Duality in Nonlinear Programs Using Augmented Lagrangian Functions*

AMNON GONEN AND MORDECAI AVRIEL

Faculty of Industrial Engineering and Management, Technion-Israel Institute of Technology, Haifa 32000, Israel

Submitted by G.-C. Rota
Received May 10, 1985

A generally nonconvex optimization problem with equality constraints is studied. The problem is introduced as an “inf sup” of a generalized augmented Lagrangian function. A dual problem is defined as the “sup inf” of the same generalized augmented Lagrangian. Sufficient conditions are derived for constructing the augmented Lagrangian function such that the extremal values of the primal and dual problems are equal. Characterization of a class of augmented Lagrangian functions which satisfy the sufficient conditions for strong duality is presented. Finally, some examples of functions and primal–dual problems in the above-mentioned class are presented. © 1987 Academic Press, Inc.

1. INTRODUCTION

Let \( f, g_1, \ldots, g_m \) be real-valued functions defined on \( S \subset \mathbb{R}^n \). The primal mathematical programming problem we are interested in is defined as

\[
(P) \quad V_p = \inf_{x \in S} f(x)
\]

subject to

\[
g_i(x) = 0, \quad i = 1, \ldots, m.
\]

Let

\[
g(x) = (g_1(x), \ldots, g_m(x))
\]

and let the generalized augmented Lagrangian function associated with \( P \) be

\[
L(x, y, r) = f(x) + \varphi(g(x), y, r),
\]

* This research was partially supported by the Fund for the Promotion of Research at the Technion and the Andre Meyer Fund for Operations Research.
where $\varphi$ is a real-valued function, called the augmented multiplier function, defined on some subset of $\mathbb{R}^{2m+1}$ and $(y, r) \in T = \mathbb{R}^{m+1}_+$, where $\mathbb{R}^k_+$ is the nonnegative orthant of $\mathbb{R}^k$. Define

$$H(y, r) = \inf_{x \in S} L(x, y, r). \quad (1.5)$$

Then, a related mathematical programming problem that under certain conditions will become a dual problem of (P) is given by

$$(D) \quad V_D = \sup_{(y, r) \in T} H(y, r). \quad (1.6)$$

The primary purpose of this paper is to characterize a class of augmented multiplier functions $\varphi$ that satisfy conditions of strong duality for a bounded stable program (P). The use of augmented multiplier functions instead of the usual fixed multipliers will lead to duality results under quite weak assumptions on the functions involved in (P).

Certain aspects of the theory presented in this paper have been anticipated by earlier works in mathematical programming. Among these we can mention the works of Dolecki and Kurcuyusz [6], where $\Phi$-convexity with applications to augmented Lagrangians was studied. Balder [2] used an extension of the conjugate function concept to obtain duality results. He introduced a Lagrangian and showed that in specialization to mathematical programming problems, this Lagrangian appears as the extended Lagrangian in exterior penalty function methods. Bertsekas [4] surveyed combined primal–dual and penalty methods for equality constrained minimization which generalize the method of multipliers. In particular, he analyzed the differentiable exact penalty method. Evans and Gould [7] considered duality theorems without convexity requirements on (P), but assuming $\varphi(u, y, r) = u^T y$, that is, $\varphi$ is the fixed (nonaugmented) multiplier function. Kort [10], Kort and Bertsekas [11], presented a class of multiplier functions for convex programs. Buys [5] represents one of many works that used $\varphi(u, y, r) = u^T y + 1/2ru^T u$ as the augmented multiplier function. Mangasarian [13] associated a class of augmented Lagrangian functions with a nonconvex, inequality and equality constrained optimization problem in such a way that unconstrained stationary points of the Lagrangian are related to Kuhn–Tucker points of the optimization problem. He obtained duality results for the case of differentiable functions and gave a list of properties that the first and second derivatives of $\varphi$ must satisfy in order to attain strong duality. Gould [8] analyzed several equivalent formulations of generalized Lagrangian functions using multiplier functions. In Rockafellar [15] it was shown that duality gaps can be eliminated by passing from the ordinary Lagrangian
function formulation to an augmented Lagrangian function formulation, involving quadratic penalty-like terms. In particular he suggested the multiplier function

$$\varphi(u, y, r) = \sum_{i=1}^{m} \left\{ y_{i} \max[u_{i}, -y_{i}/2r] + r \max^2[u_{i}, -y_{i}/2r] \right\}$$

(1.7)

for inequality constrained problems. A duality theorem was proved for this function under quite weak assumptions on the optimization problem. Pollatschek [14] derived some necessary conditions for constructing a dual program of a nonlinear program with inequality constraints such that the extremal values of both programs are the same. Tind and Wolsey [16] give a recent survey of general duality theory for nonlinear programming. Their approach is based on a price function without involving any penalty.

The present work deals with generalized augmented Lagrangian functions having as few regularity assumptions as possible such that strong duality relations can exist between primal and dual programs.

