Optim Lett (2015) 9:1433–1445
DOI 10.1007/s11590-015-0900-9

ORIGINAL PAPER

Characterizations of solution sets of cone-constrained convex programming problems

Xin-He Miao[1] · **Jein-Shan Chen**[2]

Received: 18 September 2013 / Accepted: 1 May 2015 / Published online: 12 May 2015

Abstract In this paper, we consider a type of cone-constrained convex program in finite-dimensional space, and are interested in characterization of the solution set of this convex program with the help of the Lagrange multiplier. We establish necessary conditions for a feasible point being an optimal solution. Moreover, some necessary conditions and sufficient conditions are established which simplifies the corresponding results in Jeyakumar et al. (J Optim Theory Appl 123(1), 83–103, 2004). In particular, when the cone reduces to three specific cones, that is, the p-order cone, L^p cone and circular cone, we show that the obtained results can be achieved by easier ways by exploiting the special structure of those three cones.

Keywords Convex programs · Lagrange multipliers · $\mathcal{K}$-Convex mapping · Normal cone · KKT conditions

The author's work is supported by Ministry of Science and Technology, Taiwan.

X.-H. Miao is supported by National Young Natural Science Foundation (No. 11101302) and National Natural Science Foundation of China (No. 11471241).

✉ Jein-Shan Chen
jschen@math.ntnu.edu.tw

Xin-He Miao
xinhemiao@tju.edu.cn

1 Department of Mathematics, School of Science, Tianjin University, Tianjin 300072, People's Republic of China

2 Department of Mathematics, National Taiwan Normal University, Taipei 11677, Taiwan

1 Introduction

Consider the cone-constrained convex programming problem as follows:

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & Ax = b, \\ & -g(x) \in \mathcal{K}, \end{array} \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $\mathcal{K}$ is a closed convex cone in $\mathbb{R}^r$, $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function, and $g : \mathbb{R}^n \to \mathbb{R}^r$ is a continuous $\mathcal{K}$-convex mapping which means for every $x, y \in \mathbb{R}^n$ and each $t \in [0, 1]$,

$$tg(x) + (1-t)g(y) - g(tx + (1-t)y) \in \mathcal{K}.$$

One important issue for such optimization problem is to characterize the solution set which is also a fundamental topic for many mathematical programming problems. With the help of the characterization of the solution set, we will have a deeper understanding for several important optimization problems including bi-level programming, goal programming and multiple objective programming, and so on. Moreover, it is also essential for understanding the behavior of solution methods for solving mathematical programming problems, see [3,8,13,14]. This is the main motivation to investigate characterizations of the solution set of optimization problems.

In [14], Mangasarian provides a characterization of the solution set of a convex programming problem with differentiable functions. Subsequently, Burke and Ferris [3] present another more specific characterization for the solution set. Recently, characterizations of the solution set of problem (1) where $g = 0$ and f is pseudolinear have been presented in [13]. Jeyakumar et al. [11] describe characterizations of the solution set of a general cone-constrained convex programming problem. Wu and Wu [16] characterize the solution set of a general convex program on a normed vector space.

For the problem (1), the purpose of this paper is to characterize its solution set (see Theorem 3.3) which simplifies the conclusions in [11]. Moreover, when $\mathcal{K}$ reduces to p-order cone, L^p cone or circular cone, the obtained characterizations can be reached by other ways via exploiting the special structures of these three specific cones.

Finally, we say a few words about notations which will be used in this paper. Let $\mathbb{R}$ denote the space of real numbers, $\mathbb{R}_+(\mathbb{R}_{++})$ denote the set consisting of the nonnegative (positive) reals, and $\mathbb{R}^n$ mean the n-dimensional real vector space. For the set $\mathcal{K} \subseteq \mathbb{R}^n$, int $\mathcal{K}$ denotes the interior of the set $\mathcal{K}$ and $\partial\mathcal{K}$ denotes the boundary of $\mathcal{K}$. Moreover, we write $\mathbb{B}(x, \varepsilon)$ to mean the open sphere with center $x \in \mathbb{R}^n$ and radius $\varepsilon > 0$. For the function $f : \mathbb{R}^n \to \mathbb{R}$, the convex subdifferential of the function f at $x \in \mathbb{R}^n$ is denoted by $\partial f(x)$. We denote by $\|x\|$ the 2-norm of x which induced by the inner product $\langle\cdot,\cdot\rangle$, i.e., $\|x\| = \sqrt{\langle x, x\rangle}$, where $\langle x, y\rangle$ means the inner product of x and y. We use $\|x\|_p$ to mean the p-norm of x with $1 \le p < \infty$ which is defined as $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ for any $x := (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$.

