Journal of Convex Analysis
Volume 16 (2009), No. 3&4, 845–856

Qualification-Free Optimality Conditions for Convex Programs with Separable Inequality Constraints*

G. Y. Li
Department of Applied Mathematics,
University of New South Wales, Sydney 2052, Australia
g.li@unsw.edu.au

V. Jeyakumar†
Department of Applied Mathematics,
University of New South Wales, Sydney 2052, Australia
v.jeyakumar@unsw.edu.au

Dedicated to Stephen Simons on the occasion of his 70th birthday.

Received: February 22, 2008
Revised manuscript received: October 8, 2008

In this paper, we show that separable convex functions enjoy ϵ-subdifferential sum formula as well as the Fenchel duality without a regularity assumption, and establish that for convex programs with separable convex constraints a new partially asymptotic Lagrange multiplier conditions hold without a constraint qualification. Examples are given to illustrate the results.

1. Introduction

Consider the convex program

$$\inf\{f(x) : g_i(x) \leq 0,\ i = 1, 2, \cdots, n\}, \tag{1}$$

where f and g_i $(i = 1, \ldots, n)$ are real-valued convex functions, defined on $\mathbb{R}^m$. It is known that certain technical condition on the constraints, known as constraint qualification, guarantees that the convex program enjoys the Lagrange multiplier condition,

$$(\exists \lambda \in \mathbb{R}^n_+)\ \ 0 \in \partial f(a) + \sum_{i=1}^{n} \partial(\lambda^i g_i)(a), \quad \lambda^i g_i(a) = 0,$$

which is necessary and sufficient for optimality at a. In the absence of a constraint qualification, it has recently been shown that the following limiting Lagrange multiplier condition holds (see Jeyakumar et al. [8] and also Thibault [15]):

$$u^* + v_k^* \to 0, \qquad \lambda_k^i g_i(a) \to 0, \qquad \epsilon_k \to 0,$$

*Research was partially supported by a grant from the Australian Research Council.
†Corresponding author.

ISSN 0944-6532 / $ 2.50

for some sequences $\{\epsilon_k\} \subset \mathbb{R}_+$ and $\{\lambda_k\} \subset \mathbb{R}^n_+$, and for some $u^* \in \partial f(a)$ and $v_k^* \in \partial_{\epsilon_k}(\sum_{i=1}^n \lambda_k^i g_i)(a)$. For related results see [9].

The purpose of this paper is to show that separable convex functions enjoy the ϵ-subdifferential sum formula as well as the Fenchel duality without a regularity assumption, and to establish that for convex programs with separable convex constraints the following strengthened form of the limiting Lagrange multiplier conditions holds:

$$0 \in \partial f(a) + \sum_{i=1}^n \partial_{\epsilon_k}(\lambda_k^i g_i)(a), \qquad \lambda_k^i g_i(a) \to 0, \qquad \epsilon_k \to 0,$$

for some sequences $\{\epsilon_k\} \subset \mathbb{R}_+$ and $\{\lambda_k\} \subset \mathbb{R}^n_+$. Our method of proof makes use of the fact, established recently by Tseng in [16], that there is no duality gap between (1) and its Lagrangian dual whenever the functions f and g_i are separable convex functions. Numerical examples are discussed to illustrate our results.

The class of convex programming problems with separable constraints is an extension of the standard convex quadratic programming problems (i.e., convex quadratic problems with linear constraints) and it often arises in important application areas. For instance, many classes of network optimization problems and integer programming problems can be cast as convex programming problems with separable constraints. For recent work on separable convex programming and convex programming problems with separable constraints, see [2, 10, 14, 16] and the reference therein.

The outline of the paper is as follows. Section 2 provides definitions and some basic results on conjugate functions, convex sets and functions. In Section 3, we present an ϵ-subdifferential sum formula as well as the Fenchel duality result for separable functions. Finally, in Section 4, we establish a new form of subgradient optimality conditions for convex programming with separable inequality constraint.

2. Preliminaries on Conjugate and Convex Functions

Throughout this paper, $\mathbb{R}^m$ denotes Euclidean space with dimension m. The corresponding inner product in $\mathbb{R}^m$ is defined by $\langle x, y\rangle = x^T y$ for any $x, y \in \mathbb{R}^m$. We use $\mathbb{B}(x;\epsilon)$ (resp. $\overline{\mathbb{B}}(x;\epsilon)$) to denote the open (resp. closed) ball with center x and radius ϵ. For a set A in $\mathbb{R}^m$, the interior (resp. relative interior, closure, convex hull, affine hull) of A is denoted by $\text{int}A$ (resp. $\text{ri}A$, $\overline{A}$, $\text{co}A$, $\text{aff}A$). The recession cone of A, denoted by A_∞, is defined by $A_\infty = \{d : a + td \in A \text{ for all } t \geq 0 \text{ and for all } a \in A\}$. The indicator function $\delta_A : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$\delta_A(x) := \begin{cases} 0, & \text{if } x \in A, \\ +\infty, & \text{otherwise.} \end{cases} \tag{2}$$

For a convex function f on $\mathbb{R}^m$, the effective domain and the epigraph are respectively defined by $\text{dom}f := \{x \in \mathbb{R}^m : f(x) < +\infty\}$ and $\text{epi}f := \{(x, r) \in \mathbb{R}^m \times \mathbb{R} : f(x) \leq r\}$. We say f is proper if $f(x) > -\infty$ for all $x \in X$ and $\text{dom}f \neq \emptyset$. Moreover, if $\liminf_{x' \to x} f(x') \geq f(x)$ for all $x \in \mathbb{R}^m$, we say f is a lower semicontinuous function. The (convex) subdifferential of f at $x \in \mathbb{R}^m$ is defined by

$$\partial f(x) = \begin{cases} \{x^* \in \mathbb{R}^m : \langle x^*, y - x\rangle \leq f(y) - f(x) \ \forall\, y \in \mathbb{R}^m\}, & \text{if } x \in \text{dom}f, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{3}$$

