Calculus Review

Finn Christensen *
Department of Economics
Towson University

Juergen Jung†
Department of Economics
Towson University

February 12, 2014

Abstract

This is a calculus review based on Hands (1991), Simon and Blume (1994), Dadkhah (2011), Williamson (2013) and Finn Christensen's lecture notes.

*Department of Economics, Towson University, U.S.A., *E-mail:* fchristensen@towson.edu
†Department of Economics, Towson University, U.S.A., *E-mail:* jjung@towson.edu

1 Introduction

The key uses of derivatives in Economics can be summarized as:

- As the mathematical synonym to the word marginal.
- To identify extreme values of functions.
- To perform comparative statics.

Marginal is a synonym for extra or additional in economics. The derivative can be interpreted as the rate of change in the value of a function. These two concepts are tightly related.

1.1 What is a derivative

A derivative measures the ratio of change in y to a change in x so that

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{\Delta y}{\Delta x},$$

as depicted in figure 1.

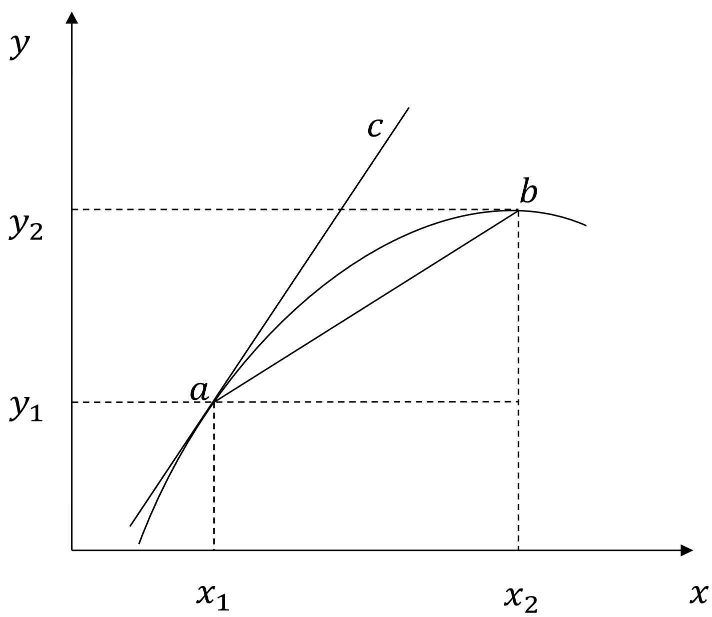

Figure 1: Line ac is the tangent line at point a. Line ab is the secant line.

If we now move point x_2 closer and closer to point x_1, point b slides down along the curve and gets closer and closer to point a. In the limit, when the distance between x_2 and x_1 approaches zero, point b moves to point a and the line ab coincides with line ac, which is the tangent to the curve at point a. If this limit exists, we will have a derivate of y with respoect to x

$$\frac{dy}{dx} = y'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

1.2 Rules of differentiation

	Rules		Examples	
	$y(x)$	$y'(x) = \frac{dy}{dx}$	$y(x)$	$y'(x) = \frac{dy}{dx}$
Power rule	ax^n	$n \times ax^{n-1}$	$6x^{1/3}$	$2x^{-2/3}$
Chain rule	$f(g(x))$	$\frac{df}{dg}\frac{dg}{dx}$	$(12-2x)^{\frac{2}{3}}$	$\frac{2}{3}(12-2x)^{\frac{-1}{3}}(-2) = -\frac{4}{3}(12-2x)^{\frac{1}{3}}$
Product rule	$f(x) \times g(x)$	$\frac{df}{dx}g(x) + \frac{dg}{dx}f(x)$	$6x^{\frac{1}{3}}(12-2x)^{\frac{2}{3}}$	$2x^{\frac{-2}{3}}(12-2x)^{\frac{2}{3}} - \frac{4}{3}(12-2x)^{\frac{-1}{3}}6x^{\frac{1}{3}}$
Division rule	$\frac{f(x)}{g(x)}$	$\frac{\frac{df}{dx}g(x) - \frac{dg}{dx}f(x)}{g(x)^2}$	$\frac{6x+1}{4x^2}$	$\frac{6\times 4x^2 - (6x+1)\times 8x}{16x^4} = \frac{-24x-8}{16x^3}$
Exponent rule	a^x	$a^x \times ln(a)$	3^x	$3^x ln(3)$
Logarithm	$ln(x)$	$\frac{dy}{dx} = \frac{1}{x}$	$ln(x^2)$	$\frac{1}{x^2} \times 2x$
Exponential	$exp(x) = e^x$	e^x	e^{x^2}	$e^{x^2} \times 2x$