2 Preliminaries

In this section, we briefly review some basic concepts and background materials about three specific closed convex cones, which will be extensively used in subsequent analysis. More details can be found in [6,9,10,17].

For problem (1), let F and S be the feasible set and the solution set, respectively, that is,

$$F := \left\{x \in \mathbb{R}^n \mid Ax = b,\ -g(x) \in \mathcal{K}\right\}$$

and

$$S := \{x \in F \mid f(x) \leq f(y),\ \forall y \in F\}.$$

The subdifferential of the function f at x is defined as

$$\partial f(x) := \left\{\xi \in \mathbb{R}^n \mid f(y) - f(x) \geq \langle \xi, y - x\rangle,\ \forall y \in \mathbb{R}^n\right\}.$$

If C is a convex set, the normal cone $\mathcal{N}_C(x)$ of C at $x \in C$ is defined by

$$\mathcal{N}_C(x) := \left\{\xi \in \mathbb{R}^n \mid \langle \xi, y - x\rangle \leq 0,\ \forall y \in C\right\}.$$

It is well known that the subdifferential of the indicator function associated with the convex set C at $x \in C$ is the normal cone $\mathcal{N}_C(x)$. Moreover, if the convex set C is the special convex set $C = \{x \in \mathbb{R}^n \mid Ax = b\}$, it is easy to verify that, for any $x \in C$, the normal cone $\mathcal{N}_C(x)$ of C at x is

$$\mathcal{N}_C(x) = \left\{A^T y \mid y \in \mathbb{R}^m\right\}.$$

In other words, the normal cone $\mathcal{N}_C(x)$ is the range space of A^T.

From the convexity of the function f, we know that the function f is continuous. Since g is a continuous $\mathcal{K}$-convex mapping again, it follows that the problem (1) is a convex optimization problem. If problem (1) satisfies the Slater condition [12], that is, there exists $\bar{x} \in \mathbb{R}^n$ such that $A\bar{x} = b$ and $-g(\bar{x}) \in \operatorname{int} \mathcal{K}$, it is known that $a \in S$ if and only if the element a satisfies the KKT conditions, i.e., $a \in F$ and there exists a Lagrange multiplier $\lambda_a \in \mathbb{R}^r$ such that

$$0 \in \partial f(a) + \partial\left(\lambda_a^T g\right)(a) + \left\{A^T y \mid y \in \mathbb{R}^m\right\},\quad \lambda_a \in \mathcal{K}^* \quad \text{and} \quad \lambda_a^T g(a) = 0,$$

where $\mathcal{K}^*$ denotes the dual cone of $\mathcal{K}$ given by

$$\mathcal{K}^* = \left\{z \in \mathbb{R}^r \mid \langle z, x\rangle \geq 0,\ \forall x \in \mathcal{K}\right\}.$$

For problem (1), we shall assume throughout that the solution set S is nonempty. Let $a \in S$. By above analysis, there exists the corresponding Lagrange multiplier λ_a

such that (a, λ_a) satisfying the KKT conditions. More specifically, we consider the Lagrange function $L_a(\cdot, \lambda_a) : \mathbb{R}^n \to \mathbb{R}$ defined by

$$L_a(x, \lambda_a) := f(x) + \lambda_a^T g(x) \quad \text{for all } x \in \mathbb{R}^n.$$

From f being convex and g being $\mathcal{K}$-convex, it follows that for all $x, y \in \mathbb{R}^n$ and each $\beta \in [0, 1]$,

$$\begin{aligned} L_a(\beta x + (1-\beta)y, \lambda_a) &= f(\beta x + (1-\beta)y) + \lambda_a^T g(\beta x + (1-\beta)y) \\ &\le \beta f(x) + (1-\beta) f(y) + \beta \lambda_a^T g(x) + (1-\beta)\lambda_a^T g(y) \\ &= \beta L_a(x, \lambda_a) + (1-\beta) L_a(y, \lambda_a). \end{aligned}$$

This demonstrates that the function $L_a(\cdot, \lambda_a)$ is also a convex function.

Next, we review the concepts of three specific closed convex cones and their dual cones.