2. GENERALIZED DUALITY

Consider the primal program as defined above:

(P) \[ V_{p} = \inf_{x \in S} f(x) \]

subject to \[ g_{i}(x) = 0, \quad i = 1, \ldots, m. \] (2.2)

Let \( V_{p} = +\infty \) if the feasible set of (P) is empty.

The generalized augmented Lagrangian associated with (P) is defined as

\[ L(x, y, r) = f(x) + \varphi(g(x), y, r). \] (2.3)

We start our derivation of duality results by finding a relationship between \( L \) and \( V_{p} \) by making certain assumptions on \( \varphi \), and without further assumptions on (P) itself:

**ASSUMPTION A.1.** For every \((y, r) \in T\) we have

\[ \varphi(0, y, r) = 0. \] (2.4)

**ASSUMPTION A.2.** (Unboundedness assumption). For every \( u \in R^{m}, u \neq 0 \) and for every \( c \in R^{1}_{+} \) there exist a \((y, r) \in T\) such that

\[ \varphi(u, y, r) > c. \] (2.5)
We then have

**Lemma 1.** If \( \varphi \) satisfies A.1 and A.2, then

\[
\inf_{x \in S} \sup_{(y, r) \in T} L(x, y, r) = V_p. \tag{2.6}
\]

**Proof.** It follows from A.1 and A.2 that

\[
\sup_{(y, r) \in T} L(x, y, r) = \begin{cases} f(x), & g(x) = 0 \\ +\infty, & g(x) \neq 0. \end{cases} \tag{2.7}
\]

Hence,

\[
\inf_{x \in S} \sup_{(y, r) \in T} L(x, y, r) = \inf_{x \in S} f(x) = V_p. \tag{2.8}
\]

Note that Lemma 1 indicates that the value of a mathematical programming problem with equality constraints can be presented as (2.6) whenever assumptions A.1 and A.2 are satisfied, regardless of configuration or properties the original problem satisfies. Assumptions A.1 and A.2 are, however, not necessary. Recall that we want to obtain a dual program of (P) defined as

\[
V_D = \sup_{(y, r) \in T} H(y, r), \tag{2.9}
\]

where

\[
H(y, r) = \inf_{x \in S} L(x, y, r). \tag{2.10}
\]

We then have

**Theorem 1** (Weak duality theorem). If \( \varphi \) satisfies A.1 and A.2, then \( V_p \geq V_D \).

**Proof.** For every \( L(x, y, r) \),

\[
\inf_{x \in S} \sup_{(y, r) \in T} L(x, y, r) \geq \sup_{(y, r) \in T} \inf_{x \in S} L(x, y, r). \tag{2.11}
\]

By Lemma 1 the left-hand side of (2.11) is \( V_p \) and by (2.9), (2.10) the right-hand side is \( V_D \). Let \( \lambda \) be an operator, \( \lambda : R^n \rightarrow R_+^n, \lambda(0) = 0 \), that satisfies

\[
|u| \leq |v| \quad \text{implies} \quad \lambda(u) \leq \lambda(v), \tag{2.12}
\]
where \(|u| = (|u_1|, \ldots, |u_m|)\). For example,
\[
\lambda(u) = (|u_1|, \ldots, |u_m|),
\]
(2.13)
or
\[
\lambda(u) = (u_1^2, \ldots, u_m^2).
\]
(2.14)
For every \(x \in S\) and \(u \in \mathbb{R}^m\) let
\[
F(x, u) = \begin{cases} f(u) & \text{if } \lambda(g(x)) \leq \lambda(u) \\ +\infty & \text{otherwise} \end{cases}
\]
(2.15)
and define the perturbation function associated with \((P)\) as
\[
W(u) = \inf_{x \in S} F(x, u).
\]
(2.16)
From (2.12), (2.15), and (2.16) it follows that for \(u^1 \geq u^2 \geq 0\) we have \(W(u^2) \geq W(u^1)\).

We also have

**Lemma 2.** If \(\varphi\) is concave in \((y, r) \in T\) for every \(u \in \mathbb{R}^m\) then \(L\) and \(H\) are also concave in \((y, r) \in T\). If \(\varphi\) is upper semicontinuous in \((y, r) \in T\) for every \(u \in \mathbb{R}^m\), then \(L\) and \(H\) are upper semicontinuous in \((y, r) \in T\).

**Proof** The concavity of \(L\) follows from the definitions, and since by (2.10) \(H\) is in this case a pointwise infimum of a collection of concave functions, it is also concave, see Avriel (1). For the second part of the lemma let \(\alpha \in \mathbb{R}\) and observe that for every \(x \in S\),
\[
\{(y, r) : L(x, y, r) \geq \alpha\} = \{(y, r) : \varphi(g(x), y, r) \geq \alpha - f(x)\}
\]
(2.17)
and by the upper semicontinuity of \(\varphi\) the right-hand side of (2.17) is a closed set. Similarly, let \(\alpha \in \mathbb{R}\). Then,
\[
\{(y, r) : H(y, r) \geq \alpha\} = \{(y, r) : \inf_{x \in S} L(x, y, r) \geq \alpha\}
\]
(2.18)
\[
= \bigcap_{x \in S} \{(y, r) : L(x, y, r) \geq \alpha\}
\]
and it follows that \(H\) is also upper semicontinuous.

