

ON THE MONOTONICITY OF THE GRADIENT OF A CONVEX FUNCTION

GEORGE J. MINTY

The object of this note is to present some elementary theorems concerning convex functions in n -dimensions and, more generally, topological vector spaces. These theorems are all essentially generalizations of the theorem "the derivative of a convex function of one real variable is monotonic non-decreasing", and appear to have been overlooked in the literature.

Let X be a topological vector space with real scalars, and Y the conjugate-space (space of continuous linear functionals) of X . We shall write $y(x)$, for $x \in X, y \in Y$, as $\langle x, y \rangle$ to facilitate applications to Hilbert space. The convex (real-valued) function ϕ will always be presumed to have convex domain $D \subset X$, and satisfies the inequality

$$\phi(sx_1 + tx_2) \leq s\phi(x_1) + t\phi(x_2)$$

for all x_1, x_2 in D , all $s \geq 0, t \geq 0, s + t = 1$. The graph G of ϕ is a subset of the topological vector space $X + R$, and it is obvious that the "set of points lying above the graph of ϕ ": $A = \{(x, r) : x \in D, r \geq \phi(x)\}$ is a convex set. (This condition is also sufficient for the convexity of ϕ .)

DEFINITION 1. A set $E \subset X \times Y$ is called a monotonic set provided that, for all (x_1, y_1) and (x_2, y_2) in E , $\langle x_1 - x_2, y_1 - y_2 \rangle \geq 0$.

DEFINITION 2. ([6]) For $D \subset X$, a function $F: D \rightarrow Y$ is called monotonic provided the graph of F is a monotonic set. Now, it is well known that the conjugate space of $X + R$ is $Y + R$, and that a closed hyperplane in $X + R$ is of the form $\{(x, r) : \langle x, y_0 \rangle + rr_0 = \alpha\}$ for some $y_0 \in Y, r_0 \in R, \alpha \in R$. (See [2], p. 26, Théorème 1.) This representation is non-unique, but if $r_0 \neq 0$, the equation $\langle x, y_0 \rangle + rr_0 = \alpha$ can be solved for r , and the resulting equation is, in an obvious sense, unique. These facts motivate the following definition:

DEFINITION 3. A gradient hyperplane H of ϕ is a closed hyperplane of support to A , the set of points lying above the graph of ϕ in $X + R$, such that H can be written in the form $\{(x, r) : r = \phi(x_0) + \langle x - x_0, y_0 \rangle\}$. (Note the analogy with the first two terms of a Taylor-series for ϕ .)

REMARK 1. This definition might be considered inappropriate if ϕ is not everywhere-defined over X ; this problem will not concern us here.

Received February 11, 1963.

DEFINITION 4. The generalized gradient of ϕ is the multiple-valued function assigning to each $x_0 \in D$, all $y_0 \in Y$ such that $r = \phi(x_0) + \langle x - x_0, y_0 \rangle$ is the equation of a gradient hyperplane of ϕ . If this function is single-valued, it will be called simply the gradient of ϕ .

REMARK 2. It is trivial to show that if X is a Banach-space, the notion of "gradient" is equivalent with that of "Fréchet differential", provided the latter exists and D is an open set (so that the Fréchet differential is unique). The advantage of the present treatment is that X need not be a Banach space, and ϕ need not be Fréchet-differentiable, in order that ϕ possess a gradient or generalized gradient. For example: $\phi(x) = \|x\|$ is a convex but not differentiable function in many of the usual Banach spaces.

THEOREM 1. *The graph E of the generalized gradient of ϕ is a monotonic set.*

Proof. Let (x_0, y_0) and (x_1, y_1) be elements of E . Then, because every point of the graph of G lies above both gradient hyperplanes, we have

$$\begin{aligned}\phi(x_1) &\geq \phi(x_0) + \langle x_1 - x_0, y_0 \rangle \\ \phi(x_2) &\geq \phi(x_1) + \langle x_0 - x_1, y_1 \rangle.\end{aligned}$$

Adding, cancelling, and rearranging, we obtain the desired relation.

COROLLARY. *If X is a Hilbert space, so that $Y = X$, and if ϕ is everywhere-defined in X and has a continuous gradient, then the graph of the gradient is a maximal monotonic set.*

Proof. Theorem 4 of [3] asserts that a continuous monotonic function from a Hilbert space to itself which has maximal, open domain has a graph which is a maximal monotonic set.

Before pursuing the interesting consequences of this corollary, we first generalize the corollary.

THEOREM 2. (Again X is a topological vector space.) *If the convex function ϕ is everywhere defined in X , and the set A of points above the graph of ϕ has nonempty interior, then the graph of the generalized gradient of ϕ is a maximal monotonic set.*

Proof. Suppose we have (x_0, y_0) such that $r = \phi(x_0) + \langle x - x_0, y_0 \rangle$ is not the equation of a gradient hyperplane of ϕ . Then there exists a point $x_1 \in X$ such that $\phi(x_1) < \phi(x_0) + \langle x_1 - x_0, y_0 \rangle$. Consider the

following real-valued function of the real variable t :

$$f(t) = \phi(x_0 + t(x_1 - x_0)) - \phi(x_0) - t\langle x_1 - x_0, y_0 \rangle .$$