More generally, the ϵ-subdifferential of f at $x \in \mathbb{R}^m$ is defined by

$$\partial_\epsilon f(x) = \begin{cases} \{x^* \in \mathbb{R}^m : \langle x^*, y - x\rangle \leq f(y) - f(x) + \epsilon \; \forall\, y \in \mathbb{R}^m\}, & \text{if } x \in \operatorname{dom} f, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{4}$$

The normal cone of a convex set A at the point $a \in A$, denoted by $N_A(a)$, is defined by

$$N_A(a) = \partial \delta_A(a) = \{x^* \in \mathbb{R}^m : \langle x^*, x - a\rangle \leq 0 \text{ for all } x \in A\}$$

Let f be a proper convex function and let $x \in \operatorname{dom} f$ be such that $\partial f(x) \neq \emptyset$. Then, one has (cf. [1, Proposition 2.5.4])

$$(\partial f(x))_\infty = N_{\operatorname{dom} f}(x).$$

Let A be a closed convex subset of $\mathbb{R}^m$. We denote $\Gamma(A)$ to be the proper lower semicontinuous convex functions on A. We also define $\Gamma_s(A)$ as follows:

$$\Gamma_s(A) = \{f \in \Gamma(A) : \partial f(a) \neq \emptyset \text{ for all } a \in A \cap \operatorname{dom} f\}.$$

It can be verify that $\Gamma_s(A)$ is a vector space (under the addition and scalar multiplication) which contains the following two important classes of convex functions: (1) $f \in \Gamma(A)$ satisfying $f(x) > \inf f \Rightarrow x \in \operatorname{ri} \operatorname{dom} f$; (2) $f = \delta_C$ for some closed convex set C where δ is the indicator function. As usual, for any $f \in \Gamma(\mathbb{R}^m)$, its conjugate function $f^* \in \Gamma(\mathbb{R}^m)$ (cf. [13]) is defined by $f^*(x^*) = \sup_{x \in \mathbb{R}^m}\{\langle x^*, x\rangle - f(x)\}$ for all $x^* \in \mathbb{R}^m$. The definition of f^* entails that $\langle x^*, x\rangle \leq f^*(x^*) + f(x)$ (Young's inequality) for any $x \in \mathbb{R}^m$ and $x^* \in \mathbb{R}^m$. Moreover, for any $\epsilon \geq 0$ and $x \in \operatorname{dom} f$

$$x^* \in \partial_\epsilon f(x) \;\Leftrightarrow\; f^*(x^*) + f(x) \leq \langle x^*, x\rangle + \epsilon \;\Leftrightarrow\; (x^*, \epsilon + \langle x^*, x\rangle - f(x)) \in \operatorname{epi} f^*. \tag{5}$$

In particular, we have the following Young's equality

$$x^* \in \partial f(x) \;\Leftrightarrow\; \langle x^*, x\rangle = f^*(x^*) + f(x).$$

It is well known that (cf. [13]) for any proper lower semicontinuous convex functions f_1, f_2,

$$f_1 \leq f_2 \;\Leftrightarrow\; f_1^* \geq f_2^* \;\Leftrightarrow\; \operatorname{epi} f_1^* \subseteq \operatorname{epi} f_2^*. \tag{6}$$

Let f_i $(1 \leq i \leq n)$ be proper lower semicontinuous convex functions on $\mathbb{R}^m$. The infimal convolution of f_i, denoted $f_1 \square \dots \square f_n$, is defined by

$$(f_1 \square \dots \square f_n)(x) = \inf\left\{\sum_{i=1}^n f_i(x_i) : \sum_{i=1}^n x_i = x\right\} \quad \text{for all } x \in \mathbb{R}^m.$$

It is well known (see [13]) that if $\bigcap_{i=1}^n \operatorname{dom} f_i \neq \emptyset$, then $(f_1 \square \dots \square f_n)^* = \sum_{i=1}^n f_i^*$. Moreover we also have

$$\left(\sum_{i=1}^n f_i\right)^* = \operatorname{cl}(f_1^* \square \dots \square f_n^*) \quad \text{and} \quad \operatorname{epi}\left(\sum_{i=1}^n f_i\right)^* = \overline{\sum_{i=1}^n \operatorname{epi} f_i^*}. \tag{7}$$

The lower semicontinuous hull in the first equation and the closure in the second equation are superfluous (see [13] for detail) if there exists $i_0 \in \{1, \ldots, n\}$ such that

$$\mathrm{dom} f_{i_0} \cap \left(\bigcap_{i \neq i_0} \mathrm{int}\, \mathrm{dom} f_i \right) \neq \emptyset. \tag{8}$$

Finally, a function $f : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ is called a separable function on $\mathbb{R}^m$ if

$$f(x) = \sum_{l=1}^{m} g_l(x_l) \quad \forall\, x = (x_1, \ldots, x_m) \in \mathbb{R}^m \tag{9}$$

for some proper lower semicontinuous function g_l on $\mathbb{R}$ $(1 \leq l \leq m)$. Clearly, an affine function f is, in particular, separable and convex.

