August 4, 2018

CHAPTER 4: HIGHER ORDER DERIVATIVES

In this chapter D denotes an open subset of $\mathbb{R}^n$.

1. Introduction

Definition 1.1. Given a function $f : D \to \mathbb{R}$ we define the second partial derivatives as

$$D_{ij}f = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)$$

Likewise, we may define the higher order derivatives.

Example 1.2. Consider the function

$$f(x, y, z) = xy^2 + e^{zx}$$

then

$$\frac{\partial f}{\partial x} = y^2 + ze^{zx} \qquad \frac{\partial f}{\partial y} = 2xy \qquad \frac{\partial f}{\partial z} = xe^{zx}$$

and, for example

$$\frac{\partial^2 f}{\partial x \partial x} = z^2 e^{zx} \qquad \frac{\partial^2 f}{\partial x \partial z} = xe^{zx} \qquad \frac{\partial^2 f}{\partial z \partial x} = xe^{zx}$$

We see that

$$\frac{\partial^2 f}{\partial x \partial z} = \frac{\partial^2 f}{\partial z \partial x}$$

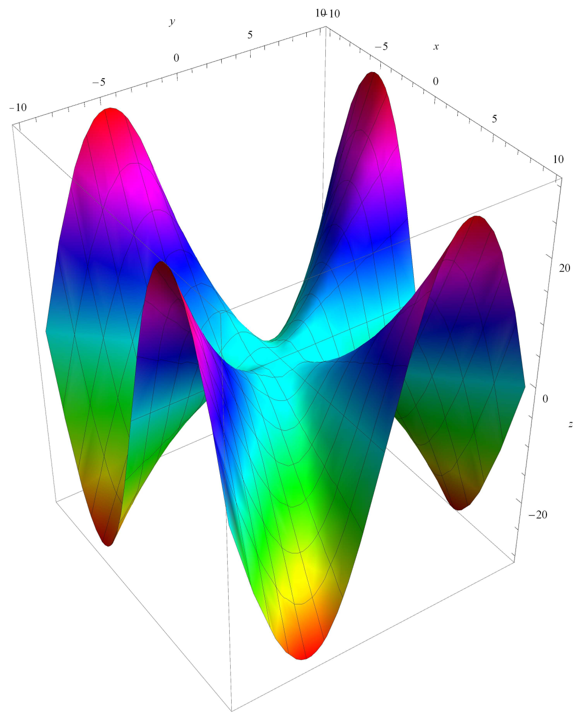
We may check that this also holds for the other variables

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \qquad \frac{\partial^2 f}{\partial y \partial z} = \frac{\partial^2 f}{\partial z \partial y}$$

Example 1.3. Consider the function

$$f(x, y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2}, & \text{si } (x, y) \neq (0, 0); \\ 0, & \text{si } (x, y) = (0, 0). \end{cases}$$

Here is the graph of f,



We may check easily that for $(x,y) \neq (0,0)$,

$$\frac{\partial f}{\partial x}(x,y) = \frac{x^4 y + 4x^2 y^3 - y^5}{(x^2+y^2)^2} \qquad \frac{\partial f}{\partial y}(x,y) = \frac{x^5 - 4x^3 y^2 - xy^4}{(x^2+y^2)^2}$$

and

$$\frac{\partial f}{\partial x}(0,0) = 0 \qquad \frac{\partial f}{\partial y}(0,0) = 0$$

Then

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{x \to 0} \frac{\frac{\partial f}{\partial y}(x,0) - \frac{\partial f}{\partial y}(0,0)}{x - 0} = \lim_{x \to 0} \frac{x}{x} = 1$$

and

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = \lim_{y \to 0} \frac{\frac{\partial f}{\partial x}(0,y) - \frac{\partial f}{\partial x}(0,0)}{y - 0} = \lim_{x \to 0} \frac{-y}{y} = -1$$

so

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0)$$

On the other hand, one can check that if $(x,y) \neq (0,0)$ then

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y)$$

The following result provides conditions under which the cross partial derivatives coincide.

Theorem 1.4 (Schwarz)**.** Suppose that for some $i, j = 1 \dots, n$ the partial derivatives

$$\frac{\partial f}{\partial x_i}, \quad \frac{\partial^2 f}{\partial x_i \partial x_j}, \quad \frac{\partial f}{\partial x_j}, \quad \frac{\partial^2 f}{\partial x_j \partial x_i}$$

exist and are continuous in some ball $B(p, r)$, with $r > 0$. Then,

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$$

for every x in the ball $B(p, r)$.

Definition 1.5. Let D be an open subset of $\mathbb{R}^n$ and let $f : D \to \mathbb{R}$. We say that f is of class

- $C^1(D)$ if all the first partial derivatives $\frac{\partial f}{\partial x_i}$ of f exist and are continuous on D for all $i = 1 \dots, n$.
- $C^2(D)$ if all the first partial derivatives
$$\frac{\partial f}{\partial x_i}$$
of f exist and are of class $C^1(D)$ for every $i = 1 \dots, n$.
- $C^k(D)$ if all the first partial derivatives
$$\frac{\partial f}{\partial x_i}$$
of f exist and are of class $C^{k-1}(D)$ for every $i = 1 \dots, n$.

We write $f \in C^k(D)$.

Definition 1.6. Let $f \in C^2(D)$. The Hessian matrix of f at p is the matrix

$$\mathrm{D}^2 f(p) = \mathrm{H}\, f(p) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(p) \right)_{i,j=1,\dots,n}$$

Remark 1.7. Note that by Schwarz's theorem, if $f \in C^2(D)$ then the matrix $\mathrm{H}\, f(p)$ is symmetric.

