Rif copy

ROCKAFELLAR, R. T.
1967
CAN. J. MATH.
Vol. 19, pp. 200–205

CONJUGATES AND LEGENDRE TRANSFORMS OF CONVEX FUNCTIONS

R. T. Rockafellar

Reprinted from

The Canadian Journal of Mathematics

CONJUGATES AND LEGENDRE TRANSFORMS OF CONVEX FUNCTIONS

R. T. ROCKAFELLAR

1. Introduction. Fenchel's conjugate correspondence for convex functions may be viewed as a generalization of the classical Legendre correspondence, as indicated briefly in **(6)**. Here the relationship between the two correspondences will be described in detail. Essentially, the conjugate reduces to the Legendre transform if and only if the subdifferential of the convex function is a one-to-one mapping. The one-to-oneness is equivalent to differentiability and strict convexity, plus a condition that the function become infinitely steep near boundary points of its effective domain. These conditions are shown to be the very ones under which the Legendre correspondence is well-defined and symmetric among convex functions. Facts about Legendre transforms may thus be deduced using the elegant, geometrically motivated methods of Fenchel. This has definite advantages over the more restrictive classical treatment of the Legendre transformation in terms of implicit functions, determinants, and the like.

2. Statement of results. Let h be a differentiable real-valued function given on a non-empty open set U in R^n. Let U^* be the image of U under the gradient map $\nabla h: x \rightarrow \nabla h(x)$. If ∇h is one-to-one, the function

$$h^*(x^*) = \langle x^*, (\nabla h)^{-1}(x^*)\rangle - h((\nabla h)^{-1}(x^*)) \tag{2.1}$$

is well-defined on U^* (where $\langle \cdot, \cdot \rangle$ denotes the ordinary inner product in R^n). The pair (U^*, h^*) is called the *Legendre transform* of (U, h).

Legendre transforms, of course, have had a long history in the calculus of variations; see for example (**3**, pp. 231–242; **4**, pp. 32–39; **8**, p. 27). More recently, they have been employed by Dennis **(5)** in the study of convex programs. Dennis assumes that U is convex, h is strictly convex, and ∇h is continuous. Unfortunately, these properties are not necessarily inherited by (U^*, h^*). In fact U^* might not be convex (see the example in §4 below). The continuity assumption on Δh is actually redundant; it is known that, if h is convex and each of its partial derivatives exists throughout U, then h is differentiable and ∇h is continuous on U (see **7**, p. 86).

We shall say that (U, h) is a *convex function of Legendre type* on R^n if U is a

Received October 19, 1965. This work was supported in part by the Air Force Office of Scientific Research.

non-empty open convex set in R^n, h is strictly convex and differentiable on U, and

$$\lim_{\lambda \downarrow 0} \frac{d}{d\lambda} h(\lambda a + (1 - \lambda)x) = -\infty \tag{2.2}$$

whenever $a \in U$ and x is a boundary point of U. The virtue of condition (2.2), which is automatically satisfied when $U = R^n$, is that it leads to a *symmetric* correspondence. We shall prove the following theorem in §3.

THEOREM 1. *Let (U, h) be a convex function of Legendre type on R^n. The Legendre transform (U^*, h^*) is then well-defined. It is another convex function of Legendre type on R^n, and $\nabla h^* = (\nabla h)^{-1}$ on U^*. The Legendre transform of (U^*, h^*) is (U, h) again.*

We shall now explain the parallel "conjugate" correspondence, introduced by Fenchel in **(6)**. Let f be a *proper convex function on* R^n, i.e. an everywhere-defined function with values in $(-\infty, +\infty]$, not identically $+\infty$, such that

$$f(\lambda x + (1 - \lambda)y) \leqslant \lambda f(x) + (1 - \lambda)f(y) \qquad \text{when } 0 < \lambda < 1.$$

Suppose also that f is lower semi-continuous (l.s.c.), in other words that $\{x | f(x) \leqslant \mu\}$ is closed in R^n for every $\mu \in R$. The function f^* on R^n defined by

$$f^*(x^*) = \sup_x \{\langle x, x^* \rangle - f(x)\} \tag{2.3}$$

is called the *conjugate* of f. Fenchel proved (in somewhat different notation) that f^* is again a l.s.c. proper convex function on R^n, and that the conjugate of f^* is in turn f.

