J. Oper. Res. Soc. China (2015) 3:421–439
DOI 10.1007/s40305-015-0101-3

On Solutions of Sparsity Constrained Optimization

Li-Li Pan[1,2] · **Nai-Hua Xiu**[1] · **Sheng-Long Zhou**[1,3]

Received: 28 March 2015 / Revised: 6 October 2015 / Accepted: 10 October 2015 /
Published online: 30 October 2015

Abstract In this paper, we mainly study the existence of solutions to sparsity constrained optimization (SCO). Based on the expressions of tangent cone and normal cone of sparsity constraint, we present and characterize two first-order necessary optimality conditions for SCO: N-stationarity and T-stationarity. Then we give the second-order necessary and sufficient optimality conditions for SCO. At last, we extend these results to SCO with nonnegative constraint.

Keywords Sparsity constrained optimization · Tangent cone · Normal cone · First-order optimality condition · Second-order optimality condition

Mathematics Subject Classification 90C26 · 90C46

The work was supported in part by the National Natural Science Foundation of China (Nos. 11431002, 71271021).

✉ Li-Li Pan
panlili1979@163.com

Nai-Hua Xiu
nhxiu@bjtu.edu.cn

Sheng-Long Zhou
longnan_zsl@163.com

1 Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044, China

2 Department of Mathematics, Shandong University of Technology, Zibo 255049, Shandong, China

3 School of Mathematics, University of Southampton, Southampton SO17 1BJ, UK

1 Introduction

Sparsity constrained optimization (SCO) is to minimize a general nonlinear function subject to sparsity constraint. It has wide applications in signal and image processing, machine learning, pattern recognition and computer vision, and so on. Recent years have witnessed a growing interest in the theory and algorithm for SCO, especially for SCO caused by the sparse recovery from linear or nonlinear observations.

SCO is actually a combinatorial optimization problem and it is generally NP-hard even for quadratic objective function [1]. So, usually continuous optimization theory is unable to cope with SCO and the research on existence of solutions to SCO is difficult. Negahban et al. [2] introduced the concept of restricted strong convexity (RSC) in convex optimization problem, and showed it is a sufficient condition for existence of unique solution for the convex program in a restricted region. Agarwal et al. [3] modified RSC and introduced the notions of restricted strong smoothness (RSS) that suffice to establish global linear convergence of a first-order method for the above problem. Later, various variants of RSC and RSS were developed and used in [4–7] to guarantee the existence of unique solution and corresponding error bound holding for SCO. Bahmani et al. [8] proposed stable restricted Hessian (SRH) for twice continuously differentiable objective functions and stable restricted linearization (SRL) for nonsmooth objective functions, and obtained that they are sufficient conditions to achieve the true unique solution to SCO in a finite number of iterations. Here, we should point out that these conditions can be regarded as some extensions or relaxations of restricted isometry property (RIP) by Candès and Tao [9] in compressed sensing (CS) [9,10], where RIP suffices to guarantee that the l_0-minimization associated with CS from linear observations has a unique solution and is polynomially solvable. Recently, Beck and Eldar [11] introduced and analyzed three kinds of first-order necessary optimality conditions for the existence of solutions to SCO: basic feasibility, L-stationarity and coordinate-wise optimality. Basic feasibility is an extension of the necessary condition that gradient is zero for unconstrained optimization. L-stationarity is based on the notion of fixed-point equation for convex constrained problems and used to derive the iterative hard thresholding (IHT) algorithm for SCO.

In this paper, we try to investigate the existence of solutions of SCO by using the concepts of tangent cone and normal cone, which are very useful in characterizing feasible region and establishing optimality condition for constrained optimization problems. We first give the expressions of the Bouligand and Clarke tangent cones and normal cones of sparsity constraint. Then we establish the concepts of N-stationarity and T-stationarity for SCO, and analyze the relationship and differences among basic feasibility, L-stationarity, N-stationarity, and T-stationarity. Our analysis shows that they play an important role in optimality theory and algorithm design for SCO. Further, we give the second-order necessary and sufficient optimality conditions under the sense of Clarke tangent cone for the existence of local optimal solutions to SCO. Especially, we conclude the existence theorems for global optimal solutions to SCO with least squares objective function from sparse recovery of linear observations. At last, we extend these results to SCO with nonnegative constraint.

This paper is organized as follows. Section 2 studies the first- and second-order optimality conditions for sparsity constrained optimization. Section 3 extends the corresponding results in Sect. 2 to SCO with nonnegative constraint. The last section gives some concluding remarks.

2 Optimality Conditions

In this section, we study the first- and second-order necessary optimality conditions of the following SCO problem

$$\min \quad f(x), \qquad \text{s.t.} \ \|x\|_0 \leqslant s, \tag{2.1}$$

where $f(x):\mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable or twice continuously differentiable function. $\|x\|_0$ is the l_0-norm of $x \in \mathbb{R}^n$, which refers to the number of nonzero elements in the vector x. $s < n$ is a positive integer. Let $S \triangleq \{x \in \mathbb{R}^n : \ \|x\|_0 \leqslant s\}$ be the feasible region of (2.1).

We first consider the projection on sparse set S, named support projection. For $\Omega \subset \mathbb{R}^n$ being nonempty and closed, we call the mapping $\mathrm{P}_\Omega : \mathbb{R}^n \to \Omega$ the projector onto Ω if

$$\mathrm{P}_\Omega(x) := \mathrm{argmin}_{y\in\Omega}\|x-y\|.$$

It is well known that the support projection $\mathrm{P}_S(x)$ sets all but s largest absolute value components of x to zero. Carefully speaking, letting $I_s(x) := \{j_1, j_2, \cdots, j_s\} \subseteq \{1, 2, \cdots, n\}$ of indices of x with $\min_{i\in I_s(x)}|x_i| \geqslant \max_{i\notin I_s(x)}|x_i|$, we have

$$\mathrm{P}_S(x) = \left\{\, y \in \mathbb{R}^n : y_i = x_i, i \in I_s(x);\ \ y_i = 0, i \notin I_s(x) \,\right\}.$$

As S is nonconvex, the support projection $\mathrm{P}_S(x)$ is not single-valued. Letting $M_i(x)$ denote the ith largest element of x, we have

$$M_1(x) \geqslant M_2(x) \geqslant \cdots \geqslant M_n(x).$$

Denote $|x| = (|x_1|, \cdots, |x_n|)^{\mathrm{T}}$. If $M_s(|x|) = 0$ or $M_s(|x|) > M_{s+1}(|x|)$, then $\mathrm{P}_S(x)$ is single-valued, in such case,

$$(\mathrm{P}_S(x))_i = \begin{cases} x_i, & |x_i| \geqslant M_s(|x|), \\ 0, & \text{otherwise.} \end{cases}$$

Otherwise,

$$(\mathrm{P}_S(x))_i = \begin{cases} x_i, & |x_i| > M_s(|x|), \\ x_i \text{ or } 0, & |x_i| = M_s(|x|), \\ 0, & \text{otherwise.} \end{cases}$$

If there are more than one x_i such that $|x_i| = M_s(|x|)$, only one element can be selected either randomly or according to some predefined ordering to be assigned itself, while others are assigned zeros.

2.1 Tangent Cone and Normal Cone

Recalling that for any nonempty set $\Omega \subseteq \mathbb{R}^n$, its Bouligand tangent cone $\mathrm{T}^B_\Omega(\bar{x})$, Clarke tangent cone $\mathrm{T}^C_\Omega(\bar{x})$ and corresponding Normal Cones $\mathrm{N}^B_\Omega(\bar{x})$ and $\mathrm{N}^C_\Omega(\bar{x})$ at point $\bar{x} \in \Omega$ are defined as [12]:

$$\mathrm{T}^B_\Omega(\bar{x}) := \left\{ d \in \mathbb{R}^n : \begin{array}{r} \exists \{x^k\} \subset \Omega, \lim_{k\to\infty} x^k = \bar{x}, \ \lambda_k \geqslant 0, \ k \in \mathbb{N} \ \text{such that} \\ \lim_{k\to\infty} \lambda_k (x^k - \bar{x}) = d \end{array} \right\},$$

$$\mathrm{T}^C_\Omega(\bar{x}) := \left\{ d \in \mathbb{R}^n : \begin{array}{r} \text{For } \forall \{x^k\} \subset \Omega, \ \forall \{\lambda_k\} \subset \mathbb{R}_+ \text{ with } \lim_{k\to\infty} x^k = \bar{x}, \\ \lim_{k\to\infty} \lambda_k = 0, \ \exists \{y^k\} \text{ such that} \\ \lim_{k\to\infty} y^k = d \text{ and } x^k + \lambda_k y^k \in \Omega, \ k \in \mathbb{N} \end{array} \right\},$$

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

$$\mathrm{N}^C_\Omega(\bar{x}) := \left\{ d \in \mathbb{R}^n : \langle d, z \rangle \leqslant 0, \ \forall z \in \mathrm{T}^C_\Omega(\bar{x}) \right\}. \tag{2.2}$$

The following two theorems give the expressions of Bouligand and Clarke tangent cones and normal cones of sparse set S.