2. The Implicit function Theorem

In this section we are going to study **non-linear** systems of equations. For example,

$$\begin{aligned} x^2 + z e^{xy} + z &= 1 \\ 3x + 2y + z &= 3 \end{aligned} \tag{2.1}$$

In general, it is extremely difficult to prove that there is a solution (and it may not exist) o to solve explicitly those systems. Nevertheless, in Economics it often happens that the model we are interested in is described by a system of equations such as, for example, the system 2.1. And we would like to be able to say something about the dependence of the solution with respect to the parameters. In this section we address this problem.

Firstly, we note that that a system of m equations and n unknowns may be written in the following form

$$\begin{aligned} f_1(u) &= 0 \\ f_2(u) &= 0 \\ &\vdots \\ f_m(u) &= 0 \end{aligned}$$

where $u \in \mathbb{R}^n$ and $f_1, f_2, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$. For example, the system 2.1 may be written as

$$\begin{aligned} f_1(u) &= 0 \\ f_2(u) &= 0 \end{aligned}$$

with $f_1(x, y, z) = x^2 + ze^{xy} + z - 1$ and $f_2(x) = 3x + 2y + z - 3$.

How are the solutions of the system 2.1. Comparing the situation with a linear system we should expect to be able to solve for two of the variables in terms of one parameter, since there are 2 equations and 3 unknowns. Suppose, for example that we want to solve for y, z as functions of x. This might be complicated and in most of the cases impossible. In this situation, the implicit function Theorem,

- provides sufficient conditions under which the system 2.1 has a unique solution, that is it to proves the existence of two functions $y(x)$ y $z(x)$ which satisfy the equations 2.1, even if we do not know how to compute these functions.
- when the system of equations 2.1 has a solution it permits us to obtain an expression for $y'(x)$ and $z'(x)$, even if we do not know how to compute $y(x)$, $z(x)$.

Let us consider a system of equations

$$\begin{aligned} f_1(u, v) &= 0 \\ f_2(u, v) &= 0 \\ &\vdots \\ f_m(u, v) &= 0 \end{aligned} \tag{2.2}$$

where $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$ are the independent variables and $v = (v_1, \ldots, v_m) \in \mathbb{R}^m$ are the variables for which we want to solve for[1] and $f_1, f_2, \ldots, f_m : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$. To this system we associate the following expression

$$\frac{\partial (f_1, f_2, \ldots, f_m)}{\partial (v_1, \ldots, v_m)} = \det \begin{pmatrix} \frac{\partial f_1}{\partial v_1} & \cdots & \frac{\partial f_1}{\partial v_m} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial v_1} & \cdots & \frac{\partial f_m}{\partial v_m} \end{pmatrix}$$

For example, for the system 2.1

$$\frac{\partial (f_1, f_2)}{\partial (y, z)} = \det \begin{pmatrix} xze^{xy} & e^{xy} + 1 \\ 2 & 1 \end{pmatrix} = xze^{xy} - 2e^{xy} - 2$$

[1] In the example 2.1 $n = 1$, $m = 2$, $u = x$, $v = (y, z)$.

Theorem 2.1 (implicit function Theorem)**.** Suppose that the functions $f_1, f_2, \ldots, f_m : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ are of class C^1 and that there is a point $(u_0, v_0) \in \mathbb{R}^n \times \mathbb{R}^m$ such that

(1) $f_1(u_0, v_0) = f_2(u_0, v_0) = \cdots = f_m(u_0, v_0) = 0$; and
(2)
$$\frac{\partial\left(f_1, f_2, \ldots, f_m\right)}{\partial\left(v_1, \ldots, v_m\right)}(u_0, v_0) \neq 0$$

Then, there are an sets $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ and functions $g_1, \ldots g_m : U \to \mathbb{R}$ such that

(1) $u_0 \in U$, $v_0 \in V$.
(2) for every $u \in U$,

$$f_1(u, g_1(u), \ldots, g_m(u)) = f_2(u, g_1(u), \ldots, g_m(u)) = \cdots = f_m(u, g_1(u), \ldots, g_m(u)) = 0$$

(3) If $u \in U$ and $v = (v_1, \ldots, v_m) \in V$ are solutions of the system of equations $f_1(u, v) = f_2(u, v) = \cdots = f_m(u, v) = 0$, then $v_1 = g_1(u), \ldots v_m = g_m(u)$.
(4) The functions $g_1, \ldots g_m : U \to \mathbb{R}$ are differentiable and for each $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$ we have that

$$\frac{\partial g_i}{\partial u_j} = -\frac{\partial\left(f_1, f_2, \ldots, f_m\right)}{\partial\left(v_1, \ldots, v_{i-1}, u_j, v_{i+1}, \ldots, v_m\right)} \Big/ \frac{\partial\left(f_1, f_2, \ldots, f_m\right)}{\partial\left(v_1, \ldots, v_m\right)} \tag{2.3}$$

Remark 2.2. Explicitly,

$$\frac{\partial\left(f_1, f_2, \ldots, f_m\right)}{\partial\left(v_1, \ldots, v_{i-1}, u_j, v_{i+1}, \ldots, v_m\right)} = \det\begin{pmatrix} \frac{\partial f_1}{\partial v_1} & \cdots & \frac{\partial f_1}{\partial v_{i-1}} & \frac{\partial f_1}{\partial u_j} & \frac{\partial f_1}{\partial v_{i+1}} & \cdots & \frac{\partial f_1}{\partial v_m} \\ \vdots & \cdots & & \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial v_1} & \cdots & \frac{\partial f_m}{\partial v_{i-1}} & \frac{\partial f_m}{\partial u_j} & \frac{\partial f_1}{\partial v_{i+1}} & \cdots & \frac{\partial f_m}{\partial v_m} \end{pmatrix}$$

Remark 2.3. The conclusion of the implicit function Theorem may be expressed in the following way,

(1) The functions
$$z_1 = g_1(u), z_2 = g_1(u), \ldots, z_m = g_m(u)$$
are a solution of the system of equations 2.2.
(2) The derivatives of the functions $g_1, \ldots g_m : U \to \mathbb{R}$ may be computed by implicitly differentiating the system of equations 2.2 and applying the chain rule.

