Abstract Convexity for Nonconvex Optimization Duality

Angelia Nedić[*], Asuman Ozdaglar[†] and Alex Rubinov[‡]

February 26, 2007

Abstract

In this paper, we use abstract convexity results to study augmented dual problems for (nonconvex) constrained optimization problems. We consider a nonincreasing function f (to be interpreted as a primal or perturbation function) that is lower semicontinuous at 0 and establish its abstract convexity at 0 with respect to a set of elementary functions defined by nonconvex augmenting functions. We consider three different classes of augmenting functions: nonnegative augmenting functions, bounded-below augmenting functions, and unbounded augmenting functions. We use the abstract convexity results to study augmented optimization duality without imposing boundedness assumptions.

Key words: Abstract convexity, augmenting functions, augmented Lagrangian functions, asymptotic directions, duality gap.

1 Introduction

The analysis of convex optimization duality relies on using linear separation results from convex analysis on the epigraph of the perturbation (primal) function of the optimization problem. This translates into dual problems constructed using traditional Lagrangian functions, which is a linear combination of the objective and constraint functions (see, for example, Rockafellar [13], Hiriart-Urruty and Lemarechal [8], Bonnans and Shapiro [5], Borwein and Lewis [6], Bertsekas, Nedić, and Ozdaglar [3, 4], Auslender and Teboulle [1]). However, linear separation results are not applicable for nonconvex optimization problems, and some recent literature considered *augmented dual problems* (see for example Rockafellar and Wets [14], Huang and Yang [9]). An augmented dual problem is constructed using an *augmented Lagrangian function*, which includes an augmenting

[*]Department of Industrial and Enterprise Systems Engineering, University of Illinois, angelia@uiuc.edu

[†]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, asuman@mit.edu

[‡]University of Ballarat, Centre for Informatics and Applied Optimization, a.rubinov@ballarat.edu.au

function representing a nonlinear penalty for violating the constraints of the problem. Geometrically, this corresponds to using nonlinear surfaces to separate the epigraph of the perturbation function from a point that does not belong to the closure of the epigraph.

The nonlinear separation results have an intimate connection to the more general notion of abstract convexity, which has proven to be a suitable unifying framework for the study of augmented Lagrangian theory in a general setting (see Burachik and Rubinov [7], Rubinov, Glover, and Yang [15], Rubinov [16], Rubinov, Huang, and Yang [17], Rubinov and Yang [18]). In previous work, the augmented optimization duality is investigated under some boundedness assumptions.

Nedić and Ozdaglar presented a geometric approach and a taxonomy of nonlinear separation results that can be used to study augmented optimization duality in [10] and [11]. There, dual problems are constructed using convex augmenting functions, and necessary and sufficient conditions are provided for zero duality gap without explicitly imposing any compactness assumptions.

In this paper, motivated by the development in [10] and [11], we present some zero duality gap results for augmented dual problems constructed with nonconvex augmenting functions, without imposing any boundedness assumptions. We establish these results by using the tools of abstract convexity and some asymptotic properties of the perturbation function of the original constrained problem. In general, the notion of abstract convexity is defined in terms of a prespecified set of elementary functions. More precisely, a function f is said to be *abstract convex* with respect to a given set of *elementary functions* H if f can be represented as the upper envelope of some functions of the set H (cf. Rubinov [16]). Here, we consider two sets of elementary functions denoted by H_σ and $\bar{H}_\sigma$, which are specified in terms of an augmenting function σ. In particular, given an augmenting function σ that satisfies certain properties, we define the sets H_σ and $\bar{H}_\sigma$ respectively by:

$$H_\sigma = \{h \mid h(x) = -r\sigma(x) + c,\ x \in \mathbb{R}^n,\ r \geq 0,\ c \in \mathbb{R}\},$$

$$\bar{H}_\sigma = \left\{h \mid \ h(x) = -\frac{1}{r}\,\sigma(rx) + c,\ x \in \mathbb{R}^n,\ r > 0,\ c \in \mathbb{R}\right\}.$$

We first analyze abstract convexity properties of the perturbation function with respect to the set of elementary functions H_σ or $\bar{H}_\sigma$. We study three different classes of augmenting functions: nonnegative augmenting functions, bounded-below augmenting functions, and unbounded augmenting functions (see Figure 1). We establish that the perturbation function p is abstract convex at 0 with respect to H_σ if σ is a nonnegative augmenting function, and is abstract convex at 0 with respect to $\bar{H}_\sigma$ if σ is a bounded-below or unbounded augmenting function. Contrary to previous studies, we do not assume that the perturbation function is bounded from below in our analysis, but instead use assumptions related to the asymptotic directions of the epigraph of the perturbation function.

We next define the augmented dual problem with arbitrary nonincreasing dualizing parametrizations. We establish an equivalent characterization of zero duality gap between the primal and augmented dual problems in terms of the relation between the values of the perturbation function and its biconjugate at 0. Using a classical result from abstract convex analysis, i.e., the Fenchel-Moreau Theorem, we translate the abstract convexity results on the perturbation function to sufficient conditions for zero duality gap.

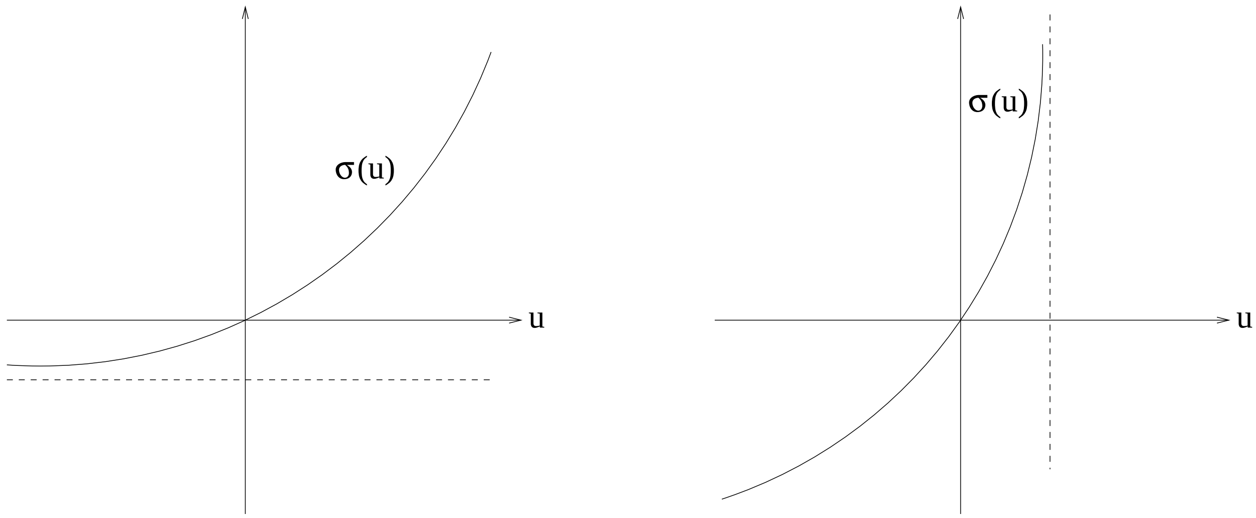

Figure 1: General augmenting functions $\sigma(u)$ for $u \in \mathbb{R}$: The figure to the left illustrates a bounded-below augmenting function, e.g., $\sigma(u) = a(e^u - 1)$ with $a > 0$. The figure to the right illustrates an unbounded augmenting function, e.g., $\sigma(u) = -\log(1-u)$ for $u < 1$.

The rest of the paper is organized as follows: In Section 2 we present some preliminaries from abstract convexity that will be used in our analysis. Section 3 contains our main results and provides various abstract convexity results with respect to sets of elementary functions parametrized by augmenting functions that satisfy certain properties. Section 4 introduces the augmented dual problem and provides sufficient conditions for zero duality gap between the primal problem and the augmented dual problem. Section 5 contains our concluding remarks.