Theorem 2.1 *For any $\bar{x} \in S$ and letting $\Gamma = \mathrm{supp}(\bar{x})$, the Bouligand tangent cone and corresponding normal cone of S at $\bar{x}$ are respectively*

$$\mathrm{T}^B_S(\bar{x}) = \{ d \in \mathbb{R}^n : \|d\|_0 \leqslant s, \ \|\bar{x} + \mu d\|_0 \leqslant s, \forall \mu \in \mathbb{R} \} \tag{2.3}$$

$$= \bigcup_{\Upsilon} \mathrm{span}\{ e_i, \ i \in \Upsilon \supseteq \Gamma, |\Upsilon| \leqslant s \}, \tag{2.4}$$

$$\mathrm{N}^B_S(\bar{x}) = \begin{cases} \{ d \in \mathbb{R}^n : d_i = 0, \ i \in \Gamma \} = \mathrm{span}\{ e_i, \ i \notin \Gamma \}, & \text{if } |\Gamma| = s, \\ \{0\}, & \text{if } |\Gamma| < s, \end{cases} \tag{2.5}$$

where $e_i \in \mathbb{R}^n$ is a vector whose ith component is one and others are zeros, $\mathrm{span}\{e_i, i \in \Gamma\}$ *denotes the subspace of $\mathbb{R}^n$ spanned by $\{ e_i : i \in \Gamma\}$, and*

$$\mathrm{supp}(x) = \{i \in \{1, \cdots, n\} : x_i \neq 0\}.$$

Proof Denote the set of the right hand of (2.3) as D. It is not difficult to verify that D is equal to (2.4), and thus we only prove (2.3). First we prove $\mathrm{T}^B_S(\bar{x}) \subseteq D$. For any $d \in \mathrm{T}^B_S(\bar{x})$, there is $\lim_{k\to\infty} x^k = \bar{x}, \ x^k \in S, \ \lambda_k \geqslant 0$ satisfies $d = \lim_{k\to\infty} \lambda_k (x^k - \bar{x})$. Since $\lim_{k\to\infty} x^k = \bar{x}$, there is k_0 when $k \geqslant k_0$, $\Gamma \subseteq \mathrm{supp}(x^k)$. In addition, $d = \lim_{k\to\infty} \lambda_k (x^k - \bar{x})$ derives $\mathrm{supp}(d) \subseteq \mathrm{supp}(x^k)$, which combining with $\|x^k\|_0 \leqslant s$ when $k \geqslant k_0$ and $\Gamma \subseteq \mathrm{supp}(x^k)$ yields that $\|d\|_0 \leqslant s$ and $\|\bar{x} + \mu d\|_0 \leqslant s, \forall \mu \in \mathbb{R}$. Next we prove $\mathrm{T}^B_S(\bar{x}) \supseteq D$. For any $\|d\|_0 \leqslant s, \|\bar{x} + \mu d\|_0 \leqslant s, \forall \mu \in \mathbb{R}$, we take any sequence $\{\lambda_k\}$ such that $\lambda_k > 0$ and $\lambda_k \to +\infty$. Then by defining $\{x^k\}$ with $x^k = \bar{x} + d/\lambda_k$, evidently $x^k \in S$, $\lim_{k\to\infty} x^k = \bar{x}$, and $d = \lim_{k\to\infty} \lambda_k (x^k - \bar{x})$, which implies $d \in \mathrm{T}^B_S(\bar{x})$.

For (2.5), by the definition of $\mathrm{N}_S^B(\bar{x})$, we obtain

$$\begin{aligned}\mathrm{N}_S^B(\bar{x}) &= \{\, d \in \mathbb{R}^n : \langle d, z\rangle \leqslant 0,\ \forall\, z \in \mathrm{T}_S^B(\bar{x})\,\}\\ &= \{\, d \in \mathbb{R}^n : \langle d, z\rangle \leqslant 0,\ \|z\|_0 \leqslant s,\ \|\bar{x}+\mu z\|_0 \leqslant s\ ,\forall\, \mu \in \mathbb{R}\,\}. \end{aligned} \tag{2.6}$$

If $|\Gamma| = s$, it yields $\mathrm{supp}(z) \subseteq \Gamma$ for any $z \in \mathrm{T}_S^B(\bar{x})$. Then

$$\begin{aligned} d \in \mathrm{N}_S^B(\bar{x}) &\Longleftrightarrow \langle d, z\rangle \leqslant 0, \quad \forall\, \mathrm{supp}(z) \subseteq \Gamma \Longleftrightarrow d_i \begin{cases} = 0, & i \in \Gamma,\\ \in \mathbb{R}, & i \notin \Gamma. \end{cases}\\ &\Longleftrightarrow d \in \mathrm{span}\,\{\, e_i,\ \ i \notin \Gamma\,\}. \end{aligned}$$

If $|\Gamma| < s$, we will prove $\mathrm{N}_S^B(\bar{x}) = \{0\}$. Assuming $d \in \mathrm{N}_S^B(\bar{x})$, we take $z = d_{i_0} e_{i_0}, \forall\, i_0 \in \{1, 2, \cdots, n\}$, then $z \in \mathrm{T}_S^B(\bar{x})$ because of $|\Gamma| < s$. By $\langle d, z\rangle = d_{i_0}^2 \leqslant 0$, we can obtain $d_{i_0} = 0$. The arbitrariness of i_0 yields that $d = 0$, henceforth $\mathrm{N}_S^B(\bar{x}) = \{0\}$.

Theorem 2.2 *For any $\bar{x} \in S$ and letting $\Gamma = \mathrm{supp}(\bar{x})$, the Clarke tangent cone and corresponding normal cone of S at $\bar{x}$ are, respectively,*

$$\mathrm{T}_S^C(\bar{x}) = \{\, d \in \mathbb{R}^n : \mathrm{supp}(d) \subseteq \Gamma\,\} = \mathrm{span}\,\{\, e_i,\ \ i \in \Gamma\,\}, \tag{2.7}$$

$$\mathrm{N}_S^C(\bar{x}) = \mathrm{span}\,\{\, e_i,\ \ i \notin \Gamma\,\}. \tag{2.8}$$

Proof Obviously, $\mathrm{span}\,\{\, e_i,\ \ i \in \Gamma\,\} = \{\, d \in \mathbb{R}^n : \mathrm{supp}(d) \subseteq \Gamma\,\}$.

We first prove $\mathrm{T}_S^C(\bar{x}) \subseteq \{\, d \in \mathbb{R}^n : \mathrm{supp}(d) \subseteq \Gamma\,\}$. For any $d \in \mathrm{T}_S^C(\bar{x})$, we have $\forall\, \{x^k\} \subset S,\ \forall\, \{\lambda_k\} \subset \mathbb{R}_+$ with $\lim_{k\to\infty} x^k = \bar{x},\ \lim_{k\to\infty} \lambda_k = 0$, there is a sequence $\{y^k\}$ such that $\lim_{k\to\infty} y^k = d$ and $x^k + \lambda_k y^k \in S,\ k = 1, 2, \cdots$. Assume that $\mathrm{supp}(d) \nsubseteq \Gamma$, namely there is an $i_0 \in \mathrm{supp}(d)$ but $i_0 \notin \Gamma$. Since $\lim_{k\to\infty} y^k = d$, it must have $y_{i_0}^k \to d_{i_0}$ which requires $y_{i_0}^k \neq 0$ when $k \geqslant k_0$. By the arbitrariness of $\{x^k\}$, we take $\{x^k\} \subset S$ such that $\lim_{k\to\infty} x^k = \bar{x}$, $\mathrm{supp}(x^k) = \Gamma \cup \Gamma_k$ with $|\Gamma \cup \Gamma_k| = s$, where $\Gamma_k \subset \{1, 2, \cdots, n\}\backslash\Gamma$, and $i_0 \notin \Gamma_k$. Because $\{y^k\}$ is fixed and the arbitrariness of $\{\lambda_k\}$, we can take $\{\lambda_k\}$ such that

$$\lambda_k = \min_{i\in\Gamma\cup\Gamma_k} \frac{|x_i^k|}{|y_i^k| + k}.$$

Then $\lambda_k (\neq 0) \to 0$ as $k \to \infty$. And thus $\forall i \in \Gamma \cup \Gamma_k$, it follows

$$|x_i^k + \lambda_k y_i^k| \geqslant |x_i^k| - \lambda_k |y_i^k| = |x_i^k| - |y_i^k| \min_{i\in\Gamma\cup\Gamma_k} \frac{|x_i^k|}{|y_i^k| + k} > 0.$$

Moreover, from $i_0 \notin \Gamma \cup \Gamma_k$ deriving $x_{i_0}^k = 0,\ y_{i_0}^k \neq 0$, we must have $\|x^k + \lambda_k' y^k\|_0 \geqslant s + 1$ for $k \geqslant k_0$, which is contradicted to $x^k + \lambda_k' y^k \in S$ for any $k = 1, 2, \cdots$. Therefore $\mathrm{supp}(d) \subseteq \Gamma$.