Remark 2.4. Applying several times the implicit function Theorem we may also compute the higher order derivatives of the dependent variables.

Example 2.5. Let us apply the implicit function Theorem to the system of equations

$$\begin{aligned} x^2 + ze^{xy} + z &= 1 \\ 3x + 2y + z &= 3 \end{aligned} \tag{2.4}$$

First we note that $x = 1$, $y = z = 0$ is a solution of the system. On the other hand, we have seen that

$$\frac{\partial\left(f_1, f_2\right)}{\partial\left(y, z\right)}(1,0,0) = \det\begin{pmatrix} xze^{xy} & e^{xy}+1 \\ 2 & 1 \end{pmatrix}\Bigg|_{x=1,y=z=0} = \left(xze^{xy} - 2e^{xy} - 2\right)\big|_{x=1,y=z=0} = -4 \neq 0$$

The implicit function Theorem guarantees that we may solve for the variables y and z as functions of x for values of x near 1. Furthermore, differentiating with respect to x in the system we obtain

$$\begin{aligned} 2x + z'e^{xy} + z(y + xy')e^{xy} + z' &= 0 \\ 3 + 2y' + z' &= 0 \end{aligned} \tag{2.5}$$

Now substitute $x = 1, y = z = 0$,

$$\begin{aligned} 2 + 2z'(1) &= 0 \\ 3 + 2y'(1) + z'(1) &= 0 \end{aligned} \tag{2.6}$$

so that $z'(1) = y'(1) = -1$. This could be computed as well using formula 2.3,

$$y'(1) = -\frac{\frac{\partial(f_1,f_2)}{\partial(x,z)}(1,0,0)}{-4} = \frac{1}{4}\det\begin{pmatrix} 2x + yze^{xy} & e^{xy}+1 \\ 3 & 1 \end{pmatrix}\Bigg|_{x=1,y=z=0} = \frac{-4}{4} = -1$$

and

$$z'(1) = -\frac{\frac{\partial(f_1,f_2)}{\partial(y,x)}(1,0,0)}{-4} = \frac{1}{4}\det\begin{pmatrix} xze^{xy} & 2x + yze^{xy} \\ 2 & 3 \end{pmatrix}\Bigg|_{x=1,y=z=0} = \frac{-4}{4} = -1$$

To compute the second derivatives $y''(x)$ y $z''(x)$, we differentiate each equation of the system 2.5 with respect to x. After simplifying we obtain

$$\begin{aligned} 2 + z''e^{xy} + 2z'(y + xy')e^{xy} + z(2y' + xy'')e^{xy} + z(y + xy')^2 e^{xy} + z'' &= 0 \\ 2y'' + z'' &= 0 \end{aligned}$$

and substituting $x = 1, y(1) = z(1) = 0$, $z'(1) = y'(1) = -1$

$$\begin{aligned} 2 + 2z''(1) &= 0 \\ 2y''(1) + z''(1) &= 0 \end{aligned}$$

from here we see that $z''(1) = -1$, $y''(1) = 1/2$. Iterated differentiation allows us to obtain the derivatives of any order $z^{(n)}(1)$, $y^{(n)}(1)$.

Example 2.6. Consider the macroecononomic model

$$\begin{aligned} Y &= C + I + G \\ C &= f(Y - T) \\ I &= h(r) \\ r &= m(M) \end{aligned} \tag{2.7}$$

where the variables are Y (national income), C (consumption), I (investment) and r (interest rate) and the parameters are M (money supply), T (taxes collected) and G (public spending). We assume that $0 < f'(z) < 1$. Compute

$$\frac{\partial Y}{\partial M}, \quad \frac{\partial Y}{\partial T}, \quad \frac{\partial Y}{\partial G}$$

The system may be written as follows

$$\begin{aligned} f_1 &= C + I + G - Y = 0 \\ f_2 &= f(Y - T) - C = 0 \\ f_3 &= h(r) - I = 0 \\ f_4 &= m(M) - r \end{aligned}$$

First we compute

$$\frac{\partial\left(f_1,f_2,f_3,f_4\right)}{\partial\left(Y,C,I,r\right)}=\det\begin{pmatrix}-1 & 1 & 1 & 0\\ f'(Y-T) & -1 & 0 & 0\\ 0 & 0 & -1 & h'(r)\\ 0 & 0 & 0 & -1\end{pmatrix}=1-f'(Y-T)\neq 0$$

By the implicit function Theorem the system 2.7 defines implicitly Y, C, I and r as functions of M, T and G. (we assume that the system has some solution). Differentiating in 2.7 with respect to M we obtain

$$\begin{aligned}\frac{\partial Y}{\partial M} &= \frac{\partial C}{\partial M}+\frac{\partial I}{\partial M}\\ \frac{\partial C}{\partial M} &= f'(Y-T)\frac{\partial Y}{\partial M}\\ \frac{\partial I}{\partial M} &= h'(r)\frac{\partial r}{\partial M}\\ \frac{\partial r}{\partial M} &= m'(M)\end{aligned}$$

