Applied Analysis and Optimization
Volume 8, Number 2, 2024, 127–144

Yokohama Publishers
ISSN 2189-1664 Online Journal

SECOND-ORDER OPTIMALITY CONDITIONS OF GROUP SPARSITY MULTIOBJECTIVE OPTIMIZATION PROBLEMS

HUASHENG SU, LA HUANG, XIAOQING OU, AND JIAWEI CHEN

Abstract. In this paper, we focus on optimality conditions of group sparsity multiobjective optimization problem (shortly, GSMOP) by the advanced variational analysis. We present some characterizations of outer second order tangent set, inner second order tangent set and secondary tangent cone of group sparse set. The first-order optimality conditions and second-order necessary and sufficient optimality conditions of GSMOP are established by using the tangent sets and the Dini directional derivatives.

1. Introduction

It is well know that most decision problems from scheduling, engineering theorem and machine learning have many objective functions and multiobjective optimization is very significant model in optimization field; see [13, 15, 16]. The research on multiobjective optimization problems has gone through more than a century, and it has achieved fruitful results in theory and applications such as optimality conditions, constraint qualification, stability, duality and applications on economy, finance, algorithm design and transportation; see, e.g. [4–6, 20]. For multiobjective optimization problems, it is almost impossible to find a point that optimizes all objective functions. Therefore, various solution notions such as proper efficient solution, efficient solution and weakly efficient solution are introduced in order to better characterize optimality of multiobjective optimization problems. Recently, a special class of nonconvex and discontinuous optimization problems, namely sparse optimization problems, have attracted widespread attention. The sparsity of a vector refers to the fact that the majority of its entries are zero, while group sparsity of a vector means that the zero and nonzero entries in the vector have a group structure. Compared to generalized sparse optimization, group sparse models are often more targeted for practical problems with group sparse structure and have broad applications in pattern recognition, image restoration, neuroimaging; see [2, 10].

Group sparsity optimization which is first proposed by Yuan and Lin [22], has been received a great attentions. In [9], the author improved the recently developed overlapping group shrinkage (OGS) algorithm for the denoising of group sparse signals. Lauer and Ohlsson [14] dealt with the problem of finding sparse solutions to systems of polynomial equations possibly perturbed by noise and showed how these solutions can be recovered from group sparse solutions of a derived system of linear equations. Peng and Chen [17] studied the first-order and second-order optimality conditions for the relaxation problems for group sparse optimization problems. Wu

2020 *Mathematics Subject Classification.* 90C29, 49J53.
Key words and phrases. Multiobjective optimization, group sparsity, optimality conditions, second-order tangent set, variational analysis.

and Peng [21] discussed the relationship among the four types of stationary points and provided the optimality conditions for group sparse optimization problems. Although group sparse optimization has made some progress so far, most of them are aimed at single objective problems. A lot of practical problems usually involve multiple objectives, such as signal recovery and compressed sensing. Such problems can be summarized as group sparse multiobjective optimization problems. As far as we know, there is little results on group sparse multiobjective optimization problems.

Motivated by the above works, we are interested in optimality conditions for group sparse multiobjective optimization problems (shortly, GSMOP) with non-diffierentiable objective functions. This paper is organized as follows. In Section 2, we introduce the optimization model and recall some basic notions, definitions and preliminary results used in this paper. In Section 3, we provide some characterizations of (inner) outer second order tangent set of group sparse set and its secondary tangent cone, and establish the relations between secondary tangent cone and (inner) outer second order tangent set. In addition, some characterizations of group sparse set are derived. In Section 4, we discuss the relationships among stationary points and obtain the first-order optimality conditions of GSMOP by stationary points. Section 5 is devoted to the study the second-order necessary and sufficient optimality conditions by using the second-order tangent set and Dini directional derivatives.

2. Preliminaries

In this section, we recall some notations and preliminaries. Throughout this paper, let $\mathbb{R}^m$ be the m-dimensional Euclidean space with the usual Euclidean norm $\|\cdot\|$ and the inner product $\langle\cdot,\cdot\rangle$, here $\langle x,y\rangle := \sum_{i=1}^m x_i y_i$ for any $x,y\in\mathbb{R}^m$. The nonnegative orthant of $\mathbb{R}^m$ is denoted by $\mathbb{R}^m_+$, i.e., $\mathbb{R}^m_+ := \{(x_1,\dots,x_m)\in\mathbb{R}^m : x_i \geq 0,\ i=1,\dots,m\}$. Set $\epsilon := \{1,1,\dots,1\}\in\mathbb{R}^m$ and $\Lambda^+ := \{\zeta\in\mathbb{R}^m_+ : \langle\zeta,\epsilon\rangle = 1\}$. It is easy to see that Λ^+ is a base of $\mathbb{R}^m_+$.

Now, we consider the following group sparsity multiobjective optimization problem (GSMOP):

$$\begin{aligned}&\min f(x) := (f_1(x), f_2(x), \dots, f_m(x))^\top\\ &\text{s.t. } \|x\|_{2,0} \leq k,\ x\in\mathbb{R}^n,\end{aligned}$$

where $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}, i=1,2,\dots,m$ are locally Lipschitz continuous functions, $\|x\|_{2,0} := |\{i\in\{1,2,\dots,p\} : \|x_i\|\neq 0\}|$ counts the number of non-zero groups in x, and $x\in\mathbb{R}^n$ is denoted by $x = (x_1^\top, x_2^\top, \dots, x_p^\top)^\top$ with $x_i := (x_{i,1}, x_{i,2}, \dots, x_{i,n_i})^\top \in \mathbb{R}^{n_i}, i=1,2,\dots,p$ and $\Sigma_{i=1}^p n_i = n$. The group sparse feasible set of GSMOP is denoted by $S := \{x\in\mathbb{R}^n : \|x\|_{2,0}\leq k\}$, where k is a positive integer with $k\leq p\leq n$.

Let $I\subseteq\{1,2,\dots,p\}$, $H := \{I : |I| = k\}$, $e_{ij}, i=1,2,\dots,p, j=1,2,\dots,n_i$ denote the m-dimensional vector in which the j-th entry in i-th group is one and the other are all zeros, and the subspace spanned by $\{e_{ij} : i\in I, j=1,2,\dots,n_i\}$ is denoted by $\mathbb{R}^n_I := span\{e_{ij} : i\in I, j=1,2,\dots,n_i\}$. It is clear that the set S can be written as

$$S = \bigcup_{I\in H}\mathbb{R}^n_I.$$

We denote by $\Gamma(x) := \{i \in \{1, 2, \ldots, p\} : \|x_i\| \neq 0\}$ the group sparsity support set of $x \in \mathbb{R}^n$. Furthermore, set $H(x^*) := \{I \subseteq H : \Gamma(x^*) \subseteq I\}$.

Next, we recall some basic definitions and results which will be used later.

Definition 2.1. $\hat{x} \in S$ is called a local Pareto efficient solution of GSMOP if, there exists a neighborhood $\mathbf{U}$ of $\hat{x}$ such that

$$f(x) - f(\hat{x}) \notin \mathbb{R}^m_+ \setminus \{0\},\ \forall\, x \in S \cap \mathbf{U}. \tag{2.1}$$

$\hat{x} \in S$ is also called a Pareto efficient solution of GSMOP if (2.1) holds when $\mathbf{U} = \mathbb{R}^n$.

Definition 2.2. $\hat{x} \in S$ is called a local weakly Pareto efficient solution of GSMOP if, there exists a neighborhood $\mathbf{U}$ of $\hat{x}$ such that

$$f(x) - f(\hat{x}) \notin \mathbb{R}^m_{++},\ \forall\, x \in S \cap \mathbf{U}. \tag{2.2}$$

where $\mathbb{R}^m_{++} := \{x \in \mathbb{R}^m : x_i > 0, i = 1, 2, \ldots, m\}$.