Next we prove $\mathrm{T}_S^C(\bar{x}) \supseteq \{ d \in \mathbb{R}^n : \operatorname{supp}(d) \subseteq \Gamma \}$. For any $d \in \mathbb{R}^n$ such that $\operatorname{supp}(d) \subseteq \Gamma$ and $\forall \{x^k\} \subset S$, $\forall \{\lambda_k\} \subset \mathbb{R}_+$ with $\lim_{k\to\infty} x^k = \bar{x}$, $\lim_{k\to\infty} \lambda_k = 0$, we have $\operatorname{supp}(d) \subseteq \Gamma \subseteq \operatorname{supp}(x^k)$ for any $k \geqslant k_0$. Let

$$\begin{array}{ll} y^k = 0, & k = 1, 2, \cdots, k_0, \\ y^k = x^k - \bar{x} + d, & k = k_0 + 1, \quad k_0 + 2, \cdots, \end{array}$$

which brings out $x^k + \lambda_k y^k \in S$ for $k = 1, 2, \cdots$ due to $x^k \in S$. In addition, $\lim_{k\to\infty} y^k = \lim_{k\to\infty} x^k - \bar{x} + d = d$. Hence $d \in \mathrm{T}_S^C(\bar{x})$.

Finally (2.8) holding is obvious. Then the whole proof is completed.

Remark 2.3 Clearly for any $\bar{x} \in S$, Bouligand tangent cone $\mathrm{T}_S^B(\bar{x})$ is closed but nonconvex, while Clarke tangent cone $\mathrm{T}_S^C(\bar{x})$ is closed and convex. In addition, $\mathrm{T}_S^C(\bar{x}) \subseteq \mathrm{T}_S^B(\bar{x})$.

To end this subsection, we give a plot to illustrate the Bouligand and Clarke tangent cones in three dimensional space (Fig. 1). One can easily verify $\mathrm{T}_S^B(x') = \mathrm{T}_S^C(x') = \{x \in \mathbb{R}^3 | \ x_1 = 0\}$, $\mathrm{T}_S^B(x'') = \{x \in \mathbb{R}^3 | \ x_1 = 0\} \cup \{x \in \mathbb{R}^3 | \ x_3 = 0\}$ and $\mathrm{T}_S^C(x'') = \{x \in \mathbb{R}^3 | \ x_1 = x_3 = 0\}$.

2.2 First-Order Optimality Conditions

When $f(x)$ is continuously differentiable on $\mathbb{R}^n$, we give the definitions of N-stationarity and T-stationarity of problem (2.1) based on the expressions of tangent cone and normal cone.

Definition 2.4 A vector $x^* \in S$ is called an $\mathrm{N}^\sharp$-stationary point and $\mathrm{T}^\sharp$-stationary point of (2.1) if it respectively satisfies the relation

$$\mathrm{N}^\sharp\text{-stationary point}: \quad 0 \in \nabla f(x^*) + \mathrm{N}_S^\sharp(x^*), \tag{2.9}$$

$$\mathrm{T}^\sharp\text{-stationary point}: \quad 0 = \|\nabla_S^\sharp f(x^*)\|, \tag{2.10}$$

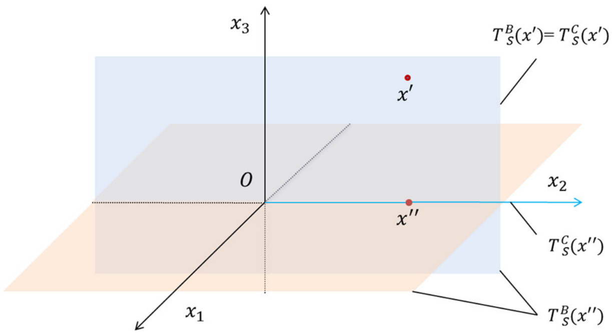

Fig. 1 Bouligand tangent cone and Clarke tangent cone in three dimensional space, where $S = \{x \in \mathbb{R}^3 | \ \|x\|_0 \leqslant 2\}$ and $x' = (0, 1, 1)^{\mathrm{T}}$, $x'' = (0, 1, 0)^{\mathrm{T}}$

where $\nabla^{\sharp}_S f(x^*) = \mathrm{argmin}\{\ \|d + \nabla f(x^*)\| :\ d \in \mathrm{T}^{\sharp}_S(x^*)\ \}$, $\sharp \in \{B, C\}$ stands for the sense of Bouligand tangent cone or Clarke tangent cone.

Note that $\mathrm{N}^{\sharp}$-stationary point is called the critical point of nonconvex optimization problem by Attouch et al. [13], $\mathrm{T}^{\sharp}$-stationary point is the zero point of projection gradient of objective function introduced by Calamai and Mor̆e in [14]. Recall that $x^* \in S$ is defined in [11] as an L-stationary point of (2.1) if

$$x^* \in \mathrm{P}_S(x^* - \nabla f(x^*)/L), \quad \forall\, L > 0. \tag{2.11}$$

L-stationary point with $L = 1$ is called a fixed point in [15].

It is well known that L-stationarity, $\mathrm{N}^{\sharp}$-stationarity and $\mathrm{T}^{\sharp}$-stationarity are equivalent in convex constrained optimization problems. But for SCO problem, it is not the case observed from the following theorem.

Theorem 2.5 *Under the concept of Bouligand tangent cone, we consider problem* (2.1)*. For $L > 0$, if the vector $x^* \in S$ satisfies $\|x^*\|_0 = s$, then*

$$L\text{-stationary point} \quad \Longrightarrow \quad \mathrm{N}^{B}\text{-stationary point} \quad \Longleftrightarrow \quad \mathrm{T}^{B}\text{-stationary point};$$

if the vector $x^* \in S$ satisfies $\|x^*\|_0 < s$, then

$$\begin{aligned} L\text{-stationary point} \quad &\Longleftrightarrow \quad \mathrm{N}^{B}\text{-stationary point} \quad \Longleftrightarrow \quad \mathrm{T}^{B}\text{-stationary point} \\ &\Longleftrightarrow \quad \nabla f(x^*) = 0. \end{aligned}$$

Proof Denote $\Gamma^* = \mathrm{supp}(x^*)$. If x^* is an L-stationary point of problem (2.1), then from Lemma 2.2 in [11], it holds for any $L > 0$,

$$x^* \in \mathrm{P}_S(x^* - \nabla f(x^*)/L) \iff \left|(\nabla f(x^*))_i\right| \begin{cases} = 0, & i \in \Gamma^*, \\ \leqslant L M_s(|x^*|), & i \notin \Gamma^*. \end{cases} \tag{2.12}$$

Case 1 First we consider the case $\|x^*\|_0 = s$. Under such circumstance, if x^* is an N^B-stationary point of problem (2.1), then by (2.5) in Theorem 2.1, we have

$$-\nabla f(x^*) \in \mathrm{N}^B_S(x^*) \iff (\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^*, \\ \in \mathbb{R}, & i \notin \Gamma^*. \end{cases} \tag{2.13}$$

Moreover, $\|x^*\|_0 = s$ produces

$$\begin{aligned} \nabla^B_S f(x^*) &= \mathrm{argmin}\{\ \|d + \nabla f(x^*)\| :\ d \in \mathrm{T}^B_S(x^*)\ \} \\ &= \mathrm{argmin}\{\ \|d + \nabla f(x^*)\| :\ \|d\|_0 \leqslant s,\ \|x^* + \mu d\|_0 \leqslant s\ , \forall\, \mu \in \mathbb{R}\ \} \\ &= \mathrm{argmin}\{\ \|d + \nabla f(x^*)\| :\ \mathrm{supp}(d) \subseteq \Gamma^*\ \}, \end{aligned} \tag{2.14}$$

where the third equality holds due to $\|x^*\|_0 = s$. (2.14) is equivalent to

$$(\nabla^B_S f(x^*))_i = \begin{cases} -(\nabla f(x^*))_i, & i \in \Gamma^*, \\ 0, & i \notin \Gamma^*. \end{cases}$$

