Duality Theory for Maximizations with Respect to Cones

H. W. Corley

Department of Industrial Engineering, The University of Texas at Arlington, Arlington, Texas 76019
Submitted by V. Lakshmikantham

1. INTRODUCTION

Generalizations of Pareto optimality have been studied by a number of authors. In finite dimensions such work is exemplified by Corley [5], DaCunha and Polak [10], Goeffrion [13], Hartley [15], Lin [17], Tanino and Sawaragi [21], Wendell and Lee [22], and Yu [23]. The optimization of functions into possibly infinite dimensions has been considered by Borwein [2], Cesari and Suryanarayana [3], Christopeit [4], Corley [6, 7], Craven [8, 9], Hurwicz [16], Neustadt [19], and Ritter [20]. An extensive bibliography on Pareto optimality, its extensions, and applications is given in [1].

In this paper a duality theory is developed using the concept of saddlepoints for a problem in which the maximization of a function into possibly infinite dimensions is defined in terms of a cone. The results here extend the work of Tanino and Sawaragi [21] to infinite dimensions. A distinction is also made here between the notions of weak and strong optimality. Distinguishing between the two concepts allows the removal of the assumption of properness in [21] in establishing a relationship between the primal and dual problems, as well as permits additional duality relationships to be proved.

2. PRELIMINARIES

Throughout the paper let $X, Y, Z$ be real normed linear spaces, each with zero element $\theta$ as clear from context, and let $A \subseteq X$, $C \subseteq Y$, $D \subseteq Z$, $f: X \rightarrow Y$, and $g: X \rightarrow Z$. The following definitions and easily established properties concerning cones are needed.

DEFINITION 1. A set $C$ in $Y$ is a cone if $\lambda y \in C$ for all $y \in C$ and $\lambda \geq 0$. 
A pointed cone $C$ is one for which $C \cap -C = \{\emptyset\}$, and a convex cone $C$ is one for which $\lambda_1 y_1 + \lambda_2 y_2 \in C$ for all $y_1, y_2 \in C$ and $\lambda_1, \lambda_2 \geq 0$.

**Definition 2.** Let $C$ be a pointed cone in $Y$ and $B \subset Y$. For $y_1, y_2 \in Y$ write $y_1 \leq_c y_2$ if $y_2 - y_1 \in C$. If $y_2 - y_1 \in C \setminus \{\emptyset\}$, write $y_1 <_c y_2$; if $y_2 - y_1 \in C^o$ (the interior of $C$), write $y_1 \ll_c y_2$. The point $y_0 \in B$ is a strong maximal element of $B$ with respect to $C$, denoted $y_0 \in \text{max } B$, if there exists no $y \in B$ for which $y_0 <_c y$. Similarly, $y_0 \in B$ is a weak maximal element of $B$ with respect to $C$, denoted $y_0 \in \text{wmax } B$, if there exists no $y \in B$ for which $y_0 \ll_c y$. The set $\text{sup } B$ of strong supremal elements of $B$ with respect to $C$ is defined as $\text{sup } B = \text{max } \overline{B}$, where $\overline{B}$ is the closure of $B$, and the set $\text{wsup } B$ of weak supremal elements as $\text{wsup } B = \text{wmax } \overline{B}$.

The cone $C$ in Definition 2 need not be pointed to define the above order relations. However, pointedness simplifies slightly the statement of the definition, and subsequent results are restricted to pointed cones. It is obvious that a strong maximal element is a weak maximal element, while in $R^1$ there is no distinction between the two.

**Definition 3.** Let $Y^*$ denote the topological dual of $Y$ and $C$ be a cone in $Y$. The nonnegative dual cone of $C$ is the cone $C^+ = \{l \in Y^*: l(y) \geq 0 \text{ for all } y \in C\}$ in $Y^*$. In addition, let $B(Z, Y)$ denote the set of all bounded (continuous) linear functions $s: Z \to Y$ and let $D \subset Z$ be a cone. A function $s \in B(Z, Y)$ is said to be nonnegative with respect to $C$ and $D$, written $s \in B^+(Z, Y)$, if $s(D) \subset C$.