(1) p**-order cone**, see [1]. It is a generalization of the second-order cone [4,5,15] and expressed as follows:

$$\mathcal{K}_p := \left\{ x \in \mathbb{R}^n \,\middle|\, x_1 \ge \left(\sum_{i=2}^n |x_i|^p \right)^{\frac{1}{p}} \right\}, \quad (1 < p < \infty).$$

If we write $x := (x_1, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{(n-1)}$ with $\bar{x} := (x_2, \ldots, x_n)^T \in \mathbb{R}^{(n-1)}$, the p-order cone $\mathcal{K}_p$ can be expressed as

$$\mathcal{K}_p = \left\{ x \in \mathbb{R}^n \mid x_1 \ge \|\bar{x}\|_p \right\}, \quad (1 < p < \infty).$$

Indeed, $\mathcal{K}_p$ is a solid (i.e., int $\mathcal{K}_p \neq \emptyset$), closed and convex cone, and its dual cone is given by

$$\mathcal{K}_p^* = \left\{ y \in \mathbb{R}^n \,\middle|\, y_1 \ge \left(\sum_{i=2}^n |y_i|^q \right)^{\frac{1}{q}} \right\}$$

or equivalently

$$\mathcal{K}_p^* = \left\{ y = (y_1, \bar{y}) \in \mathbb{R} \times \mathbb{R}^{(n-1)} \mid y_1 \ge \|\bar{y}\|_q \right\} = \mathcal{K}_q,$$

where q satisfies the condition $q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, and $\bar{y} := (y_2, y_3, \ldots, y_n)^T \in \mathbb{R}^{(n-1)}$. Note that the dual cone $\mathcal{K}_p^*$ is also a convex cone.

(2) L^p **cone**, see [10]. Let $n \in \mathbb{N}$ and $p := (p_1, p_2, \ldots, p_n)^T \in \mathbb{R}^n$ with $p_i > 1$. The L^p cone is defined by

$$L^p := \left\{ (z, \theta, k) \in \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{R}_+ \;\middle|\; \sum_{i=1}^{n} \frac{|z_i|^{p_i}}{p_i \theta^{p_i - 1}} \leq k \right\},$$

where $|z_i|/0 := \infty$ if $z_i \neq 0$; 0 if $z_i = 0$. As shown in [10], we know that L^p is a solid, closed and convex cone, and its dual cone is the switched cone L_s^q given by

$$(w, h, \phi) \in L_s^q \iff (w, \phi, h) \in L^q,$$

where $q := (q_1, q_2, \ldots, q_n)^T \in \mathbb{R}_{++}^n$ such that $\frac{1}{p_i} + \frac{1}{q_i} = 1$ for each i.

(3) The circular cone L_θ, see [7,18]. The circular cone L_θ is defined as follows:

$$L_\theta := \left\{ x = (x_1, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{(n-1)} \mid \|x\| \cos\theta \leq x_1 \right\},$$

where $\theta \in (0, \frac{\pi}{2})$. Again, as shown in [7,18], we know that L_θ is a solid, closed and convex cone, and its dual cone L_θ^* is given by

$$L_\theta^* = \left\{ z = (z_1, \bar{z}) \in \mathbb{R} \times \mathbb{R}^{(n-1)} \;\middle|\; \|z\| \cos\left(\frac{\pi}{2} - \theta\right) \leq z_1 \right\} = L_{\frac{\pi}{2} - \theta}.$$

By direct calculation or reference to [18], the circular cone L_θ and its dual cone L_θ^* can also be expressed as follows, respectively,

$$L_\theta = \left\{ x = (x_1, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{(n-1)} \mid \|\bar{x}\| \cot\theta \leq x_1 \right\}$$

and

$$L_\theta^* = \left\{ z = (z_1, \bar{z}) \in \mathbb{R} \times \mathbb{R}^{(n-1)} \mid \|\bar{z}\| \tan\theta \leq z_1 \right\}.$$

Remark 2.1 (a) When $p = 2$, $\mathcal{K}_p$ is exactly the second-order cone which says that p-order cone is a generalization of the second-order cone.

(b) When $p_i = 2$ for all i, we have that the L^p cone is the hyperbolic or rotated second-order cone which is a transformation of the standard second-order cone.

(c) Clearly, the circular cone L_θ includes second-order cone as a special case when the rotation angle is 45°.

3 Characterizations of solution set with Lagrange multiplier

In this section, we will establish some results which characterizes the solution set of problem (1) in terms of Lagrange multiplier of a solution and subgradients of Lagrange function for the problem (1). we first show a necessary condition for the solution set of problem (1). Then, when $\mathcal{K}$ reduces to the aforementioned specific cones, we show the same results can be obtained by other ways by exploiting the special structure of those three cones.

Theorem 3.1 *For problem* (1), *let* $a \in S$. *Suppose that the corresponding Lagrange multiplier* $\lambda_a \in \mathbb{R}^r$ *satisfies the conditions*:

$$0 \in \partial L_a(a, \lambda_a) + \{A^T y \mid y \in \mathbb{R}^m\}, \quad \lambda_a \in \mathcal{K}^*, \quad \textit{and} \quad \lambda_a^T g(a) = 0. \tag{2}$$

Then, the following hold.

