Economics Letters 22 (1987) 251–255
North-Holland

CONVEX FUNCTIONS ON NON-CONVEX DOMAINS

H.J.M. PETERS

University of Limburg, 6200 MD Maastricht, The Netherlands

P.P. WAKKER

University of Nijmegen, 6500 HC Nijmegen, The Netherlands

Received 28 May 1986

It is shown that a convex function, defined on an arbitrary, possibly finite, subset of a linear space, can be extended to the whole space. An application to decision making under risk is given.

1. Introduction

The definitions of convexity of a function, customary in literature, assume or imply convexity of the domain. In applications however, observations about the values of a function, and inequalities verified by these, are often available only on non-convex, e.g. finite, domains.

This paper considers a way to define convexity on non-convex domains. The main theorem, in section 2, shows that a function is convex, if and only if it is the restriction of a convex function defined on the whole space. Section 3 gives an application to decision making under risk, and section 4 concludes.

2. The extension theorem

Let V be a linear space over the reals. Let T be some arbitrary subset of V, and f a function from T to $\mathbb{R} \cup \{-\infty, \infty\}$. The following definition adapts the definitions of convexity, given in literature only for convex sets T [see for instance Rockafellar (1970, sect. 4)], to general, possibly finite, sets T. For compact sets it was given in Peters and Tijs (1981), and for general sets in Wakker, Peters and Van Riel (1985).

Definition 1. The function is *convex* if for all convex combinations $\sum_{j=1}^{n} p_j x^j$ of elements x^j of T, for which not both $-\infty$ and $+\infty$ are contained in $\{f(x^j)\}_{j=1}^{n}$, we have

$$\sum_{j=1}^{n} p_j f(x^j) \geq f\left(\sum_{j=1}^{n} p_j x^j\right), \tag{1}$$

whenever $\sum p_j x^j$ *is in* T.

0165-1765/86/$3.50

As usual we take: $\lambda\infty := \infty$ for $\lambda \in \mathbb{R}_{++}$, $\lambda\infty := 0$ for $\lambda = 0$, $\lambda\infty := -\infty$ for $\lambda \in \mathbb{R}_{--}$; $\lambda + \infty := \infty$ for $\lambda \in \mathbb{R}$ or $\lambda = \infty$, and $\lambda + \infty$ is undefined for $\lambda = -\infty$; $\lambda(-\infty) := -\infty$ for $\lambda \in \mathbb{R}_{++}$, $\lambda(-\infty) := 0$ for $\lambda = 0$, $\lambda(-\infty) := \infty$ for $\lambda \in \mathbb{R}_{--}$; $\lambda - \infty := -\infty$ for $\lambda \in \mathbb{R}$ or $\lambda = -\infty$, and $\lambda - \infty$ is left undefined for $\lambda = \infty$.

A function f is *concave* if $-f$ is convex. All results, derived in the sequel for convex functions f, can be reformulated for concave functions g, by setting $g := -f$.

Theorem 1 (extension theorem). Let V be a linear space over $\mathbb{R}$; let $T \subset V$. Let $f: T \to \mathbb{R} \cup \{-\infty, \infty\}$ be convex. Then there exists a convex function $\tilde{f}: V \to \mathbb{R} \cup \{-\infty, \infty\}$ which extends f.

Proof. We define $\tilde{f}$ on V as follows:

$$\tilde{f}: x \mapsto \inf\{\mu : \mu = \textstyle\sum_{j=1}^{n} p_j f(x^j),\ x \text{ is a convex combination } \sum p_j x^j \text{ of elements } x^j \text{ of } T \text{ such that not } f(x^j) = \infty \text{ for some } j, \text{ and } f(x^j) = -\infty \text{ for some other } j\}. \tag{2}$$

Note that $\tilde{f}(x) \leqq f(x)$ for all $x \in T$; by convexity of f, $\tilde{f}(x) \geqq f(x)$ for all $x \in T$. So indeed $\tilde{f}$ extends f. Convexity of $\tilde{f}$ remains to be demonstrated. So let y be a convex combination $\sum_{i=1}^{m} q_i y^i$, with all $q_i > 0$. We have to prove

$$\tilde{f}(y) \leqq \sum_{i=1}^{m} q_i \tilde{f}(y^i) \tag{3}$$

whenever not $\tilde{f}(y^i) = \infty$ for some i, $\tilde{f}(y^i) = -\infty$ for some other i.

