given as

$$\begin{aligned}M(a^\theta):=\Big\{&(\lambda^\theta,u^\theta,v^\theta)\in\mathbb{R}^m_+\times\prod_{j\in J}\mathbb{R}^{q_j}\times\prod_{i\in I}\mathbb{R}^{q_i}:\\&0\in\sum_{j\in J}\theta_j\partial_x f_j(a^\theta,u^\theta_j)+\sum_{i\in I}\lambda^\theta_i\partial^\circ_x g_i(a^\theta,v^\theta_i),\\&\lambda^\theta_i g_i(a^\theta,v^\theta_i)=0,\ \forall i\in I \text{ and}\\&f_j(a^\theta,u^\theta_j)=\max_{u_j\in\mathcal{U}_j} f_j(a^\theta,u_j),\ \forall j\in J\Big\},\end{aligned}$$

is nonempty. Let further

$$\begin{aligned}&\tilde{I}_1(a^\theta,\lambda^\theta):=\{i\in I_1(a^\theta):\lambda^\theta_i>0\},\\&X_1(\lambda^\theta,v^\theta):=\{x\in C_R:\forall i\in\tilde{I}_1(a^\theta,\lambda^\theta),\ \exists\eta_i\in\partial^\circ_x g_i(a^\theta,v^a_i),\ \langle\eta_i,x-a^\theta\rangle=0\},\\&X_1(\lambda^\theta,u^\theta,v^\theta):=\{x\in X_1(\lambda^\theta,v^\theta):f_j(x,u^\theta_j)=\max_{u_j\in\mathcal{U}_j} f_j(x,u_j),\ \forall j\in J\}.\end{aligned}$$

By means of linear scalarization applied in Theorem 3.2.11, we can get characterizations of the weakly robust efficient solution sets $WR(C_R)$ and properly robust efficient solution set $PR(C_R)$ of the problem (UMP) immediately.

Theorem 3.3.4. [76] *For the problem* (UMP), *assume all conditions of Theorem 3.2.4 hold, and the set C_R is convex.*

(i) *Suppose further that for each* $\theta \in \mathbb{R}^p_+ \setminus \{0\}$, S^θ_R *is nonempty. Let* $a^\theta \in S^\theta_R$ *and the robust basic constraint qualification holds at* a^θ. *Let* $(\lambda^\theta, u^\theta, v^\theta) \in M(a^\theta)$. *Then*

$$WR(C_R) = \bigcup_{\theta \in \mathbb{R}^p_+ \setminus \{0\}} A_\theta,$$

where

$$A_\theta := \Big\{ x \in X_1(\lambda^\theta, u^\theta, v^\theta) : \exists \zeta^\theta \in \sum_{j\in J} \theta_j \partial_x f_j(x, u^\theta_j) \cap \sum_{j \in J} \theta_j \partial_x f_j(a^\theta, u^\theta_j), \langle \zeta^\theta, x - a^\theta \rangle = 0 \Big\}.$$

(ii) *If for each* $\theta \in int\mathbb{R}^p_+$, S^θ_R *is nonempty,* $a^\theta \in S^\theta_R$ *is fulfilled the robust basic constraint qualification,* $(\lambda^\theta, u^\theta, v^\theta) \in M(a^\theta)$, *then*

$$PR(C_R) = \bigcup_{\theta \in int\mathbb{R}^p_+} A_\theta.$$

To close this section, we give an example illustrating Theorem 3.3.4 which is indicated to be conveniently applied is applicable while the aforementioned result, due to Sun et al. [45, Theorem 4.7], are not. It means that at least one of the constraint functions $g_i(\cdot, v_i)$ for some $v_i \in \mathcal{V}_i$, is not convex while the robust feasible set is convex.

Example 3.3.5. Let $x := (x_1, x_2) \in \mathbb{R}^2$, $u_1 := (u_{1,1}, u_{1,2})$, $u_2 := (u_{2,1}, u_{2,2})$, $v_1 := (v_{1,1}, v_{1,2})$, $v_2 := (v_{2,1}, v_{2,2})$, $v_3 := (v_{3,1}, v_{3,2})$, $v_4 := (v_{4,1}, v_{4,2})$, $\mathcal{U}_1 = \mathcal{U}_2 := [0, 1]$, $\mathcal{V}_1 := \{(v_1, v_2) \in \mathbb{R}^2 : v_1^2 + v_2^2 \leq 1\}$, $\mathcal{V}_2 := [-2, -1] \times [-2, -1]$, $\mathcal{V}_3 := [4, 5] \times [2, 3]$ and $\mathcal{V}_4 := [0, 1] \times [0, 1]$.

We now consider the following constrained multiobjective optimization problem with uncertainty data:

$$\min_{x\in\mathbb{R}^2} \{(f_1(x, u_1), f_2(x, u_2)) : g_i(x, v_1) \leq 0, \ i \in I := \{1, 2, 3, 4\}\},$$

where

$$f_1(x, u_1) := u_1 x_1,$$

$$f_2(x, u_2) := u_2 x_2,$$

$$g_1(x, v_1) := v_{1,1}x_1 + v_{1,2}x_2 + x_1^3 - 2,$$
$$g_2(x, v_2) := v_{2,1}x_1 + v_{2,2}x_2 + 1,$$
$$g_3(x, v_3) := -v_{3,1}x_1 - v_{3,2}x_2 + 3,$$
$$g_4(x, v_4) := -v_{4,1}x_1 - v_{4,2}x_2^2,$$

and its robust counterpart

$$\min_{x\in\mathbb{R}^2} \left\{ (\max_{u_1\in\mathcal{U}_1} f_1(x, u_1), \max_{u_2\in\mathcal{U}_2} f_2(x, u_2)) : x \in C_R \right\}.$$

We obtain that for every $x \in \mathbb{R}^2$,

- $\max_{u_1\in\mathcal{U}_1} f_1(x, u_1) = \max\{x_1, 0\}$,
- $\max_{u_2\in\mathcal{U}_2} f_2(x, u_2) = \max\{x_2, 0\}$,
- $C_R = \{x \in \mathbb{R}^2 : \sqrt{x_1^2 + x_2^2} + x_1^3 - 2 \leq 0,\ -x_1 - x_2 + 1 \leq 0,\ -4x_1 - 2x_2 + 3 \leq 0,\ -x_1 \leq 0\}$,

as a straightforward calculation shows. Let us denote

$$\Lambda := \{(\theta_1, \theta_2) \in \mathbb{R}^2_+ : \theta_1 + \theta_2 = 1\},$$
$$\text{int}\,\Lambda := \{(\theta_1, \theta_2) \in \mathbb{R}^2_+ : \theta_1 > 0,\ \theta_2 > 0,\ \theta_1 + \theta_2 = 1\}.$$

We now consider the following possibilities:

(i) If $\theta := (\theta_1, \theta_2) = (1, 0)$ then (RP_θ) reads as follows:

$$\min_{x\in\mathbb{R}^2}\{x_1 : x \in C_R\}.$$

As $\theta_1 x_1 + \theta_2 x_2 = x_1 \geq 0$, $\forall x \in C_R$, we can take $a^\theta := (0, 2) \in S_\theta$ and so, $I_1(a^\theta) = \{1, 4\}$. Let us choose $\lambda^\theta := (0, 0, 0, 1)$, $u^\theta := (1, 1)$, $v^\theta := (v_1^\theta, v_2^\theta, v_3^\theta, v_4^\theta) = ((0, 1), (-1, -1), (4, 2), (1, 0))$, and we also have $\tilde{I}_1(a^\theta, \lambda^\theta) = \{4\}$. Then A_θ can be easily calculated $A_\theta = \{x \in \mathbb{R}^2 : x_1 = 0,\ \frac{3}{2} \leq x_2 \leq 2\}$.

(ii) Similarly, if $\theta := (0, 1)$ then we can take $a^\theta := (1, 0) \in S_\theta$ and so, $I_1(a^\theta) = \{1, 2, 4\}$. Let us choose $\lambda^\theta := (\frac{1}{4}, 1, 0, 0)$, $u^\theta := (1, 1)$, $v^\theta := (v_1^\theta, v_2^\theta, v_3^\theta, v_4^\theta) = ((1, 0), (-1, -1), (4, 2), (0, 1))$, and we also have $\tilde{I}_1(a^\theta, \lambda^\theta) = \{1, 2\}$. Thus $A_\theta = \{(1, 0)\}$.

(iii) If $\theta \in \{(\theta_1,\theta_2) \in \mathbb{R}^2 : \frac{2}{3} < \theta_1 < 1,\ \theta_2 > 0,\ \theta_1+\theta_2 = 1\}$, then (RP_θ) becomes

$$\min_{x\in\mathbb{R}^2}\{\theta_1 x_1 + \theta_2 x_2 : x \in C_R\}.$$

As $\theta_1 x_1 + \theta_2 x_2 \geq \theta_1 x_1 - 2\theta_2 x_1 + \frac{3}{2}\theta_2 = (3\theta_1 - 2)x_1 + \frac{3}{2}\theta_2 \geq \frac{3}{2}\theta_2$, $\forall x \in C_R$, then we can take $a^\theta := (0, \frac{3}{2}) \in S_\theta$ and so, $I_1(a^\theta) = \{2,4\}$. Let us choose $\lambda^\theta := (0, 0, \frac{\theta_2}{2}, \theta_1 - 2\theta_2)$ (note that $\theta_1 > \frac{2}{3}$ and $\theta_1+\theta_2 = 1$ imply $\theta_1 - 2\theta_2 > 0$), $u^\theta := (1,1)$, $v^\theta := (v_1^\theta, v_2^\theta, v_3^\theta, v_4^\theta) = ((1,0),(-1,-1),(4,2),(1,0))$, and we also have $\tilde{I}_1(a^\theta, \lambda^\theta) = \{2,4\}$. In this case, it is easy to see that $A_\theta = \{(0,\frac{3}{2})\}$.

(iv) Similarly, if $\theta \in \{(\theta_1,\theta_2) \in \mathbb{R}^2 : \frac{1}{2} < \theta_1 < \frac{2}{3},\ \theta_2 > 0,\ \theta_1+\theta_2 = 1\}$ then $A_\theta = \{(0,1)\}$ and if $\theta \in \{(\theta_1,\theta_2) \in \mathbb{R}^2 : 0 < \theta_1 < \frac{1}{2},\ \theta_2 > 0,\ \theta_1+\theta_2 = 1\} = \{(\theta_1,\theta_2) \in \mathbb{R}^2 : \frac{1}{2} < \theta_2 < 1,\ \theta_1 > 0,\ \theta_1+\theta_2 = 1\}$ then $A_\theta = \{(1,0)\}$.

(v) If $\theta := (\frac{1}{2}, \frac{1}{2})$, then $\theta_1 x_1 + \theta_2 x_2 \geq \frac{1}{2}$, $\forall x \in C_R$. Take $a^\theta := (\frac{1}{2}, \frac{1}{2}) \in S_\theta$ and so, $I_1(a^\theta) = \{2,3,4\}$. Let us choose $\lambda^\theta := (0, \frac{1}{2}, 0, 0)$, $u^\theta := (1,1)$, $v^\theta := (v_1^\theta, v_2^\theta, v_3^\theta, v_4^\theta) = ((1,0),(-1,-1),(4,2),(1,0))$, and we also have $\tilde{I}_1(a^\theta,\lambda^\theta) = \{2\}$. Then elementary calculations give us $A_\theta = \{x \in \mathbb{R}^2 : -x_1 - x_2 + 1 = 0,\ x_1 \geq 0,\ x_2 \geq 0\}$.

(vi) Similarly, if $\theta = (\frac{2}{3}, \frac{1}{3})$ then we can take $a^\theta = (\frac{1}{4}, 1) \in S_\theta$, $\lambda^\theta = (0,0,\frac{1}{6},0)$ and so, $A_\theta = \{x \in \mathbb{R}^2 : -4x_1 - 2x_2 + 3 = 0,\ x_1 \geq 0,\ x_2 \geq 0\}$.

Therefore, by Theorem 3.3.4, weakly and properly robust efficient solution sets of (UMP) look like

$$\begin{aligned} WR(C_R) &= \bigcup_{\theta\in\mathbb{R}^2_+\setminus\{0\}} A_\theta = \bigcup_{\theta\in\Lambda} A_\theta \\ &= \{x \in \mathbb{R}^2 : x_1 = 0,\ 1 \leq x_2 \leq 2\} \\ &\quad \cup\{x \in \mathbb{R}^2 : -4x_1 - 2x_2 + 3 = 0,\ x_1 \geq 0,\ x_2 \geq 0\} \\ &\quad \cup\{x \in \mathbb{R}^2 : -x_1 - x_2 + 1 = 0,\ x_1 \geq 0,\ x_2 \geq 0\} \end{aligned}$$

and

$$PR(C_R) = \bigcup_{\theta\in\mathrm{int}\mathbb{R}^2_+} A_\theta = \bigcup_{\theta\in\mathrm{int}\Lambda} A_\theta$$

$$= \{x \in \mathbb{R}^2 : -4x_1 - 2x_2 + 3 = 0,\ x_1 \geq 0,\ x_2 \geq 0\}$$
$$\cup \{x \in \mathbb{R}^2 : -x_1 - x_2 + 1 = 0,\ x_1 \geq 0,\ x_2 \geq 0\}.$$

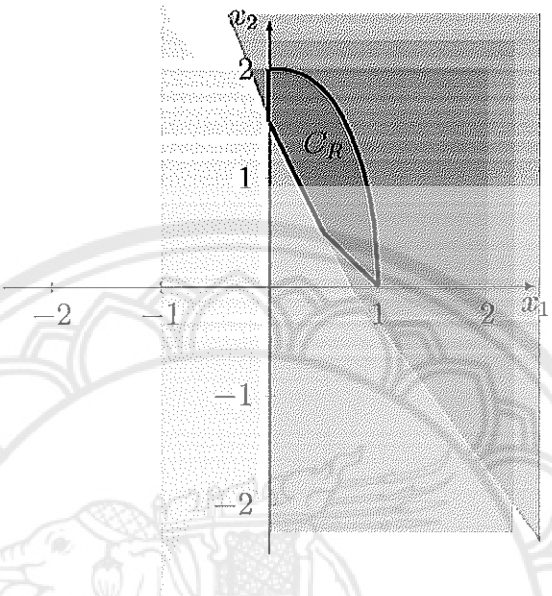


Figure 25: Illustration of the robust feasible set C_R in Example **3.3.5**.

