Journal of Sciences, Islamic Republic of Iran 35(3): 267 - 277 (2024) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

Second-ordered Characterization of Generalized Convex Functions and Their Applications in Optimization Problems

M. T. Nadi[1*], J. Zafarani[2]

[1] *School of Mathematics, Institute for Research in Fundamental Sciences (IPM), P. O. Box 19395-5746, Tehran, Islamic Republic of Iran*
[2] *Department of Mathematics, Sheikhbahaee University and University of Isfahan, Isfahan, Islamic Republic of Iran*

Received: 26 January2025 / Revised: 16 April 2025 / Accepted: 20 May 2025

Abstract

This survey investigates some developments in the second-order characterization of generalized convex functions using the coderivative of subdifferential mapping. More precisely, it presents the second-order characterization for quasiconvex, pseudoconvex and invex functions. Furthermore, it gives some applications of the second-order subdifferentials in optimization problems such as constrained and unconstrained nonlinear programming.

Keywords: Second-order subdifferential; Positive semidefinite property; Regular second-order subdifferential; Second-order optimality conditions.
Mathematics Subject Classification (2010): 26B25, 49J40, 49J52, 49J53, 90C33

Introduction

Second-order subdifferentials and their application in the optimization and characterization of various kinds of convexity have attracted the attention of the literature. It is well known that the second-order differential of a twice continuously differentiable function $g: \mathbb{R}^n \to \mathbb{R}$ is convex if and only if $\nabla^2 g$ (its Hessian matrix) is positive semidefinite and g is strictly convex when $\nabla^2 g$ is positive definite everywhere.

This result is true even in normed spaces:

Theorem 1.1 (Flett, 1980) Let X be a real normed space and let $g: X \to \mathbb{R}$ be a twice Fréchet differentiable function, then g is convex if and only if $d^2 g(x)(y)^2 \geq 0$ for all $x, y \in X$.

Convex functions and their generalizations have many applications in optimization, economy, control theory and several other sciences; thus the characterization of convex functions is fundamental and useful. We know that when a C^2 function $g: \mathbb{R}^n \to \mathbb{R}$ attains its minimum at x, its Hessian is positive semidefinite and conversely, the positive definiteness of its Hessian is sufficient for g to reach its minimum at x when $(\nabla g)(x) = 0$. Indeed the strict local convexity of g guaranteed by positive definiteness of $\nabla^2 g(x)$. Some authors have studied the characterization of convex functions and their generalizations by their subdifferentials. Also, the second-order optimality conditions have received much attention in optimization theory, in recent years; see (1,2,3) for example.

Theorem 1.2 (4, Rockafellar 1970) The maximal monotonicity of Fréchet subdifferential of a lower semicontinuous function is a necessary and sufficient condition for its convexity.

Characterization of generalized convex functions by second-order subdifferentials can be more useful, especially in optimization.

[*] Corresponding Author: Tel: +989394520191; Email: mt_nadi@yahoo.com

Second-order characterization of convex functions by generalized second-order directional derivatives have studied by some authors.

The upper Dini-directional derivative of g at $x \in X$ in direction $v \in X$ is defined as an element of $\mathbb{R}$ by

$$g'_+(x; v) = \limsup_{t \downarrow 0} t^{-1}(g(x + tv) - g(x)). \quad (1)$$

The second-order upper Dini-directional derivative of g at $x \in X$ in direction $v \in X$ for which $g'_+(x; v)$ is defined by

$$g''_+(x; v) = \limsup_{t \downarrow 0} 2t^{-2}(g(x + tv) - g(x) - tg'_+(x; v)). \quad (2)$$

In the case of an infinite $g'_+(x; v)$, the derivative $g''_+(x; v)$ will not be considered.

Theorem 1.3 (5, Ginchev and Ivanov 2003) Let $g: X \to \mathbb{R}$ be u.s.c. Then g is convex on X if and only if the following Conditions (C_1) and (C_2) hold for each $x, u \in X$:

(C_1) $g'_+(x; v) + g'_+(x; -v) \geq 0$,

if the expression on the left-hand side has the sense

(C_2) $g'_+(x; v) + g'_+(x; -v) = 0$, implies that $g''_+(x; u) \geq 0$.

Example 1.1 The function $g(x) = -|x|, x \in \mathbb{R}$, satisfies the equality $g''_+(x; v) = 0$ for all $x, v \in \mathbf{R}$. It is continuous, but not convex. Obviously, $g'_+(x; v) + g'_+(x; -v) = -2$.

Example 1.2 The function $g: \mathbb{R} \to \mathbb{R}$ defined as

$$g(x) = \begin{cases} x^2, & \text{if x is rational ;} \\ 0, & \text{otherwise} \end{cases}$$

satisfies conditions (C_1) and (C_2), but g is not convex. This function is not u.s.c.

Some other authors used the second-order Fréchet (Second-order regular subdifferentials) and Mordukhovich (limiting) subdifferentials defined by the coderivative of the subdifferential mappings. See (6,7) for the following definitions and more details.

Let X be a Banach space endowed with a norm $\|.\|$, X^* its dual space, X^{**} its second dual space and $\langle .,. \rangle$ be the dual pairing between X and X^*. For a set-valued mapping $T: X \rightrightarrows Y$ between Banach spaces, we define the effective domain and the graph of T by

$$dom T = \{x \in X: T(x) \neq \emptyset\}, \quad gph T = \{(x, y) \in X \times Y: y \in T(x)\}.$$

The sequential Painlevé-Kuratowski upper limit of T at x in the topology of Y is defined by

$$\limsup_{x \to \bar{x}} T(x) = \{y \in Y: \exists \text{ sequences } x_k \to \bar{x}, y_k \to y \text{ with } y_k \in T(x_k), \ \forall k = 1,2, \dots\}.$$

Given $\varepsilon \geq 0$ and $\Omega \subseteq X$, the $\varepsilon - normals$ to Ω at $\bar{x} \in cl(\Omega)$ is defined by

$$\widehat{N}_\varepsilon(\bar{x}; \Omega) := \{x^* \in X^*: \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x - \bar{x} \rangle}{||x - \bar{x}||} \leq \varepsilon\},$$

where the symbol $x \xrightarrow{\Omega} \bar{x}$ means that $x \to \bar{x}$ with $x \in \Omega$. When $\varepsilon = 0$, the set $\widehat{N}_0(\bar{x}, \Omega) = \widehat{N}(\bar{x}, \Omega)$ is named the prenormal cone or Fréchet normal to Ω at $\bar{x}$.

The limiting or Mordukhovich normal cone to Ω at $\bar{x}$ is

$$N(\bar{x}; \Omega) := \limsup_{x \to \bar{x}, \varepsilon \downarrow 0} \widehat{N}_\varepsilon(x; \Omega),$$

where the sequential Painlevé-Kuratowski upper limit is taking in the $weak^*$ topology of X^*. When X is an Asplund Banach space and Ω is closed, we can put $\varepsilon = 0$.

Definition 1.1 (6) The Fréchet or regular coderivative of T at $(\bar{x}, \bar{y})$ is

$$\widehat{D}^*T(\bar{x}, \bar{y})(y^*) = \{x^* \in X^*: (x^*, -y^*) \in \widehat{N}((\bar{x}, \bar{y}), \text{gph } T)\} \ \forall y^* \in Y^*.$$

The limiting or Mordukhovich coderivative of T at $(\bar{x}, \bar{y})$ is

$$D^*T(\bar{x}, \bar{y})(y^*) = \{x^* \in X^*: (x^*, -y^*) \in N((\bar{x}, \bar{y}), \text{gph } T)\} \ \forall y^* \in Y^*.$$

i.e.,

$$D^*T(\bar{x}, \bar{y})(y^*) = \{x^* \in X^*: \exists \varepsilon_k \downarrow 0, (x_k, y_k) \text{ with } (x_k^*, -y_k^*) \in \widehat{N}_{\varepsilon_k}((x_k, y_k), \text{gph } T) \text{ as } k \to \infty\}.$$

Definition 1.2 (6) The mixed coderivative of T at $(\bar{x}, \bar{y})$ is

$$D_M^*T(\bar{x}, \bar{y})(y^*) = \{x^* \in X^*: \exists \varepsilon_k \downarrow 0, (x_k, y_k, y_k^*) \to (\bar{x}, \bar{y}, y^*), x_k^* \xrightarrow{w^*} x^* \text{ with } (x_k^*, -y_k^*) \in \widehat{N}_{\varepsilon_k}((x_k, y_k), \text{gph } T) \text{ as } k \to \infty\}.$$

Definition 1.3 (6) A single-valued mapping $g: X \to Y$ is said to be strictly differentiable at $\bar{x}$ if there is a linear continuous operator $\nabla g(\bar{x}): X \to Y$ such that

$$\lim_{u,x \to \bar{x}} \frac{g(x) - g(u) - \langle \nabla g(\bar{x}), u - x \rangle}{\| x - u \|} = 0.$$