1.3 Second- and higher-order derivatives

The derivative $y'(x)$ of a function $y(x)$ is itself a function. As such function $y'(x)$ may also be differentiable which we denote as

$$y''(x) = \frac{d^2y}{dx^2}.$$

Example: Let $y(x) = 5x^6 - 3e^{2x}$, then

$$\begin{aligned} y' &= \frac{dy}{dx} = 30x^5 - 6e^{2x}, \text{ and} \\ y'' &= \frac{d^2y}{dx^2} = 150x^4 - 12e^{2x}, \text{ and} \\ y''' &= \frac{d^3y}{dx^3} = 600x^3 - 24e^{2x}, \ldots \end{aligned}$$

1.4 Differential

So far we have treated dy/dx as one entity, so that we have applied the operator d/dx to a function

$$y = f(x)$$

to obtain its derivative. We can separte the two and call dy the *differential* of y which measures the change in y as a result of an infinitesimal change in x. We write this as

$$dy = f'(x)\,dx.$$

We next illustrate this with a few examples.

Example: To find the differential of the function

$$y = 3(x-9)^2$$

we note that $f'(x) = 6(x-9)$. Therefore the differential is

$$dy = f'(x)\,dx = 6(x-9)dx.$$

So the rules of differential from above apply equally to differentials.

1.5 Higher order differential

Similar to derivates, we can compute second- and higer-order differentials as well:

$$\begin{aligned} d^2y &= f''(x)\,dx^2, \\ &\quad ... \\ d^ny &= f^n(x)\,dx^n. \end{aligned}$$

1.6 Economic examples

- Suppose the cost of producing skirts is given by $C(q) = 100 + 5q$. It is easy to see that the marginal cost of another skirt is \$5. But also notice that $\frac{dC}{dq} = 5$.
- Now suppose the cost of producing q skirts is given by $C(q) = 100 + q^2$. What is the marginal cost of producing another skirt? $MC = \frac{dC}{dq} = 2q \rightarrow$ increasing marginal cost. If the quantity of skirts increases, the marginal cost of skirts increases linearly.
- Joe's utility of eating pizza slices is $u(z) = z^{1/2}$. What is his marginal utility of another slice?

$$MU_z = \frac{du}{dz} = \frac{1}{2}z^{-1/2} = \frac{1}{2}\frac{1}{\sqrt{z}}$$

  This illustrates diminishing marginal utility. As the quantity of pizzas consumed increases, the marginal utility decreases.

Exercise 1: Joanne's utility of consuming good x is given by $u(x) = 6x^{1/3}(12-2x)^{2/3}$. What is the marginal utility of good x?

Exercise 2: Bob's utility of consuming goods x and y is $u(x,y) = x^{1/4}y^{3/4}$. What is the marginal utility of good x? What is the marginal utility of good y?

2 Differentiation of functions of several variables

In economics many functions have more than one variable. A typical utility function is a function of more than just one consumption good

$$U(c_1, c_2) = c_1^{0.1} + c_2^{0.4}.$$

A typical production function is a function of capital K and labor L

$$Y = F(K, L) = AK^{0.5}L^{0.5}.$$

2.1 Partial derivatives

The idea of differentiation of a function of one variable can be extended to functions of several variables. Consider the general function

$$y = f(x_1, x_2, ..., x_n),$$

then the *partial derivative* of y with respect to x_1 is the derivative of y with respect to x_1 when x_2 remains constant and, therefore, can be treated as such. The same is true for the partial derivative with respect to x_2. In this case x_1 is treated like a constant. The partial derivatives are written as

$$\begin{aligned} \frac{\partial y}{\partial x_1} &= f_1(x_1, x_2, ..., x_n), \\ \frac{\partial y}{\partial x_2} &= f_2(x_1, x_2, ..., x_n), \\ &\vdots \\ \frac{\partial y}{\partial x_n} &= f_n(x_1, x_2, ..., x_n). \end{aligned}$$

We can "collect" these partial derivatives in a vector called the gradient of function f which we denote as:

$$\nabla f(x_1, x_2, ...x_n) = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix}.$$