2 Notation, Terminology, and Basics

Consider the n-dimensional space $\mathbb{R}^n$ with the coordinate-wise order relation $\geq$. We view a vector as a column vector, and we denote the inner product of two vectors x and y by $x'y$. We denote the nonpositive orthant in $\mathbb{R}^n$ by $\mathbb{R}^n_-$, i.e., $\mathbb{R}^n_- = \{x \in \mathbb{R}^n \mid x \leq 0\}$.

For any vector $x \in \mathbb{R}^n$, we can write

$$x = x^+ + x^- \qquad \text{with} \;\; x^+ \geq 0 \;\; \text{and} \;\; x^- \leq 0,$$

where the vector x^+ is the component-wise maximum of x, i.e.,

$$x^+ = (\max\{0, x_1\}, ..., \max\{0, x_n\})',$$

and the vector x^- is the component-wise minimum of x, i.e.,

$$x^- = (\min\{0, x_1\}, ..., \min\{0, x_n\})'.$$

For a function $f : \mathbb{R}^n \mapsto [-\infty, \infty]$, we denote the epigraph of f by $\text{epi}(f)$, i.e.,

$$\text{epi}(f) = \{(x, w) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq w\}.$$

We consider sets of functions defined on $\mathbb{R}^n$ with the pointwise order relations: $f_1 \geq f_2$ means that $f_1(x) \geq f_2(x)$ for all $x \in \mathbb{R}^n$. We say that a function $f : \mathbb{R}^n \to [-\infty, \infty]$ is *non-increasing* if $x \geq y$ implies $f(x) \leq f(y)$.

Definition 1 (*Abstract Convexity at a Point*) Let H be a set of extended real-valued proper functions $h : \mathbb{R}^n \mapsto (-\infty, \infty]$. We say that a function $f : \mathbb{R}^n \mapsto (-\infty, \infty]$ is *abstract convex at a point* $\bar{x} \in \mathbb{R}^n$ *with respect to* H when the following relation holds

$$f(\bar{x}) = \sup\{h(\bar{x}) \mid h \in H, \ h \leq f\}.$$

In this paper, we are interested in abstract convexity with respect to special classes of functions

$$H_\sigma = \{h \mid h(x) = -r\sigma(x) + c, \ x \in \mathbb{R}^n, \ r \geq 0, \ c \in \mathbb{R}\} \tag{1}$$

or

$$\bar{H}_\sigma = \left\{h \mid \ h(x) = -\frac{1}{r}\,\sigma(rx) + c, \ x \in \mathbb{R}^n, \ r > 0, \ c \in \mathbb{R}\right\} \tag{2}$$

specified in terms of an augmenting function σ. In particular, we define an augmenting function as follows:

Definition 2 A function $\sigma : \mathbb{R}^n \mapsto (-\infty, \infty]$ is called an *augmenting function* if it is not identically equal to 0 and it takes the zero value at the origin, i.e.,

$$\sigma \neq 0 \qquad \text{and} \qquad \sigma(0) = 0.$$

This definition of an augmenting function is motivated by the convex augmenting functions introduced by Rockafellar and Wets [14] (see Definition 11.55); however note that we do not restrict ourselves to convex functions here. By definition, an augmenting function is a proper function.

When establishing abstract convexity results for functions that are unbounded from below, we use the notion of an asymptotic cone of a set. In particular, the asymptotic cone of a set C is denoted by C^∞ and is defined as follows.

Definition 3 (*Asymptotic Cone*) The asymptotic cone C^∞ of a nonempty set C is given by

$$C^\infty = \{d \mid \lambda_k x_k \to d \ \text{ for some } \{x_k\} \subset C \text{ and } \{\lambda_k\} \subset \mathbb{R} \text{ with } \lambda_k \geq 0, \ \lambda_k \to 0\}.$$

A direction $d \in C^\infty$ is referred to as a *asymptotic direction* of the set C.

3 Main Results

In this section, we discuss sufficient conditions on augmenting functions σ and the function f that guarantee abstract convexity of f with respect to a set H_σ or $\bar{H}_\sigma$, defined in (1) or (2), respectively. We establish these sufficient conditions by separating the epigraph epi(f) of the function f and the half-line $\{(0, w) \mid w \leq f(0) - \epsilon\}$ for some $\epsilon > 0$. The separation of these two sets is realized through some augmenting function σ.

For the separation results, an important characteristic of the function f is the "bottom-shape" of the epigraph of f. In particular, it is desirable that f does not decrease faster than a linear function i.e., the ratio of $f(x)$ and $\|x\|$-values is asymptotically finite, as $f(x)$ decreases to infinity. To characterize this, we use the notion of asymptotic directions and asymptotic cone of a nonempty set (see Section 2). In particular, we impose the condition that the direction $(0, -1)$ is not an asymptotic direction of epi(f), i.e.,

$$(0, -1) \notin (\text{epi}(f))^\infty.$$

More precisely, we consider functions f that satisfy the following assumption.

Assumption 1 Let $f : \mathbb{R}^n \mapsto (-\infty, \infty]$ be a function with the following properties:

(a) The function f is nonincreasing and the value $f(0)$ is finite.

(b) The function f is lower semicontinuous at $x = 0$, i.e., for all sequences $\{x_k\} \subset \mathbb{R}^n$ with $x_k \to 0$, we have

$$f(0) \leq \liminf_{k\to\infty} f(x_k).$$

(c) The vector $(0, -1)$ is not an asymptotic direction of epi(f), i.e.,

$$(0, -1) \notin (\text{epi}(f))^\infty.$$

As mentioned earlier, Assumption 1(c) plays a crucial role in establishing the separation of the epigraph of the function f and the half-line $\{(0, w) \mid w \leq f(0) - \epsilon\}$ for some $\epsilon > 0$. To provide more insights into Assumption 1(c), we give a simpler equivalent characterization of the relation $(0, -1) \notin (\text{epi}(f))^\infty$ in the following lemma.

Lemma 1 Let $f : \mathbb{R}^n \mapsto (-\infty, \infty]$ be a function. Then $(0, -1) \notin (\text{epi}(f))^\infty$ if and only if for any sequence $\{x_k\} \subset \mathbb{R}^n$ with $f(x_k) \to -\infty$, we have

$$\liminf_{k\to\infty} \frac{f(x_k)}{\|x_k\|} > -\infty.$$

Proof. Assume first that $(0, -1) \notin (\text{epi}(f))^\infty$. Furthermore, assume to arrive at a contradiction that there exists a sequence $\{x_k\} \subset \mathbb{R}^n$ with $f(x_k) \to -\infty$ such that

$$\liminf_{k\to\infty} \frac{f(x_k)}{\|x_k\|} = -\infty.$$

By restricting attention to a subsequence if necessary, we can assume without loss of generality that $\frac{f(x_k)}{\|x_k\|} \to -\infty$. Note that we can write

$$\left(\frac{x_k}{|f(x_k)|}, \frac{f(x_k)}{|f(x_k)|} \right) = \left(\frac{x_k}{\|x_k\|} \frac{\|x_k\|}{|f(x_k)|}, \frac{f(x_k)}{|f(x_k)|} \right).$$

Taking the limit as $k \to \infty$ in the preceding relation and using the fact $\frac{f(x_k)}{\|x_k\|} \to -\infty$, we obtain

$$\lim_{k\to\infty} \left(\frac{x_k}{|f(x_k)|}, \frac{f(x_k)}{|f(x_k)|} \right) = (0, -1).$$

Since $\frac{1}{|f(x_k)|} \to 0$, it follows by the definition of an asymptotic direction (cf. Definition 3) that $(0, -1) \in (\mathrm{epi}(f))^\infty$ – a contradiction.