CHAPTER IV

KKT CONDITIONS IN NON-CONVEX OPTIMIZATION PROBLEMS

In a recent work, Ho [52] established some necessary and sufficient KKT optimality conditions for the differentiable problem (P) by imposing the convexity of the level sets of the given function f without the convexity of the constraint set and of the functions f and each g_i. In the same context we succeed to weaken the differentiability and the convexity assumptions considered in [52, Theorem 1] and [53, Theorem 3.1], respectively, in a way that the KKT conditions still remain hold for a given global minimizer and for a given weak Pareto minimum, respectively. In order to present these results, we first recall the concepts of nearly convex set and quasi-convexity.

Definition 4.0.6. The set A is **nearly convex** at the point $x \in A$ if

$$\forall y \in A,\ \exists t_l \downarrow 0 \text{ such that } x + t_l(y-x) \in A. \tag{4.0.3}$$

Definition 4.0.7. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be **quasiconvex** if for any $\alpha \in \mathbb{R}$,

$$L(f, \leq, \alpha) := \{x \in \mathbb{R}^n : f(x) \leq \alpha\}$$

is a convex set, or equivalently, for any $\alpha \in \mathbb{R}$,

$$L(f, <, \alpha) := \{x \in \mathbb{R}^n : f(x) < \alpha\}$$

is a convex set.

Theorem 4.0.8. [78, Theorem 11] *Let f be a continuous quasiconvex function from $\mathbb{R}^n$ to $\mathbb{R}$, and $\alpha \in \mathbb{R}$. If* $\operatorname{int} L(f, =, \beta) = \emptyset$ *and* $\operatorname{int} L(f, <, \beta) \neq \emptyset$ *for some $\beta \in \mathbb{R}$, then we have*

(i) $L(f, <, \beta) = \operatorname{int} L(f, \leq, \beta)$,

(ii) $\operatorname{cl} L(f, <, \beta) = L(f, \leq, \beta)$.

4.1 Non-smooth optimization problems

The aim of this section is to provide an affirmative answer to the question that whether Ho's result can be extended to non-differentiable case. More precisely, consider the problem (P) where a constraint set is given by

$$\mathcal{X} := \{x \in \mathbb{R}^n : g_i(x) \leq 0,\ i = 1, 2, \ldots, m\},$$

we will show that when Slater's constraint qualification holds and a non-degeneracy condition holds at the feasible point x without both the convexity and differentiability of f and g_i as well as the convexity of the feasible set, the KKT necessary optimality condition becomes globally sufficient provided that the set $\mathcal{X}$ is nearly convex at x, the strict level set of f at x, $L(f, <, f(x))$, is convex and the additional condition $x \in \text{cl } L(f, <, f(x))$ is satisfied. It is remarkable that the condition $x \in \text{cl } L(f, <, f(x))$ can be absent in the differentiable case. To do these, we need the following lemma which plays a significant role in our subsequent analysis.

Lemma 4.1.1. *Let s be a real-valued sublinear function on $\mathbb{R}^n$. If there exists $x_0 \in \mathbb{R}^n$ such that $s(x_0) < 0$, then we have*

$$\text{cl}\{x \in \mathbb{R}^n : s(x) < 0\} = \{x \in \mathbb{R}^n : s(x) \leq 0\}.$$

Proof. By assumption, we can check that $\inf_{x \in \mathbb{R}^n} s(x) < 0$ and $x_0 \in \text{int}\{x \in \mathbb{R}^n : s(x) < 0\} \neq \emptyset$. We claim that

$$\text{int}\{x \in \mathbb{R}^n : s(x) = 0\} = \emptyset.$$

Suppose on contrary that there exists $\hat{x} \in \text{int}\{x \in \mathbb{R}^n : s(x) = 0\}$. Then $\mathbb{B}(\hat{x}, \varepsilon) \subseteq \{x \in \mathbb{R}^n : s(x) = 0\}$ for some $\varepsilon > 0$, and hence, by Lemma 2.2.25, $s(x) \geq 0$ for all $x \in \mathbb{R}^n$. This is a contradiction, and therefore the conclusion follows by applying Theorem 4.0.8. □

According to technical approach given in [52], we begin with an extension of Proposition 2.2.(i) in [79] which will play a key role to derive sufficient KKT optimality condition in our main result for non-differentiable problem.

Lemma 4.1.2. [80] *Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous and tangentially convex at $x \in \mathbb{R}^n$. Suppose that the set $L(f, <, f(x))$ is nonempty and convex. If $x \in \text{cl } L(f, <, f(x))$, then*

$$N(L(f, <, f(x)), x)^\circ \subseteq \{d \in \mathbb{R}^n : f'(x; d) \leq 0\}.$$

If $0 \notin \partial_T f(x)$, then also $x \in \text{cl } L(f, <, f(x))$ and the above containment becomes equality.

Proof. Consider $x \in \text{cl } L(f, <, f(x))$. Firstly, we will show that

$$\partial_T f(x) \subseteq N(L(f, <, f(x)), x).$$

Now take any $\xi \in \partial_T f(x)$. Given any y in $L(f, <, f(x))$. Let us notice that, by continuity of f at x, the set $L(f, <, f(x))$ is (relative) open. So, $y \in L(f, <, f(x)) = \text{ri } L(f, <, f(x))$ and, by the line segment principle (Proposition 2.2.7(ii)),

$$x + t(y - x) \in \text{ri } L(f, <, f(x)) = L(f, <, f(x)) \text{ for } t \in]0, 1].$$

For values t sufficiently small, $f'(x; y - x) \leq 0$, and hence $\langle \xi, y - x \rangle \leq 0$ by the definition of the tangential subdifferential of f at x. Therefore, for any $y \in L(f, <, f(x))$, $\langle \xi, y - x \rangle \leq 0$, thereby showing that $\xi \in N(L(f, <, f(x)), x)$. Thus, $\partial_T f(x) \subseteq N(L(f, <, f(x)), x)$ as required. It then follows from Proposition 2.2.13(i) that

$$N(L(f, <, f(x)), x)^\circ \subseteq \partial_T f(x)^\circ.$$

Further, it can be checked that

$$\partial_T f(x)^\circ = \{d \in \mathbb{R}^n : f'(x; d) \leq 0\}.$$

Hence, the desired result is obtained.

Assuming $0 \notin \partial_T f(x)$, we now demonstrate that x is contained in $\text{cl } L(f, <, f(x))$. As $0 \notin \partial_T f(x)$, there exists $d_0 \in \mathbb{R}^n$ such that $f'(x; d_0) < 0$. By the definition of directional derivative, there exists a positive real number δ such that

$$f(x + td_0) < f(x) \text{ for all } t \in]0, \delta[.$$

In particular, for each $k \in \mathbb{N}$, the vector $x_l := x + \frac{\delta}{l+1} d_0$ belongs to $L(f, <, f(x))$, and $x_l \to x$ as $l \to +\infty$. This means that $x \in \text{cl } L(f, <, f(x))$. To establish the remaining inclusion $\{d \in \mathbb{R}^n : f'(x; d) \leq 0\} \subseteq N(L(f, <, f(x)), x)^\circ$, we argue first by applying Lemma 4.1.1 that

$$\{d \in \mathbb{R}^n : f'(x; d) \leq 0\} = \text{cl}\{d \in \mathbb{R}^n : f'(x; d) < 0\}.$$

Moreover, by Proposition 2.2.19, convexity of $L(f, <, f(x))$ asserts that

$$N(L(f, <, f(x)), x)^\circ = T(L(f, <, f(x)), x).$$

Thus it is enough to show that, as $T(L(f, <, f(x)), x)$ is closed due to Theorem 2.2.17,

$$\{d \in \mathbb{R}^n : f'(x; d) < 0\} \subseteq T(L(f, <, f(x)), x).$$

Given any $d \in \mathbb{R}^n$ such that $f'(x; d) < 0$. As seen before, we can find two sequences $\{r_l\} \subset]0, +\infty[$ and $\{x_l\} \subset L(f, <, f(x))$ such that

$$x_l = x + r_l d \in L(f, <, f(x)) \text{ and } r_l \to 0 \text{ as } l \to +\infty.$$

Consequently, $x_l \to x$ and $\frac{1}{r_l}(x_l - x) \to d$ as $l \to +\infty$. Therefore, $d \in T(L(f, <, f(x)), x)$, thereby establishing the requisite result. ☐

Remark 4.1.3. In Lemma 4.1.2, if f is differentiable at $x \in \mathbb{R}^n$ such that $\nabla f(x) \neq 0$, then $f'(x; d) = \langle \nabla f(x), d \rangle$ for all $d \in \mathbb{R}^n$. In this case, we obtain the following result.

Corollary 4.1.4. [79, Proposition 2.2.(i)] *Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable at x with $\nabla f(x) \neq 0$. Then*

$$N(L(f, <, f(x)), x) = \{d \in \mathbb{R}^n : d = r\nabla f(x), \text{ for some } r \geq 0\}$$

provided that $L(f, <, f(x))$ is convex.

Proof. As the set $N(L(f, <, f(x)), x)$ and $\text{cone}\{\nabla f(x)\}$ are closed convex cone, by the polar cone theorem [54, Proposition 3.1.1(b)], $(N(L(f, <, f(x)), x)^\circ)^\circ = N(L(f, <, f(x)), x)$ and $((\text{cone co } \{\nabla f(x)\})^\circ)^\circ = \text{cone co } \{\nabla f(x)\}$. Therefore, owing to Proposition 2.2.12 and Lemma 4.1.2,

$$N(L(f, <, f(x)), x) = (N(L(f, <, f(x)), x)^\circ)^\circ$$

$$
\begin{aligned}
&= \{d \in \mathbb{R}^n : \langle \nabla f(x), d\rangle \leq 0\}^\circ \\
&= ((\text{cone co } \{\nabla f(x)\})^\circ)^\circ \\
&= \text{cone co } \{\nabla f(x)\} \\
&= \{d \in \mathbb{R}^n : d = r\nabla f(x), \text{ for some } r \geq 0\}.
\end{aligned}
$$

□

Remark 4.1.5. The condition $0 \notin \partial_T f(x)$ given in Lemma 4.1.2 ensuring $x \in \text{cl } L(f,<,f(x))$ can be replaced by $L(f,\leq,f(x)) = \text{cl}L(f,<,f(x))$. However, the following two examples show that the validity of the condition $0 \notin \partial_T f(x)$ does not necessarily imply the condition $L(f,\leq,f(x)) = \text{cl } L(f,<,f(x))$ and vise versa.

Example 4.1.6. Consider a non-differentiable function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$
f(x) := \begin{cases} \max\{\frac{1}{2}x(x+1)(x+2), 2x\}, & \text{if } x \in [-1,+\infty[\ ; \\ \max\{-\frac{1}{2}(x+1)(x+2), 0\}, & \text{ohterwise.} \end{cases}
$$

For $x := 0$, we have $\partial_T f(x) = [1,2]$, and hence $0 \notin \partial_T f(x)$. However, cl $L(f,<,f(x)) = [-1,0] \neq\ ]-\infty,-2] \cup [-1,0] = L(f,\leq,f(x))$.

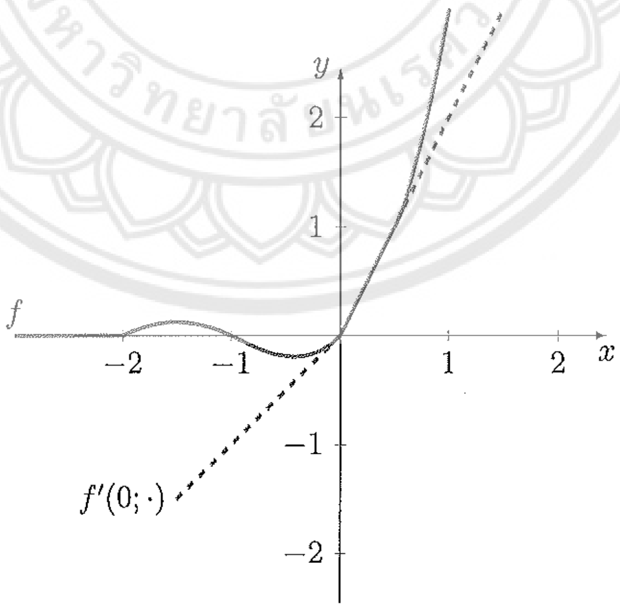


Figure 26: Plots of function f and its directional derivative in Example 4.1.6.

Example 4.1.7. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as $f(x) := \max\{-x^2, x\}$, $\forall x \in \mathbb{R}$ and $x := 0$. So, $L(f, \leq, f(x)) =]-\infty, 0] = \text{cl } L(f, <, f(x))$, while $\partial_T f(x) = [0, 1]$.

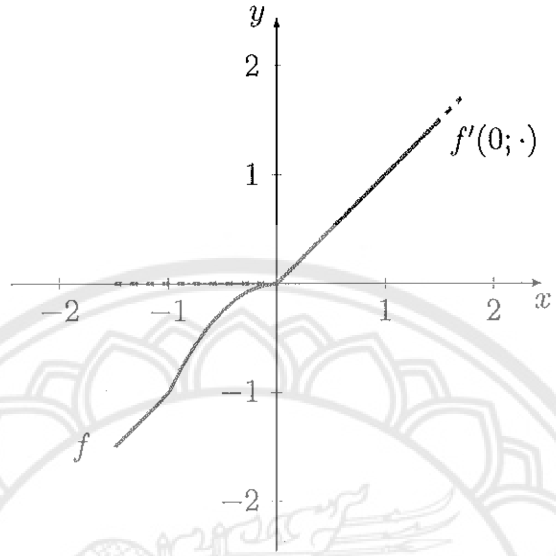


Figure 27: Plots of function f and its directional derivative in Example 4.1.7.