To find connections between \(F, H, L,\) and \(W\) we introduce a further assumption on \(\varphi\):
ASSUMPTION A.3 (Isotonicity assumption). For every $u^1 \in \mathbb{R}^m$, $u^2 \in \mathbb{R}^m$ and $(y, r) \in T$,
\[ \lambda(u^1) \geq \lambda(u^2) \Rightarrow \varphi(u^1, y, r) \geq \varphi(y^2, y, r). \]

**Lemma 3.** If $\varphi$ satisfies A.3, then
\[ L(x, y, r) = \inf_{u \in \mathbb{R}^m} [F(x, u) + \varphi(u, y, r)] \] (2.20)
for every $x \in S$, and $(y, r) \in T$.

**Proof.** If $x \in S$, $u \in \mathbb{R}^m$ such that $\lambda(g(x)) \leq \lambda(u)$, then by A.3
\[ \varphi(g(x), y, r) \leq \varphi(u, y, r) \] (2.21)
for every $(y, r) \in T$. If $\lambda(g(x)) \leq \lambda(u)$, then $F(x, u) = +\infty$. Hence,
\[ L(x, y, r) = f(x) + \varphi(g(x), y, r) \leq F(x, u) + \varphi(u, y, r) \] (2.22)
and
\[ L(x, y, r) \leq \inf_{u \in \mathbb{R}^m} [F(x, u) + \varphi(u, y, r)]. \] (2.23)
But,
\[ L(x, y, r) = F(x, g(x)) + \varphi(g(x), y, r) \]
\[ \geq \inf_{u \in \mathbb{R}^m} [F(x, u) + \varphi(u, y, r)] \] (2.24)
and (2.20) holds.

**Lemma 4.** If $\varphi$ satisfies A.3, then
\[ H(y, r) = \inf_{u \in \mathbb{R}^m} [W(u) + \varphi(u, y, r)] \] (2.25)
for every $(y, r) \in T$.

**Proof.** From (2.10) and Lemma 3 we obtain
\[ H(y, r) = \inf_{x \in S} L(x, y, r) = \inf_{x \in S} \inf_{u \in \mathbb{R}^m} [F(x, u) + \varphi(u, y, r)]. \] (2.26)
Hence
\[ H(y, r) = \inf_{u \in \mathbb{R}^m} \inf_{x \in S} [F(x, u) + \varphi(u, y, r)] = \inf_{u \in \mathbb{R}^m} [W(u) + \varphi(u, y, r)]. \] (2.27)
Turning now to establishing strong duality relations between programs (P) and (D) we need some stronger assumptions on \( \varphi \) which will be stated below.

**Definition 1.** A real-valued function \( \varphi \) defined on \( M \times \mathbb{R}^i \subset \mathbb{R}^m \times \mathbb{R}^i \) is said to be uniformly unbounded on \( M \) if for every \( c > 0 \) there is an \( r \in \mathbb{R}^i \) such that
\[
\varphi(u, r) > c \tag{2.28}
\]
for every \( u \in M \).

Note that the function \( \varphi(u, r) = ru^T u \) is uniformly unbounded on \( M_n = \{ u : \|u\| > 1/n \} \), but it is not uniformly unbounded on \( M = \bigcup_{n=1}^{\infty} \{ u : \|u\| > 1/n \} \) although for each \( u \in M \), \( \lim_{r \to \infty} ru^T u = +\infty \). Here \( \|u\| \) denotes the Euclidean norm of \( u \). It can be verified that \( \varphi \) is uniformly unbounded on \( M \) if and only if the family of functions \( \{ \xi_r(u) = 1/\varphi(u, r) \} \) contains a subsequence \( \{ \xi_{r_i}(u) : i = 1, 2, \ldots \} \) such that
\[
\lim_{i \to \infty} \sup_{u \in M} \{ \xi_{r_i}(u), i = 1, 2, \ldots \} = 0. \tag{2.29}
\]

Let us introduce now additional assumptions on \( \varphi \):

**Assumption (A.4)** (Behavior of \( \varphi \) for \( y = 0 \)). Assume that \( r \geq 0 \). Then,

(A.4.1). For \( u \neq 0 \), \( \varphi(u, 0, r) \) is nondecreasing in \( r \).

(A.4.2). For \( u \in \mathbb{R}^m \), \( u \neq 0 \), \( \lim_{r \to +\infty} \varphi(u, 0, r) = +\infty \).

(A.4.3). Let \( N(0) \subset \mathbb{R}^m \) be any open spherical neighborhood of the origin in \( \mathbb{R}^m \) and let \( M \subset \mathbb{R}^m \) be the complement of \( N(0) \). Then the function \( \theta_s \), defined by
\[
\theta_s(u, r) = \varphi(u, 0, r) - \varphi(u, 0, s) \tag{2.30}
\]
is uniformly unbounded on \( M \) for every \( s < r \).