Assume next that $(0, -1) \in (\mathrm{epi}(f))^\infty$. We show that there exists a sequence $\{x_k\}$ with $f(x_k) \to -\infty$ such that

$$\liminf_{k\to\infty} \frac{f(x_k)}{\|x_k\|} = -\infty.$$

Since $(0, -1) \in (\mathrm{epi}(f))^\infty$, by the definition of an asymptotic cone (cf. Definition 3), there exist a scalar sequence $\{\lambda_k\}$ with $\lambda_k \geq 0$ and $\lambda_k \to 0$, and a vector sequence $\{(x_k, y_k)\} \subset \mathrm{epi}(f)$ such that

$$\lim_{k\to\infty} \lambda_k(x_k, y_k) = (0, -1).$$

Since $\lambda_k \to 0$ and $\lambda_k y_k \to -1$, we have that

$$\lim_{k\to\infty} y_k = \lim_{k\to\infty} \frac{-1}{\lambda_k} = -\infty.$$

Furthermore, since $\lambda_k x_k \to 0$ and $\lambda_k y_k \to -1$, we obtain

$$\liminf_{k\to\infty} \frac{\lambda_k y_k}{\lambda_k \|x_k\|} = \liminf_{k\to\infty} \frac{y_k}{\|x_k\|} = -\infty.$$

Because $\{(x_k, y_k)\} \subset \mathrm{epi}(f)$, we have $f(x_k) \leq y_k$, which together with the preceding relation yields $\liminf_{k\to\infty} \frac{f(x_k)}{\|x_k\|} = -\infty$, thus completing the proof. **Q.E.D.**

By using Lemma 1, we can see that $(0, -1)$ is not an asymptotic direction of f when one of the following holds:

(1) The function f is bounded from below over $\mathbb{R}^n$, i.e., $\inf_{x\in\mathbb{R}^n} f(x) > -\infty$,

(2) For some vectors $a_i \in \mathbb{R}^n$ and scalars b_i, $i = 1, \ldots, r$, the function f is given by

$$f(x) = \max_{1\leq i\leq r} \{a_i'x + b_i\} \qquad \text{for all } x,$$

or

$$f(x) = \min_{1\leq i\leq r} \{a_i'x + b_i\} \qquad \text{for all } x.$$

(3) For some vectors $a_i \in \mathbb{R}^n$ and scalars b_i, $i = 1, \ldots, r$, the function f satisfies

$$f(x) \geq \max_{1\leq i\leq r} \{a_i'x + b_i\} \qquad \text{for all } x,$$

or

$$f(x) \geq \min_{1\leq i\leq r} \{a_i'x + b_i\} \qquad \text{for all } x.$$

3.1 Nonnegative Augmenting Functions

Here, we establish an abstract convexity result for an augmenting function σ that is nonnegative. In particular, we consider a class of augmenting functions σ satisfying the following assumption.

Assumption 2 Let σ be an augmenting function with the following properties:

(a) The function σ is nonnegative,

$$\sigma(x) \geq 0 \qquad \text{for all } x.$$

(b) For any sequence $\{x_k\} \subset \mathbb{R}^n$, the convergence of $\sigma(x_k)$ to zero implies the convergence of the nonnegative part of the sequence $\{x_k\}$ to zero, i.e.,

$$\sigma(x_k) \to 0 \quad \Rightarrow \quad x_k^+ \to 0.$$

(c) For any sequence $\{x_k\} \subset \mathbb{R}^n$ and any positive scalar sequence $\{\lambda_k\}$ with $\lambda_k \to \infty$, if the relation $\lim_{k\to\infty} \frac{\sigma(\lambda_k x_k)}{\lambda_k} = 0$ holds, then the nonnegative part of the sequence $\{x_k\}$ converges to zero, i.e.,

$$\lim_{k\to\infty} \frac{\sigma(\lambda_k x_k)}{\lambda_k} = 0 \ \text{ with } \{x_k\} \subset \mathbb{R}^n \text{ and } \lambda_k \to \infty \quad \Rightarrow \quad x_k^+ \to 0.$$

It can be seen that Assumption 2(b) is equivalent to the following condition: for all $\delta > 0$, there holds

$$\inf_{\{x \ \mid \ \mathrm{dist}(x, \mathbb{R}^n_-) \geq \delta\}} \sigma(x) > 0. \tag{3}$$

To see this, assume first that Assumption 2(b) holds and assume to arrive at a contradiction that there exists some $\delta > 0$ such that

$$\inf_{\{x \ \mid \ \mathrm{dist}(x, \mathbb{R}^n_-) \geq \delta\}} \sigma(x) = 0.$$

This implies that there exists a sequence $\{x_k\}$ such that $\sigma(x_k) \to 0$ and $\|x_k^+\| \geq \delta$ for all k, contradicting Assumption 2(b). Conversely, assume that condition (3) holds. Let $\{x_k\}$ be a sequence with $\sigma(x_k) \to 0$, and assume that $\limsup_{k\to\infty} \|x_k^+\| > 0$. This implies the existence of some $\delta > 0$ such that along a subsequence, we have $\mathrm{dist}(x_k, \mathbb{R}^n_-) > \delta$ for all k sufficiently large. Since $\sigma(x_k) \to 0$, this contradicts condition (3).

Assumption 2(b) is related to the *peak at zero condition*, which can be expressed as follows: for all $\delta > 0$, there holds

$$\inf_{\{x \ \mid \ \|x\| \geq \delta\}} \sigma(x) > 0.$$

This condition was studied by Rubinov et al. [17] to provide zero duality gap results for arbitrary dualizing parametrizations.

The following are some examples of augmenting functions σ that satisfy Assumption 2:

$$\sigma(x) = \|x\|_p^\gamma \qquad \text{or} \qquad \sigma(x) = \|x^+\|_p^\gamma \tag{4}$$

for some scalars $\gamma \geq 1$ and p with $0 < p \leq \infty$, where $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ for $p < \infty$ and $\|x\|_\infty = \max_i |x_i|$ for $p = \infty$;

$$\sigma(x) = \|Ax\|_p^\gamma \qquad \text{or} \qquad \sigma(x) = \|Ax^+\|_p^\gamma \tag{5}$$

for a scalar $\gamma \geq 1$ and an m by n matrix A with a full column rank;

$$\sigma(x) = (x'Qx)^\gamma \qquad \text{or} \qquad \sigma(x) = \big((x^+)'Qx^+\big)^\gamma \tag{6}$$

for a scalar $\gamma \geq 1/2$ and a symmetric positive definite n by n matrix Q;

$$\sigma(x) = |x_1|^{\gamma_1}|x_2|^{\gamma_2}\cdots|x_n|^{\gamma_n} + \sigma_1(x) \tag{7}$$

or

$$\sigma(x) = (x_1^+)^{\gamma_1}(x_2^+)^{\gamma_2}\cdots(x_n^+)^{\gamma_n} + \sigma_1(x) \tag{8}$$

for some scalars $\gamma_1 \geq 0, \ldots, \gamma_n \geq 0$ with $\gamma_1 + \cdots + \gamma_n \geq 1$, and for a function σ_1 being one of the preceding examples of augmenting functions given in Eqs. (4)–(6).

We note that the augmenting functions given in Eqs. (4)–(5) are nonconvex for $p < 1$. Furthermore, the augmenting functions of the form as in Eqs. (7)–(8) can also be nonconvex. As a simple example, consider the case when σ_1 is convex but $\gamma_1 = \gamma_2 = 1/2$ and $\gamma_3 = \ldots = \gamma_n = 0$, in which case the functions $x \to |x_1|^{\gamma_1}|x_2|^{\gamma_2}\cdots|x_n|^{\gamma_n}$ and $x \to (x_1^+)^{\gamma_1}(x_2^+)^{\gamma_2}\cdots(x_n^+)^{\gamma_n}$ are nonconvex.