When g is single-valued and strictly differentiable at $\bar{x}$ or continuously differentiable around $\bar{x}$, with the adjoint operator $\nabla g(\bar{x})^*: Y^* \to X^*$, we have

$$D^*g(\bar{x})(y^*) = \widehat{D}^*g(\bar{x})(y^*) = \{\nabla g(\bar{x})^* y^*\} \text{ for all } y^* \in Y^*.$$

Let $g: X \to \overline{\mathbb{R}} = [-\infty, +\infty]$ be an extended real-valued function. We define

$$\text{dom} g = \{x \in X: |g(x)| < \infty\} \text{ and } \text{epi}(g) = \{(x, \mu) \in (X \times \mathbb{R}): \mu \geq g(x)\}.$$

The Fréchet subdifferential or presubdifferential of g at $\bar{x} \in \text{dom } g$ is defined by

$$\hat{\partial}g(\bar{x}) = \{x^* \in X^*: (x^*, -1) \in \hat{N}((\bar{x}, g(\bar{x})), \text{epi } g)\}$$

and the basic or Mordukhovich limiting subdifferential is defined by

$$\partial g(\bar{x}) = \{x^* \in X^*: (x^*, -1) \in N((\bar{x}, f(\bar{x})), \text{epi } g)\}.$$

For $\bar{x} \notin \text{dom } f$, we put $\hat{\partial}g(\bar{x}) = \partial g(x) = \emptyset$. Also, g is said to be lower regular at $\bar{x}$ if $\hat{\partial}g(\bar{x}) = \partial g(\bar{x})$.

Definition 1.4 (6) Let $g: X \longrightarrow \bar{\mathbb{R}}$ be a function and its value at $\bar{x}$ is finite,

(i) For any $\bar{y} \in \partial g(\bar{x})$, the mapping $\partial^2 g(\bar{x}, \bar{y}): X^{**} \rightrightarrows X^*$ with the values

$$\partial^2 g(\bar{x}, \bar{y})(v) = (D^* \partial g)(\bar{x}, \bar{y})(v), (v \in X^{**}),$$

is called the limiting or Mordukhovich second-order subdifferential of g at $\bar{x}$ relative to $\bar{y}$.

(ii) For any $\bar{y} \in \hat{\partial}g(\bar{x})$, the mapping $\hat{\partial}^2 g(\bar{x}, \bar{y}): X^{**} \rightrightarrows X^*$ with the values

$$\hat{\partial}^2 g(\bar{x}, \bar{y})(v) = (\hat{D}^* \hat{\partial}g)(\bar{x}, \bar{y})(v), (v \in X^{**}),$$

is called the Fréchet second-order subdifferential of g at $\bar{x}$ relative to $\bar{y}$.

(iii) For any $\bar{y} \in \partial g(\bar{x})$, the mapping $\breve{\partial}^2 g(\bar{x}, \bar{y}): X^{**} \rightrightarrows X^*$ with the values

$$\breve{\partial}^2 g(\bar{x}, \bar{y})(v) = (\hat{D}^* \partial g)(\bar{x}, \bar{y})(v), (v \in X^{**}),$$

is called the Combined second-order subdifferential of g at $\bar{x}$ relative to $\bar{y}$.

(iv) For any $\bar{y} \in \partial g(\bar{x})$, the mapping $\partial^2_M g(\bar{x}, \bar{y}): X^{**} \rightrightarrows X^*$ with the values

$$\partial^2_M g(\bar{x}, \bar{y})(v) = (D^*_M \partial g)(\bar{x}, \bar{y})(v), (v \in X^{**}),$$

is called the mixed second-order subdifferential of g at $\bar{x}$ relative to $\bar{y}$. When the function g is C^2 around $\bar{x}$ and $v \in X^{**}$, we have

$$\hat{\partial}^2 g(\bar{x})(v) = \partial^2 g(\bar{x})(v) = \partial^2_M g(\bar{x})(v) = \breve{\partial}^2 g(\bar{x})(v) = \{(\nabla^2 g(\bar{x}))^* v\},$$

where $(\nabla^2 g(\bar{x}))^*$ is the adjoint operator of the Hessian $\nabla^2 g(\bar{x})$.

Definition 1.5 (PSD) holds for $g: X \longrightarrow \bar{\mathbb{R}}$, in the Fréchet sense, when $\langle z, v\rangle \geq 0$ for every $v \in X^{**}$ and $z \in \hat{\partial}^2 g(x, y)(v)$ with $(x, y) \in \text{gph } \hat{\partial}g$.

When $\langle z, v\rangle > 0$ whenever $v \neq 0$, (PD) holds in the Fréchet sense for g.

Also, (PSD) holds in the limiting sense, when $\langle z, v\rangle \geq 0$ for every $v \in X^{**}$ and $z \in \partial^2 g(x, y)(v)$ with $(x, y) \in \text{gph } \partial g$.

When $\langle z, v\rangle > 0$ whenever $v \neq 0$, (PD) holds in the limiting sense for g.

Chieu and Huy considered these cases and extended those results for the class of C^1 functions $g: X \to \mathbb{R}$, where X is a Hilbert space or an Asplund space.

Theorem 1.4 (8, Chieu, Huy 2011) Let $g: X \to R$ be a C^1 function and X be an Asplund space. Then g is convex if the following condition holds:

$$\langle z, v\rangle \geq 0 \text{ for all } v \in X^{**}, z \in \hat{\partial}^2 g(x, y)(v) \text{with } (x, y) \in \text{gph } \hat{\partial}g.$$

Results

Definition 2.1 A proper subdifferentials and their application in the optimization and characterization of various kinds of convexity have attracted the attention of the literature. It is well known that the second-order differential of a twice continuously differentiable function $g: \mathbb{R}^{\mathbf{n}} \to \mathbb{R}$ is convex if and only if $\nabla^2 g$ (its Hessian matrix) is positive semidefinite and g is strictly convex when $\nabla^2 g$ is positive definite everywhere.

This result is true even in normed spaces:

Theorem 1.1 (Flett, 1980) Let X be a real normed space and let $g: X \to \mathbb{R}$ be a twice Fréchet differentiable function, then g is convex if and only if $d^2 g(x)(y)^2 \geq 0$ for all $x, y \in X$.

Convex functions and their generalizations have many applications in optimization, economy, control theory and several other sciences; thus the characterization of convex functions is fundamental and useful. We know that when a C^2 function $g: \mathbb{R}^{\mathbf{n}} \to \mathbb{R}$ attains its minimum at x, its Hessian is positive semidefinite and conversely, the positive definiteness of its Hessian is sufficient for g to reach its minimum at x when $(\nabla g)(x) = 0$. Indeed the strict local convexity of g guaranteed by positive definiteness of $\nabla^2 g(x)$. Some authors have studied the characterization of convex functions and their generalizations by their subdifferentials. Also, the second-order optimality conditions have received much attention in optimization theory, in recent years; see (1,2,3) for example.

Theorem 1.2 (4, Rockafellar 1970) The maximal monotonicity of Fréchet subdifferential of a lower semicontinuous function is a necessary and sufficient condition for its convexity.

Characterization of generalized convex functions by second-order subdifferentials can be more useful, especially in optimization.

Second-order characterization of convex functions by generalized second-order directional derivatives have studied by some authors.

The upper Dini-directional derivative of g at $x \in X$ in direction $v \in X$ is defined as an element of $\mathbb{R}$ by

$$g'_+(x; v) = \limsup_{t \downarrow 0} t^{-1}(g(x + tv) - g(x)). \quad (1)$$

The second-order upper Dini-directional derivative of g at $x \in X$ in direction $v \in X$ for which $g'_+(x; v)$ is defined by

$$g''_+(x; v) = \limsup_{t \downarrow 0} 2t^{-2}(g(x + tv) - g(x) - tg'_+(x; v)). \quad (2)$$

In the case of an infinite $g'_+(x; v)$, the derivative

$g''_+(x;v)$ will not be considered.

Theorem 1.3 (5, Ginchev and Ivanov 2003) Let $g: X \to \mathbb{R}$ be u.s.c. Then g is convex on X if and only if the following Conditions (C_1) and (C_2) hold for each $x, u \in X$:

(C_1) $g'_+(x;v) + g'_+(x;-v) \geq 0$,

if the expression on the left-hand side has the sense

(C_2) $g'_+(x;v) + g'_+(x;-v) = 0$, implies that $g''_+(x;u) \geq 0$.

Example 1.1 The function $g(x) = -|x|, x \in \mathbb{R}$, satisfies the equality $g''_+(x;v) = 0$ for all $x, v \in \mathbf{R}$. It is continuous, but not convex. Obviously, $g'_+(x;v) + g'_+(x;-v) = -2$.