Solving these equations, we obtain

$$\frac{\partial Y}{\partial M}=\frac{h'(r)m'(M)}{1-f'(Y-T)}$$

Differentiating in 2.7 with respect to T we obtain

$$\begin{aligned}\frac{\partial Y}{\partial T} &= \frac{\partial C}{\partial T}+\frac{\partial I}{\partial T}\\ \frac{\partial C}{\partial T} &= f'(Y-T)(\frac{\partial Y}{\partial T}-1)\\ \frac{\partial I}{\partial T} &= h'(r)\frac{\partial r}{\partial T}\\ \frac{\partial r}{\partial T} &= 0\end{aligned}$$

Solving these equations, we obtain

$$\frac{\partial Y}{\partial t}=\frac{-f'(Y-T)}{1-f'(Y-T)}$$

Differentiating in 2.7 with respect to G we obtain

$$\begin{aligned}\frac{\partial Y}{\partial G} &= \frac{\partial C}{\partial G}+\frac{\partial I}{\partial G}+1\\ \frac{\partial C}{\partial G} &= f'(Y-T)\frac{\partial Y}{\partial G}\\ \frac{\partial I}{\partial G} &= h'(r)\frac{\partial r}{\partial G}\\ \frac{\partial r}{\partial G} &= 0\end{aligned}$$

Solving these equations, we obtain

$$\frac{\partial Y}{\partial T}=\frac{1}{1-f'(Y-T)}$$

Example 2.7 (Indifference curves). Suppose that there are two consumption goods and a consumer whose preferences a represented by a utility function $u(x, y)$. The indifference curves of the consumer are the sets

$$\{(x, y) \in \mathbb{R}^2 : x, y > 0, \quad u(x, y) = C\}$$

with $C \in \mathbb{R}$. Suppose that the function $u(x, y)$ is differentiable and that

$$\frac{\partial u}{\partial x} > 0 \qquad \frac{\partial u}{\partial y} > 0$$

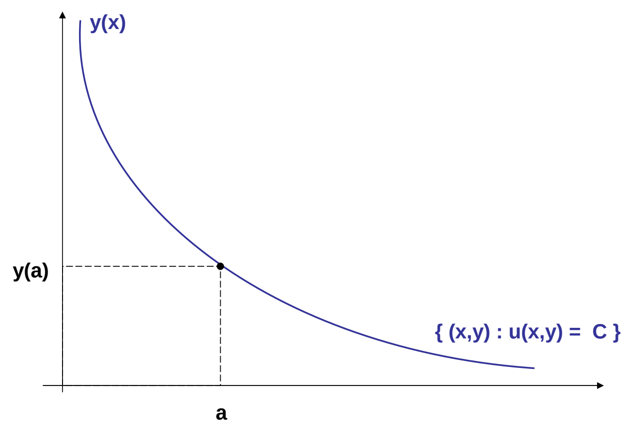
Applying the implicit function Theorem, we see that the equation

$$u(x, y) = C$$

defines y as a function of x. The set

$$\{(x, y) \in \mathbb{R}^2 : x, y > 0, \quad u(x, y) = C\}$$

may be represented as the graph of the function $y(x)$.



Differentiating implicitly, we may compute the derivative y'

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial x} y'(x) = 0$$

so that

$$y'(x) = -\frac{\partial u / \partial x}{\partial u / \partial y}$$

We see that $y(x)$ is a decreasing function. The absolute value of $y'(x)$ (that is, the absolute value of the slope of the straight line tangent to the indifference curve) is the marginal rate of substitution of the consumer. Therefore, we define the marginal rate of substitution of the consumer as

$$\mathrm{MRS}(x, y) = \frac{\partial u / \partial x}{\partial u / \partial y}(x, y)$$

Suppose that a consumer has a bundle of consumption goods $(a, b = y(a))$. Recalling the interpretation of the derivative $y'(a)$, we see that the marginal rate of substitution $\mathrm{MRS}(a, b)$ of the agent measures (approximately) the maximum amount of good y that the agent would be willing to exchange for an additional consumption of one unit of good x.

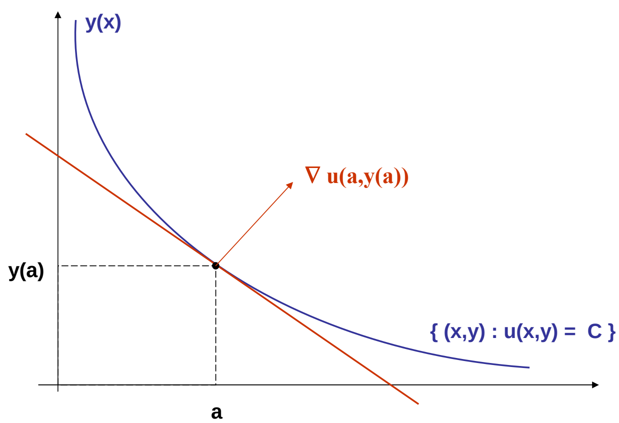
For example, if the consumer has a Cobb-Douglas utility function $u(x, y) = x^2 y^4$, the marginal rate of substitution is

$$\mathrm{MRS}(x, y) = \frac{\partial u / \partial x}{\partial u / \partial y} = \frac{2xy^4}{4x^2y^3} = \frac{y}{2x}$$

On the other hand, recall that the slope of the straight line tangent to the graph of $y(x)$ at the point $(a, y(a))$ is $y'(a)$. That is, the director vector of the straight line tangent to the graph of $y(x)$ at the point $(a, y(a))$ is the vector $(1, y'(a))$. Performing the scalar product of this vector with the gradient vector of u at the point $(a, y(a))$ we obtain that

$$(1, y'(a)) \cdot \nabla u(a, y(a)) = \left(1, -\frac{\partial u / \partial x}{\partial u / \partial y}\right) \cdot \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) = 0$$

And we have checked again that the gradient vector ∇u is perpendicular to the straight line tangent the indifference curve of the consumer.



