Mathematics for Economists
Aalto BIZ
Spring 2021
JuusoVälimäki

Nonlinear functions in economics

Utility functions

The utility function $u : \mathbb{R}^n_+ \to \mathbb{R}$ of a consumer assigns a numerical value $u(\boldsymbol{x})$ for each possible (positive) consumption vector $\boldsymbol{x} \in \mathbb{R}^n_+ \Delta = \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{x} \geq 0\}$. The economic interpretation of u is that $u(\boldsymbol{x}) \geq u(\boldsymbol{y})$ if and only if the consumer considers $\boldsymbol{x}$ at least as good as $\boldsymbol{y}$. It is not a strong assumption that the consumer's preferences can be represented by such a function (more on this in future courses).

Marginal utilities and the marginal rate of substitution

In order to have a nontrivial consumer problem, $n \geq 2$. In the simplest case, there are only two goods and we can draw illustrative pictures. We start with the interpretation of the partial derivatives of the utility function called the *marginal utilities* denoted by MU_{x_i}.

$$MU_{x_i}(\hat{x}_1, \hat{x}_2) := \frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_i}.$$

If $\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1} > 0$, then $u(\hat{x}_1 + h, \hat{x}_2) > u(\hat{x}_1, \hat{x}_2)$ for small $h > 0$, and we say that utility is strictly increasing in good 1 at $(\hat{x}_1, \hat{x}_2)$. If this holds at all (x_1, x_2), we say simply that utility is strictly increasing. Typically it is assumed that utility is strictly increasing in all goods.

For small consumption changes $(\Delta x_1, \Delta x_2)$, we can approximate the change in utility by using the derivative $D_{\boldsymbol{x}} u(\hat{x}_1, \hat{x}_2)$:

$$u(\hat{x}_1 + \Delta x_1, \hat{x}_2 + \Delta x_2) - u(\hat{x}_1, \hat{x}_2) = D_{\boldsymbol{x}} u(\hat{x}_1, \hat{x}_2)(\Delta x_1, \Delta x_2)$$

$$= \frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1} \Delta x_1 + \frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_2} \Delta x_2.$$

Recall from Principles 1 that we say that $(\hat{x}_1+\Delta x_1, \hat{x}_2+\Delta x_2)$ and $(\hat{x}_1, \hat{x}_2)$ are on the same indifference curve if they are equally good to the consumer. In terms of the utility function, this means that $u(\hat{x}_1 + \Delta x_1, \hat{x}_2 + \Delta x_2) = u(\hat{x}_1, \hat{x}_2)$. But then we have:

$$\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1}\Delta x_1 + \frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_2}\Delta x_2 = 0,$$

or

$$\Delta x_2 = -\frac{\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1}}{\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_2}}\Delta x_1.$$

In other words, the consumer is willing to give up $\frac{\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1}}{\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_2}}$ units of good 2 to get an additional unit of good 1 at $(\hat{x}_1, \hat{x}_2)$. We have then argued that the familiar concept of marginal rate of substitution at $(\hat{x}_1, \hat{x}_2)$ is captured in the ratio of marginal utilities:

$$MRS_{x_1,x_2}(\hat{x}_1, \hat{x}_2) = \frac{\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1}}{\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_2}} = \frac{MU_{x_1}(\hat{x}_1, \hat{x}_2)}{MU_{x_2}(\hat{x}_1, \hat{x}_2)}.$$

Nothing in the above discussion depended on having only two goods. If $n > 2$ we can ask how many (small) units of good j the consumer would be willing to give up in order to get an additional (small) unit of good i. All the other goods remain fixed at $\hat{\boldsymbol{x}}$. If $u(\hat{\boldsymbol{x}} + \Delta x_i \boldsymbol{e}^i + \Delta x_j \boldsymbol{e}^j) = u(\hat{\boldsymbol{x}})$, then

$$\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_i}\Delta x_i + \frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_j}\Delta x_j = 0,$$

and we have:

$$MRS_{x_i,x_j}(\hat{\boldsymbol{x}}) = \frac{\frac{\partial u(\hat{\boldsymbol{x}})}{\partial x_i}}{\frac{\partial u(\hat{\boldsymbol{x}})}{\partial x_j}} = \frac{MU_{x_i}(\hat{\boldsymbol{x}})}{MU_{x_j}(\hat{\boldsymbol{x}})}.$$

Computing marginal utilities and MRS

1. Linear utility:

$$u(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{x} = \sum_{i=1}^{n} a_i x_i.$$

   Then $MU_{x_i}(\boldsymbol{x}) = a_i$ for all i and all $\boldsymbol{x}$, and $MRS_{x_i,x_j} = \frac{a_i}{a_j}$ for all $i \neq j$ and all $\boldsymbol{x}$.

