DUALITY THEOREMS FOR CONVEX FUNCTIONS

BY

R. T. ROCKAFELLAR

Reprinted from the
BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY
January, 1964, Vol. 70, No. 1
Pp. 189-192

DUALITY THEOREMS FOR CONVEX FUNCTIONS

BY R. T. ROCKAFELLAR[1]

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

$$f(\lambda x_1 + (1-\lambda)x_2) \leqq \lambda f(x_1) + (1-\lambda)f(x_2)$$

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

$$(1) \qquad f^*(x^*) = \sup\{(x, x^*) - f(x) \mid x \in F\} \qquad \text{for each } x^* \in F^*,$$

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 ri C denotes the relative interior of a convex set C, then

$$\inf\{f(x) - g(x) \mid x \in F\} = \max\{g^*(x^*) - f^*(x^*) \mid x^* \in F^*\}.$$

This was proved by Fenchel [3, p. 108] (reproduced in [5, 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 $x \geqq 0$ and $Ax \geqq y$ for some $y \in \text{ri}\,(\text{dom}\, g)$, then*

$$(2) \qquad \begin{aligned} &\inf\{f(x) - g(y) \mid x \geqq 0,\ Ax \geqq y\} \\ &\qquad = \max\{g^*(y^*) - f^*(x^*) \mid y^* \geqq 0,\ A^*y^* \leqq x^*\}, \end{aligned}$$

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 $(x, x^*) \geqq 0$ whenever $x \geqq 0$, 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* $Ax \in \mathrm{ri}(\mathrm{dom}\, g)$ *for at least one* $x \in \mathrm{ri}(\mathrm{dom}\, f)$, *then*

$$(2') \quad \inf\{f(x) - g(Ax) \mid x \in F\} = \max\{g^*(y^*) - f^*(A^*y^*) \mid y^* \in G^*\}.$$

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* dom f, dom f^*, dom g *and* dom g^* *are all linear manifolds. If any one of the following is true,*

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

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

(c) $\{\langle x, y\rangle \mid 0 \leqq x \in \mathrm{dom}\, f,\ Ax \geqq y \in \mathrm{dom}\, g\} \neq \emptyset$ *and*

$$\{\langle y^*, x^*\rangle \mid 0 \leqq y^* \in \mathrm{dom}\, g^*,\ A^*y^* \leqq x^* \in \mathrm{dom}\, f^*\} \neq \emptyset,$$

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 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* $\mathrm{ri}(\mathrm{dom}\ h)$ *intersects K, then*

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

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

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 $\mathrm{ri}(\mathrm{dom}\ h)$ actually intersects $\mathrm{ri}\ K$. 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 $\dim K = 0$, then $\mathrm{ri}\ K = K$ trivially, so (3) is true. Assume next that (3) is true for cones of dimension less than r, and that $\dim K = r$. It may be supposed that $\mathrm{ri}(\mathrm{dom}\ h)$ does not intersect $\mathrm{ri}\ K$, since the other case has been covered. A separation argument then produces a $z_0^* \in K^*$ such that $-z_0^* \notin K^*$ and

$$(z, z_0^*) \leqq 0 \qquad \text{for all } z \in \mathrm{dom}\ h. \tag{4}$$

Let $K_0 = \{z \in K \mid (z, z_0^*) = 0\}$. Then K_0 is a polyhedral convex cone, and $\dim K_0 < r$. Hence by the induction hypothesis

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

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,

$$K_0^* = \{z^* - \lambda z_0^* \mid z^* \in K^*, \lambda \geqq 0\}.$$

Moreover, (4) and definition (1) imply that $h^*(z^* - \lambda z_0^*) \geqq h^*(z^*)$ for all $z^* \in E^*$ and $\lambda \geqq 0$. 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

$$E = \{z = \langle x, y\rangle \mid x \in F, y \in G\}, \quad h(z) = f(x) - g(y),$$
$$K = \{\langle x, y\rangle \mid x \geqq 0, Ax \geqq y\}.$$

REFERENCES

1. E. Eisenberg, *Duality in homogeneous programming*, Proc. Amer. Math. Soc. 12 (1961), 783–787.

2. W. Fenchel, *On conjugate convex functions*, Canad. J. Math. 1 (1949), 73–77.

3. ———, *Convex cones, sets and functions*, multilith lecture notes, Princeton Univ., Princeton, N. J., 1953.

4. A. J. Goldman and A. W. Tucker, *Theory of linear programming*, pp. 53–98, Annals of Mathematics Studies no. 38, Princeton Univ. Press, Princeton, N. J., 1956.

5. S. Karlin, *Mathematical methods and theory in games, programming and economics*, Vol. I, Addison-Wesley, Reading, Mass., 1960.

6. J.-J. Moreau, *Fonctions convexes en dualité*, Faculté des Sciences de Montpellier, Séminaires de Mathématiques, 1962 (multigraph).

7. ———, *Fonctions convexes duales et points proximaux dans un espace hilbertien*, C.R. Acad. Sci. Paris **255** (1962), 2897–2899.

University of Texas