Example 2.8. Suppose that there are two consumption goods and the agent has preferences over theses which might be represented by a utility function $u(x, y)$. Suppose the prices of the goods are p_x and p_y. Consuming the bundle (x, y) costs

$$p_x x + p_y y$$

to the agent. If his income is I then

$$p_x x + p_y y = I$$

That is, if the agent buys x units of the first good, then the maximum amount he can consume of the second good is

$$\frac{I}{p_y} - \frac{p_x}{p_y} x$$

so his utility is

$$u\left(x, \frac{I}{p_y} - \frac{p_x}{p_y} x\right) \tag{2.8}$$

In Economic Theory one assumes that the agent chooses the bundle of goods (x, y) that maximizes his utility. That is the agent maximizes the function 2.8. Differentiating implicitly with respect to x we obtain

$$\frac{\partial u}{\partial x} - \frac{\partial u}{\partial y}\frac{p_x}{p_y} = 0 \tag{2.9}$$

Thus, the first order condition is

$$\mathrm{MRS}(x, y) = \frac{p_x}{p_y}$$

The above equation together with the budget restriction

$$p_x x + p_y y = I$$

determines the demand of the agent.

For example if the preferences of the consumer may be represented by a Cobb-Douglas utility function

$$u(x, y) = x^2 y$$

the MRS is

$$\mathrm{MRS}(x, y) = \frac{2xy}{x^2} = \frac{2y}{x}$$

and the demand of the agent is determined by the system of equations

$$\begin{aligned} \frac{2y}{x} &= \frac{p_x}{p_y} \\ p_x x + p_y y &= I \end{aligned}$$

from these we obtain the demand of the agent

$$\begin{aligned} x(p_x, p_y, I) &= \frac{2I}{3p_x} \\ y(p_x, p_y, I) &= \frac{I}{3p_y} \end{aligned}$$

Example 2.9 (Isoquants and the marginal rate of technical substitution). Suppose that a firm uses the production function $Y = f(x_1, x_2)$ where (x_1, x_2) are the units of inputs used in manufacturing of$Y$ units of the product. Given a fixed level of production $\bar{y}$, the corresponding isoquant is the level set

$$\{(x_1, x_2) \in \mathbb{R}^2 : x_1, x_2 > 0, \quad f(x_1, x_2) = \bar{y}\}$$

As in the previous exercise, we see that on the isoquants we may write x_2 as a function of x_1 and that

$$x_2'(x_1) = -\frac{\partial f / \partial x_1}{\partial f / \partial x_2}$$

The marginal rate of technical substitution is defined as

$$\mathrm{RMST} = -x_2'(x_1) = \frac{\partial f / \partial x_1}{\partial f / \partial x_2}$$

For example, if the production function of the firm is $Y = x_1^{1/3} x_2^{1/2}$ then the marginal rate of technical substitution is

$$\mathrm{RMST} = \frac{\partial Y / \partial x_1}{\partial Y / \partial x_2} = \frac{\frac{1}{3} x_1^{-2/3} x_2^{1/2}}{\frac{1}{2} x_1^{1/3} x_2^{-1/2}} = \frac{2x_2}{3x_1}$$

3. Taylor's Approximations of First and Second order

Definition 3.1. Let $f \in C^1(D)$, $p \in D$. The first order Taylor polynomial at p is

$$P_1(x) = f(p) + \nabla f(p) \cdot (x - p)$$

Remark 3.2. If $f(x, y)$ is a function of two variables and $p = (a, b)$then Taylor's first order polynomial for the function f around the point $p = (a, b)$ is the polynomial

$$P_1(x, y) = f(a, b) + \frac{\partial f}{\partial x}(a, b) \cdot (x - a) + \frac{\partial f}{\partial y}(a, b) \cdot (y - b)$$

Definition 3.3. If $f \in C^2(D)$ we define the Taylor polynomial of order 2 around the point p as

$$P_2(x) = f(p)+\nabla f(p)\cdot(x-p)+\frac{1}{2}(x-p)\,\mathrm{H}\,f(p)(x-p) = P_1(x)+\frac{1}{2}(x-p)\,\mathrm{H}\,f(p)(x-p)$$

Remark 3.4. If $f(x, y)$ is a function of two variables and $p = (a, b)$ then Taylor's second order polynomial for the function f around the point $p = (a, b)$ is the polynomial

$$\begin{aligned} P_2(x, y) &= f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b) + \\ &+ \frac{1}{2}\left(\frac{\partial^2 f}{\partial x \partial x}(x - a)^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x - a)(y - b) + \frac{\partial^2 f}{\partial y \partial y}(y - b)^2\right) \end{aligned}$$

Remark 3.5. These are good approximations to $f(x)$ in the sense that if f is of class $C^1(D)$, then,

$$\lim_{x \to p} \frac{f(x) - P_1(x)}{\|x - p\|} = 0$$

and if f is of class $C^2(D)$, then,

$$\lim_{x \to p} \frac{f(x) - P_2(x)}{\|x - p\|^2} = 0$$

4. Quadratic forms

Definition 4.1. A **quadratic form** of order n is a function $Q : \mathbb{R}^n \to \mathbb{R}$ of the form

$$Q(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

for some real numbers $a_{ij} \in \mathbb{R}$ $i, j = 1, \ldots, n$

Example 4.2. $Q(x, y, z) = x^2 - 2xy + 4xz + 6yz + 5z^2$

Remark 4.3. A quadratic form can be expressed in matrix notation. For example,

$$Q(x, y, z) = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ -1 & 0 & 3 \\ 2 & 3 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 - 2xy + 4xz + 6yz + 5z^2$$

It is possible to do this in general (infinitely many) ways. For example, the previous quadratic form can also be expressed as

$$Q(x,y,z) = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 1 & -2 & 1 \\ 0 & 0 & 4 \\ 3 & 2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

(The condition is that

$$xAx^t = xBx^t$$

as long as $a_{ij} + a_{ji} = b_{ij} + b_{ji}$ for every $i, j = 1, 2, \ldots, n$)

But, there is a unique way if we require that A be symmetric.