(a) *If* $\lambda_a = 0$, *then, for each* $x \in S$, *there exists* $y \in \mathbb{R}^m$ *such that*

$$-A^T y \in \partial f(x).$$

(b) *If* $\lambda_a \neq 0$, *then, for each* $x \in S$ *and* $g(x) \neq 0$, *we have*

$$-g(x) \in \partial \mathcal{K}, \quad \lambda_a \in \partial \mathcal{K}^*, \quad \textit{and} \quad \lambda_a^T g(x) = 0.$$

Proof (a) When $\lambda_a = 0$, since the Lagrange multiplier λ_a satisfies the condition $0 \in \partial L_a(a, \lambda_a) + \{A^T y \mid y \in \mathbb{R}^m\}$, there exists $y \in \mathbb{R}^m$ such that

$$-A^T y \in \partial L_a(a, \lambda_a) = \partial f(a) + \partial(\lambda_a^T g)(a) = \partial f(a).$$

Applying the properties of convex functions, for each $x \in S$ and every $z \in \mathbb{R}^n$, it follows that

$$\begin{aligned} f(z) - f(x) &= f(z) - f(a) \\ &\geq -(A^T y)^T (z - a) \\ &= -(A^T y)^T (z - x + x - a) \\ &= -(A^T y)^T (z - x) - (A^T y)^T (x - a) \\ &= -(A^T y)^T (z - x), \end{aligned}$$

where the first and last equalities respectively follow from $f(x) = f(a)$ and $Ax = Aa = b$ due to $x, a \in S$, which implies that $-A^T y \in \partial f(x)$.

(b) When $\lambda_a \neq 0$, from the conditions (2), i.e.,

$$0 \in \partial L_a(a, \lambda_a) + \left\{A^T y \mid y \in \mathbb{R}^m\right\}, \quad \lambda_a \in \mathcal{K}^*, \quad \text{and} \quad \lambda_a^T g(a) = 0,$$

we know there exists $y \in \mathbb{R}^m$ such that

$$-A^T y \in \partial L_a(a, \lambda_a).$$

Because f is convex and g is $\mathcal{K}$-convex, the function $L_a(\cdot, \lambda_a)$ is convex as shown earlier in Section 2. Therefore, for every $x \in S$, we have

$$\begin{aligned} f(x) + \lambda_a^T g(x) &= L_a(x, \lambda_a) \\ &\geq L_a(a, \lambda_a) - (A^T y)^T (x - a) \\ &= L_a(a, \lambda_a) \\ &= f(a) + \lambda_a^T g(a). \end{aligned} \tag{3}$$

This together with $f(x) = f(a)$ and $\lambda_a^T g(a) = 0$ yield $\lambda_a^T g(x) \geq 0$. On the other hand, noting that for every $x \in S$, $\lambda_a \in \mathcal{K}^*$ and $-g(x) \in \mathcal{K}$, by the definition of the dual cone $\mathcal{K}^*$, we obtain $\lambda_a^T g(x) \leq 0$. Hence, this together with $\lambda_a^T g(x) \geq 0$ give $\lambda_a^T g(x) = 0$.

Next, we argue that $\lambda_a \in \partial\mathcal{K}^*$ and $-g(x) \in \partial\mathcal{K}$ for every $x \in S$ and $g(x) \neq 0$. We prove $-g(x) \in \partial\mathcal{K}$ only. Similar arguments will apply to the case of $\lambda_a \in \partial\mathcal{K}^*$. Indeed, we will prove it by contradiction. Suppose that $-g(x) \in \text{int } \mathcal{K}$. Then, there exists $\varepsilon > 0$ such that the open ball $\mathbb{B}(-g(x), \varepsilon) \subset \mathcal{K}$. Thus, for any $y \in \mathbb{R}^r$, there exists $\alpha > 0$ such that $-g(x) + \alpha y \in \mathbb{B}(-g(x), \varepsilon) \subset \mathcal{K}$, which gives

$$\langle -g(x) + \alpha y, \lambda_a \rangle \geq 0.$$

Then, it follows from $\langle \lambda_a, -g(x) \rangle = -\lambda_a^T g(x) = 0$ that $\alpha \langle y, \lambda_a \rangle \geq 0$. By the arbitrariness of y in $\mathbb{R}^r$, we see that $\lambda_a = 0$. This contradicts the condition $\lambda_a \neq 0$. Thus, $-g(x) \in \partial\mathcal{K}$. □

Remark 3.1 (i) By Theorem 3.1, for every $x, y \in S$, we have

$$f(x) + \lambda_a^T g(x) = f(a) + \lambda_a^T g(a) = f(y) + \lambda_a^T g(y).$$

This explains that Lagrange function $L_a(\cdot, \lambda_a)$ is constant on the solution set S of problem (1).