**Definition 4.** Let $C$ be a pointed cone in $Y$ and $f: X \to Y$. The function $f$ is $C$-concave on the convex set $A \subset X$ if $\lambda f(x_1) + (1 - \lambda) f(x_2) \leq_c f[\lambda x_1 + (1 - \lambda) x_2]$ for all $x_1, x_2 \in A$ and $\lambda \in [0, 1]$.

**Property 1.** Let $C$ be a pointed cone in $Y$, $B \subset Y$, and $y_0 \in B$. Then $y_0 \in \text{max } B$ if and only if $B \cap [C + y_0] = \{y_0\}$, and $y_0 \in \text{wmax } B$ if and only if $B \cap [C^o + y_0] = \emptyset$.

**Property 2.** Let $C$ be a pointed convex cone in $Y$ and $B \subset Y$. Then $\text{max } B = \text{max } (B - C)$ and $\text{wmax } B \subset \text{wmax } (B - C)$.

**Property 3.** Let $C$ be a pointed convex cone in $Y$. If $y_1 \in C^o$ and $y_2 \in C$, then $y_1 + y_2 \in C^o$.

3. The Primal, Saddlepoint, and Dual Problems

**Definition 5 (Primal).** Let $Y$ be ordered by the pointed cone $C \subset Y$ and let $D \subset Z$, $A \subset X$, $f: A \to Y$, $g: A \to Z$. Then the primal problem, written

\[
(P) \text{ maximize } f(x) \text{ subject to } g(x) \in D,
\]

\[
x \in A
\]
is to find all \( x_0 \in A \cap g^{-1}(D) \) for which \( f(x_0) \) is a maximal element of \( f[A \cap g^{-1}(D)] \). Any such \( x_0 \) for which \( f(x_0) \) is a weak maximal element of \( f[A \cap g^{-1}(D)] \) is called a weak maximal point for \( P \). A strong maximal point for \( P \) is similarly defined.

\( D \) is frequently taken to be a pointed cone in \( Z \) in Definition 5, in which case the constraint becomes \( g(x) \geq_D \theta \) and thus resembles a standard nonlinear programming constraint.

**Definition 6 (Saddlepoint).** Under the assumptions of Definition 5 and with \( D \) being also a cone in \( Z \), define the Lagrangian function \( L: X \times B(Z, Y) \to Y \) by \( L(x, s) = f(x) + sg(x) \). The point \( (x_0, s_0) \) is said to be a weak saddle point of \( L(x, s) \) if

\[
x_0 \in A \quad \text{and} \quad s_0 \in B^+(Z, Y);
\]

there does not exist \( s \in B^+(Z, Y) \) such that

\[
L(x_0, s) \leq_c L(x_0, s_0);
\]

there does not exist \( x \in A \) such that

\[
L(x_0, s_0) \leq_c L(x, s_0).
\]

Similarly, \( (x_0, s_0) \) is said to be a strong saddlepoint of \( L(x, s) \) if the inequality \( <_c \) replaces the inequality \( \leq_c \) in (2) and (3).

A saddlepoint in Definition 6 can be explained as follows. Condition (3) states that \( x_0 \) is a (weak or strong) maximal point of \( A \) for the objective function \( L(x, s_0) \) with respect to the cone \( C \). Furthermore, according to (2), \( s_0 \) is a maximal point of \( B^+(Z, Y) \) for \( L(x_0, s) \) with respect to the cone \( -C \). This last statement might be interpreted as \( s_0 \) being a minimal point with respect to \( C \). Definition 6 thus describes a generalization of the saddlepoint problem associated with the maximization of a real-valued objective functions such as presented in [18].