Example 1.2 The function $g: \mathbb{R} \to \mathbb{R}$ defined as

$$g(x) = \begin{cases} x^2, & \text{if x is rational;} \\ 0, & \text{otherwise} \end{cases}$$

satisfies conditions (C_1) and (C_2), but g is not convex. This function is not u.s.c.

Some other authors used the second-order Fréchet (Second-order regular subdifferentials) and Mordukhovich (limiting) subdifferentials defined by the coderivative of the subdifferential mappings. See (6,7) for the following definitions and more details.

Let X be a Banach space endowed with a norm $\|.\|$, X^* its dual space, X^{**} its second dual space and $\langle .,. \rangle$ be the dual pairing between X and X^*. For a set-valued mapping $T: X \rightrightarrows Y$ between Banach spaces, we define the effective domain and the graph of T by

$$\text{dom}T = \{x \in X: T(x) \neq \emptyset\}, \quad \text{gph}T = \{(x,y) \in X \times Y: y \in T(x)\}.$$

The sequential Painlevé-Kuratowski upper limit of T at x in the topology of Y is defined by

$$\text{limsup}_{x \to \bar{x}} T(x) = \{y \in Y: \exists \text{ sequences } x_k \to \bar{x}, y_k \to y \text{ with } y_k \in T(x_k),\ \forall k = 1,2,\ldots\}.$$

Given $\varepsilon \geq 0$ and $\Omega \subseteq X$, the $\varepsilon - normals$ to Ω at $\bar{x} \in cl(\Omega)$ is defined by

$$\widehat{N}_\varepsilon(\bar{x};\Omega) := \{x^* \in X^*: \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x - \bar{x} \rangle}{||x - \bar{x}||} \leq \varepsilon\},$$

where the symbol $x \xrightarrow{\Omega} \bar{x}$ means that $x \to \bar{x}$ with $x \in \Omega$. When $\varepsilon = 0$, the set $\widehat{N}_0(\bar{x},\Omega) = \widehat{N}(\bar{x},\Omega)$ is named the prenormal cone or Fréchet normal to Ω at $\bar{x}$.

The limiting or Mordukhovich normal cone to Ω at $\bar{x}$ is

$$N(\bar{x};\Omega) := \limsup_{x \to \bar{x}, \varepsilon \downarrow 0} \widehat{N}_\varepsilon(x;\Omega),$$

where the sequential Painlevé-Kuratowski upper limit is taking in the $weak^*$ topology of X^*. When X is an Asplund Banach space and Ω is closed, we can put $\varepsilon = 0$.

Definition 1.1 (6) *T*he Fréchet or regular coderivative of T at $(\bar{x},\bar{y})$ is

$$\widehat{D}^*T(\bar{x},\bar{y})(y^*) = \{x^* \in X^*: (x^*, -y^*) \in \widehat{N}((\bar{x},\bar{y}), \text{gph } T)\} \ \forall y^* \in Y^*.$$

The limiting or Mordukhovich coderivative of T at $(\bar{x},\bar{y})$ is

$$D^*T(\bar{x},\bar{y})(y^*) = \{x^* \in X^*: (x^*, -y^*) \in N((\bar{x},\bar{y}), \text{gph } T)\} \ \forall y^* \in Y^*.$$

i.e.,

$$D^*T(\bar{x},\bar{y})(y^*) = \{x^* \in X^*: \exists \varepsilon_k \downarrow 0, (x_k, y_k) \to (\bar{x},\bar{y}), (x_k^*, y_k^*) \xrightarrow{w^*} (x^*, y^*) \text{ with } (x_k^*, -y_k^*) \in \widehat{N}_{\varepsilon_k}((x_k, y_k), \text{gph } T) \text{ as } k \to \infty\}.$$

Definition 1.2 (6) The mixed coderivative of T at $(\bar{x},\bar{y})$ is

$$D_M^*T(\bar{x},\bar{y})(y^*) = \{x^* \in X^*: \exists \varepsilon_k \downarrow 0, (x_k, y_k, y_k^*) \to (\bar{x},\bar{y}, y^*), x_k^* \xrightarrow{w^*} x^* \text{ with } (x_k^*, -y_k^*) \in \widehat{N}_{\varepsilon_k}((x_k, y_k), \text{gph } T) \text{ as } k \to \infty\}.$$

Definition 1.3 (6) A single-valued mapping $g: X \to Y$ is said to be strictly differentiable at $\bar{x}$ if there is a linear continuous operator $\nabla g(\bar{x}): X \to Y$ such that

$$\lim_{u,x \to \bar{x}} \frac{g(x) - g(u) - \langle \nabla g(\bar{x}), u - x \rangle}{\| x - u \|} = 0.$$

When g is single-valued and strictly differentiable at $\bar{x}$ or continuously differentiable around $\bar{x}$, with the adjoint operator $\nabla g(\bar{x})^*: Y^* \to X^*$, we have

$$D^*g(\bar{x})(y^*) = \widehat{D}^*g(\bar{x})(y^*) = \{\nabla g(\bar{x})^* y^*\} \text{ for all } y^* \in Y^*.$$

Let $g: X \to \overline{\mathbb{R}} = [-\infty, +\infty]$ be an extended real-valued function. We define

$$\text{dom}g = \{x \in X: |g(x)| < \infty\} \text{ and } \text{epi}(g) = \{(x,\mu) \in (X \times \mathbb{R}): \mu \geq g(x)\}.$$

The Fréchet subdifferential or presubdifferential of g at $\bar{x} \in \text{dom } g$ is defined by

$$\hat{\partial} g(\bar{x}) = \{x^* \in X^*: (x^*, -1) \in \widehat{N}((\bar{x}, g(\bar{x})), \text{epi } g)\}$$

and the basic or Mordukhovich limiting subdifferential is defined by

$$\partial g(\bar{x}) = \{x^* \in X^*: (x^*, -1) \in N((\bar{x}, f(\bar{x})), \text{epi } g)\}.$$

For $\bar{x} \notin \text{dom } f$, we put $\hat{\partial} g(\bar{x}) = \partial g(x) = \emptyset$. Also, g is said to be lower regular at $\bar{x}$ if $\hat{\partial} g(\bar{x}) = \partial g(\bar{x})$.

Definition 1.4 (6) Let $g: X \longrightarrow \overline{\mathbb{R}}$ be a function and its value at $\bar{x}$ is finite,

(i) For any $\bar{y} \in \partial g(\bar{x})$, the mapping $\partial^2 g(\bar{x},\bar{y}): X^{**} \rightrightarrows X^*$ with the values

$$\partial^2 g(\bar{x},\bar{y})(v) = (D^*\partial g)(\bar{x},\bar{y})(v), (v \in X^{**}),$$

is called the limiting or Mordukhovich second-order subdifferential of g at $\bar{x}$ relative to $\bar{y}$.

(ii) For any $\bar{y} \in \hat{\partial} g(\bar{x})$, the mapping $\hat{\partial}^2 g(\bar{x},\bar{y}): X^{**} \rightrightarrows X^*$ with the values

$$\hat{\partial}^2 g(\bar{x},\bar{y})(v) = (\hat{D}^*\hat{\partial} g)(\bar{x},\bar{y})(v), (v \in X^{**}),$$

is called the Fréchet second-order subdifferential of g at $\bar{x}$ relative to $\bar{y}$.

(iii) For any $\bar{y} \in \partial g(\bar{x})$, the mapping $\breve{\partial}^2 g(\bar{x},\bar{y}): X^{**} \rightrightarrows X^*$ with the values

$$\breve{\partial}^2 g(\bar{x},\bar{y})(v) = (\hat{D}^*\partial g)(\bar{x},\bar{y})(v), (v \in X^{**}),$$

is called the Combined second-order subdifferential of g at $\bar{x}$ relative to $\bar{y}$.

(iv) For any $\bar{y} \in \partial g(\bar{x})$, the mapping $\partial_M^2 g(\bar{x},\bar{y}): X^{**} \rightrightarrows X^*$ with the values

$$\partial_M^2 g(\bar{x},\bar{y})(v) = (D_M^*\partial g)(\bar{x},\bar{y})(v), (v \in X^{**}),$$

is called the mixed second-order subdifferential of g at $\bar{x}$ relative to $\bar{y}$. When the function g is C^2 around $\bar{x}$ and $v \in X^{**}$, we have

$$\hat{\partial}^2 g(\bar{x})(v) = \partial^2 g(\bar{x})(v) = \partial_M^2 g(\bar{x})(v) = \breve{\partial}^2 g(\bar{x})(v) = \{(\nabla^2 g(\bar{x}))^* v\},$$

where $(\nabla^2 g(\bar{x}))^*$ is the adjoint operator of the Hessian $\nabla^2 g(\bar{x})$.