(ii) By Theorem 3.1, for every $a, x \in S$, the Lagrange multiplier λ_a and the vector $-g(x)$ solve the complementarity problem [9]:

$$-g(x) \in \mathcal{K}, \quad \lambda_a \in \mathcal{K}^*, \quad \lambda_a^T g(x) = 0.$$

Now, we show that Theorem 3.1(b) can be verified by other ways when $\mathcal{K}$ reduces to the p-order cone, L^p cone or the circular cone. We present the three cases as below.

(1) For the case where $\mathcal{K}$ is the p-order cone $\mathcal{K}_p$, let $0 \neq -g(x) := (h_1, \bar{h}) \in \mathcal{K}_p \subset \mathbb{R}_+ \times \mathbb{R}^{r-1}$ and $0 \neq \lambda_a := (\lambda_1, \bar{\lambda}) \in \mathcal{K}_q \subset \mathbb{R}_+ \times \mathbb{R}^{r-1}$. Note that $h_1 > 0$ and $\lambda_1 > 0$. By the definitions of $\mathcal{K}_p$ and its dual cone $\mathcal{K}_q$, we have

$$h_1 \geq \|\bar{h}\|_p \quad \text{and} \quad \lambda_1 \geq \|\bar{\lambda}\|_q.$$

Hence, it follows from $\lambda_a^T g(x) = 0$ that

$$\begin{aligned} 0 &= h_1\lambda_1 + \bar{h}^T\bar{\lambda} \\ &\geq \|\bar{h}\|_p\|\bar{\lambda}\|_q + \bar{h}^T\bar{\lambda} \\ &\geq 0, \end{aligned}$$

where the last inequality follows by Hölder's inequality. This leads to $h_1 = \|\bar{h}\|_p$ and $\lambda_1 = \|\bar{\lambda}\|_q$ due to $h_1\lambda_1 > 0$, which says $\lambda_a \in \partial\mathcal{K}_p^*$ and $-g(x) \in \partial\mathcal{K}_p$.

(2) For the case where $\mathcal{K}$ is the L^p cone, let $0 \neq -g(x) := (z, \theta, k) \in L^p \subset \mathbb{R}^{r-2} \times \mathbb{R}_+ \times \mathbb{R}_+$ and $0 \neq \lambda_a := (w, h, \phi) \in L_s^q \subset \mathbb{R}^{r-2} \times \mathbb{R}_+ \times \mathbb{R}_+$. By the definitions of L^p cone and its dual cone L_s^q, we obtain that

$$\sum_{i=1}^{r-2} \frac{|z_i|^{p_i}}{p_i\theta^{p_i-1}} \leq k \quad \text{and} \quad \sum_{i=1}^{r-2} \frac{|w_i|^{q_i}}{q_i\phi^{q_i-1}} \leq h.$$

We discuss two subcases.

Case 1: $\theta = 0$ or $\phi = 0$. If $\theta = 0$, then by definition, $z = 0$ follows. Then $\lambda_a^T g(x) = 0$ becomes $k\phi = 0$, and hence $\phi = 0$ because $-g(x) = (0, 0, k) \neq 0$. Also, $\phi = 0$ yields $w = 0$, so that $\lambda_a = (0, h, 0)$. Therefore, $-g(x) \in \partial L^p$ and $\lambda_a \in \partial L_s^q$.

Case 2: $\theta > 0$ and $\phi > 0$. Then, $\lambda_a^T g(x) = 0$ that

$$\begin{aligned} 0 &= z^T w + \theta h + k\phi \\ &\geq z^T w + \theta \sum_{i=1}^{r-2} \frac{|w_i|^{q_i}}{q_i\phi^{q_i-1}} + \phi \sum_{i=1}^{r-2} \frac{|z_i|^{p_i}}{p_i\theta^{p_i-1}} \\ &= z^T w + \theta\phi \sum_{i=1}^{r-2} \left(\frac{1}{q_i} |\frac{w_i}{\phi}|^{q_i} + \frac{1}{p_i} |\frac{z_i}{\theta}|^{p_i} \right) \\ &\geq z^T w + \theta\phi \sum_{i=1}^{r-2} \left|\frac{w_i}{\phi}\right| \cdot \left|\frac{z_i}{\theta}\right| \\ &\geq z^T w - \sum_{i=1}^{r-2} w_i z_i \\ &= 0, \end{aligned}$$

where the second inequality follows from Young's inequality. This implies $\sum_{i=1}^{r-2} \frac{|z_i|^{p_i}}{p_i\theta^{p_i-1}} = k$ and $\sum_{i=1}^{r-2} \frac{|w_i|^{q_i}}{q_i\phi^{q_i-1}} = h$, which says $\lambda_a \in \partial L_s^q$ and $-g(x) \in \partial L^p$.