Therefore, if x^* is a T^B-stationary point of problem (2.1), then from above

$$\nabla_S^B f(x^*) = 0 \iff \big(\nabla f(x^*)\big)_i \begin{cases} = 0, \ i \in \Gamma^*, \\ \in \mathbb{R}, \ i \notin \Gamma^*. \end{cases} \tag{2.15}$$

Henceforth, from (2.12), (2.13) and (2.15), one can easily check that when $\|x^*\|_0 = s$,

$$L\text{-stationary point} \implies \mathrm{N}^B\text{-stationary point} \iff \mathrm{T}^B\text{-stationary point}.$$

Case 2 Now we consider the case $\|x^*\|_0 < s$. Under such circumstance, $M_s(|x^*|) = 0$, and thus if x^* is an L-stationary point of problem (2.1), then from (2.12), it holds

$$x^* \in \mathrm{P}_S(x^* - \nabla f(x^*)/L) \iff \nabla f(x^*) = 0. \tag{2.16}$$

Then when $\|x^*\|_0 < s$, $\mathrm{N}_S^B(x^*) = \{0\}$ from (2.5), which implies $\nabla f(x^*) = 0$. Therefore, if x^* is an N^B-stationary point of problem (2.1), then

$$0 \in \nabla f(x^*) + \mathrm{N}_S^B(x^*) \iff \nabla f(x^*) = 0. \tag{2.17}$$

Finally, we prove $\nabla f(x^*) = 0 \iff \nabla_S^B f(x^*) = 0$ when $\|x^*\|_0 < s$. On the one hand, if $\nabla f(x^*) = 0$, then

$$\begin{aligned} \nabla_S^B f(x^*) &= \operatorname{argmin}\{ \|d + \nabla f(x^*)\| : \|d\|_0 \leqslant s, \ \|x^* + \mu d\|_0 \leqslant s \ , \forall \ \mu \in \mathbb{R} \} \\ &= \operatorname{argmin}\{ \|d\| : \|d\|_0 \leqslant s, \ \|x^* + \mu d\|_0 \leqslant s \ , \forall \ \mu \in \mathbb{R} \} = 0. \end{aligned}$$

On the other hand, if $\nabla_S^B f(x^*) = 0$, then

$$0 = \nabla_S^B f(x^*) = \operatorname{argmin}\{ \|d + \nabla f(x^*)\| : \|d\|_0 \leqslant s, \ \|x^* + \mu d\|_0 \leqslant s \ , \forall \ \mu \in \mathbb{R} \}$$

leads to $\|\nabla f(x^*)\| \leqslant \|d + \nabla f(x^*)\|$ for any $\|d\|_0 \leqslant s, \ \|x^* + \mu d\|_0 \leqslant s \ , \forall \ \mu \in \mathbb{R}$. Particularly, for any $i_0 \in \{1, 2, \cdots, n\}$, we take d with $\mathrm{supp}(d) = \{i_0\}$. Apparently, $\|x^* + \mu d\|_0 \leqslant s \ , \forall \ \mu \in \mathbb{R}$ owing to $\|x^*\|_0 < s$. Then by valuing $d_{i_0} = -(\nabla f(x^*))_{i_0}$ and $d_i = 0, \ i \neq i_0$, we immediately get $(\nabla f(x^*))_{i_0} = 0$ because of $\|\nabla f(x^*)\| \leqslant \|\nabla f(x^*) - (\nabla f(x^*))_{i_0}\|$. Then by the arbitrariness of i_0, it holds $\nabla f(x^*) = 0$. Therefore, if x^* is a T^B-stationary point of problem (2.1), then

$$\nabla_S^B f(x^*) = 0 \iff \nabla f(x^*) = 0. \tag{2.18}$$

Henceforth, from (2.16), (2.17) and (2.18), one can easily check that when $\|x^*\|_0 < s$

$$\begin{aligned} L\text{-stationary point} \iff \mathrm{N}^B\text{-stationary point} &\iff \mathrm{T}^B\text{-stationary point} \\ &\iff \nabla f(x^*) = 0. \end{aligned}$$

Overall, the whole proof is finished.

Based on the proof of Theorem 2.5, we use Table 1 to illustrate the relationship among these three stationary points under the concept of Bouligand tangent cone.

Theorem 2.6 *Under the concept of Clarke tangent cone, we consider problem* (2.1). *For* $L > 0$*, if* $x^* \in S$ *then*

$$L\text{-stationary point} \implies \mathrm{N}^C\text{-stationary point} \iff \mathrm{T}^C\text{-stationary point}.$$

Proof Denote $\Gamma^* = \mathrm{supp}(x^*)$. If x^* is an L-stationary point of problem (2.1), for any $L > 0$, we have (2.12).

If x^* is an N^C-stationary point of problem (2.1), then by (2.8), we have

$$-\nabla f(x^*) \in \mathrm{N}_S^C(x^*) \iff \big(\nabla f(x^*)\big)_i \begin{cases} = 0, & i \in \Gamma^*, \\ \in \mathbb{R}, & i \notin \Gamma^*. \end{cases} \tag{2.19}$$

Moreover, by (2.7), it follows

$$\begin{aligned} \nabla_S^C f(x^*) &= \mathrm{argmin}\{\, \|d + \nabla f(x^*)\| : \ d \in \mathrm{T}_S^C(x^*) \,\} \\ &= \mathrm{argmin}\{\, \|d + \nabla f(x^*)\| : \ \mathrm{supp}(d) \subseteq \Gamma^* \,\}, \end{aligned}$$

which is equivalent to

$$\Big(\nabla_S^C f\big(x^*\big)\Big)_i = \begin{cases} -(\nabla f(x^*))_i, & i \in \Gamma^*, \\ 0, & i \notin \Gamma^*. \end{cases}$$

Therefore, if x^* is a T^C-stationary point of problem (2.1), then from above

$$\nabla_S^C f(x^*) = 0 \iff \big(\nabla f(x^*)\big)_i \begin{cases} = 0, & i \in \Gamma^*, \\ \in \mathbb{R}, & i \notin \Gamma^*. \end{cases} \tag{2.20}$$

Henceforth, from (2.12), (2.19) and (2.20), one can easily check

$$L\text{-stationary point} \implies \mathrm{N}^C\text{-stationary point} \iff \mathrm{T}^C\text{-stationary point}.$$

Overall, the whole proof is finished.

Based on the proof of Theorem 2.6, we use Table 2 to illustrate the relationship among these three stationary points under the concept of Clarke tangent cone.

From Tables 1 and 2, we can see that N^C-stationarity and T^C-stationarity are weaker than N^B-stationarity and T^B-stationarity. While N^B- and T^B-stationarity coincide with basic feasibility property in [11] for problem (2.1), that is $\nabla f(x^*) = 0$ if $\|x^*\|_0 < s$, $\nabla_i f(x^*) = 0$ for $i \in \mathrm{supp}(x^*)$ if $\|x^*\|_0 = s$. The following theorem shows that they are all the necessary optimality conditions for SCO under some assumptions.

Table 1 The relationship of the stationary points for (2.1) under Bouligand tangent cone

	$\|x^*\|_0 = s$	$\|x^*\|_0 < s$
L-stationary point	$\left\|(\nabla f(x^*))_i\right\| \begin{cases} = 0, & i \in \Gamma^* \\ \leqslant LM_s(\|x^*\|) & i \notin \Gamma^* \end{cases}$	$\nabla f(x^*) = 0$
N^B-stationary point	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}, & i \notin \Gamma^* \end{cases}$	$\nabla f(x^*) = 0$
T^B-stationary point	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}, & i \notin \Gamma^* \end{cases}$	$\nabla f(x^*) = 0$

Table 2 The relationship of the stationary points for (2.1) under Clarke tangent cone

	$\|x^*\|_0 = s$	$\|x^*\|_0 < s$
L-stationary point	$\left\|(\nabla f(x^*))_i\right\| \begin{cases} = 0, & i \in \Gamma^* \\ \leqslant LM_s(\|x^*\|), & i \notin \Gamma^* \end{cases}$	$\nabla f(x^*) = 0$
N^C-stationary point	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}, & i \notin \Gamma^* \end{cases}$	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}, & i \notin \Gamma^* \end{cases}$
T^C-stationary point	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}, & i \notin \Gamma^* \end{cases}$	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}, & i \notin \Gamma^* \end{cases}$