**Definition 7.** The weak dual problem for \( P \) is the problem

\[
(WD) \text{ minimize } \bigcup_{s \in B^+(Z, Y)} \text{wsup}(f + sg)(A).
\]

Minimization here refers to determining the set of weak maximal elements of the set \( \bigcup_{s \in B^+(Z, Y)} W(s) \) with respect to the cone \( -C \), where \( W(s) = \)
wsup(\(f + sg\))(A). This set of weak maximal elements is written wmin[\(\bigcup_{s \in B^+(Z, Y)} W(s)\)]. The strong dual problem SD for \(P\) is the problem

\[
\text{(SD) minimize } \left[ \bigcup_{s \in B^+(Z, Y)} \sup(f + sg)(A) \right],
\]

where now strong maximal elements are to be determined with respect to \(-C\).

When \(Y = R^1\), \(W(s)\) in WD contains a single number (perhaps \(-\infty\) with some minor changes in definitions). Therefore \(\bigcup_{s \in B^+(Z, Y)} W(s)\) is the range of the (extended) real-valued function \(\sup_{x \in A}[f(x) + sg(x)]\), where \(\sup\) is now used in the usual sense. An analogous statement holds for SD. Both WD and SD thus reduce to the standard Lagrangian dual problem in mathematical programming as presented, say, in [14, 18]. Most subsequent duality relations, however, involve only WD.

4. SADDLEPOINT OPTIMALITY CONDITIONS

Sufficient conditions are first stated.

THEOREM 1. Suppose that \(C\) is a nontrivial pointed convex cone in \(Y\) and \(D\) is a closed convex cone in \(Z\). If \((x_0, s_0)\) is a strong saddlepoint of \(L(x, s)\), then \(x_0\) is a strong maximal point for \(P\).

Proof. It is first shown that \(g(x_0) \in D\). From (2) there is no \(s \in B^+(Z, Y)\) such that

\[
(s_0 - s) g(x_0) \in C \setminus \{\theta\}. \tag{4}
\]

To arrive at a contradiction to (4) assume that \(g(x_0) \in D\). Applying a separation theorem [11, p. 417] to \(D\) and the compact set \(\{g(x_0)\}\) yields the existence of constants \(\alpha\) and \(\delta > 0\) and a \(v \in Z^*\) such that for every \(z \in D\), \(v(z) \leq \alpha - \delta < \alpha \leq v[g(x_0)]\). If there exists \(z \in D\) such that \(v(z) > 0\), then for some \(\lambda > 0\), \(\lambda z \in D\) and \(\lambda v(z) > \alpha - \delta\). Hence \(\alpha - \delta \leq 0\). But \(\theta \in D\), so \(\alpha = \delta > 0\). Choose \(u = -v\) to obtain \(u[g(x_0)] < 0\) and \(u(z) \geq 0\) for all \(z \in D\). Now fix \(\theta \neq y_1 \in C\) and let \(s_1 \in B^+(Z, Y)\) be defined by \(u(z) y_1\). Set \(s = s_1 + s_0 \in B^+(Z, Y)\). Since \(u(g(x_0)) < 0\), \(\theta \neq (s_0 - s) g(x_0) = -s_1 g(x_0) \in -C\) in violation of (4). Thus

\[
g(x_0) \in D. \tag{5}
\]

Hence \(x_0 \in A \cap g^{-1}(D)\) from (1).
We next prove that \( x_0 \) is a strong maximal point for \( P \). To do so, first let \( s \) be the zero functional. Then from (5) and (2)

\[
  s_0 g(x_0) = \theta. \quad (6)
\]

Now assume that the feasible point \( x_0 \) is not a strong maximal point. Then there exists \( x \in A \) with \( g(x) \in D \) for which \( f(x) - f(x_0) \in C - \{\theta\} \). Using (6), the fact that \( s_0 g(x) \in C \), and Definition 1 we conclude that \( L(x, s_0) - L(x_0, s_0) \in C - \{\theta\} \) in contradiction to (3). Consequently \( x_0 \) is a strong maximal point for \( P \).

