Natalia Lazzati

Mathematics for Economics

Note 10: Quasiconcave and Pseudoconcave Functions

Note 10 is based on Madden (1986, Ch. 13, 14) and Simon and Blume (1994, Ch. 21).

Monotonic transformations: Cardinal Versus Ordinal Utility

A utility function could be said to measure the level of satisfaction associated to each commodity bundle. Nowadays, no economist really believes that a real number can be assigned to each commodity bundle which expresses (in utils?) the consumer's level of satisfaction with this bundle. Economists believe that consumers have well-behaved preferences over bundles and that, given any two bundles, a consumer can indicate a preference of one over the other or the indifference between the two. Although economists work with utility functions, they are concerned with the level sets of such functions, not with the number that the utility function assigns to any given level set. In consumer theory these level sets are called indifference curves. A property of utility functions is called ordinal if it depends only on the shape and location of a consumer's indifference curves. It is alternatively called cardinal if it also depends on the actual amount of utility the utility function assigns to each indifference set.

In the modern approach, we say that two utility functions are equivalent if they have the same indifference sets, although they may assign different numbers to each level set. For instance, let $u : \mathbb{R}^2_+ \to \mathbb{R}$ where $u(\mathbf{x}) = x_1 x_2$ be a utility function, and let $v : \mathbb{R}^2_+ \to \mathbb{R}$ be the utility function $v(\mathbf{x}) = u(\mathbf{x}) + 1$. These two utility functions represent the same preferences and are therefore equivalent. Here v is just a monotonic transformation of u.

Definition 1. (Monotonic transformation) Let I be an interval of the real line. Then $g : I \to \mathbb{R}$ is a monotonic transformation of I if g is a strictly increasing function on I. Furthermore, if g is a monotonic transformation and u is a real-valued function of n variables, then we say that

$$g(u) : \mathbf{x} \to g[u(\mathbf{x})]$$

is a monotonic transformation of u.

If g is differentiable, then it is a monotonic transformation if $g'(x) > 0$ for all $x \in I$. (**Show that the functions $3z + 2$, z^4 and $\ln z$ are monotonic transformations of $\mathbb{R}_{++}$.)**

We next define ordinal properties in a precise way.

Definition 2. (Ordinal property) A characteristic of a function is called ordinal if every monotonic transformation of a function with this characteristic still has this characteristic.

The next example sheds light on this concept.

Example 1. Let $u : \mathbb{R}^2_+ \to \mathbb{R}$ represent a differentiable utility function. In economics the marginal rate of substitution (MRS) at a bundle $\mathbf{x} = \mathbf{x}^*$ is defined as

$$\frac{\partial u}{\partial x_1}(\mathbf{x}^*) / \frac{\partial u}{\partial x_2}(\mathbf{x}^*).$$

Let us consider a monotonic transformation of u. In particular, let us define $v(\mathbf{x}) = g[u(\mathbf{x})]$ with $g'(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathbb{R}^2_+$. At $\mathbf{x} = \mathbf{x}^*$ the MRS for v is

$$g'[u(\mathbf{x}^*)] \frac{\partial u}{\partial x_1}(\mathbf{x}^*) / g'[u(\mathbf{x}^*)] \frac{\partial u}{\partial x_2}(\mathbf{x}^*) = \frac{\partial u}{\partial x_1}(\mathbf{x}^*) / \frac{\partial u}{\partial x_2}(\mathbf{x}^*)$$

and coincides with the MRS for u. Therefore, MRS is an ordinal concept. ▲

Quasiconcave and Pseudoconcave Functions

We showed in previous notes that concavity plays a key role in optimization, as it guarantees that local and global maxima coincide. Unfortunately, concavity is not an ordinal property and it is then a very strong assumption in some economic models, e.g. consumer theory. Fortunately, many of the proofs we provided still hold with weaker inequalities. Quasiconcavity and pseudoconcavity are two of the most well-known generalizations of concavity with applications in mathematical economics and mathematical programming. The definition of quasiconcavity is as follows.

Definition 3. (Quasiconcavity) A function $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconcave if, for all $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ and all $\lambda \in [0, 1]$,

$$f\left[\lambda \mathbf{x} + (1 - \lambda)\mathbf{x}'\right] \geq \min\left\{f(\mathbf{x}), f\left(\mathbf{x}'\right)\right\}.$$

We say that f is strictly quasiconcave if, for all $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ such that $\mathbf{x} \neq \mathbf{x}'$ and all $\lambda \in (0, 1)$,

$$f\left[\lambda \mathbf{x} + (1 - \lambda)\mathbf{x}'\right] > \min\left\{f(\mathbf{x}), f\left(\mathbf{x}'\right)\right\}.$$

The function f is said to be (strictly) quasiconvex if $-f$ is (strictly) quasiconcave.

