DUALITY THEOREMS FOR CONVEX FUNCTIONS

BY R. T. ROCKAFELLAR

Communicated by A. M. Gleason, September 19, 1963

Let \( F \) be a finite-dimensional real vector space. A proper convex function on \( F \) is an everywhere-defined function \( f \) such that \(-\infty < f(x) \) for all \( x \), \( f(x) < \infty \) for at least one \( x \), and

\[
f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)
\]

for all \( x_1 \) and \( x_2 \) when \( 0 < \lambda < 1 \). Its effective domain is the convex set \( \text{dom } f = \{ x | f(x) < \infty \} \). Its conjugate \([2; 3; 6; 7]\) is the function \( f^* \) defined by

\[
(1) \quad f^*(x^*) = \sup\{ (x, x^*) - f(x) | x \in F \}
\]

for each \( x^* \in F^* \), where \( F^* \) is the space of linear functionals on \( F \). The conjugate function is proper convex on \( F^* \), and is always lower semi-continuous. If \( f \) itself is l.s.c., then \( f \) coincides with the conjugate \( f^{**} \) of \( f^* \) (where \( F^{**} \) is identified with \( F \)). These facts and definitions have obvious analogs for concave functions, with “inf” replacing “sup” in (1).

Suppose \( f \) is l.s.c. proper convex on \( F \) and \( g \) is u.s.c. proper concave on \( F \). If

\[
\text{ri } (\text{dom } f) \cap \text{ri } (\text{dom } g) \neq \emptyset,
\]

where \( \text{ri } C \) denotes the relative interior of a convex set \( C \), then

\[
\inf\{ f(x) - g(x) | x \in F \} = \max\{ g^*(x^*) - f^*(x^*) | x^* \in F^* \}.
\]

This was proved by Fenchel \([3, \text{p. 108}]\) (reproduced in \([5, \text{p. 228}]\)). The purpose of this note is to announce the following more general fact.

**Theorem 1.** Let \( F \) and \( G \) be finite-dimensional partially-ordered real vector spaces in which the nonnegative cones \( P(F) \) and \( P(G) \) are polyhedral. Let \( A \) be a linear transformation from \( F \) to \( G \). Let \( f \) be a proper convex function on \( F \) and let \( g \) be a proper concave function on \( G \). If there exists at least one \( x \in \text{ri } (\text{dom } f) \) such that \( x \geq 0 \) and \( Ax \geq y \) for some \( y \in \text{ri } (\text{dom } g) \), then

\[
\inf \{ f(x) - g(y) | x \geq 0, Ax \geq y \} = \max\{ g^*(y^*) - f^*(x^*) | y^* \geq 0, A^*y^* \leq x^* \},
\]

where \( A^* \) is the adjoint of \( A \).

\[1\] The material in this note stems from the author’s recent doctoral dissertation at Harvard. Support was provided under grant AF-AFOSR-62-348 at the Computation Center, Massachusetts Institute of Technology.
The partial-orderings are, of course, assumed to be compatible with the vector structure. The orderings in \( F^* \) and \( G^* \) are dual to those in \( F \) and \( G \), i.e. \( P(F^*) \) consists of the \( x^* \) such that \( (x, x^*) \geq 0 \) whenever \( x \geq 0 \), etc.

In particular, any \( F \) and \( G \) can be supplied with the degenerate partial-orderings in which \( P(F) = F \) and \( P(G) = \{0\} \), so that \( P(F^*) = \{0\} \) and \( P(G^*) = G^* \). If Theorem 1 is then invoked, one obtains

**Corollary 1.** Assume the notation of Theorem 1, but omit the partial-ordering of \( F \) and \( G \). If \( Ax \in \text{ri}(\text{dom } g) \) for at least one \( x \in \text{ri}(\text{dom } f) \), then

\[
\inf \{f(x) - g(Ax) \mid x \in F\} = \max \{g^*(y^*) - f^*(A^*y^*) \mid y^* \in G^*\}. 
\]

When \( F = G \) and \( A = I \), Corollary 1 furnishes a slightly generalized version of Fenchel's theorem not requiring semi-continuity.

Another new result is the following.