Remark 4.1.8. In Example 4.1.6 tell us that the fulfillment of $x \in \text{cl } L(f, <, f(x))$ is not sufficient to ensure that the equality $L(f, \leq, f(x)) = \text{cl } L(f, <, f(x))$.

Next, we will see how the condition $0 \notin \partial_T f(x)$ is not necessarily to be assumed when considering KKT optimality conditions. Before doing so let us formally state the notion of a KKT point of (P) in terms of tangential subdifferentials. A feasible point $\bar{x}$ of (P) is called a **KKT point** if there exist scalars $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$, such that

i) $0 \in \partial_T f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \partial_T g_i(\bar{x})$,

ii) $\lambda_i g_i(\bar{x}) = 0, \quad i = 1, 2, \ldots, m.$

The above feasible point is also called a **non-trivial KKT point** if the corresponding $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_m)$ is a non-zero vector.

We now turn our attention to state our main result of this section.

Theorem 4.1.9. [80] *Given the nonlinear programming problem* (P) *fulling Slater's constraint qualification. Let $\bar{x} \in \mathcal{X}$ be a feasible solution, the functions $g_i : \mathbb{R}^n \to$*

$\mathbb{R}$, $i = 1, 2, \ldots, m$, *be continuous such that for every* $i \in I(\bar{x})$ *the functions* g_i *is tangentially convex at* $\bar{x}$. *Suppose that* $\bar{x}$ *satisfies the condition* (4.0.3), *and* $0 \notin \partial_T g_i(\bar{x})$ *for all* $i \in I(\bar{x})$. *Assume further that the functions* $f : \mathbb{R}^n \to \mathbb{R}$, *is continuous and tangentially convex at* $\bar{x}$.

(i) *If* $\bar{x}$ *is a global minimizer then* $\bar{x}$ *is a KKT point.*

(ii) *Conversely, if* $\bar{x}$ *is a non-trivial KKT point such that* $\bar{x} \in \text{cl } L(f, <, f(\bar{x}))$, *and* $L(f, <, f(\bar{x}))$ *is convex then* $\bar{x}$ *is a global minimizer.*

Proof. (i) If $\bar{x}$ is a global minimizer, then, by the Fritz-John optimality conditions (Theorem 2.3.7), there exist real numbers $\lambda_i \geq 0$, $i \in I$, not all zero, satisfying ii) and

$$\lambda_0 f'(\bar{x}; d) + \sum_{i \in I} \lambda_i g_i'(\bar{x}; d) \geq 0, \ \forall d \in \mathbb{R}^n. \tag{4.1.1}$$

We shall now show that $\lambda_0 > 0$. Let us assume that $\lambda_0 = 0$. Hence there exists some $i \in I$ such that $\lambda_i > 0$ which, by ii), implies the non-emptyness of $I(\bar{x})$. Taking (4.1.1) into account we actually have

$$\sum_{i \in I(\bar{x})} \lambda_i g_i'(\bar{x}; d) \geq 0, \ \forall d \in \mathbb{R}^n. \tag{4.1.2}$$

As $\sum_{i \in I(\bar{x})} \lambda_i > 0$, setting $\bar{\lambda}_i := \frac{\lambda_i}{\sum_{i \in I(\bar{x})} \lambda_i}$ for each $i \in I(\bar{x})$. Multiplying both sides of (4.1.2) by $\frac{1}{\sum_{i \in I(\bar{x})} \lambda_i}$ for each $d \in \mathbb{R}^n$, we would have from the obtained inequality that

$$0 \in \partial_T \left(\sum_{i \in I(\bar{x})} \bar{\lambda}_i g \right) (\bar{x}) = \sum_{i \in I(\bar{x})} \bar{\lambda}_i \partial_T g_i(\bar{x}).$$

Then there exist $\xi_i \in \partial_T g_i(\bar{x})$, $i \in I(\bar{x})$, such that

$$\sum_{i \in I(\bar{x})} \bar{\lambda}_i \xi_i = 0. \tag{4.1.3}$$

It follows from the non-degeneracy condition at $\bar{x}$ that $\xi_i \neq 0$, $\forall i \in I(\bar{x})$. This together with the fact that $\sum_{i \in I(\bar{x})} \bar{\lambda}_i = 1$, we get

$$0 < \sum_{i \in I(\bar{x})} \bar{\lambda}_i \|\xi_i\|. \tag{4.1.4}$$

On the one hand, as the functions g_i are continuous, the set of Slater points, which contained in $\mathcal{X}$, is open. It means that, for a Slater point x_0, there exists a positive real number ρ such that $\mathbb{B}(x_0, \rho) \subseteq \mathcal{X}$. Thus, $x_0 + \frac{\rho}{2\|\xi_i\|}\xi_i \in \mathbb{B}(x_0, \rho) \subseteq \mathcal{X}$, $\forall i \in I(\bar{x})$. It is worth noting that

$$g_i'(\bar{x}; y - \bar{x}) \leq 0 \text{ for all } i \in I(\bar{x}),\ y \in \mathcal{X}. \tag{4.1.5}$$

Otherwise, $g_i'(\bar{x}; y - \bar{x}) > 0$ for some $y \in \mathcal{X}$ and for some $i \in I(\bar{x})$. Then by the definition of directional derivative, there exists $\delta > 0$ such that

$$\left| \frac{g_i(\bar{x} + t(y - \bar{x}))}{t} - g_i'(\bar{x}; y - \bar{x}) \right| < g_i'(\bar{x}; y - \bar{x}) \text{ whenever } t \in]0, \delta[.$$

Subsequently, $g_i(\bar{x} + t(y - \bar{x})) > 0$ for all $t \in]0, \delta[$. This contradicts to the condition (4.1.1) that we can find some t_l small enough such that $\bar{x} + t_l(y - \bar{x}) \in \mathcal{X}$. Therefore, by the definition of tangential subdifferentials, for each $i \in I(\bar{x})$ we have

$$\langle \xi_i, x_0 - \bar{x} \rangle + \frac{\rho}{2}\|\xi_i\| = \left\langle \xi_i, x_0 + \frac{\rho}{2\|\xi_i\|}\xi_i - \bar{x} \right\rangle \leq g_i'\left(\bar{x}; x_0 + \frac{\rho}{2\|\xi_i\|}\xi_i - \bar{x}\right) \leq 0.$$

Multiplying both sides of inequality above by $\bar{\lambda}_i$, $i \in I(\bar{x})$, and summing up the obtained inequalities together with (4.1.3) we get

$$\frac{\rho}{2} \sum_{i \in I(\bar{x})} \lambda_i \|\xi_i\| \leq 0,$$

which is in turn a contrast to (4.1.4). Hence $\lambda_0 > 0$, and without loss of generality we can set $\lambda_0 = 1$. Then, for every $d \in \mathbb{R}^n$ we have

$$\left(f + \sum_{i=1}^{m} \lambda_i g_i \right)' (\bar{x}; d) = f'(\bar{x}; d) + \sum_{i=1}^{m} \lambda_i g_i'(\bar{x}; d) \geq 0,$$

which is noting else than

$$0 \in \partial_T \left(f + \sum_{i=1}^{m} \lambda_i g_i \right)(\bar{x}) = \partial_T f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \partial_T g_i(\bar{x}),$$

i) as required.

(ii) Let $\bar{x} \in \mathcal{X}$ be an arbitrary non-trivial KKT point. We see that for every $y \in L(f, <, f(\bar{x}))$ we get the following inequality $\langle x, y - \bar{x} \rangle \leq 0$, $\forall x \in N(L(f, <, f(\bar{x})), \bar{x})$, which means that $y - \bar{x} \in N(L(f, <, f(\bar{x})), \bar{x})^{\circ}$. By Lemma 4.1.2, we obtain that

$$f'(\bar{x}; y - \bar{x}) \leq 0 \text{ for all } y \in L(f, <, f(\bar{x})). \tag{4.1.6}$$

On the other hand, employing i) and ii), we obtain

$$f'(\bar{x}; d) + \sum_{i \in I(\bar{x})} \lambda_i g_i'(\bar{x}; d) \geq 0 \text{ for any } d \in \mathbb{R}^n.$$

In particular, by using (4.1.6),

$$\sum_{i \in I(\bar{x})} \lambda_i g_i'(\bar{x}; y - \bar{x}) \geq 0 \text{ for any } y \in L(f, <, f(\bar{x})). \tag{4.1.7}$$

Next, we claim that $L(f, <, f(\bar{x})) \cap \mathcal{X} = \emptyset$. Arguing by contradiction, suppose that there exists $w \in L(f, <, f(\bar{x})) \cap \mathcal{X}$. Then, from (4.1.5) and (4.1.7),

$$\sum_{i \in I(\bar{x})} \lambda_i g_i'(\bar{x}; w - \bar{x}) = 0. \tag{4.1.8}$$

Furthermore, since $L(f, <, f(\bar{x}))$ is open, for each $d \in \mathbb{R}^n$ there exists some $t > 0$ small enough such that $w + td \in L(f, <, f(\bar{x}))$. Hence, by using (4.1.7), (4.1.8) and sub-linearity of $g_i'(\bar{x}, \cdot)$, $i \in I(\bar{x})$,

$$\begin{aligned} 0 &\leq \sum_{i \in I(\bar{x})} \lambda_i g_i'(\bar{x}; w + td - \bar{x}) \\ &\leq \sum_{i \in I(\bar{x})} \lambda_i g_i'(\bar{x}; w - \bar{x}) + t \sum_{i \in I(\bar{x})} \lambda_i g_i'(\bar{x}; d) \\ &= t \sum_{i \in I(\bar{x})} \lambda_i g_i'(\bar{x}; d). \end{aligned}$$

As $\bar{x}$ is a non-trivial KKT point, we have $\sum_{i \in I(\bar{x})} \lambda_i > 0$. In a similar manner of the first argument, the last inequality arrives at a contradiction. So, our claim $L(f, <, f(\bar{x})) \cap \mathcal{X} = \emptyset$ holds. This means that $\bar{x}$ is a global minimizer, and the proof is completed. □

Remark 4.1.10. It is worth observing that the Slater's constraint qualification along with a non-degeneracy at $\bar{x}$ arrives at the assertion

$$0 \notin \sum_{i \in I(\bar{x})} \lambda_i \partial_T g_i(\bar{x}) \text{ whenever } \lambda_i \geq 0,\ i \in I(\bar{x}) \text{ such that } \sum_{i \in I(\bar{x})} \lambda_i = 1,$$

or equivalently, $0 \notin \operatorname{co}\left(\bigcup_{i \in I(x)} \partial_T g_i(x)\right)$.

It can be seen that the condition $0 \notin \partial_T f(\bar{x})$ will follows from the nonemptiness of $L(f, <, f(\bar{x}))$ and the pseudoconvexity of f at $\bar{x}$. In this context together with Lemma 4.1.2, pseudoconvexity of f at $\bar{x}$ provided sufficient condition for $\bar{x} \in \text{cl } L(f, <, f(\bar{x}))$ whenever $L(f, <, f(\bar{x}))$ is a nonempty and convex set. However, the following example shows that pseudoconvexity of f at $\bar{x}$ does not necessarily imply convexity of $L(f, <, f(\bar{x}))$, and hence Theorem 4.1.9 cannot be applied in this situation.

Example 4.1.11. Let $f : \mathbb{R} \to \mathbb{R}$ be a non-differentiable function defined by

$$f(x) := \begin{cases} \max\{x^3 + x, 2x\}, & \text{if } x \in [0, +\infty[\ ; \\ \frac{1}{2}x(x+1)(x+2), & \text{if } x \in]-\infty, 0[. \end{cases}$$

Then, for $\bar{x} := 0$, $f'(\bar{x}; d) = \max\{d, 2d\}$ for any $d \in \mathbb{R}$, from which we can obtain that f is pseudoconvex at $\bar{x}$, while $L(f, <, f(\bar{x})) =]-\infty, -2[\cup]-1, 0[$ is not convex.

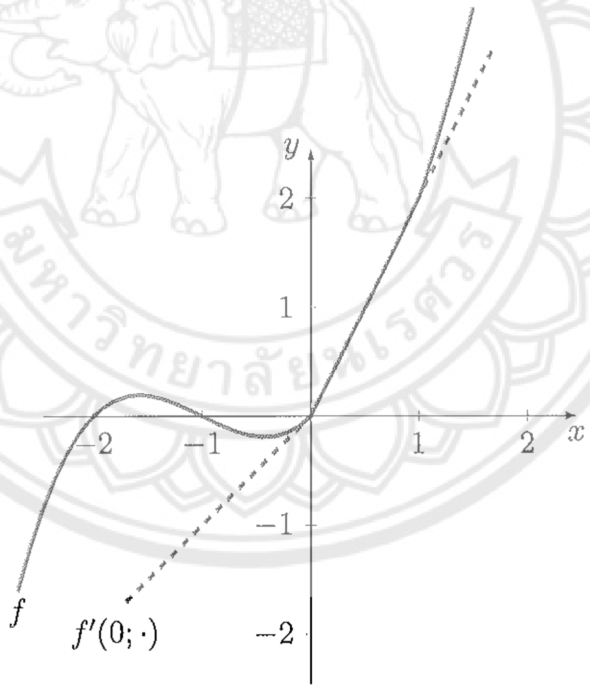


Figure 28: Plots of function f and its directional derivative in Example 4.1.11.

Here we give an example to illustrate that Theorem 4.1.9 is indicated to be conveniently applied in some cases where Theorem 9 of [48] cannot be used even when the feasible set $\mathcal{X}$ is convex. Namely, the objective function is not pseudoconvex at considered point.