If \( s_0 \) is identically zero in the saddlepoint \((x_0, s_0)\) of Theorem 1, then just as in the real-valued case \( x_0 \) is a strong maximal point for the problem \( \max_{x \in A} f(x) \). Likewise, the following corollary to the proof of Theorem 1 should be recognized as a generalization of the procedure in mathematical programming of maximizing the Lagrangian function.

**Corollary 1.** Suppose that \( C \) is a pointed convex cone in \( Y \), \( D \) is a closed convex cone in \( Z \), and there exists \( s_0 \in B^+(Z, Y) \) such that (i) \( x_0 \) is a strong maximal point for the problem \( \max_{x \in A} [f(x) + s_0 g(x)] \), (ii) \( g(x_0) \in D \), and (iii) \( s_0 g(x_0) = \theta \). Then \( x_0 \) is a strong maximal point for \( P \).

Necessary saddlepoint optimality conditions are now given. Theorem 2 involves weak maximal points as compared to Theorem 1.

**Theorem 2.** Let \( C \) be a pointed convex cone in \( Y \) with \( C^0 \neq \emptyset \). Suppose that \( D \) is a pointed convex cone in \( Z \) with \( D^0 \neq \emptyset \), \( A \) is a convex set, \( f \) is \( C \)-concave on \( A \), \( g \) is \( D \)-concave on \( A \), and \( A \cap g^{-1}(D^0) \neq \emptyset \). If \( x_0 \) is a weak maximal point for \( P \), then there exists \( s_0 \in B^+(Z, Y) \) such that \((x_0, s_0)\) is a weak saddle point of \( L(x, s) \) and (6) holds.

**Proof.** Let \( E = A \cap g^{-1}(D) \), which is convex. Apply Property 2 to obtain \( f(x_0) \in \text{wmax}[f(E) - C] \), where \( f(E) - C \) is convex from the convexity of \( E \) and \( C \)-concavity of \( f \). From Property 1, \([f(E) - C] \cap [C^0 + f(x_0)] = \emptyset\) for the convex sets \( f(E) - C \) and \( C^0 + f(x_0) \). A separation theorem \([12, p. 118]\) now yields the existence of a nonzero \( l \in Y^* \) and a real number \( \alpha \) such that \( l(y) \leq \alpha \) for all \( y \in f(E) - C \) and \( l(y) > \alpha \) for all \( y \in C^0 + f(x_0) \). Since \( \theta \in C \), we have that

\[
  l(f(x_0)) \leq \alpha. \quad (7)
\]

Let \( y \in C^0 \). Then

\[
  l[y + f(x_0)] = l(y) + l(f(x_0)) > \alpha. \quad (8)
\]
It follows from (7) and (8) that \( l(y) > 0 \) for all \( y \in C^o \). An argument utilizing the continuity of \( l \) and the convexity of \( C \) further establishes that \( l \in C^+ \). Next write

\[
l[f(x) - y_1] \leq l[f(x_0) + y_2] \quad \text{for all } x \in E, \, y_1 \in C, \, y_2 \in C^o. \tag{9}
\]

Take \( y_1 = y_2 \in C^o \) in (9) to obtain that \( l[f(x)] \leq l[f(x_0)] \) for all \( x \in E \), so in \( R^1 \) \( x_0 \) maximizes \( l[f(x)] \) subject to \( x \in A \) and \( g(x) \in D \). By a standard Lagrange multiplier theorem in [18], there exists \( u_0 \in D^+ \) with \( u_0 g(x_0) = 0 \) for which

\[
l[f(x_0)] + u_0 g(x_0) \geq l[f(x)] + u_0 g(x) \quad \text{for all } x \in A. \tag{10}
\]

Since \( C^o \neq \emptyset \) and \( l(y) > 0 \) for all \( y \in C^o \), choose \( y_0 \in C^o \) such that \( l(y_0) = 1 \). This choice is possible since \( C^o \cup \{ \theta \} \) is also a cone. Define \( s: Z \to Y \) by \( s_0(z) = u_0(z) y_0 \). Then \( s_0 \in B^+(Z, Y) \) and \( s_0 g(x_0) = u_0[g(x_0)] y_0 = \theta \), so (6) is established. Moreover, \( l[f(x)] - l[f(x_0)] + s_0 g(x) - s_0 g(x_0) \leq 0 \) for all \( x \in A \) from (10). Since \( l(y) > 0 \) for all \( y \in C^o \), then for all \( x \in A \) we have that \( f(x) - f(x_0) + s_0 g(x) - s_0 g(x_0) \in C^o \). Thus (3) is proved.