$\hat{x} \in S$ is also called a weakly Pareto efficient solution of GSMOP if (2.2) holds when $\mathbf{U} = \mathbb{R}^n$.

It is easy to see that if $\hat{x} \in S$ is a (local) Pareto efficient solution of GSMOP, then $\hat{x} \in S$ is a (local) weakly Pareto efficient solution of GSMOP.

Definition 2.3. For a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and $x \in \mathbb{R}^n$, the set

$$S_f(x^*) = \{x \in S : f(x) - f(x^*) \in -\mathbb{R}^m_+\}$$

is called the level set of f restricted in the set S.

Definition 2.4 ([8])**.** Let $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a locally Lipschitz function. The Clarke subdifferential of g at $x \in \mathbb{R}^n$ is defined as

$$\partial_c g(x) = \{\xi \in \mathbb{R}^n : \langle \xi, v\rangle \leq g^0(x; v),\ \forall\, v \in \mathbb{R}^n\},$$

where $g^0(x; v) = \limsup\limits_{y \longrightarrow x, t \searrow 0} \frac{g(y+tv)-g(y)}{t}$ is the Clarke directional derivative of g at x in the direction v.

Definition 2.5 ([7])**.** The Clarke subdifferential of $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ at $x \in \mathbb{R}^n$ is defined as

$$\partial_c f(x) = \{(\xi_1, \xi_2, \ldots, \xi_m)^\top : \xi_i \in \partial_c f_i(x),\ i = 1, 2, \ldots, m\},$$

where $\partial_c f_i(x)$ is the Clarke subdifferential of $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ at $x \in \mathbb{R}^n$.

Definition 2.6 ([1])**.** A locally Lipschitz function $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ is said to be ∂_c-pseudoconvex at $\hat{x} \in A \subseteq \mathbb{R}^n$ if, for any $x \in A$ and $x \neq \hat{x}$,

$$g(x) < g(\hat{x}) \Longrightarrow \langle \xi, x - \hat{x}\rangle < 0\ \forall\, \xi \in \partial_c g(\hat{x}).$$

Definition 2.7. A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is said to be ∂_c-pseudoconvex at $\hat{x} \in A \subseteq \mathbb{R}^n$ if, f_i, $i = 1, 2, \ldots, m$ are ∂_c-pseudoconvex at $\hat{x}$.

Remark 2.1. Clearly, it is obvious that $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is ∂_c-pseudoconvex on $\hat{x}$ if and only if $\lambda^\top f$ is ∂_c-pseudoconvex on $\hat{x}$ for all $\lambda \in \Lambda^+$.

Definition 2.8 ([19])**.** Let $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a locally Lipschitzian function and $x, u, v \in \mathbb{R}^n$. The lower Dini derivatives of g at x in the direction $u \in \mathbb{R}^n$ is defined as

$$D^- g(x; u) = \liminf_{t \searrow 0} \frac{g(x + tu) - g(x)}{t}.$$

Furthermore, if $D^- g(x; u)$ exists, the second-order Dini derivatives of g at (x, u) in the direction $v \in \mathbb{R}^n$ is defined as

$$D^2 g(x; u, v) = \liminf_{t \searrow 0} \frac{g(x + tu + t^2 v) - g(x) - tD^- g(x; u)}{t^2}.$$

Definition 2.9 ([18])**.** Let Ω be a nonempty closed subset of $\mathbb{R}^n$ and $x \in \Omega$.

(i) The Bouligand tangent cone and Clarke tangent cone of Ω at x are respectively defined as

$$T^B_\Omega(x) = \left\{ d \in \mathbb{R}^n : \exists\, x^k \xrightarrow{\Omega} x, t_k \searrow 0, \lim_{k \longrightarrow \infty} \frac{x^k - x}{t_k} = d \right\},$$

and

$$T^C_\Omega(x) = \left\{ d \in \mathbb{R}^n : \forall x^k \xrightarrow{\Omega} x, t_k \searrow 0, \exists\, y^k \longrightarrow x, \lim_{k \longrightarrow \infty} \frac{x^k - y^k}{t_k} = d \right\},$$

where $x^k \xrightarrow{\Omega} x$ means $\lim_{k \longrightarrow \infty} x^k = x$ and $x^k \in \Omega$ for each $k = 1, 2, \ldots$.

(ii) The Fréchet normal cone and Clarke normal cone of Ω at x are respectively defined as

$$N^B_\Omega(x) = \left\{ d \in \mathbb{R}^n : \langle d, z \rangle \leq 0,\ \forall\, z \in T^B_\Omega(x) \right\},$$

and

$$N^C_\Omega(x) = \left\{ d \in \mathbb{R}^n : \langle d, z \rangle \leq 0,\ \forall\, z \in T^C_\Omega(x) \right\}.$$

Obviously, $T^C_\Omega(x) \subseteq T^B_\Omega(x)$ for $x \in \Omega$. It is directly obtained from Proposition 6.5 of [18] that

$$N^B_\Omega(x) \subseteq N^C_\Omega(x). \tag{2.3}$$

Definition 2.10 ([3])**.** Let Ω be a nonempty closed subset of $\mathbb{R}^n$ and $x \in \Omega$. The outer second-order tangent set and the inner second-order tangent set of Ω at x in the direction $d \in \mathbb{R}^n$ are respectively defined as

$$T^2_\Omega(x, d) = \{ w \in \mathbb{R}^n : \exists\, t_k \searrow 0, w^k \longrightarrow w, \quad x + t_k d + \frac{1}{2} t_k^2 w^k \in \Omega \},$$

$$T^{2,i}_\Omega(x, d) = \{ w \in \mathbb{R}^n : \forall t_k \searrow 0, \exists\, w^k \longrightarrow w, \quad x + t_k d + \frac{1}{2} t_k^2 w^k \in \Omega \}.$$

Definition 2.11 ([12])**.** Let $\Omega \subseteq \mathbb{R}^n$ be a nonempty closed set. Ω is called pseudo-convex set at x with respect to $T_\Omega(x)$ if

$$\Omega \subseteq x + T_\Omega(x),$$

where $T_\Omega(x)$ is a tangent cone to Ω at $x \in \Omega$.

Definition 2.12. Let $\Omega \subseteq \mathbb{R}^n$ be a nonempty closed cone. The dual cone of Ω is defined by

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

The following results show the quality of Clarke subdifferentials of locally Lipschitz function.

Lemma 2.1 ([8])**.** *If $g_i : \mathbb{R}^n \longrightarrow \mathbb{R}(i = 1, \ldots, m)$ are locally Lipschitz functions, then the following relations hold:*

(i) $\partial_c(g_1 + \cdots + g_m)(x) \subseteq \partial_c g_1(x) + \cdots + \partial_c g_m(x)$;

(ii) $\partial_c(tg)(x) = t\partial_c g(x)$ *for all* $t > 0$.

The following lemma gives the expression of the tangent cones and normal cones of the group sparse feasible set S.

