CONVEX FUNCTIONS ON CONVEX POLYTOPES

DAVID GALE, VICTOR KLEE[1] AND R. T. ROCKAFELLAR[2]

Abstract. The behavior of convex functions is of interest in connection with a wide variety of optimization problems. It is shown here that this behavior is especially simple, in certain respects, when the domain is a polytope or belongs to certain classes of sets closely related to polytopes; moreover, the polytopes and related classes are actually characterized by this simplicity of behavior.

The following corollary is useful in mathematical economics: If D is a boundedly polyhedral set and ϕ is a convex function on the relative interior of D such that ϕ is bounded on bounded sets, then ϕ can be extended in a unique way to a continuous convex function on D.

Introduction. Throughout the paper, E denotes a finite-dimensional Euclidean space. A subset of E is called a *polytope* provided that it is the convex hull of a finite set of points, and is called *boundedly polyhedral* provided that its intersection with any polytope is a polytope. A function ϕ, whose domain D is a subset of E, is called *convex* provided that ϕ is real-valued, D is convex, and

$$\phi(\lambda x + (1 - \lambda)y) \leq \lambda\phi(x) + (1 - \lambda)\phi(y)$$

for all $\lambda \in [0, 1]$ and $x, y \in D$. It is a familiar fact that ϕ is continuous at interior points of D but need not be so at boundary points where it can "jump upward," as is easily seen in the one-dimensional case. In this case, on the other hand, it is clear that ϕ cannot jump downward and hence is upper semicontinuous. Even this property fails to hold generally in higher dimensions, but it does carry over if D is boundedly polyhedral. In fact, as we shall prove, the property characterizes the boundedly polyhedral sets among the closed convex subsets of E. From this there follows another characteristic property of boundedly polyhedral sets which is useful in mathematical economics and other applications of optimization theory:

If D is boundedly polyhedral and ϕ is a convex function on the relative interior of D which is bounded on bounded sets, then ϕ can be extended in a unique way to a continuous convex function on D. In general, our

results show that the polytopes and closely related sets are exactly those domains whose convex functions have nice boundary behavior.

The following is a consequence of more detailed results established below.

For a bounded closed convex subset D of E, the following five conditions are equivalent:

(P) *D is a polytope.*

(L) *Every convex function ϕ on D is upper Lipschitzian at each point x of D; that is, there exists $L_x < \infty$ such that $\phi(y) - \phi(x) \leq L_x \|y - x\|$ for all $y \in D$.*

(S) *Every convex function on D is upper semicontinuous.*

(M) *Every convex function on D attains a maximum.*

(B) *Every convex function on D is bounded.*

If A represents L, S, M or B, the weakened condition which results from restricting (A) to convex functions which are

{lower semicontinuous, bounded, lower semicontinuous and bounded}

is denoted by $\{(A_l), (A_b), (A_{lb})\}$.

For example, (B_l) asserts that every lower semicontinuous convex function on D is bounded. Conditions (B_b) and (B_{lb}) are tautologies, but it will be shown that *for an arbitrary bounded closed convex set D in E, all the other conditions are equivalent to* (P). For unbounded sets the situation is similar but somewhat more complicated.

Convex functions on convex polytopes. Let us begin by showing that (P)$\Rightarrow$(M) and (P)$\Rightarrow$(L). Suppose that D is a polytope and ϕ is a convex function on D. If μ is the maximum of ϕ on the (finite) set of all extreme points of D, then $\sup \phi D = \mu < \infty$ and thus (P)$\Rightarrow$(M). If x is a point of D then D is the union of finitely many simplices, each having x as a vertex; plainly (L) (or, for that matter, (S) or (M) or (B)) holds for D if it holds for each of these simplices. But then the desired conclusion ((P)$\Rightarrow$(L)) follows easily, because a convex function on a simplex is majorized by an affine function coinciding with it at all vertices of the simplex.

For a more quantitative proof that (P)$\Rightarrow$(L), let δ denote the minimum of the distances from x to the various faces of D missing x, and let $L_x = (\mu - \phi(x))/\delta$. Consider an arbitrary point y of $D \sim \{x\}$, and let F denote the smallest face of D which includes both y and x. The ray from x through y meets the relative boundary of F at a point z, and $\|z - x\| \geq \delta$ because x is not in the smallest face of F which includes z. With $y = \lambda x + (1 - \lambda)z$ for some $\lambda \in [0, 1]$, it follows that

$$\frac{\phi(y) - \phi(x)}{\|y - x\|} = \frac{\phi(\lambda x + (1 - \lambda)z) - \phi(x)}{\|\lambda x + (1 - \lambda)z - x\|} \leqq \frac{\lambda\phi(x) + (1 - \lambda)\phi(z) - \phi(x)}{(1 - \lambda)\|z - x\|}$$

$$= \frac{\phi(z) - \phi(x)}{\|z - x\|} \leqq \frac{\mu - \phi(x)}{\delta} = L_x.$$