Lemma 2.1 (see [5]). *Let f be a proper lower semicontinuous function on $\mathbb{R}^m$. Then for each $x \in \mathrm{dom} f$,*

$$\mathrm{epi} f^* = \bigcup_{\epsilon \geq 0} \{(x^*, \epsilon + \langle x^*, x \rangle - f(x)) : x^* \in \partial_\epsilon f(x)\}.$$

Lemma 2.2 (cf. [8]). *Let I be an arbitrary index set and let f_i $(i \in I)$ be proper lower semicontinuous functions on $\mathbb{R}^m$. Suppose that there exists $x_0 \in \mathbb{R}^m$ such that $\sup_{i \in I} f_i(x_0) < \infty$. Then*

$$\mathrm{epi} \left(\sup_{i \in I} f_i \right)^* = \overline{\mathrm{co} \bigcup_{i \in I} \mathrm{epi} f_i^*},$$

where $\sup_{i \in I} f_i : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ is defined by $(\sup_{i \in I} f_i)(x) = \sup_{i \in I} f_i(x)$ for all $x \in \mathbb{R}^m$.

3. Separable Convex Functions and ϵ-subdifferential formulas

In this section, we establish ϵ-subdifferential sum formulas as well as some Fenchel duality results for separable convex functions. To do this , we recall the following results which is essentially due to [4, 5]. For related results, see [11, 12]. We state it in a version that is convenient to us.

Theorem 3.1. *Let $n \in \mathbb{N}$ and let $f_i \in \Gamma(\mathbb{R}^m)$ $(1 \leq i \leq n)$ with $\bigcap_{i=1}^n \mathrm{dom} f_i \neq \emptyset$. Let $f = \sum_{i=1}^n f_i$. Then the following statements are equivalent:*

(i) *(ϵ-sum rule)*

$$\partial_\epsilon f(x) = \bigcup \left\{ \sum_{i=1}^{n} \partial_{\epsilon_i} f_i(x) : \sum_{i=1}^{n} \epsilon_i = \epsilon,\ \epsilon_i \geq 0\, (i \in I) \right\} \ \forall\, \epsilon \geq 0 \text{ and } x \in \mathrm{dom} f; \tag{10}$$

(ii) $\mathrm{epi} f^* = \sum_{i=1}^{n} \mathrm{epi} f_i^*$;

(iii) *(Stable Fenchel duality) For any $x^* \in \mathbb{R}^m$ we have*

$$\inf_{x \in \mathbb{R}^m} \{f(x) - \langle x^*, x \rangle\} = \max \left\{ -\sum_{i=1}^{n} f_i^*(x_i^*) : \sum_{i=1}^{n} x_i^* = x^* \right\}.$$

(*iv*) *The infimal convolution is exact for* $f_1, \ldots, f_n$ *in the sense that*

$$\left(\sum_{i=1}^{n} f_i\right)^*(x^*) = (f_1^* \Box \ldots \Box f_n^*)(x^*) \quad \text{for all } x^* \in \mathbb{R}^m.$$

Moreover, if one of the statements (i)–(iv) holds, then the following assertion holds

(*v*) *(Fenchel duality)*

$$\inf_{x\in X} f(x) = \max\left\{-\sum_{i=1}^{n} f_i^*(x_i^*) : \sum_{i=1}^{n} x_i^* = 0\right\}.$$

Lemma 3.2. *Let* $n \in \mathbb{N}$ *and* $f_i \in \Gamma_s(\mathbb{R})$ $(1 \le i \le n)$ *with* $\bigcap_{i=1}^n \mathrm{dom} f_i \neq \emptyset$. *Then we have* $\sum_{i=1}^n \mathrm{epi} f_i^*$ *is closed and hence*

$$\mathrm{epi}\left(\sum_{i=1}^{n} f_i\right)^* = \sum_{i=1}^{n} \mathrm{epi} f_i^*.$$

Proof. Since $\bigcap_{i=1}^n \mathrm{dom} f_i \neq \emptyset$, it follows that $\sum_{i=1}^n f_i$ is a proper lower semicontinuous convex function on $\mathbb{R}$. By induction, we only need to consider the case when $n = 2$. Note that $\overline{\mathrm{dom} f_i}$ $(i = 1, 2)$ are closed convex subsets of $\mathbb{R}$. Without loss of generality, we may assume that $\overline{\mathrm{dom} f_i} = [a_i, b_i]$ for some $a_i, b_i \in \mathbb{R}$ $(i = 1, 2)$ with $b_i > a_i$. Let $\overline{x} \in [a_0, b_0] \cap [a_1, b_1]$ (this is possible since $\mathrm{dom} f_1 \cap \mathrm{dom} f_2 \neq \emptyset$). Denote $[\overline{a}, \overline{b}] := [a_0, b_0] \cap [a_1, b_1]$. If $\overline{a} < \overline{b}$, it follows that $\mathrm{int}(\mathrm{dom} f_1 \cap \mathrm{dom} f_2) \neq \emptyset$ and hence the conclusion follows (see (8)). Therefore, we assume without loss of generality that $\overline{a} = \overline{b}$. In this case, $\mathrm{dom} f_1 \cap \mathrm{dom} f_2 = \{\overline{x}\}$ for some $\overline{x} \in \mathbb{R}$. We may further assume that $b_0 = a_1 = \overline{x}$. Next, we claim that

$$\partial f_1(\overline{x}) + \partial f_2(\overline{x}) = \mathbb{R}. \tag{11}$$

Granting this, it follows that for any $\epsilon \ge 0$

$$\partial_\epsilon (f_1 + f_2)(\overline{x}) \subseteq \mathbb{R} = \partial f_1(\overline{x}) + \partial f_2(\overline{x}) \subseteq \bigcup \{\partial_{\epsilon_0} f_1(\overline{x}) + \partial_{\epsilon_1} f_2(\overline{x}) : \epsilon_0 + \epsilon_1 = \epsilon\}$$

Note that $\partial_\epsilon (f_1 + f_2)(x) = \emptyset$ for any $x \neq \overline{x}$ (since $\mathrm{dom}(f_1 + f_2) = \{\overline{x}\}$). It follows from $(i) \Leftrightarrow (iii)$ of Theorem 3.1 that the conclusion holds.