We now provide an abstract convexity result for augmenting functions that satisfy Assumption 2.

Proposition 1 Let $f : \mathbb{R}^n \mapsto (-\infty, \infty]$ be a function that satisfies Assumption 1 and let σ be an augmenting function that satisfies Assumption 2. Then, for every $\epsilon > 0$, there exist scalars c and $\bar{r} > 0$ such that

$$f(x) + r\sigma(x) \geq c > f(0) - \epsilon \qquad \text{for all } x \in \mathbb{R}^n \text{ and all } r \geq \bar{r}. \tag{9}$$

As a particular consequence of the preceding relation, we have that the function f is abstract convex at $x = 0$ with respect to H_σ, where

$$H_\sigma = \{h \mid h(x) = -r\sigma(x) + c,\ x \in \mathbb{R}^n,\ r \geq 0,\ c \in \mathbb{R}\}.$$

Proof. Assume to arrive at a contradiction that relation (9) does not hold. Then, there exist a positive scalar sequence $\{r_k\}$ with $r_k \to \infty$ and a vector sequence $\{x_k\} \subset \mathbb{R}^n$ such that

$$f(x_k) + r_k\sigma(x_k) \leq f(0) - \epsilon \qquad \text{for all } k. \tag{10}$$

Because of the nonnegativity of σ [cf. Assumption 2(a)], it follows that

$$\liminf_{k\to\infty} f(x_k) \leq f(0) - \epsilon. \tag{11}$$

We now consider separately the following two cases: the sequence $\{f(x_k)\}$ is bounded from below, and $\{f(x_k)\}$ is unbounded from below.

Case 1: The sequence $\{f(x_k)\}$ is bounded from below.

We have $f(x_k) \geq K$ for some scalar K and for all k. Then, from Eq. (10) and the nonnegativity of the augmenting function σ [cf. Assumption 2(a)], it follows that

$$0 \leq \sigma(x_k) \leq \frac{f(0) - \epsilon - f(x_k)}{r_k} \leq \frac{f(0) - \epsilon - K}{r_k} \qquad \text{for all } k.$$

Since $r_k \to \infty$, the preceding relation implies that $\sigma(x_k) \to 0$. Therefore, by Assumption 2(b), it follows that $x_k^+ \to 0$. Furthermore, since $x_k \leq x_k^+$ and the function f is nonincreasing [cf. Assumption 1(a)], we have $f(x_k^+) \leq f(x_k)$ for all k. Combining these with the assumption that f is lower semicontinuous at 0 [cf. Assumption 1(b)], we obtain

$$f(0) \leq \liminf_{k\to\infty} f(x_k^+) \leq \liminf_{k\to\infty} f(x_k).$$

Since $\liminf_{k\to\infty} f(x_k) \leq f(0) - \epsilon$ [cf. Eq. (11)], this yields a contradiction.

Case 2: The sequence $\{f(x_k)\}$ is unbounded from below.

Assume without loss of generality that $f(x_k) \to -\infty$, and consider the sequence $\{x_k^+\}$. Since $x_k \leq x_k^+$ for all k and the function f is nonincreasing, it follows that $f(x_k^+) \leq f(x_k)$ for all k. Because $f(x_k) \to -\infty$, we have $f(x_k^+) \to -\infty$.

Suppose that the sequence $\{x_k^+\}$ is bounded. Then, we have

$$\liminf_{k\to\infty} \frac{f(x_k^+)}{\|x_k^+\|} = -\infty.$$

By Lemma 1, it follows that $(0, -1) \in (\mathrm{epi}(f))^\infty$, thus contradicting Assumption 1(c). Hence, the sequence $\{x_k^+\}$ must be unbounded, and without loss of generality, we may assume that $\|x_k^+\| \to \infty$ with $\|x_k^+\| > 0$ for all k.

Dividing by $\|x_k^+\|$ in Eq. (10) and using the fact $f(x_k^+) \leq f(x_k)$ for all k, we obtain

$$\frac{f(x_k^+)}{\|x_k^+\|} + r_k \frac{\sigma(x_k)}{\|x_k^+\|} \leq \frac{f(0) - \epsilon}{\|x_k^+\|}.$$

By rearranging the terms and taking the limit superior as $k \to \infty$ in the preceding relation, we obtain

$$\limsup_{k\to\infty} r_k \frac{\sigma(x_k)}{\|x_k^+\|} \leq -\liminf_{k\to\infty} \frac{f(x_k^+)}{\|x_k^+\|}.$$

By Assumption 1(c) and Lemma 1, we have $\liminf_{k\to\infty} \frac{f(x_k^+)}{\|x_k^+\|} > -\infty$, implying that

$$\limsup_{k\to\infty} r_k \frac{\sigma(x_k)}{\|x_k^+\|} < \infty.$$

Since $r_k \to \infty$, it further follows that $\limsup_{k\to\infty} \frac{\sigma(x_k)}{\|x_k^+\|} \leq 0$, which by the nonnegativity of the augmenting function σ [cf. Assumption 2(a)] yields $\frac{\sigma(x_k)}{\|x_k^+\|} \to 0$. Therefore,

$$\lim_{k\to\infty} \frac{\sigma(\lambda_k v_k)}{\lambda_k} = 0 \qquad \text{with} \;\; \lambda_k = \|x_k^+\| \;\; \text{and} \;\; v_k = \frac{x_k}{\|x_k^+\|},$$

where $\|x_k^+\| \to \infty$. Hence, from Assumption 2(c) and the preceding relations, we have $v_k^+ \to 0$, implying that $\frac{x_k^+}{\|x_k^+\|} \to 0$ - a contradiction. Hence, relation (9) holds.

As a particular consequence of relation (9), we have that

$$f(x) + \bar{r}\sigma(x) \geq c > f(0) - \epsilon \qquad \text{for all } x \in \mathbb{R}^n.$$

By letting $h_{\bar{r},c}(x) = -\bar{r}\sigma(x) + c$ for all x, the inequality $f(x) + \bar{r}\sigma(x) \geq c$ for all x is equivalent to the following

$$f(x) \geq h_{\bar{r},c}(x) \qquad \text{for all } x \in \mathbb{R}^n,$$

i.e., $h_{\bar{r},c} \leq f$ with $h_{\bar{r},c} \in H_\sigma$, where $H_\sigma = \{h \mid h(x) = -r\sigma(x) + c,\; x \in \mathbb{R}^n,\; r \geq 0,\; c \in \mathbb{R}\}$. By the definition of the augmenting function [cf. Definition 2], we have $\sigma(0) = 0$, implying that $h_{\bar{r},c}(0) = c$. Hence, the relation $c > f(0) - \epsilon$ is equivalent to $h_{\bar{r},c}(0) > f(0) - \epsilon$. Since such scalars $\bar{r}$ and c exist for any $\epsilon > 0$, it follows that f is abstract convex at $x = 0$ with respect to H_σ. **Q.E.D.**

We note here that the analysis of Case 1 in the preceding proof does not use Assumption 2(c) on augmenting functions.

3.2 Bounded-Below Augmenting Functions

In this section, we establish an abstract convexity result for an augmenting function σ that is bounded from below but not necessarily nonnegative. In particular, we consider augmenting functions σ satisfying the following assumption.