2. Quasilinear utility:

$$u(x_1, x_2) = v(x_1) + x_2,$$

for some increasing function v. In this case, $MU_{x_2} = 1, MU_{x_1} = MRS_{x_1,x_2} = v'(x_1)$.

For example if $v(x_1) = \ln(x_1)$, then $MU_{x_1} = MRS_{x_1,x_2} = \frac{1}{x_1}$.

3. Cobb-Douglas utility:

$$u(x_1, x_2) = x_1^\alpha x_2^{1-\alpha} \text{ for } \alpha \in (0,1).$$

In this case, $MU_{x_1}(x_1, x_2) = \alpha(\frac{x_2}{x_1})^{1-\alpha}$, $MU_{x_1}(x_1, x_2) = (1-\alpha)(\frac{x_1}{x_2})^\alpha$, and therefore:

$$MRS_{x_1,x_2} = \frac{\alpha(\frac{x_2}{x_1})^{1-\alpha}}{(1-\alpha)(\frac{x_1}{x_2})^\alpha} = \frac{\alpha x_2}{(1-\alpha)x_1}.$$

For $n > 2$, we have:

$$u(\boldsymbol{x}) = \prod_{i=1}^n x_i^{\alpha_i} \text{ for } \alpha_i > 0, \sum_{i=1}^n \alpha_i = 1.$$

Denote $y = \prod_{i=1}^n x_i^{\alpha_i}$. Then $MU_{x_i} = \alpha_i \frac{y}{x_i}$ and therefore

$$MRS_{x_i,x_j}(\boldsymbol{x}) = \frac{\alpha_i x_j}{\alpha_j x_i}.$$

Exercise: Compute the marginal utilities for $u(\boldsymbol{x}) = \sum_{i=1}^n \alpha_i \ln x_i$.

4. Constant elasticity of substitution utility (CES utility function): for $\rho < 1$ and $\rho \neq 0$,

$$u(x_1, x_2) = (a_1 x_1^\rho + a_2 x_2^\rho)^{\frac{1}{\rho}}.$$

Denote $y(x_1, x_2) = (a_1 x_1^\rho + a_2 x_2^\rho)$ and we have $u(x_1, x_2) = y(x_1, x_2)^{\frac{1}{\rho}}$. We get by chain rule that

$$MU_{x_i}(x_1, x_2) = \frac{1}{\rho} y(x_1, x_2)^{\frac{1-\rho}{\rho}} \rho a_i x_i^{\rho-1}.$$

Therefore the marginal rates of substitution are quite simple:

$$MRS_{x_1,x_2}(x_1, x_2) = (\frac{a_i}{a_j})(\frac{x_i}{x_j}^{\rho-1}.$$

Exercise: Can you connect the limit of the MRS as $\rho \to 1$ to some other utility function?

The gradient

The gradient of the utility function denoted by $\nabla u(\boldsymbol{x})$ is the transpose of its derivative. In other words, it is the column vector of partial derivatives.

$$\nabla u(\boldsymbol{x}) = (\frac{\partial u(\boldsymbol{x})}{\partial x_1}, ..., \frac{\partial u(\boldsymbol{x})}{\partial x_n}).$$

Does the gradient have any particular interpretation? A first observation is that when $n = 2$, the gradient is orthogonal to the indifference curve. In other words, the inner product of the gradient and any direction that leaves the utility level unchanged is zero.

$$(1, -\frac{\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1}}{\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_2}}) \cdot (\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1}, \frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_2}) = \frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1} - \frac{\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1}}{\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_2}} \frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_2} = 0.$$

Therefore the gradient is orthogonal to indifference curves.

The gradient at $\boldsymbol{x}$ gives the direction in which the utility function increases the fastest near $\boldsymbol{x}$. To see this, consider the change in the utility using the linear approximation:

$$u(\hat{\boldsymbol{x}} + \Delta \boldsymbol{x}) - u(\hat{\boldsymbol{x}}) = D_{\boldsymbol{x}}(\hat{\boldsymbol{x}}) \Delta \boldsymbol{x} + h.o.t.$$

For a unit length (or norm) of $\Delta \boldsymbol{x}$, the change in utility is maximized at $\Delta \boldsymbol{x} = \frac{1}{\|\nabla u(\hat{\boldsymbol{x}})\|} \nabla u(\hat{\boldsymbol{x}})$ by Cauchy's inequality stating that for all $\boldsymbol{x}, \boldsymbol{y}$, $\boldsymbol{x} \cdot \boldsymbol{y} \leq \|\boldsymbol{x}\| \|\boldsymbol{y}\|$.