To demonstrate (2), suppose to the contrary that there exists \( s \in B^+(Z, Y) \) for which

\[
f(x_0) + s_0 g(x_0) - f(x_0) - sg(x_0) \in C^o. \tag{11}
\]

But then (6), (11), and Property 3 imply that \( sg(x_0) \in -C^o \). Since \( s \in B^+(Z, Y) \) and \( g(x_0) \in D \), we thus arrive at a contradiction to complete the proof.

The following corollary is stated for reference in the next section.

**Corollary 2.** Suppose that the hypotheses of Theorem 2 are satisfied. If \( x_0 \) is a weak maximal point for \( P \), then there exists \( s_0 \in B^+(Z, Y) \) such that (i) \( x_0 \) is a weak maximal point for the problem maximize \( x \in A \) \( [f(x) + s_0 g(x)] \), (ii) \( g(x_0) \in D \), and (iii) \( s_0 g(x_0) = \theta \).

**5. Duality Relationships**

Theorem 3 below is a generalization of the weak duality theorem of mathematical programming which states that the value of the primal objective function at any feasible point is never larger than the value of the dual objective function at any feasible point.

**Theorem 3.** Suppose that \( C \) is a pointed convex cone in \( Y \), \( x_0 \) is feasible to \( P \) (i.e., \( x_0 \in A \cap g^{-1}(D) \)), and \( s_0 \) is feasible to \( WD \) (i.e., \( s_0 \in B^+(Z, Y) \)).
Then there does not exist $y_0 \in W(s_0)$ such that $y_0 \prec_c f(x_0)$. A similar relationship involving $\prec_c$ holds for SD.

Proof. The result is proved only for WD. Suppose to the contrary that there exists $y_0 \in W(s_0)$ such that

$$f(x_0) - y_0 \in C^\circ. \tag{12}$$

Since $y_0 \in \text{wsup}(f + s_0 g)(A)$, there does not exist $x \in A$ for which $f(x) + s_0 g(x) - y_0 \in C^\circ$. But $x_0 \in A$, so in particular

$$f(x_0) + s_0 g(x_0) - y_0 \notin C^\circ. \tag{13}$$

Upon adding $s_0 g(x_0) \in C$ to (12) and using Property 3, a contradiction to (13) is obtained and the result is established.

$D$ need not be a cone, $A$ does not have to be a convex, and no concavity restrictions are placed on $f$ or $g$ in Theorem 3. A stronger result requires these assumptions. Theorem 4(a) below may be very loosely interpreted in the context of mathematical programming as follows. A solution to the primal implies the existence of a solution to the dual, and the values of the two objective functions are equal. Theorem 4(b) is somewhat similar; it should be obvious that the conclusion of (b) is equality in $R^1$.

**Theorem 4.** Let $C$ be a pointed convex cone in $Y$ with $C^\circ \neq \emptyset$. Suppose that $D$ is a pointed convex cone in $Z$ with $D^\circ \neq \emptyset$, $A$ is a convex set, $f$ is $C$-concave on $A$, $g$ is $D$-concave on $A$, and $A \cap g^{-1}(D^\circ) \neq \emptyset$.

(a) If $x_0$ is a weak maximal point for $P$, then

$$f(x_0) \in \text{wmin} \left[ \bigcup_{s \in B^+(Z,Y)} W(s) \right]. \tag{14}$$

(b) Suppose $y_0 \in \text{wmin} [\bigcup_{s \in B^+(Z,Y)} W(s)]$. Then for every weak maximal point $x_0$ for $P$, neither $f(x_0) \prec_c y_0$ nor $y_0 \prec_c f(x_0)$.