Assumption 2.7 The gradient of the objective function $f(x)$ is Lipschitz with constant L_f over $\mathbb{R}^n$:

$$\|\nabla f(x) - \nabla f(y)\| \leqslant L_f \|x - y\|, \quad \forall\, x, y \in \mathbb{R}^n. \tag{2.21}$$

Theorem 2.8 *If x^* be an optimal solution of (2.1), then x^* is an N^B-stationary point and hence N^C-stationary point. Further, if Assumption 2.7 holds and $L > L_f$, then x^* is an L-stationary point of (2.1).*

Proof The first conclusion comes from the equivalence of N^B-stationarity and basic feasibility and Theorem 2.1 in [11]. The second conclusion is also proven in Theorem 2.2 in [11], and here we give a short proof. Suppose to the contrary that there exists a vector

$$y \in \mathrm{P}_S(x^* - \nabla f(x^*)/L).$$

It follows that

$$\|y - x^* + \nabla f(x^*)/L\|^2 \leqslant \|x^* - x^* + \nabla f(x^*)/L\|^2,$$

which yields that

$$\langle \nabla f(x^*), y - x^* \rangle \leqslant -(L/2)\|y - x^*\|^2.$$

By Lipschitz continuity property of ∇f and $L > L_f$, we have

$$\begin{aligned} f(y) &\leqslant f(x^*) + \langle \nabla f(x^*), y - x^* \rangle + (L_f/2)\|y - x^*\|^2 \\ &\leqslant f(x^*) + (L_f/2 - L/2)\|y - x^*\|^2 \\ &< f(x^*), \end{aligned}$$

contradicting the optimality of x^*. This completes the proof.

This theorem indicates that $L > L_f$ is sufficient to guarantee the optimal solution being an L-stationary point. Next example shows that it is not necessary for the case of $\|x^*\|_0 = s$, where x^* denotes a solution to SCO.

Example 2.9 Consider the problem

$$\begin{aligned} &\min \ f(x) = 1.5x_1^2 + 3x_2^2 + 4x_1x_2 - 0.5x_1 - 12x_2 \\ &\text{s.t.} \quad \|x\|_0 \leqslant 1. \end{aligned} \tag{2.22}$$

The gradient of the objective function is Lipschitz with constant $L_f = \lambda_{\max}\begin{pmatrix} 3 & 4 \\ 4 & 6 \end{pmatrix} = 8$. Obviously the optimal solution of the problem is $x^* = (0, 2)^{\mathrm{T}}$. By (2.12), x^* is an L-stationary point if and only if $L \geqslant \frac{|\nabla_1 f(x^*)|}{M_1(|x^*|)} = \frac{15}{4}$. Therefore, the condition $L > L_f$ is sufficient but not necessary to guarantee the optimal solution being an L-stationary point.

2.3 Second-Order Optimality Conditions

In this subsection, we will study the second-order necessary and sufficient optimality conditions of problem (2.1).

Theorem 2.10 (Second-Order Necessary Optimality) *Assume $f(x)$ is twice continuously differentiable on $\mathbb{R}^n$. If $x^* \in S$ is the optimal solution of (2.1), we have*

$$d^{\mathrm{T}}\nabla^2 f(x^*)d \geqslant 0, \quad \forall\, d \in \mathrm{T}_S^C(x^*), \tag{2.23}$$

where $\nabla^2 f(x^*)$ is the Hessian matrix of f at x^*.

Proof If $x^* \in S$ is the optimal solution of (2.1), it is also an N^C-stationary point from Theorem 2.8. From (2.7) and Table 2, one can easily verify that

$$d^{\mathrm{T}}\nabla f(x^*) = 0, \quad \forall\, d \in \mathrm{T}_S^C(x^*).$$

Moreover, for any $\tau > 0$ and $d \in \mathrm{T}_S^C(x^*)$, by the optimality of x^* and the equality above, we have

$$\begin{aligned} 0 &\leqslant f(x^* + \tau d) - f(x^*) \\ &= f(x^*) + \tau d^{\mathrm{T}}\nabla f(x^*) + \frac{\tau^2}{2}d^{\mathrm{T}}\nabla^2 f(x^*)d + o(\|\tau d\|^2) - f(x^*) \\ &= \frac{\tau^2}{2}d^{\mathrm{T}}\nabla^2 f(x^*)d + o(\|\tau d\|^2), \end{aligned}$$

which implies that

$$d^{\mathrm{T}}\nabla^2 f(x^*)d \geqslant 0, \quad \forall\, d \in \mathrm{T}_S^C(x^*).$$

The desired result is acquired.

Theorem 2.11 (Second-Order-Sufficient Optimality) *If* $x^* \in S$ *with* $\Gamma^* = \mathrm{supp}(x^*)$ *is an* N^C*-stationary point of* (2.1) *and* $\nabla^2 f(x^*)$ *is restricted positive definite, that is*

$$d^{\mathrm{T}} \nabla^2 f(x^*) d > 0, \quad \forall\, d \in \mathrm{T}_S^C(x^*), \quad d \neq 0, \tag{2.24}$$

then x^* *is the strictly locally optimal solution of* (2.1) *restricted on the subspace* $\mathbb{R}^n_{\Gamma^*}$*, where* $\mathbb{R}^n_{\Gamma^*} := \mathrm{span}\{e_i,\ i \in \Gamma^*\}$*. Moreover, there are* $\eta > 0$ *and* $\delta > 0$*, for any* $x \in B(x^*, \delta) \cap \mathbb{R}^n_{\Gamma^*}$ *such that*

$$f(x) \geqslant f(x^*) + \eta \|x - x^*\|^2. \tag{2.25}$$

Proof We only prove the second conclusion because it can yield the first one. From Table 2, one can easily check

$$d^{\mathrm{T}} \nabla f(x^*) = 0, \ \ \forall\, d \in \mathrm{T}_S^C(x^*). \tag{2.26}$$

Suppose the conclusion does not hold. There must be a sequence $\{x^k\} \in B(x^*, \delta) \cap \mathbb{R}^n_{\Gamma^*}$ with

$$\lim_{k \to \infty} x^k = x^* \quad \text{and} \quad \mathrm{supp}(x^k) = \mathrm{supp}(x^*)$$

such that

$$f(x^k) - f(x^*) \leqslant \frac{1}{k} \|x^k - x^*\|^2. \tag{2.27}$$

Denote $d^k = \frac{x^k - x^*}{\|x^k - x^*\|}$. Due to $\|d^k\| = 1$, there exists a convergent subsequence, without loss of generality, assuming $d^k \to \bar{d}$. Then $d^k \in \mathbb{R}^n_{\Gamma^*}$ and $\bar{d} \in \mathbb{R}^n_{\Gamma^*}$ by $\mathrm{supp}(x^k) = \mathrm{supp}(x^*)$. From (2.26) and $\mathrm{T}_S^C(x^*) = \mathbb{R}^n_{\Gamma^*}$, we get $d^{k\mathrm{T}} \nabla f(x^*) = 0$. It follows from (2.27) that

$$\begin{aligned}
\frac{1}{k} &\geqslant \frac{1}{\|x^k - x^*\|^2} \left(f\left(x^k\right) - f\left(x^*\right) \right) \\
&= \frac{1}{\|x^k - x^*\|^2} \Bigg(\left(x^k - x^*\right)^{\mathrm{T}} \nabla f(x^*) + \frac{1}{2} \left(x^k - x^*\right)^{\mathrm{T}} \\
&\quad \times \nabla^2 f\left(x^*\right) \left(x^k - x^*\right) + o\left(\|x^k - x^*\|^2\right) \Bigg) \\
&= \frac{1}{2} d^{k\mathrm{T}} \nabla^2 f(x^*) d^k + \frac{1}{\|x^k - x^*\|} d^{k\mathrm{T}} \nabla f\left(x^*\right) + o(1) \\
&= \frac{1}{2} d^{k\mathrm{T}} \nabla^2 f\left(x^*\right) d^k + o(1).
\end{aligned}$$

Then by taking the limit of both sides of the above inequality, we obtain

$$0 = \lim_{k \to \infty} \frac{1}{k} \geqslant \lim_{k \to \infty} \left(\frac{1}{2} d^{k\mathrm{T}} \nabla^2 f(x^*) d^k + o(1) \right) = \frac{1}{2} \bar{d}^{\mathrm{T}} \nabla^2 f(x^*) \bar{d} > 0, \quad \bar{d} \in \mathbb{R}^n_{\Gamma^*},$$

which is contradicted. Therefore the conclusion does hold.