Proposition 4.4. Every quadratic form $Q : \mathbb{R}^n \to \mathbb{R}$, can be written in a unique way $Q(x) = xAx^t$ with $A = A^t$ a symmetric matrix.

Remark 4.5. Observe that the symmetric matrix

$$A = (a_{ij})$$

is associated with the following quadratic form

$$Q(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{1 \le i < j \le n} x_i x_j$$

We will identify the quadratic form $Q(x) = xAx^t$ with the matrix A.

4.1. Classification of quadratic forms.

Definition 4.6. A quadratic form $Q : \mathbb{R}^n \to \mathbb{R}$ is

(1) **Positive definite** if $Q(x) > 0$ for every $x \in \mathbb{R}^n$, $x \neq 0$.
(2) **Negative definite** if $Q(x) < 0$ for every $x \in \mathbb{R}^n$, $x \neq 0$.
(3) **Positive semidefinite** if $Q(x) \ge 0$ for every $x \in \mathbb{R}^n$ and $Q(x) = 0$ for some $x \neq 0$.
(4) **Negative semidefinite** if $Q(x) \le 0$ for every $x \in \mathbb{R}^n$ and $Q(x) = 0$ for some $x \neq 0$.
(5) **Indefinite** if there are some $x, y \in \mathbb{R}^n$ such that $Q(x) > 0$ and $Q(y) < 0$.

Example 4.7. $Q_1(x,y,z) = x^2 + 3y^2 + z^2$ is positive definite.

Example 4.8. $Q_2(x,y,z) = -2x^2 - y^2$ is negative semidefinite.

Example 4.9. $Q_3(x,y) = -2x^2 - y^2$ is negative definite.

Example 4.10. $Q_4(x,y,z) = x^2 - y^2 + 3z^2$ is indefinite.

The previous quadratic forms are easy to classify because they are in diagonal form, i.e.

$$Q_1 \Leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad Q_2 \Leftrightarrow \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$Q_3 \Leftrightarrow \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix} \quad Q_4 \Leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Proposition 4.11. Consider the matrix

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Then, the quadratic form

$$Q(x) = xAx^t = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

es,

(1) positive definite if and only if $\lambda_i > 0$ for every $i = 1, 2, \ldots, n$;
(2) negative definite if and only if $\lambda_i < 0$ for every $i = 1, 2, \ldots, n$;
(3) positive semidefinite if and only if $\lambda_i \geq 0$ for every $i = 1, 2, \ldots, n$ and $\lambda_k = 0$ for some $k = 1, 2, \ldots, n$;
(4) negative semidefinite if and only if $\lambda_i \leq 0$ for every $i = 1, 2, \ldots, n$ and $\lambda_k = 0$ for some $k = 1, 2, \ldots, n$;
(5) indefinite if and only if there is some $\lambda_i > 0$ and some $\lambda_i < 0$.

We are going to study some methods to determine whether a quadratic form is positive/negative definite, semidefinite or indefinite. They are based on making a change of variables that puts the quadratic form in diagonal form.

Let $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{14} \\ a_{12} & a_{22} & \cdots & a_{24} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}$ be a symmetric matrix.

Let $D_1 = a_{11}$, $D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, $D_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}$, $\ldots$, $Dn = |A|$ be the leading principal minors of A. Suppose $D_1 \neq 0, D_2 \neq 0, \ldots, D_{n-1} \neq 0$. Then, there is a change of variables $Tx = z$ such that the quadratic form $Q(x) = xAx^t$ becomes

$$\tilde{Q}(z) = D_1 z_1^2 + \frac{D_2}{D_1} z_2^2 + \frac{D_3}{D_2} z_3^2 + \cdots + \frac{D_n}{D_{n-1}} z_n^2$$

This helps to remember the following.

Proposition 4.12. Let $Q(x) = xAx^t$ with A symmetric and suppose that $|A| \neq 0$. Then,

(1) A is positive definite if and only $D_i > 0$ for every $i = 1, 2, \ldots, n$;
(2) A is negative definite if and only $(-1)^i D_i > 0$ for every $i = 1, 2, \ldots, n$;
(3) if and (1) and (2) do not hold, then Q is indefinite.

The previous proposition applies when $|A| \neq 0$. What can we say if $|A| = 0$? The following is a partial answer.

Proposition 4.13. Let $Q(x) = xAx^t$ with A symmetric and suppose $D_n = 0$ and $D_1 \neq 0, D_2 \neq 0, \ldots, D_{n-1} \neq 0$. Then A is

(1) positive semidefinite if and only $D_1, D_2, \ldots, D_{n-1} > 0$;
(2) negative semidefinite if and only $D_1 < 0, D_2 > 0, \ldots, (-1)^{n-1} D_{n-1} > 0$;
(3) indefinite otherwise.

Next, we study some examples that illustrate some of thing one can do when $D_n = 0$ and some of the $D_1 \ \ldots, D_{n-1}$ vanish.