Lemma 2.2 ([21])**.** *Let $x^* \in S$. The Bouligand tangent cone, the Clarke tangent cone and their normal cones of S at x^* are respectively given as follows:*

(i) *Bouligand tangent cone:*

$$\begin{aligned} T_S^B(x^*) &= \{d \in \mathbb{R}^n : \|d\|_{2,0} \leq k, \|x^* + \mu d\|_{2,0} \leq k,\ \forall\, \mu \in \mathbb{R}\} \\ &= \bigcup_{I \in H(x^*)} \{d \in \mathbb{R}^n : d_i = 0, i \notin I\} \\ &= \bigcup_{I \in H(x^*)} span\{e_{ij}, i \in I, j = 1, 2, \ldots, n_i\}, \end{aligned}$$

where $d_i \in \mathbb{R}^{n_i}$ is the i-th group of $d \in \mathbb{R}^n$;

(ii) *Clarke tangent cone:*

$$\begin{aligned} T_S^C(x^*) &= \{d \in \mathbb{R}^n : d_i = 0,\ i \notin \Gamma(x^*)\} \\ &= \{d \in \mathbb{R}^n : \Gamma(d) \subseteq \Gamma(x^*)\} \\ &= span\{e_{ij}, i \in \Gamma(x^*), j = 1, 2, \ldots, n_i\}; \end{aligned}$$

(iii) *Fréchet normal cone:*

$$N_S^B(x^*) = \begin{cases} span\{e_{ij},\ i \notin \Gamma(x^*),\ j = 1, 2, \ldots, n_i\},\ \ if\ \ \|x\|_{2,0} = k, \\ \{0\},\ \ if\ \ \|x\|_{2,0} < k; \end{cases}$$

(iv) *Clarke normal cone:*

$$N_S^C(x^*) = \{d \in \mathbb{R}^n : d_i = 0,\ i \in \Gamma(x^*)\}.$$

3. Variational analysis of the group sparsity set S

In this section, we present the characterizations of (inner) outer second order tangent set of the group sparse set S and its secondary tangent cone, and the relations between secondary tangent cone and (inner) outer second order tangent set. In addition, we also give some characterizations of the group sparse set S.

The following results present a characterization of the outer second-order tangent set of the group sparse set S via the Bouligand tangent cone.

Proposition 3.1. *Let $x^* \in S$ and $d \in T^B S(x^*)$. Then*

$$T_S^2(x^*, d) = T_S^B(x^*) \cap T_S^B(d).$$

Proof. Obviously, the following equality holds:

$$T_S^2(x^*, d) = T^2_{\bigcup_{I\in H} \mathbb{R}_I^n}(x^*, d) = \bigcup_{I\in H} T^2_{\mathbb{R}_I^n}(x^*, d), \tag{3.1}$$

where the second equality holds comes from [3, Proposition 3.37]. Take any $\mathbb{R}_I^n$ satisfying $I \notin H(x^*)$. Then, $x^* \notin \mathbb{R}_I^n$, $T^2_{\mathbb{R}_I^n}(x^*, d) = \emptyset$. Therefore, it follows from (3.1) that

$$\bigcup_{I\in H} T^2_{\mathbb{R}_I^n}(x^*, d) = \bigcup_{I\in H(x^*)} T^2_{\mathbb{R}_I^n}(x^*, d). \tag{3.2}$$

Obverse that $\mathbb{R}_I^n = T^B_{\mathbb{R}_I^n}(x^*)$. Then if $d \notin \mathbb{R}_I^n$, $T^2_{\mathbb{R}_I^n}(x^*, d) = \emptyset$. This combined with (3.2) yields that

$$\bigcup_{I\in H(x^*)} T^2_{\mathbb{R}_I^n}(x^*, d) = \bigcup_{I\in H(x^*), d\in\mathbb{R}_I^n} T^2_{\mathbb{R}_I^n}(x^*, d). \tag{3.3}$$

Therefore, the desire result is as follows

$$\begin{aligned}
T_S^2(x^*, d) &= \bigcup_{I\in H(x^*), d\in\mathbb{R}_I^n} T^2_{\mathbb{R}_I^n}(x^*, d) \\
&= \bigcup_{I\in H(x^*), d\in\mathbb{R}_I^n} (\mathbb{R}_I^n + \mathbb{R}d) \\
&= \bigcup_{I\in H(x^*), d\in\mathbb{R}_I^n} \mathbb{R}_I^n \\
&= \bigcup_{|I|=k, \Gamma(x^*)\subseteq I, \Gamma(d)\subseteq I} \mathbb{R}_I^n \\
&= (\bigcup_{|I|=k, \Gamma(x^*)\subseteq I} \mathbb{R}_I^n) \cap (\bigcup_{|I|=k, \Gamma(d)\subseteq I} \mathbb{R}_I^n) \\
&= T_S^B(x^*) \cap T_S^B(d),
\end{aligned}$$

where the first equality holds due to (3.1)and (3.3), the second equality comes from [18, Proposition 13.12]. $\square$

Remark 3.1. The outer second-order tangent set of the group sparse feasible set S at $x^* \in S$ in the direction $d \in T_S^B(x^*)$ can be also written as

$$T_S^2(x^*, d) = \{w \in \mathbb{R}^n : \|x^* + \alpha w\|_{2,0} \leq k,\ \|d + \beta w\|_{2,0} \leq k, \forall \alpha, \beta \in \mathbb{R}\}.$$

The following result show that $T_S^2(x^*, d)$ has a different form when $d \in T_S^C(x^*)$.

Proposition 3.2. *Let $x^* \in S$ and $d \in T_S^C(x^*)$. Then*

$$T_S^2(x^*, d) = \{w \in \mathbb{R}^n : \|w\|_{2,0} \leq k,\ \|x^* + \mu d + \gamma w\|_{2,0} \leq k,\ \forall\ \mu, \gamma \in \mathbb{R}\}.$$

Proof. Set $D := \{w \in \mathbb{R}^n : \|w\|_{2,0} \leq k,\ \|x^* + \mu d + \gamma w\|_{2,0} \leq k,\ \forall\ \mu, \gamma \in \mathbb{R}\}$. Take any $w \in T_S^2(x^*, d)$. From Remark 3.1, we obtain that $\|w\|_{2,0} \leq k$ and

$$\|x^* + \gamma w\|_{2,0} \leq k,\ \forall \gamma \in \mathbb{R}. \tag{3.4}$$

If $\|x^*\|_{2,0} = k$, then it follows from (3.4) that $\Gamma(w) \subseteq \Gamma(x^*)$. Since $d \in T_S^C(x^*)$, by virtue of Lemma 2.2, we have

$$\Gamma(d) \subseteq \Gamma(x^*). \tag{3.5}$$

Note that $\Gamma(w) \subseteq \Gamma(x^*)$. Then combined with (3.5), we obtain

$$\Gamma(x^* + \mu d + \gamma w) \subseteq \Gamma(x^*),\ \forall\ \mu, \gamma \in \mathbb{R},$$

which yields that

$$\|x^* + \mu d + \gamma w\|_{2,0} \leq k,\ \forall\ \mu, \gamma \in \mathbb{R}. \tag{3.6}$$

If $\|x^*\|_{2,0} < k$, without loss of generality, we can assume that for any $j \in \mathbb{R}$, $\|x^*\|_{2,0} = k - j,\ 0 < j \leq k$. Then we can divide $1, \ldots, p$ into

$$i \in \begin{cases} I_1, & \text{if } w_i \neq 0, x_i^* = 0 \\ I_2, & \text{if } w_i \neq 0, x_i^* \neq 0 \\ I_3, & \text{if } w_i = 0, x_i^* = 0 \\ I_4, & \text{if } w_i = 0, x_i^* \neq 0. \end{cases}$$

According to (3.4), one has $|I_1| = 0, 1, \ldots, j$. It is obvious that $\Gamma(w) \subseteq \Gamma(x^*)$ if $|I_1| = 0$. Thus, (3.6) always holds. when $|I_1| \neq 0$, it stems from (3.5) that the index ℓ satisfying $\Gamma(x^* + \mu d + \gamma w)_\ell \neq 0$ is up to $|\Gamma(x^*) \cup I_1|$ for all $\mu, \gamma \in \mathbb{R}$. This implies that (3.6) holds. Altogether, $T_S^2(x^*, d) \subseteq D$.

Conversely, take any $w \in D$. Let $\mu = 0$. Then (3.4) holds. Again using (3.5), one has $\|d + \beta w\|_{2,0} \leq k$ for all $\beta \in \mathbb{R}$. Therefore, we conclude that $D \subseteq T_S^2(x^*, d)$. Consequently, the conclusion holds. □

Similarly, we next give a characterization of the inner second-order tangent set of the group sparse set S.

