14.102, Math for Economists
Fall 2004
Lecture Notes, 10/12/2004

These notes are primarily based on those written by Andrei Bremzen for 14.102 in 2002/3, and by Marek Pycia for the MIT Math Camp in 2003/4. I have made only minor changes to the order of presentation, and added a few short examples. The usual disclaimer applies; questions and comments are welcome.

Nathan Barczi
nab@mit.edu

11 More Optimization Results

11.1 Envelope Theorem

When there is a parameter in the optimization problem, how does the value function (the value of f at the optimum) depend of it? Let's start with the simplest case: Unconstrained optimization:

Theorem 184 *$f : U \times I \to R$ where $U \subset R^n$ open and $I \subset R$ interval is C^1 :*

$$f(x, q)$$

Suppose that for each q, there is a solution $x(q)$. If $V(q) = f(x^*(q), q) = \max_{x \in R^n} f(x, q)$, Suppose that $q \to x^*(q)$ is of class C^1, then:

$$\frac{dV(q)}{dq} = \frac{\partial f}{\partial q}(x^*(q), q)$$

Proof. Take the first order condition: $\frac{\partial f}{\partial x}(x^*(q), q) = 0$. Now, $V'(q) = \frac{\partial f}{\partial q}(x^*(q), q) + x^{*\prime}(q)\frac{\partial f}{\partial x}(x^*(q), q) = \frac{\partial f}{\partial q}(x^*(q), q)$ ■

We can generalize this:

Theorem 185 *Generalization:*

Let $f(x; q)$ be a continuous function, and $x^*(q)$ denote the solution to the problem of maximizing $f(x; q)$ on the constraint set $h_i(x; q) = 0$, $i = 1, ..., k$. The Lagrangian is $\Lambda = f(x; q) - \lambda h(x; q)$ (where the constraints are written as a vector). Then

$$\begin{aligned} \frac{df(x^*(q); q)}{dq} &= \left.\frac{\partial \Lambda}{\partial q}\right|_{x=x^*(q)} \\ &= \left.\frac{\partial f(x; q)}{\partial q}\right|_{x=x^*(q)} - \lambda \left.\frac{\partial h(x; q)}{\partial q}\right|_{x=x^*(q)} \end{aligned}$$

Like in the unconstrained case, this says that we can "ignore" the effect of x changing as q changes, and focus on the direct effect of changes in q.
Proof. *Define the 'value function' as the value of the objective function at the maximum. The value function is written as a function of q, not x, because it is assumed that for any given q, the x's are simply whichever maximize the objective function for the given q (this is why we write $x^*(q)$). Denote this as*

$$M(q) = f(x^*(q), q)$$

Then differentiating both sides with respect to q, we have

$$\frac{dM(q)}{dq} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial q} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial q} + ... + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial q} + \frac{\partial f}{\partial q}$$

From the Lagrangian we have

$$\frac{\partial f}{\partial x_i} = \lambda\frac{\partial h}{\partial x_i}$$

Substituting, we have

$$\frac{dM(q)}{dq} = \lambda\frac{\partial h}{\partial x_1}\frac{\partial x_1}{\partial q} + \lambda\frac{\partial h}{\partial x_2}\frac{\partial x_2}{\partial q} + ... + \lambda\frac{\partial h}{\partial x_n}\frac{\partial x_n}{\partial q} + \frac{\partial f}{\partial q}$$

Since we know that at the optimum, the constraints bind, we can differentiate the following identity:

$$\begin{aligned} h(x; q) &= 0 \\ \frac{\partial h}{\partial x_1}\frac{\partial x_1}{\partial q} + \frac{\partial h}{\partial x_2}\frac{\partial x_2}{\partial q} + ... + \frac{\partial h}{\partial x_n}\frac{\partial x_n}{\partial q} + \frac{\partial h}{\partial q} &= 0 \end{aligned}$$

Substituting again, we have the desired result:

$$\frac{dM(q)}{dq} = -\lambda\frac{\partial h}{\partial q} + \frac{\partial f}{\partial q}$$

■

Application: Consider maximization subject to budget constraints

$$\mathbf{V}(I) = \max_x U(x)$$
$$px \le I$$

In this example

$$\frac{d\mathbf{V}(I)}{dI} = \lambda,$$

so λ is the marginal utility of wealth (called also the shadow price of the constraint).

11.2 Homogeneous/Homothetic Functions

Definition 186 *A function f of non-negative real numbers $(x_1, x_2, ..., x_n)$ is called* ***homogeneous of degree r*** *if*

$$f(kx_1, kx_2, ..., kx_n) = k^r f(x_1, x_2, ..., x_n)$$

for all $k > 0$.

We mostly run across cases with r equal to 0 or 1.

Example 187 $f(x) = \frac{x_1}{x_2}$ *is homogeneous of degree zero.*

Example 188 $f(x) = x_1^{\alpha} x_2^{1-\alpha}$ *is homogeneous of degree one.*

Example 189 *Neither $f(x) = \alpha_0 + \alpha_1 x_1$ nor $f(x) = x_1 + x_2^2$ are homogeneous functions.*

If f is a production function then the degree of homogeneity refers to the degree of returns to scale ($r = 1$ indicates a CRS production function). An important theorem for CRS functions (more generally, functions with degree of homogeneity one) is:

Theorem 190 *(Euler's Theorem): if $f(x_1, x_2, ..., x_n)$ is homogeneous of degree one then*

$$\sum x_i f_i(x_1, x_2, ..., x_n) = f(x_1, x_2, ..., x_n)$$

where f_i refers to the first derivative of f with respect to the i^{th} component.