There is an alternative characterization of quasiconcavity in terms of upper level sets, that sheds light on the reason for which this kind of functions is so important in optimization.

Theorem 4. A function $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconcave if and only if $UC_f(y)$ is a convex set for all y in the range of f.

Proof. (**Prove this result.**) ■

There are many functions that satisfy Definition 3 (and Theorem 4) but that are not concave, e.g. $f : \mathbb{R} \to \mathbb{R}$ where $f(x) = x^3$. In Note 5 we described a concave function as a function that has a convex hypograph and, as a by product of this characteristic, its upper level sets are convex. The graph of a quasiconcave function need not have the former property. It follows that the class of quasiconcave functions includes, but is larger than, the class of concave functions.

The geometry of quasiconcave functions can be visualized with $f : \mathbb{R}^2_{++} \to \mathbb{R}$ where $f(\mathbf{x}) = x_1 x_2$. The range of this function is $\mathbb{R}_{++}$ and for any $y \in \mathbb{R}_{++}$, $UC_f(y) = \left\{\mathbf{x} \in \mathbb{R}^2_{++} : x_1 \geq y/x_2\right\}$ is certainly a convex set (**draw a picture!**). However, this function is not concave; for instance the Hessian is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, which is not negative semi-definite. The visual way to realize about the latter is with the hypograph of f, which is a subset of $\mathbb{R}^3$. To avoid the three dimensional problem, let us consider $x = x_1 = x_2$. The hypograph itself is $\left\{(y, \mathbf{x}) \in \mathbb{R}^3_{++} : y \leq x_1 x_2\right\}$ and the subset of the hypograph where $x = x_1 = x_2$ is $\left\{(y, \mathbf{x}) \in \mathbb{R}^3_{++} : y \leq x^2\right\}$. It is easy to check that $\left\{(y, \mathbf{x}) \in \mathbb{R}^3_{++} : y \leq x^2\right\}$ is not a convex set, which means that HG_f is not convex.

A nice example of a quasiconave function in Statistics is the probability density function of a random variable that is normally distributed.(**Show this statement.**)

As in the concave case, the theory of differentiable quasiconcave functions can be very useful.

Theorem 5. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is C^1. Then f is quasiconcave if and only if, for every $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$,

$$f(\mathbf{x}) \geq f(\mathbf{x}') \text{ implies } Df(\mathbf{x}')(\mathbf{x} - \mathbf{x}') \geq 0.$$

If, moreover,

$$\mathbf{x} \neq \mathbf{x}' \text{ and } f(\mathbf{x}) \geq f(\mathbf{x}') \text{ implies } Df(\mathbf{x}')(\mathbf{x} - \mathbf{x}') > 0$$

then f is strictly quasiconcave, but the converse statement is not necessarily true.

Proof. See Madden (1986, pp. 210-211). ■

There are also some useful second-derivative characterizations of quasiconcavity.

Theorem 6. Let $f : D \to \mathbb{R}$ be a C^2 function defined on an open and convex set $D \subseteq \mathbb{R}^n$, and let us define the following bordered Hessian matrix

$$H'_r = \begin{pmatrix} 0 & \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_r} \\ \frac{\partial f}{\partial x_1} & \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_r} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial x_r} & \frac{\partial^2 f}{\partial x_r \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_r^2} \end{pmatrix}.$$

(a) A necessary condition for the quasiconcavity of f is that

$$(-1)^r \left|H'_r\right| \geq 0 \text{ for all } r = 1, ..., n \text{ and all } \mathbf{x} \in D.$$

(b) A sufficient condition for the quasiconcavity of f is that

$$(-1)^r \left|H'_r\right| > 0 \text{ for all } r = 1, ..., n \text{ and all } \mathbf{x} \in \mathbb{R}^n_+.$$

(c) If $D \subseteq \mathbb{R}^n_{++}$, f is monotonically increasing and

$$(-1)^r \left|H'_r\right| > 0 \text{ for all } r = 1, ..., n \text{ and all } \mathbf{x} \in D$$

then f is strictly quasiconcave.

Let f be a C^1 quasiconcave function, and $\mathbf{x}$ and $\mathbf{x}'$ two points in its domain with $f(\mathbf{x}) \geq f(\mathbf{x}')$. Theorem 5 above says that the directional derivative of f at the "lower" point, $\mathbf{x}'$, in the direction of the "higher" point, $\mathbf{x}$, in nonnegative. Roughly speaking, the derivative gives us the correct signal about the direction of change of the function, which is quite helpful when looking for a maximum. Notice, however, that plain quasiconcavity is compatible with a zero directional derivative at $\mathbf{x}'$ even if $f(\mathbf{x}) > f(\mathbf{x}')$. Hence, a zero gradient could send the wrong signal that $\mathbf{x}'$ is a candidate for a maximum. Strict quasiconcavity rules out this possibility, as does pseudoconcavity a concept we now define.