Proof. (a) From Corollary 2 there exists $s_0 \in B^+(Z,Y)$ for which $f(x_0) \in W(s_0)$. To verify (14), assume the contrary. Then there exists $s_1 \in B(Z,Y)$ and $y_1 \in W(s_1)$ such that

$$f(x_0) - y_1 \in C^\circ. \tag{15}$$

Add $s_1 g(x_0) \in C$ to (15) to get

$$f(x_0) + s_1 g(x_0) - y_1 \in C^\circ \tag{16}$$

by Property 3. But (16) contradicts the assumption that $y_1 \in W(s_1)$, so (14) is established.
(b) Since $y_0 \in \text{wmin}[\bigcup_{s \in B^+(Z, Y)} W(s)]$, there exists $s \in B^+(S, Y)$ for which $y_0 \in W(s)$. Let $x_0$ be a weak maximal point for $P$. It follows immediately from Theorem 3 that $f(x_0) - y_0 \in C^\circ$. On the other hand, from Corollary 2 there is an $s_0 \in B^+(Z, Y)$ for which $f(x_0) \in W(s_0)$. Since $y_0 \in \text{wmin}[\bigcup_{s \in B^+(Z, Y)} W(s)]$, $y_0 - f(x_0) \in C^\circ$. The conclusion of (b) now follows.

**Corollary 3.** Suppose that the assumptions of Theorem 4 are satisfied. If $\text{wmin}[\bigcup_{s \in B^+(Z, Y)} W(s)] = \emptyset$ for $WD$, then problem $P$ has neither weak nor strong maximal points.

**Proof.** Apply Theorem 4(a). Then $P$ has no weak maximal points and hence no strong maximal points.

**Example**

The following simple example in $R^2$ illustrates the geometric nature of the various definitions and duality relations. In the primal problem $P$ let $X = Y = Z = R^2$, $C = D = R^2_+ = \{(a_1, a_2): a_1, a_2 \geq 0\}$, $A = \{(a_1, a_2): -1 \leq a_1 \leq 1, a_2 = 0\}$, and $f = g = I$ (the identity function). Since $A \cap g^{-1}(D) = \{(a_1, a_2): 0 \leq a_1 \leq 1, a_2 = 0\}$, the set of weak maximal elements of $A \cap g^{-1}(D)$, as well as the set of weak maximal points for $P$, is

$$\{(a_1, a_2): 0 \leq a_1 \leq 1, a_2 = 0\}.$$  \hfill (17)

$B^+(X, Y)$ in $WD$ can be identified with the set of real $2 \times 2$ matrices with nonnegative components; i.e.,

$$B^+(Z, Y) = \left\{ \begin{bmatrix} a \\ c \\ b \\ d \end{bmatrix} : a, b, c, d \geq 0 \right\}.$$

It follows that for fixed $s = [a \\ c \\ b \\ d] \in B^+(Z, Y)$, $(f + sg)(A) = \{(a_1(a + 1, c): -1 \leq a_1 \leq 1\}$, and therefore

$$\text{wsup}(f + sg)(A) = \{(a_1, a_2): a_2 = 0\} \quad \text{if} \quad c > 0$$

$$\{(a_1, 0): -(a + 1) \leq a_1 \leq a + 1\} \quad \text{if} \quad c = 0.$$

Thus $\bigcup_{s \in B^+(Z, Y)} W(s) = \{(a_1, a_2): a_1 > 1, a_2 > 0\} \cup \{(a_1, a_2): a_2 = 0\}$ from which

$$\text{wmin}\left[ \bigcup_{s \in B^+(Z, Y)} W(s) \right] = \{(a_1, a_2): a_2 = 0\}.$$ \hfill (18)

By comparing (17) and (18), Theorems 3 and 4 are easily verified for this example.
REFERENCES


7. H. W. CORLEY, Optimality conditions for maximizations with respect to cones, submitted for publication.