Example: Consider $y = f(x_1, x_2) = 3x_1^2 - 6x_1x_2 + 4x_2$, then the partial derivatives are

$$\begin{aligned} \frac{\partial y}{\partial x_1} &= 6x_1 - 6x_2, \\ \frac{\partial y}{\partial x_2} &= -6x_1 + 4, \end{aligned}$$

so that the gradient becomes

$$\nabla f(x_1, x_2) = \begin{bmatrix} 6x_1 - 6x_2 \\ -6x_1 + 4 \end{bmatrix}.$$

2.2 Economic examples

Example 1: The utility function $U(x_1, x_2, ..., x_n)$ measures total satisfaction from consuming a certain consumption bundle that can be made up of multiple goods where x_1 denotes the quantity of good 1, x_2 denotes the quantity of good 2 etc. The *marginal utility* of good i is the additional satisfaction derived from the consumption of an incremental amount of that good, or

$$U_i = \frac{\partial U}{\partial x_i}.$$

Marginal utilities are assumed to be positive:

$$U_i > 0,\ i = 1, ..., n$$

Example 2: Consider the production $Y = F(K, L) = AK^{0.5}L^{0.5}$. The marginal product of capital K measures the additional output that can be generated by a small increment in the amount of capital used. This additional output can be expressed as

$$F_K = \frac{\partial F}{\partial K} = 0.5AK^{-0.5}L^{0.5} = 0.5A\left(\frac{L}{K}\right)^{0.5}.$$

2.3 Second-order partial derivatives

In the same manner that we took second derivatives of functions of one variable we can take second derivatives of multivariate functions. Note however that with multivariate functions the number of second derivatives is larger. In the case of a function of two variables

$$y = f(x_1, x_2)$$

we end up with four second derivates:

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} \frac{\partial^2 y}{\partial x_1 \partial x_1} = \frac{\partial^2 y}{\partial x_1^2} = f_{11}(x_1, x_2) & \frac{\partial^2 y}{\partial x_1 \partial x_2} = f_{12}(x_1, x_2) \\ \frac{\partial^2 y}{\partial x_2 \partial x_1} = f_{21}(x_1, x_2) & \frac{\partial^2 y}{\partial x_2 \partial x_2} = \frac{\partial^2 y}{\partial x_2^2} = f_{22}(x_1, x_2) \end{bmatrix}.$$

The terms in the off-diagonal are called cross derivatives. This Matrix of second partial derivatives is called the *Hessian* matrix. For a more general function

$$y = f(x_1, x_2, ..., x_n)$$

we have the following Hessian matrix:

$$\nabla^2 f(x_1, x_2, ..., x_n) = \begin{bmatrix} \frac{\partial^2 y}{\partial x_1 \partial x_1} = f_{11} & \frac{\partial^2 y}{\partial x_1 \partial x_2} = f_{12} & \cdots & \frac{\partial^2 y}{\partial x_1 \partial x_n} = f_{1n} \\ \frac{\partial^2 y}{\partial x_2 \partial x_1} = f_{21} & \frac{\partial^2 y}{\partial x_2 \partial x_2} = f_{22} & \cdots & \frac{\partial^2 y}{\partial x_2 \partial x_n} = f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 y}{\partial x_n \partial x_1} = f_{n1} & \frac{\partial^2 y}{\partial x_n \partial x_2} = f_{n2} & \cdots & \frac{\partial^2 y}{\partial x_n \partial x_n} = f_{nn} \end{bmatrix}.$$

Example : Consider again function $y = f(x_1, x_2) = 3x_1^2 - 6x_1x_2 + 4x_2$ from the previous example. The first and second derivatives are:

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 6x_1 - 6x_2 \\ -6x_1 + 4 \end{bmatrix},$$

and

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} \frac{\partial^2 y}{\partial x_1 \partial x_1} & \frac{\partial^2 y}{\partial x_1 \partial x_2} \\ \frac{\partial^2 y}{\partial x_2 \partial x_1} & \frac{\partial^2 y}{\partial x_2 \partial x_2} \end{bmatrix} = \begin{bmatrix} 6 & -6 \\ -6 & 0 \end{bmatrix}.$$

2.4 Differentials of functions with multiple variables

The differential

$$dy = f'(x)\,dx$$

measures how y changes due to an infinitesimal change in x. With multivariate functions $y = f(x_1, x_2, ..., x_n)$ the differential measures the change in the dependent variable y when all independent variables $x_1, x_2, ..., x_n$ change infinitesimally and simultaneously

$$dy = \overbrace{\begin{bmatrix} \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial x_2} & \cdots & \frac{\partial y}{\partial x_n} \end{bmatrix}}^{\nabla f(x_1,x_2)'} \begin{bmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{bmatrix} = \frac{\partial y}{\partial x_1}dx_1 + \frac{\partial y}{\partial x_2}dx_2 + ... + \frac{\partial y}{\partial x_n}dx_n. \tag{1}$$