We now establish (11). Since proper lower semicontinuous function on $\mathbb{R}$ is continuous on the closure of its domain (cf. [13]), we have f_1 is continuous on $[a_0, \overline{x}]$ and f_2 is continuous on $[\overline{x}, b_1]$. This together with $a_0 < \overline{x}$, $\overline{x} < b_1$ and $\partial f_i(\overline{x}) \neq \emptyset$ (by $f_i \in \Gamma_s(\mathbb{R}^m)$) $(i = 1, 2)$ implies that

$$(\partial f_1(\overline{x}))_\infty = N_{\mathrm{dom} f_1}(\overline{x}) = N_{\overline{\mathrm{dom} f_1}}(\overline{x}) = N_{[a_0, \overline{x}]}(\overline{x}) = \mathbb{R}_+$$

and

$$(\partial f_2(\overline{x}))_\infty = N_{\mathrm{dom} f_2}(\overline{x}) = N_{\overline{\mathrm{dom} f_2}}(\overline{x}) = N_{[\overline{x}, b_1]}(\overline{x}) = \mathbb{R}_-.$$

Thus, one has $\partial f_1(\overline{x}) = [a_0^*, +\infty)$ and $\partial f_2(\overline{x}) = (-\infty, a_1^*]$ for some $a_i^* \in \mathbb{R}$ $(i = 1, 2)$. Therefore, (11) holds and this finishes the proof. □

The following example shows that our assumption "each $f_i \in \Gamma_s(\mathbb{R})$" cannot be dropped.

Example 3.3. Consider $f_1(x) = \delta_{(-\infty,-1]}(x)$ and $f_2(x) = \delta_{[-1,1]}(x) - \sqrt{1-x^2}$. It is clear that $\partial f_2(-1) = \emptyset$ but $\partial(f_1 + f_2)(-1) = \mathbb{R}$. Thus the sum rule fails. Therefore, from Theorem 3.1 $(i) \Leftrightarrow (ii)$, we see that $\mathrm{epi} f_1^* + \mathrm{epi} f_2^*$ is not closed.

Theorem 3.4. *Let $n, m \in \mathbb{N}$ and $f_i \in \Gamma_s(\mathbb{R}^m)$ $(1 \le i \le n)$ be separable convex functions on $\mathbb{R}^m$ with $\bigcap_{i=1}^n \mathrm{dom} f_i \neq \emptyset$. Then we have*

$$\mathrm{epi}\left(\sum_{i=1}^n f_i\right)^* = \sum_{i=1}^n \mathrm{epi} f_i^*.$$

Proof. Define $f : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ by

$$f(x) = \sum_{i=1}^n f_i(x) \quad \forall\, x \in \mathbb{R}^m.$$

Since $\bigcap_{i=1}^n \mathrm{dom} f_i \neq \emptyset$, it follows that f is a proper lower semicontinuous convex function on $\mathbb{R}^m$. In view of the second relation of (7), it suffices to show that

$$\mathrm{gph} f^* \subseteq \sum_{i=1}^n \mathrm{epi} f_i^*,$$

where $\mathrm{gph} f^* := \{(x^*, r) : r = f^*(x^*)\}$. To do this, let $(x^*, f^*(x^*)) \in \mathrm{gph} f^*$. Since f_i $(1 \le i \le n)$ are separable convex functions $\mathbb{R}^m$, we may assume that for each $i \in \{1, \dots, n\}$

$$f_i(x) = \sum_{l=1}^m g_l^i(x_l) \quad \forall\, x = (x_1, \dots, x_m) \in \mathbb{R}^m, \tag{12}$$

for some proper lower semicontinuous convex function g_l^i $(1 \le l \le m)$ on $\mathbb{R}$ and hence

$$f(x) = \sum_{i=1}^n \sum_{l=1}^m g_l^i(x_l) = \sum_{l=1}^m \sum_{i=1}^n g_l^i(x_l) \quad \forall\, x = (x_1, \dots, x_m) \in \mathbb{R}^m.$$

Note that each $g_l^i \in \Gamma_s(\mathbb{R})$ $(l \in \{1, \dots, m\}, i \in \{0, \dots, n\})$ since $f_i \in \Gamma_s(\mathbb{R}^m)$. For each $l \in \{1, \dots, m\}$, let $g_l : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be defined by $g_l(x) = \sum_{i=1}^n g_l^i(x)$. Then we have

$$f(x) = \sum_{l=1}^m g_l(x_l) \quad \forall\, x = (x_1, \dots, x_m) \in \mathbb{R}^m.$$

It follows that for any $x^* = (x_1^*, \dots, x_m^*) \in \mathbb{R}^m$,

$$\begin{aligned} f^*(x^*) = \sup_{x \in \mathbb{R}^m} \{\langle x^*, x\rangle - f(x)\} &= \sup_{x=(x_1,\dots,x_m)\in\mathbb{R}^m} \left\{ \sum_{l=1}^m (\langle x_l^*, x_l\rangle - g_l(x_l)) \right\} \\ &= \sum_{l=1}^m \sup_{x_l \in \mathbb{R}} \{\langle x_l^*, x_l\rangle - g_l(x_l)\} = \sum_{l=1}^m g_l^*(x_l^*). \end{aligned} \tag{13}$$