Example 4.1.12. Let $f : \mathbb{R} \to \mathbb{R}$ be a non-differentiable function defined by

$$f(x) := \begin{cases} (x-1)(x-2)(x-3)+1, & \text{if } x \in [1, +\infty[\ ; \\ \max\{x^3, x\}, & \text{if } x \in [-\infty, 1[\ , \end{cases}$$

and the feasible set $\mathcal{X} = \{x \in \mathbb{R} : g_1(x) \leq 0\}$, where

$$g_1(x) := \max\left\{-x, -\frac{1}{2}(x-1)^2(x-2) - 1\right\}.$$

Evidently, the function f is not pseudoconvex at $\bar{x} := 0$. We can verify that $\mathcal{X} = [0, +\infty[$ and the feasible point $\bar{x}$ satisfies non-trivial KKT conditions with $\lambda_1 := 1$, non-degeneracy condition, and $\bar{x} \in \text{cl}\ L(f, <, f(\bar{x}))$ in which $L(f, <, f(\bar{x})) =]-\infty, 0[$ is convex. Then, by Theorem 4.1.9, $\bar{x}$ is a global minimizer.

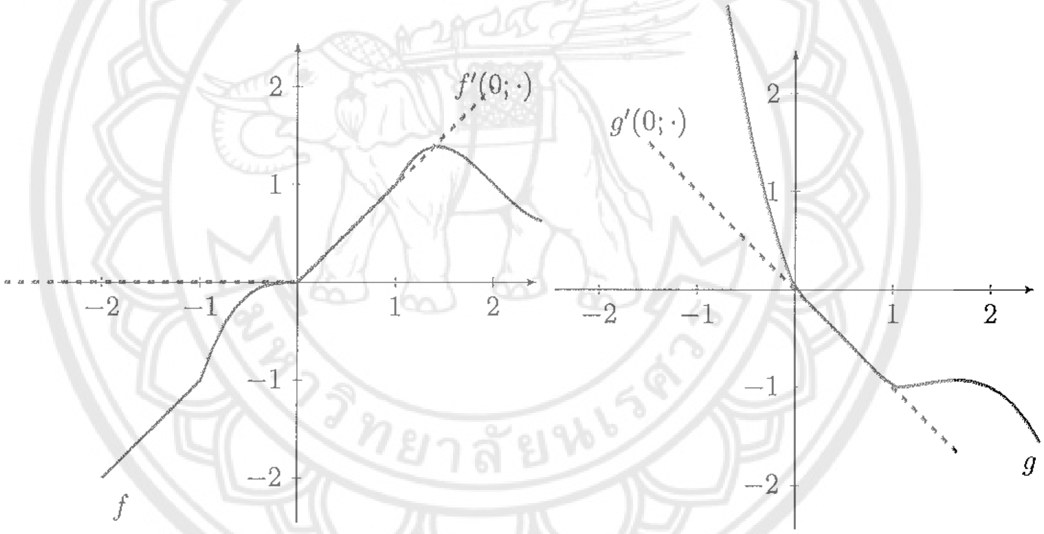


Figure 29: Plots of functions f and g, and their directional derivatives in Example 4.1.12.

Example 4.1.13. *Let $f : \mathbb{R} \to \mathbb{R}$ be defined as in Example 4.1.12, and the feasible set $\mathcal{X} = \{x \in \mathbb{R} : g_1(x) \leq 0\}$, where*

$$g_1(x) := \max\{-x, -x(x-1)(x-2)\}.$$

We can check that $\mathcal{X} = [0, 1] \cup [2, +\infty[$ and the feasible point $\bar{x} := 0$ satisfies non-trivial KKT conditions with $\lambda_1 := 1$, $L(f, <, f(\bar{x})) =]-\infty, 0[$ is convex such that $\bar{x} \in \text{cl}\ L(f, <, f(\bar{x}))$. In addition, a non-degeneracy condition and condition (4.0.3) *hold at $\bar{x}$. Theorem 4.1.9 then indicates that $\bar{x}$ is a global minimizer.*

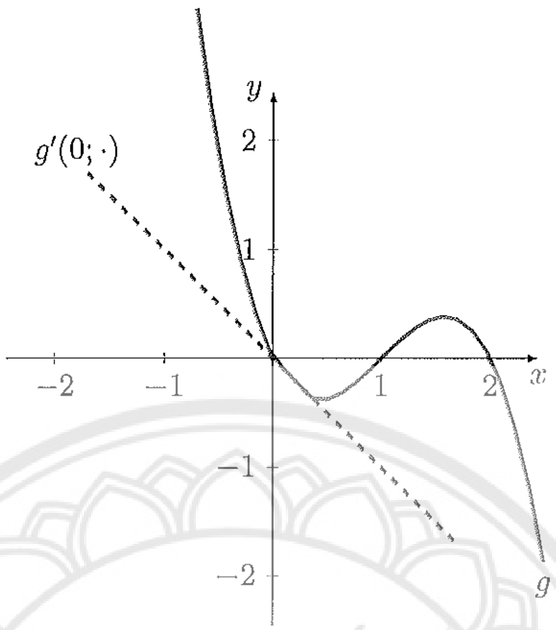


Figure 30: Plots of function g and its directional derivative in Example 4.1.13.

It is worth mentioning that from Remark 4.1.5, the following result can be deduced.

Corollary 4.1.14. *If we replace the condition $\bar{x} \in \operatorname{cl} L(f, <, f(\bar{x}))$ by $\operatorname{cl} L(f, <, f(\bar{x})) = L(f, \leq, f(\bar{x}))$ or $0 \notin \partial_T f(\bar{x})$, Theorem 4.1.9 is also true.*

As tangential convexity collapses to regularly locally Lipschitz setting and differentiability, the following two corollaries are immediately direct consequences as a special case of Theorem 4.1.9. We will also see how the condition $x \in \operatorname{cl} L(f, <, f(x))$ can be absent in differentiable case.

Corollary 4.1.15. *Given the nonlinear programming problem* (P) *and let the Slater's constraint qualification holds. Let $\bar{x} \in \mathcal{X}$ be a feasible solution satisfying the condition* (4.0.3) *and the functions $f, g_i : \mathbb{R}^n \to \mathbb{R}$, $i \in I(\bar{x})$, be locally Lipschitz and regular in the sense of Clarke at $\bar{x}$. Assume that $0 \notin \partial^o g_i(\bar{x})$ for all $i \in I(\bar{x})$.*

(i) *If $\bar{x}$ is a global minimizer then there exist $\lambda_i \geq 0$, $i \in I$ such that*

i) $0 \in \partial^o f(\bar{x}) + \sum_{i \in I} \lambda_i \partial^o g_i(\bar{x})$,

ii) $\lambda_i g_i(\bar{x}) = 0, \ \forall i \in I$.

(ii) *Conversely, if $\bar{x}$ is a non-trivial KKT point such that $\bar{x} \in \operatorname{cl} L(f,<,f(\bar{x}))$, and $L(f,<,f(\bar{x}))$ is convex then $\bar{x}$ is a global minimizer.*

Corollary 4.1.16. [52, Theorem 1] *Given the nonlinear programming problem* (P) *and let the Slater's constraint qualification holds. Let $\bar{x} \in \mathcal{X}$ be a feasible solution satisfying the condition* (4.0.3) *and the functions $f, g_i : \mathbb{R}^n \to \mathbb{R}$, $i \in I$, be differentiable functions. Assume that $\nabla g_i(\bar{x}) \neq 0$ for all $i \in I(\bar{x})$.*

(i) *If $\bar{x}$ is a global minimizer then there exist $\lambda_i \geq 0$, $i \in I$ such that*

i) $\nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) = 0$,

ii) $\lambda_i g_i(\bar{x}) = 0, \ \forall i \in I$.

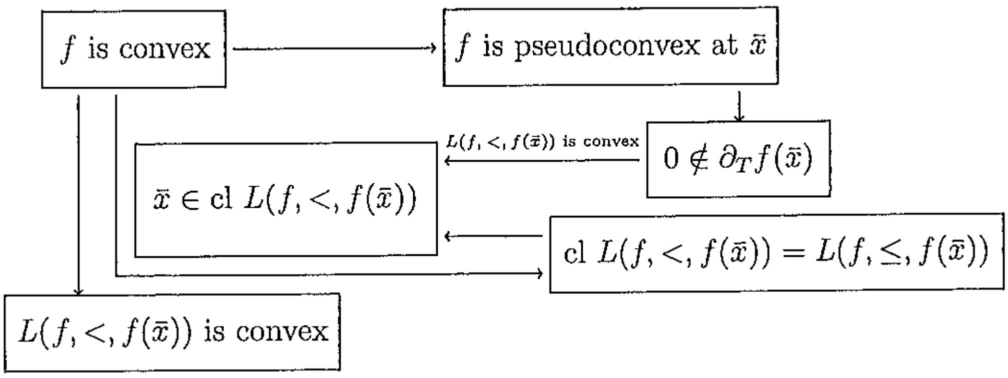
(ii) *Conversely, if $\bar{x}$ is a non-trivial KKT point, and $L(f,<,f(\bar{x}))$ is convex then $\bar{x}$ is a global minimizer.*

Proof. Owing to $\bar{x}$ is a non-trivial KKT point,

$$-\frac{1}{\sum_{i \in I(\bar{x})} \lambda_i} \nabla f(\bar{x}) = \sum_{i \in I(\bar{x})} \frac{\lambda_i}{\sum_{i \in I(\bar{x})} \lambda_i} \nabla g_i(\bar{x}) \in \operatorname{co}\left(\bigcup_{i \in I(x)} \{\nabla g_i(x)\} \right).$$

In view of Remark 4.1.10, we obtain that $\nabla f(\bar{x}) \neq 0$. The desired result will follows by the virtue of Lemma 4.1.2. $\square$

To this end of this section, we would like to summarize the relationship of the several conditions, which were considered in this paper, for KKT optimality conditions whenever $f : \mathbb{R}^n \to \mathbb{R}$, $\bar{x} \in \mathbb{R}^n$ and $L(f,<,f(\bar{x})) \neq \emptyset$:



4.2 Multi-objective optimization problems with cone constraints

In this section, we deal with a class of differentiable multi-objective optimization problems (MOPC) over cone constraints without the convexity of the feasible set, and the cone-convexity of objectives and constraint functions. Precisely stated, we will be mainly concerned with the multi-objective optimization problem with cone constraint (MOPC) given as

$$K - \min_{x \in \mathbb{R}^n} \{f(x) : -g(x) \in Q\}, \tag{MOPC}$$

where $f := (f_1, f_2, \ldots, f_p) : \mathbb{R}^n \to \mathbb{R}^p$ and $g := (g_1, g_2, \ldots, g_m) : \mathbb{R}^n \to \mathbb{R}^m$, are differentiable functions, K and Q are closed convex cones with non-empty interiors in $\mathbb{R}^p$ and $\mathbb{R}^m$, respectively. Let

$$\mathcal{X}_C := \{x \in \mathbb{R}^n : -g(x) \in Q\} \tag{4.2.1}$$

be the set of all feasible solutions of (MOPC). The notation "K – Minimize "refers to the **weak Pareto minimum** (resp. **Pareto minimum**) with respect to the ordering cone K for the problem (MOPC), namely a point $x^* \in \mathcal{X}_C$ such that for every $x \in \mathcal{X}_C$,

$$f(x^*) - f(x) \notin \text{int}K \text{ (resp. } f(x^*) - f(x) \notin K \backslash \{0\}).$$

For a closed convex cone $K \subseteq \mathbb{R}^p$, let us recall some known definitions in the literature that a vector valued function $f := (f_1, f_2, \ldots, f_p) : \mathbb{R}^n \to \mathbb{R}^p$ is said to be

(i) K-**convex** at a point $x^* \in \mathbb{R}^n$ if for every $x \in \mathbb{R}^n$,

$$f(x) - f(x^*) - \nabla f(x^*)(x - x^*) \in K.$$

(ii) K-**pseudoconvex** [81,82] at a point $x^* \in \mathbb{R}^n$ if for every $x \in \mathbb{R}^n$,

$$-\nabla f(x^*)(x - x^*) \notin \text{int}K \Rightarrow f(x^*) - f(x) \in \text{int}K,$$

where $\nabla f(x^*) := (\nabla f_1(x^*), f_2(x^*), \ldots, \nabla f_p(x^*))^T$ signifies the Jacobian of f. If f is K-convex (K-pseudoconvex) at every point $x^* \in \mathbb{R}^n$ then f is said to be K-convex (resp. K-pseudoconvex) on $\mathbb{R}^n$.

Recall that a feasible point $x^* \in \mathcal{X}_C$ is said to be a **KKT point** if there exist multipliers $\theta \in K^*\backslash\{0\}$ and $\lambda \in Q^*$ such that the following Karush-Kuhn-Tucker (KKT) optimality conditions hold:

(i) $\nabla f(x^*)^T\theta + \nabla g(x^*)^T\lambda = 0$;

(ii) $\langle \lambda, g(x^*) \rangle = 0$.

Also, the above feasible point x^* is also called a **non-trivial** KKT point if the corresponding λ is a non-zero vector.

The following constraint qualification for the problem (MOPC) is well known (see [55]).

Definition 4.2.1. The feasible set $\mathcal{X}_C$ as in (4.2.1) is said to satisfy **Slater-type cone constraint qualification** at $x \in \mathcal{X}_C$ if there exists $\hat{x} \in \mathbb{R}^n$ such that

$$g(x) + \nabla g(x)(\hat{x} - x) \in -\text{int}Q.$$

It is worth noting that, in contrast with the Slater-type constraint qualification to be introduced next, Slater-type cone constraint qualification is associated with a given feasible solution.

This part of the work has been motivated by a paper of Suneja et al. [53]. With the introduction of scalar convex optimization without convexity of constraint functions by Lasserre [46], the authors have established KKT optimality conditions for weak Pareto minimum (resp. Pareto minimum) of some classes of multi-objective convex optimization problems. In fact, the authors have shown that even if the convex feasible set is not necessarily described by cone-convex constraint, the Slater-type cone constraint qualification renders the KKT optimality conditions both necessary and sufficient.