Example 1: Consider function $y = x_1^2 x_2^2$. The differential is

$$dy = 2x_1x_2^2dx_1 + 2x_1^2x_2dx_2.$$

2.5 Total derivative

The differential measures the change in the dependent variable when all independent variables change. Now suppose that the change in the independent variables is caused by a change in another variable. More specifically let all independent variables x be a function of another variable t (e.g. time) so that

$$x_1 = x_1(t), x_2 = x_2(t), ..., x_n = x_n(t).$$

This means that

$$\begin{aligned}
dx_1 &= x_1'(t)\,dt,\\
dx_2 &= x_2'(t)\,dt,\\
\\
dx_n &= x_n'(t)\,dt.
\end{aligned}$$

Substituting these expressions into (1) we get

$$dy = \frac{\partial y}{\partial x_1}x_1'(t)\,dt + \frac{\partial y}{\partial x_2}x_2'(t)\,dt + ... + \frac{\partial y}{\partial x_n}x_n'(t)\,dt,$$

and dividing by dt we then have

$$\frac{dy}{dt} = \frac{\partial y}{\partial x_1}x_1'(t) + \frac{\partial y}{\partial x_2}x_2'(t) + ... + \frac{\partial y}{\partial x_n}x_n'(t)\,.$$

The total derivative measures the rato of change in the dependent varaible to an infinitesimal change in the ultimate independent variable t. It is called total derivative as opposed to partial derivative. The latter measures the ratio of the change in y to an infinitesimal change in one of the indpendent variables x. Variable t could be time in which case dy/dt signifies the instantaneous change in y as all independent variables $x_1(t), x_2(t), ..., x_n(t)$ change with time t and thereby affect y.

3 Linear equation systems

A general system of linear equations can be written as

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2, \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n = b_n, \end{cases}$$

which can be written in matrix notation as

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

or

$$Ax = b.$$

3.1 Cramer's rule

Here is a convenient rule for solving linear systems of equations of the form

$$Ax = b.$$

In order to find the component x_i of the solution vector x to this system of linear equations, replace the i^{th} column of the matrix A with the column vector b to form a matrix A_i. Then x_i is the determinant of A_i divided by the determinant of A :

$$x_i = \frac{|A_i|}{|A|}.$$

Example 1: Consider the following system of equations

$$\begin{cases} 6x - 7y = 5, \\ 5x + 3y = 13. \end{cases}$$

In matrix form this is

$$\begin{bmatrix} 6 & -7 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 13 \end{bmatrix}.$$

The determinants are then

$$|A| = \begin{vmatrix} 6 & -7 \\ 5 & 3 \end{vmatrix} = 53; \ |A_1| = \begin{vmatrix} 5 & -7 \\ 13 & 3 \end{vmatrix} = 106; \ |A_2| = \begin{vmatrix} 6 & 5 \\ 5 & 13 \end{vmatrix} = 53,$$

so that

$$x = \frac{|A_1|}{|A|} = 2 \text{ and } y = \frac{|A_2|}{|A|} = 1.$$

4 The derivative and comparative statics

- Comparative statics is the process of evaluating the impact of a change in the value of an exogenous variable on an endogenous variable.
- The value of an **endogenous variable** is determined by the economic model.
- The value of an **exogenous variable** is taken as given.

In a supply and demand model, the equilibrium price and quantity are endogenous variables while consumer income, consumer tastes, and producer costs (and everything else) are exogenous variables.

Examples: We already saw some simple examples when we discussed the interpretation of the derivative as marginal. More examples will follow in the next lecture.

4.1 Differentiating systems of equations

Consider the system of equations

$$\begin{cases} f_1\left(y_1, ..., y_m, x_1, ..., x_n\right) = 0, \\ \vdots \\ f_m\left(y_1, ..., y_m, x_1, ..., x_n\right) = 0. \end{cases}$$

Here we have m endogenous variables y_i, $i = 1, ..., m$ for which we can in principle solve for in terms of the exogenous variables x_i, $i = 1, ..., n$. Even if we cannot find explicit solutions for the endogeneous variables, we may still be able to analyze how the endogenous variables y change with respect to small changes in the exogenous variables x.