We next give an example to show that for problem (2.1), under the condition (2.24), the N^C-stationary point may not be the local minimizer of f on S.

Example 2.12 Consider the problem

$$\begin{aligned} \min \quad & f(x) = \frac{1}{2}\left((x_1+1)^2 + (x_2-1)^2 + (x_3-1)^2\right) \\ \text{s.t.} \quad & \|x\|_0 \leqslant 2. \end{aligned}$$

The gradient and Hessian matrix of f are $\nabla f(x) = (x_1+1, x_2-1, x_3-1)^{\mathrm{T}}$ and $\nabla^2 f(x) = I$ respectively. From Table 2, $\bar{x} = (0,0,1)^{\mathrm{T}}$ with $\nabla f(\bar{x}) = (1,-1,0)^{\mathrm{T}}$ is an N^C-stationary point of the above problem, but not the local minimizer because $f((0,\varepsilon,1)^{\mathrm{T}}) < f(\bar{x}), 0 < \varepsilon \leqslant 1$.

In the end of this section, we will give more details about the special case of $f(x) \equiv \frac{1}{2}\|Ax-b\|^2$, where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $L_f = \lambda_{\max}(A^{\mathrm{T}}A)$ is the largest eigenvalue of $A^{\mathrm{T}}A$. From Theorem 2.11, we easily derive the following result.

Corollary 2.13 *Let* $f(x) \equiv \frac{1}{2}\|Ax-b\|^2$. *If* $x^* \in S$ *with* $\Gamma^* = \mathrm{supp}(x^*)$ *is an* N^C*-stationary point of* (2.1) *and*

$$d^{\mathrm{T}}A^{\mathrm{T}}Ad > 0, \quad \forall\, d \in \mathrm{T}_S^C(x^*), \quad d \neq 0, \tag{2.28}$$

then x^* *is the globally optimal solution of* (2.1) *restricted on* $\mathbb{R}^n_{\Gamma^*}$.

Note that condition (2.28) is weaker than s-regularity of matrix A in [11], which means that every s column of A is linearly independent. Using the s-regularity of matrix A, we can obtain the global uniqueness of solution of problem (2.1).

Corollary 2.14 *Let* $f(x) \equiv \frac{1}{2}\|Ax-b\|^2$. *If matrix* A *is* s*-regular, the number of* N^C*-stationary points is finite, and any* N^C*-stationary point, say* $x^* \in S$ *with* $\Gamma^* = \mathrm{supp}(x^*)$, *is uniquely global minimizer of problem* (2.1) *on* $\mathbb{R}^n_{\Gamma^*}$.

Proof Since x^* is an N^C-stationary point of (2.1) and $\Gamma^* = \mathrm{supp}(x^*)$, we denote $x^* = (x^{*\mathrm{T}}_{\Gamma^*}, 0)^{\mathrm{T}}$ and let A_{Γ^*} be the submatrix of A made up of the columns corresponding to the set Γ^*. By Tables 1 and 2, we have $0 = (\nabla f(x))_i = (A^{\mathrm{T}}(Ax-b))_i = 0, \forall\, i \in \Gamma^*$, namely,

$$A^{\mathrm{T}}_{\Gamma^*}A_{\Gamma^*}x^*_{\Gamma^*} = A^{\mathrm{T}}_{\Gamma^*}b.$$

The s-regularity of A yields that $x^*_{\Gamma^*} = (A^{\mathrm{T}}_{\Gamma^*}A_{\Gamma^*})^{-1}A^{\mathrm{T}}_{\Gamma^*}b$ is unique. Since the number of subsets Γ^* of $\{1,2,\cdots,n\}$ with $|\Gamma^*| \leqslant s$ is finite, the number of N^C-stationary points is finite.

The s-regularity of A implies that $f(x)$ is strongly convex on subspace $\mathbb{R}^n_{\Gamma^*}$. By Corollary 2.13, every stationary point is uniquely global minimizer of (2.1) on subspace $\mathbb{R}^n_{\Gamma^*}$. This completes the proof.

Furthermore, we obtain the result on globally unique solution to SCO with $f(x) \equiv \frac{1}{2}\|Ax - b\|^2$.

Theorem 2.15 *Let $f(x) \equiv \frac{1}{2}\|Ax - b\|^2$. If matrix A is s-regular, and both A and b guarantee a unique solution to*

$$\Gamma_0 \triangleq \underset{|\Gamma| \leqslant s}{\operatorname{argmin}} \ \|\Pi_\Gamma b\|, \tag{2.29}$$

where $\Pi_\Gamma b = A_\Gamma (A_\Gamma^{\mathrm{T}} A_\Gamma)^{-1} A_\Gamma^{\mathrm{T}} b$, then problem (2.1) has a unique solution.

Proof We will prove it by contradiction. Let x and y be two different solutions of (2.1) with $\Gamma_1 = \operatorname{supp}(x)$ and $\Gamma_2 = \operatorname{supp}(y)$. Denote $x = (x_{\Gamma_1}^{\mathrm{T}},\ 0)^{\mathrm{T}}$ and $y = (y_{\Gamma_2}^{\mathrm{T}},\ 0)^{\mathrm{T}}$. We easily verify that (2.1) is equivalent to

$$\min_{|\Gamma| \leqslant s} \min_{x_\Gamma} \|A_\Gamma x_\Gamma - b\|^2.$$

Since A is s-regular, we have $\Gamma_1 \neq \Gamma_2$ provided $x \neq y$. It is easy to see that $x_{\Gamma_1} = (A_{\Gamma_1}^{\mathrm{T}} A_{\Gamma_1})^{-1} A_{\Gamma_1}^{\mathrm{T}} b$, $y_{\Gamma_2} = (A_{\Gamma_2}^{\mathrm{T}} A_{\Gamma_2})^{-1} A_{\Gamma_2}^{\mathrm{T}} b$. From $\|Ax - b\|^2 = \|Ay - b\|^2$, we have

$$\|A_{\Gamma_1}(A_{\Gamma_1}^{\mathrm{T}} A_{\Gamma_1})^{-1} A_{\Gamma_1}^{\mathrm{T}} b - b\|^2 = \|A_{\Gamma_2}(A_{\Gamma_2}^{\mathrm{T}} A_{\Gamma_2})^{-1} A_{\Gamma_2}^{\mathrm{T}} b - b\|^2,$$

that is

$$\|\Pi_{\Gamma_1} b - b\|^2 = \|\Pi_{\Gamma_2} b - b\|^2.$$

Because $\|Ax - b\|^2 = \|Ay - b\|^2$ is the optimal value of (2.29) and $\Pi_\Gamma^{\mathrm{T}} = \Pi_\Gamma$ and $\Pi_\Gamma^2 = \Pi_\Gamma$, the above equation means

$$\|\Pi_{\Gamma_1} b\| = \|\Pi_{\Gamma_2} b\| = \min_{|\Gamma| \leqslant s} \ \|\Pi_\Gamma b\|.$$

Hence it violates the assumption.

3 Extensions

In this section, we mainly aim at specifying the results in Sect. 2 for SCO with nonnegative constraint:

$$\min \quad f(x), \qquad \text{s.t.} \ \ \|x\|_0 \leqslant s, \quad x \geqslant 0. \tag{3.1}$$

First, we give the explicit expression of the projection on $S \cap \mathbb{R}_+^n$, which is named the nonnegative support projection.

Proposition 3.1 $\mathrm{P}_{S \cap \mathbb{R}_+^n}(x) = \mathrm{P}_S \cdot \mathrm{P}_{\mathbb{R}_+^n}(x)$.

Proof Denote $I_+(x) = \{i : x_i > 0\}$, $I_0(x) = \{i : x_i = 0\}$, $I_-(x) = \{i : x_i < 0\}$, and let $y \in \mathrm{P}_{S\cap\mathbb{R}^n_+}(x)$. For $i \in I_0(x) \cup I_-(x)$, it is easy to see $y_i = 0$. There are two cases: Case 1, $|I_+(x)| \leqslant s$, then $y = \mathrm{P}_{\mathbb{R}^n_+}(x) = \mathrm{P}_S \cdot \mathrm{P}_{\mathbb{R}^n_+}(x)$. Case 2, $|I_+(x)| > s$, we should choose no more than s coordinates from $I_+(x)$ to minimize $\|x - y\|$. For $i, j \in I_+(x)$ and $x_i > x_j$,

$$\begin{aligned}(x_i - x_i)^2 + (x_j - x_j)^2 &< (x_i - x_i)^2 + (x_j - 0)^2 \\ &< (x_i - 0)^2 + (x_j - x_j)^2 < (x_i - 0)^2 + (x_j - 0)^2.\end{aligned}$$

Then the projection on $S \cap \mathbb{R}^n_+$ sets all but s largest elements of $\mathrm{P}_{\mathbb{R}^n_+}(x)$ to zero, which is $\mathrm{P}_S \cdot \mathrm{P}_{\mathbb{R}^n_+}(x)$.