Example 4.14. Consider the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a \end{pmatrix}$$

We see that $D_1 = 1$, $D_2 = D_3 = 0$. Clearly, the eigenvalues are $\lambda_1 = 1$, $\lambda_1 = 0$, $\lambda_1 = a$. So, the associated quadratic form is positive semidefinite if and only if $a \geq 0$ and indefinite if and only if $a < 0$. But, this is impossible to tell from $D_1 = 1$, $D_2 = D_3 = 0$.

Remark that in the previous example, if we exchange the variables y and z then the associated matrix becomes

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and then, the above propositions apply. This is formalized in the following observation.

Definition 4.15. A principal minor is centered if includes the **same rows and columns**. For example, the minor

$$\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}$$

is a centered minor, because it includes rows and columns 1, 3. But, the minor,

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

is NOT centered, because it includes rows 1, 3 and columns 1, 2.

Proposition 4.16. Proposition 4.13 still holds if we replace the chain of leading principal minors by any other chain consisting of principal centered minors.

Remark 4.17. The methods above are especially useful for symmetric 2×2 matrices. For example if A is 2×2 matrix and $|A| < 0$, then the associated quadratic form is indefinite. Why?

5. Concavity and convexity

In this section, we assume that: $D \subset \mathbb{R}^n$ is a convex, open set.

Definition 5.1. We say that

(1) f is **concave** on D if for every $\lambda \in [0, 1]$ and $x, y \in D$ we have that

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda) f(y)$$

(2) f is **convex** on D if for every $\lambda \in [0, 1]$ and $x, y \in D$ we have that

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y)$$

Remark 5.2. Note that f is convex on D if and only if $-f$ is concave on D.

Definition 5.3. We say that

(1) f is **strictly concave** on D if for every $\lambda \in (0,1)$ and $x, y \in D$, $x \neq y$ we have that
$$f(\lambda x + (1-\lambda)y) > \lambda f(x) + (1-\lambda) f(y)$$
(2) f is **strictly convex** on D if for every $\lambda \in (0,1)$ and $x, y \in D$, $x \neq y$ we have that
$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y)$$

Remark 5.4. Note that f is strictly convex on D if and only if $-f$ is strictly concave on D.

Proposition 5.5. Let D be a convex, open subset of $\mathbb{R}^n$. Then,

(1) f is concave $\Leftrightarrow$ the set $\{(x,y) : x \in D, y \leq f(x)\}$ is convex.
(2) f is convex $\Leftrightarrow$ the set $\{(x,y) : x \in D, y \geq f(x)\}$ is convex.
(3) f is strictly convex $\Leftrightarrow$ the set $\{(x,y) : x \in D, y \geq f(x)\}$ is convex and the graph of f contains no segments.
(4) f is strictly concave $\Leftrightarrow$ the set $\{(x,y) : x \in D, y \leq f(x)\}$ is convex and the graph of f no contains segments.
(5) If f is convex, then the lower contour set $\{x \in D : f(x) \leq \alpha\}$ is convex for every $\alpha \in \mathbb{R}$
(6) If f is concave, then the upper contour set $\{x \in D : f(x) \geq \alpha\}$ is convex for every $\alpha \in \mathbb{R}$

Example 5.6. $f(x,y) = x^2 + y^2$ is strictly convex.

Example 5.7. $f(x,y) = (x-y)^2$ is convex, but not strictly convex.

Remark 5.8. The conditions in (5) and (6) are necessary but not sufficient. For example, any monotone function $f : \mathbb{R} \to \mathbb{R}$ satisfies that both sets
$$\{x \in D : f(x) \leq \alpha\} \quad \text{and} \quad \{x \in D : f(x) \geq \alpha\}$$
are convex.

6. First order conditions for concavity and convexity

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is concave and differentiable on a convex set D. Then, the plane tangent to the graph of f at $p \in D$ is above the graph of f. Recall that the tangent plane is the set of points $(x_1, \ldots, x_n, x_{n+1}) \in \mathbb{R}^{n+1}$ that satisfy the equation,
$$x_{n+1} = f(p) + \nabla f(p) \cdot (x - p)$$
Hence, if is concave and differentiable on D, then we have that,
$$f(x) \leq f(p) + \nabla f(p) \cdot (x - p)$$
for every $x \in D$.

Proposition 6.1. Suppose $f \in C^1(D)$. Then,

(1) f is concave on D if and only if for all $u, v \in D$ we have that
$$f(u) \leq f(v) + \nabla f(v) \cdot (u - v)$$
(2) f is strictly concave on D if and only if for all $u, v \in D$, $u \neq v$, we have that
$$f(u) < f(v) + \nabla f(v) \cdot (u - v)$$

(3) f is convex on D if and only if for all $u, v \in D$ we have that
$$f(u) \geq f(v) + \nabla f(v) \cdot (u - v)$$
(4) f is strictly convex on D if and only if for all $u, v \in D$, $u \neq v$, we have that
$$f(u) > f(v) + \nabla f(v) \cdot (u - v)$$

7. Second order conditions for concavity and convexity

Proposition 7.1. Let $D \subset \mathbb{R}^n$ be open and convex. Let $f \in C^2(D)$ Let $\mathrm{H}\, f(p)$ be the Hessian matrix of f at p. Then,

(1) f is concave on D if and only if for every $p \in D$, $\mathrm{H}\, f(p)$ is negative semidefinite or negative definite. That is, f is concave on D if and only if for every $p \in D$ and $x \in \mathbb{R}^n$ we have that $x \cdot \mathrm{H}\, f(p)x \leq 0$.
(2) f is convex on D if and only if for every $p \in D$, $\mathrm{H}\, f(p)$ is positive semidefinite or positive definite. That is, f is convex on D if and only if for every $p \in D$ and $x \in \mathbb{R}^n$ we have that $x \cdot \mathrm{H}\, f(p)x \geq 0$.
(3) If $\mathrm{H}\, f(p)$ is definite negative for every $p \in D$, then f is strictly concave on D.
(4) If $\mathrm{H}\, f(p)$ is positive negative for every $p \in D$, then f is strictly convex on D.