Let us now prove the first italicized statement of the introduction. Suppose that D is boundedly polyhedral and the relative interior of D is the domain of a convex function ϕ which is bounded on bounded sets. As Fenchel has observed [1, pp. 74–75], ϕ can be extended in a unique way to a lower semicontinuous convex function $\bar{\phi}$ on D. On each polytope in D, $\bar{\phi}$ is upper Lipschitzian (for (P)$\Rightarrow$(L)) and hence upper semicontinuous. With D boundedly polyhedral, this implies that $\bar{\phi}$ is actually continuous and completes the proof.

For the second italicized statement of the introduction, let us suppose that D is a bounded closed convex subset of E. We have shown that (P)$\Rightarrow$(L) and (P)$\Rightarrow$(M), while it is evident that (L)$\Rightarrow$(S) and (M)$\Rightarrow$(B). To complete the proof, we observe that (S) and (B) fail if D is not a polytope. Indeed, there is an infinite sequence $x_1, x_2 \cdots$ of distinct extreme points of D, and with $\phi(x_i) = i$ while $\phi(y) = 0$ for all $y \in D \sim \{x_1, x_2, \cdots\}$, ϕ is a convex function on D which is neither upper semicontinuous nor bounded. (Note, however, that ϕ is not lower semicontinuous.)

The characterization theorems. We shall see that the polytopes, the boundedly polyhedral sets, and the vector sums of polytopes and closed convex cones can all be characterized in terms of the boundary behavior of their convex functions. It is convenient to employ the following notation for the three conditions.

(P) D is a polytope.

(BP) The intersection of D with any polytope is a polytope.

(P+C) D is the vector sum of a polytope and a closed convex cone.

THEOREM 1. *For a closed convex subset D of E the conditions* (P), (L), (L$_l$), (M), (M$_l$), (B) *and* (B$_l$) *are equivalent.*

THEOREM 2. *For a closed convex subset D of E the conditions* (BP), (L$_b$), (L$_{lb}$), (S), (S$_l$), (S$_b$) *and* (S$_{lb}$) *are equivalent.*

THEOREM 3. *For a closed convex subset D of E the conditions* (P+C), (M$_b$) *and* (M$_{lb}$) *are equivalent.*

The proofs of these theorems are all based on the following lemma, which refines the construction in the last paragraph of the preceding section.

LEMMA. *If D is a closed convex subset of E whose set ext D of extreme points is infinite, then D admits lower semicontinuous convex functions η and ζ, both nonnegative, such that η is unbounded and ζ is bounded but attains no maximum. If some bounded subset of ext D is infinite, then η and ζ can be constructed so as not to be upper semicontinuous.*

PROOF OF THE LEMMA. A point x of D is said to be an *exposed point* provided that $\{x\}$ is the intersection of D with a supporting hyperplane or, equivalently, provided that E admits a linear form f_x whose maximum on D is attained precisely at x. For any such f_x and for an arbitrary positive ϵ, let

$$M_D(f_x, \epsilon) = \{y \in D : f_x(y) \geqq f_x(x) - \epsilon\}.$$

A standard argument based on the local compactness of D shows that the family of sets $\{M_D(f_x, \epsilon) : \epsilon > 0\}$ is a basis for the neighborhoods of x in D. From [3, p. 91] it follows that the set of all exposed points of D is dense in the set of all extreme points, and hence under the present hypothesis D has infinitely many exposed points. Let $x_1, x_2, \cdots$ be a sequence of distinct exposed points of D such that no x_i is a cluster point of the sequence, and for each i let f_i be a linear form whose maximum on D is attained precisely at x_i. For each i let

$$\delta_i = \min\{\|x_j - x_i\| : j \neq i\} > 0$$

and choose $\epsilon_i > 0$ so that

$$M_D(f_i, \epsilon_i) \subset \{y \in D : \|y - x_i\| < \delta_i/3\}.$$

The sets $M_D(f_i, \epsilon_i)$ and $M_D(f_j, \epsilon_j)$ are disjoint except when $i = j$. Let

$$\xi_i(y) = 0 \qquad\qquad \text{for } y \in D \sim M_D(f_i, \epsilon_i),$$
$$\xi_i(y) = (f_i(y) - f_i(x_i) + \epsilon_i)/\epsilon_i \qquad \text{for } y \in M_D(f_i, \epsilon_i),$$

so that ξ_i is a continuous convex function on D, ξ_i is nonnegative, and the maximum of ξ_i is 1, attained only at x_i. Finally, let

$$\eta = \sum_{i=1}^{\infty} i\xi_i \quad \text{and} \quad \zeta = \sum_{i=1}^{\infty} \frac{i}{i+1} \xi_i.$$

The functions η and ζ have the desired properties. When some bounded subset of ext D is infinite, the sequence x_α of exposed points may be chosen to converge to a point x_0 of D; then η and ζ are not upper semicontinuous, for $\eta(x_0) = \zeta(x_0) = 0$ while $\lim \eta(x_\alpha) = \infty$ and $\lim \zeta(x_\alpha) = 1$.