Definition 1.5 (PSD) holds for $g: X \longrightarrow \overline{\mathbb{R}}$, in the Fréchet sense, when $\langle z, v\rangle \geq 0$ for every $v \in X^{**}$ and $z \in \hat{\partial}^2 g(x,y)(v)$ with $(x,y) \in \text{gph } \hat{\partial} g$.

When $\langle z, v\rangle > 0$ whenever $v \neq 0$, (PD) holds in the Fréchet sense for g.

Also, (PSD) holds in the limiting sense, when $\langle z, v\rangle \geq 0$ for every $v \in X^{**}$ and $z \in \partial^2 g(x,y)(v)$ with $(x,y) \in \text{gph } \partial g$.

When $\langle z, v\rangle > 0$ whenever $v \neq 0$, (PD) holds in the limiting sense for g.

Chieu and Huy considered these cases and extended those results for the class of C^1 functions $g: X \to \mathbb{R}$, where X is a Hilbert space or an Asplund space.

Theorem 1.4 (8, Chieu, Huy 2011) Let $g: X \to R$ be a C^1 function and X be an Asplund space. Then g is convex if the following condition holds:

$$\langle z, v\rangle \geq 0 \text{ for all } v \in X^{**}, z \in \hat{\partial}^2 g(x,y)(v) \text{with } (x,y) \in \text{gph } \hat{\partial} g.$$

Convex case

The following questions were raised (8, Chieu, Huy 2011):

1. Is it true that, for any Fréchet differentiable function $g: X \longrightarrow \overline{\mathbb{R}}$, PSD implies convexity?
2. Which class of locally Lipschitz functions does PSD, imply the convexity of the corresponding function?
3. How to extend the characterizations to a general Banach setting?

We proved in (11), that (PSD) holds for any function $g: X \to \overline{\mathbb{R}}$, defined on an arbitrary Banach space, where g is a lower semicontinuous strongly convex.

Theorem 2 .1 (11, Nadi, Yao, Zafarani) Let X be a Banach space and $g: X \longrightarrow \overline{\mathbb{R}}$ be a lower semicontinuous strongly convex function. Then (PSD) holds.

The foregoing result also holds when we replace second-order Fréchet coderivative with mixed second-order coderivative:

Corollary 2.1 (11, Nadi, Yao, Zafarani) Let X be a Banach space and $g: X \to \overline{\mathbb{R}}$ be a lower semicontinuous strongly convex function. Then (PSD) holds in the mixed second-order sense, that is

$\langle z, v\rangle \geq 0$ for any $v \in X^{**}$ and $z \in D_M^* \partial g(\bar{x},\bar{y})(v) = \partial_M^2 g(\bar{x},\bar{y})(v)$.

Also, (PSD) guarantees the convexity of $g: X \to \mathbb{R}$ for some classes of functions. For example, (PSD) guarantees convexity for the class of continuously differentiable functions (C^1 functions) defined on Asplund spaces. Theorem 2.1 of (8, Chieu, Huy, 2011) and (PSD) imply convexity of lower-C^2 functions on $\mathbb{R}^n$ (12, Theorem 4.1). In the following, we illustrate that (PSD) is not a sufficient condition for convexity, when the function is differentiable at a point.

Example 2.1 (11, Nadi, Yao, Zafarani) Consider the function $g: \mathbb{R} \longrightarrow \mathbb{R}$ as follows:

$$g(x) = \begin{cases} \frac{1}{n^2}, & x \in]0,1], \frac{1}{n+1} < x \leq \frac{1}{n}, n \in \mathbb{N} \\ 0, & x \leq 0 \\ 2, & x > 1, \end{cases}$$

It is clear that g is differentiable at zero, but is not convex. Also, by an easy calculation, we can show that (PSD) holds for g.

In the following theorem, we showed that (PD) guarantees the convexity of $g: X \to \mathbb{R}$ when g is differentiable on X and $\hat{\partial} g$ is non-empty on X.

We proved it for $X = \mathbb{R}$ and afterwards for Banach spaces.

Theorem 2.2 (11, Nadi, Yao, Zafarani) Let $g: \mathbb{R} \to \mathbb{R}$ be a differentiable function and (PSD) holds in the Fréchet sense and $\hat{\partial} g'$ be nonempty on $\mathbb{R}$. Then g is convex.

We concluded the following corollary for g on Banach spaces by using the above argument. For arbitrary $a, v \in X$, $g: X \to \mathbb{R}$ and $s \in \mathbb{R}$, define $g_{a,v}(s) = g(a + sv)$. We know that g is convex on X if and only if $g_{a,v}$ is convex on $\mathbb{R}$ for any $a, v \in X$; See,

(13) for more details.

Corollary 2.2 Let $g: X \to \mathbb{R}$ be a differentiable function on X, $(\widehat{D}^*\nabla g)(x)(v)$ be non empty for any $x, v \in X$ and $\langle z, v\rangle \geq 0$ for every $x, v \in X$ and $z \in (\widehat{D}^*\nabla g)(x)(v)$. Then g is convex.

Corollary 2.3 Let $g: X \to \mathbb{R}$ be a differentiable function on X, $(\widehat{D}^*\nabla g)(x)(v)$ be non empty for any $x, v \in X$ and $\langle z, v\rangle \geq 0$ for every $x, v \in X$ and $z \in (D^*\nabla g)(x)(v)$ or $z \in (D_M^*\nabla g)(x)(v)$. Then g is convex.

We concluded that (PSD) and differentiability, imply the continuity of differential mapping.

Corollary 2.4 Let $g: X \to \mathbb{R}$ be differentiable on X and $\widehat{D}^*(\nabla g)(x)(v)$ be non-empty for any $x, v \in X$. If (PSD) holds in the Fréchet sense, then g is of class C^1.

Theorem 2.3 (14, Nadi, Zafarani) Let $g: X \to \mathbb{R}$ be a locally Lipschitz approximately convex function and X be an Asplund Banach space. Then g is convex, if (PSD) holds in the regular sense:

$$\langle z, v\rangle \geq 0, \forall v \in X \text{ and } z \in \widehat{\partial}^2 g(x,y)(v) \text{ with } (x,y) \in \text{gph } \widehat{\partial} g$$

Theorem 2.4 (15, Nadi, Zafarani) Let $g: X \to \mathbb{R}$ be a lower semicontinuous approximately convex function, X be an Asplund space and (PSD) holds. Then g is convex.

For $X = \mathbb{R}^n$, two classes of lower-C^1 functions and lower semicontinuous approximately convex functions are the same (16, Daniilidis, Georgiev, 2004). The class of lower- C^1 functions was initially introduced by Spingarn (1981) and afterwards, the smaller class of lower- C^k functions was introduced in 1982 by Rockafellar. The function $g: \mathbb{R}^n \to \mathbb{R}$ is said to be lower-C^k for ($k \in \mathbb{N}$) if, for each $\bar{x} \in \mathbb{R}^n$, there exists a neighbourhood of $\bar{x}$ as V such that g has the representation

$$g(x) = \max_{s \in S} g_s(x),$$

where the index set S is compact, the functions g_s are of class C^k on V, and $g_s(x)$ and all of the partial derivatives of the functions g_s of order k are jointly continuous on (s, x).

Definition 2.4 We say that a locally Lipschitz function $g: X \to \mathbb{R}$ is directionally Clarke regular (d-regular) at z if, for every $v \in X$, the Clarke directional derivative of g at z in the direction v coincides with $d^- g(z, v)$, where

$$d^- g(z, v) := \liminf_{t \to 0^+} \frac{g(z + tv) - g(z)}{t}.$$

Remark 2.1 *The above Theorem is the lower-C^1 version of Theorem 4.1 (12, Chieu, Lee, Mordukhovich, Nghia, 2016). We know that in finite dimensional spaces, a lower-C^1 function g is approximately convex and locally Lipschitz* (16, Daniilidis, Georgiev, 2004)*. Also, we answer question 2 posed in* (8, Chieu, Huy, 2011) *by this result. By a similar proof, we concluded that* (PSD) *holds for d-regular and semismooth functions defined on $X = \mathbb{R}^n$.*

We show by the following example that in the foregoing theorem, approximate convexity is essential. It means that, the class which was asked in question 2 of (8, Chieu, Huy, 2011) is approximately convex functions (the class of lower-C^1 functions when the space is finite-dimensional). We show that the following function which is Lipschitz and was given in (8, Chieu, Huy, 2011), theorem 4.2, is not approximately convex (lower-C^1).