Our intention is to weaken the convexity assumptions to the problem (MOPC). The key feature of this study that distinguishes from the scalar problems is that only non-degeneracy at the point x^* under consideration (see Assumption 1) is assumed, without Slater-type cone constraint qualification as well as Slater-type constraint qualification, which is given as follows:

Definition 4.2.2. The feasible set $\mathcal{X}_C$ as in (4.2.1) is said to satisfy **Slater-type constraint qualification** if there exists $\hat{x} \in \mathbb{R}^n$ such that $-g(\hat{x}) \in \text{int}Q$.

So, at first, it should be to investigate the connections among non-degeneracy condition, Slater-type cone constraint qualification, and Slater-type constraint qualification. Afterwards, we will establish necessary and sufficient KKT optimality conditions for a weak Pareto minimum of (MOPC). In addition, we also establish sufficient conditions for guaranteeing a weak Pareto minimum to be a Pareto minimum of the problem (MOPC). Further, illustrative examples are also provided to demonstrate that our results generalize and improve the corresponding known results obtained in [53] for the problem (MOPC) in some appropriate situations. To arrive there, we need the following lemma, which will be crucial in the sequel.

Lemma 4.2.3. [83] *Let $\mathcal{X}_C$ be as in* (4.2.1). *Assume that $\mathcal{X}_C$ is nearly convex at a feasible point $x^* \in \mathcal{X}_C$. Then for every $\lambda \in Q^*\backslash\{0\}$ for which $\langle \lambda, g(x^*)\rangle = 0$, one has*

$$\langle \nabla g(x^*)^T\lambda, v - x^*\rangle \leq 0 \text{ for all } v \in \mathcal{X}_C.$$

Proof. Suppose on contrary that there exist $\lambda \in Q^*\backslash\{0\}$ for which $\langle \lambda, g(x^*)\rangle$ and $v \in \mathcal{X}_C$ such that

$$\langle \nabla g(x^*)^T\lambda, v - x^*\rangle > 0.$$

By defining $h(x) := \langle \lambda, g(x)\rangle$ for all $x \in \mathbb{R}^n$, we have by (2.1.1) that $\nabla h(x) = \nabla g(x^*)^T\lambda$ for all $x \in \mathbb{R}^n$. So, in view of Remark 2.1.16, there exists $\delta > 0$ such that

$$\langle \lambda, g(x^* + t(v - x^*))\rangle = h(x^* + t(v - x^*)) > 0, \ \forall t \in]0, \delta[.$$

This together with the condition (4.0.3) in turn gives us that there exists some t_l small enough such that

$$\langle \lambda, g(x^* + t_l(v - x^*))\rangle > 0 \tag{4.2.2}$$

and $x^* + t_l(v - x^*) \in \mathcal{X}_C$. The latter means that $-g(x^* + t_l(v - x^*)) \in Q$ and consequently, $\langle \lambda, g(x^* + t_l(v - x^*))\rangle \leq 0$, which contradicts (4.2.2). □

4.2.1 Constraint qualifications

In this subsection, we present the constraint qualifications that are used to derive the KKT conditions for (MOPC) and their connections. At first, we recall one of constraint qualifications the so-called **non-degeneracy condition** at some feasible point $x^* \in \mathcal{X}_C$ in the vector setting, which introduced in [53].

Assumption 1: (Non-degeneracy condition [53]) Consider (MOP), for every $\lambda \in Q^*\backslash\{0\}$,

$$\nabla g(x^*)^T \lambda \neq 0 \text{ whenever } \langle \lambda, g(x^*) \rangle = 0.$$

Remark 4.2.4 (Sufficient condition for non-degeneracy condition to be valid). Note that if the Slater-type cone constraint qualification at x^* holds, then the non-degeneracy condition is satisfied at x^*. Indeed, if there exists $\hat{x} \in \mathbb{R}^n$ such that $g(x^*)+\nabla g(x^*)(\hat{x}-x^*) \in -\text{int}Q$, then for every $\lambda \in Q^*\backslash\{0\}$ for which $\langle \lambda, g(x^*) \rangle = 0$, one has $\langle \nabla g(x^*)^T\lambda, \hat{x} - x^* \rangle = \langle \lambda, g(x^*) \rangle + \langle \nabla g(x^*)^T\lambda, \hat{x} - x^* \rangle < 0$ which implies that $\nabla g(x^*)^T \lambda \neq 0$.

Remark 4.2.5. The Slater's condition can also be guaranteed by the Slater-type cone constraint qualification at some point x^* as well. To see this, we have from the Slater-type cone constraint qualification that $\nabla g(x^*)(\hat{x} - x^*) \in -\text{int}Q - g(x^*)$ for some $\hat{x} \in \mathbb{R}^n$. So, for any $\lambda \in Q^*\backslash\{0\}$, it then follows from Lemma 2.2.15(ii) that

$$\langle \lambda, g(x^*) \rangle + \langle \lambda, \nabla g(x^*)(\hat{x} - x^*) \rangle < 0.$$

This together with the fact that

$$\frac{\langle \lambda, g(x^* + t(\hat{x} - x^*)) \rangle - \langle \lambda, g(x^*) \rangle}{t} \to \langle \nabla g(x^*)^T \lambda, \hat{x} - x^* \rangle \text{ as } t \to 0^+,$$

in turn implies that for some $t_0 > 0$ sufficiently small, it holds

$$\langle \lambda, g(x^* + t_0(\hat{x} - x^*)) \rangle < (1 - t_0)(\langle \lambda, g(x^*) \rangle + \langle \lambda, \nabla g(x^*)(\hat{x} - x^*) \rangle) < 0.$$

Again, by using Lemma (2.2.15)(ii), we get that $-g(x^* + t_0(\hat{x} - x^*)) \in \text{int } Q$, and hence, the Slater's constraint qualification has been justified.

Now, we present some sufficient conditions for the Slater-type cone constraint qualification to be valid.

Theorem 4.2.6. [83] *Let $\mathcal{X}_C$ be as in* (4.2.1). *Assume that the Slater-type constraint qualification holds and the condition* (4.0.3) *is satisfied at a feasible point $x^* \in \mathcal{X}_C$. If the non-degeneracy condition holds at x^*, then the Slater-type cone constraint qualification also holds at x^*.*

Proof. Suppose that the non-degeneracy condition holds at x^*. Assume on contrary that for every $x \in \mathbb{R}^n$, one has $g(x^*) + \nabla g(x^*)(x - x^*) \notin -\mathrm{int}Q$, equivalently,

$$-[g(x^*) + \nabla g(x^*)(\mathbb{R}^n - x^*)] \cap \mathrm{int}Q = \emptyset.$$

So, by the convex separation theorem (Theorem 2.2.3), there exists $\lambda \in \mathbb{R}^m \backslash \{0\}$ such that

$$\langle \lambda, g(x^*) \rangle + \langle \lambda, \nabla g(x^*)(x - x^*) \rangle + \langle \lambda, y \rangle \geq 0, \ \forall x \in \mathbb{R}^n, \ \forall y \in Q. \tag{4.2.3}$$

By taking $x := x^*$ and $y := 0$ in (4.2.3), we would have $\langle \lambda, g(x^*) \rangle = 0$. Hence, with regard to (4.2.3) with $x := x^*$, we get $\lambda \in Q$. Therefore, in view of (4.2.3), we find a vector $\lambda \in Q^* \backslash \{0\}$ with $\langle \lambda, g(x^*) \rangle = 0$ such that

$$\langle \nabla g(x^*)^T \lambda, x - x^* \rangle \geq 0, \ \forall x \in \mathbb{R}^n. \tag{4.2.4}$$

On the other hand, by assumption, there exists $\hat{x} \in \mathbb{R}^n$ such that $-g(\hat{x}) \in \mathrm{int}Q$. Then, since g is continuous at $\hat{x}$, there exists $r > 0$ such that $g(\hat{x} + ru) \in -Q$ for all $u \in \mathbb{B}(0, 1)$. Consequently, $\hat{x} + ru \in \mathcal{X}_C$ for all $u \in \mathbb{B}(0, 1)$. So, as $x^* \in \mathcal{X}_C$ and x^* satisfies the condition (4.0.3), we conclude from Lemma 4.2.3 that

$$\langle \nabla g(x^*)^T \lambda, \hat{x} + ru - x^* \rangle \leq 0, \ \forall u \in \mathbb{B}(0, 1). \tag{4.2.5}$$

In particular, put $u := 0 \in \mathbb{B}(0, 1)$, one has $\langle \nabla g(x^*)^T \lambda, \hat{x} - x^* \rangle \leq 0$. Thus, with regard to (4.2.4), $\langle \nabla g(x^*)^T \lambda, \hat{x} - x^* \rangle = 0$, and hence we deduce from (4.2.5) that

$$\langle \nabla g(x^*)^T \lambda, u \rangle \leq 0, \ \forall u \in \mathbb{B}(0, 1).$$

So, $\nabla g(x^*)^T \lambda$ must ultimately be zero vector, which contradicts the validity of non-degeneracy condition at x^*. □

Remark 4.2.7. In the absence of the condition (4.0.3) at x^*, the validity of both Slater-type constraint qualification and the non-degeneracy condition at x^* does not guarantee the validity of Slater-type cone constraint qualification at x^*, for instance, let $x := (x_1, x_2) \in \mathbb{R}^2$, $Q := \mathbb{R}^2_+$ and $g(x) := (x_2^3 + x_2 - x_1, x_1 - x_2)$. We see that $g(-3,-2) = (-7,-1) \in -\text{int}Q$, that is, Slater's condition holds. Also, one has

$$\nabla g(x) = \begin{pmatrix} -1 & 3x_2^2+1 \\ 1 & -1 \end{pmatrix}, \ \forall x \in \mathbb{R}^2,$$

and a short calculation shows that the non-degeneracy holds at $x^* := (0,0) \in \mathcal{X}_C$, while the condition (4.0.3) together with the Slater-type cone constraint qualification is invalid at x^*. In fact, let us consider $x_0 := (-2,-1) \in \mathcal{X}_C$ and arbitrary sequence $\{t_l\}_{l\in\mathbb{N}} \subset]0,+\infty[$ such that $t_l \to 0$ as $l \to +\infty$. So, $t_{l_0} < 1$ for some $l_0 \in \mathbb{N}$ and $x^* + t_{l_0}(x_0 - x^*) = t_{l_0}x_0 \notin \mathcal{X}_C$. Otherwise, we have that

$$t_{l_0}(1-t_{l_0})(1+t_{l_0}) = (-t_{l_0})^3 + (-t_{l_0}) - (-2t_{l_0}) \le 0,$$

whence, $1 \le t_{l_0}$. This contradicts to the fact that $t_{l_0} < 1$. In addition, we can not find out $\hat{x} := (\hat{x}_1, \hat{x}_2) \in \mathbb{R}^2$ such that

$$\begin{pmatrix} -\hat{x}_1 + \hat{x}_2 \\ \hat{x}_1 - \hat{x}_2 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} = g(x^*) + \nabla g(x^*)(\hat{x} - x^*) \in -\text{int}Q.$$

Remark 4.2.8.

(i) It is worth noticing that there is a partial overlapping between Slater-type constraint qualification and non-degeneracy condition at a given point x^* in general. For example, it is easy to check that Slater-type constraint qualification fails to hold for $\mathcal{X}_C := \{x \in \mathbb{R}^n : -g(x) \in Q\}$, where $Q := \mathbb{R}^2_+$ and $g(x) := (-x_1 + x_2, x_1 - x_2)$ for all $x \in \mathbb{R}^2$, while non-degeneracy condition holds at $x^* := (0,0)$. In contrast, redefining $g(x) := (x_1^3 - x_2 + 1, -x_1^2 + x_2 - 1)$ for all $x \in \mathbb{R}^2$, we get $-g(-1,1) = (1,1) \in \text{int}Q$ and so, Slater-type constraint qualification holds. Now we see that non-degeneracy does not hold at x^*. Indeed, taking $\mu_0 := (1,1) \in Q^*\backslash\{0\}$ entails that $\langle \lambda_0, g(x^*) \rangle = 0$ and

$$\nabla g(x^*)^T \lambda_0 = \begin{pmatrix} 0 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

showing that non-degeneracy fails to hold at x^*.

(ii) In addition to the Q-convexity of g at a given point x^*, if Slater-type constraint qualification holds, then non-degeneracy condition is satisfied at x^*. To see this, suppose now by contradiction that there exists $\lambda_0 \in Q^*\backslash\{0\}$ satisfying $\langle \lambda_0, g(x^*)\rangle = 0$ and $\nabla g(x^*)^T\lambda_0 = 0$. It then follows from Q-convexity of g at x^* that $\langle \lambda_0, g(\hat{x})\rangle - \langle \lambda_0, g(x^*)\rangle = \langle \lambda_0, g(\hat{x})\rangle - \langle \lambda_0, g(x^*)\rangle - \langle \nabla g(x^*)^T\lambda_0, \hat{x} - x^*\rangle \geq 0$ for a Slater's point $\hat{x}$. This contradicts to the fact that $\langle \lambda_0, g(\hat{x})\rangle < 0 = \langle \lambda_0, g(x^*)\rangle$.

Let us recall that the set $\{x \in \mathbb{R}^n : \ g_i(x) \leq 0, \ i = 1, 2, \ldots, m\}$ is said to satisfy the

(i) **Mangasarian-Fromovitz constraint qualification** [84] at x^* if there exists $v \in \mathbb{R}^n$ such that $\langle \nabla g_i(x^*), v\rangle < 0$ for each $i \in I(x^*) := \{i \in \{1, 2, \ldots, m\} : g_i(x^*) = 0\}$.