**Corollary 2.** Assume the notation of Theorem 1, and suppose also that \( \text{dom } f, \text{dom } f^*, \text{dom } g \) and \( \text{dom } g^* \) are all linear manifolds. If any one of the following is true,

(a) \( \inf \{f(x) - g(y) \mid x \geq 0, Ax \geq y\} \) is finite,

(b) \( \sup \{g^*(y^*) - f^*(x^*) \mid y^* \geq 0, A^*y^* \leq x^*\} \) is finite,

(c) \( \{\langle x, y \rangle \mid 0 \leq x \in \text{dom } f, Ax \geq y \in \text{dom } g\} \neq \emptyset \) and

\[
\{\langle y^*, x^* \rangle \mid 0 \leq y^* \in \text{dom } g^*, A^*y^* \leq x^* \in \text{dom } f^*\} \neq \emptyset,
\]

then all three are true. Moreover, then the "\( \inf \)" and "\( \sup \)" are equal and both are attained.

This corollary is deduced from Theorem 1 and its dual (in which the roles of the starred and unstarred elements are reversed), using the trivial fact that \( \text{ri } C = C \) when \( C \) is a linear manifold. The appropriate semi-continuity of \( f \) and \( g \), which one needs in order that \( f^{**} = f \) and \( g^{**} = g \) in the dual of Theorem 1, is also a consequence of the hypothesis, because a convex or concave function is actually continuous on any relatively open set where it is finite-valued.

Fix any \( b^* \in F^* \) and \( c \in G \). Let \( f(x) = (x, b^*) \). Let \( g(y) = 0 \) if \( y = c \) and \( g(y) = -\infty \) if \( y \neq c \). Then \( f^*(x^*) = 0 \) if \( x^* = b^* \), \( f^*(x^*) = \infty \) if \( x^* \neq b^* \), and \( g^*(y^*) = (c, y^*) \). In this situation, Corollary 2 yields the important existence and duality theorems of Gale, Kuhn and Tucker for linear programs (see [4]). Many other convex programming results, both new and old, are also contained in the theorem and its corollaries. The common extremum value can be characterized as a minimax.
Theorem 1 is proved by way of a simpler theorem of some interest in itself.

**Theorem 2.** Let $h$ be a proper convex function on a finite-dimensional real vector space $E$ and let $K$ be a polyhedral convex cone in $E$. If $\text{ri}(\text{dom } h)$ intersects $K$, then

$$(3) \quad \inf \{ h(z) \mid z \in K \} = - \min \{ h^*(z^*) \mid z^* \in K^* \},$$

where $K^* = \{ z^* \in E^* \mid (z, z^*) \geq 0 \text{ for all } z \in K \}$.

An outline of the proof of Theorem 2 follows. One shows first that no generality is lost if $h$ is assumed l.s.c. Then one observes that (3) holds whenever $\text{ri}(\text{dom } h)$ actually intersects $\text{ri } K$. This is obtained from Fenchel's theorem by taking $f(z) = h(z)$, $g(z) = 0$ if $z \in K$, $g(z) = -\infty$ if $z \notin K$. The proof proceeds now by induction on the dimension of $K$. If $\text{dim } K = 0$, then $\text{ri } K = K$ trivially, so (3) is true. Assume next that (3) is true for cones of dimension less than $r$, and that $\text{dim } K = r$. It may be supposed that $\text{ri}(\text{dom } h)$ does not intersect $\text{ri } K$, since the other case has been covered. A separation argument then produces a $z^* \in K^*$ such that $-z^* \in K^*$ and

$$(4) \quad (z, z^*) \leq 0 \quad \text{for all } z \in \text{dom } h.$$

Let $K_0 = \{ z \in K \mid (z, z^*_0) = 0 \}$. Then $K_0$ is a polyhedral convex cone, and $\text{dim } K_0 < r$. Hence by the induction hypothesis

$$(5) \quad \inf \{ h(z) \mid z \in K_0 \} = - \min \{ h^*(z^*) \mid z^* \in K_0^* \}.$$

It is easy to see from the properties of $z^*_0$ that the left sides of (3) and (5) are the same. On the other hand, because $K$ is polyhedral,

$$K_0^* = \{ z^* - \lambda z^*_0 \mid z^* \in K^*, \lambda \geq 0 \}.$$

Moreover, (4) and definition (1) imply that $h^*(z^* - \lambda z^*_0) \geq h^*(z^*)$ for all $z^* \in E^*$ and $\lambda \geq 0$. Therefore the minimum of $h^*$ on $K_0^*$ can be achieved on $K^*$ itself, so that the right sides of (3) and (5) are equivalent, too.

Theorem 1 is deduced from Theorem 2 by choosing

$$E = \{ z = \langle x, y \rangle \mid x \in F, y \in G \}, \quad h(z) = f(x) - g(y),$$

$$K = \{ \langle x, y \rangle \mid x \geq 0, Ax \geq y \}.$$


University of Texas