Proposition 3.3. *Let $x^* \in S$ and $d \in T_S^B(x^*)$. Then*

$$T_S^{2,i}(x^*, d) = \{w \in \mathbb{R}^n : \|w\|_{2,0} \leq k,\ \|x^* + \mu d + \gamma w\|_{2,0} \leq k,\ \forall\, \mu, \gamma \in \mathbb{R}\}. \tag{3.7}$$

Proof. Set $D := \{w \in \mathbb{R}^n : \|w\|_{2,0} \leq k,\ \|x^* + \mu d + \gamma w\|_{2,0} \leq k,\ \forall\ \mu, \gamma \in \mathbb{R}\}$. Take arbitrary $w \in T_S^{2,i}(x^*, d)$. Then for any $t_k \searrow 0$, there exists $w^k \longrightarrow w$ such that $x^* + t_k d + \frac{1}{2} t_k^2 w^k = x^k \in S$, which yields that $x^k \xrightarrow{S} x^*$ and $w = \lim_{k \longrightarrow \infty} \frac{x^k - x^* - t_k d}{\frac{1}{2} t_k^2}$. Then we have

$$\Gamma(x^*) \subseteq \Gamma(x^k) \tag{3.8}$$

and

$$\Gamma(w) \subseteq \Gamma(x^k - x^* - t_k d). \tag{3.9}$$

We next claim that

$$\Gamma(d) \subseteq \Gamma(x^k - x^*). \tag{3.10}$$

Suppose that $\Gamma(d) \nsubseteq \Gamma(x^k - x^*)$. Then there is at least one i that satisfies $d_i > 0$ and $(x^k - x^*)_i = 0$. Note that $x^* + t_k d + \frac{1}{2}t_k^2 w^k = x^k$. Then, $t_k d + \frac{1}{2}t_k^2 w^k = x^k - x^*$. Hence, we derive that

$$(t_k d + \frac{1}{2}t_k^2 w^k)_i = (x^k - x^*)_i = 0,$$

which is a contradiction since $d_i > 0$ and $w_i^k \geq 0$. Then combined (3.8), (3.9) with (3.10), we obtain

$$\Gamma(w) \subseteq \Gamma(x^k - x^*) \subseteq \Gamma(x^k),$$

which implies that $\|w\|_{2,0} \leq k$ and $\|x^* + \mu d + \gamma w\|_{2,0} \leq k,\ \forall\ \mu, \gamma \in \mathbb{R}$. Thus, $w \in D$. So $T_S^{2,i}(x^*, d) \subseteq D$.

Conversely, take any $w \in D$. For any $t_k \searrow 0$, there exists $w^k = w$ such that $x^* + t_k d + \frac{1}{2}t_k^2 w^k = x^* + t_k d + \frac{1}{2}t_k^2 w \in S$, which means that $D \subseteq T_S^{2,i}(x^*, d)$. Altogether, the conclusion holds. □

The following proposition gives a equivalent characterization of the Bouligand tangent cone of $T_S^B(x^*)$ (i.e., secondary tangent cone) at $d \in T_S^B(x^*)$ by (3.7).

Proposition 3.4. *Let* $d \in T_S^B(x^*)$. *Then*

$$T^B_{T_S^B(x^*)}(d) = \{w \in \mathbb{R}^n : \|w\|_{2,0} \leq k, \|x^* + \mu d + \gamma w\|_{2,0} \leq k, \forall \mu, \gamma \in \mathbb{R}\}. \tag{3.11}$$

Proof. Set $D := \{w \in \mathbb{R}^n : \|w\|_{2,0} \leq k,\ \|x^* + \mu d + \gamma w\|_{2,0} \leq k,\ \forall \mu, \gamma \in \mathbb{R}\}$. According to the definition of Bouligand tangent cone, we have

$$T^B_{T_S^B(x^*)}(d) = \{w \in \mathbb{R}^n : \exists\ d^n \stackrel{T_S^B(x^*)}{\longrightarrow} d,\ t_n \searrow 0,\ \lim_{n \longrightarrow \infty} \frac{d^n - d}{t_n} = w\}.$$

For any $w \in T^B_{T_S^B(x^*)}(d)$, there exists $d^n \subseteq T_S^B(x^*)$ such that $d^n \longrightarrow d$. Then

$$\Gamma(d) \subseteq \Gamma(d^n) \tag{3.12}$$

for any sufficiently large n. Note that $\lim_{n \longrightarrow \infty} \frac{d^n - d}{t_n} = w$ and $t_n \searrow 0$. Then $\Gamma(w) \subseteq \Gamma(d^n - d)$. This together with (3.12) yields that

$$\Gamma(w) \subseteq \Gamma(d^n - d) \subseteq \Gamma(d^n) \tag{3.13}$$

for any sufficiently large n.

Since $d^n \in T_S^B(x^*)$, it follows from (3.13) that

$$\|x^* + (\mu + \gamma) d^n\|_{2,0} \leq k,\ \forall\ \mu, \gamma \in \mathbb{R}. \tag{3.14}$$

and

$$\|w\|_{2,0} = |\Gamma(w)| \leq |\Gamma(d^n)| = \|d^n\|_{2,0} \leq k \tag{3.15}$$

Then, by virtue of (3.12), (3.13) and (3.14), one has

$$\|x^* + \mu d + \gamma w\|_{2,0} \leq k,\ \forall\ \mu, \gamma \in \mathbb{R}.$$

This together with (3.15) yields that $T^B_{T_S^B(x^*)}(d) \subseteq D$.

Conversely, take arbitrary $w \in D$ and any given the sequence $\{t_n\} \searrow 0$. Let $d_n = d + t_n w$. Then we directly have $\lim_{n \longrightarrow \infty} d_n = d$ and $\lim_{n \longrightarrow \infty} \frac{d_n - d}{t_n} = w$. Besides, it follows from $d_n = d + t_n w$ and $w \in D$ that

$$\|x^* + \mu d^n\|_{2,0} = \|x^* + \mu d + \mu t_n w\|_{2,0} \leq k,\ \forall\ \mu \in \mathbb{R},$$

then $d^n \in T_S^B(x^*)$. Thus there exists $d^n \overset{T_S^B(x^*)}{\longrightarrow} d,\ t_n \searrow 0$ such that $\lim_{n\longrightarrow\infty} \frac{d_n-d}{t_n} = w$. This means that $w \in T^B_{T_S^B(x^*)}(d)$ and $D \subseteq T^B_{T_S^B(x^*)}(d)$. Altogether, the conclusion holds. □

In the following, we state the relations among the inner, outer second-order tangent set of the group sparse feasible set S at $x^* \in S$ in the direction $d \in T_S^B(x^*)$ and secondary tangent cone at $d \in T_S^B(x^*)$.

Theorem 3.1. *If $x^* \in S$ and $d \in T_S^B(x^*)$, then*

$$T_S^{2,i}(x^*, d) = T^B_{T_S^B(x^*)}(d) \subseteq T_S^2(x^*, d).$$

Proof. It follows from Proposition 3.3 and Proposition 3.4 that the equation obviously holds. By virtue of Proposition 3.4 and Remark 3.1, we obtain

$$\begin{aligned} T_S^2(x^*, d) =& \{w \in \mathbb{R}^n : \|w\|_{2,0} \le k,\ \|x^* + \alpha w\|_{2,0} \le k, \\ & \|d + \beta w\|_{2,0} \le k,\ \forall\ \alpha, \beta \in \mathbb{R}\} \end{aligned} \tag{3.16}$$

and

$$T^B_{T_S^B(x^*)}(d) = \{w \in \mathbb{R}^n : \|w\|_{2,0} \le k,\ \|x^* + \mu d + \gamma w\|_{2,0} \le k,\ \forall\ \mu, \gamma \in \mathbb{R}\}. \tag{3.17}$$

Take any $w \in T^B_{T_S^B(x^*)}(d)$. Then let $\mu = 0$ in (3.17) and we have $\|w\|_{2,0} \le k$ and $\|x^* + \alpha w\|_{2,0} \le k$, $\forall \alpha \in \mathbb{R}$. Now, let's take $\mu^* > 0$ small enough. Then when $d_i \ne 0$, we derive that $x_i^* + \mu^* d_i \ne 0$. This together with the definition of the group sparsity support set implies that

$$\Gamma(d) \subseteq \Gamma(x^* + \mu^* d). \tag{3.18}$$

Note that $w \in T^B_{T_S^B(x^*)}(d)$. Thus by (3.17) and (3.18), we arrive at

$$\|d + \beta w\|_{2,0} \le \|x^* + \mu^* d + \beta w\|_{2,0} \le k,\ \forall \beta \in \mathbb{R}.$$

Therefore, according to (3.16), $w \in T_S^2(x^*, d)$ and $T^B_{T_S^B(x^*)}(d) \subseteq T_S^2(x^*, d)$. Consequently, the conclusion holds. □

We now give an example to show that the anti-inclusion relationship does not hold for the above result.