Notice that the order of the nonnegative support projection on $S \cap \mathbb{R}^n_+$ can not be changed. For example $x = (-2, 1)^{\mathrm{T}}$, $s = 1$. $\mathrm{P}_{S\cap\mathbb{R}^2_+}(x) = \mathrm{P}_S \cdot \mathrm{P}_{\mathbb{R}^2_+}(x) = (0, 1)^{\mathrm{T}}$, while $\mathrm{P}_{\mathbb{R}^2_+} \cdot \mathrm{P}_S(x) = (0, 0)^{\mathrm{T}}$.

The direct result of Theorems 2.1 and 2.2 is the following theorem.

Theorem 3.2 *For any* $\bar{x} \in S \cap \mathbb{R}^n_+$, *by denoting* $\mathbb{R}^n_+(\bar{x}) := \{\, d \in \mathbb{R}^n : d_i \geqslant 0, i \notin \mathrm{supp}(\bar{x}) \,\}$, *it follows*

$$\mathrm{T}^B_{S\cap\mathbb{R}^n_+}(\bar{x}) = \mathrm{T}^B_S(\bar{x}) \cap \mathbb{R}^n_+(\bar{x}), \quad \mathrm{N}^B_{S\cap\mathbb{R}^n_+}(\bar{x}) = \mathrm{N}^B_S(\bar{x}) \cup (-\mathbb{R}^n_+(\bar{x})), \tag{3.2}$$

$$\mathrm{T}^C_{S\cap\mathbb{R}^n_+}(\bar{x}) = \mathrm{T}^C_S(\bar{x}), \quad \mathrm{N}^C_{S\cap\mathbb{R}^n_+}(\bar{x}) = \mathrm{N}^C_S(\bar{x}). \tag{3.3}$$

Clearly, one can check that

$$\begin{aligned}&\mathrm{T}^B_{S\cap\mathbb{R}^n_+}(\bar{x}) = \mathrm{T}^B_S(\bar{x}), \quad \mathrm{N}^B_{S\cap\mathbb{R}^n_+}(\bar{x}) = \mathrm{N}^B_S(\bar{x}), && \text{if } \|\bar{x}\|_0 = s \text{ and } \bar{x} \geqslant 0,\\ &\mathrm{T}^B_{S\cap\mathbb{R}^n_+}(\bar{x}) = \mathrm{T}^B_S(\bar{x}) \cap \mathbb{R}^n_+(\bar{x}), \quad \mathrm{N}^B_{S\cap\mathbb{R}^n_+}(\bar{x}) = -\mathbb{R}^n_+(\bar{x}), && \text{if } \|\bar{x}\|_0 < s \text{ and } \bar{x} \geqslant 0.\end{aligned}$$

For problem (3.1), the corresponding definitions of L-stationary point, $\mathrm{N}^\sharp$-stationary point and $\mathrm{T}^\sharp$-stationary point can be obtained from Definition 2.3 in [11] and Definition 2.4, just by substituting $S \cap \mathbb{R}^n_+$ for S.

$$L\text{-stationary point}: \quad x^* \in \mathrm{P}_{S\cap\mathbb{R}^n_+}(x^* - \nabla f(x^*)/L), \tag{3.4}$$

$$\mathrm{N}^\sharp\text{-stationary point}: \quad 0 \in \nabla f(x^*) + \mathrm{N}^\sharp_{S\cap\mathbb{R}^n_+}(x^*), \tag{3.5}$$

$$\mathrm{T}^\sharp\text{-stationary point}: \quad 0 = \|\nabla^\sharp_{S\cap\mathbb{R}^n_+} f(x^*)\|. \tag{3.6}$$

In order to facilitate the discussion next, we describe a more explicit representation of L-stationary point for (3.1).

Theorem 3.3 *For* $L > 0$, *a vector* $x^* \in S \cap \mathbb{R}^n_+$ *is* L-*stationary point of* (3.1) *if and only if*

$$(\nabla f(x^*))_i \begin{cases} = 0, & \text{if } i \in \mathrm{supp}(x^*), \\ \in [-LM_s(x^*), +\infty), & \text{if } i \notin \mathrm{supp}(x^*). \end{cases} \tag{3.7}$$

Proof Suppose $x^* \in S \cap \mathbb{R}^n_+$ is an L-stationary point, namely, satisfying (3.4). For simplicity, we write $\nabla_i f(x^*) = (\nabla f(x^*))_i$. Note that the component of $\mathrm{P}_{S\cap\mathbb{R}^n_+}(x^* - \nabla f(x^*)/L)$ is either zero or itself. If $i \in \mathrm{supp}(x^*)$, then $x_i^* = x_i^* - \nabla_i f(x^*)/L$, so that $\nabla_i f(x^*) = 0$; If $i \notin \mathrm{supp}(x^*)$, there are two cases: either $x_i^* - \nabla_i f(x^*)/L \leqslant 0$, that is $\nabla_i f(x^*) \geqslant x_i^* = 0$, or $0 \leqslant x_i^* - \nabla_i f(x^*)/L \leqslant M_s(x^*)$, that is $-LM_s(x^*) \leqslant \nabla_i f(x^*) \leqslant 0$.

On the contrary, assume (3.7) holds. If $\|x^*\|_0 < s$, we get $M_s(x^*) = 0$, then for $i \in \mathrm{supp}(x^*)$, $\nabla_i f(x^*) = 0$, then $x^* - \nabla_i f(x^*)/L = x_i^*$ or for $i \notin \mathrm{supp}(x^*)$, $x_i^* - \nabla_i f(x^*)/L \leqslant 0$, therefore, (3.4) holds. If $\|x^*\|_0 = s$, that is $M_s(x^*) > 0$. By (3.7), for $i \in \mathrm{supp}(x^*)$, $\nabla_i f(x^*)/L = 0$; for $i \notin \mathrm{supp}(x^*)$, $x^* - \nabla_i f(x^*)/L \leqslant 0$ or $0 \leqslant x_i^* - \nabla_i f(x^*)/L \leqslant M_s(x^*)$, so that (3.4) holds as well.

The following theorem is derived by Theorems 2.5, 2.6 and 3.3 directly.

Theorem 3.4 *For the problem* (3.1) *and* $L > 0$, *the following assertions hold.*

(i) *Under the concept of Bouligand tangent cone, if* $\|x^*\|_0 = s, x^* \geqslant 0$, *then*

$$L\text{-stationary point} \implies \mathrm{N}^B\text{-stationary point} \iff \mathrm{T}^B\text{-stationary point}.$$

If $\|x^*\|_0 < s, x^* \geqslant 0$, *then*

$$L\text{-stationary point} \iff \mathrm{N}^B\text{-stationary point} \iff \mathrm{T}^B\text{-stationary point}.$$

(ii) *Under the concept of Clarke tangent cone, if* $\|x^*\|_0 \leqslant s, x^* \geqslant 0$, *then*

$$L\text{-stationary point} \implies \mathrm{N}^C\text{-stationary point} \iff \mathrm{T}^C\text{-stationary point}.$$

Combining Theorems 3.3 and 3.4, we have the following table to illustrate the relationship among the three stationary points for (3.1) under the concepts of Bouligand and Clarke tangent cones (Table 3).

Combining Theorems 2.10 and 2.11, we derive the following second-order optimality result.