PROOF of THEOREM 1. Having proved already that $(P) \Rightarrow (M)$ and $(P) \Rightarrow (L)$, we remark that the following implications are obvious:

$$(M) \quad \begin{matrix} \Rightarrow (M_l) \Rightarrow \\ \Rightarrow (B) \Rightarrow \end{matrix} (B_l). \qquad (L) \Rightarrow (L_l).$$

To complete the proof of Theorem 1 it suffices to show that $(L_l) \Rightarrow (P)$ and $(B_l) \Rightarrow (P)$. If D is unbounded then D contains a ray $x + [0, \infty [u$ for some $x \in D$ and some $u \in E \sim \{0\}$. For each $y \in D$ let $\phi(y) = \langle u, y \rangle^2$, the square of the inner product of u and y. Then ϕ is continuous and convex, but is neither bounded on D nor upper Lipschitzian at x. Hence the boundedness of D is implied by (B_l) and also by (L_l). If D is a bounded closed convex set which is not a polytope, the Lemma implies the existence on D of an unbounded lower semicontinuous convex function, thus contradicting (B_l), and of a lower semicontinuous convex function which is not upper semicontinuous and hence not upper Lipschitzian, thus contradicting (L_l). The proof of Theorem 1 is complete.

PROOF OF THEOREM 2. We show first that $(BP) \Rightarrow (L_b)$ and $(BP) \Rightarrow (S)$. Consider a boundedly polyhedral set D, a point x of D, and a convex function ϕ on D. Let D_0 be a polytope such that $y \in D_0$ whenever $y \in D$ and $\|y - x\| \leq 1$. Let ϕ_0 denote the restriction of ϕ to D_0, whence (by the implication $(P) \Rightarrow (L)$ already established) there exists $L_x < \infty$ such that $\phi(y) - \phi(x) \leq L_x \|y - x\|$ for all $y \in D_0$. In particular, ϕ is upper semicontinuous at x and we conclude $(BP) \Rightarrow (S)$. If $\sup \phi D = \bar{\mu} < \infty$, then of course

$$\phi(y) - \phi(x) \leq \max(L_x, \bar{\mu} - \phi(x)) \|y - x\|$$

for all $y \in D$, whence $(BP) \Rightarrow (L_b)$.

The following implications are obvious:

$$(L_b) \Rightarrow (L_{lb}) \Rightarrow (S_{lb}). \qquad (S) \quad \begin{matrix} \Rightarrow (S_l) \Rightarrow \\ \Rightarrow (S_b) \Rightarrow \end{matrix} (S_{lb}).$$

To complete the proof of Theorem 2 it suffices to show that if the closed convex subset D of E is not boundedly polyhedral, then D admits a bounded convex function which is lower semicontinuous but not upper semicontinuous. Let ∂D denote the relative boundary of D, and for each point x of ∂D, let cone (x, D) denote the union of all rays which issue from x and pass through the various points of $D \sim \{x\}$. From a known characterization of boundedly polyhedral sets [4, p. 95] there follows the existence of a point x_0 of ∂D such that cone (x_0, D) is not closed, and then by [4, p. 88] there is a 2-dimensional flat Q through x_0 such that the set $Q \cap D$ is not polyhedral at x_0. With Q 2-dimensional, this implies that x_0 is an accumulation point of the set ext $(Q \cap D)$, whence the Lemma guarantees the existence of a convex

function ϕ on $Q \cap D$ such that $\phi(Q \cap D) \times [0, 1]$ and ϕ is lower semi-continuous but not upper semicontinuous. Let $\bar\phi$ be the supremum of all affine forms α on D such that $\alpha(q) \leq \phi(q)$ for all $q \in Q \cap D$ and $\alpha(x) \leq 2$ for all $x \in D$. Plainly $\bar\phi D \subset [0, 2]$ and $\bar\phi$, as a real-valued supremum of affine forms, is convex and lower semicontinuous. It remains only to show that $\bar\phi$ is not upper semicontinuous, for which it suffices to show that $\bar\phi(q_0) = \phi(q_0)$ for all $q_0 \in Q \cap D$.