Example 3.1. Consider the group sparse set

$$S := \{x = ((x_{1,1}, x_{1,2}), (x_{2,1}, x_{2,2}, x_{2,3}), x_3, x_4)^\top \in \mathbb{R}^7 : \|x\|_{2,0} \le 3\}.$$

Let $x^* = ((1, 1), (0, 0, 0), 0, 0)^\top$ and $d = ((0, 0), (1, 1, 0), 0, 0)^\top$. It is obvious that $d \in T_S^B(x^*)$. Take $w = ((0, 0), (1, 0, , 0), 1, 1)^\top$. It is easy to get that

$$w \in T^B_{T_S^B(x^*)}(d).$$

After calculation, $\|x^* + d + w\|_{2,0} = 4 > 3$. This yields that

$$w \notin T_S^2(x^*, d).$$

Therefore, we derive that $T_S^2(x^*, d) \not\subseteq T^B_{T_S^B(x^*)}(d)$.

We now present some characterizations of the group sparse set S.

Theorem 3.2. (i) *The group sparse set S is a cone and its dual cone $S^* = \{0\}$.*

(ii) *If $x^* \in S$, then $x^* + T_S^B(x^*) \subseteq S$.*
Moreover, if S is pseudoconvex set at $x^* \in S$ with respect to $T_S^B(x^*)$, then $x^* + T_S^B(x^*) = S$.

(iii) *If $x^* \in S$, then $x^* + T_S^C(x^*) \subseteq S$. Moreover, if S is pseudoconvex set at $x^* \in S$ with respect to $T_S^C(x^*)$, then $x^* + T_S^C(x^*) = S$.*

Proof. (i) It is clear that S is cone. Due to Lemma 2.2 and $|\Gamma(0)| < k$, we have

$$T_S^B(0) = \{d \in \mathbb{R}^n : \|d\|_{2,0} \le k,\ \|0 + \mu d\|_{2,0} \le k,\ \forall\ \mu \in \mathbb{R}\} = S$$

and $N_S^B(0) = \{0\}$. Furthermore, one has by Definition 2.9 and 2.12

$$S^* = T_S^B(0)^* = -N_S^B(0) = \{0\}.$$

(ii) Taking $\mu = 1$ for all $d \in T_S^B(x^*)$, then it follows from Lemma 2.2 that $|\Gamma(x^*+d| \le k$. This yields that $x^* + T_S^B(x^*) \subseteq S$. Besides, according to Definition 2.11, the equality obviously holds.

(iii) The proof is similar to (ii). The proof is completed. □

4. First-order optimality conditions

In this section, we discuss the relationships among stationary points, and study the first-order necessary and sufficient optimality conditions of GSMOP by stationary points. To do this, We firstly introduce the definitions of stationary points.

Definition 4.1. $x^* \in S$ is called:

(i) an $N^\sharp$-stationary point of GSMOP if there exists $\lambda^* \in \Lambda^+$ such that

$$0 \in \partial_c f(x^*)^\top \lambda^* + N_S^\sharp(x^*);$$

(ii) a strict $N^\sharp$-stationary point of GSMOP if there exists $\lambda^* \in \Lambda^+$ such that

$$\partial_c f(x^*)^\top \lambda^* \subseteq N_S^\sharp(x^*);$$

(iii) an $T^\sharp$-stationary point of GSMOP if there exists $\lambda^* \in \Lambda^+$ such that

$$0 \in \partial_S^\sharp f(x^*);$$

where $\sharp \in \{B, C\}$ represent the mean of Bouligand or Clarke, and

$$\partial_S^\sharp f(x^*) := \arg\min_{d \in T_S^\sharp(x^*)} \{\|d + \xi\| : \xi \in \partial_c f(x^*)^\top \lambda^*\}.$$

The following theorem presents the relationship between T^B-stationary point and N^B-stationary point of GSMOP.

Theorem 4.1. *$x^* \in S$ is an N^B-stationary point of GSMOP if and only if $x^* \in S$ is an T^B-stationary point of GSMOP.*

Proof. In order to prove the desired result holds, we split the proof into two cases.

Case 1: Let $\|x^*\|_{2,0} = k$.

" $\Leftarrow$ " Assume that $x^* \in S$ is an T^B-stationary point. Then there exists $\bar{\lambda} \in \Lambda^+$ such that $0 \in \partial_S^B f(x^*)$. On the other hand, according to Lemma 2.2, we have

$$d \in T_S^B(x^*) \Leftrightarrow \Gamma(d) \subseteq \Gamma(x^*).$$

Hence, it follows that

$$0 \in \partial_S^B f(x^*) = \arg\min_{d\in T_S^B(x^*)}\{\|d+\xi\| : \xi \in \partial_c f(x^*)^\top \bar{\lambda}\}$$
$$= \arg\min_{\Gamma(d)\subseteq\Gamma(x^*)}\{\|d+\xi\| : \xi \in \partial_c f(x^*)^\top \bar{\lambda}\}.$$

If $i \notin \Gamma(x^*)$, then by the above equality one has

$$d_i = 0,\ \forall\ d \in \partial_S^B f(x^*). \tag{4.1}$$

If $i \in \Gamma(x^*)$, then for any $\xi \in \partial_c f(x^*)^\top \bar{\lambda}$, there exists $d \in \partial_S^B f(x^*) \subseteq T_S^B(x^*)$ such that $d_i = -\xi_i$. This combined with (4.1) and $0 \in \partial_S^B f(x^*)$ implies that there exists $\bar{\xi} \in \partial_c f(x^*)^\top \bar{\lambda}$ such that

$$\bar{\xi}_i \begin{cases} = 0, & \text{if}\quad i \in \Gamma(x^*) \\ \in \mathbb{R}^{n_i}, & \text{if}\quad i \notin \Gamma(x^*). \end{cases} \tag{4.2}$$

Hence, by Lemma 2.2 and $\|x^*\|_{2,0} = k$, we obtain

$$N_S^B(x^*) = \{d \in \mathbb{R}^n : d_i = 0, i \in \Gamma(x^*)\}. \tag{4.3}$$

Then it follows from (4.2) and (4.3) that

$$0 \in \partial_c f(x^*)^\top \bar{\lambda} + N_S^B(x^*),$$

which means that $x^* \in S$ is an N^B-stationary point of GSMOP.