Production functions

A firm produces output Y by combining inputs K, L interpreted as capital and labor according to a production function $Y = F(K, L)$. The partial derivatives are interpreted as marginal products so that $\frac{\partial F(\hat{K}, \hat{L})}{\partial K} =: MP_K(\hat{K}, \hat{L})$ is the marginal product of capital and $\frac{\partial F(\hat{K}, \hat{L})}{\partial L} =: MP_L(\hat{K}, \hat{L})$ is the marginal product of labor. These are naturally assumed to be nonnegative and typically strictly positive.

When considering the substitutability between the factors in production, we may ask how many (small) units of labor can be given up if one more (small) unit of capital is used if the level of production is to remain unchanged. In other words, we are asking what are the $\Delta K, \Delta L$ such that:

$$\frac{\partial F(\hat{K},\hat{L})}{\partial K}\Delta K + \frac{\partial F(\hat{K},\hat{L})}{\partial L}\Delta L = 0.$$

Solving for ΔL, we get

$$\Delta L = -\frac{\frac{\partial F(\hat{K},\hat{L})}{\partial K}}{\frac{\partial F(\hat{K},\hat{L})}{\partial L}}\Delta K.$$

We call $frac\frac{\partial F(\hat{K},\hat{L})}{\partial K}\frac{\partial F(\hat{K},\hat{L})}{\partial L}$ the *marginal rate of technical substitution* between capital and labor, $MRTS_{K,L}$. The corresponding notion to indifference curves is given by an isoquant. An isoquant traces those combinations of (K, L) that result in the same output.

Since the relevant production functions are: linear, Cobb-Douglas and the CES-function, I will not go over the same calculations as above. Just substitute (K, L) for (x_1, x_2) in the appropriate places and you have the results.

Homogenous functions

Homogenous functions are an important class of functions for many economic applications. A function $f : \mathbb{R}^n \to \mathbb{R}$ is homogenous of degree k if for all $\lambda > 0$ and $\boldsymbol{x}$,

$$f(\lambda\boldsymbol{x}) = \lambda^k f(\boldsymbol{x}). \tag{1}$$

We say that a production function has constant returns to scale if it is homogenous of degree 1. In this case, scaling all inputs in the same proportion scales the output in the same proportion. Euler's theorem for homogenous functions yields an important observation.

Proposition 1. If f is homogenous of degree 1, then

$$D_{\boldsymbol{x}} f(\boldsymbol{x})\boldsymbol{x} = kf(\boldsymbol{x}),$$

or writing out the partial. derivatives:

$$\sum_{i=1}^{n} x_i \frac{\partial f(\boldsymbol{x})}{\partial x_i} = kf(\boldsymbol{x}).$$

Proof. Since equation (1) holds for all λ, we can take derivatives with respect to λ on both sides to get:

$$\sum_{i=1}^{n} x_i \frac{\partial f(\lambda \boldsymbol{x})}{\partial x_i} = k\lambda^{k-1} f(\boldsymbol{x}).$$

Setting $\lambda = 1$ gives the claim. □

Why is this an important result? We shall see that if capital and labor are chosen optimally in a firm, then $MP_K = r$ and $MP_L = w$, where w is the market wage rate and r is the market rate for renting capital. We get from Euler's theorem that $Y = rK + wL$. In other words, capital and labor compensations exhaust the entire output.

For another application, take the consumers problem of dividing income I between two goods, x_1, x_2 at prices p_1, p_2. As you recall from Principles 1, the budget set of the consumer does not change if all prices and the income are multiplied by the same number. Since the choice set is the same, optimal choices should not change either. If we write $x_1(p_1, p_2, I), x_2(p_1, p_2, I)$ as the demands for the two goods, we have argued that $x_i(p_1, p_2, I) = x_i(\lambda p_1, \lambda p_2, \lambda I)$, i.e. the demands are homogenous of degree 0 in prices and income.

Euler's theorem gives then:

$$p_i \frac{\partial x_i(p_1, p_2, I)}{\partial p_1} + \ p_2 \frac{\partial x_i(p_1, p_2, I)}{\partial p_2} + I \frac{\partial x_i(p_1, p_2, I)}{\partial I} = 0.$$

Dividing both sides by x_i, we get:

$$\frac{p_1}{x_i} \frac{\partial x_i(p_1, p_2, I)}{\partial p_1} + \frac{p_2}{x_i} \frac{\partial x_i(p_1, p_2, I)}{\partial p_2} + \frac{I}{x_i} \frac{\partial x_i(p_1, p_2, I)}{\partial I} = 0,$$

and writing with price and income elasticities, we get:

$$\varepsilon_{p_1}^{x_i} + \varepsilon_{p_2}^{x_i} = \varepsilon_{I}^{x_i},$$

where $\varepsilon_{p_j}^{x_i}$ is the price elasticity of good i with respect to price j and $\varepsilon_{I}^{x_i}$ is its income elasticity.