(A.4.4). For every \( r \) there exists an \( N(0) \subset \mathbb{R}^m \) where \( \varphi(u, 0, r) \) is nonnegative.

Assumption (A.4.3) can be presented in a more simple way when the function \( \varphi(u, 0, r) \) is differentiable with respect to \( r \):

**Lemma 5.** Let \( N(0) \) and \( M \) be as in (A.4.3) and assume that \( \varphi(u, 0, r) \) is differentiable with respect to \( r \) for all \( u \in M \). If there exists \( r_0 \in \mathbb{R}^i \) and \( c > 0 \) such that \( \partial \varphi/\partial r(u, 0, r) > c \) for all \( u \in M \) and \( r > r_0 \) then \( \varphi \) satisfies (A.4.3).

**Proof.** Let \( r > s > r_0 \) and \( u \in \mathbb{R}^m \). Then by the Mean Value Theorem (Bartle \[3\])
\[
\varphi(u, 0, r) - \varphi(u, 0, s) = (r - s) \frac{\partial \varphi}{\partial r}(u, 0, t). \tag{2.31}
\]
where \( t \) is some point between \( s \) and \( r \). Hence

\[
\theta_s(u, r) = \varphi(u, 0, r) - \varphi(u, 0, s) \geq (r-s) e^{-\alpha s} \to \infty.
\]

It is interesting to observe that for any differentiable function that satisfies (A.4.1), it follows that \( \frac{\partial \varphi}{\partial r}(u, 0, r) \geq 0 \), which makes the conditions of Lemma 5 to be generally satisfied.

**Assumption A.5.** Let \( N(0) \) and \( M \) be as in (A.4.3). There exists a real nonnegative function \( \Psi \) defined on \( \mathbb{R}^m \times (0, +\infty) \), satisfying the following properties:

(A.5.1). For every \( u \in \mathbb{R}^m \), \( (y, r) \in T \), \( (z, s) \in T \), \( r > s \),

\[
\varphi(u, y, r) - \varphi(u, z, s) \geq -\Psi(y-z, r-s).
\]

(A.5.2). For every \( a \in \mathbb{R}^m \),

\[
\lim_{b \to +\infty} \Psi(a, b) = 0.
\]

The class of multiplier functions that satisfies assumptions A.1–A.5 includes some of the functions proposed in previous works on augmented Lagrangian functions. These will be discussed in the next section.

We can prove now some more results on the properties of the functions \( H \) and \( L \).

**Lemma 6.** If \( \varphi \) satisfies A.1 and (A.4.1), then \( L(x, 0, r) \) and \( H(0, r) \) are nondecreasing in \( r \).

**Proof.** From A.1 we have that \( \varphi(g(x), 0, r) = 0 \) for every \( x \) such that \( g(x) = 0 \) and for every \( r \). If \( g(x) \neq 0 \), then by (A.4.1) we know that \( \varphi(g(x), 0, r) \) is nondecreasing in \( r \). Hence \( L(x, 0, r) \) is also nondecreasing in \( r \). Similar arguments can be applied to show that \( H(0, r) \) is also non-decreasing in \( r \). \( \blacksquare \)

**Lemma 7.** If \( \varphi \) satisfies A.3, and (A.5.1), then for every \( r > 0 \),

\[
H(y, r) \geq \sup_{(z, s) \in T, r > s > 0} \left[ H(z, s) - \Psi(y-z, r-s) \right].
\]

**Proof.** From (A.5.1) we obtain for every \( u \in \mathbb{R}^m \), \( (y, r) \in T \), \( (z, s) \in T \), and \( r > s \),

\[
W(u) + \varphi(u, y, r) \geq W(u) + \varphi(u, z, s) - \Psi(y-z, r-s).
\]
Taking infimum with respect to \( u \in \mathbb{R}^m \) on both sides of (2.36) we obtain by Lemma 4,
\[
H( y, r) \geq H(z, s) - \Psi( y - z, r - s). \tag{2.37}
\]
Since (2.37) holds for every \((z, s) \in T\) and \( r > s \), it follows that (2.35) holds.

**Corollary 1.** If \( \phi \) satisfies A.3 and (A.5.1), and there exists a \((z, s) \in T\) such that \( H(z, s) \) is finite, then \( H( y, r) \neq -\infty \), for every \((y, r) \in T\) satisfying \( r > s \).

**Proof.** It follows from the assumptions and the proof of Lemma 7 that for every \( r > s \), (2.37) holds; that is,
\[
H( y, r) \geq H(z, s) - \Psi[(y - z), (r - s)]. \tag{2.38}
\]
Since \( \Psi \) is real valued, \( H( y, r) \) is bounded below for every \((y, r) \) such that \( r > s \).

**Corollary 2.** If \( \phi \) satisfies A.3, (A.5.1) and (A.5.2), then
\[
V_D = \sup_{(z, s) \in T} H(z, s) = \lim_{r \to +\infty} H( y, r) \tag{2.39}
\]
for every \( y \in \mathbb{R}^n \).