" $\Rightarrow$ " Assume that $x^* \in S$ is an N^B-stationary point. Then there exists $\lambda^* \in \Lambda^+$ such that

$$0 \in \partial_c f(x^*)^\top \lambda^* + N_S^B(x^*).$$

Thus there exists $\xi^* = (\xi_1^*, \xi_2^*, \ldots, \xi_p^*) \in \partial_c f(x^*)^\top \lambda^*$ such that

$$-\xi^* \in N_S^B(x^*).$$

Then according to Lemma 2.2, we have

$$\xi_i^* \begin{cases} = 0, & \text{if}\quad i \in \Gamma(x^*) \\ \in \mathbb{R}^{n_i}, & \text{if}\quad i \notin \Gamma(x^*). \end{cases}$$

On the other hand, for all $d \in T_S^B(x^*)$, it follows from $\|x^*\|_{2,0} = k$ and (4.1) that for $i \notin \Gamma(x^*)$, $d_i = 0$, and for $i \in \Gamma(x^*)$, there exists $d^* \in T_S^B(x^*)$ satisfying $d_i^* = -\xi_i^*$. Therefore, we get

$$0 = d_i^* = \arg\min_{d\in T_S^B(x^*)}\{\|d+\xi^*\|\}.$$

Consequently, by virtue of the definition of T^B-stationary point, we conclude that there exists $\lambda^* \in \Lambda^+$ such that $0 \in \partial_S^B f(x^*)$. This yields that $x^* \in S$ is an T^B-stationary point of GSMOP.

Case2: Let$\|x^*\|_{2,0} < k$.

" $\Leftarrow$ " Assume that $x^* \in S$ is an T^B-stationary point. Then there exists $\hat{\lambda} \in \Lambda^+$ such that

$$0 \in \partial_S^B f(x^*) = \arg\min_{d\in T_S^B(x^*)}\{\|d+\xi\| : \xi \in \partial_c f(x^*)^\top \hat{\lambda}\}.$$

Hence, there exists $\hat{\xi} \in \partial_c f(x^*)^\top \hat{\lambda}$ such that

$$\|\hat{\xi}\| = \|0+\hat{\xi}\| \le \|d+\hat{\xi}\|,\quad \forall\ d \in T_S^B(x^*). \tag{4.4}$$

Take any $i \in \{1, 2, \dots, p\}$. Let $\Gamma(\hat{d}) = i$ and $\hat{d}_i = -\hat{\xi}_i$. Then taking into account $\|x^*\|_{2,0} < k$, we have $\|\hat{d}\|_{2,0} \leq 1$ and $\|x^* + \mu\hat{d}\|_{2,0} \leq \|x^*\|_{2,0} + 1 \leq k$. Thus it stems from Lemma 2.2 that $\hat{d} \in T_S^B(x^*)$. According to (4.4), we get

$$|\hat{\xi}_i| \leq |\hat{d}_i + \hat{\xi}_i| = | - \hat{\xi}_i + \hat{\xi}_i|,$$

which yields that $\hat{\xi}_i = 0$. Then by virtue of the arbitrariness of i, one has

$$\hat{\xi} = 0. \tag{4.5}$$

Then it follows from Lemma 2.2 that

$$N_S^B(x^*) = \{0\}. \tag{4.6}$$

Together with (4.5) and (4.6), we derive that

$$0 \in \partial_c f(x^*)^\top \hat{\lambda} + N_S^B(x^*),$$

which implies that $x^* \in S$ is an N^B-stationary point of GSMOP.

" $\Rightarrow$ " Assume that $x^* \in S$ is an N^B-stationary point. Then there exists $\tilde{\lambda} \in \Lambda^+$ such that

$$0 \in \partial_c f(x^*)^\top \tilde{\lambda} + N_S^B(x^*).$$

It follows from (4.6) that there exists $\tilde{\xi} \in \partial_c f(x^*)^\top \tilde{\lambda}$ such that $\tilde{\xi} = 0$. Hence, we obtain

$$0 = \arg\min_{d \in T_S^B(x^*)} \{\|d + 0\|\} \in \arg\min_{d \in T_S^B(x^*)} \{\|d + \tilde{\xi}\|\}.$$

This implies that $0 \in \partial_S^B f(x^*)$ and so $x^* \in S$ is an T^B-stationary point of GSMOP. □

The equivalent relationship between T^C-stationary point and N^C-stationary point of GSMOP is established by the following theorem.

Theorem 4.2. *$x^* \in S$ is an N^C-stationary point of GSMOP if and only if $x^* \in S$ is an T^C-stationary point of GSMOP.*

Proof. " $\Leftarrow$ " Assume that $x^* \in S$ is an T^C-stationary point. Then there exists $\bar{\lambda} \in \Lambda^+$ such that

$$0 \in \partial_S^C f(x^*) = \arg\min_{d \in T_S^C(x^*)} \{\|d + \xi\| : \xi \in \partial_c f(x^*)^\top \bar{\lambda}\}. \tag{4.7}$$

According to Lemma 2.2, we have

$$N_S^C(x^*) = \{d \in \mathbb{R}^n : d_i = 0, i \in \Gamma(x^*)\} \tag{4.8}$$

and

$$T_S^C(x^*) = \{d \in \mathbb{R}^n : d_i = 0, i \notin \Gamma(x^*)\}. \tag{4.9}$$

If $i \notin \Gamma(x^*)$, then from (4.7) and (4.9) we obtain,

$$d_i = 0, \ \forall \ d \in \partial_S^C f(x^*). \tag{4.10}$$

If $i \in \Gamma(x^*)$, then by Lemma 2.2 (ii), for any $\xi \in \partial_c f(x^*)^\top \bar{\lambda}$, there exists $d \in \partial_S^C f(x^*)$ such that

$$d_i = -\xi_i. \tag{4.11}$$

Note that $0 \in \partial_S^C f(x^*)$. Then it follows from (4.10) and (4.11) that there exists $\bar{\xi} \in \partial_c f(x^*)^\top \bar{\lambda}$ such that

$$\bar{\xi}_i \begin{cases} = 0, & \text{if } \ i \in \Gamma(x^*) \\ \in \mathbb{R}^{n_i}, & \text{if } \ i \notin \Gamma(x^*). \end{cases} \tag{4.12}$$

Hence, according to (4.8) and (4.12), we have $-\bar{\xi} \in N_S^C(x^*)$, which implies that $0 \in \partial_c f(x^*)^\top \bar{\lambda} + N_S^C(x^*)$. This implies that $x^* \in S$ is an N^C-stationary point of GSMOP.

" $\Rightarrow$ " Assume that $x^* \in S$ is an N^C-stationary point. Then there exists $\lambda^* \in \Lambda^+$ such that

$$0 \in \partial_c f(x^*)^\top \lambda^* + N_S^C(x^*).$$

Thus there exists $\xi^* = (\xi_1^*, \xi_2^*, \ldots, \xi_p^*) \in \partial_c f(x^*)^\top \lambda^*$ such that

$$-\xi^* \in N_S^C(x^*). \tag{4.13}$$

Together with (4.8) and (4.13), we get

$$\xi_i^* \begin{cases} = 0, & \text{if } \ i \in \Gamma(x^*) \\ \in \mathbb{R}^{n_i}, & \text{if } \ i \notin \Gamma(x^*). \end{cases}$$

This combined with (4.9) yields that (4.10) holds and there exists $d^* \in T_S^C(x^*)$ such that

$$d_i^* = -\xi_i^* = 0, \quad i \in \Gamma(x^*).$$

Therefore,

$$d^* = 0 \in \arg\min_{d \in T_S^C(x^*)} \{\|d + \xi\| : \xi \in \partial_c f(x^*)^\top \lambda^*\},$$

which yields that $x^* \in S$ is an T^C-stationary point of GSMOP. □

We now pay the attention to the first-order optimality conditions. The following result presents the first-order necessary optimality condition of GSMOP under some suitable conditions.

Theorem 4.3. *Let $x^* \in S$ be a local weakly Pareto efficient solution of GSMOP. Then there exists $\lambda^* \in \Lambda^+$ such that x^* is an N^C-stationary point of GSMOP. Moreover, if $\partial_c f(x^*)^\top \lambda^* \subseteq T_S^C(x^*)$, then x^* is an N^B-stationary point of GSMOP.*

Proof. According to [11, Theorem 2.2 and Remark 2.3] and Λ^+ is a base of $\mathbb{R}^m_+$, there exists $\lambda^* \in \Lambda^+$ such that

$$0 \in \partial_c(\lambda^{*\top} f)(x^*) + N_S^C(x^*).$$

By virtue of Lemma 2.1, we have

$$0 \in \partial_c f(x^*)^\top \lambda^* + N_S^C(x^*), \tag{4.14}$$

which yields that x^* is an N^C-stationary point of GSMOP.