If we are for instance interested in how the endogenous variables react to changes in the exogenous variables x we can start by taking the the total derivate of this system w.r.t. all variables (endogenous y and exogenous x):

$$\begin{cases} \frac{\partial f_1}{\partial y_1} dy_1 + \frac{\partial f_1}{\partial y_2} dy_2 + ... + \frac{\partial f_1}{\partial y_m} dy_m + \frac{\partial f_1}{\partial x_1} dx_1 + ... + \frac{\partial f_1}{\partial x_n} dx_n = 0, \\ \vdots \\ \frac{\partial f_m}{\partial y_1} dy_1 + \frac{\partial f_m}{\partial y_2} dy_2 + ... + \frac{\partial f_m}{\partial y_m} dy_m + \frac{\partial f_m}{\partial x_1} dx_1 + ... + \frac{\partial f_m}{\partial x_n} dx_n = 0. \end{cases}$$

This can be collected as

$$\begin{cases} \frac{\partial f_1}{\partial y_1} dy_1 + \frac{\partial f_1}{\partial y_2} dy_2 + ... + \frac{\partial f_1}{\partial y_m} dy_m + \frac{\partial f_1}{\partial x_1} dx_1 + ... + \frac{\partial f_1}{\partial x_n} dx_n = 0, \\ \vdots \\ \frac{\partial f_m}{\partial y_1} dy_1 + \frac{\partial f_m}{\partial y_2} dy_2 + ... + \frac{\partial f_m}{\partial y_m} dy_m + \frac{\partial f_m}{\partial x_1} dx_1 + ... + \frac{\partial f_m}{\partial x_n} dx_n = 0. \end{cases}$$

If we are now interested in how the dependent variables are affected by independent variable x_i holding all other independent variables constant (i.e. $dx_1 = dx_2 = ... = dx_{i-1} = dx_{i+1} = ... =$

$dx_n = 0$) we have

$$\begin{cases} \frac{\partial f_1}{\partial y_1} dy_1 + \frac{\partial f_1}{\partial y_2} dy_2 + ... + \frac{\partial f_1}{\partial y_m} dy_m = -\frac{\partial f_1}{\partial x_i} dx_i, \\ \vdots \\ \frac{\partial f_m}{\partial y_1} dy_1 + \frac{\partial f_m}{\partial y_2} dy_2 + ... + \frac{\partial f_m}{\partial y_m} dy_m = -\frac{\partial f_m}{\partial x_i} dx_i. \end{cases}$$

Then dividing by dx_i and collecting the system in matrix notation we get

$$\begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial y_1} & \frac{\partial f_m}{\partial y_2} & \cdots & \frac{\partial f_m}{\partial y_m} \end{bmatrix} \begin{bmatrix} \frac{dy_1}{dx_i} \\ \frac{dy_2}{dx_i} \\ \vdots \\ \frac{dy_m}{dx_i} \end{bmatrix} = \begin{bmatrix} -\frac{\partial f_1}{\partial x_i} \\ \vdots \\ -\frac{\partial f_m}{\partial x_i} \end{bmatrix}.$$

This is a linear equation system of the form $Ax = b$ that we can now solve with Cramer's rule from above. We then get

$$\frac{dy_1}{dx_i} = \frac{\begin{vmatrix} \boxed{\begin{matrix} -\frac{\partial f_1}{\partial x_i} \\ \vdots \\ -\frac{\partial f_m}{\partial x_i} \end{matrix}} & \begin{matrix} \cdots \\ \ddots \\ \cdots \end{matrix} & \begin{matrix} \frac{\partial f_1}{\partial y_m} \\ \vdots \\ \frac{\partial f_m}{\partial y_m} \end{matrix} \end{vmatrix}}{\begin{vmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_m} \end{vmatrix}}; \frac{dy_2}{dx_i} = \frac{\begin{vmatrix} \begin{matrix} \frac{\partial f_1}{\partial y_1} \\ \vdots \\ \frac{\partial f_m}{\partial y_1} \end{matrix} & \boxed{\begin{matrix} -\frac{\partial f_1}{\partial x_i} \\ \vdots \\ -\frac{\partial f_m}{\partial x_i} \end{matrix}} & \begin{matrix} \cdots \\ \ddots \\ \cdots \end{matrix} & \begin{matrix} \frac{\partial f_1}{\partial y_m} \\ \vdots \\ \frac{\partial f_m}{\partial y_m} \end{matrix} \end{vmatrix}}{\begin{vmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_m} \end{vmatrix}}, ..., \frac{dy_m}{dx_i} = \frac{\begin{vmatrix} \begin{matrix} \frac{\partial f_1}{\partial y_1} \\ \vdots \\ \frac{\partial f_m}{\partial y_1} \end{matrix} & \begin{matrix} \cdots \\ \ddots \\ \cdots \end{matrix} & \boxed{\begin{matrix} -\frac{\partial f_1}{\partial x_i} \\ \vdots \\ -\frac{\partial f_m}{\partial x_i} \end{matrix}} \end{vmatrix}}{\begin{vmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_m} \end{vmatrix}}.$$

