Reprinted from The International Journal Of Applied Analysis

Volume 1 Number 1 1977

HIGHER DERIVATIVES OF CONJUGATE CONVEX FUNCTIONS

R. T. Rockafellar*

Department of Mathematics

University of Washington

Seattle, WA 98195, USA

Properties of convex functions are important in many models in economics and operations research, and sometimes it is desirable to know whether the conjugate f^* of a given function f is once, twice or three-times differentiable, even if there is no direct formula for f^* making this clear. For instance, a method in econometrics or the convergence of some optimization algorithm may depend on the degree of differentiability of f^*. The question arises as to just what this entails for the function f, which is more accessible.

Essential first-order differentiability of f^* is equivalent to essential strict convexity of f, as shown in [I, art (26)] The word "essential" here covers certain details about the behavior of f and f^* at the boundaries of their effective domains. The purpose of the present brief note is to point out how this result can be combined with the classical inverse function theorem to get a corresponding equivalence of higher order.

Let f denote an extended-real-valued convex function on R^n which is proper and closed, in the sense that its effective domain $\text{dom}\, f = \{ x \mid f(x) < \infty \}$ is nonempty, f is finite on dom f, and all the level sets of the form $\{ x \mid f(x) \leq \alpha \}$ are closed. The conjugate of f is the function f^* on R^n defined by

$$(1) \quad f^*(x^*) = \sup_{x \in R^n} \{ x.x^* - f(x) \}, \qquad x^* \in R^n.$$

It is likewise convex, proper and closed, and $(f^*)^* = f$.

*This work was supported in part by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under AFOSR grant number 72-2269.

We shall say f is regular of order m (where $m \geq 2$) if the following four conditions are satisfied :

(a) the (open convex) set $C = \text{int}\,(\text{dom } f)$ is nonempty ;

(b) f is m times continuously differentiable on C ;

(c) for every $x \in C$, the second derivative matrix $\nabla^2 f(x)$ is positive definite ;

(d for every sequence $\{x^k\}$ in C that converges to some boundary point x of C. one has $\| \nabla f(x^k) \| \to \infty$.

Remarks. Condition (d) is satisfied vacuously if f is finite everywhere ($C = R^n$). Note that (b), (c) and (d) only depend on the restriction of f to C. As a matter of fact, any function f defined on an open convex set $C \subset R^n$ and having these properties can be extended in a unique manner to be a closed proper convex function on all of R^n with $C = \text{int}\,(\text{dom } f$. To do this, simply select any $\bar{x} \in C$, and for each boundary point x of C define

$$(2) \qquad f(x) = \lim_{\lambda \downarrow 0} f((1-\lambda)x + \lambda \bar{x})$$

(This limit always exists.) For $x \notin C$, define $f(x) = +\infty$. The proof of this work is given in [1, p. 57]. Incidentally, (d) can be replaced by the condition that

$$(3) \qquad \lim_{\lambda \downarrow 0} \nabla f((1-\lambda)x + \lambda \bar{x}) \cdot (\bar{x} - x) = -\infty$$

for x and $\bar{x}$ as in (2) ; see [1, Lemma 26.2].

THEOREM. f is regular of order m if and only if f^* is regular of order m.

Proof. By the symmetry of the conjugacy correspondence, it suffices to show that conditions (a), (b), (c), (d) for f imply the same for f^*. Certainly (b) and (c) imply :

(e) f is strictly convex and differentiable on C.

According to [1, Theorem 26.5], conditions (a), (e) and (c) for f imply (a), (e) and (c) for f^*. Moreover, in this case the gradient mapping ∇f carries C one-to-one onto $C^* = \text{int}$ $(\text{dom } f^*)$, and its inverse is the mapping f^*. Thus the equations $x^* = \nabla f(x)$ and $x = \nabla f^*(x^*)$ are equivalent.

The inverse function theorem (which of course is a corollary of the implicit function theorem, cf. [2, p. 172]), asserts that if a mapping $F : C \rightarrow R^n$ is r-times continuously differentiable ($r \geq 1$) and if $\bar{x}$ and $\bar{x}^*$ are points such that $\nabla F(\bar{x}) = \bar{x}^*$ and the derivative matrix (Jacobian) $\nabla F(\bar{x})$ is nonsingular, then for certain neighborhoods X of $\bar{x}$ and X^* of $\bar{x}^*$ there is a unique mapping $G : X^* \rightarrow X$ such that

$$x = G(x^*) <-> x^* = F(x), x \varepsilon X.$$

Moreove , G is likewise r-times continuously differentiable, and $G(\bar{x}^*) = [\nabla F(\bar{x})]^{-1}$.

This theorem can be applied to the mapping $F = \nabla f$ with $r = m-1$. As seen above F is invertible with $F^{-1} = \nabla f^*$. Moreover $\nabla F(x)$ is the matrix $\nabla^2 f(x)$, which is nonsingular for every $x \varepsilon C$ by assumption. For any choice of $\bar{x}^*$ in C^*, the point $\bar{x} = \nabla f^*(\bar{x}^*)$ is such that $\bar{x}^* = F(\bar{x})$. The local inverse G of F described by the theorem must be the restriction of ∇f^* to a neighborhood of $\bar{x}^*$. Therefore ∇f^* is m-1 times continuously differentiable in some neighborhood of $\bar{x}^*$, and

$$\nabla^2 f^*(\bar{x}^*) = \nabla G(\bar{x}^*) = [\nabla F(\bar{x})]^{-1} = [\nabla^2 f(\bar{x})]^{-1}.$$

In particular, $\nabla^2 f^*(\bar{x}^*)$ is nonsingular, and f^* is m times continuously differentiable in some neighborhood of $\bar{x}^*$. Since this is true for arbitrary $\bar{x}^* \varepsilon C$, we may conclude that f^* satisfies (b) and (c).

Acknowledgement. I wish to thank Professor L. J. Lau of Stanford University for raising this question in connection with a problem in economic analysis.

REFERENCES

1. R. T. Rockafellar, Convex Analysis, Princeton University Press, 1970.
2. M. R. Hestenes, Optimization Theory : the Finite-Dimensional Case, Wiley & Sons, 1975.