Assumption 3 Let σ be an augmenting function with the following properties:

(a) The function σ is bounded-below, i.e.,

$$\sigma(x) \geq \sigma_0 \quad \text{for some scalar } \sigma_0 \text{ and for all } x.$$

(b) For any sequence $\{x_k\} \subset \mathbb{R}^n$ and any positive scalar sequence $\{\lambda_k\}$ with $\lambda_k \to \infty$, if the relation $\limsup_{k\to\infty} \frac{\sigma(\lambda_k x_k)}{\lambda_k} < \infty$ holds, then the nonnegative part of the sequence $\{x_k\}$ converges to zero, i.e.,

$$\limsup_{k\to\infty} \frac{\sigma(\lambda_k x_k)}{\lambda_k} < \infty \;\; \text{with } \{x_k\} \subset \mathbb{R}^n \text{ and } \lambda_k \to \infty \quad \Rightarrow \quad x_k^+ \to 0.$$

Clearly, the examples of nonnegative augmenting functions given in Eqs. (4)–(8) satisfy Assumption 3(a) with $\sigma_0 = 0$. Furthermore, it can be seen that these functions also satisfy Assumption 3(b). Also, the following is an augmenting function that satisfies Assumption 3:

$$\sigma(x) = a_1(e^{x_1} - 1) + \cdots + a_n(e^{x_n} - 1)$$

for some scalars $a_1 > 0, \ldots, a_n > 0$.

We next provide an abstract convexity result for bounded-below augmenting functions that satisfy Assumption 3.

Proposition 2 Let $f : \mathbb{R}^n \mapsto (-\infty, \infty]$ be a function that satisfies Assumption 1. Let σ be an augmenting function that satisfies Assumption 3. Then, for every $\epsilon > 0$, there exist scalars c and $\bar{r} > 0$ such that

$$f(x) + \frac{1}{r}\sigma(rx) \geq c > f(0) - \epsilon \qquad \text{for all } x \in \mathbb{R}^n \text{ and all } r \geq \bar{r}. \tag{12}$$

As a particular consequence, we have that the function f is abstract convex at $x = 0$ with respect to $\bar{H}_\sigma$, where

$$\bar{H}_\sigma = \left\{ h \mid h(x) = -\frac{1}{r}\sigma(rx) + c, \ x \in \mathbb{R}^n, \ r > 0, \ c \in \mathbb{R} \right\}.$$

Proof. Assume to arrive at a contradiction that relation (12) does not hold. Then, there exist a positive scalar sequence $\{r_k\}$ with $r_k \to \infty$ and a vector sequence $\{x_k\} \subset \mathbb{R}^n$ such that

$$f(x_k) + \frac{1}{r_k}\sigma(r_k x_k) \leq f(0) - \epsilon \qquad \text{for all } k. \tag{13}$$

Because $\sigma(x) \geq \sigma_0$ [cf. Assumption 3(a)] and $r_k \to \infty$, it follows that

$$\liminf_{k \to \infty} f(x_k) \leq f(0) - \epsilon. \tag{14}$$

Now, consider separately the following two cases: the sequence $\{f(x_k)\}$ is bounded from below, and $\{f(x_k)\}$ is unbounded from below.

Case 1: The sequence $\{f(x_k)\}$ is bounded from below.

We have $f(x_k) \geq K$ for some scalar K and all k. Then, from Eq. (13) it follows that

$$\limsup_{k \to \infty} \frac{\sigma(r_k x_k)}{r_k} \leq \limsup_{k \to \infty} \Big(f(0) - \epsilon - f(x_k) \Big) \leq f(0) - \epsilon - K < \infty.$$

By Assumption 3(b), the preceding relation implies that $x_k^+ \to 0$. Furthermore, since $x_k \leq x_k^+$ and the function f is nonincreasing, we have $f(x_k^+) \leq f(x_k)$ for all k. Combining these relations with the assumption that f is lower semicontinuous at $x = 0$ [cf. Assumption 1(b)], we obtain

$$f(0) \leq \liminf_{k \to \infty} f(x_k^+) \leq \liminf_{k \to \infty} f(x_k).$$

Since $\liminf_{k\to\infty} f(x_k) \le f(0) - \epsilon$ [cf. Eq. (14)], this yields a contradiction.

Case 2: The sequence $\{f(x_k)\}$ is unbounded from below.

Assume without loss of generality that $f(x_k) \to -\infty$, and consider the sequence $\{x_k^+\}$. Since $x_k \le x_k^+$ for all k and the function f is nonincreasing, it follows that $f(x_k^+) \le f(x_k)$ for all k. Because $f(x_k) \to -\infty$, we have $f(x_k^+) \to -\infty$.

Suppose that the sequence $\{x_k^+\}$ is bounded. Then, we have

$$\liminf_{k\to\infty} \frac{f(x_k^+)}{\|x_k^+\|} = -\infty.$$

By Lemma 1, it follows that $(0, -1) \in (\mathrm{epi}(f))^\infty$, thus contradicting Assumption 1(c). Hence, the sequence $\{x_k^+\}$ must be unbounded, and without loss of generality, we may assume that $\|x_k^+\| \to \infty$ with $\|x_k^+\| > 0$ for all k.

By using the fact $f(x_k^+) \le f(x_k)$ for all k, and dividing by $\|x_k^+\|$ in Eq. (13), we obtain

$$\frac{f(x_k^+)}{\|x_k^+\|} + \frac{\sigma(r_k x_k)}{r_k\|x_k^+\|} \le \frac{f(0) - \epsilon}{\|x_k^+\|}. \tag{15}$$

Since $f(x_k^+) \to -\infty$, by Assumption 1(c) and Lemma 1, we have $\liminf_{k\to\infty} \frac{f(x_k^+)}{\|x_k^+\|} > -\infty$. By rearranging the terms in Eq. (15) and by taking the limit superior as $k \to \infty$, we further obtain

$$\limsup_{k\to\infty} \frac{\sigma(r_k x_k)}{r_k\|x_k^+\|} \le -\liminf_{k\to\infty} \frac{f(x_k^+)}{\|x_k^+\|} < \infty.$$

Therefore,

$$\limsup_{k\to\infty} \frac{\sigma(\lambda_k v_k)}{\lambda_k} < \infty \qquad \text{with } \lambda_k = r_k\|x_k^+\| \text{ and } v_k = \frac{x_k}{\|x_k^+\|}.$$

Since $\lambda_k \to \infty$, by Assumption 3(b) we have $v_k^+ \to 0$, implying that $\frac{x_k^+}{\|x_k^+\|} \to 0$ - a contradiction. Thus, relation (12) holds.

As a special case of relation (12), we have that

$$f(x) + \frac{1}{\bar{r}}\sigma(\bar{r}x) \ge c > f(0) - \epsilon \qquad \text{for all } x \in \mathbb{R}^n.$$

Let $h_{\bar{r},c}(x) = -\frac{1}{\bar{r}}\sigma(\bar{r}x) + c$ for all x. Then, the inequality $f(x) + \frac{1}{\bar{r}}\sigma(\bar{r}x) \ge c$ for all x is equivalent to the following

$$f(x) \ge h_{\bar{r},c}(x) \qquad \text{for all } x \in \mathbb{R}^n,$$

i.e., $h_{\bar{r},c} \le f$ with $h_{\bar{r},c} \in \bar{H}_\sigma$ for $\bar{H}_\sigma = \big\{h \mid h(x) = -\frac{1}{r}\sigma(rx) + c,\ x \in \mathbb{R}^n,\ r > 0,\ c \in \mathbb{R}\big\}$. By the definition of the augmenting function [cf. Definition 2], we have $\sigma(0) = 0$, so that $h_{\bar{r},c}(0) = c$. Thus, the relation $c > f(0) - \epsilon$ is equivalent to $h_{\bar{r},c}(0) > f(0) - \epsilon$. Since such scalars $\bar{r}$ and c exist for any $\epsilon > 0$, it follows that f is abstract convex at $x = 0$ with respect to $\bar{H}_\sigma$. **Q.E.D.**

3.3 Unbounded Augmenting Functions

In this section, we present an abstract convexity result for an augmenting function σ that is unbounded from below. In particular, we consider a class of augmenting functions σ satisfying the following assumption.