Remark 7.2. One can show that if f is strictly convex, then $\mathrm{H}\, f(x, y)$ is positive definite except on a "small" set. For example, $f(x, y) = x^4 + y^4$ is strictly convex and
$$\mathrm{H}\, f(x, y) = \begin{pmatrix} 12x^2 & 0 \\ 0 & 12y^2 \end{pmatrix}$$
is positive definite if $xy \neq 0$, that is, it is positive definite on all of $\mathbb{R}^2$ except on the two axis $\{(x, y) \in \mathbb{R}^2 : xy = 0\}$. For points on the two axis (that is for points $(x, y) \in \mathbb{R}^2$ such that $xy = 0$) the Hessian matrix is positive semidefinite.

8. Applications to convex sets

Proposition 8.1. If $X_1, \ldots, X_k$ are convex subsets of $\mathbb{R}^n$, then $X_1 \cap X_2 \cap \cdots \cap X_k$ is also a convex subset.

Example 8.2. Using the theory of this chapter, prove that the set $\{(x, y) \in \mathbb{R}^2 : 3x^2 + 10y^2 \leq 10, x \geq 0, y \leq 0\}$ is convex.

Example 8.3 (Concavity, convexity and preferences). In example 2.7 we considered a consumer whose preferences (over two consumption goods) are represented by the utility function $u(x, y)$. The indifference curves of the consumer are the sets
$$\{(x, y) \in \mathbb{R}^2 : x, y > 0, \quad u(x, y) = C\}$$
with $C \in \mathbb{R}$. Suppose that the function $u(x, y)$ is differentiable and that
$$\frac{\partial u}{\partial x} > 0 \qquad \frac{\partial u}{\partial y} > 0$$
Applying the implicit function Theorem, we see that the equation
$$u(x, y) = C$$
defines y as a function of x. The set
$$\{(x, y) \in \mathbb{R}^2 : x, y > 0, \quad u(x, y) = C\}$$

may be represented as the graph of the function $y(x)$. Differentiation implicitly the equation $u(x,y) = C$ we may compute the derivative y'

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} y'(x) = 0$$

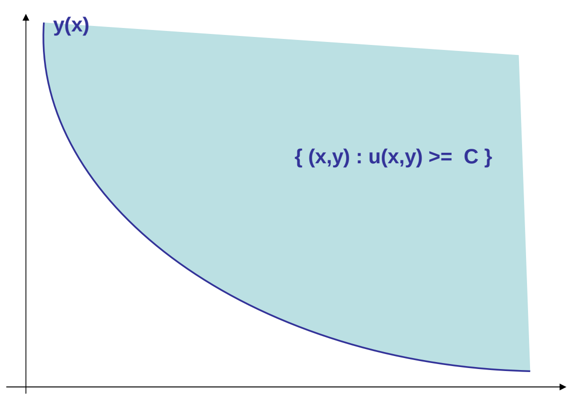
Applying again the implicit function Theorem we obtain an equation for the second derivative y''

$$\frac{\partial^2 u}{\partial x \partial x} + 2\frac{\partial^2 u}{\partial x \partial y} y'(x) + \frac{\partial^2 u}{\partial y \partial y} \left(y'(x)\right)^2 + \frac{\partial u}{\partial x} y''(x) = 0 \tag{8.1}$$

One of l standard assumptions in Economic Theory is that the set which consist of all the consumption bundles which are preferred to a given consumption bundle is a convex set. In terms of the utility function this means that the set

$$\{(x,y) \in \mathbb{R}^2 : x, y > 0, \quad u(x,y) > C\}$$

is convex.



By Proposition 5.5, the set $\{(x,y) \in \mathbb{R}^2 : x, y > 0, \quad u(x,y) > C\}$ is convex if we assume that the function $u(x,y)$ is concave[2]. Suppose that the function $u(x,y)$ is concave and of class C^2. According to the definition of concavity, this means that for every $h, k \in \mathbb{R}$ we have that

$$\frac{\partial^2 u}{\partial x \partial x} h^2 + 2\frac{\partial^2 u}{\partial x \partial y} hk + \frac{\partial^2 u}{\partial y \partial y} k^2 \leq 0$$

If in this equation we plug in $h = 1$, $k = y'(x)$ we obtain that

$$\frac{\partial^2 u}{\partial x \partial x} + 2\frac{\partial^2 u}{\partial x \partial y} y'(x) + \frac{\partial^2 u}{\partial y \partial y} \left(y'(x)\right)^2 \leq 0$$

and solving for y'' in the equation 8.1 we obtain

$$y''(x) = -\frac{\frac{\partial^2 u}{\partial x \partial x} + 2\frac{\partial^2 u}{\partial x \partial y} y'(x) + \frac{\partial^2 u}{\partial y \partial y} \left(y'(x)\right)^2}{\partial u / \partial x} \geq 0$$

[2]But, that the set $\{(x,y) \in \mathbb{R}^2 : x, y > 0, \quad u(x,y) \geq C\}$ is convex does not necessarily imply that the function u is concave

that is, the function $y(x)$ is convex, so that $y'(x)$ is increasing. Since $\mathrm{MRS}(x, y(x)) = -y'(x)$, we see that if the preferences of the consumer are convex his marginal rate of substitution is decreasing.