**Proof.** For every \((z, s) \in T\) and \( \varepsilon > 0 \) and for every \( y \in \mathbb{R}^n \) there exist by (A.5.2) a sufficiently large \( r \in \mathbb{R}^1_+, r > s \) such that
\[
\Psi[(y - z), (r - s)] < \varepsilon. \tag{2.40}
\]
Hence, by Lemma 7
\[
H( y, r) \geq H(z, s) - \varepsilon \tag{2.41}
\]
and since (2.41) holds for every sufficiently large \( r \), we obtain
\[
\lim_{r \to +\infty} H( y, r) \geq H(z, s) - \varepsilon. \tag{2.42}
\]
Since (2.42) holds for every \((z, s) \in T\) and \( \varepsilon > 0 \), we have
\[
\lim_{r \to +\infty} H( y, r) \geq \sup_{(z, s) \in T} H(z, s) = V_D. \tag{2.43}
\]
Clearly,
\[
\sup_{(z, s) \in T} H(z, s) \geq \lim_{r \to +\infty} H( y, r) \tag{2.44}
\]
and (2.39) holds.
Equation (2.39) shows the connection between the optimal value of the dual program $V_D$, and the solution of the dual program by a penalty function method. This result was earlier stated in Rockafellar [15].

So far we have not imposed any conditions on the functions appearing in the original problem (P). However, for proving strong duality relations between (P) and (D) we need certain assumptions on (P).

**Definition 2.** Program (P) is said to satisfy the boundedness condition if there exists an $r \in R^1_+$ such that $L(x, 0, r)$ is bounded below for every $x \in S$.

The boundedness condition certainly holds if $\varphi$ is bounded below on $R^m \times T$ and $f$ is bounded below on $S$, or, alternatively, if $S$ is compact and $f$ is lower semicontinuous.

**Lemma 8.** Suppose that $\varphi$ satisfies A.1, A.3, (A.4.1), (A.5.1), and (A.5.2). Then $V_D \neq -\infty$ if and only if (P) satisfies the boundedness condition.

**Proof.** Assume that (P) satisfies the boundedness condition. Then there exists an $r \in R^1_+$ such that

$$H(0, r) = \inf_{x \in S} L(x, 0, r) = q > -\infty.$$  \hspace{1cm} (2.45)

From Lemma 6 we obtain that $H(0, r)$ is nondecreasing in $r$. Hence, from Corollary 2,

$$V_D = \lim_{s \to +\infty} H(0, s) \geq H(0, r) = q > -\infty.$$  \hspace{1cm} (2.46)

Conversely, if (P) does not satisfy the boundedness condition, then $L(x, 0, r)$ is unbounded below and $H(0, r) = -\infty$ for every $r \in R^1_+$. Form Corollary 2 it follows that $V_D = -\infty$. \hfill \square

Next we have

**Theorem 2.** If $\varphi$ satisfies A.1, A.3, (A.4.2), (A.4.3), (A.4.4), (A.5.1), and (A.5.2) and (P) satisfies the boundedness condition, then

$$V_P = W(0) \geq \lim_{u \to 0} \{ \inf \{ W(u) \} \} = V_D.$$  \hspace{1cm} (2.47)

**Proof.** Form the definitions of $W$ and $V_P$ it follows that $V_P = W(0)$. From Lemma 4 we obtain

$$H(0, r) = \inf_{u \in R^m} [W(u) + \varphi(u, 0, r)]$$  \hspace{1cm} (2.48)

$$\leq \inf_{u \in R^m} [W(u) + \varphi(0, 0, r)] \leq \lim_{u \to 0} \{ \inf \{ W(u) \} \},$$
where the last inequality follows from A.1. By Corollary 2 we conclude that

\[ V_D = \lim_{r \to +\infty} H(0, r) \leq \lim_{u \to 0} \{ \inf \{ W(u) \} \}. \quad (2.49) \]

Suppose now that (P) satisfies the boundedness condition; that is, suppose \( L(x, 0, \bar{r}) \) is bounded below for every \( x \in S \). Hence, there is a \( \bar{q} \) such that

\[ H(0, \bar{r}) \geq \bar{q}, \quad (2.50) \]

and by Lemma 4

\[ W(u) + \varphi(u, 0, \bar{r}) \geq \bar{q} \quad (2.51) \]

for every \( u \in \mathbb{R}^m \). Let \( q \) be a real number such that

\[ \lim_{u \to 0^+} \{ \inf \{ W(u) \} \} > q. \quad (2.52) \]