Assumption 4 Let σ be an augmenting function with the following properties:

(a) For any sequence $\{x_k\} \subset \mathbb{R}^n$ with $x_k \to \bar{x}$ and for any positive scalar sequence $\{\lambda_k\}$ with $\lambda_k \to \infty$, the relation $\limsup_{k\to\infty} \frac{\sigma(\lambda_k x_k)}{\lambda_k} < \infty$ implies that the vector $\bar{x}$ is nonpositive, i.e.,

$$\limsup_{k\to\infty} \frac{\sigma(\lambda_k x_k)}{\lambda_k} < \infty \text{ with } x_k \to \bar{x} \text{ and } \lambda_k \to \infty \qquad \Rightarrow \qquad \bar{x} \leq 0.$$

(b) For any sequence $\{x_k\} \subset \mathbb{R}^n$ with $x_k \to \bar{x}$ and $\bar{x} \leq 0$, and for any positive scalar sequence $\{\lambda_k\}$ with $\lambda_k \to \infty$, we have

$$\liminf_{k\to\infty} \frac{\sigma(\lambda_k x_k)}{\lambda_k} \geq 0.$$

Here, we note that the augmenting functions given in Eqs. (4)–(8) satisfy Assumption 4, some of which are nonconvex as discussed there. Also, Assumption 4 is satisfied for an augmenting function σ of the form (see Nedić and Ozdaglar [11]):

$$\sigma(x) = \sum_{i=1}^{n} \theta(x_i)$$

with the following choices of the scalar function θ:

$$\theta(t) = \begin{cases} -\log(1-t) & t < 1, \\ +\infty & t \geq 1, \end{cases}$$

(cf. modified barrier method of Polyak [12]),

$$\theta(t) = \begin{cases} \frac{t}{1-t} & t < 1, \\ +\infty & t \geq 1, \end{cases}$$

(cf. hyperbolic modified barrier method of Polyak [12]),

$$\theta(t) = \begin{cases} t + \frac{1}{2}t^2 & t \geq -\frac{1}{2}, \\ -\frac{1}{4}\log(-2t) - \frac{3}{8} & t < -\frac{1}{2}, \end{cases}$$

(cf. quadratic logarithmic method of Ben-Tal and Zibulevski [2]).

To establish an abstract convexity result for an augmenting function that may be unbounded from below, we use an additional assumption on the function f.

Assumption 5 For any $\bar{x} \in \mathbb{R}^n$ with $\bar{x} \leq 0$ and $\bar{x} \neq 0$, the vector $(\bar{x}, 0)$ is not an asymptotic direction of $\text{epi}(f)$, i.e.,

$$(\bar{x}, 0) \notin (\text{epi}(f))^\infty \qquad \text{for any } \bar{x} \leq 0 \text{ with } \bar{x} \neq 0.$$

For example, a decreasing function f satisfies the preceding assumption. Informally speaking, any nonincreasing function f that is not "flat" along any rays of the form $\{x + \lambda\bar{x} \mid \lambda \geq 0\}$, for $x, \bar{x} \in \mathbb{R}^n$ with $\bar{x} \leq 0$ and $\bar{x} \neq 0$, satisfies the preceding assumption.

We next state our abstract convexity result for unbounded augmenting functions. The proof uses a similar line of analysis to that of Proposition 2.

Proposition 3 Let $f : \mathbb{R}^n \mapsto (-\infty, \infty]$ be a function that satisfies Assumption 1 and Assumption 5. Let σ be an augmenting function that satisfies Assumption 4. Then, for every $\epsilon > 0$, there exist scalars c and $\bar{r} > 0$ such that

$$f(x) + \frac{1}{r}\,\sigma(rx) \geq c > f(0) - \epsilon \qquad \text{for all } x \in \mathbb{R}^n \text{ and all } r \geq \bar{r}. \tag{16}$$

As a special consequence of the preceding relation, we have that the function f is abstract convex at $x = 0$ with respect to $\bar{H}_\sigma$, where

$$\bar{H}_\sigma = \left\{ h \mid h(x) = -\frac{1}{r}\,\sigma(rx) + c,\ x \in \mathbb{R}^n,\ r > 0,\ c \in \mathbb{R} \right\}.$$

Proof. Assume to arrive at a contradiction that relation (16) does not hold. Then, there exist a positive scalar sequence $\{r_k\}$ with $r_k \to \infty$ and a vector sequence $\{x_k\} \subset \mathbb{R}^n$ such that

$$f(x_k) + \frac{1}{r_k}\,\sigma(r_k x_k) \leq f(0) - \epsilon \qquad \text{for all } k. \tag{17}$$

Now, we consider separately the following two cases: the sequence $\{x_k\}$ is bounded, and $\{x_k\}$ is unbounded.

Case 1: The sequence $\{x_k\}$ is bounded.

We may assume without loss of generality that $x_k \to \bar{x}$. In view of Assumption 1(c) and Lemma 1, it follows that the sequence $\{f(x_k)\}$ is bounded from below, i.e., $f(x_k) \geq K$ for some scalar K and for all k. Hence, it follows from Eq. (17) that

$$\frac{\sigma(r_k x_k)}{r_k} \leq f(0) - \epsilon - f(x_k) \leq f(0) - \epsilon - K,$$

and therefore

$$\limsup_{k\to\infty} \frac{\sigma(r_k x_k)}{r_k} < \infty.$$

Since $r_k \to \infty$ and $x_k \to \bar{x}$, by Assumption 4(a), we have $\bar{x} \leq 0$. Consequently, by Assumption 4(b), we further have

$$\liminf_{k\to\infty} \frac{\sigma(r_k x_k)}{r_k} \geq 0.$$

Taking the limit inferior in Eq. (17) as $k \to \infty$, and using the preceding relation, we obtain

$$\liminf_{k\to\infty} f(x_k) \leq f(0) - \epsilon. \tag{18}$$

Since $x_k \to \bar{x}$ and $\bar{x} \leq 0$, it follows that $x_k^+ \to 0$. Furthermore, since $x_k \leq x_k^+$ and the function f is nonincreasing, we have $f(x_k^+) \leq f(x_k)$ for all k. Combining these with the assumption that f is lower semicontinuous at 0 [cf. Assumption 1(b)], we obtain

$$f(0) \leq \liminf_{k\to\infty} f(x_k^+) \leq \liminf_{k\to\infty} f(x_k).$$

Since $\liminf_{k\to\infty} f(x_k) \leq f(0) - \epsilon$ [cf. Eq. (18)], this yields a contradiction.

Case 2: The sequence $\{x_k\}$ is unbounded.

We may assume without loss of generality that $\|x_k\| \to \infty$ and $\|x_k\| > 0$ for all k. Dividing by $\|x_k\|$ in Eq. (17), we obtain

$$\frac{f(x_k)}{\|x_k\|} + \frac{\sigma(\lambda_k v_k)}{\lambda_k} \leq \frac{f(0) - \epsilon}{\|x_k\|} \qquad \text{for all } k, \tag{19}$$

where

$$\lambda_k = r_k \|x_k\| \qquad \text{and} \qquad v_k = \frac{x_k}{\|x_k\|}.$$

Note that v_k is bounded, and we may assume without loss of generality that $v_k \to \bar{v}$ for some vector $\bar{v} \neq 0$.

