14.102, Math for Economists
Fall 2005
Lecture Notes, 10/18/2005

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, mostly from Rudin. The usual disclaimer applies; questions and comments are welcome.

Nathan Barczi
nab@mit.edu

11 Static Optimization II

11.1 Inequality Constrained Optimization

Similar logic applies to the problem of maximizing $f(x)$ subject to inequality constraints $h_i(x) \leq 0$. At any point of the feasible set some of the constraints will be binding (i.e., satisfied with equality) and others will not. For the first order conditions only binding constraints matter and only their gradients play a role; this can be captured by allowing only multipliers corresponding to binding constraints to be nonzero in the first order condition for an optimum.

Consider again the two-dimensional example discussed above. Now we will maximize $f(x)$ subject to $g(x) \leq 0$. In an optimum where the constraint is not binding the problem locally looks like an unconstrained problem and the first order condition will be $\nabla f = 0$. In an optimum where the constraint is binding (and $\nabla g \neq 0$), by the theorem of Lagrange it must be the case that $\nabla f = \lambda \nabla g$ for some λ (note that the case $\nabla f = 0$ also satisfied this condition for $\lambda = 0$). But now we can say more: if λ were negative, we could move slightly from the prospective maximum in the direction of ∇f, and that will not violate the constraint (we would be moving in the direction opposite to ∇g, so g would decrease and hence still remain nonpositive). Therefore, at any local optimum it must be the case that $\lambda \geq 0$.

The intuition of the example above is summarized by the following

Theorem 257 *(The Theorem of Kuhn and Tucker) Let $f, h_i : \mathbb{R}^n \to \mathbb{R}$ be C^1 functions, $i = 1, ..., l$. Suppose x^* is a local maximum of f on the set*

$$Z = U \cap \{x \in \mathbb{R}^n | h_i(x) \leq 0, i = 1, ..., l\},$$

where U is an open set in $\mathbb{R}^n$.Suppose that all constraints that are binding at x^* have linearly independent gradients at x^*. Then there exist real numbers $\lambda_i^*, i = 1, ..., l$, such that:

- $\lambda_i^* \geq 0$ *and* $\lambda_i^* h_i(x^*) = 0, \ \ i = 1, ..., l$

- $\nabla f(x^*) = \sum_{i=1}^{l} \lambda_i^* h_i(x^*)$.

The conditions that $\lambda_i^* h_i(x^*) = 0$ are called *complementary slackness* conditions. Essentially they state that nontrivial Lagrange multipliers ($\lambda_i^* \neq 0$) may come only with constraints that are binding at x^* ($h_i(x^*) = 0$). Constraint qualification is similar to that in the Theorem of Lagrange with the obvious modification that only gradients of binding constraints count. Again, it is *almost* always safe to ignore them, but *generally* it is not.

Exercise 258 *Consider the consumer's utility maximization problem:* $\max u(x_1, ..., x_n)$ *subject to* $x_i \geq 0, i = 1, ..., n$ *and* $\sum_{i=1}^{n} p_i x_i \leq I$ *where* $p_1, ..., p_n, I > 0$. *Show that the constraint qualification condition is satisfied at any feasible point.*

Similar to the equality constraint optimization, one can set up Lagrangean

$$L(x_1, ..., x_n, \lambda_1, ..., \lambda_k) = f(x_1, ..., x_n) - \sum_{i=1}^{l} \lambda_i h_i(x_1, ..., x_n).$$

The theorem of Kuhn and Tucker then gives conditions on L that must be satisfied at a local optimum (x^*, λ^*) :

- $\frac{\partial L}{\partial x_j} = 0,\ j = 1, ..., n$
- $\lambda_i \frac{\partial L}{\partial \lambda_i} = 0, i = 1, ..., l$
- $\lambda_i \geq 0, \frac{\partial L}{\partial \lambda_i} \geq 0, i = 1, ..., l$

The first two of the above conditions constitute $n + l$ equations on $n + l$ unknowns. Solving this (non-linear) system gives all points (typically, finitely many) that are candidates for a local maximum.

Unfortunately, there is generally no way of telling ex ante which constraints will end up binding at the optimum and which will not. For example, if there are five constraints, there will be $2^5 = 32$ possible combinations of binding constraints. However, some conclusions will typically follow from the economics: for instance, if one of the constraints is the budget constraint, one can argue that it will be binding (you want to use up all your resources).

Finally, let us consider the case of mixed constraints: some $g_i(x) = 0, i = 1, ..., k$, and some $h_j(x) \leq 0, i = k + 1, ..., k + l$. Combining the Theorem of Lagrange with the Theorem of Kuhn and Tucker gives the following

Theorem 259 *Let* $f, g_i, h_j : \mathbb{R}^n \to \mathbb{R}$ *be* C^1 *functions,* $i = 1, ..., l, \ \ j = k + 1, ..., k + l$. *Suppose* x^* *is a local maximum of* f *on the set*

$$Z = U \cap \{x \in \mathbb{R}^n | g_i(x) = 0, i = 1, ..., k, h_j(x) \leq 0, j = k + 1, ..., k + l\},$$

where U *is an open set in* $\mathbb{R}^n$.*Suppose that all constraints that are binding at* x^* *have linearly independent gradients at* x^*.*Then there exist real numbers* $\lambda_i^*, i = 1, ..., k + l$, *such that:*

- $\lambda_i^* \geq 0$ *and* $\lambda_i^* h_i(x^*) = 0, \;\; i = k+1, ..., k+l$
- $\nabla f(x^*) = \sum\limits_{i=1}^{k} \lambda_i^* g_i(x^*) + \sum\limits_{i=k+1}^{k+l} \lambda_i^* h_i(x^*)$.

11.2 Homogeneous/Homothetic Functions

Definition 260 *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 261 $f(x) = \frac{x_1}{x_2}$ *is homogeneous of degree zero.*

Example 262 $f(x) = x_1^{\alpha} x_2^{1-\alpha}$ *is homogeneous of degree one.*

Example 263 *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 264 *(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 265 *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 266 *$f(x) = \alpha_0 + \alpha_1 x_1$ 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 267 *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 268 *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 269 *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 270 *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.