The existence of such a \( q \) is assured by (2.46). Now choose a sufficiently small \( \varepsilon > 0 \) such that

\[ W(u) > q \quad (2.53) \]

for every \( u \) satisfying \( \| u \| < \varepsilon \). By (A.4.2) and (A.4.3) there is a sufficiently large \( r \) such that

\[ \varphi(u, 0, r) - \varphi(u, 0, \bar{r}) \geq q - \bar{q} \quad (2.54) \]

for every \( u \in \mathbb{R}^m, \| u \| \geq \varepsilon \). Thus

\[ W(u) \geq \bar{q} - \varphi(u, 0, \bar{r}) \geq q - \varphi(u, 0, r) \quad (2.55) \]

and

\[ W(u) + \varphi(u, 0, r) \geq q \quad (2.56) \]

for every \( u \in \mathbb{R}^m, \| u \| \geq \varepsilon \). For \( u \) such that \( \| u \| < \varepsilon \) the case is easier. Inequality (2.53) holds and from (A.4.4) it follows that \( \varphi(u, 0, r) \geq 0 \), therefore (2.56) holds, that is it holds for every \( u \in \mathbb{R}^m \). Hence,

\[ V_D \geq H(0, r) \geq q. \quad (2.57) \]

Since, if (2.52) holds, then also (2.57) holds—it must be that

\[ V_D \geq \lim_{u \to 0} \{ \inf \{ W(u) \} \} \quad (2.58) \]

and the proof is complete. \( \square \)
Following Rockafellar [15] we have

**DEFINITION 3.** Program (P) is said to be stable of degree 0 if there is a real function \( \theta \) defined on \( N(0) \), an open neighborhood of the origin in \( R^n \), such that \( \theta \) is continuous and

1. \( W(u) \geq \theta(u) \) for every \( u \in N(0) \), (2.59)
2. \( W(0) = \theta(0) \). (2.60)

We then have

**THEOREM 3.** Program (P) is stable of degree 0 if and only if

\[
W(0) = \lim_{u \to 0} \{ \inf W(u) \} \quad (2.61)
\]

whenever \( W(0) \) is finite.

**Proof.** Following Rockafellar [15], suppose that (P) is stable of degree 0. Hence

\[
\lim_{u \to 0} \{ \inf W(u) \} \geq \lim_{u \to 0} \{ \inf \theta(u) \} = \theta(0) = W(0). \quad (2.62)
\]

As mentioned earlier, \( W(0) \geq W(u) \) for \( u \in R^n \), thus

\[
W(0) \geq \lim_{u \to 0} \{ \inf W(u) \}. \quad (2.63)
\]

and (2.61) holds.

Conversely, define the real function \( \xi : R^+ \to R^+ \) as

\[
\xi(s) = \inf_{\|u\| < s} W(u). \quad (2.64)
\]

It can be seen that \( \xi \) is nonincreasing on \( R^+ \) and

\[
W(0) = \lim_{u \to 0} \{ \inf W(u) \} \Rightarrow \lim_{s \to 0} \xi(s) = \xi(0). \quad (2.65)
\]

Choose any \( \varepsilon > 0 \) and define the function \( \xi \) on \( [0, \varepsilon/2] \) as

\[
\xi(0) = \xi(0), \quad \xi(\varepsilon/j + 1) = \xi(\varepsilon/j) \quad (2.66)
\]

for every positive \( j \). On the interval \( (\varepsilon/j + 1, \varepsilon/j) \) we define \( \xi \) as a linear interpolation of its values at \( \varepsilon/j + 1 \) and \( \varepsilon/j \). It follows that \( \xi \) is continuous and satisfies \( \xi \leq \xi \). Let us define

\[
\theta(u) = \xi(\|u\|). \quad (2.67)
\]
It can be easily shown that $W(u) \geq \theta(u)$ and $W(0) = \theta(0)$, hence (P) is stable of degree 0.

From Theorems 2 and 3 we immediately obtain

**THEOREM 4.** (Strong duality theorem). Let $\varphi$ satisfy A.1, A.3, A.4, and A.5. If (P) satisfies the boundedness condition and is stable of degree 0, then $V_P = V_D$.

An alternative characterization of stability is given in the following theorem.

**THEOREM 5.** Program (P) is stable of degree 0 if and only if the perturbation function $W : R^m \to R^1$ is continuous at $u = 0$.

**Proof.** If $W$ is continuous at $u = 0$, then

$$W(0) = \lim_{u \to 0} \{\inf \{W(u)\}\}$$

and, by Theorem 3, (P) is stable of degree 0. Conversely, if (P) is stable of degree 0, then by Theorem 3 (2.68) holds. Since $W(0) \geq W(u)$ for every $u \in R^m$ it follows that

$$W(0) = \lim_{u \to 0} \{\inf \{W(u)\}\}.$$  

Clearly,

$$\lim_{u \to 0} \{\sup \{W(u)\}\} \geq \lim_{u \to 0} \{\inf \{W(u)\}\}.$$  

Combining (2.68), (2.69), (2.70) we obtain

$$W(0) = \lim_{u \to 0} \{\sup \{W(u)\}\} = \lim_{u \to 0} \{\inf \{W(u)\}\}$$

and the continuity of $W$ at $u = 0$ follows.

Let $\Phi$ denote the family of all functions $\varphi : R^{2m+1} \to R^1$ that satisfy assumptions A.1–A.5. Then we have the following closedness properties of $\Phi$:

**LEMA 9.** If $\varphi \in \Phi$ and $\alpha$ is a positive number, then also $\alpha \cdot \varphi \in \Phi$. If $\varphi_1 \in \Phi$ and $\varphi_2 \in \Phi$ then also $\varphi_1 + \varphi_2 \in \Phi$.