On the other hand, for each $l \in \{1, \dots, m\}$ (by the preceding Lemma)

$$(x_l^*, g_l^*(x_l^*)) \in \mathrm{epi} g_l^* = \mathrm{epi}\left(\sum_{i=1}^n g_l^i\right)^* = \sum_{i=1}^n \mathrm{epi}(g_l^i)^*.$$

Hence, for each $l \in \{1, \dots, m\}$ there exists $(x_l^{i*}, r_l^i) \in \mathrm{epi}(g_l^i)^*$ such that

$$(x_l^*, g_l^*(x_l^*)) = \sum_{i=1}^n (x_l^{i*}, r_l^i). \tag{14}$$

This together with (13) implies that

$$f^*(x^*) = \sum_{l=1}^m g_l^*(x_l^*) = \sum_{l=1}^m \sum_{i=1}^n r_l^i = \sum_{i=1}^n \sum_{l=1}^m r_l^i = \sum_{i=1}^n s_i,$$

where $s_i := \sum_{l=1}^m r_l^i$. Note from $(x_l^i, r_l^i) \in \mathrm{epi}(g_l^i)^*$ and (12) that

$$\begin{aligned} s_i &= \sum_{l=1}^m r_l^i \geq \sum_{l=1}^m (g_l^i)^*(x_l^{i*}) \\ &= \sum_{l=1}^m \sup_{x_i \in \mathbb{R}} \{\langle x_l^{i*}, x_i\rangle - g_l^i(x_i)\} \\ &= \sup_{(x_1,\dots,x_m)\in\mathbb{R}^m} \{\langle (x_1^{i*}, \dots, x_m^{i*}), (x_1, \dots, x_m)\rangle - f_i(x_1, \dots, x_m)\} \\ &= f_i^*(x_1^{i*}, \dots, x_m^{i*}). \end{aligned}$$

It now follows from (14) that $(x^*, f^*(x^*)) \in \sum_{i=1}^n \mathrm{epi} f_i^*$. This completes the proof. $\square$

The following example shows that our assumption, "each f_i is separable", cannot be dropped.

Example 3.5. Consider $I = \{1, 2\}$, $X = \mathbb{R}^2$. Let C_1, C_2 be closed convex subsets of X defined by $C_1 = \overline{B}((0, 1), 1)$ and $C_2 = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \leq 0\}$. Let $f_1 = \delta_{C_1}$, $f_2 = \delta_{C_2}$. Then $f := f_1 + f_2 = \delta_{\{0\}}$ and hence $\inf_{x\in X} f(x) = 0$. Moreover, for any $x^* := (x_1^*, x_2^*) \in \mathbb{R}^m$, we have

$$f_1^*(x^*) = \sigma_{C_1}(x^*) = \|(x_1^*, x_2^*)\| + x_2^* \text{ and } f_2^*(x^*) = \sigma_{C_2}(x^*) = \delta_{\{0\}\times[0,\infty)}(x^*). \tag{15}$$

Noting that $(f_1 + f_2)^* = \sigma_{\{0\}} \equiv 0$, we have $(1, 0, 0) \in \mathrm{epi} f^*$. We claim that $(1, 0, 0) \notin (\mathrm{epi} f_1^* + \mathrm{epi} f_2^*)$. To see this, we proceed by contradiction. Suppose there exist $(a_1^*, a_2^*, r) \in \mathrm{epi} f_1^*$ and $(b_1^*, b_2^*, s) \in \mathrm{epi} f_2^*$ such that

$$(a_1^*, a_2^*, r) + (b_1^*, b_2^*, s) = (1, 0, 0).$$

Note that $b_1^* = 0$ and r, s, b_2^* are all nonnegative (see (15)). It follows that $a_1^* = 1$ and $r = s = 0$. This together with $(a_1^*, a_2^*, r) \in \mathrm{epi} f_1^*$ implies that

$$\sqrt{1 + (a_2^*)^2} + a_2^* \leq 0.$$

However this is impossible and hence $(1, 0, 0) \in \mathrm{epi} f^* \backslash (\mathrm{epi} f_1^* + \mathrm{epi} f_2^*)$. In particular, we have $\mathrm{epi} f^* \neq (\mathrm{epi} f_1^* + \mathrm{epi} f_2^*)$. This together with Theorem 3.1 $(ii) \Leftrightarrow (iii)$ implies that the stable Fenchel duality fails in this case.

Using the preceding Theorem, the following corollary follows directly from the implication $(i) \Leftrightarrow (ii) \Leftrightarrow (iii) \Leftrightarrow (iv)$ in Theorem 3.1.

Corollary 3.6. *Let $n, m \in \mathbb{N}$ and $f_i \in \Gamma_s(\mathbb{R}^m)$ $(1 \leq i \leq n)$ be separable convex functions on $\mathbb{R}^m$ with $\bigcap_{i=1}^n \mathrm{dom} f_i \neq \emptyset$. Let $f := \sum_{i=1}^n f_i$. Then the following statements hold.*

(*1*) *For each $\epsilon \geq 0$ and for each $x \in \mathrm{dom} f$,*

$$\partial_\epsilon f(x) = \bigcup \left\{ \sum_{i=1}^n \partial_{\epsilon_i} f_i(x) : \sum_{i=1}^n \epsilon_i = \epsilon,\ \epsilon_i \geq 0\, (i \in I) \right\}.$$

(*2*) *The infimal convolution is exact for $f_1, \ldots, f_n$ in the sense that*

$$\left(\sum_{i=1}^n f_i \right)^* (x^*) = (f_1^* \square \ldots \square f_n^*)(x^*) \quad \textit{for all } x^* \in \mathbb{R}^m.$$

(*3*) *The following stable Fenchel duality holds: for any $x^* \in \mathbb{R}^m$ one has*

$$\inf_{x \in \mathbb{R}^m} \{f(x) - \langle x^*, x \rangle\} = \max \left\{ -\sum_{i=1}^n f_i^*(x_i^*) : \sum_{i=1}^n x_i^* = x^* \right\}.$$

4. Convex Programs with Separable Constraints

In this section, we establish ϵ-subgradient optimality conditions for convex programming problems with separable inequality constraints without a constraint qualification.