Example: Consider the system of equations

$$\begin{cases} y_1 + y_2 - 2x_1 + 4x_2 = 0, \\ y_1 - y_2 + 5 - 2x_2 = 0. \end{cases}$$

We are interested in the effect of exogenous variable x_1 on the endogenous variables y_1 and y_2. We start by totally differentiating this system:

$$\begin{cases} 1dy_1 + 1dy_2 - 2dx_1 + 4dx_2 = 0, \\ 1dy_1 - 1dy_2 - 2dx_2 = 0. \end{cases}$$

We then hold all other exogenous variables constant (i.e. $dx_2 = 0$) and divide by dx_1 :

$$\begin{cases} 1\frac{dy_1}{dx_1} + 1\frac{dy_2}{dx_1} = 2, \\ 1\frac{dy_1}{dx_1} - 1\frac{dy_2}{dx_1} = 0. \end{cases}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{dy_1}{dx_1} \\ \frac{dy_2}{dx_1} \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

We can now solve for the partial derivatives using Cramer's rule to get

$$\frac{dy_1}{dx_1} = \frac{\begin{vmatrix} 2 & 1 \\ 0 & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}} = \frac{-2}{-2} = 1; \ \frac{dy_2}{dx_1} = \frac{\begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}} = \frac{-2}{-2} = 1.$$

4.2 Economic example

Consider the following system of equations

$$\begin{cases} w\left(h-l\right)+\pi-T-C=0, \\ U_2\left(C,l\right)-wU_1\left(C,l\right)=0, \end{cases}$$

where w is the exogenous wage rate, h is the maximum amount of time a consumer has available, π is divident income, T is a lump sum tax, C is consumption and $U\left(C,l\right)$ is a twice differentiable utility function in consumption C and leisure l. This system is a typical system of first order conditions that define a maximum (see next section).

We are interested in measuring how the exogenous wage rate w affects optimal consumption C and leisure l. We start by totally differentiating the system:

$$\begin{cases} -1dC-wdl+\left(h-l\right)dw+1d\pi-1dT=0, \\ U_{12}\left(C,l\right)dC+U_{22}\left(C,l\right)dl-wU_{11}\left(C,l\right)dC-wU_{12}\left(C,l\right)dl-U_{12}\left(C,l\right)dw=0, \end{cases}$$

which we can collect as

$$\begin{cases} -dC-wdl+\left(h-l\right)dw+d\pi-dT=0, \\ \left(U_{12}-wU_{11}\right)dC+\left(U_{22}-wU_{12}\right)dl-U_{12}dw=0. \end{cases}$$

We next collect the endogenous variables on the LHS and the exogenous variables on the RHS

$$\begin{cases} -dC-wdl=-\left(h-l\right)dw-d\pi+dT, \\ \left(U_{12}-wU_{11}\right)dC+\left(U_{22}-wU_{12}\right)dl=U_{12}dw. \end{cases}$$

In matrix notation this is

$$\begin{bmatrix} -1 & -1 \\ \left(U_{12}-wU_{11}\right) & \left(U_{22}-wU_{12}\right) \end{bmatrix} \begin{bmatrix} dC \\ dl \end{bmatrix} = \begin{bmatrix} -\left(h-l\right) & -1 & 1 \\ U_{12} & 0 & 0 \end{bmatrix} \begin{bmatrix} dw \\ d\pi \\ dT \end{bmatrix}.$$