A vector x^* is said to be a *subgradient* of f at x if

$$f(y) \geqslant f(x) + \langle y - x, x^* \rangle \qquad \text{for all } y \in R^n. \tag{2.4}$$

The set of these subgradients x^* is denoted by $\partial f(x)$. The (multiple-valued) mapping $\partial f: x \rightarrow \partial f(x)$ is the *subdifferential* of f. By (2.3) and (2.4),

$$f^*(x^*) = \langle x, x^* \rangle - f(x) \qquad \text{if and only if } x^* \in \partial f(x). \tag{2.5}$$

Since the condition on the left in (2.5) can be expressed symmetrically, we have

$$x^* \in \partial f(x) \qquad \text{if and only if } x \in \partial f^*(x^*). \tag{2.6}$$

Formula (2.5) can be viewed as a generalization of (2.1). We shall see below that the two formulas are equivalent when ∂f is *one-to-one* (from its domain to its range), i.e. when $x_1^* \in \partial f(x_1)$ and $x_2^* \in \partial f(x_2)$ imply either that $x_1 \neq x_2$ and $x_1^* \neq x_2^*$, or that $x_1 = x_2$ and $x_1^* = x_2^*$.

The two function correspondences can be tied together by Fenchel's closure operation. Given a finite convex function h on an open convex set U in R^n and any $a \in U$, the function f on R^n defined by

$$f(x) = \begin{cases} h(x) & \text{if } x \in U, \\ \lim_{\lambda \downarrow 0} h(\lambda a + (1 - \lambda)x) & \text{if } x \in (\mathrm{cl}\, U) \backslash U, \\ +\infty & \text{if } x \notin \mathrm{cl}\, U, \end{cases} \tag{2.7}$$

will be called the *closed extension* of (U, h). It is a l.s.c. proper convex function. It does not depend on the particular a chosen; see (**6** and **7**, pp. 74–79). The conjugate of f can be compared with the Legendre transform of (U, h) when the latter is defined. In this connection we have the following result.

THEOREM 2. *Let f be any l.s.c. proper convex function on R^n. Then ∂f is one-to-one if and only if f is the closed extension of a convex function of Legendre type (U, h). In that case $\partial f = \nabla h$, i.e. $x^* \in \partial f(x)$ if and only if $x \in U$ and $x^* = \nabla h(x)$. Furthermore, the conjugate function f^* is then likewise the closed extension of the Legendre transform (U^*, h^*).*

Finally, we have a characterization of the case where the Legendre transform is everywhere-defined.

THEOREM 3. *Let h be a (real-valued) differentiable convex function defined on all of R^n. Then ∇h is a continuous one-to-one mapping of R^n onto itself, if and only if h is strictly convex and*

$$\lim_{\lambda\to\infty} h(\lambda x)/\lambda = +\infty \qquad \textit{for every } x \neq 0. \tag{2.8}$$

The conjugate of h is then the same as its Legendre transform.

3. Proofs. If f is a l.s.c. proper convex function on R^n whose subdifferential ∂f is one-to-one, the same is true of the conjugate function f^* by (2.5). The conjugate of f^* is f. Thus Theorem 1 is a corollary of Theorem 2.

We shall now prove Theorem 2. Let f be any l.s.c. proper convex function on R^n. Obviously $\partial f(x)$ is empty when $f(x)$ is not finite, i.e., when x does not belong to the *effective domain* of f, which is the convex set

$$\operatorname{dom} f = \{x \mid f(x) < +\infty\}.$$

On the other hand, if $x \in \operatorname{dom} f$ the directional derivative

$$f'(x; z) = \lim_{\lambda\downarrow 0}[f(x + \lambda z) - f(x)]/\lambda \tag{3.1}$$

exists for every $z \in R^n$. The following facts about the directional derivative function were proved essentially by Fenchel in (**7**, pp. 79–88 and 102–104). If $\partial f(x)$ is empty, $f'(x; z) = -\infty$ for every z in the relative interior of the convex cone generated by $(\operatorname{dom} f) - x$. If $\partial f(x)$ is not empty, $f'(x; \cdot)$ is a proper convex function on R^n such that

$$\liminf_{y\to z} f'(x; y) = \sup\{\langle z, x^*\rangle \mid x^* \in \partial f(x)\} \tag{3.2}$$

for every $z \in R^n$. Therefore $\partial f(x)$ contains exactly one vector if and only if the left side of (3.2) is a linear function of z. Trivially $f'(x; z) = +\infty$ when z does not belong to the convex cone generated by $(\operatorname{dom} f) - x$. Hence the left side of (3.2) is $+\infty$ when z lies outside the closure of this cone. Of course, the only convex cone dense in R^n is R^n itself. Thus the left side of (3.2) cannot be linear in z unless the convex cone generated by $(\operatorname{dom} f) - x$ is all of R^n. Since $\operatorname{dom} f$ is a convex set, this can happen only if x is an interior point of $\operatorname{dom} f$, as can easily be shown using standard separation theorems. Now, when x is an interior point of $\operatorname{dom} f$, it is a classical fact (see **2**) that $f'(x; z))$ is finite for every z, and that $f'(x; \cdot)$ is linear if and only if f is differentiable at x. In that event

$$(3.3) \qquad f'(x; z) = \langle z, \nabla f(x) \rangle \qquad \text{for all } z.$$