Proof. *We know that*

$$f(kx_1, kx_2, ..., kx_n) = kf(x_1, x_2, ..., x_n)$$

Now regard k as a variable and look at the derivative of both sides of this equation with respect to k:

$$\sum x_i f_i(x_1, x_2, ..., x_n) = f(x_1, x_2, ..., x_n)$$

which is simply the statement of the theorem. ■

The idea here is that for a CRS production function, total production is simply the sum of each input multiplied by that input's marginal product ('product exhaustion').

Homogeneous functions are very regular in the sense that if we know the value of the function at a single point x we know its value at all points kx proportional to x; it is k^r times the first. In particular, if x and x' are on the same level curve, kx and kx' are on the same level curve as well. Thus, from one level curve, all level curves can be constructed (example: indifference curves). Another way to say this is that level curves are radial expansions and contractions of one another

Functions with this property are called *homothetic*. Homotheticity is a weaker condition than homogeneity, in that a homogeneous function is necessarily homothetic, but the converse need not be true.

Definition 191 *A function $v : \mathbb{R}^n_+ \to \mathbb{R}$ is called* ***homothetic*** *if it is a monotone transformation of a homogeneous function, that is, if there is a monotonic transformation $z \to g(z)$ of $\mathbb{R}_+$ and a homogeneous function $u : \mathbb{R}^n_+ \to \mathbb{R}_+$ such that $v(\mathbf{x}) = g(u(\mathbf{x}))$ for all $\mathbf{x}$ in the domain.*

Example 192 *$f(x) = \alpha_0 + \alpha_1 x_1$ is homothetic, but not homogeneous, as previously noted.*

The definition of homotheticity reveals its attraction. Modern utility theory is founded on ordinal concepts: we do not want to concern ourselves with the actual level of utility experienced by, say, a consumer as he consumes some bundle of goods - we want to focus on how he ranks the utility derived from this bundle, relative to the utility he would get from another bundle. Homothetic functions are attractive because they preserve all the ordinal properties of homogeneous functions (such as the level curve property mentioned above), and retain their homotheticity under monotone transformations (unlike homogeneous functions). This makes them ideal for representing utility.

Remark 193 *Our examples in discussing homogeneity were drawn from production theory; now that we are talking about homotheticity we are referring to utility theory. This is because in production theory, relabeling the quantities along an isoquant really is changing the story - production is more cardinal than utility. So it is interesting to ask if a production function is homogeneous; on the other hand, it is less interesting to ask this about a utility function, which is meant to be entirely ordinal. More interesting is whether a utility function is homothetic.*

Let us first show that a monotone transformation of a homothetic function remains homothetic. In other words, let $z \to h(z)$ be a monotone transformation, and let $v(\mathbf{x})$ be a homothetic function; we want to show that $h(v(\mathbf{x}))$ is homothetic. But we know that $v(\mathbf{x}) = g(u(\mathbf{x}))$ for some homogeneous u and some monotone g, from the definition of homotheticity. So if we can show that $(h \circ g)$ is a monotone transformation - that is, that a monotone transformation of a monotone transformation is still monotone! - then we will have shown that $h(v(\mathbf{x})) = (h \circ g)(u(\mathbf{x}))$ is a monotone transformation of a homogeneous function, that is, homothetic. But this is not hard: suppose $z_1 > z_2$. Then since g is monotone (increasing, let's say - the proof is the same for decreasing, or for nonstrict monotonicity), $g(z_1) > g(z_2)$. And since h is also monotone, $h(g(z_1)) > h(g(z_2))$, so that $(h \circ g)$ is in fact monotone.

We now want to characterize the result that for homothetic functions, as for homogeneous functions, level curves are radial expansions and contractions of one another. First, a few definitions to generalize the idea of monotonicity to higher dimensions:

Definition 194 *If $x, y \in \mathbb{R}^n$, write*

$$\begin{aligned} x &\geq y \text{ if } x_i \geq y_i \text{ for } i = 1, ..., n \\ x &> y \text{ if } x_i > y_i \text{ for } i = 1, ..., n \end{aligned}$$

*A function $u : \mathbb{R}^n_+ \to \mathbb{R}_+$ is **monotone** if for all $x, y \in \mathbb{R}^n_+$,*

$$x \geq y \Rightarrow u(x) \geq u(y)$$

*The function u is **strictly monotone** if for all $x, y \in \mathbb{R}^n_+$,*

$$x > y \Rightarrow u(x) > u(y)$$

Now, the promised characterization of homothetic functions:

Theorem 195 *Let $u : \mathbb{R}^n_+ \to \mathbb{R}_+$ be a strictly monotonic function. Then u is homothetic if and only if for all $x, y \in \mathbb{R}^n_+$,*

$$u(x) \geq u(y) \Leftrightarrow u(\alpha x) \geq u(\alpha y) \text{ for all } \alpha > 0$$

Another important property of homothetic and homogeneous functions is that the slope of level sets is constant along rays from the origin. Formally:

Theorem 196 *Let u be a C^1 function on $\mathbb{R}^n_+$. If u is homothetic, then the slopes of the tangent planes to the level sets of u are constant along rays from the origin; in other words, for every i, j and for every $x \in \mathbb{R}^n_+$,*

$$\frac{\frac{\partial u}{\partial x_i}(tx)}{\frac{\partial u}{\partial x_j}(tx)} = \frac{\frac{\partial u}{\partial x_i}(x)}{\frac{\partial u}{\partial x_j}(x)} \text{ for all } t > 0$$

This is of interest to us in economics because it states that if u is homothetic, then its marginal rate of substitution is a homogeneous function of degree zero.