Consider the following primal convex programming problem (P):

$$(P) \qquad v(P) = \inf\{f(x) : g_i(x) \leq 0 \text{ for all } 1 \leq i \leq n\},$$

where f and g_i $(i = 1, \ldots, n)$ are continuous convex functions on $\mathbb{R}^m$. Its dual problem can be written as follows:

$$\begin{aligned} (D) \qquad v(D) &= \sup_{(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n_+} \inf_{x \in \mathbb{R}^m} \left\{ f(x) + \sum_{i=1}^n \lambda_i g_i(x) \right\} \\ &= \sup_{\lambda \in \mathbb{R}^n_+} \inf_{x \in \mathbb{R}^m} \{f(x) + \langle \lambda, g \rangle (x)\} \end{aligned}$$

where $g(x) = (g_1(x), \ldots, g_n(x))$ and $\langle \lambda, g \rangle : \mathbb{R}^m \to \mathbb{R}$ is defined by $\langle \lambda, g \rangle(x) = \sum_{i=1}^n \lambda_i g_i(x)$ $(\lambda \in \mathbb{R}^n)$. It is easy to see that $v(P) \geq v(D)$, i.e., the weak duality always holds. We say the zero duality gap holds if $v(P) = v(D)$. Define the feasible set C by $C := \mathrm{dom} f \cap [u \leq 0])$ where $[u \leq 0] := \{x : u(x) \leq 0\}$ and $u = \max_{1 \leq i \leq n} g_i$. Throughout this section, we always assume $C \neq \emptyset$. We say x is a feasible point of P if $x \in C$. Moreover, we say x is a solution of (P) if x is feasible and $f(x) = v(P)$.

Recently, Tseng [16] established the following zero duality gap result for a separable convex programming problem, which will be used to derive subgradient conditions characterizing optimality.

Lemma 4.1 (cf. [16]). *Let f and g_i $(1 \le i \le n)$ be proper lower semicontinuous separable convex functions with* $\mathrm{dom}f \subseteq \bigcap_{i=1}^n \mathrm{dom}g_i$. *Let* $u := \max_{1\le i\le n} g_i$ *with* $\mathrm{dom}f \bigcap [u \le 0] \neq \emptyset$. *Then*

$$\inf_{x\in[u\le 0]} f(x) = \sup_{\lambda_i\ge 0, 1\le i\le n} \inf_{x\in\mathbb{R}^m} \left\{ f(x) + \sum_{i=1}^n \lambda_i g_i(x) \right\}.$$

Using Lemma 4.1, we are now able to represent the normal cone, $N_{[u\le 0]}(x)$ at x in terms of ϵ-subgradients, which allows us to derive optimality conditions characterizing optimality without a constraint qualification.

Theorem 4.2. *Let g_i $(1 \le i \le n)$ be continuous separable convex functions on $\mathbb{R}^m$ and let $g := (g_1, \ldots, g_n)$. Let $u := \max_{1\le i\le n} g_i$ with $[u \le 0] \neq \emptyset$. Then, for each $x \in [u \le 0]$,*

$$N_{[u\le 0]}(x) = \bigcap_{\epsilon>0} \bigcup_{\substack{\lambda\in\mathbb{R}^n_+ \\ (\langle\lambda,g\rangle)(x)\in[-\epsilon,0]}} \partial_\epsilon(\langle\lambda, g\rangle)(x).$$

Proof. For each $x \in [u \le 0]$, the inclusion

$$N_{[u\le 0]}(x) \supseteq \bigcap_{\epsilon>0} \bigcup_{\substack{\lambda\in\mathbb{R}^n_+ \\ (\langle\lambda,g\rangle)(x)\in[-\epsilon,0]}} \partial_\epsilon(\langle\lambda, g\rangle)(x) \tag{16}$$

always holds. To show the reverse inclusion, let $x \in [u \le 0]$ and let $x^* \in N_{[u\le 0]}(x)$. Then x solves the following minimization problem

$$\begin{array}{lll} (P) & \min & \langle -x^*, z\rangle \\ & s.t. & z \in [u \le 0]. \end{array}$$

Its dual problem is (D) $\sup_{\lambda\in\mathbb{R}^n_+} \inf_{z\in\mathbb{R}^m} \{\langle -x^*, z\rangle + (\langle\lambda, g\rangle)(z)\}$. Since $\langle -x^*, \cdot\rangle$ is affine, $[u \le 0] \neq \emptyset$ and each g_i is continuous on $\mathbb{R}^m$, it follows from Lemma 4.1 that

$$\langle -x^*, x\rangle = \inf_{z\in[u\le 0]} \langle -x^*, z\rangle = \sup_{\lambda\in\mathbb{R}^n_+} \inf_{z\in\mathbb{R}^m} \{\langle -x^*, z\rangle + (\langle\lambda, g\rangle)(z)\}.$$