By rearranging the terms and taking the limit superior in relation (19), we obtain

$$\limsup_{k\to\infty} \frac{\sigma(\lambda_k v_k)}{\lambda_k} \leq -\liminf_{k\to\infty} \frac{f(x_k)}{\|x_k\|}. \tag{20}$$

If the sequence $\{f(x_k)\}$ is bounded from below, then we have

$$\liminf_{k\to\infty} \frac{f(x_k)}{\|x_k\|} > -\infty.$$

If the sequence $\{f(x_k)\}$ is unbounded from below, the preceding relation still holds in view of Assumption 1(c) and Lemma 1. Hence, it follows from Eq. (20) that

$$\limsup_{k\to\infty} \frac{\sigma(\lambda_k v_k)}{\lambda_k} < \infty.$$

Since $\lambda_k \to \infty$ and $v_k \to \bar{v}$, by Assumption 4(a), we have $\bar{v} \leq 0$. By Assumption 4(b), we further have

$$\liminf_{k\to\infty} \frac{\sigma(\lambda_k v_k)}{\lambda_k} \geq 0.$$

By taking limit inferior in relation (19) and by using the preceding inequality, we obtain

$$\liminf_{k\to\infty} \frac{f(x_k)}{\|x_k\|} \leq 0.$$

Without loss of generality, we may assume that $\frac{f(x_k)^+}{\|x_k\|} \to 0$ along some subsequence. Therefore, we have

$$\frac{1}{\|x_k\|}(x_k, f(x_k)^+) \to (\bar{v}, 0) \qquad \text{for some } \bar{v} \le 0 \;\text{ with }\; \bar{v} \neq 0.$$

Since $\frac{1}{\|x_k\|} \to 0$ and $(x_k, f(x_k)^+) \in \mathrm{epi}(f)$ for all k, by the definition of an asymptotic direction [cf. Definition 3] it follows that $(\bar{v}, 0) \in (\mathrm{epi}(f))^\infty$ for some $\bar{v} \le 0$ and $\bar{v} \neq 0$. This, however, contradicts Assumption 5. Hence, relation (16) holds. The proof that relation (16) implies the abstract convexity of f at $x = 0$ with respect to $\bar{H}$ is similar to that of Proposition 2. **Q.E.D.**

4 Application to Constrained Optimization Duality

In this section, we use the abstract convexity results of Section 3 to study duality for constrained (nonconvex) optimization problems. We consider the following optimization problem

$$\begin{array}{ll} \min & F_0(x) \\ \text{s.t.} & x \in X,\; F(x) \le 0, \end{array}$$

where X is a nonempty subset of $\mathbb{R}^n$,

$$F(x) = (F_1(x), \ldots, F_m(x)),$$

and $F_i : \mathbb{R}^n \mapsto (-\infty, \infty]$ for $i = 0, 1, \ldots, m$. We refer to this as the *primal problem*, and denote its optimal value by F^*.

For the primal problem, we consider a *dualizing parametrization function* $\bar{F} : \mathbb{R}^n \times \mathbb{R}^m \mapsto (-\infty, \infty]$ that satisfies $\bar{F}(x, 0) = F_0(x)$ for all $x \in X$ and $F(x) \le 0$. One particular example of a dualizing parametrization is the following:

$$\bar{F}(x, u) = \begin{cases} F_0(x) & \text{if } F(x) \le u, \\ +\infty & \text{otherwise}, \end{cases} \tag{21}$$

(considered by Nedić and Ozdaglar in [10] and [11]). The parametrization function induces the *perturbation* or *primal function* given by

$$p(u) = \inf_{x \in X} \bar{F}(x, u). \tag{22}$$

We next define the augmented dual problem through the use of coupling functions (see Burachik and Rubinov [7]). In particular, for $\Omega = \mathbb{R}_+ \times \mathbb{R}^m$ and a given augmenting function σ, we consider coupling functions ρ of the following two forms:

$$\rho(u, w) = -r\sigma(u) - \mu' u \qquad \text{for all } u \in \mathbb{R}^m \text{ and all } w = (r, \mu) \in \Omega,$$

and

$$\rho(u, w) = -\frac{1}{r}\,\sigma(ru) - \mu' u \qquad \text{for all } u \in \mathbb{R}^m \text{ and all } w = (r, \mu) \in \Omega.$$

Note that the preceding coupling functions satisfy $\rho(0, w) = 0$ for all $w \in \Omega$.

For any coupling function ρ, we define the *augmented Lagrangian function* as

$$l(x, w) = \inf_{u \in \mathbb{R}^m} \left\{ \bar{F}(x, u) - \rho(u, w) \right\}, \tag{23}$$

and the *augmented dual function* as

$$q(w) = \inf_{x \in X} l(x, w). \tag{24}$$

We consider the problem

$$\begin{array}{ll} \max & q(w) \\ \text{s.t.} & w \in \Omega. \end{array} \tag{25}$$

We refer to this problem as the *augmented dual problem*, and denote its optimal value by q^*. We say that *there is zero duality gap* when $q^* = F^*$, and we say that *there is a duality gap* when $q^* < F^*$.

We establish our duality results through the use of Fenchel-Moreau theory involving conjugate functions (see Burachik and Rubinov [7]). Specifically, let $p : \mathbb{R}^m \mapsto [-\infty, \infty]$ be an arbitrary function. We define the *Fenchel-Moreau conjugate* to p by

$$p^\rho(w) = \sup_{u \in \mathbb{R}^m} \{\rho(u, w) - p(u)\} \qquad \text{for all } w \in \Omega. \tag{26}$$

We also define the *Fenchel-Moreau biconjugate* to p by

$$p^{\rho\rho}(u) = \sup_{w \in \Omega} \{\rho(u, w) - p^\rho(w)\} \qquad \text{for all } \ u \in \mathbb{R}^m. \tag{27}$$

In the subsequent development, we use the following classical result of abstract convex analysis (see Rubinov [16]).

Theorem 1 *(Fenchel-Moreau Theorem)* Let H be a set of functions given by

$$H = \{g \mid g(u) = \rho(u, w) + c, \ u \in \mathbb{R}^m, \ w \in \Omega, \ c \in \mathbb{R}\}. \tag{28}$$

Then, a function $p : \mathbb{R}^m \mapsto [-\infty, \infty]$ is abstract convex at a point $\bar{u} \in \mathbb{R}^m$ with respect to H if and only if

$$p(\bar{u}) = p^{\rho\rho}(\bar{u}).$$

We next provide an equivalent characterization of zero duality gap in terms of the perturbation function and its biconjugate (see also [17]).

Proposition 4 There is zero duality gap if and only if

$$p(0) = p^{\rho\rho}(0),$$

where $p^{\rho\rho}$ is the Fenchel-Moreau biconjugate of p.

Proof. Combining the relations in (23)-(25), we obtain

$$
\begin{aligned}
q^* &= \sup_{w\in\Omega} \inf_{x\in X} \inf_{u\in\mathbb{R}^m} \{\bar{F}(x,u) - \rho(u,w)\} \\
&= \sup_{w\in\Omega} \inf_{u\in\mathbb{R}^m} \inf_{x\in X} \{\bar{F}(x,u) - \rho(u,w)\} \\
&= \sup_{w\in\Omega} \inf_{u\in\mathbb{R}^m} \{p(u) - \rho(u,w)\} \\
&= \sup_{w\in\Omega} \Big[- \sup_{u\in\mathbb{R}^m} \{\rho(u,w) - p(u)\}\Big].
\end{aligned}
$$

By using the definition of Fenchel-Moreau conjugate of p [cf. Eq. (26)], we have

$$
\begin{aligned}
q^* &= \sup_{w\in\Omega}(-p^{\rho}(w)) \\
&= \sup_{w\in\Omega}\{\rho(0,w) - p^{\rho}(w)\},
\end{aligned}
$$

where the second equality follows from the assumption on the coupling function that $\rho(0,w) = 0$ for all $w \in \Omega$. By the definition of Fenchel-Moreau conjugate of p [cf. Eq. (27)], we obtain

$$q^* = p^{\rho\rho}(0).$$

By definition of the perturbation function, we have $p(0) = f^*$, thus implying that there is zero duality gap if and only if $p(0) = p^{\rho\rho}(0)$. **Q.E.D.**

For a given coupling function ρ, we define the set of functions

$$H = \{h \mid h(u) = \rho(u,w) + c,\ u \in \mathbb{R}^m,\ w \in \Omega,\ c \in \mathbb{R}\}$$

[cf. Eq. (28)]. From Proposition 4 and Theorem 1 it follows that there is zero duality gap if and only if the perturbation function p is abstract convex at $u = 0$ with respect to the set H. This, together with the abstract convexity results of Section 3, yields the following sufficient conditions for zero duality gap.