The case where $\tilde{f}(y^i) = \infty$ for some i is immediate anyhow, so we suppose $\tilde{f}(y^i) < \infty$ for all i. This implies that every y^i is in conv(T). Say y^i is the convex combination $\sum_k p_{ik} x^{ik}$ of elements x^{ik} of T. Then $y = \sum_{i=1}^{m} \sum_k q_i p_{ik} x^{ik}$. By definition of $\tilde{f}$,

$$\tilde{f}(y) \leqq \sum_{i=1}^{m} q_i \sum_k p_{ik} f(x^{ik}). \tag{4}$$

Now first suppose $\tilde{f}(y^l) = -\infty$, for some l. Then, for every $M \in \mathbb{N}$, we can take the x^{lk} above such that $\sum_k p_{lk} f(x^{lk}) < -M$. Since $q_l > 0$, by (4) we get $\tilde{f}(y) = -\infty$, and (3) follows.

Next suppose $\tilde{f}(y^i) > -\infty$ for all i. Then, for any $\epsilon > 0$ and every i, we can take the x^{ik} such that $\sum_k p_{ik} f(x^{ik}) \leqq \tilde{f}(y^i) + \epsilon$. This again implies (3). Q.E.D.

Note that the $\tilde{f}$ defined in (2) is the maximal convex extension. If f in Theorem 1 is real-valued, then even on conv(T), it may be impossible to have $\tilde{f}$ real-valued, as the following example shows.

Example 1. Let $V = \mathbb{R}^2$, $T = \{(-1, -j)\}_{j \in \mathbb{N}} \cup \{(1, j)\}_{j \in \mathbb{N}}$, $f(x_1, x_2) = -|x_2|$ for all $(x_1, x_2) \in T$. Then considering the pair $(-1, -j)$ and $(1, j)$, one sees that $\tilde{f}(0, 0) \leqq -j$ must hold, for all $j \in \mathbb{N}$.

For bounded f, $\tilde{f}$ can be taken real-valued on conv(T):

Corollary 1. For the $\tilde{f}$, defined in (2), $\sup\tilde{f}(\text{conv}(T)) = \sup(f(T))$, and $\inf\tilde{f}(\text{conv}(T)) = \inf f(T)$.

Proof. Obvious from (2).

If $V = \mathbb{R}$, then a real-valued f in Theorem 1 can be extended to a real-valued $\tilde{f}$ on conv(T).

Corollary 2. *Let* $V=\mathbb{R}$, $T\subset V$, $f: T\to\mathbb{R}$ *convex. Then* f *has a convex extension* $\tilde{f}:\operatorname{conv}(T)\to\mathbb{R}$.

Proof. Obvious if T contains no more than two elements, then $\tilde{f}$ can be taken affine (i.e., both convex and concave). So let $x^1>x^2>x^3$ in T. Let l^1 be the affine function through $(x^1, f(x^1))$ and $(x^2, f(x^2))$, l^2 be the one through $(x^2, f(x^2))$ and $(x^3, f(x^3))$. Then, with $\tilde{f}$ as in (2), for all $x\in\mathbb{R}$ we have $f(x)\geqq\min\{l^1(x), l^2(x)\}>-\infty$.

Further, for all $x\in\operatorname{conv}(T)$, $x\in[x^4, x^5]$ for some x^4, x^5 in T. Hence $\tilde{f}(x)\leqq\max\{f(x^4), f(x^5)\}<\infty$ for such x. Q.E.D.

The following example shows that, even if $V=\mathbb{R}$, and f is real valued and bounded on T, then still no convex real-valued extension $\tilde{f}$ of f to all of V may exist.

Example 2. Let $T=[0, 1]$, $f: x\mapsto -\sqrt{x}$ on T. Then any convex extension $\tilde{f}$ of f on V can be seen to assign ∞ to all of $\mathbb{R}_{--}$.