Example 2.2 (15, Nadi, Zafarani) For all $x \in \mathbb{R}$; define

$g(x) = \int_0^x \chi_E(t)dt$, where E is a subset of $\mathbb{R}$ which is measurable and the intersection of both E and its complement with each nonempty open interval of $\mathbb{R}$ has positive Lebesgue measure. The function g is Lipschitz, and (PSD) holds but it is not convex.

Corollary 2.5 (15, Nadi, Zafarani) Let $g: X \to \mathbb{R}$ be a lower semicontinuous approximately convex function and X be a Hilbert space. Then the function g is strongly convex (with modulus $\kappa > 0$) if and only if

$$\langle z, v\rangle \geq \kappa \| v \|^2, \forall v \in X \text{ and } z \in \widehat{\partial}^2 g(x,y)(v) \text{ with } (x,y) \in \text{gph } \widehat{\partial} g. \quad (3)$$

Convex mappings

We assume that the spaces X and Y are Banach spaces and X is reflexive, $K \subseteq Y$ is a closed convex and pointed cone ($K \cap -K = 0$) and K^* is the positive dual cone of K; that is $K^* = \{y^* \in Y^*: y^*(k) \geq 0, \text{for all } k \in K\}$.

Definition 2.5 Let $g: X \longrightarrow Y$ be a vector valued function. g is K-convex on X if for any $x_1, x_2 \in X$ and $\lambda \in [0,1]$,

$$g(\lambda x_1 + (1-\lambda)x_2) \leq_K \lambda g(x_1) + (1-\lambda)g(x_2).$$

Theorem 2.5 (15, Nadi, Zafarani) Let $g: X \to Y$ be a C^1 mapping. If (PSD) holds in the limiting sense, then g is K-convex.

Also, the converse holds for twice continuously differentiable case:

Theorem 2.6 (15, Nadi, Zafarani) Let Y and X be Banach spaces and $g: X \to Y$ be a C^2 mapping. Then *(PSD)* holds if and only if g is K-convex.

The following example illustrates the foregoing theorem.

Example 2.3 (15, Nadi, Zafarani) Consider $g: \mathbb{R}^2 \to \mathbb{R}^2$ with $g(z) = g(z_1, z_2) = (z_1^2 + z_2^2, z_1^2 + z_1)$ and $C = \{(z_1, z_2) \in \mathbb{R}^2 : z_1, z_2 \geq 0 \text{ and } z_2 \leq z_1\}$. Then g is a C-convex mapping, twice continuously differentiable and *(PSD)* holds because for every $z = (z_1, z_2)$ and $v = (v_1, v_2) \in \mathbb{R}^2$ we have:

$$\nabla^2 g(z)(v) = 2v_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 2v_2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 2v_1 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} v_1 + v_2 \\ v_1 \end{pmatrix}.$$

But this means that

$$\nabla^2 g(z)(v) = \begin{pmatrix} 2v_1^2 + 2v_2^2 \\ 2v_1^2 \end{pmatrix} \in C.$$

Quasi convex functions

Characterization of pseudoconvexity and quasiconvexity by their second-order subdifferentials and their applications are studied in the literature. For twice differentiable pseudoconvex and quasiconvex functions $g: C \subseteq \mathbb{R}^n \to \mathbb{R}$, where ∇g is locally Lipschitz, the second-order characterization has been extended by (13, Crouziex and Ferland, 1996).

Given a normed linear space X and a convex subset K of X, a function $g: K \to \mathbb{R}$ is called

(i) quasiconvex on K, where for every $x, y \in K$ and $t \in]0,1[$,

$$g(x + t(y - x)) \leq \max\{g(x), g(y)\},$$

or equivalently where its level sets ($Lev_\alpha g$) are convex, i.e.,

for every $\alpha \in \mathbb{R}$, $\mathrm{Lev}_\alpha g =: \{x \in K : g(x) \leq \alpha\}$ is convex,

(ii) pseudoconvex on K if for every $x, y \in K$, $x \neq y$ and $x^* \in \hat{\partial} g(x)$,

$$\langle x^*, y - x \rangle \geq 0 \Longrightarrow g(y) \geq g(x).$$

Definition 2.6 [(14), Nadi, Zafarani] Let X be a Banach space and $F: X \rightrightarrows X^*$ be a set-valued mapping and, for every $\bar{x} \in X$ and $v \in X^{**}$, define:

$$\widehat{D}_+ F(\bar{x}, v) := \sup\{\langle z, v \rangle : z \in \widehat{D}^* F(x, y)(v), x \to \bar{x}, y \to \bar{y}, y \in F(x)\}.$$

$$\widehat{D}_- F(\bar{x}, v) := \inf\{\langle z, v \rangle : z \in \widehat{D}^* F(x, y)(v), x \to \bar{x}, y \to \bar{y}, y \in F(x)\}.$$

Here we mention a result for the quasiconvex case:

Theorem 2.7 (14, Nadi, Zafarani) Let $g: X \to R$ be a locally Lipschitz function. If the following assertions hold for every $\bar{x}, u \in X$:

(i) $\varphi_u(\bar{x}) = \inf\{\langle y, v \rangle : y \in \partial_c g(\bar{x})\} = 0$ implies that $\widehat{D}_+ \partial_c g(\bar{x}, v) \geqslant 0$;

(ii) $\varphi_u(\bar{x}) = 0$, $\widehat{D}_+ \partial_c g(\bar{x}, v) \geq 0, \widehat{D}_- \partial_c g(\bar{x}, v) \leq 0$ and $\langle y_{\bar{t}}, v \rangle > 0$ (for some $\bar{t} < 0$ and $y_{\bar{t}} \in \partial_c g(\bar{x} + \bar{t}v)$), implies that there exists $\hat{t} > 0$ such that $\langle y_t, v \rangle \geq 0$ for every $t \in [0, \hat{t}]$ and $y_t \in \partial_c g(\bar{x} + tv)$.

(iii) g is approximately convex.

Then g is quasiconvex.

Example 2.4 (14, Nadi, Zafarani) Consider the function $g: S = \{z : \| z \| < \frac{1}{2}\} \subseteq \mathbb{R}^2 \to \mathbb{R}$ defined as

$$g(z_1, z_2) = f(z) = -\| z \|^2 + \| z \|.$$

It is easy to see that g is continuously differentiable on $S \backslash \{(0,0)\}$. Also, the Clarke subdifferential at (0,0) is

For every $0 \neq v \in \mathbb{R}^2$, we have $\inf\{\langle y, v \rangle : y \in \partial_c g((0,0))\} < 0$, because the closed unit ball is a balanced subset of $\mathbb{R}^2$. Therefore, clearly (i) holds.

For (ii), assume that $v \neq (0,0)$ is arbitrary. Now, an easy calculation shows that

$$\langle \nabla g(tv), v \rangle = (v_1^2 + v_2^2)(-2t + \frac{1}{\sqrt{v_1^2 + v_2^2}}) \geq 0,$$

for every $t \in [0, \hat{t}]$ with $\hat{t} := 2(v_1^2 + v_2^2)^{-\frac{1}{2}}$, which means that (ii) holds.

Pseudo convex functions

A similar result holds for the pseudoconvex case:

Theorem 2.8 (14, Nadi, Zafarani) Let $g: X \to \mathbb{R}$ be a locally Lipschitz function. Suppose that the following conditions hold for every $\bar{x}, v \in X$:

(i) $\varphi_v(\bar{x}) = \inf\{\langle y, v \rangle : y \in \partial_c g(\bar{x})\} = 0$ implies that $\widehat{D}_+ \partial_c g(\bar{x}, v) \geqslant 0$;

(ii) $\varphi_v(\bar{x}) = 0$, $\widehat{D}_+ \partial_c g(\bar{x}, v) \geq 0$ and $\widehat{D}_- \partial_c g(\bar{x}, v) \leq 0$, implies that: there exists $\hat{t} > 0$ such that $\langle y_t, u \rangle \geq 0$ for every $t \in [0, \hat{t}]$ and $y_t \in \partial_c f(\bar{x} + tu)$.

(iii) g is approximately convex.

Then g is pseudoconvex.

For the case of strictly pseudoconvex functions, the following result is interesting:

Theorem 2.9 (17, Khanh Phat 2020) Let $g: \mathbb{R}^n \to \mathbb{R}$ be a $C^{1,1}$-smooth function satisfying

$$x \in \mathbb{R}^n, v \in \mathbb{R}^n \backslash \{0\}, \langle \nabla g(x), v \rangle = 0, \Rightarrow \langle z, v \rangle > 0, \text{for all } z \in \partial^2 g(x)(v).$$

Then g is a strictly pseudoconvex function.