The proof is immediate and will be omitted. These closedness properties of $\Phi$ under addition and scalar multiplication will be useful in the next section where some specific examples of members of $\Phi$ will be examined.
3. EXAMPLES OF AUGMENTED MULTIPLIER FUNCTIONS

In this section we examine several augmented multiplier functions with respect to satisfying the assumptions introduced in the previous section. Let us choose \( \lambda(u) = |u| = (|u_1|, \ldots, |u_m|) \).

**EXAMPLE 1.** Consider the augmented multiplier function

\[
\phi(u, y, r) = |u|^T y + ru'u. \tag{3.1}
\]

Assumptions A.1 and A.2 are obviously satisfied. Since \( y \geq 0 \), also A.3 holds. Assumptions (A.4.1), (A.4.2), and (A.4.4) are also satisfied. Regarding (A.4.3), let

\[
\theta_{r_1}(u, r) = \phi(u, 0, r) - \phi(u, 0, r_1) = (r - r_1) u'u. \tag{3.2}
\]

For every \( c > 0 \) and \( \varepsilon > 0 \) we can take \( r > (c/\varepsilon^2) + r_1 \). Then, for every \( u \) such that \( \|u\| \geq \varepsilon \),

\[
\theta_{r_1}(u, r) > \frac{c u'u}{\varepsilon^2} \geq c \tag{3.3}
\]

and (A.4.3) holds.

To show that (A.5.1), (A.5.2), are satisfied, consider

\[
\phi(u, y, r) - \phi(u, z, s) = |u|^T(y - z) = (r - s) u'u. \tag{3.4}
\]

Minimizing the right-hand side of (3.4) with respect to \( u \) it can be verified that

\[
\phi(u, y, r) - \phi(u, z, s) \geq -\frac{1}{4} \frac{(y - z)^T(y - z)}{r - s} \tag{3.5}
\]

for every \( u \in \mathbb{R}^m \) and \( r > s \). By defining

\[
\psi(y - z, r - s) = \frac{1}{4} \frac{(y - z)^T(y - z)}{r - s} \tag{3.6}
\]

we obtain (A.5.1) and (A.5.2). Thus, this type of augmented multiplier function satisfies all the assumptions for strong duality (provided, of course, that the boundedness and stability conditions on (P) are also satisfied).

**EXAMPLE 2.** Our second example is the ordinary penalty function

\[
\phi(u, y, r) = ru'u. \tag{3.7}
\]
Assumptions A.1, A.2, A.3 can be immediately verified. Assumption A.4 is satisfied similar to Example 1, and A.5 can be established by defining $\psi(a, b) = 0$. Thus, if (P) satisfies the boundedness and stability conditions, strong duality holds for this type of augmented multiplier functions.

**Example 3.** Consider the following family of augmented multiplier functions:

$$
\varphi_p(u, y, r) = \sum_{i=1}^{m} y_i |u_i| + r|u_i|^p,
$$

where $p$ is a positive number. The augmented multiplier function in Example 1 is a member of this family with $p = 2$; Mangasarian [12] used $\varphi_4(u, y, r)$ without the absolute value, as an augmented multiplier function for inequality constrained problems.

It can be easily verified that A.1, A.2, A.3, (A.4.1), (A.4.2), and (A.4.4) are satisfied. To show that (A.4.3) holds, consider

$$
\theta_{r_1}(u, r) = \varphi(u, 0, r) - \varphi(u, 0, r_1) = (r - r_1) \sum_{i=1}^{m} |u_i|^p. \tag{3.9}
$$

For every $\varepsilon > 0$, there exists a $\delta = 0$ such that $\sum_{i=1}^{m} (u_i)^p \geq \delta \varepsilon$ for $\|u\| \geq \varepsilon$. This can be seen from the following relations:

$$
\sum_{i=1}^{m} |u_i|^p \geq \max_{i} |u_i|^p = \max_{i} (|u_i|^2)^{p/2} = \left[ \sum_{i=1}^{m} (u_i)^2 \right]^{p/2} \geq \frac{\|u\|^2}{m} \geq \varepsilon^{p-1} m^{p/2} \varepsilon. \tag{3.10}
$$

Letting $\delta = \varepsilon^{p-1}/m^{p/2}$ we obtain that for every $c > 0$ and $\varepsilon > 0$ taking $r > (c/\delta \varepsilon) + r_1$ assures that $\theta_{r_1}(u, r) > c$ for every $u \in \mathbb{R}^n$ such that $\|u\| \geq \varepsilon$. Thus (A.4.3) holds.