Then, for each $\epsilon > 0$, there exists $\lambda_\epsilon \in \mathbb{R}^n_+$ such that

$$\langle -x^*, z\rangle + \langle\lambda_\epsilon, g(z)\rangle \ge \langle -x^*, x\rangle - \epsilon/2 \;\text{ for all } z \in \mathbb{R}^m. \tag{17}$$

Letting $z = x$ and noting that $x \in [u \le 0]$, $\lambda_\epsilon \in \mathbb{R}^n_+$, one has $(\langle\lambda, g\rangle)(x) \in [-\epsilon/2, 0]$. This together with (17) implies that for each $z \in \mathbb{R}^m$ $\langle x^*, z - x\rangle \le (\langle\lambda_\epsilon, g\rangle)(z) + \epsilon/2 \le (\langle\lambda_\epsilon, g\rangle)(z) - (\langle\lambda_\epsilon, g\rangle)(x) + \epsilon$. That is to say, $x^* \in \partial_\epsilon(\langle\lambda_\epsilon, g\rangle)(x)$. Therefore, one has

$$x^* \in \bigcap_{\epsilon>0} \bigcup_{\substack{\lambda\in\mathbb{R}^n_+ \\ (\langle\lambda,g\rangle)(x)\in[-\epsilon,0]}} \partial_\epsilon(\langle\lambda, g\rangle)(x).$$

This completes the proof. □

The characterization of optimality now follows from the preceding Lemma.

Theorem 4.3. *Let f be a continuous convex function on $\mathbb{R}^m$. Let g_i $(1 \le i \le n)$ be continuous convex separable functions on $\mathbb{R}^m$. Let a be a feasible point of (P). Then, a is a solution of (P) if and only if there exist sequences $\{\epsilon_k\} \subset \mathbb{R}_+$ and $\{\lambda_k\} \subset \mathbb{R}^n_+$ such that $\epsilon_k \to 0$ and $\lambda^i_k g_i(a) \to 0, (i = 1, \ldots, n)$, as $k \to \infty$, and*

$$\forall k,\ 0 \in \partial f(a) + \sum_{i=1}^{n} \partial_{\epsilon_k}(\lambda^i_k g_i)(a).$$

Proof. [(*1*) $\Rightarrow$ (*2*).] Since a is a solution of the problem (P), $a \in [u \le 0]$ and for all $x \in [u \le 0]$

$$f(a) + \delta_{[u \le 0]}(a) \le f(x) + \delta_{[u \le 0]}(x).$$

So, $0 \in \partial(f + \delta_{[u \le 0]})(a)$. Since f is continuous on $\mathbb{R}^m$, it follows from the standard convex sum rule that

$$0 \in \partial f(a) + N_{[u \le 0]}(a).$$

It now follows from the preceding Lemma that

$$0 \in \partial f(a) + \bigcap_{\epsilon > 0} \bigcup_{\substack{\lambda \in \mathbb{R}^n_+ \\ (\langle \lambda, g \rangle)(a) \in [-\epsilon, 0]}} \partial_\epsilon(\langle \lambda, g \rangle)(a).$$

Let $\{\epsilon_k\}$ be a sequence satisfying $\epsilon_k > 0$ and $\epsilon_k \to 0$. Then, for each $k \in \mathbb{N}$, there exists $\lambda_k = (\lambda^1_k, \ldots, \lambda^n_k) \in \mathbb{R}^n_+$ such that

$$\sum_{i=1}^{n} \lambda^i_k g_i(a) = (\langle \lambda_k, g \rangle)(a) \in [-\epsilon_k, 0]$$

and

$$0 \in \partial f(a) + \partial_{\epsilon_k}(\langle \lambda_k, g \rangle)(a) = \partial f(a) + \sum_{i=1}^{n} \partial_{\frac{\epsilon_k}{n}}(\lambda^i_k g_i)(a),$$

where the last equality follows from Corollary 3.6 (*1*). Note that $\lambda^i_k g_i(a) \le 0$ (since $\lambda^i_k \ge 0$ and a is a solution). Thus statement (*2*) follows.

[(*2*) $\Rightarrow$ (*1*)] Note from Corollary 3.6 (*1*) that

$$\sum_{i=1}^{n} \partial_{\epsilon_k}(\lambda^i_k g_i)(a) = \partial_{n\epsilon_k}\left(\sum_{i=1}^{n} \lambda^i_k g_i\right)(a).$$

This direction follows directly from the definitions of convexity and ϵ-subdifferentials. $\Box$

We now present two examples. The first example illustrates the case where a convex program with separable constraints satisfies our ϵ-subgradient optimality conditions whereas the standard Lagrange multiplier rule fails. The second example shows that the assumption, "g_i is separable", in the preceding theorem cannot be dropped.

Example 4.4. Let $m = n = 2$ and let $f(x_1, x_2) = x_1$ and $g(x_1, x_2) = (g_1(x_1, x_2), g_2(x_1, x_2))$ where $g_1(x_1, x_2) = x_2$ and $g_2(x_1, x_2) = x_1^2 - x_2$. It is clear that f is a real-valued convex function and g_i $(i = 1, 2)$ are real-valued convex separable functions. Let $a_0 := (0, 0)$ and let $\epsilon_k = 1/4k$ and $\lambda_k = (\lambda_k^1, \lambda_k^2) = (k, k)$. It is clear that $\epsilon_k \to 0$ and $\lambda_k \in \mathbb{R}^2_+$ and $(\langle \lambda_k, g\rangle)(a_0) = 0$. Since, for all $\epsilon \geq 0$, $\mu > 0$, $\partial_\epsilon f(a_0) = (1, 0)$, $\partial_\epsilon(\mu g_1)(a_0) = (0, \mu)$ and

$$\partial_\epsilon(\mu g_2)(a_0) = [-2\sqrt{\mu\epsilon}, 2\sqrt{\mu\epsilon}] \times \{-\mu\}. \tag{18}$$