Also, for the case of strictly quasiconvex functions, the following result is interesting:

Theorem 2.10 (17, Khanh Phat 2020) Let $g: \mathbb{R}^n \to \mathbb{R}$ be a $C^{1,1}$-smooth function satisfying

$$x \in \mathbb{R}^n, v \in \mathbb{R}^n \backslash \{0\}, \langle \nabla g(x), v \rangle = 0, \Rightarrow \langle z, v \rangle > 0, \text{for all } z \in \hat{\partial}^2 g(x)(v) \cup -\hat{\partial}^2 g(x)(-v).$$

Then g is a strictly quasicoconvex function.

Invex function

In recent years, the mathematical landscape has witnessed numerous extensions and generalizations of classical convexity, particularly through the invex functions by Hanson in 1981(18). This pivotal advancement sparked a wave that has substantially enriched the applications of invexity within nonlinear optimization and related fields. Notably, Hanson demonstrated that the Kuhn-Tucker conditions, which are fundamental in optimization theory, serve as sufficient criteria for optimality when dealing with invex functions. This revelation has prompted further exploration into the properties and applications of generalized convexity.

Definition 2.7 A set C is said to be invex with respect to $\eta: X \times X \to X$, when for any $x, y \in C$ and $0 \leq \lambda \leq 1$,

$$y + \lambda\eta(x, y) \in C.$$

Definition 2.8 A vector valued function $\eta: X \times X \to X$ is said to be skew, if

$$\eta(x, y) + \eta(y, x) = 0, \text{ for any } x, y \in X.$$

The following assumptions are frequently used in the literature:

ASSUMPTION **A**: Let C be an invex set with respect to η, and $g: C \to \mathbb{R}$. Then

$$g(y + \eta(x, y)) \leq g(x) \text{ for any } x, y \in C.$$

ASSUMPTION **C**: Let $\eta: X \times X \to X$. Then, for any $x, y \in X$ and for any $\delta \in [0,1]$,

$$\eta(y, y + \delta\eta(x, y)) = -\delta\eta(x, y),$$
$$\eta(x, y + \delta\eta(x, y)) = (1 - \delta)\eta(x, y)$$

Definition 2.9 A differentiable function $g: X \to \mathbb{R}$ is said to be invex with respect to η, if for any $x, y \in C$, one has

$$\langle \nabla g(y), \eta(x, y)\rangle \leq g(x) - g(y).$$

Definition 2.10 A locally Lipschitz function $g: C \subseteq X \to \mathbb{R}$ is called invex with respect to η, if for any $x, y \in C$ and any $\xi \in \partial g(x)$, one has

$$\langle \xi, \eta(x, y)\rangle \leq g(x) - g(y).$$

Remark 2.2 Note that, in the above definitions by letting $\eta(x, y) = x - y$, we reduce to the convex case. Indeed, invex functions reduce to convex functions, and invex sets, to convex sets.

Proposition 2.1 (19, Nadi, Zafarani) Let $g: \mathbb{R}^n \to \mathbb{R}$ be an invex function with respect to a skew $\eta: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, be twice differentiable at $x \in \mathbb{R}^n$ and $\eta(., x)$ be differentiable at x. Then $\langle \eta_x(x, x)v, D^2 g(x)v\rangle \geq 0$ for any $v \in \mathbb{R}^n$.

Theorem 2.11 (19, Nadi, Zafarani) Suppose that $g: \mathbb{R}^n \to \mathbb{R}$ is $C^{1,1}$, invex function with respect to a skew $\eta: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, where η is differentiable in the first argument at x and continuous. Then $\langle \eta_x(x, x)v, x^* v\rangle \geq 0$, for any $v \in \mathbb{R}^n$ and $x^* \in \partial g'(x)$.

Remark 2.3 The above results are the natural extensions of the convex case. In fact, by replacing $\eta(x, y)$ with $x - y$, we have the classical form of Hessian.

In the following example, we show that sometimes characterizing the invexity of a function by the second-order condition is easier than using the first order condition.

Example 2.4 (19, Nadi, Zafarani) Consider the following $C^{1,1}$ function $g: \mathbb{R} \to \mathbb{R}$,

$$g(x) = \begin{cases} -x^2 + x, & x \leq 0 \\ x^2 + x, & x > 0. \end{cases}$$

Consider, also $\eta(x, y) = x^3 - y^3$. An easy calculation implies that

$$\partial g'(x) = \begin{cases} -2, & x < 0 \\ \{[-2,2]\}, & x = 0 \\ 2, & x > 0, \end{cases}$$

which means that $\langle \eta_x(x, x)v, x^* v\rangle = 3x^2 x^* < 0$, by letting $x = -1$ and any arbitrary $v \in \mathbb{R}$.

Theorem 2.12 (19, Nadi, Zafarani) Let $g: \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function, g and η satisfy Assumptions A and C, $\eta(., y)$ be onto for any $y \in \mathbb{R}^n$ and skew. If $\langle \eta_x(x, x)v, \nabla^2 g(x)v\rangle \geq 0$, for any $x, v \in \mathbb{R}^n$, then g is invex with respect to η.

Optimization

Consider the nonlinear programming (NLP) as follows, with C^1 data ($f, g_i: X \to \mathbb{R}$ for $1 \leq i \leq n$ are continuously differentiable):

$$\text{minimize } f(x) \text{ subject to}$$
$$g_i(x) = 0, \text{ for } i \in E \text{ and } g_i(x) \leq 0 \text{ for } i \in I,$$

Where for the constrains, $E := \{1, \ldots, n_1\}$ and $I := \{n_1 + 1, \ldots, n_1 + n_2\}$ are finite index sets and $n := n_1 + n_2$. The point x is called a feasible point of the foregoing (NLP) problem if

$$x \in \Gamma := \{y \in X: g_i(y) = 0 \text{ for } i \in E \text{ and } g_i(y) \leq 0 \text{ for } i \in I\}.$$

Also, the classical Lagrange function is:

$$L(x, \lambda) := f(x) + \langle \lambda, g\rangle(x), \text{ for } x \in X \text{ and } \lambda \in \mathbb{R}^l.$$

When $\bar{x}$ is a solution for (NLP), the first order necessary condition is that there exist λ_i for $i = 1, \ldots n$, which are said to be the Lagrange multipliers, with

$\lambda_i g_i(\bar{x}) = 0$ (for $i = 1,\dots,n$) and $\nabla f(\bar{x}) + \sum_{i=1}^{n} \lambda_i \nabla g_i(\bar{x}) = 0$

and the standard second-order sufficient condition (SSOSC) is that there exists $k > 0$ such that

$$\nabla_x^2 L(\bar{x},\bar{\lambda})(v,v) \geq k \parallel v \parallel^2 \quad \text{with } \bar{\lambda} = (\bar{\lambda}_1,\dots,\bar{\lambda}_l) \quad (4)$$

for all $v \in X$, with $\langle \nabla g_i(\bar{x}), v\rangle = 0$ for $i \in E \cup I^+(\bar{\lambda})$ and $\langle \nabla g_i(\bar{x}), v\rangle \leq 0$ for $i \in I^0(\bar{\lambda})$, where

$$I^+(\bar{\lambda}) = \{i \in I: \bar{\lambda}_i > 0\} \text{ and } I^0(\bar{\lambda}) = \{i \in I: \lambda_i = 0\}.$$

Also, when X is finite-dimensional, we can change the inequality (4) as follows:

$$\nabla_x^2 L(\bar{x},\bar{\lambda})(v,v) > 0 \quad \text{with } \bar{\lambda} = (\bar{\lambda}_1,\dots,\bar{\lambda}_l). \quad (5)$$

Indeed, when X is finite-dimensional, the second-order sufficient condition implies optimality of $\bar{x}$ (the critical point) for Lagrange multipliers $\bar{\lambda}$, when $\nabla_x^2 L(\bar{x},\bar{\lambda})$ is positive definite on the critical cone of (NLP) at $(\bar{x},\bar{\lambda})$; it means that

$$C(\bar{x}) =: \{v: \langle \nabla g_i(\bar{x}), v\rangle = 0 \text{ for } I^+(\bar{\lambda}) \cup E \text{ and } \langle \nabla g_i(\bar{x}), v\rangle \leq 0 \text{ for } i \in I^0(\bar{\lambda})\}.$$

We continue with the following second-order sufficient condition for optimality of a KKT-point of (NLP). In the following, X is a reflexive Banach space.

Theorem 3.1 (20, Nadi, Zafarani) *(*Point-based sufficient condition) Assume the foregoing stated (NLP) problem with $\bar{z} \in \Gamma$ a KKT-point of (NLP) and Lagrange multipliers $\bar{\lambda}$. Suppose that the second-order condition holds:

$$\widehat{D}_{-}\nabla L(\bar{z},\bar{\lambda},v) > 0 \quad \text{for all } v \in C(\bar{z})\backslash\{0\}. \quad (6)$$

Then $\bar{z}$ is a strictly local minimum for (NLP).