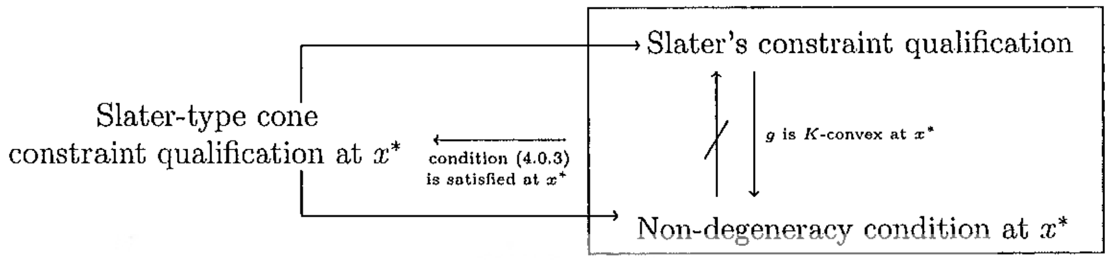
(ii) **Robinson constraint qualification** at x^* if $0 \in \text{int}\{g(x^*) + \nabla g(x^*)(\mathbb{R}^n - x^*) + \mathbb{R}^m_+\}$ where $g(x) := (g_1(x), g_2(x), \ldots, g_m(x))$.

Remark 4.2.9. In the case of $Q = \mathbb{R}^m_+ := \{(x_1, x_2, \ldots, x_m) \in \mathbb{R}^m : x_i \geq 0, \ \forall i = 1, 2, \ldots, m\}$, non-degeneracy conditions at x^* can be view as the Mangasarian-Fromovitz constraint qualification at x^* and non-degeneracy conditions at x^* in [46, 52] as well. Indeed,

$$\begin{aligned} & \exists v \in \mathbb{R}^n \text{ such that } \langle \nabla g_i(x^*), v\rangle < 0, \ \forall i \in I(x^*) \\ \Leftrightarrow \quad & 0 \notin \text{co}\{\nabla g_i(x^*) : i \in I(x^*)\} \\ \Leftrightarrow \quad & \forall \lambda := (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{R}^m_+\backslash\{0\} \text{ with } \lambda_i g_i(x^*) = 0, \ i = 1, 2, \ldots, m, \\ & \quad \text{one has } \sum_{i=1}^m \lambda_i \nabla g_i(x^*) \neq 0, \end{aligned}$$

and for each $i \in \{1, 2, \ldots, m\}$, by taking $\lambda := e_i$, where e_i is the unit vector in $\mathbb{R}^m$ with the ith component is 1 and the others 0, one has $\nabla g_i(x^*) \neq 0$ whenever $i \in I(x^*)$. Note that Slater-type cone constraint qualification at x^* also is equivalent to the Robinson constraint qualification at x^* [84, Lemma 2.99, p. 69]. Then, as the considered set $\{x \in \mathbb{R}^n : \ g_i(x) \leq 0, \ i = 1, 2, \ldots, m\}$ is not necessarily convex, one can notice that Theorem 4.2.6 extends [73, Theorem 2.1] to non-convex setting on the set $\{x \in \mathbb{R}^n : \ g_i(x) \leq 0, \ i = 1, 2, \ldots, m\}$.

To this end of this subsection, the relationship between proposed constraint qualifications can be summarized in following diagram whenever $x^* \in \mathcal{X}_C$:



4.2.2 KKT optimality conditions

In this subsection, we establish necessary and sufficient KKT optimality conditions for a weak Pareto minimum of (MOPC). In addition, we also establish sufficient conditions for guaranteeing a weak Pareto minimum to be a Pareto minimum of the problem (MOPC). We begin by the following lemma.

Lemma 4.2.10. [85, Lemma 1] Consider the problem (MOPC). If $x^* \in \mathcal{X}_C$ is a weak Pareto minimum of (MOPC), then there exist $\theta \in K^*$ and $\lambda \in Q^*$ not both zero such that

$$\langle \nabla f(x^*)^T \theta + \nabla g(x^*)^T \lambda, x - x^* \rangle \geq 0, \ \forall x \in \mathbb{R}^n$$

and

$$\langle \lambda, g(x^*) \rangle = 0.$$

Now, we also recall the following important result which can be found in [79] and will play a key role in deriving a feasible point to be a weak Pareto minimum as well as a Pareto minimum of (MOPC).

Proposition 4.2.11. [79, Proposition 2.2.] Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable at x^* with $\nabla f(x^*) \neq 0$. Then:

(i) $N(L(f, <, f(x^*)), x^*) = \{d \in \mathbb{R}^n : d = r\nabla f(x^*), \text{ for some } r \geq 0\}$ provided that $L(f, <, f(x^*))$ is convex.

(ii) $N(L(f, \leq, f(x^*)), x^*) = \{d \in \mathbb{R}^n : d = r\nabla f(x^*), \text{ for some } r \geq 0\}$ provided that $L(f, \leq, f(x^*))$ is convex.

Now, we are in the position to give necessary and sufficient KKT optimality conditions for a weak Pareto minimum of (MOPC).

Theorem 4.2.12. [83] *Consider the problem* (MOPC) *and let both Assumption 1 and the condition* (4.0.3) *be satisfied at a feasible point* x^*.

(i) *If* x^* *is a weak Pareto minimum of* (MOPC), *then* x^* *is a KKT point.*

(ii) *Conversely, if* x^* *is a non-trivial KKT point with multipliers* θ *and* λ, *and* $L(\langle \theta, f \rangle, <, \langle \theta, f(x^*) \rangle)$ *is convex then* x^* *is a weak Pareto minimum of* (MOPC), *where* $\langle \theta, f \rangle$ *is defined by*

$$\langle \theta, f \rangle(x) := \langle \theta, f(x) \rangle, \ \forall x \in \mathbb{R}^n.$$

Proof. (i) Let $x^* \in \mathcal{X}_C$ be a weak Pareto minimum of (MOPC). By Lemma 4.2.10, there exist $\theta \in K^*$ and $\lambda \in Q^*$ not both zero such that $\langle \lambda, g(x^*) \rangle = 0$ and

$$\langle \nabla f(x^*)^T \theta + \nabla g(x^*)^T \lambda, x - x^* \rangle \geq 0, \ \forall x \in \mathbb{R}^n. \tag{4.2.6}$$

As the inequality (4.2.6) holds for every $x \in \mathbb{R}^n$, we conclude that

$$\nabla f(x^*)^T \theta + \nabla g(x^*)^T \lambda = 0 \text{ and } \langle \lambda, g(x^*) \rangle = 0.$$

Moreover, we assert that $\theta \neq 0$. Otherwise, it follows in turn that $\lambda \neq 0$, which stands in contradiction to Assumption 1, and therefore, $\theta \neq 0$.

(ii) Let $x^* \in \mathcal{X}_C$ be an arbitrary non-trivial KKT point, i.e.,

$$\nabla f(x^*)^T \theta + \nabla g(x^*)^T \lambda = 0, \ \langle \lambda, g(x^*) \rangle = 0,$$

for some non-zero vectors $\theta \in \mathbb{R}^p$, $\lambda \in \mathbb{R}^m$. This together with Assumption 1 implies that $\nabla f(x^*)^T \theta$ must ultimately be non-zero vector. It can be seen that if the set $L(\langle \theta, f \rangle, <, \langle \theta, f(x^*) \rangle)$ is empty, then x^* actually is a weak Pareto minimum of (MOPC). In fact, if x^* is not a weak Pareto minimum of (MOPC), there exists $x \in \mathcal{X}_C$ such that $f(x^*) - f(x) \in \text{int } K$. So, by the virtue of Lemma 2.2.15(ii),

$\langle\theta, f(x^*)\rangle > \langle\theta, f(x)\rangle$, which contradicts to the fact that $L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle) = \emptyset$. Let us consider in the case $L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle) \neq \emptyset$. Applying Proposition 4.2.11(i) with $h(x) := \langle\theta, f(x)\rangle$, we obtain that $\nabla h(x^*) = \nabla f(x^*)^T\theta$ and

$$\langle\nabla f(x^*)^T\theta, u - x^*\rangle \leq 0,\ \forall u \in L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle). \tag{4.2.7}$$

Therefore, by Lemma 4.2.3,

$$\langle\nabla f(x^*)^T\theta, v - x^*\rangle = -\langle\nabla g(x^*)^T\theta, v - x^*\rangle \geq 0,\ \forall v \in \mathcal{X}_C. \tag{4.2.8}$$

Note that,

$$\{y \in \mathbb{R}^n : f(x^*) - f(y) \in \text{int } K\} \subseteq L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle).$$

Thus, in order to obtain that x^* is a weak Pareto minimum of (MOP), it suffices to show that $\mathcal{X}_C \subseteq \mathbb{R}^n \backslash L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle)$ or consequently,

$$L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle) \cap \mathcal{X}_C = \emptyset.$$

Suppose, ad absurdum, $L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle) \cap \mathcal{X}_C \neq \emptyset$. Thus, from (4.2.7) and (4.2.8) we get the assertion $\langle\nabla f(x^*)^T\theta, w - x^*\rangle = 0$ for any $w \in L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle) \cap \mathcal{X}_C$. Furthermore, as the set $L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle)$ being open, for each $d \in \mathbb{R}^n$ we can find $t > 0$ small enough such that $w + td \in L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle)$. Hence,

$$t\langle\nabla f(x^*)^T\theta, d\rangle = \langle\nabla f(x^*)^T\theta, w + td - x^*\rangle - \langle\nabla f(x^*)^T\theta, w - x^*\rangle \leq 0.$$

This means $\nabla f(x^*)^T\theta = 0$, a contradiction. Thus, $L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle) \cap \mathcal{X}_C = \emptyset$, and x^* is a weak Pareto minimum of (MOPC) as desired. □

Remark 4.2.13. It is worth mentioning here that Proposition 4.2.11 plays a significant role in Theorem 4.2.12(ii) for ensuring a feasible point x^* to be a weak Pareto minimum of (MOPC). Beside, non-degeneracy condition (Assumption 1) at x^* need to be assumed for guaranteeing $\theta^T\nabla f(x^*) \neq 0$ with correspond to multiplier vector $\theta \in K^*\backslash\{0\}$. In contrast, it generally does not need constraint qualification to establish the sufficient optimality conditions. Therefore, it might be reasonably assumed the assertion $\nabla f(x^*)^T\theta \neq 0$ instead of assuming the non-degeneracy condition at x^*. However, keeping in mind the fact that we need to justify the convexity of $L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle)$ with the same choice θ, and so in this case the multiplier vector θ turn out to be difficult to determine for which satisfying all conditions in Theorem 4.2.12(ii) simultaneously. This being a reason why non-degeneracy condition make used in Theorem 4.2.12(ii). Another reason is that non-degeneracy condition is actually justified to check a feasible point that can be a weak Pareto minimum of (MOPC) or not before to justify sufficient optimality conditions.

We now demonstrate with the following example to guarantee that Theorem 4.2.12 is indicated to be conveniently applied in some cases where Theorem 3.1 and Theorem 3.2 of [53] cannot be used even when the feasible set $\mathcal{X}_C$ is convex.

Example 4.2.14. Let us consider the following muti-objective optimization problem (MOPC) over cones:

$$K - \min_{x\in\mathbb{R}}\{f(x) := (x+1, x^3 - 5x^2 + 8x - 3) : x \in \mathcal{X}_C\},$$

where $g(x) := (x-1, x^2 - x - 1)$, $K := \mathbb{R}^2_+$ and $Q := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \leq 0,\ x_2 \leq x_1\}$. A straightforward calculation shows that:

- $\mathcal{X}_C = [2, +\infty[$,
- $K^* = K$,
- $Q^* = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \leq 0,\ x_2 \leq -x_1\}$,
- $x^* := 2$ satisfies the non-trivial KKT conditions by taking $\theta := (2, 0)$ and $\lambda := (1, -1)$,
- $L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle) =]-\infty, 2[$ is convex,
- It is easily to seen that Assumption 1 and the condition (4.0.3) are satisfied.

Applying Theorem 4.2.12 (ii), we can conclude that x^* is a weak Pareto minimum of (MOPC). However, it can be checked that g is not Q-convex, i.e.

$$g(1) - g(2) - \nabla g(2)(1-2) = (0, 1) \notin Q,$$

but the feasible set $\mathcal{X}_C$ is convex. Furthermore, the function f is not K-pseudoconvex at $x^* := 2$, because if we take $x := 0$ then

$$-\nabla f(x^*)(x - x^*) = (2, 0) \notin \text{int}K, \text{ but } f(x^*) - f(x) = (2, 4) \in \text{int}K.$$

Hence, the corresponding results [53] is not applicable.

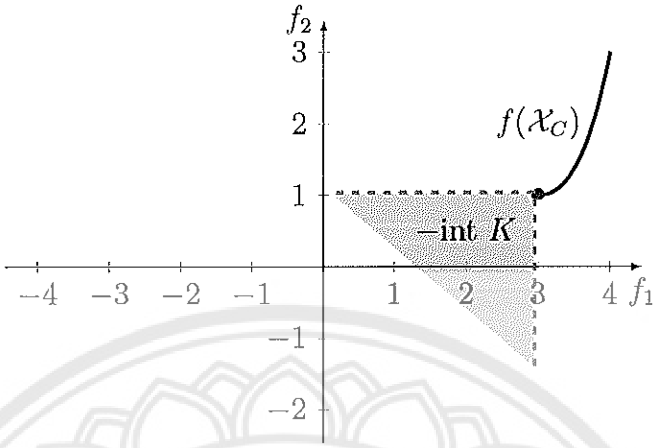


Figure 31: Illustration of the behavior of weak Pareto minimality in Example 4.2.14.

Note that the multiplier vector μ is assumed to be non-zero vector (the non-triviality of the KKT conditions) in order to ensure that $\nabla f(x^*)^T\theta \neq 0$ in Theorem 4.2.12(ii). The following example demonstrates that this assumption cannot be dropped.