It is well known that a convex function finite on all of R^n is automatically continuous. The "lim inf" is therefore unnecessary in (3.2) when x is an interior point of $\operatorname{dom} f$. We are thus led to the following conclusion:

$\partial f(x)$ contains exactly one vector x^*, if and only if x is an interior point of $\operatorname{dom} f$ *and f is differentiable at x; then $x^* = \nabla f(x)$.*

Next we show that, when x_1 and x_2 are interior points of $\operatorname{dom} f$,

$$\partial f(x_1) \cap \partial f(x_2) \neq \emptyset$$

if and only if

$$(3.4) \qquad f(\lambda x_1 + (1 - \lambda)x_2) = \lambda f(x_1) + (1 - \lambda) f(x_2)$$

holds for some λ with $0 < \lambda < 1$. Suppose first that x^* is common to both $\partial f(x_1)$ and $\partial f(x_2)$, so that

$$(3.5) \qquad f(y) \geqslant f(x_1) + \langle y - x_1, x^* \rangle \quad \text{and} \quad f(y) \geqslant f(x_2) + \langle y - x_2, x^* \rangle$$

for every $y \in R^n$. This implies that

$$(3.6) \qquad f(x_1) - f(x_2) = \langle x_1 - x_2, x^* \rangle.$$

If $0 < \lambda < 1$, we have by (3.5) and (3.6)

$$f(\lambda x_1 + (1 - \lambda)x_2) \geqslant f(x_2) + \lambda \langle x_1 - x_2, x^* \rangle = \lambda f(x_1) + (1 - \lambda) f(x_2).$$

The opposite inequality is true because f is convex, so (3.4) holds. Conversely, suppose (3.4) holds for some λ with $0 < \lambda < 1$. Then $x = \lambda x_1 + (1 - \lambda)x_2$ is another interior point of $\operatorname{dom} f$. Hence $f'(x; z)$ is finite for all z, so that $\partial f(x) \neq \emptyset$, as in the preceding paragraph. Let x^* be any element of $\partial f(x)$. For every $y \in R^n$,

$$\begin{aligned}(3.7) \quad f(y) &\geqslant f(\lambda x_1 + (1 - \lambda)x_2) + \langle y - \lambda x_1 - (1 - \lambda)x_2, x^* \rangle \\ &= \lambda[f(x_1) + \langle y - x_1, x^* \rangle] + (1 - \lambda)[f(x_2) + \langle y - x_2, x^* \rangle].\end{aligned}$$

Taking first $y = x_1$ and then $y = x_2$ in (3.7), we see that (3.6) holds. When this fact is substituted into (3.7), we get (3.5). Thus x^* belongs to $\partial f(x_1)$ and $\partial f(x_2)$.

Of course, (3.4) holds for some λ with $0 < \lambda < 1$ and two different interior points x_1 and x_2 of $\operatorname{dom} f$, if and only if f fails to be strictly convex on the interior of $\operatorname{dom} f$.

Our arguments have shown so far that ∂f is one-to-one if and only if f is differentiable and strictly convex on the interior of dom f, with $\partial f(x) = \emptyset$ for non-interior points of dom f. We have also seen that then ∂f coincides with the ordinary gradient mapping of f defined on the interior of dom f. It is known **(10)** that ∂f determines f up to an additive constant. Hence $\partial f(x)$ certainly cannot be empty for every x. Consequently, if ∂f is one-to-one, the interior U of dom f is a *non-empty* open convex set, and the restriction h of f to U is differentiable and strictly convex. Moreover, f is then the closed extension of (U, h), because

$$f(x) = \lim_{\lambda \downarrow 0} f(\lambda a + (1 - \lambda)x)$$

for all x when a is any point of dom f. (This follows from the fact that

$$g(0) = \lim_{\lambda \downarrow 0} g(\lambda) \tag{3.8}$$

when g is a l.s.c. proper convex function on R such that $g(\lambda) < +\infty$ for some $\lambda > 0$; cf. **(7**, p. 78**)**.)