In condition (6), we use the coderivative of the differential mapping and it is more efficient than the other similar second-order optimality conditions which have been introduced by the various kinds of generalized second-order directional derivatives. As illustrated by the following example, the following theorem due to (21, Ben-Tal and Zowe) and its constrained version can not be used for the C^1 data case.

Let $g: \mathbb{R}^n \to \mathbb{R}$ be differentiable at $\bar{x}$. We denote by $g''(\bar{x},v)$, the second-order directional derivative of g at x in direction $v \in \mathbb{R}^n$ which is defined as an element of $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$; that is

$$g''(\bar{x},v) := \lim_{t\to+\infty} \frac{2}{t^2}(g(\bar{x}+tv) - g(\bar{x}) - t\nabla g(\bar{x})v).$$

Theorem 3.2 (21, Ben-Tal and Zowe) Suppose that $g \in C^{1,1}(\mathbb{R}^n)$, $\nabla g(\bar{x}) = 0$ and $g''(\bar{x},v) > 0$ for all $v \in \mathbb{R}^n\backslash\{0\}$. Then $\bar{x}$ is a strict local minimizer of g.

Example 3.1 (20, Nadi, Zafarani) Consider the function $g: \mathbb{R}^2 \to \mathbb{R}$ defined as

$$g(z_1,z_2) := (\max(0, z_2 - 2z_1^{\frac{4}{3}}))^{\frac{3}{2}} + (\max(0, z_1^{\frac{4}{3}} - z_2))^{\frac{3}{2}}.$$

One can show that $g''(\bar{x},v) > 0$ for $\bar{x} = (0,0)$ and all nonzero direction v, but $\bar{x}$ is not a strict local minimum of g since $g(z) = 0$ for all z between the curves $z_2 = z_1^{\frac{4}{3}}$ and $z_2 = 2z_1^{\frac{4}{3}}$.

Letting (z_k) be an arbitrary sequence which converges to zero, we have $(z_k, \frac{3}{2}z_k^{\frac{4}{3}}) \to (0,0)$. It is trivial that $\nabla g(z_k, \frac{3}{2}z_k^{\frac{4}{3}}) = 0$ because g is equal to zero in a neighbourhood of $(z_k, \frac{3}{2}z_k^{\frac{4}{3}})$.

Now, it is easy to see that $0 \in \widehat{D}^*\nabla g(z_k, \frac{3}{2}z_k^{\frac{4}{3}})(v)$ for all $v \in \mathbb{R}^2$, which implies that $\widehat{D}_{-}\nabla g(\bar{x},v) \leq 0$. This means that condition (6) in the above theorem does not hold.

Pseudoconvexity of the cost function in addition to the quasiconvexity of constrained functions implies the optimality of the point that satisfies the Karush Kahn-Tucker conditions. More precisely, if the cost function or one of the active constrained functions with positive Lagrange multipliers is pseudoconvex and the rest are quasiconvex, then the Lagrange function is pseudoconvex. Booth of quasiconvexity and pseudoconvexity of constrained functions imply the convexity of the feasible set and optimality of a KKT-point will be obtained. But we know that the convexity of the feasible set is not necessary in (NLP). As mentioned below, the pseudoconvexity of the cost function and quasiconvexity of constraint functions at a KKT-point is sufficient for its optimality.

Theorem 3.3 (Mangasarian) Let the set constraint be open. The functions f and g_i for $i = 1,\dots,n_1$ are the functions defined on X and $\bar{x}$ is a feasible point. Assume that f is pseudoconvex at $\bar{x}$, f and g_i for $i \in I(\bar{x})$ are differentiable at $\bar{x}$, and g_i for $i \in I(\bar{x})$ are quasiconvex at $\bar{x}$. If there exist Lagrange nonnegative multipliers $\lambda_1,\dots,\lambda_{l_1}$ with $\lambda_i g_i(\bar{x}) = 0$ for $i = 1,\dots,n_1$ and $\nabla L(\bar{x}) = 0$ where $L = f + \sum_{i=1}^{l_1} \lambda_i g_i$, then $\bar{x}$ is a global minimizer of (NLP).

The following example shows that the pseudoconvexity at a point for the cost function in the foregoing Theorem is more than what is required.

Example 3.2 (20, Nadi, Zafarani) Consider the following (NLP) with $C^{1,1}$ data:

$$\text{minimize}\quad f(x) := -\frac{1}{2}z_1|z_1| + z_1z_2 - z_1 + z_2 \ \text{ for } z = (z_1, z_2)$$
$$\text{subject to}\quad g_1(z) := z_1^2 + z_1 - z_2 \le 0,$$
$$g_2(z) := z_1 + z_2 - 1 \le 0,\quad g_3(z) := -2z_1 + z_2 \le 0.$$

The Lagrangian function for $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ is

$$L(z,\lambda) = -\frac{1}{2}z_1|z_1| + z_1z_2 + \lambda_1(z_1^2 + z_2 - 1) + \lambda_2(z_1 + z_2 - 2) - \lambda_3 z_1.$$

Now, we can show that $\bar{z} = (0,0)$ is a KKT-point for (NLP) with Lagrange multipliers $\bar{\lambda} = (1,0,0)$. Also, for all $z \in \mathbb{R}^2$ we have

$$\nabla L(z) = (-|z_1| + z_2 + 2z_1, z_1).$$

For $z_1 > 0$ and $v = (v_1, v_2) \in \mathbb{R}^2$,

$$\nabla^2 L(z)(v) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 + v_2 \\ v_1 \end{pmatrix}$$

and for $z_1 < 0$ and $v = (v_1, v_2) \in \mathbb{R}^2$ we have

$$\nabla^2 L(z)(v) = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 + v_2 \\ v_1 \end{pmatrix}.$$

Thus, for $z_1 > 0$ and $p \in \widehat{D}^*(\nabla L)(z)(v) = \nabla^2 L(z)(v)$ we derive

$$\langle p, v\rangle = v_1^2 + 2v_2v_1.$$

Also, for $z_1 < 0$ and $p \in \widehat{D}^*(\nabla L)(z)(v) = \nabla^2 L(z)(v)$ we deduce

$$\langle p, v\rangle = 3v_1^2 + 2v_1v_2.$$

On the other hand, the set of active indexes in $\bar{z}$ is $I(\bar{z}) = \{1,3\}$ and $I^+(\bar{\lambda}) = \{1\}$ and $I^0(\bar{\lambda}) = \{3\}$. Therefore, by an easy calculation, we conclude that the critical direction cone at $\bar{x}$ is

$$C(\bar{z}) = \{v : \langle \nabla g_1(\bar{z}), v\rangle = 0 \text{ and } \langle \nabla g_3(\bar{z}), v\rangle \le 0\}$$
$$= \{(v_1, v_2) : v_1 = v_2 \text{ and } -2v_1 + v_2 \le 0\}$$
$$= \{(v_1, v_2) : v_1 = v_2 \text{ and } v_1, v_2 \ge 0\}.$$

This means that $\langle p, v\rangle > 0$ for all $p \in \widehat{D}^*(\nabla L)(z)(v)$ with $z \neq 0$ and $v \in C(\bar{z})\backslash\{0\}$. It is not difficult to see that $\widehat{D}^*(\nabla L)(z)(v) = \emptyset$ for all $z = (0, z_2)$. Therefore, the second-order sufficient condition $\widehat{D}_-(\nabla L)(z, v) > 0$ holds for all $v \in C(\bar{z})\backslash\{0\}$ by our Theorem. Moreover, it is easy to see that the cost function f is strictly pseudoconvex in direction $v = (1,1) \in C(\bar{z})$, because for all $t > 0$:

$$f(\bar{z} + tv) = f(t,t) = -\frac{1}{2}t^2 + t^2 - t + t = \frac{1}{2}t^2 > f(\bar{z}) = 0.$$

But for $u = (1,0) \notin C(\bar{z})$ and all $t > 0$:

$$f(\bar{z} + tu) = f(t, 0) = -\frac{1}{2}t^2 - t < f(\bar{z}) = 0$$

This means that f is not pseudoconvex at $\bar{z}$ in the direction u. Therefore, f is not pseudoconvex at $\bar{z}$, but $\bar{z}$ is a minimizer for (NLP).

Instead of pseudoconvexity and quasiconvexity at a point, we use the pseudoconvexity and quasiconvexity at a point in a direction and present the following extension of Mangasarian's theorem in the case of local solution.