It follows from (4.14) and Lemma 2.2 that there exists $\xi^* \in \partial_c f(x^*)^\top \lambda^*$ such that

$$\xi_i^* = 0, \ \forall \ i \in \Gamma(x^*).$$

Note that $\partial_c f(x^*)^\top \lambda^* \subseteq T_S^C(x^*)$. Then by Lemma 2.2, we derive that for all $\xi \in \partial_c f(x^*)^\top \lambda^*$

$$\xi_i = 0, \ \forall \ i \notin \Gamma(x^*),$$

which means that $\xi^* = 0$. Therefore, we have

$$0 = \xi^* \in \partial_c f(x^*)^\top \lambda^*. \tag{4.15}$$

Due to (4.15) and Lemma 2.2 (iii), one has

$$0 \in \partial_c f(x^*)^\top \lambda^* + N_S^B(x^*),$$

which implies that x^* is an N^B-stationary point of GSMOP. □

The following theorem establishes the sufficient optimality condition for GSMOP.

Theorem 4.4. *Let $x^* \in S$ be a strict N^B-stationary point of GSMOP and f be ∂_c-pseudoconvex at x^*. If $S_f(x^*)$ is a pseudoconvex set at x^* with respect to $T_{S_f(x^*)}(x^*)$, then x^* is a weakly Pareto efficient solution of GSMOP.*

Proof. Suppose that x^* is not a weakly Pareto efficient solution of GSMOP. Then there exists a feasible point $x' \neq x^*$ such that

$$f(x') - f(x^*) \in -\mathbb{R}^m_{++}.$$

This yields that

$$\lambda^\top f(x') < \lambda^\top f(x^*), \forall \lambda \in \Lambda^+. \tag{4.16}$$

Since $x^* \in S$ is a strict N^B-stationary point of GSMOP, it follows from Lemma 2.1 that there exists $\lambda^* \in \Lambda^+$ such that

$$-\partial_c({\lambda^*}^\top f)(x^*) \subseteq -\partial_c f(x^*)^\top \lambda^* \subseteq N_S^B(x^*). \tag{4.17}$$

By virtue of Lemma 2.2, we conclude that

$$T_S^B(x^*) = \{d \in \mathbb{R}^n : \|d\|_{2,0} \leq k, \|x^* + \mu d\|_{2,0} \leq k, \forall \mu \in \mathbb{R}\}$$

and

$$N_S^B(x^*) = \begin{cases} \{d \in \mathbb{R}^n : d_i = 0, i \in \Gamma(x^*)\}, & \text{if } \ \|x\|_{2,0} = k, \\ \{0\}, & \text{if } \ \|x\|_{2,0} < k. \end{cases}$$

Take any $\xi \in N_S^B(x^*)$ and $\eta \in T_S^B(x^*)$. If $|\Gamma(x^*)| = s$, then $\xi = 0$ and so $\langle \xi, \eta \rangle = 0$. If $|\Gamma(x^*)| < s$, then $\xi_i = 0, i \in \Gamma(x^*)$ and $\eta_i = 0, i \notin \Gamma(x^*)$ and so $\langle \xi, \eta \rangle = 0$. Therefore, we have

$$\langle \xi, \eta \rangle = 0, \quad \forall\ \xi \in N_S^B(x^*),\ \eta \in T_S^B(x^*).$$

This together with (4.17) implies that

$$\langle \xi', \eta \rangle = 0, \quad \forall \xi' \in \partial_c({\lambda^*}^\top f)(x^*),\ \eta \in T_S^B(x^*), \tag{4.18}$$

where $\xi' = -\xi$. It follows from (4.16) that $x' \in S_f(x^*)$. Note that $S_f(x^*) \subseteq S$. Then

$$T_{S_f(x^*)}(x^*) \subseteq T_S^B(x^*).$$

Since $S_f(x^*)$ is a pseudoconvex set at x^* with respect to $T_{S_f(x^*)}(x^*)$, we have

$$S_f(x^*) \subseteq x^* + T_{S_f(x^*)}(x^*),$$

which yields that

$$x' - x^* \in T_{S_f(x^*)}(x^*) \subseteq T_S^B(x^*).$$

Consequently, this together with (4.18) implies that

$$\langle \xi', x' - x^* \rangle = 0, \quad \forall \xi' \in \partial_c ({\lambda^*}^\top f)(x^*). \tag{4.19}$$

Besides, observe that f is ∂_c-pseudoconvex at x^*. By Remark 2.1, we get

$$\langle \xi', x' - x^* \rangle < 0, \forall \xi' \in \partial_c ({\lambda^*}^\top f)(x^*),$$

which contradicts with (4.19). Therefore, x^* is a weakly Pareto efficient solution of GSMOP. □

5. Second-order optimality conditions

In this section, we discuss the second-order necessary and sufficient optimality conditions by using the second-order tangent set and Dini directional derivatives.

We now give a second-order necessary optimality condition of GSMOP.

Theorem 5.1. *Let $x^* \in S$ be a local weakly efficient solution of GSMOP. Then there exists $\lambda^* \in \Lambda^+$ such that for each $v \in T_S^C(x^*)$ with $D^-({\lambda^*}^\top f)(x^*; v) = 0$, the following statement holds:*

$$D^2({\lambda^*}^\top f)(x^*; v, w) \geq 0, \quad \forall\, w \in T_S^2(x^*, v).$$

Proof. On the contrary, suppose that for any $\lambda \in \Lambda^+$, there exists some $v^* \in T_S^C(x^*)$ satisfying $D^-(\lambda^\top f)(x^*; v^*) = 0$ and $w^* \in T_S^2(x^*, v^*)$ such that

$$D^2({\lambda^*}^\top f)(x^*; v^*, w^*) < 0.$$

According to the definition of the second-order Dini derivatives, we have

$$\begin{aligned} &D^2({\lambda^*}^\top f)(x^*; v^*, w^*) \\ &= \liminf_{t \searrow 0} \frac{\lambda^\top f(x^* + tv^* + t^2 w^*) - \lambda^\top f(x^*) - tD^-(\lambda^\top f)(x^*; v^*)}{t^2} < 0. \end{aligned}$$

This together with $D^-(\lambda^\top f)(x^*; v^*) = 0$ yields that there exists a sufficiently little $t^* > 0$ such that

$$\frac{\lambda^\top f(x^* + tv^* + t^2 w^*) - \lambda^\top f(x^*)}{t^2} < 0, \quad \forall\, t \in (0, t^*].$$

Hence, we get

$$\lambda^\top (f(x^* + tv^* + t^2 w^*) - f(x^*)) < 0, \quad \forall\, t \in (0, t^*].$$

Note that $x^* + tv^* + t^2 w^* \subseteq S$ by Proposition 3.2. Then it follows from the arbitrariness of $\lambda \in \Lambda^+$ and the above inequality that

$$f(x^* + tv^* + t^2 w^*) - f(x^*) \in -\mathbb{R}^m_{++}, \quad \forall\, t \in (0, t^*],$$

which is a contradiction duo to x^* is a local weakly efficient solution of GSMOP. □

The following theorem presents a second-order sufficient optimality conditions of GSMOP in terms of the first and second-order Dini directional derivatives.

Theorem 5.2. *Assume that $x^* \in S$ and there exists $\lambda^* \in \Lambda^+$ such that the following condition (5.1) holds for each $v \in T_S^B(x^*)$:*

$$D^-(\lambda^{*\top} f)(x^*; v) \geq 0,\ D^2(\lambda^{*\top} f)(x^*; v, w) > 0, \quad \forall\ w \in T_S^2(x^*, v). \tag{5.1}$$

Then x^* is a local weakly Pareto efficient solution of GSMOP.