If we are interested in how w affects C and l we hold all other exogenous variables constant, that is $d\pi=dT=0$. After dividing by dw the system reduces to

$$\begin{bmatrix} -1 & -1 \\ \left(U_{12}-wU_{11}\right) & \left(U_{22}-wU_{12}\right) \end{bmatrix} \begin{bmatrix} \frac{dC}{dw} \\ \frac{dl}{dw} \end{bmatrix} = \begin{bmatrix} -\left(h-l\right) \\ U_{12} \end{bmatrix},$$

which we can solve with Cramer's rule so that

$$\frac{dC}{dw} = \frac{\begin{vmatrix} -(h-l) & -1 \\ U_{12} & (U_{22}-wU_{12}) \end{vmatrix}}{\begin{vmatrix} -1 & -1 \\ (U_{12}-wU_{11}) & (U_{22}-wU_{12}) \end{vmatrix}} = \frac{-(h-l)(U_{22}-wU_{12})+U_{12}}{-(U_{22}-wU_{12})+(U_{12}-wU_{11})},$$

$$\frac{dl}{dw} = \frac{\begin{vmatrix} -1 & -(h-l) \\ (U_{12}-wU_{11}) & U_{12} \end{vmatrix}}{\begin{vmatrix} -1 & -1 \\ (U_{12}-wU_{11}) & (U_{22}-wU_{12}) \end{vmatrix}} = \frac{-U_{12}+(h-l)(U_{12}-wU_{11})}{-(U_{22}-wU_{12})+(U_{12}-wU_{11})}.$$

Whether $\frac{dC}{dw}$ and $\frac{dl}{dw}$ are positive or negative now depends on the functional form of the utility function. If you assume that $U_1 > 0, U_2 > 0, U_{11} < 0, U_{22} < 0$, and $U_{12} > 0$ (which are standard assumptions) you will be able to determine the direction of the effect of w.

Exercise: Calculate the marginal effect of dividend income on consumption and leisure, i.e. $\frac{dC}{d\pi}$ and $\frac{dl}{d\pi}$.

5 Unconstrained optimization

5.1 Extrema: minima and maxima

Figure 2 shows a continuous function with multiple local minima and maxima.

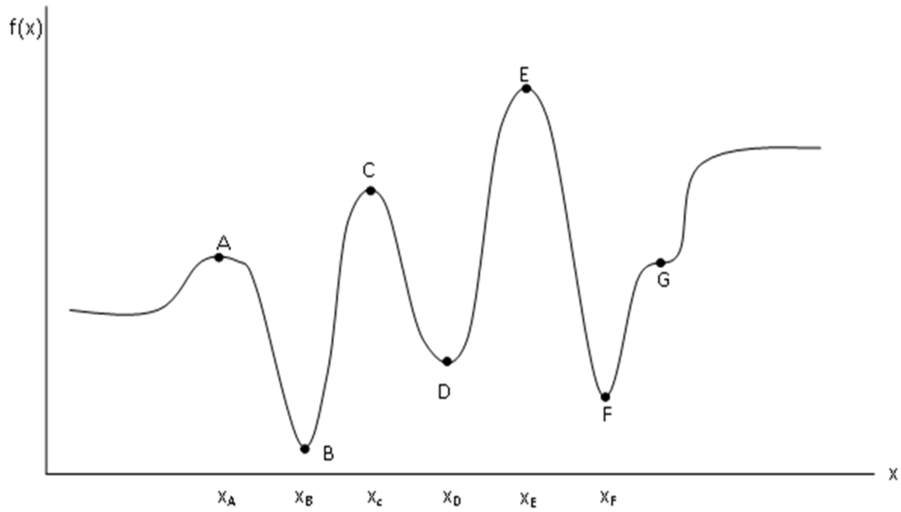

Figure 2: A,C,E are local maxima; E is a global maximum; B,D,F are local minima; B is the global minimum. G is a saddle point (not an extremum).

5.2 Necessary and sufficient conditions for maxima

Necessary condition For any (differentiable) function $f(x)$, the point x^* is a minimum or maximum if

$$\boxed{\frac{df(x^*)}{dx} = 0.} \tag{2}$$

This is called the *First Order Condition (FOC)*.

Sufficient condition for local extrema. For any (differentiable) function $f(x)$, the point x^* is a local minimum or local maximum if

$$\boxed{\begin{array}{c} \dfrac{df(x^*)}{dx} = 0 \text{ and} \\ \text{Local max: } \dfrac{d^2 f(x^*)}{dx^2} < 0; \text{ Local min: } \dfrac{d^2 f(x^*)}{dx^2} > 0. \end{array}} \tag{3}$$

Conditions involving the second derivatives are called *Second Order Conditions (SOC)*.
Note: At a saddle point we have $\frac{d^2 f(x^*)}{dx^2}$ <= but $\frac{d^2 f(x^*)}{dx^2} = 0$ does not imply that we have found a saddle point.