Example 4.2.15. Let $f(x) := (x+1, -(x-2)^3)$, $g(x) := (x^2-1, 2x-1)$, $K := \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq -x_1,\ x_1 \geq 0\}$ and $Q := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq x_2,\ x_1 \geq 0\}$. It is not hard to check that $\mathcal{X}_C = [1, 2]$, $x^* := 2$ is a KKT point with $\theta := (0, -1)$ and $\lambda := (0, 0)$, and all the conditions in Theorem 4.2.12 (ii) are fulfilled. However x^* is not even a weak Pareto minimum, i.e., if we take $x := \frac{3}{2}$ then $f(x^*) - f(x) = (3, 0) - (\frac{5}{2}, \frac{1}{8}) = (\frac{1}{2}, -\frac{1}{8}) \in \text{int}K$. The main reason is that x^* is not a non-trivial KKT point.

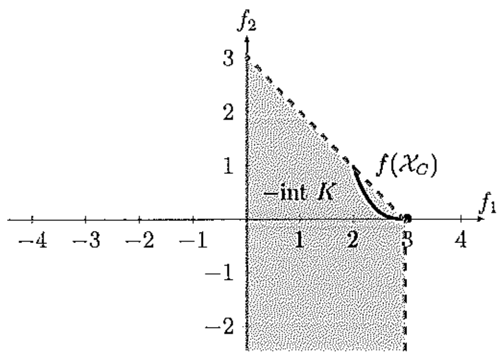


Figure 32: Illustration of Example 4.2.15.

To appreciate Theorem 4.2.12 we present an example that is applicable while the aforementioned result in [53] is not.

Example 4.2.16. Consider the following multi-objective optimization problem (MOPC) over cones:

$$K - \min_{x\in\mathbb{R}}\{f(x) := (x^2 - 1, -x^3 + 5x^2 - 8x + 5) : x \in \mathcal{X}_C\},$$

where $g(x) := (x^3+x^2+x, x^3+2x^2-5x+8)$, $K := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0,\ x_2 \leq x_1\}$ and $Q := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \leq 0,\ x_2 \leq x_1\}$. Evidently, f, and g are not K, and Q-convex, respectively. Indeed, $f(1) - f(0) - \nabla f(0)(1-0) = (1, 4) \notin K$, and $g(1) - g(0) - \nabla g(0)(1-0) = (2, 3) \notin Q$. It is easy to verify that $\mathcal{X}_C = [0, 2] \cup [4, +\infty[$. Then we have already seen that the feasible set $\mathcal{X}_C$ is not convex. Therefore, the results in [53] cannot be applicable. However, it is not hard to verify that

- $K^* = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \leq 0,\ x_2 \geq -x_1\}$,
- $Q^* = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \leq 0,\ x_2 \leq -x_1\}$,
- $x^* := 0$ satisfies the non-trivial KKT conditions by taking $\theta := (1, -1)$ and $\lambda := (-8, 0)$,
- Assumption 1 and the condition (4.0.3) are satisfied,
- $L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle) =]-\infty, 0[$, which is convex.

Hence, Theorem 4.2.12 (ii) indicates that x^* is a weak Pareto minimum of (MOPC).

Next, we will see now how the convexity of $L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle)$ together with the strict convexity of $L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle)$ at a non-trivial KKT point x^* possess x^* to be a Pareto minimum of (MOPC). To do this, we recall first the following notion of strict convexity.

Definition 4.2.17. A set $A \subseteq \mathbb{R}^n$ is called **strictly convex** at $x \in A$ if $\langle\xi, y - x\rangle < 0$ for every $y \in A\backslash\{x\}$ and $\xi \in N(A, x)\backslash\{0\}$.

It is worth noting that the strict convexity of A at some point x does not guarantee the convexity of A. For instance, the set $A := \{(x_1, x_2) \in \mathbb{R}^2 : x_2 > 0\} \cup \{(0, 0)\}$ is strictly convex at $(0, 0)$ while A is not convex.

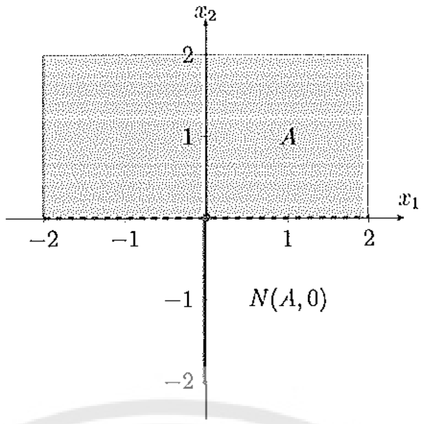


Figure 33: Illustration of nonconvex set that it is strictly convex.

Theorem 4.2.18. [83] *Consider the problem* (MOPC) *and let both Assumption 1 and the condition* (4.0.3) *be satisfied at a feasible point x^*. If x^* is a non-trivial KKT point with multipliers θ and λ, $L(\langle\theta, f\rangle, \leq, \langle\theta, f(x^*)\rangle)$ is convex, and additionally $L(\langle\theta, f\rangle, \leq, \langle\theta, f(x^*)\rangle)$ is strictly convex at x^*, then x^* is a Pareto minimum of* (MOPC).

Proof. In a similar manner of the second argument as the proof of Theorem 4.2.12, by the KKT conditions and Proposition 4.2.11(ii), we arrive at the following assertion

$$\langle\nabla f(x^*)^T\theta, v-x^*\rangle \geq 0 \geq \langle\nabla f(x^*)^T\theta, u-x^*\rangle, \forall v \in \mathcal{X}_C, \forall u \in L(\langle\theta, f\rangle, \leq, \langle\theta, f(x^*)\rangle), \tag{4.2.9}$$

and $\nabla f(x^*)^T\theta \neq 0$. To establish the desired results, we argue first by using Lemma 2.2.15(i) that

$$\{y \in \mathbb{R}^n : f(x^*) - f(y) \in K\backslash\{0\}\} \subseteq L(\langle\theta, f\rangle, \leq, \langle\theta, f(x^*)\rangle)\backslash\{x^*\}.$$

Thus, we only need to justify this containment

$$\mathcal{X}_C \subseteq \mathbb{R}^n\backslash\Big(L(\langle\theta, f\rangle, \leq, \langle\theta, f(x^*)\rangle)\backslash\{x^*\}\Big)$$

We argue by contradiction that there exists some $w \in \mathcal{X}_C$ such that $w \neq x^*$ and $w \in L(\langle\theta, f\rangle, \leq, \langle\theta, f(x^*)\rangle)$. Taking (4.2.9) into account we actually have

$$\langle\nabla f(x^*)^T\theta, w - x^*\rangle = 0.$$

Furthermore, as $\nabla f(x^*)^T\theta \in N(L(\langle\theta, f\rangle, \leq, \langle\theta, f(x^*)\rangle), x^*)\backslash\{0\}$ (by the second inequality in (4.2.9)) and $L(\langle\theta, f\rangle, \leq, \langle\theta, f(x^*)\rangle)$ is strictly convex, then $\langle\nabla f(x^*)^T\theta, w - x^*\rangle < 0$. This is a contradiction, and thereby implying that x^* is a Pareto minimum of (MOPC). □

Remark 4.2.19. In Example 4.2.16 with $\theta := (1, -1)$, it is evident that $L(\langle\theta, f\rangle, \leq, \langle\theta, f(x^*)\rangle)$ is strictly convex at $x^* := 0$, by Theorem 4.2.18, and hence x^* is a Pareto minimum of (MOPC) (see the below figure).

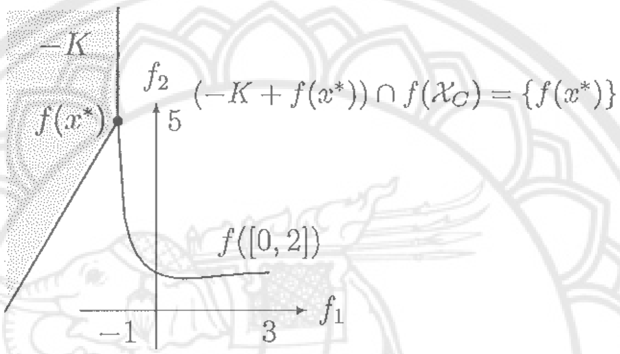


Figure 34: Illustration of Pareto minimality in Example 4.2.16.

Remark 4.2.20. It should be noted that to obtain a Pareto minimum from a drawback (see [53,55] and other references therein), the multiplier vector λ in KKT conditions need to be taken from the strict positive dual cone of K, K^{s^*}, which defined as

$$K^{s^*} := \{\xi \in \mathbb{R}^n : \langle\xi, x\rangle > 0 \text{ for all } x \in K\backslash\{0\}\}.$$

However, in this case study the multiplier vector θ is not necessarily to take from the strict positive dual cone. In fact, as K defined in Example 4.2.16 and $\theta := (1, -1)$, Then elementary calculations give us

$$K^{s^*} = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 > 0,\ x_2 > -x_1\}$$

and so, $\theta \notin K^{s^*}$.

To this end, we now give an example showing that the strict convexity of $L(\langle\theta, f\rangle, \leq, \langle\theta, f(x^*)\rangle)$ with corresponding multiplier θ is essential for x^* under the question to be a Pareto minimum of (MOPC) in Theorem 4.2.18.

Example 4.2.21. Let $x := (x_1, x_2) \in \mathbb{R}^2$, $f(x) := (x_1^2, x_2 - x_1)$, $g(x) := (-x_1^3 + 3x_1 + x_2, x_1 - x_2)$ and $K = Q := \{x \in \mathbb{R}^2 : x_1 \leq 0,\ x_2 \geq 0\}$. It is easy to check that the feasible set $\mathcal{X}_C$ is not convex and the condition (4.0.3) is valid at $x^* := (1, 1) \in \mathcal{X}_C$. Then elementary calculations give us

- $K^* = Q^* = K$,
- $g(x^*) = (3, 0)$, $\nabla g(x^*) = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$, $f(x^*) = (1, 0)$, $\nabla f(x^*) = \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix}$,
- x^* satisfies Assumption 1 and the non-trivial KKT conditions by taking $\theta = \lambda := (0, 1)$,
- $L(\langle \theta, f \rangle, <, \langle \theta, f(x^*) \rangle) = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 < x_1\}$ and $L(\langle \theta, f \rangle, \leq, \langle \theta, f(x^*) \rangle) = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \leq x_1\}$ are convex sets.

By Theorem 4.2.12 (ii), we can conclude that x^* is a weak Pareto minimum of (MOPC). However, the set $L(\langle \theta, f \rangle, \leq, \langle \theta, f(x^*) \rangle)$ is not a strictly convex set at x^*, i.e., it is clear that $N(L(\langle \theta, f \rangle, \leq, \langle \theta, f(x^*) \rangle), x^*) = \{(-r, r) \in \mathbb{R}^2 : r \geq 0\}$. So, by taking $\xi := (-1, 1) \in N(L(\langle \theta, f \rangle, \leq, \langle \theta, f(x^*) \rangle), x^*) \backslash \{(0, 0)\}$ and $y := (2, 2) \in L(\langle \theta, f \rangle, \leq, \langle \theta, f(x^*) \rangle) \backslash \{(0, 0)\}$, we have $\langle \xi, y - x^* \rangle = 0$. Actually, a point x^* is not even a Pareto minimum, i.e., if we take $x := (-2, -2) \in \mathcal{X}_C$, one has

$$f(x^*) - f(x) = (-3, 0) \in K \backslash \{(0, 0)\}.$$

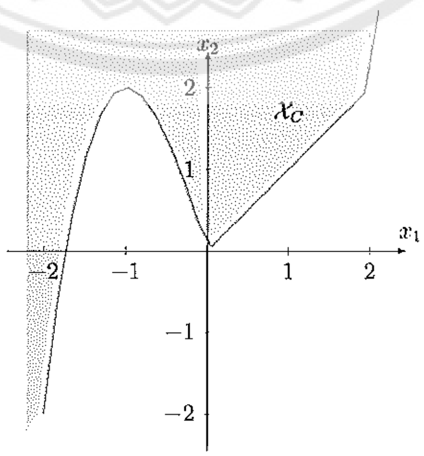


Figure 35: Illustration of constraint set $\mathcal{X}_C$ in Example 4.2.21.

Remark 4.2.22. It is worth noting that the convexity of $L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle)$ (resp. $L(\langle\theta, f\rangle, \leq, \langle\theta, f(x^*)\rangle))$ in Theorem 4.2.12 (resp. in Theorem 4.2.18) can be viewed as a generalized quasiconvexity of f at x^* due to the notion of $*$-*quasiconvexity* [86] in the sense that for each $\theta \in K^*$ the function $\langle\theta, f\rangle : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex. It is quite clear from the definition that $*$-quasiconvexity of f guarantees the convexity of the level set $L(\langle\theta, f\rangle, <, \langle\theta, f(x^*)\rangle)$ or of $L(\langle\theta, f\rangle, \leq, \langle\theta, f(x^*)\rangle)$. In fact, the function f in Example 4.2.21 is not $*$-quasiconvex, i.e., by taking $\theta := (-1, 1) \in K^*$ and $x := (1, 1)$, the set $L(\langle\theta, f\rangle, \leq, \langle\theta, f(x^*)\rangle)$ is non-convex. For related conditions for cone quasiconvex mappings we refer the reader to [68,87,88].

CHAPTER V

CONCLUSION

In this thesis, following the framework of robust optimization, we mainly concerned to characterize the robust optimal solution sets for uncertain convex optimization problems without convexity of constraint data uncertainty. We achieve this by investigating convex optimization problems without convexity of constraint in the absence of data uncertainty. We provide a new pseudo Lagrangian-type function which is constant on the optimal solution set. This property is still valid in the case of a pseudoconvex locally Lipschitz objective function. We also obtain some characterizations of the optimal solution set of all optimal solutions of a given problem. Afterwards, with slight consideration, characterizations of the robust optimal solution set for uncertain convex optimization problems with a robust convex constraint set described by locally Lipschitz constraints are obtained. Furthermore, by employing the linear scalarization, characterizations of weakly robust efficient solution set and properly robust efficient solution set of uncertain convex multi-objective optimization problems are also obtained.