Consider an arbitrary $\epsilon > 0$. In the 3-flat $Q \times R$, the point $(q_0, \phi(q_0) - \epsilon)$ is at positive distance from the closed convex set $\{(q, r) : q \in Q \cap D, r \geq \phi(q)\}$, and hence (by a well-known separation theorem) the point is strictly separated from the set by a 2-flat F in $Q \times R$. Since $\phi(Q \cap D) \subset [0, 1]$ it is easily seen that F is at positive distance from the closed convex set

$$\{(d, 2) : d \in D\} \subset E \times R,$$

whence F is contained in a hyperplane H at positive distance from the set. Following a standard procedure, we define an affine form α on E by the condition that $(y, \alpha(y)) \in H$ for all $y \in E$. It is then verified that

$$\alpha(q_0) > \phi(q_0) - \epsilon, \qquad \alpha \leq \phi \text{ on } Q \cap D, \text{ and } \alpha \leq 2 \text{ on } D.$$

This shows that $\bar\phi(q_0) = \phi(q_0)$ and completes the proof.

An alternative construction of $\bar\phi$ is possible, avoiding the Lemma. First observe that with the aid of the Decomposition Theorem stated below, attention may be restricted to the case in which D contains no line. With x_0 as above, let V denote the union of $\{x_0\}$ with all rays in ∂D which issue from x_0 and let W denote the set of all points w of ∂D such that x_0 lies in every hyperplane supporting D at w. Then $x_0 \in V \subset W \subset \partial D$. Let $\bar\phi$ be the supremum of all affine forms α on D such that $\bar\phi(v) \leq 0$ for all $v \in V$ and $\bar\phi(y) \leq 1$ for all $y \in \partial D \sim W$. Then $\bar\phi$ is a bounded convex function on D which is lower semicontinuous but not upper semicontinuous. The proof of this fact is left to the reader. (For the proof it is helpful to know that D is the closed convex hull of its exposed points together with its exposed rays [3, p. 91].)

In the proof of Theorem 3, the following result justifies a reduction to the case of closed convex sets containing no line.

DECOMPOSITION THEOREM. *Suppose that D is a closed convex subset of E, with $0 \in D$. Let the union of all lines through 0 in D be denoted by L_D (a linear subspace of E), the orthogonal supplement of L_D by $L_D^\perp$, the intersection $D \cap L_D^\perp$ by K_D, the union of all rays from 0 in K_D by C_D (a pointed closed convex cone), and the set of all extreme points of K_D by E_D. Then*

$$D = L_D + K_D \quad and \quad K_D = C_D + \mathrm{con}\ E_D.$$

Each point y of D admits a unique expression in the form $y = y' + y''$, with $y' \in L_D$ and $y'' \in K_D$. Any convex function ψ on K_D can be extended to a convex function ϕ on D by setting $\phi(y) = \psi(y'')$ for all $y \in D$. Every bounded convex function on D arises in this way from a bounded convex function on K_D.

PROOF OF THEOREM 3. $(P+C) \Rightarrow (M_b) \Rightarrow (M_{lb})$. The second implication is obvious. For the first, suppose that $D = P + C$, where P is a polytope and C is a closed convex cone, and where we may assume without loss of generality that $0 \in D$ and 0 is the vertex of C. With notation as in the Decomposition Theorem, let π denote the orthogonal projection of E onto $L_D^\perp$. Then πP is a polytope and $L_D + P = L_D + \pi P$. It can be verified that $L_D + C_D = C$, whence

$$D = C + P = C_D + (L_D + P) = C_D + L_D + \pi P = L_D + (C_D + \pi P)$$

and it follows that $K_D = C_D + \pi P$. But then E_D, the set of all extreme points of K_D, is contained in the (finite) set of all extreme points of πP, so of course E_D is finite. By a theorem of Hirsch and Hoffman [2, p. 364] (for an alternate proof, see [5]) it follows that every convex function on D which is bounded above on D attains its D-supremum at some point of E_D.

$(M_{lb}) \Rightarrow (P+C)$. If D does not satisfy the condition $(P+C)$, then neither does K_D. But K_D has infinitely many extreme points and hence, by the Lemma, admits a lower semicontinuous bounded convex function not attaining a maximum. This can be extended to such a function on D.

REFERENCES

1. W. Fenchel, *On conjugate convex functions*, Canad. J. Math. 1 (1949), 73–77.
2. W. M. Hirsch and A. J. Hoffman, *Extreme varieties, concave functions, and the fixed charge problem*, Comm. Pure Appl. Math. 14 (1961), 355–369.
3. V. Klee, *Extremal structure of convex sets*. II, Math. Z. 69 (1958), 90–104.
4. ———, *Some characterizations of convex polyhedra*, Acta Math. 102 (1959), 79–107.
5. ———, Review of [2], Math. Reviews 24 (1962), 307.

UNIVERSITY OF CALIFORNIA, BERKELEY AND
    UNIVERSITY OF WASHINGTON