Proof. Suppose on the contrary that x^* is not local weakly Pareto efficient solution of GSMOP. Then there exists a sequence $x^k \xrightarrow{S} x^*$ such that

$$f(x^k) - f(x^*) \in -\mathbb{R}^m_+ \setminus 0.$$

Then we have

$$\lambda^\top (f(x^k) - f(x^*)) < 0, \quad \forall\ \lambda \in \Lambda^+. \tag{5.2}$$

Set $v^k = \frac{x^k - x^*}{\|x^k - x^*\|}$. Then without loss of generality, we can assume that $v^k \longrightarrow v$. Since $x^k \xrightarrow{S} x^*$, it follows from Lemma 2.2 that $v \in T_S^B(x^*)$. Set $t_k = \|x^k - x^*\|$. Then due to (5.2), we obtain

$$\lambda^\top (f(x^* + t_k v^k) - f(x^*)) < 0, \forall \lambda \in \Lambda^+, \tag{5.3}$$

which implies that for any $\lambda \in \Lambda^+$,

$$\begin{aligned} D^-(\lambda^\top f)(x^*; v) &= \liminf_{t \searrow 0} \frac{\lambda^\top f(x^* + tv) - \lambda^\top f(x^*)}{t} \\ &\leq \liminf_{k \longrightarrow \infty} \frac{\lambda^\top f(x^* + t_k v^k) - \lambda^\top f(x^*)}{t_k} \\ &\leq 0. \end{aligned}$$

This combined with $D^-(\lambda^{*\top} f)(x^*; v) \geq 0$ yields that

$$D^-(\lambda^{*\top} f)(x^*; v) = 0.$$

Set $q_k = \frac{\sqrt{1+4t_k} - 1}{2}$. Then $q_k^2 + q_k = t_k$. Consequently, $x^k = x^* + t_k v^k = x^* + (q_k^2 + q_k) v^k$ and so,

$$\begin{aligned} & D^2(\lambda^{*\top} f)(x^*; v, v) \\ &= \liminf_{t \searrow 0} \frac{\lambda^{*\top} f(x^*(t + t^2)v) - \lambda^{*\top} f(x^*) - tD^-(\lambda^{*\top} f)(x^*; v)}{t^2} \\ &= \liminf_{t \searrow 0} \frac{\lambda^{*\top} f(x^* + (t + t^2)v) - \lambda^{*\top} f(x^*)}{t^2} \\ &\leq \liminf_{k \longrightarrow \infty} \frac{\lambda^{*\top} f(x^* + (q_k^2 + q_k)v^k) - \lambda^{*\top} f(x^*)}{q_k^2} \\ &= \liminf_{k \longrightarrow \infty} \frac{\lambda^{*\top} f(x^* + t_k v^k) - \lambda^{*\top} f(x^*)}{q_k^2} \\ &\leq 0, \end{aligned} \tag{5.4}$$

where the last inequality comes from (5.3). Note that $v \in T_S^B(x^*)$. Then one has

$$\|v\|_{2,0} \leq k,\ \|x^* + \alpha v\|_{2,0} \leq k,\ \|v + \beta v\|_{2,0} \leq k, \quad \forall\ \alpha, \beta \in \mathbb{R}.$$

Therefore, from Remark 3.1 we have

$$v \in T_S^2(x^*, v). \tag{5.5}$$

Hence, it follows from (5.4) and (5.5) that there exists $v \in T_S^2(x^*, v)$ such that $D^2(\lambda^{*\top} f)(x^*; v, v) \leq 0$, which contradicts the fact that for all $w \in T_S^2(x^*, v)$, $D^2(\lambda^{*\top} f)(x^*; v, w) > 0$. □

Acknowledgments

This paper was partially supported by the Youth Project of Science and Technology Research Program of Chongqing Education Commission of China (No. KJQN202201802), and the Natural Science Foundation of Chongqing (No. cstc2021jcyj-msxmX0925, cstc2022ycjh-bgzxm0097).

References

[1] D. Aussel, *Subdifferential properties of quasiconvex and pseudoconvex functions*, J. Optim. Theory Appl. **97** (1998), 29–45.

[2] A. Beck and N. Hallak, *Optimization involving group sparsity terms*, Math. Program. **178** (2018), 39–67.

[3] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, Berlin, 2000.

[4] J. W. Chen and Y. Dai, *Multiobjective optimization with least constraint violation: optimality conditions and exact penalization*, J. Global Optim. **87** (2023), 807–830.

[5] J. W. Chen, L. Huang and S. J. Li, *Separations and optimality of constrained multiobjective optimization via improvement sets*, J. Optim. Theory Appl. **178** (2018), 794–823.

[6] J. W. Chen, S. J. Li and J. C. Yao, *Vector-valued separation functions and constrained vector optimization problems: Optimality and saddle points*, J. Ind. Manag. Optim. **16** (2020), 707–724.

[7] J. W. Chen, J. Li and J. N. Wang, *Duality and optimality conditions for nonsmooth multiobjective optimization problems with cone constraints*, Acta Math. Sci. **32** (2012), 1–12.

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

[9] P. Y. Chen and I. W. Selesnick, *Group-sparse signal denoising: non-convex regularization, convex optimization*, IEEE Trans. Signal Process. **62** (2014), 3464–3478.

[10] Y. H. Hu, C. Li, K. W. Meng, J. Qin and X. Q. Yang, *Group sparse optimization via $\ell_{p,q}$ regularization*, J. Mach. Learn. Res. **18** (2017), 1–52.

[11] X. X. Huang, *Optimality conditions and approximate optimality conditions in locally Lipschitz vector optimization*, Optim. **51** (2002), 309–321.

[12] V. I. Ivanov, *On the functions with pseudoconvex sublevel sets and optimality conditions*, J. Math. Anal. Appl. **345** (2008), 964–974.

[13] J. Jahn, *Vector Optimization Theory, Applications and Extensions*, Springer, Berlin, 2004.

[14] F. Lauer, H. Ohlsson, *Finding sparse solutions of systems of polynomial equations via group-sparsity optimization*, J. Glob. Optim. **62** (2015), 319–349.

[15] T. M. Leschine, H. Wallenius and W. A. Verdini, *Interactive multiobjective analysis and assimilative capacitybased ocean disposal decisions*, Eur. J. Oper. Res. **56** (1992), 278–289.

[16] D. T. Luc, *Theory of Vector Optimization*, Lecture Notes in Economics and Mathematical Systems, Springer, Berlin, 1989.

[17] D. T. Peng and X. J. Chen, *Computation of second-order directional stationary points for group sparse optimization*, Optim. Meth. Softw. **35** (2020), 348–376.

[18] R.T. Rockafellar, R.J. Wets, *Variational Analysis*, Springer, Berlin, 1998.

[19] M. Studniarski, *Second-order necessary conditions for optimality in nonsmooth nonlinear programming*, J. Math. Anal. Appl. **154** (1991), 303–317.

[20] M. M. Wiecek, *Advances in cone-based preference modeling for decision making with multiple criteria*, Decis. Mark. Manuf. Serv. **1** (2007), 153–173.
[21] W. Y. Wu and D. T. Peng, *Optimality conditions for group sparse constrained optimization problems*, Math. **9** (2021): 84.
[22] M. Yuan, Y. Lin, *Model selection and estimation in regression with grouped variables*, J. R. Stat. Soc. Ser. B **68** (2006), 49–67.

Manuscript received November 20 2023
revised April 18 2024

Huasheng Su
School of Mathematics and Statistics, Southwest University, Chongqing 400715, China
E-mail address: SHS13516355736@163.com

La Huang
School of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China
E-mail address: angelabjy111@163.com

Xiaoqing Ou
College of Management, Chongqing College of Humanities, Science & Technology, Chongqing 401524, China
E-mail address: ouxiaoqing413@163.com

Jiawei Chen
School of Mathematics and Statistics, Southwest University, Chongqing 400715, China
E-mail address: j.w.chen713@163.com