Lemma 1. *Let* $V=\mathbb{R}$, *and* f *non-decreasing and convex. Then* $\tilde{f}$, *as defined by (2), is non-decreasing on* $\operatorname{conv}(T)$.

Proof. By Corollary 1, $\tilde{f}$ has the same infimum on $\operatorname{conv}(T)$, as f on T. Hence the infimum of $\tilde{f}$ can be found on the 'left side' of $\operatorname{conv}(T)$: the convex $\tilde{f}$ must be non-decreasing. Q.E.D.

3. An application to risk aversion

Let $\mathscr{C}$ be a non-empty set of *consequences*. $\mathscr{L}^s(\mathscr{C})$ is the set of (*simple*) *lotteries* on $\mathscr{C}$; a (simple) lottery on $\mathscr{C}$ is a probability measure on $(\mathscr{C}, 2^{\mathscr{C}})$, assigning probability one to a finite subset of $\mathscr{C}$. By $(p_j;\ x^j)_{j=1}^n$ we denote the simple lottery, assigning probability p_j to every x^j. Implicit in this is that $x^j\in\mathscr{C}$ for all j, $p_j\geqq 0$ for all j, and $\Sigma p_j=1$. For any $\alpha\in\mathscr{C}$, $\bar{\alpha}$ denotes $(1;\ \alpha)$, i.e., the lottery which with probability one results in consequence α.

We assume there are two *decision makers* (persons, players, etc.) T^k, $k=1, 2$, with preference relations $\succcurlyeq^k$, i.e., binary relations on $\mathscr{L}^s(\mathscr{C})$. The interpretation of $l\succcurlyeq^k l'$ is that T^k, when having to choose one element from $\{l, l'\}$, is willing to choose l.

We assume that there exist von Neumann–Morgenstern (vNM) utility functions U^1, U^2, i.e., for $k=1, 2$, $U^k:\mathscr{C}\to\mathbb{R}$ is such that

$$\left(p_j;\ x^j\right)_{j=1}^n\succcurlyeq\left(q_i;\ y^i\right)_{i=1}^m\Leftrightarrow\sum_{j=1}^n p_jU^k(x^j)\geqq\sum_{i=1}^m q_iU^k(y^i). \tag{5}$$

Here $\Sigma p_jU^k(x^j)$ is the *expected utility of* $(p_j;\ x^j)_{j=1}^n$.

The following definition gives a way to compare T^1 and T^2 with respect to their 'risk aversion'. It has been introduced in Yaari (1969, p. 316, in terms of so-called 'acceptance sets'), and is a variation on earlier definitions of Pratt (1964) and Arrow (1971).

Definition 2. T^1 is *more risk averse* (MRA) *than* T^2 if

$$\left(p_j;\ sx^j\right)_{j=1}^n\succcurlyeq^1\bar{\alpha}\Rightarrow\left(p_j;\ x^j\right)_{j=1}^n\succcurlyeq^2\bar{\alpha}. \tag{6}$$

The above definition applies if T^2 is willing to take a risky lottery $(p_j;\ x^j)_{j=1}^n$ instead of a certain consequence α, whenever T^1 is willing to do so.

Definition 2, and the results to be derived below, have first been studied in literature for the case $\mathscr{C}=\mathbb{R}$. Later Kihlstrom and Mirman (1974) extended these to the case $\mathscr{C}=\mathbb{R}^m_+$. All this was done under differentiability and monotonicity assumptions about the vNM utility functions. In Wakker, Peters and Van Riel (1985, Theorem 3.1.b) the results were extended to arbitrary spaces $\mathscr{C}$, without any restriction on the vNM utility functions. Motivation for this was, firstly, that differentiability of the utility functions does not always have a clear behavioural meaning; secondly, that it seems desirable for applications to be able to handle cases where one has information about the choices of decision makers, only with respect to finitely many consequences; and thirdly, that it seems desirable to be able to handle cases where one does not (yet) have a quantification of the consequences.

The following theorem characterizes the 'more risk averse than' relation.

Theorem 2. The following two statements are equivalent:

(i) T^1 is more risk averse than T^2.
(ii) $U^2=\psi\circ U^1$ for a convex non-decreasing function $\psi: U^1(\mathscr{C})\to U^2(\mathscr{C})$.