(3) For the case where $\mathcal{K}$ is the circular cone L_θ, let $0 \neq -g(x) := (h_1, \bar{h}) \in L_\theta \subset \mathbb{R}_+ \times \mathbb{R}^{r-1}$ and $0 \neq \lambda_a := (\lambda_1, \bar{\lambda}) \in L_\theta^* \subset \mathbb{R}_+ \times \mathbb{R}^{r-1}$. By the expressions of the circular cone L_θ and its dual cone L_θ^+, we have

$$\|\bar{h}\| \cot\theta \leq h_1 \quad \text{and} \quad \|\bar{\lambda}\| \tan\theta \leq \lambda_1.$$

Then, from $\lambda_a^T g(x) = 0$, we have

$$\begin{aligned} 0 &= h_1\lambda_1 + \bar{h}^T\bar{\lambda} \\ &\geq \|\bar{h}\| \cot\theta \cdot \|\bar{\lambda}\| \tan\theta + \bar{h}^T\bar{\lambda} \\ &= \|\bar{h}\|\|\bar{\lambda}\| + \bar{h}^T\bar{\lambda} \\ &\geq 0. \end{aligned}$$

This leads to $\|\bar{h}\| \cot\theta = h_1$ and $\|\bar{\lambda}\| \tan\theta = \lambda_1$, which yields $\lambda_a \in \partial L_\theta^*$ and $-g(x) \in \partial L_\theta$.

From [16, Theorem 3.1], we have the following theorem which will give the form of the solution set of problem (1) in terms of subgradients.

Theorem 3.2 *For problem* (1)*, let $a \in S$. Then*

$$\begin{aligned} S &= \{x \in F \mid \langle \xi, x-a\rangle = 0, \exists\, \xi \in \partial f(x) \cap \partial f(a)\} \\ &= \{x \in F \mid \langle \xi, x-a\rangle = 0, \exists\, \xi \in \partial f(x)\} \\ &= \{x \in F \mid \langle \xi, x-a\rangle \leq 0, \exists\, \xi \in \partial f(x)\}. \end{aligned}$$

Proof Let C_1, C_5 and C_6 be the following sets, respectively,

$$\begin{aligned} C_1 &:= \{x \in F \mid \langle \xi, x-a\rangle = 0, \exists\, \xi \in \partial f(x) \cap \partial f(a)\}, \\ C_5 &:= \{x \in F \mid \langle \xi, x-a\rangle = 0, \exists\, \xi \in \partial f(x)\} \end{aligned}$$

and

$$C_6 := \{x \in F \mid \langle \xi, x-a\rangle \leq 0, \exists\, \xi \in \partial f(x)\}.$$

Then, the sets C_1,C_5 and C_6 correspond to those in [16, Theorem 3.1], from which the results follow immediately. □

Theorem 3.3 *For problem* (1)*, let $a \in S$ and let λ_a be the corresponding Lagrange multiplier satisfying the conditions:*

$$0 \in \partial L_a(a, \lambda_a) + \left\{A^T y \mid y \in \mathbb{R}^m\right\}, \quad \lambda_a \in \mathcal{K}^*, \quad \textit{and} \quad \lambda_a^T g(a) = 0.$$

(a) *If $\lambda_a = 0$, then*

$$S = \left\{x \in F \mid \partial f(a) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\} = \partial f(x) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\}\right\}.$$

(b) *If $\lambda_a \neq 0$, then*

$$S = \left\{x \in F \mid \lambda_a^T g(x) = 0,\ 0 \in \partial L_a(a, \lambda_a) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\}\right\}.$$

Proof (a) For convenience, we denote

$$\bar{S} = \left\{x \in F \mid \partial f(a) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\} = \partial f(x) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\}\right\}.$$

Then, we need to argue that $S = \bar{S}$ as below.

We first verify the direction $S \subset \bar{S}$. Let $C := \{x \mid Ax = b\}$. By the analysis of section 2, we know that $\mathcal{N}_C(x) = \{A^T y \mid y \in \mathbb{R}^m\}$. If $\lambda_a = 0$, it follows that $L_a(x, \lambda_a) = f(x)$. Then, we have $\partial L_a(x, \lambda_a) = \partial f(x)$. Hence, for any $x \in S$, by proposition 2.1 of [11] again, it is easy to obtain that

$$\partial f(a) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\} = \partial f(x) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\}.$$

This yields $S \subset \bar{S}$.