Thus, one has $\partial f(a_0) = (1, 0)$, $\partial_{\epsilon_k}(\lambda_k^1 g_1)(a_0) = (0, k)$ and $\partial_{\epsilon_k}(\lambda_k^2 g_2)(a_0) = [-1, 1] \times \{-k\}$. Therefore, one has

$$(0, 0) = (1, 0) + (0, k) + (-1, -k) \in \partial f(a_0) + \partial_{\epsilon_k}(\lambda_k^1 g_1)(a_0) + \partial_{\epsilon_k}(\lambda_k^2 g_2)(a_0).$$

Thus, from the preceding Theorem, a_0 is a minimizer of the corresponding problem (P). Moreover, we note that the standard Lagrange multiplier rule fails at a_0. Indeed, since $\partial f(a_0) = (1, 0)$, $\partial g_1(a_0) = (0, 1)$ and $\partial g_2(a_0) = (0, -1)$. We see that there do not exist $\lambda_1, \lambda_2 \geq 0$ such that $0 \in \partial f(a_0) + \lambda_1 \partial g_1(a_0) + \lambda_2 \partial g_2(a_0)$.

Example 4.5. Let $m = 2$, $n = 1$. Let $f(x_1, x_2) = 2x_2$ and $g_1(x_1, x_2) = \sqrt{x_1^2 + x_2^2} - x_1$. It is clear that g_1 is not separable. It is clear that $a_0 = (0, 0)$ is the unique feasible point and hence is the unique solution of the corresponding problem (P). Note that $\|\cdot\|$ is sublinear and hence $\partial_\epsilon \|\cdot\|\big((0, 0)\big) = \partial\|\cdot\|\big((0, 0)\big) = \overline{B}((0, 0); 1)$. Thus, for all $\lambda \geq 0$, one has

$$\partial_\epsilon(\lambda g_1)\big((0, 0)\big) := \begin{cases} \lambda\big(\partial_{\epsilon/\lambda}\|\cdot\|\big((0, 0)\big) + (-1, 0)\big) = \lambda\overline{B}((-1, 0); 1), & \text{if } \lambda > 0, \\ (0, 0) & \text{if } \lambda = 0. \end{cases} \tag{19}$$

Thus for any $\epsilon, \lambda \geq 0$, one has

$$\partial f(a_0) + \partial_\epsilon(\lambda g_1)(a_0) \subseteq (0, 2) + \lambda\overline{B}((-1, 0); 1).$$

Since $(0, 0) \notin (0, 2) + \lambda\overline{B}((-1, 0); 1)$ for all $\lambda \geq 0$. It follows that the ϵ-subgradient optimality condition fails in this case.

Acknowledgements. Authors are grateful to the referees for their helpful comments which have contributed to the final preparation of the paper.

References

[1] A. Auslender, M. Teboulle: Asymptotic Cones and Functions in Optimization and Variational Inequalities, Springer Monographs in Mathematics, Springer, New York (2003).

[2] D. P. Bertsekas: Extended monotropic programming and duality, J. Optim. Theor. Appl. 139 (2008) 209–226.

[3] R. I. Bot, S. M. Grad, G. Wanka: On strong and total Lagrange duality for convex optimization problems, J. Math. Anal. Appl 337 (2008) 1315–1325.

[4] R. S. Burachik, V. Jeyakumar, Z. Y. Wu: Necessary and sufficient conditions for stable conjugate duality, Nonlinear Anal., Theory Methods Appl. 64A(9) (2006) 1998–2006.

[5] R. S. Burachik, V. Jeyakumar: A new geometric condition for Fenchel's duality in infinite dimensional spaces, Math. Program. B 104(2–3) (2005) 229–233.

[6] R. S. Burachik, V. Jeyakumar: A dual condition for the convex subdifferential sum formula with applications, J. Convex Analysis 12(2) (2005) 279–290.

[7] V. Jeyakumar: Constraint qualifications characterizing Lagrangian duality in convex optimization, J. Optim. Theor. Appl. 136 (2008) 31-41.

[8] V. Jeyakumar, G. M. Lee, N. Dinh: New sequential Lagrange multiplier conditions characterizing optimality without constraint qualification for convex programs, SIAM J. Optim. 14(2) (2003) 534–547.

[9] V. Jeyakumar, Z. Y. Wu, G. M. Lee, N. Dinh: Liberating the subgradient optimality conditions from constraint qualifications, J. Glob. Optim. 36 (2006) 127–137.

[10] D. Kalpana, K. S. Surjeet, V. Vanita: Convex programming with single separable constraint and bounded variables, Comput. Optim. Appl. 36 (2007) 67–82.

[11] G. Y. Li, K. F. Ng: On extension of Fenchel duality and its application, SIAM J. Optim. 19 (2008) 1489–1509.

[12] S. Simons: Minimax and Monotonicity, Lecture Notes in Mathematics 1693, Springer, Berlin (1998).

[13] R. T. Rockafellar: Convex Analysis, Princeton Univ. Press, Princeton (1970).

[14] S. M. Stefanov: Separable Programming. Theory and Methods, Applied Optimization 53, Kluwer, Dordrecht (2001).

[15] L. Thibault: On subdifferentials of optimal value functions, SIAM J. Control Optimization 29 (1991) 1019–1036.

[16] P. Tseng: Some convex programs without a duality gap, Math. Progam. B 116 (2009) 553–578.