Concerning the fruitful theories of quasi-convexity, we have established necessary and sufficient KKT optimality conditions of non-smooth optimization problems with inequality constraints without the presence of convexity of objective function, of constraint functions and of feasible set. In addition, for the differentiable multi-objective optimization problem (MOPC) over cone constraint, we have proposed constraint qualifications and discussed the relationship between them without the convexity of the feasible set, and the cone-convexity of objective and constraint functions. Finally, necessary and sufficient the Karush-Kuhn-Tucker optimality conditions for weak Pareto minimum as well as Pareto minimum of the problem (MOPC) are also obtained.

REFERENCES

REFERENCES

1. John F. Extremum problems with inequalities as subsidiary conditions: studies and essays presented to Courant R. on his 60th birthday. New York: Interscience Publishers; 1948. p. 187–204.
2. Kuhn HW, Tucker AW. Nonlinear programming. Proceedings of the Second Berkley Symposium on Mathematical Statistics and Probability. Berkeley: University of California Press; 1951. p. 481–92.
3. Karush W. Minima of functions of several variables with inequalities as side conditions [Master's thesis]. Chicago: University of Chicago; 1939.
4. Mangasarian OL. A simple characterization of solution sets of convex programs. Oper. Res. Lett. 1988;7:21–6.
5. Mangasarian OL. Error bounds for nondegenerate monotone linear complementarity problems. Math. Program. 1990;48:437–45.
6. Burke JV, Ferris M. Characterization of solution sets of convex programs. Oper. Res. Lett. 1991;10:57–60.
7. Burke JV, Ferris M. Weak sharp minima in mathematical programming. SIAM J. Control Optim. 1993;31:1340–59.
8. Feltenmark S, Kiwiel K. Dual applications of proximal bundle methods, including Lagrangian relaxation of nonconvex problems. SIAM J. Optim. 2000;10(3):697–721.
9. Jeyakumar V, Lee GM, Dinh N. Lagrange multiplier conditions characterizing optimal solution sets of cone-constrained convex programs. J. Optim. Theory Appl. 2004;123:83–103.
10. Miao XH, Chen JS. Characterizations of solution sets of cone-constrained convex programming problems. Optim. Lett. 2015;9:1433–45.
11. Jeyakumar V, Yang XQ. Characterizing the solution sets of pseudo-linear programs. J. Optim. Theory Appl. 1995;87:747–55.
12. Dinh N, Jeyakumar V, Lee GM. Lagrange multiplier characterizations of solution sets of constrained pseudolinear optimization problems. Optimization 2006;55:241–50.

13. Lalitha CS, Mehta M. Characterizations of solution sets of mathematical programs in terms of Lagrange multipliers. Optimization 2009;58:995–1007.

14. Mishra SK, Upadhyay BB, An, LTH. Lagrange multiplier characterizations of solution sets of constrained nonsmooth pseudolinear optimization problems. J. Optim. Theory Appl. 2014;160:763–77.

15. Ivanov VI. First order characterizations of pseudoconvex functions. Serdica Math. J. 2001;27:203–18.

16. Ivanov VI. Characterizations of the solution sets of generalized convex minimization problems. Serdica Math. J. 2003;29:1–10.

17. Castellani M, Giuli M. A characterization of the solution set of pseudoconvex extremum problems. J. Convex Anal. 2012;19:113–23.

18. Wu ZL, Wu SY. Characterizations of the solution sets of convex programs and variational inequality problems. J. Optim. Theory Appl. 2006;130:339–58.

19. Ivanov VI. Optimality conditions and characterizations of the solution sets in generalized convex problems and variational inequalities. J. Optim. Theory Appl. 2013;158:65–84.

20. Penot JP. Characterization of solution sets of quasiconvex programs. J. Optim. Theory Appl. 2003;117:627–36.

21. Suzuki S, Kuroiwa D. Characterizations of the solution set for quasiconvex programming in terms of Greenberg-Pierskalla subdifferential. J. Glob. Optim. 2015;62:431–41.

22. Suzuki S, Kuroiwa D. Characterizations of the solution set for non-essentially quasiconvex programming. Optim. Lett. 2017;11:1699–712.

23. Yang XM. On characterizing the solution sets of pseudoinvex extremum problems. J. Optim. Theory Appl. 2009;140:537–42.

24. Zhao KQ, Yang XM. Characterizations of the solution set for a class of nonsmooth optimization problems. Optim. Lett. 2013;7:685–94.

25. Liu CP, Yang XM. Characterizations of the approximate solution sets of nonsmooth optimization problems and its applications. Optim. Lett. 2015;9:755–68.

26. Son TQ, Dinh N. Characterizations of optimal solution sets of convex infinite programs. TOP. 2008;16:147–63.

27. Son TQ, Kim DS. A new approach to characterize the solution set of a pseudoconvex programming problem. J. Comput. Appl. Math. 2014;261:333–40.

28. Long XJ, Peng ZY, Wang XF. Characterizations of the solution set for nonconvex semi-infinite programming problems. J. Nonlinear Convex Anal. 2016;17:251–65.

29. Jeyakumar V, Lee GM, Dinh N. Characterizations of solution sets of convex vector minimization problems. Eur. J. Oper. Res. 2006;174:1380–95.

30. Ben-Tal A, Nemirovski A. Robust Optimization-methodology and applications. Math. Program. Ser. B 2002;92:453–80.

31. Beck A, Ben-Tal A. Duality in robust optimization: primal worst equals dual best. Oper. Res. Lett. 2009;37:1–6.

32. Ben-Tal A, Ghaoui LE, Nemirovski A. Robust Optimization. Princeton: Princeton University Press; 2009.

33. Bertsimas D, Brown DS, Caramanis C. Theory and applications of robust optimization. SIAM Rev. 2011;53:464–501.

34. Boţ RI, Jeyakumar V, Li GY. Robust duality in parametric convex optimization. Set-Valued Var. Anal. 2013;21:177–89.

35. Goberna MA, Jeyakumar V, Li GY, Lopez M. Robust linear semi-infinite programming duality. Math. Program. Ser. B 2013;139:185–203.

36. Lee GM, Son PT. On nonsmooth optimality theorems for robust optimization problems. Bull. Korean Math. Soc. 2014;51:287–301.

37. Lee GM, Lee JH. On nonsmooth optimality theorems for robust multiobjective optimization problems. J. Nonlinear Convex Anal. 2015;16:2039–52.

38. Jeyakumar V, Li GY. Characterizing robust set containments and solutions of uncertain linear programs without qualifications. Oper. Res. Lett. 2010;38:188–94.

39. Jeyakumar V, Li GY. Strong duality in robust convex programming: complete characterizations. SIAM J. Optim. 2010;20:3384–407.

40. Jeyakumar V, Li GY, Wang JH. Some robust convex programs without a duality gap. J. Convex Anal. 2013;20:377–94.

41. Jeyakumar V, Lee GM, Li GY. Characterizing robust solution sets of convex programs under data uncertainty. J. Optim. Theory Appl. 2015;164:407–35.

42. Li XB, Wang S. Characterizations of robust solution set of convex programs with uncertain data. Optim. Lett. 2018;12:1387–402.

43. Lee GM, Yao JC. On solution sets for robust optimization problems. J. Nonlinear Convex Anal. 2016;17(5):957–66.

44. Sun XK, Peng ZY, Guo XL. Some characterizations of robust optimal solutions for uncertain convex optimization problems. Optim. Lett. 2016;10:1463–78.

45. Sun XK, Long XJ, Fu HY, Li XB. Some characterizations of robust optimal solutions for uncertain fractional optimization and applications. J. Ind. Manag. Optim. 2017;13:803–24.

46. Lasserre JB. On representations of the feasible set in convex optimization. Optim. Lett. 2010;4:1–5.

47. Dutta J, Lalitha CS. Optimality conditions in convex optimization revisited. Optim. Lett. 2013;7(2):221–9.

48. Martinez-Legaz JE. Optimality conditions for pseudoconvex minimization over convex sets defined by tangentially convex constraints. Optim. Lett. 2015;9:1017–23.

49. Yamamoto S, Kuroiwa D. Constraint qualifications for KKT optimality condition in convex optimization with locally Lipschitz inequality constraints. Linear Nonlinear Anal. 2016;2(1):101–11.

50. Kabgani A, Soleimani-damaneh M, Zamani M. Optimality conditions in optimization problems with convex feasible set using convexificators. Math. Meth. Oper. Res. 2017;86:103–21.

51. Clarke FH. Optimization and nonsmooth analysis. New York: Wiley; 1983.

52. Ho Q. Necessary and sufficient KKT optimality conditions in non-convex optimization. Optim. Lett. 2017;11(1):41–6.

53. Suneja SK, Sharma S, Grover MB, Kapoor M. A different approach to cone-convex optimization. Am. J. Oper. Res. 2013;3:536–41.

54. Bertsekas DP. Convex analysis and optimization. Belmont: Athena Scientific; 2003.

55. Jahn J. Vector optimization: theory, applications and extensions. New York: Springer; 2004.

56. Rockafellar RT. Convex analysis. Princeton: Princeton University Press; 1970.

57. Hiriart-Urruty JB. Convex analysis and minimization algorithms I.: fundamentals principles of mathematical sciences. Berlin: Springer-Verlag; 1993.

58. Clarke FH. Generalized gradients and applications. Trans. Amer. Math. Soc. 1975;205:247–62.

59. Clarke FH. Generalized gradients of Lipschitz functions. Adv. in Math. 1981;40:52–67.

60. Pshenichnyi BN. Necessary conditions for an extremum. New York: Marcel Dekker Inc; 1971.

61. Lemaréchal C. An introduction to the theory of nonsmooth optimization. Optimization 1986;17(6):827–58.

62. Slater M. Lagrange multipliers revisited: a contribution to non-linear programming. Cowles Commission Discussion Paper 403. New Haven: The Cowles Foundation; 1950.

63. Stadler W. Multicriteria optimization in mechanics: a survey. Appl. Mech. Rev. 1984;37:277–86.

64. Yu PL. Multiple-criteria decision making: concepts, techniques and extensions. New York: Plenum Press; 1985.

65. Stadler W. Multicriteria optimization in engineering and in the sciences. New York: Plenum Press; 1988.

66. Ehrgott M. Multicriteria optimization. 2nd ed. Berlin: Springer; 2005.

67. Sawaragi Y, Nakayama H, Tanino T. Theory of multiobjective optimization. Orlando: Academic Press; 1985.

68. Luc DT. Theory of vector optimization. New York: Springer; 1989.

69. Ansari QH, Yao JC. Recent developments in vector optimization. Heidelberg: Springer-Verlag; 2012.

70. Pareto V. Course d'economie politique. Lausanne: F. Rouge; 1896.

71. Geoffrion AM. Proper efficiency and the theory of vector maximization. J. Math. Anal. Appl. 1968;22:618–30.

72. Jeyakumar V, Zaffaroni A. Asymptotic conditions for weak and proper optimality in infinite dimensional convex vector optimization. Numer. Func. Anal. Opt. 1996;17:323–43.

73. Chieu NH, Jeyakumar V, Li G, Mohebi H. Constraint qualifications for convex optimization without convexity of constraints: new connections and applications to best approximation. Eur. J. Oper. Res. 2018;265(1):19–25.

74. Sisarat N, Wangkeeree R. Characterizing the solution set of convex optimization problems without convexity of constraints. Optim. Lett. 2019. doi: 10.1007/s11590-019-01397-x.

75. Ivanov VI. Characterizations of pseudoconvex functions and semistrictly quasiconvex ones. J. Global Optim. 2013;57:677–93.

76. Sisarat N, Wangkeeree R, Lee GM. Some characterizations of robust solution sets for uncertain convex optimization problems with locally Lipschitz inequality constraints. J. Ind. Manag. Optim. 2018. doi: 10.3934/jimo.2018163.

77. Kuroiwa D, Lee GM. On robust convex multiobjective optimization. J. Nonlinear Convex Anal. 2014;15:1125–36.

78. Suzuki S, Kuroiwa D. Set containment characterization for quasiconvex programming. J. Glob. Optim. 2009;45:551–63.

79. Khanh PQ, Quyen HT, Yao JC. Optimality conditions under relaxed quasi-convexity assumptions using star and adjusted subdifferentials. Eur. J. Oper. Res. 2011;212:235–41.

80. Sisarat N, Wangkeeree R, Lee GM. KKT optimality conditions in nonsmooth, non-convex optimization. J. Nonlinear Convex Anal. 2018;19(8):1319–29.

81. Aggarwal S. Optimality and duality in mathematical programming involving generalized convex functions [dissertation]. Delhi: University of Delhi; 1998.

82. Weir T, Mond B, Craven BD. Weak minimization and duality. Numer. Func. Anal. Optim. 1987;9(1-2):181–92.

83. Sisarat N, Wangkeeree R. Necessary and sufficient KKT optimality conditions in non-convex multi-objective optimization problems with cone constraints. Pacific J. Optim. In press 2019.

84. Bonnans JF, Shapiro A. Perturbation analysis of optimization problems. New York: Springer-Verlag; 2000.

85. Suneja SK, Aggarwal S, Davar S. Multiobjective symmetric duality involving cones. Eur. J. Oper. Res. 2002;141(3):471–9.

86. Jeyakumar V, Oettli W, Natividad M. A solvability theorem for a class of quasiconvex mappings with applications to optimization. J. Math. Anal. Appl. 1993;179(2):537–46.

87. Benoist J, Borwein JM, Popovici N. A characterization of quasiconvex vector-valued functions. Proc. Am. Math. Soc. 2003;131(4):1109–13.

88. Sach PH. Characterization of scalar quasiconvexity and convexity of locally Lipschitz vector-valued maps. Optimization 1999;46(3):283–310.