Proposition 5 *(Sufficient Conditions for Zero Duality Gap)* Assume that the perturbation function p satisfies Assumption 1. Furthermore, assume that one of the following holds:

(a) The augmenting function σ satisfies Assumption 2, and the coupling function ρ is given by

$$\rho(u,w) = -r\sigma(u) - \mu' u.$$

(b) The augmenting function σ satisfies Assumption 3, and the coupling function ρ is given by

$$\rho(u,w) = -\frac{1}{r}\,\sigma(ru) - \mu' u.$$

(c) The perturbation function p satisfies Assumption 5. The augmenting function σ satisfies Assumption 4, and the coupling function ρ is given by

$$\rho(u,w) = -\frac{1}{r}\,\sigma(ru) - \mu' u.$$

Then, there is zero duality gap, i.e., $q^* = f^*$.

Proof. It suffices to show that p does not take the value $-\infty$. Once this is established, the result in part (a) [(b) and (c), respectively] follows from Proposition 4, Theorem 1, and Proposition 1 [Proposition 2 and Proposition 3, respectively].

Assume to obtain a contradiction that $p(\bar{u}) = -\infty$ for some $\bar{u} \in \mathbb{R}^m$. Consider the following sequence

$$u_k = \bar{u} + ke \qquad \text{with } e = (1, \ldots, 1) \text{ and } e \in \mathbb{R}^m.$$

Note that we have $u_k \geq \bar{u}$ for all k. Since the perturbation function p is nonincreasing [cf. Eqs. (21)-(22)] and $p(\bar{u}) = -\infty$, it follows that

$$p(u_k) \leq p(\bar{u}) = -\infty \qquad \text{for all } k.$$

Thus, the vectors $(u_k, -k\|u_k\|)$ for $k \geq 1$ belong to the epigraph epi(p). Define

$$\lambda_k = \frac{1}{k\|u_k\|} \qquad \text{for } k \geq 1,$$

and note that $\lambda_k > 0$ and $\lambda_k \to 0$ as $k \to \infty$. Furthermore, we have

$$\lambda_k \left(u_k, -k\|u_k\|\right) = \left(\frac{u_k}{k\|u_k\|}, -1\right).$$

Thus, it follows that $\lambda_k \left(u_k, -k\|u_k\|\right)$ converges to $(0, -1)$. By the definition of the asymptotic cone [cf. Definition 3], it further follows that $(0, -1)$ is an asymptotic direction of epi(p), i.e., $(0, 1) \in (\text{epi}(p))^\infty$. This, however, contradicts the assumption that $(0, 1) \notin (\text{epi}(p))^\infty$ [cf. Assumption 1(c)]. **Q.E.D.**

We now provide an example of a perturbation function satisfying Assumption 1. At first, note that p is always nonincreasing [cf. Eqs. (21)-(22)]. Consider an optimization problem where the constraint set X is compact and the functions F_i, $i = 0, 1, \ldots, m$ are lower semicontinuous over $\mathbb{R}^n$. In this case, clearly $p(0)$ is finite, and thus p satisfies Assumption 1(a). Furthermore, using compactness of X and lower semicontinuity of F_i's, one can show that $p(u)$ is lower semicontinuous at $u = 0$, thus verifying that p satisfies Assumption 1(b). Moreover, by the compactness of X and the lower semicontinuity of F_0 it follows that $p(u) \geq \inf_{x \in X} F_0(x)$ for all u. In view of this and Lemma 1, it further follows that $(0, 1) \notin (\text{epi}(p))^\infty$, thus showing that p satisfies Assumption 1(c).

5 Conclusions

In this paper, we provided some zero duality gap results for constrained nonconvex optimization problems using the framework of abstract convexity. In particular, we have considered three different types of augmenting functions: nonnegative augmenting functions, bounded-below augmenting functions, and unbounded augmenting functions. Using these augmenting functions, we have defined two different sets of elementary functions and used them to analyze the abstract convexity properties of the perturbation function of the constrained problem. In our analysis, we have assumed some asymptotic direction properties of the perturbation function which are less restrictive than compactness assumptions used in previous work.

We have considered augmented dual problems defined in terms of nonconvex augmenting functions. We have connected the abstract convexity results with the zero duality gap properties of the augmented dual problems through the use of the well-known Fenchel-Moreau Theorem. The zero duality gap results established here have potential use in the development of dual algorithms for solving nonconvex constrained optimization problems. In particular, for such problems, one may consider relaxing some or all of the constraints by using the augmented Lagrangian scheme. Our results provide sufficient conditions guaranteeing the convergence of dual values to the primal optimal value without convexity assumptions for augmented Lagrangian functions.

References

[1] Auslender A. and Teboulle, M., Asymptotic Cones and Functions in Optimization and Variational Inequalities, Springer-Verlag, New York, 2003.

[2] Ben-Tal A. and Zibulevski M., "Penalty /barrier multiplier methods: A new class of augmented Lagrangian algorithms for large scale convex prorgramming problems," Research report 4/93, Optimization Laboratory, Faculty of Industrial Engineering and Management, Technion-Israel Institute of Technology, 1993.

[3] Bertsekas D. P., Nedić A., and Ozdaglar A., Convex Analysis and Optimization, Athena Scientific, Belmont, MA, 2003.

[4] Bertsekas D. P., Nedić A., and Ozdaglar A., "Min Common/Max Crossing Duality: A Simple Geometric Framework for Convex Optimization and Minimax Theory," Report LIDS-P-2536, Jan. 2002.

[5] Bonnans J. F. and Shapiro A., Perturbation Analysis of Optimization Problems, Springer-Verlag, New York, NY, 2000.

[6] Borwein J. M. and Lewis A. S., Convex Analysis and Nonlinear Optimization, Springer-Verlag, New York, NY, 2000.

[7] Burachik R. S. and Rubinov A., "Abstract convexity and augmented Lagrangians," preprint, 2006.

[8] Hiriart-Urruty J.-B. and Lemarechal C., Convex Analysis and Minimization Algorithms, vols. I and II, Springer-Verlag, Berlin and NY, 1993.

[9] Huang X. X. and Yang X. Q., "A unified augmented Lagrangian approach to duality and exact penalization," *Mathematics of Operations Research*, vol. 28, no. 3, pp. 533-552, 2003.

[10] Nedić A. and Ozdaglar A., "A Geometric Framework for Nonconvex Optimization Duality using Augmented Lagrangian Functions," forthcoming in *Journal of Global Optimization*, 2006.

[11] Nedić A. and Ozdaglar A., "Separation of Nonconvex Sets with General Augmenting Functions," submitted for publication, 2006.

[12] Polyak R., "Modified barrier functions: theory and methods," *Mathematical Programming*, vol. 54, pp. 177-222, 1992.

[13] Rockafellar R. T., Convex Analysis, Princeton University Press, Princeton, N.J., 1970.

[14] Rockafellar R. T. and Wets R. J.-B., Variational Analysis, Springer-Verlag, New York, 1998.

[15] Rubinov A. M., Glover B. M., and Yang X. Q., "Decreasing functions with applications to penalization," *SIAM Journal on Optimization*, vol. 10, no. 1, pp. 289-313, 1999.

[16] Rubinov A. M., Abstract Convexity and Global Optimization, Kluwer Academic Publishers, Dordrecht, 2000.

[17] Rubinov A. M., Huang X. X., and Yang X. Q., "The zero duality gap property and lower semicontinuity of the perturbation function," *Mathematics of Operations Research*, vol. 27, no. 4, pp. 775-791, 2002.

[18] Rubinov A. M. and Yang X. Q., Lagrange-type Functions in Constrained Nonconvex Optimization, Kluwer Academic Publishers, MA, 2003.