Proof. First suppose (i). Then $\bar\beta \succcurlyeq^1 \bar\alpha \Rightarrow \bar\beta \succcurlyeq^2 \bar\alpha$, so $U^1(\beta)\geqq U^1(\alpha)\Rightarrow U^2(\beta)\geqq U^2(\alpha)$. So there must exist a non-decreasing $\psi: U^1(\mathscr{C})\to U^2(\mathscr{C})$ such that $U^2=\psi\circ U^1$. Now let some element $U^1(\mu)$ of $U^1(\mathscr{C})$ be a convex combination $\Sigma_{j=1}^n p_j U^1(\mu^j)$ of other elements of $U^1(\mathscr{C})$. Then $(p_j;\ \mu^j)_{j=1}^n \succcurlyeq^1 \bar\mu$, so by (i): $(\beta_j;\ \mu^j)_{j=1}^n \succcurlyeq^2 \bar\mu$, i.e., $\Sigma p_j U^2(\mu^j)\geqq U^2(\mu)$. Substituting ψ gives

$$\Sigma p_j \psi\big(U^1(\mu^j)\big) \geqq \quad \big[\psi\big(U^1(\mu)\big)=\big] \quad \psi\big(\Sigma p_j U^1(\mu^j)\big), \tag{7}$$

so ψ is convex on $U^1(\mathscr{C})$, and (ii) is proved.

Next suppose (ii). Let $(p_j;\ \mu^j)_{j=1}^n \succcurlyeq^1 \bar\mu$, so $\Sigma_{j=1}^n p_j U^1(\mu^j)\geqq U^1(\mu)$. Let $\tilde\psi$ be as defined in (2); by Lemma 1, $\tilde\psi$ is non-decreasing on conv(T). We get $\Sigma_{j=1}^n p_j U^2(\mu^j)=\Sigma p_j\psi(U^1(\mu^j))\geqq\tilde\psi(\Sigma p_j U^1(\mu^j))$ $\geqq \psi(U^1(\mu))=U^2(\mu)$. This implies $(p_j;\ \mu^j)_{j=1}^n \succcurlyeq^2 \bar\mu$, which is what (i) requires. Q.E.D.

Theorem 2 shows a natural occurrence of convex functions on non-convex domains: if observations of choices of T^1 and T^2 are available only with respect to consequences from $\mathscr{C}$, then (inequalities concerning) ψ will only be observed on $U^1(\mathscr{C})$, which may be any arbitrary subset of $\mathbb{R}$. In the last occurrence of $\tilde\psi$ in the proof we essentially use the extension theorem: the argument of $\tilde\psi$ there does not have to be in the domain of ψ. Because of this complication, in Wakker, Peters and Van Riel (1985) a more complicated proof (without the extension theorem available) was given, using first-order-difference results derived in the appendix there.

4. Conclusion

Convexity and concavity have proved to be fruitful concepts in economics. Therefore, it seems to be worthwhile to define convexity and concavity of a function also on non-convex domains, since these may occur in economics in a natural way. The main theorem presented in this paper, which extends a convex function to at least the convex hull of its domain, enables one to apply results known for convex functions on convex domains to such functions on non-convex domains. An example of an application, to decision making under risk, has been given.

References

Arrow, K.J., 1971, Essays in the theory of risk bearing (Markham, Chicago, IL).
Kihlstrom, R.E. and R.J. Mirman, 1974, Risk aversion with many commodities, Journal of Economic Theory 8, 361–388.
Peters, H. and S.H. Tijs, 1981, Risk sensitivity of bargaining solutions, Methods of Operations Research 44, 409–420.
Pratt, J.W., 1964, Risk aversion in the small and in the large, Econometrica 32, 122–136.
Rockafellar, R.T., 1970, Convex analysis (Princeton University Press, Princeton, NJ).
Wakker, P.P., H.J.M. Peters and T. van Riel, 1985, Comparisons of risk aversion, with an application to bargaining, Report (Department of Mathematics, University of Nijmegen, Nijmegen).
Yaari, M.E., 1969, Some remarks on measures of risk aversion and on their uses, Journal of Economic Theory 1, 315–329.