Table 3 The relationship of the stationary points for (3.1) under Bouligand and Clarke tangent cones

	$\|x^*\|_0 = s, x^* \geqslant 0$	$\|x^*\|_0 < s, x^* \geqslant 0$
L-stationary point	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \geqslant -LM_s(x^*), & i \notin \Gamma^* \end{cases}$	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}_+, & i \notin \Gamma^* \end{cases}$
N^B-stationary point	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}, & i \notin \Gamma^* \end{cases}$	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}_+, & i \notin \Gamma^* \end{cases}$
N^C-stationary point	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}, & i \notin \Gamma^* \end{cases}$	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}, & i \notin \Gamma^* \end{cases}$
T^B-stationary point	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}, & i \notin \Gamma^* \end{cases}$	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}_+, & i \notin \Gamma^* \end{cases}$
T^C-stationary point	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}, & i \notin \Gamma^* \end{cases}$	$(\nabla f(x^*))_i \begin{cases} = 0, & i \in \Gamma^* \\ \in \mathbb{R}, & i \notin \Gamma^* \end{cases}$

Theorem 3.5 (Second-Order Optimality) *Let $f(x)$ be twice differentiable. If $x^* \in S \cap \mathbb{R}^n_+$ with $\Gamma^* = \mathrm{supp}(x^*)$ is the optimal solution of* (3.1)*, then* (i) *x^* is an* N^B-*stationary point of* (3.1) *and hence an* N^C-*stationary point;* (ii) *x^* is also an L-stationary point of* (3.1) *provided that Assumption* 2.7 *holds and $L > L_f$; and* (iii)

$$d^{\mathrm{T}}\nabla^2 f(x^*)d \geqslant 0, \quad \forall\, d \in \mathrm{T}^C_{S\cap\mathbb{R}^n_+}(x^*). \tag{3.8}$$

Conversely, if $x^* \in S \cap \mathbb{R}^n_+$ is an N^C-*stationary point of* (3.1) *and satisfies*

$$d^{\mathrm{T}}\nabla^2 f(x^*)d > 0, \quad \forall\, d \in \mathrm{T}^C_{S\cap\mathbb{R}^n_+}(x^*), \quad d \neq 0, \tag{3.9}$$

then x^* is the strictly locally optimal solution of (3.1) *restricted on $\mathbb{R}^n_{\Gamma^*} \cap \mathbb{R}^n_+$. Moreover, there are $\gamma > 0$ and $\delta > 0$, for any $x \in B(x^*, \delta) \cap \mathbb{R}^n_{\Gamma^*} \cap \mathbb{R}^n_+$, it holds*

$$f(x) \geqslant f(x^*) + \gamma \|x - x^*\|^2. \tag{3.10}$$

In the same way, Corollaries 2.13 and 2.14 and Theorem 2.15 can be extended to the problem of (3.1) with the case of $f(x) \equiv \frac{1}{2}\|Ax - b\|^2$.

Theorem 3.6 *Let $f(x) \equiv \frac{1}{2}\|Ax - b\|^2$. We have the following conclusions.*

(i) If $x^* \in S$ with $\Gamma^* = \mathrm{supp}(x^*)$ is an N^C- *stationary point of* (3.1) *and $d^{\mathrm{T}}A^{\mathrm{T}}Ad > 0$, $\forall\, d \in \mathrm{T}^C_{S\cap\mathbb{R}^n_+}(x^*)$, $d \neq 0$, then x^* is the strictly globally optimal solution of* (3.1) *on $\mathbb{R}^n_{\Gamma^*} \cap \mathbb{R}^n_+$.*
(ii) If matrix A is s-regular, then the number of N^C-*stationary points of* (3.1) *is finite, and every* N^C-*stationary point, say x^* with $\Gamma^* = \mathrm{supp}(x^*)$, is uniquely global minimizer of problem* (3.1) *on $\mathbb{R}^n_{\Gamma^*} \cap \mathbb{R}^n_+$. Furthermore, if both A and b guarantee a unique solution of $\Gamma_0 \triangleq \mathrm{argmin}_{|\Gamma|\leqslant s} \|\Pi_\Gamma b\|$, where $\Pi_\Gamma b = A_\Gamma(A_\Gamma^{\mathrm{T}}A_\Gamma)^{-1}A_\Gamma^{\mathrm{T}}b$, then problem* (3.1) *has a unique solution.*

4 Conclusions

In this paper, we have established the first- and second-order optimality conditions for sparsity constrained optimization with the help of tangent cone and normal cone. Furthermore, we extended the results to sparsity constrained optimization with non-negative constraint. However, there are many practical problems with both sparsity constraint and other additional constraints, such as [16–19]. In the future, we will use the tangent cone and normal cone to consider the optimality conditions for this kind of problems and derive algorithms using the conditions.

This paper is based on the technical report [20] uploaded on http://arxiv.org/abs/1406.7178 on 27th Jun 2014. Over the past year, some results on optimality conditions for sparsity constrained optimization were found. Bauschke et al. [21] gave the expressions of proximal and Mordukhovich normal cones and Bouligand tangent cone to sparse set S. While in [20], we showed explicitly the Bouligand and Clarke tangent cones and normal cones to sparse set S by the definitions. Lu and Zhang [22] gave the first-order necessary optimality condition of a more general form of sparsity constrained optimization. If there is only the sparsity constraint, this condition turns out to be an N^M-stability in the sense of Mordukhovich normal cone. For $x^* \in S$ satisfying $\|x^*\|_0 = s$, N^M-stability is equivalent to N^B- and N^C- stability for sparsity constrained optimization. Beck and Hallak extended the results in [11] to the optimization problem over sparse symmetric sets and gave three types of optimality conditions and the hierarchy between the optimality conditions using the orthogonal projection operator. And then they presented and analyzed algorithms satisfying the various optimality conditions.

Acknowledgments The authors are grateful to Dr. Cai-Hua Chen in Nanjing University for his helpful advice and to the anonymous referees who have contributed to improve the quality of the paper.

References

[1] Natarajan, B.K.: Sparse approximate solutions to linear systems. SIAM J. Comput. **24**, 227–234 (1995)
[2] Negahban, S., Ravikumar, P., Wainwright, M., Yu, B.: A unified framework for highdimensional analysis of m-estimators with decomposable regularizers. Stat. Sci. **27**, 538–557 (2012)
[3] Agarwal, A., Negahban, S., Wainwright, M.: Fast global convergence rates of gradient methods for high-dimensional statistical recovery. Adv. Neural Inf. Process. Syst. **23**, 37–45 (2010)
[4] Blumensath, T.: Compressed sensing with nonlinear observations and related nonlinear optimisation problems. IEEE Trans. Inf. Theory **59**, 3466–3474 (2013)
[5] Jalali, A., Johnson, C.C., Ravikumar, P.K.: On learning discrete graphical models using greedy methods. Adv. Neural Inf. Process. Syst. **24**, 1935–1943 (2011)
[6] Tewari, A., Ravikumar, P.K., Dhillon, I.S.: Greedy algorithms for structurally constrained high dimensional problems. Adv. Neural Inf. Process. Syst. **24**, 882–890 (2011)
[7] Yuan, X., Li, P., Zhang, T.: Gradient hard thresholding pursuit for sparsity-constrained optimization. ICML (2014)
[8] Bahmani, S., Raj, B., Boufounos, P.: Greedy sparsity-constrained optimization. J. Mach. Learn. Res. **14**, 807–841 (2013)
[9] Candès, E.J., Tao, T.: Decoding by linear programming. IEEE Trans. Inf. Theory **51**, 4203–4215 (2005)
[10] Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theory **52**, 1289–1306 (2006)
[11] Beck, A., Eldar, Y.: Sparsity constrained nonlinear optimization: optimality conditions and algorithms. SIAM J. Optim. **23**, 1480–1509 (2013)
[12] Rockafellar, R.T., Wets, R.J.: Variational Analysis. Springer, Berlin (1998)
[13] Attouch, H., Bolte, J., Svaiter, B.F.: Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. Math. Program. **137**, 91–129 (2013)
[14] Calamai, P.H., More, J.J.: Projection gradient methods for linearly constrained problems. J. Math. Program. **39**, 93–116 (1987)
[15] Blumensath, T., Davies, M.E.: Iterative thresholding for sparse approximations. J. Fourier Anal. Appl. **14**, 626–654 (2008)
[16] Blumensath, T.: Non-linear compressed sensing and its application to beam hardening correction in X-ray tomography. In: Proceedings of Inverse Problems-From Theory to Application. Bristol (2014)
[17] Chen, A.I., Graves, S.C.: Sparsity-constrained transportation problem. arXiv:1402.2309 (2014)
[18] Smith, N.A., Tromble, R.W.: Sampling uniformly from the unit simplex. Technical Report, Johns Hopkins University, 1–6 (2004)
[19] Takeda, A., Niranjan, M., Gotoh, J., Kawahara, Y.: Simultaneous pursuit of out-of-sample performance and sparsity in index tracking portfolios. Comput. Manag. Sci. **10**, 21–49 (2012)
[20] Pan, L., Xiu, N., Zhou, S.: Gradient support projection algorithm for affine feasibility problem with sparsity and nonnegativity. http://arxiv-web3.library.cornell.edu/pdf/1406.7178v1

[21] Bauschke, H.H., Luke, D.R., Phan, H.M., Wang, X.: Restricted normal cones and sparsity optimization with affine constraints. Found. Comput. Math. **14**, 63–83 (2014)
[22] Lu, Z., Zhang, Y.: Sparse approximation via penalty decomposition methods. SIAM J. Optim. **23**, 2448–2478 (2013)