To show that A.5 is satisfied, consider

$$
\varphi(u, y, r) - \varphi(u, z, s) = \sum_{i=1}^{m} \{ |u_i| (y_i - z_i) + (r - s)(u_i)^p \}, \tag{3.11}
$$

where $r > s > 0$. Clearly,

$$
\varphi(u, y, r) - \varphi(u, z, s) \geq \sum_{i=1}^{m} \{ -|y_i - z_i| |u_i| + (r - s)|u_i|^p \}. \tag{3.12}
$$

Letting $p > 1$ and minimizing the right-hand side of (3.12) with respect to $u$ it can be verified that

$$
\varphi(u, y, r) - \varphi(u, z, s) \geq (1 - p)(p)^{p/(p-1)}(r - s)^{1-1/p} \sum_{i=1}^{m} |y_i - z_i|^p (p - 1) \tag{3.13}
$$
for every $u \in R^n$ and $r > s$. By defining $\psi(y - z, r - s)$ as the negative of the right-hand side in (3.13) we obtain that A.5 holds. For $p \leq 1$ it can be shown that A.5 does not hold.

**Example 4.** Consider the augmented multiplier function [6]

$$\varphi(u, y, r) = \sum_{i=1}^{m} |u_i| y_i + r(e^{\|y_i\|} - 1). \tag{3.14}$$

Assumptions A.1, A.2, A.3, (A.4.1), (A.4.2), (A.4.4), and A.5 are clearly satisfied. Assumption (A.4.3) can be shown to hold by observing that for $a \geq 0$,

$$e^a - 1 \geq a^2 \tag{3.15}$$

and by using arguments similar to those in Example 1.

Let us examine now some functions which do not satisfy all the sufficient conditions for strong duality, developed earlier in this work.

**Example 5.** Consider the Lagrangian function without penalty term

$$\varphi(u, y, r) = |u|^T y. \tag{3.16}$$

This function obviously does not satisfy assumptions (A.4.2), (A.4.3), and A.5.

**Example 6.** Consider now a so-called barrier function [14]:

$$\varphi(u, y, r) = \begin{cases} r \sum_{i=1}^{m} \frac{1}{yi - u_i} - \frac{1}{y_i} & \text{if } 0 \leq u_i < y_i, i = 1, \ldots, m \\ +\infty & \text{otherwise.} \end{cases} \tag{3.17}$$

This augmented multiplier function satisfies assumptions A.1, A.2, A.3. It does not satisfy A.4 since $\varphi(u, 0, r) = +\infty$.

**Example 7.** Let us consider the function

$$\varphi(u, y, r) = \sum_{i=1}^{m} |u_i| y_i + r \ln(1 + |u_i|). \tag{3.18}$$

This function satisfies assumptions A.1–A.4. We have

$$\varphi(u, y, r) - \varphi(u, z, s) = \sum_{i=1}^{m} |u_i|(y_i - z_i) + (r - s) \ln(1 + |u_i|). \tag{3.19}$$
Take $z_i = y_i + 1$ for $i = 1, \ldots, m$ and choose any $c > 0$. One can find $u \in \mathbb{R}^m$ such that

$$
\sum_{i=1}^m -|u_i| + (r-s) \ln(1 + |u_i|) < -c \tag{3.20}
$$

and, therefore, there is no function $\psi(y-z, r-s)$ with the properties required for A.5.

**Example 8.** It is of interest to observe that the multipliers can multiply a power of the constraints as in the following example:

$$
\varphi(u, y, r) = \sum_{i=1}^m |u_i|^p y_i + r|u_i|^q, \tag{3.21}
$$

where $q > \max\{p, 1\}$ and $p > 0$. It can be easily verified that assumptions A.1, A.2, A.3, and A.4 are satisfied. To show that A.5 is satisfied, consider

$$
\varphi(u, y, r) - \varphi(u, z, s) = \sum_{i=1}^m |u_i|^p (y_i - z_i) + (r-s)|u_i|^q. \tag{3.22}
$$

Minimizing the right-hand side of (3.22) with respect to $u$ it can be verified that

$$
\varphi(u, y, r) - \varphi(u, z, s) \geq - (r-s)^{-p/(q-p)} \left[ \left[ \frac{p}{q} \right]^{p/(q-p)} + \left[ \frac{p}{q} \right]^{q/(q-p)} \right] \times \sum_{i=1}^m |y_i - z_i|^{q/(q-p)} \tag{3.23}
$$

for every $u \in \mathbb{R}^m$ and $r > s$. By defining

$$
\psi(y-z, r-s) = (r-s)^{-p/(q-p)} \left[ \left[ \frac{p}{q} \right]^{p/(q-p)} \right. \left. + \left[ \frac{p}{q} \right]^{q/(q-p)} \right] \sum_{i=1}^m |y_i - z_i|^{q/(q-p)} \tag{3.24}
$$

we obtain (A.5.1) and (A.5.2).

**Example 9.** Our final example is a noncontinuous augmented Lagrangian function. Let $\delta(x)$ be the integer value of the absolute value of $x$, that is

$$
\delta(x) = \inf\{n : n \text{ is integer, } n \leq |x|\} \tag{3.25}
$$
and define

$$\varphi(u, y, r) = \sum_{i=1}^{n} r\delta(u_i)^2.$$ \hspace{1cm} (3.26)

It is easy to verify as in Example 2, that assumptions A.1, A.2, A.3, A.4, and A.5 are satisfied.

REFERENCES