Definition 7. (**Pseudoconcavity**) A C^1 function $f : \mathbb{R}^n \to \mathbb{R}$ is pseudoconcave if for every $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ we have

$$f(\mathbf{x}) > f(\mathbf{x}') \text{ implies } Df(\mathbf{x}')(\mathbf{x} - \mathbf{x}') > 0 \tag{1}$$

or, equivalently,

$$Df(\mathbf{x}')(\mathbf{x} - \mathbf{x}') \leq 0 \text{ implies } f(\mathbf{x}) \leq f(\mathbf{x}'). \tag{2}$$

The function f is said to be pseudoconvex if $-f$ is pseudoconcave.

A pseudoconcave function can be characterized as a quasiconcave function having a maximum at $\mathbf{x}^*$ whenever $Df(\mathbf{x}^*) = (0, ..., 0)$.

Note that strict quasiconcavity implies pseudoconcavity, but quasiconcavity does not.

Remark. It can be shown that for nonstationary functions, quasiconcavity does in fact imply pseudoconcavity.

Some Properties of Quasiconcave Functions

Quasiconcavity is often hard to prove. The following composition results simplify the test in some circumstances.

Theorem 8. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are two functions such that $f(\mathbf{x}) = kg(\mathbf{x})$ where $k \in \mathbb{R}$. If g is quasiconcave then f is quasiconcave if $k > 0$ and quasiconvex if $k < 0$.

Proof. (Show this result.) ■

What is a bit surprising is that the sum of quasiconcave functions need not be quasiconcave. (**For instance, suppose** $h(x) = x^3 + x$ **and** $g(x) = -2x$**, both with domains** $\mathbb{R}$. **Show that** $h(x)$ **and** $g(x)$ **are both quasiconcave, while** $h(x) + g(x)$ **is not.)**

Let us now consider functions of functions, i.e. $f(\mathbf{x}) = h[g(\mathbf{x})]$. To be precise suppose $g : D \to \mathbb{R}$ where $D \subset \mathbb{R}^n$ has range $E \subset \mathbb{R}$, and suppose that $h : E \to \mathbb{R}$. Then f is defined by

$$f : D \to \mathbb{R} \text{ with values } f(\mathbf{x}) = h[g(\mathbf{x})].$$

The next theorem states that any monotonic transformation of a quasiconcave function is quasiconcave. This means that quasiconcavity is in fact an ordinal property!

Theorem 9. Suppose $D \subset \mathbb{R}^n$ is convex, $g : D \to \mathbb{R}$ has range E and $h : E \to \mathbb{R}$. Define $f : D \to \mathbb{R}$ by $f(\mathbf{x}) = h[g(\mathbf{x})]$. If h is a monotonic transformation and g is quasiconcave, then f is quasiconcave.

The operation of taking monotonic transformations of a function has the important property of leaving the level sets unchanged in a certain sense. The proof of the previous theorem follows as a direct implication of the next result.

Theorem 10. If f is a monotonic transformation (h) of the function g (under the assumptions of Theorem 9 above) then for all y in the range of g we get $UC_g(y) = UC_f[h(y)]$, $LC_g(y) = LC_f[h(y)]$ and $C_g(y) = C_f[h(y)]$.

Proof. To show the first statement recall that $UC_g(y) = \{\mathbf{x} \in D : g(\mathbf{x}) \geq y\}$, and $UC_f[h(y)] = \{\mathbf{x} \in D : f(\mathbf{x}) = h[g(\mathbf{x})] \geq h(y)\}$. Since h is monotone increasing, $h[g(\mathbf{x})] \geq h(y)$ if and only if $g(\mathbf{x}) \geq y$. Hence $UC_f[h(y)] = \{\mathbf{x} \in D : g(\mathbf{x}) \geq y\} = UC_g(y)$. The other two claims are similar. ■

When a function undergoes a monotonic transformation, its level map remains the same, except that the value labelling on a level set changes from y to $h(y)$. Similarly for upper and lower level sets. Theorem 9 is an immediate implication of this result. If g is quasiconcave, then $UC_g(y)$ is a convex set. By Theorem 10 the upper level sets of f are the same as those of g (except for the labelling) and therefore must be convex sets. So f is quasiconcave.

Consider for instance a function $f : \mathbb{R}^n_{++} \to \mathbb{R}$ where $f(\mathbf{x}) = \left(\sum_{i=1}^n x_i^{1/2}\right)^2$. Define $g : \mathbb{R}^n_{++} \to \mathbb{R}$ by $g(\mathbf{x}) = \sum_{i=1}^n x_i^{1/2}$ and $h : \mathbb{R}_{++} \to \mathbb{R}$ by $h(y) = y^2$. Now $g(.)$ is a concave function (and then quasiconcave), and h is a monotonic transformation. Then f is a quasiconcave function.