Suppose now that f is the closed extension of (U, h), where U is a non-empty open convex set and h is a differentiable strictly convex function on U. Let $a \in U$ and let x be a boundary point of U. We shall prove that $\partial f(x)$ is empty if and only if (2.2) holds. This will establish all but the last assertion of Theorem 2. Since a is an interior point of dom f, the vector $a - x$ belongs to the interior of the convex cone generated by $(\operatorname{dom} f) - x$. Thus, according to the facts about directional derivatives reviewed at the beginning of the proof of Theorem 2, $\partial f(x)$ is empty if and only if either $f(x) = +\infty$, or $f(x) < +\infty$ but $f'(x; a - x) = -\infty$. Setting

$$g(\lambda) = f(\lambda a + (1 - \lambda)x),$$

we can pass to a one-dimensional context. Here g is finite, differentiable, and strictly convex on an open interval including $\{\lambda \mid 0 < \lambda \leqslant 1\}$, and (3.8) holds. The derivative function g' is then increasing and continuous on the same open interval. The problem is to show that

$$\lim_{\lambda \downarrow 0} g'(\lambda) = -\infty$$

if and only if either $g(0) = +\infty$, or $g(0)$ is finite but

$$\lim_{\lambda \downarrow 0} [g(\lambda) - g(0)]/\lambda = -\infty .$$

This is an elementary exercise in the calculus.

The last assertion of Theorem 2 will now be verified. Assume that f is the closed extension of a convex function of Legendre type (U, h). By (2.6) and the part of the theorem already proved, f^* must also be the closed extension of some convex function of Legendre type (V, k) with $\nabla k = (\nabla h)^{-1}$. The latter implies $V = U^*$. For each $x^* \in U^*$, we have

$$k(x^*) = f^*(x^*) = \langle x, x^* \rangle - h(x) \qquad \text{if } x^* = \nabla h(x),$$

by (2.5). Thus k is the function h^* defined in (2.1). This completes the proof of Theorem 2.

Theorem 3 is immediate from Theorem 2 and (**9**, 5B and 5C), which says that (2.8) holds if and only if the conjugate of $f = h$ is everywhere finite. (Continuity is automatic, as pointed out in §2.)

4. Counterexample. Let U be the open upper half-plane in R^2, and let

$$h(\xi_1, \xi_2) = [(\xi_1^2/\xi_2) + \xi_1^2 + \xi_2^2]/4$$

on U. This differentiable function has a positive definite Hessian matrix throughout U. Hence it is strictly convex; see (**1**). But (2.2) fails for $a = (0, 1)$ and $x = (0, 0)$. The closed extension f of h coincides with h on U, is 0 at the origin, and is $+\infty$ elsewhere. The conjugate f^* is finite everywhere; it actually gives the square of the distance of each point in R^2 from the parabolic convex set

$$C = \{(\xi_1^*, \xi_2^*) \mid \xi_2^* \leqslant -(\xi_1^*)^2\}.$$

But the Legendre transform h^* is not everywhere defined. In fact U^* is the complement of C, and therefore is not even convex. Incidentally, f^* is continuously differentiable on all of R^n, too. The range of its gradient mapping is the union of U with the origin, which is convex but not open.

Note that the non-convexity of U^* did not result here because the given function was arbitrarily restricted to a smaller domain so that some of its gradient mapping was lost. Since h approaches $+\infty$ as one nears the ξ_1-axis, except at the origin, it is clear that h is not the restriction of any finite convex function defined on an open set larger than U. This example also shows that condition (2.2) is generally not "constructive." A pair (U, h) satisfying all the requirements for being a convex function of Legendre type except (2.2) cannot always be extended to a convex function of Legendre type.

REFERENCES

1. B. Bernstein and R. A. Toupin, *Some properties of the Hessian matrix of a strictly convex function*, J. Reine Angew. Math., *210* (1962), 65–72.

2. T. Bonnesen and W. Fenchel, *Konvexe Körper* (Berlin, 1934).

3. R. Courant and D. Hilbert, *Methods of mathematical physics*, Vol. I (New York, 1953).

4. ——— *Methods of mathematical physics*, Vol. II (New York, 1962).

5. J. B. Dennis, *Mathematical Programming and Electrical Networks* (New York, 1959).

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

7. ——— *Convex cones, sets and functions*, lecture notes (Princeton, 1953).

8. G. W. Mackey, *The mathematical foundations of quantum mechanics* (New York, 1963).

9. R. T. Rockafellar, *Level sets and continuity of conjugate convex functions*, Trans. Amer. Math. Soc. *123* (1966), 46–63.

10. ——— *Characterization of the subdifferentials of convex functions*, Pacific J. Math. *17* (1966), 497–510.

Princeton University

PRINTED IN CANADA