Theorem 3.4 (20, Nadi, Zafarani) Let the set constraint be open. The functions f and g_i for $i = 1, \ldots, n_1$ are defined on X and $\bar{z}$ is a feasible point. Suppose that there exist Lagrange nonnegative multipliers $\lambda_1, \ldots, \lambda_{l_1}$ with $\lambda_i g_i(\bar{z}) = 0$ for $i = 1, \ldots, n_1$ and $\nabla L(\bar{z}) = 0$ where $L = f + \sum_{i=1}^{l_1} \lambda_i g_i$. If f and g_i for $i \in I(z)$ are differentiable at z, f is pseudoconvex at $\bar{z}$ in all critical directions $v \in C(\bar{z})$ and g_i for $i \in I(\bar{z})$ are quasiconvex at $\bar{x}$ in all critical directions $v \in C(\bar{z})$, then $\bar{z}$ is a local minimizer of (NLP).

Now, we give some applications in tilt-stability theory, as an application of our results in classical optimization.

Proposition 3.1 (14, Nadi, Zafarani) Let (PSD) hold for $g : X \to \mathbb{R}$ that is a differentiable function and $\widehat{D}^*(\nabla g)(z)(v)$ be non-empty for any $z, v \in X$. If $\nabla g(\bar{z}) = 0$, then $\bar{z}$ is a global minimizer of g.

Definition 3.1 (22, Tilt Stability, Poliquin-Rockafellar 1998) Given $g : X \to \overline{\mathbb{R}}$, a point $\bar{z} \in \operatorname{dom} f$ is a tilt-stable local minimizer of g, if there is $\gamma > 0$ such that the mapping

$$M_\gamma : z^* \longmapsto \operatorname{argmin}\{f(z) - \langle z^*, z\rangle : z \in B_\gamma(\bar{z})\}$$

is a single-valued ma*pping* and Lipschitz continuous on some vicinity of $0 \in X^*$ with $M_\gamma(0) = \bar{z}$.

Proposition 3.2 (14, Nadi, Zafarani) Let $g : X \longrightarrow \mathbb{R}$ be a strongly convex lower semicontinuous function and X be a Banach space. Then the following conditions hold:

(i) If $\bar{z}$ is a global minimizer for g, then it is the tilt-stable local minimum of g.

(ii) The point $\bar{z}$ is a local minimizer for g when X is an Asplund space. Also, there exist numbers $r \in (0, \frac{1}{\kappa})$ and $\varepsilon > 0$ such that

$$g(x) \ge g(\bar{x}) + \langle \bar{y}, z - \bar{y}\rangle - \frac{r}{2\kappa} \| z - \bar{z} \|^2 \quad \text{whenever } z \in B_\varepsilon(\bar{z}).$$

Proposition 3.3 (20, Nadi, Zafarani) Let $g : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function which satisfies Assumption **A** with respect to some η and $\nabla g(\bar{z}) = 0$. Moreover, suppose that one of the following holds:

(i) $\langle \eta_z(z,z)v, \nabla^2 g(z)(z)v\rangle \ge 0$, for any $z, v \in \mathbb{R}^n$, where η is skew and satisfies Assumption C and $\eta(., y)$ is onto for any $y \in \mathbb{R}^n$.

(ii) $\langle \eta(y,\bar{z}), \nabla^2 g(\bar{z})\eta(y,\bar{z})\rangle \ge 0$, for any $y \in \mathbb{R}^n$.

Then $\bar{z}$ is a local minimizer of g.

Consider the following constrained optimization problem:

min $g_0(z)$ subject to $g_i(z) \leq 0 \quad (i = 1, \ldots, m)$,(7)

which $g_0, g_1, \ldots, g_m$ are twice differentiable functions defined on $\mathbb{R}^n$.

Let $g(z) = (g_0(z), \ldots, g_n(z))$. We know that the existence of a vector $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ which satisfies the following conditions, (Kuhn-Tucker conditions) is necessary for $\bar{z}$ to solve this problem:

$\nabla g(\bar{z}) + \langle \lambda, \nabla g(\bar{z}) \rangle = 0$ (8)

$\langle \lambda, g(\bar{z}) \rangle = 0$ (9)

$\lambda_1, \ldots, \lambda_n \geq 0$. (10)

Hanson 1981 showed that the Kuhn-Tucker conditions are also sufficient for $\bar{z}$ to be a solution of (4), when each g_i is invex with respect to the same η. Indeed, only the invexity in a neighbourhood of $\bar{z}$ for each g_i guarantees that the foregoing conditions are sufficien (Craven 1982).

Now, we give some second-order sufficient conditions for constrained optimization problems, by using our results.

Proposition 3.4 (20, Nadi, Zafarani) Suppose we have the constrained optimization problem (4). If the Kuhn-Tucker conditions hold in $\bar{z}$, each g_i satisfies Assumption **A**, and one of the following second-order conditions holds (with respect to the same η):

(i) $\langle \eta_z(z, z)v, \nabla^2 g_i(z)v \rangle \geq 0$, for any $z, v \in \mathbb{R}^n$, where η is skew and satisfies Assumption **C** and $\eta(., y)$ is onto for any $y \in \mathbb{R}^n$,

(ii) $\langle \eta(y, \bar{z}), \nabla^2 g_i(z)\eta(y, \bar{z}) \rangle \geq 0$ for any $y \in \mathbb{R}^n$,

then $\bar{z}$ is a solution for the constrained optimization problem (4).

Acknowledgement

This research was in part supported by a grant from IPM (No.1402260047).

References

1. Khanh PD, Mordukhovich BS, Phat VT. Variational convexity of functions and variational sufficiency in optimization. SIAM Journal on Optimization. 2023;33:1121-1158.
2. Benku M, Gfrerer H, Ye JJ, Zhang J, Zhou J. Second-order optimality conditions for general nonconvex optimization problems and variational analysis of disjunctive systems. SIAM Journal on Optimization. 2023;33(4):2625-2653.
3. Gfrerer H. On Second-Order variational analysis of variational convexity of prox-regular functions. Set-Valued Variational Analysis. 2025;33(8).
4. Rockafellar RT. On the maximal monotonicity of subdifferential mappings. Pacific Journal of Mathematics. 1970;33:209-216.
5. Ginchev I, Ivanov VI. Second-order characterizations of convex and pseudoconvex functions. Journal of Applied Analysis. 2003;9:261-273.
6. Mordukhovich BS. Variational analysis and generalized differential I. II: Springer, NewYork, 2006.
7. Mordukhovich BS. Second-order variational analysis in optimization, variational Stability, and Control: Theory, Algorithms, Applications. Series in Operations Research and Financial Engineering, Springer, Cham, 2024.
8. Chieu NH, Huy NQ. Second-order subdifferentials and convexity of real-valued functions. Nonlinear Anallysis. 2011;74:154-160.
9. Ngai HV, Luc DT, Thera M. Approximate convex functions. Journal of Nonlinear and Convex Analysis. 2000;1(2):155-176.
10. Aussel D, Daniilidis A, Thibault L. Subsmooth sets: functional characterizations and related concepts. Transactions of the American Mathematical Society. 2005;357(4):1275-1301.
11. Nadi MT, Yao JC, Zafarani J. Second-order characterization of convex functions and its applications. Journal of Applied Analysis. 2019;25:49-58.
12. Chieu NH, Lee GM, Mordukhovich BS, Nghia TTA. Coderivative characterizations of maximal monotonicity for set-valued mappings. Journal of Convex Analysis. 2016;23:461-480.
13. Crouzeix JP, Ferland JA. Criteria for differentiable generalized monotone maps. Mathematical Programming. 1996;75:399-406.
14. Nadi MT, Zafarani J. Characterizations of quasiconvex and pseudoconvex functions by their second-order regular subdifferentials. Journal of the Australian Mathematical Society. 2020;109:217-229.
15. Nadi MT, Zafarani J. Second-order characterization of convex mappings in Banach spaces and its applications. Journal of Global Optimization. 2023;86:1005-1023.
16. Daniilidis A, Georgiev P. Approximate convexity and submonotonicity. Journal of Mathematical Analaysis and Applictions. 2004;291(1):292–301.
17. Khanh PD, Phat VT. Second-order characterizations of quasiconvexity and pseudoconvexity for differentiable functions with Lipschitzian derivatives. Optimization Letters. 2020(14): 2413-2427.
18. Hanson MA. On sufficiency of Kuhn-Tucker conditions. Journal of Mathematical Analaysis and Applictions. 1981;80:545–550.
19. Nadi MT, Zafarani J. Second-order characterization of invex functions and its applications in optimization problems. Indian Journal of Industrial and Applied Mathematics. 2019;10:59-70.
20. Nadi MT, Zafarani J. Second-order optimality conditions for constrained optimization problems with C^1 data via regular and limiting subdifferentials. Journal of the Optimization Theory and Applications. 2022;193:158-179.
21. Ben-Tal A, Zowe J. Directional derivatives in nonsmooth optimization. Journal of the Optimization Theory and Applications. 1985;47:483-490.
22. Poliquin RA, Rockafellar RT. Tilt stability of a local minimum. SIAM Journal of Optimization. 1998;287-299.