Conversely, let $x \in \bar{S}$. Then, we know that $x \in F$ and

$$\partial f(a) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\} = \partial f(x) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\}.$$

Since $a \in S$ and its corresponding Lagrange multiplier λ_a satisfy the condition

$$0 \in \partial L_a(a, \lambda_a) + \left\{A^T y \mid \ y \in \mathbb{R}^m\right\},$$

we have

$$-A^T y \in \partial f(a) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\} = \partial f(x) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\},$$

for some $y \in \mathbb{R}^m$. Then, it is easy to see that $-y^T A(x - a) = 0$. This together with Theorem 3.2 implies $x \in S$. Hence, the conclusion holds.

(b) Let $\lambda_a \neq 0$. By the Remark 3.1, we know that the Lagrange function $L_a(\cdot, \lambda_a)$ is constant on the solution set S of the problem (1). Hence, for any $x \in S$ and $a \in S$, we have $L_a(x, \lambda_a) = L_a(a, \lambda_a)$ and then for each $\xi \in \partial L_a(x, \lambda_a) \cap \{-A^T y \mid y \in \mathbb{R}^m\}$, there exists $y \in \mathbb{R}^m$ such that $-A^T y = \xi$. Moreover, we get also that

$$\begin{aligned} L_a(x, \lambda_a) - L_a(a, \lambda_a) &= 0 \\ &= L_a(a, \lambda_a) - L_a(x, \lambda_a) \\ &\geq -(A^T y)^T (a - x) \\ &= -y^T A(a - x) \\ &= 0 \\ &= -y^T A(x - a), \end{aligned}$$

where the fourth equality holds due to $a, x \in S \subset F$. This shows that $\xi = -A^T y \in \partial L_a(a, \lambda_a) \cap \{-A^T y \mid y \in \mathbb{R}^m\}$. Thus, we have

$$\partial L_a(x, \lambda_a) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\} \subset \partial L_a(a, \lambda_a) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\}.$$

Similarly, with the same arguments, we may verify that

$$\partial L_a(a, \lambda_a) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\} \subset \partial L_a(x, \lambda_a) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\}.$$

Therefore,

$$\partial L_a(a, \lambda_a) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\} = \partial L_a(x, \lambda_a) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\}.$$

Combining with Corollary 2.6 of [11], this implies

$$S = \left\{x \in F \,\middle|\, \lambda_a^T g(x) = 0,\ 0 \in \partial L_a(a, \lambda_a) \cap \left\{-A^T y \mid y \in \mathbb{R}^m\right\}\right\},$$

which is the desired result. □

Remark 3.2 In the setting of the Banach space and $\mathcal{K}$ is a closed convex cone, the corresponding conclusions of Theorem 3.3 have been obtained, see [11, Corollary 2.5] and [16, Corollary 3.1]. However, in [11, Corollary 2.5], the expression of the solution set $\bar{S}$ is more complicated than that given in Theorem 3.3. Here, we provide a simplified expression for the solution set S.

Example 3.1 Consider the following nonlinear convex programming problem:

$$\begin{aligned} \min \quad & f(x) = \sqrt{x_1^2 + x_2^2} + x_2 \\ \text{s.t.} \quad & -g(x) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} \in \mathcal{K}_p, \end{aligned}$$

where $x := (x_1, x_2)^T \in \mathbb{R}^2$.

Let F and S be the feasible set and the solution set of the considered problem, respectively. For any $x = (x_1, x_2)^T \in F$, we have

$$f(x) = \sqrt{x_1^2 + x_2^2} + x_2 \geq |x_2| + x_2 \geq 0.$$

Thus, we know that $a = (0, 0)^T$ is a solution of the considered problem, i.e., $a \in S$. Note that

$$\partial f(a) = \{(0, 1)^T\} + \mathbb{B},$$

where $\mathbb{B}$ denotes the closed unit ball of $\mathbb{R}^n$, and

$$\partial f(x) = \left\{\left(\frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \frac{x_2}{\sqrt{x_1^2 + x_2^2}} + 1\right)^T\right\}$$

for any $x \neq a$. For the solution $a = (0,0)^T \in S$, it is easy to see that the corresponding Lagrange multiplier $\lambda_a = (0,0)^T \in \mathcal{K}_q$. Moreover, we also obtain that $(0,0)^T \in \partial L_a(a, \lambda_a) = \partial f(a)$. Therefore, it follows that

$$(0,0)^T \in \partial f(x) \Longleftrightarrow x_1 = 0, \quad x_2 \leq 0.$$

With this, we see that the solution set can be simplified as

$$S = \{x = (x_1, x_2)^T \in \mathbb{R}^2 \mid x_1 = 0, x_2 \leq 0\}.$$

To close this section, combining Theorem 3.3, [16, Corollary 3.1] and the contents of [2, page 267], we immediately obtain the following corollary as a special case.