Clearly f is convex, and also $f(0) = 0, f(1) < 0$. As is well known ([1], p. 46) f has left- and right-hand derivatives for every t , and it is not hard to show that $f^r(t)$ is a monotonic non-decreasing function and

$$f(t) = \int_0^t f^r(t) dt .$$

Thus there exists t_1 , with $0 < t_1 < 1$, such that $f^r(t_1) < 0$. From the fact that A has nonempty interior, we see that A is a convex *body*, and hence, ([2], p. 72, Prop. 3) there exists a hyperplane of support to A at every boundary-point of A , and in particular at $(x_2, \phi(x_2))$, where $x_2 = t_1 x_1$. It is easy to show that it is a gradient hyperplane characterized by x_2 and a suitable y_2 , so that for any x ,

$$\phi(x) \geq \phi(x_2) + \langle x - x_2, y_2 \rangle .$$

In particular, putting $x = x_0 + t(x_1 - x_0)$, we have

$$\phi(x_0 + t(x_1 - x_0)) \geq \phi(x_2) + \langle x_0 + t(x_1 - x_0) - x_2, y_2 \rangle$$

(equality holds here if $t = t_1$) or, rearranging,

$$f(t) \geq \langle x_0 - x_2, y_2 \rangle + t \langle x_1 - x_0, y_2 - y_0 \rangle$$

for all t , with equality holding for $t = t_1$. Thus the right-hand side is a linear function of t , and $r = \langle x_0 - x_2, y_2 \rangle + t \langle x_1 - x_0, y_2 - y_0 \rangle$ is the equation of a "supporting line" to the graph of f at t_1 . Hence

$$\frac{dr}{dt} = \langle x_1 - x_0, y_2 - y_0 \rangle \leq f^r(t_1) < 0$$

from which $\langle t_1(x_2 - x_0), y_2 - y_0 \rangle < 0$. Dividing by t_1 , we see that (x_0, y_0) is not monotonically related to (x_2, y_2) , which is an element of the graph of the generalized gradient of ϕ .

To make this theorem useful, we need a sufficient condition for A to have nonempty interior. We state such a condition in the form of a theorem:

THEOREM 3. *For the everywhere-defined convex function $\phi: X \rightarrow Y$ to have the property that the set A of points above the graph has nonempty interior it is sufficient that*

- (i) ϕ is continuous, or
- (ii) ϕ is bounded in some (nonempty) open set in X

Proof. (i) is equivalent to (ii) by [2], p. 92, Prop. 2. Now, if ϕ

is bounded above by the real constant M over the nonempty open set $B \subset X$, then $B \times (M, \infty)$ is a nonempty open set in A .

REMARK 3. In applications of Theorems 2 and 3 to Banach spaces, it is preferable to use the strong topology, since the hypotheses of Theorem 3 are easier to satisfy than with the weak topology.

We now cite some interesting consequences of these theorems:

COROLLARY 1. *If X is a Hilbert space, and ϕ is an everywhere-defined, continuous convex function, then for any $u \in X$, there exists a unique point of the graph of the generalized gradient such that $x_0 + y_0 = u$. In particular: if ϕ is in addition Fréchet differentiable, then the equation $x + \phi'(x) = u$ is always solvable for x , and the solution depends continuously on u .*

Proof. By Theorem 3 of [3], which asserts that the map $(x, y) \rightarrow x + y$ carries a maximal monotonic set onto the Hilbert-space, and is a homeomorphism.

REMARK 4. The latter part of this theorem was known to E. Rothe, and exists (in modified form) in his paper [5].

COROLLARY 2. *If X is finite-dimensional, and ϕ is everywhere-defined, convex, and continuous, then the range of the generalized gradient of ϕ is an almost-convex set (contains the interior of its convex hull).*

Proof. Follows from the theorem of [4].

Added in proof. Further recent information on monotonic functions can be found in [7] and [8]. Also, J.-J. Moreau has pointed out to the writer that the notion of "gradient" developed here is extremely closely related to the "application prox" whose theory has been expounded by Prof. Moreau in a series of papers, *e.g.* [9].

REFERENCES

1. N. Bourbaki, *Fonctions d'une variable réelle* (Théorie élémentaire). Paris: Hermann et Cie, 1949.
2. ———, *Espaces vectorielles topologiques*, Chs. I-II. Paris: Hermann et Cie, 1953.
3. G. J. Minty, *Monotone (nonlinear) operators in Hilbert space*, Duke Math. J., **29** (1962), 341-346.
4. ———, *On the maximal domain of a 'monotone' function*, Mich. Math. J., **8** (1961), 135-137.

5. E. H. Rothe, *Gradient mappings and extrema in Banach spaces*, Duke Math. J., **2** (1948), 421-431.
6. E. H. Zarantonello, *Solving functional equations by contractive averaging*, U. S. Army Math. Res. Ctr. Rpt. No. 160, June 1960.
7. C. L. Dolph and G. J. Minty, *On Nonlinear Integral Equations of the Hammerstein Type*, to appear in Proceedings of the Seminar on Nonlinear Integral Equations of the U.S. Army Math. Res. Ctr.
8. G. J. Minty, *On a "Monotonicity" Method for the Solution of Nonlinear Equations in Banach Spaces*, to appear in Proc. Nat. Acad. Sci. U.S.A., December 1963.
9. J.-J. Moreau, *Proprietes des Applications "Prox"*, C. R. Acad. Sci., 256, (1963), 1069-1071.

UNIVERSITY OF MICHIGAN