Corollary 3.1 *For problem (1), let $a \in S$. If the $\mathcal{K}$-convex mapping g is an identity mapping, i.e., $g(x) = x$ for all $x \in \mathbb{R}^n$, then the following hold.*

(a) *If the solution $a \in \operatorname{int} \mathcal{K}$, then*

$$S = \{x \in F \mid \partial f(x) = \partial f(a)\}.$$

(b) *If the solution $a \in \partial \mathcal{K}$, then*

$$S = \left\{ x \in F \,\middle|\, \partial f(x) \cap \left[\mathcal{N}_{C_1}(x) - \{\lambda \mid \lambda \in \mathcal{K}^*, \lambda^T x = 0\}\right] \right.$$
$$\left. = \partial f(a) \cap \left[\mathcal{N}_{C_1}(a) - \{\lambda_a \mid \lambda_a \in \partial \mathcal{K}^*, \lambda_a^T x = 0\}\right] \right\},$$

where $\mathcal{N}_{C_1}(x) = \mathcal{N}_{C_1}(a) = \{A^T y \mid y \in \mathbb{R}^m\}$ with $C_1 = \{x \in \mathbb{R}^n \mid Ax = b\}$.

Acknowledgments The authors are very grateful to the referees for their constructive comments, which have considerably improved the paper.

References

1. Andersen, E.D., Roos, C., Terlaky, T.: Notes on duality in second order and p-order cone optimization. Optimization **51**(4), 627–643 (2002)
2. Bertsekas, D.P., Nedić, A., Ozdaglar, A.E.: Convex Analysis and Optimization. Athena Scientific (2003)
3. Burke, J.V., Ferris, M.C.: Characterization of solution sets of convex programs. Oper. Res. Lett. **10**, 57–60 (1991)
4. Chen, J.-S.: Conditions for error bounds and bounded level sets of some merit functions for the second-order cone complementarity problem. J. Optim. Theory Appl. **135**, 459–473 (2007)
5. Chen, J.-S., Tseng, P.: An unconstrained smooth minimization reformulation of second-order cone complementarity problem. Math. Program. **104**, 293–327 (2005)
6. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley-Interscience, New York (1983)
7. Pznto Da Costa, A., Seeger, A.: Numerical resolution of cone-constrained eigenvalue problems. Comput. Appl. Math. **28**(1), 37–61 (2009)
8. Deng, S.: Characterizations of the nonemptiness and compactness of solution sets in convex vecter optimization. J. Optim. Theory Appl. **96**, 123–131 (1998)

9. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, vol. I. Springer, New York (2003)
10. Glineur, F., Terlaky, T.: Conic formulation for l_p-norm optimization. J. Optim. Theory Appl. **122**(2), 285–307 (2004)
11. Jeyakumar, V., Lee, G.M., Dinh, N.: Lagrange multiplier conditions characterizing the optimal solution sets of cone-constrained convex programs. J. Optim. Theory Appl. **123**(1), 83–103 (2004)
12. Jeyakumar, V., Wolkowicz, H.: Generalizations of Slater's constraint qualification for infinite convex programs. Math. Program. **57**, 85–101 (1992)
13. Jeyakumar, V., Yang, X.-Q.: characterizing the solution sets of pseudolinear programs. J. Optim. Theory Appl. **87**, 747–755 (1995)
14. Mangasarian, O.L.: A simple characterization of solution sets of convex programs. Oper. Res. Lett. **7**(1), 21–26 (1988)
15. Pan, S.-H., Kum, S., Lim, Y., Chen, J.-S.: On the generalized Fischer–Burmeister merit function for the second-order cone complementarity problem. Math. Comput. **83**(287), 1143–1171 (2014)
16. Wu, Z.-L., Wu, S.-Y.: Characterizations of the solution sets of convex programs and variational inequality problems. J. Optim. Theory Appl. **130**(2), 339–358 (2006)
17. Xue, G., Ye, Y.: An efficient algorithm for minimizing a sum of p-norms. SIAM J. Optim. **10**(2), 315–330 (1999)
18. Zhou, J.-C., Chen, J.-S.: Properties of circular cone and spectral factorization associated with circular cone. J. Nonlinear Convex Anal. **14**(4), 807–816 (2013)