Sufficient conditions for a global maximum. Point x^* maximized the function if

$$\boxed{\frac{df(x^*)}{dx} = 0 \text{ and } \frac{d^2 f(x^*)}{dx^2} < 0 \text{ for all } x.} \tag{4}$$

5.3 Economic examples

Profit Maximization

Suppose the inverse demand for skits is given by $p(q) = 1000 - 9q$. The cost of producing q skirts is $C(q) = 100 + q^2$. How many skirts should the skirt manufacturer produce to maximize profits?

Step 1: Write down the maximization problem The profit function is $\pi(q) = R(q) - C(q) = (100 - 9q)q - (100 + q^2)$. The firm wants to find the q^* that maximized the value of this function. In short,

$$max_q (100 - 9q)q - (100 + q^2).$$

Step 2: Solve the FOC. This finds values of q that might maximize profits.

$$\begin{aligned} FOC \quad &: \quad \frac{d\pi(q^*)}{dq} = 0 \\ &\rightarrow \quad 1000 - 18q^* - 2q^* = 0 \\ &\rightarrow \quad q^* = 50. \end{aligned}$$

Step 3: Solve the SOC. This identifies what kind of extremum you have found in Step 2 (if any).

$$SOC : \frac{d^2\pi}{dq^2} = -20 \text{ for all } q.$$

This is a sufficient condition for a global maximizer, so that $q^* = 50$ skirts maximizes profits.

Unconstrained utility maximization

Joe already paid to get into Newell dining hall. HE only eats pizza and his utility from pizza slices is $u(z) = -z^2 + 8z$. How many slices will Joe eat? Note that since Joe is already in the dining hall, pizza is "free" in the sense that th marginal cost of an additional slice is zero.

Step 1: Write down th maximization problem.

$$max_z - z^2 + 8z$$

Step 2: Solve the FOC. This fins values of z that might maximize utility.

$$\begin{aligned} FOC \quad &: \quad \frac{du(z^*)}{dz} = 0 \\ &\rightarrow \quad -2z^* + 8 = 0 \\ &\rightarrow \quad z^* = 4 \end{aligned}$$

Step 3: Solve the SOC. This identifies what kind of extremum you have found in Step 2 (if any).

$$SOC : \frac{d^2 u}{dz^2} = -2 \text{ for all } z.$$

This is a sufficient condition for a global maximizer, so Joe will eat $z^* = 4$ slices (until his marginal utility is zero).

6 Constrained optimization

In economics we tend to study the optimal allocation of scarce resources. A typical optimization problem can therefore by written as follows

$$\begin{aligned} \max_{x_1,...,x_n} u\left(x_1, ..., x_n\right) &\quad : \quad \text{s.t.} \\ p_1 x_1 + ... + p_n x_n &\quad = \quad I, \end{aligned}$$

where so called Inada conditions on the utility function u (i.e. $\lim_{x_i \to 0} \frac{\partial u}{\partial x_i} = \infty$ for $i = 1, ..., n$) ensure interior solutions so that the consumer consumes a positive quantity of each good x_i. Variable p_i is the price of good i and I is income.

We first look at a two goods case so that

$$\begin{aligned} \max_{x_1,x_2} u\left(x_1, x_2\right) &\quad : \quad \text{s.t.} \\ p_1 x_1 + p_2 x_2 &\quad = \quad I. \end{aligned}$$

You can draw this problem into a graph (see figure 3).

At the optimum it has to hold that the slope of the indifference curve equals the slope of the budget constraint

$$\frac{\partial u / \partial x_1}{\partial u / \partial x_2} = \frac{p_1}{p_2}.$$

We can rewrite this optimality conditions as

$$\frac{\partial u / \partial x_1}{p_1} = \frac{\partial u / \partial x_2}{p_2},$$

which equals the "bang per buck" spent on good 1 to the "bang per buck" spent on good 2. The "bang per buck" is measured as marginal utility increase per dollar spent.

Note: If $\frac{u_{x_1}}{p_1} > \frac{u_{x_2}}{p_2}$ you could make yourself better off by shifting some money spent on good 2 towards purchasing more of good 1. If $\frac{u_{x_1}}{p_1} < \frac{u_{x_2}}{p_2}$ you could make yourself better off spending more on good 2. Once the two expressions are equal, you cannot make yourself better off by further shifting of funds.

We next set this ratio equal to a constant λ so that

$$\frac{u_{x_1}}{p_1} = \frac{u_{x_2}}{p_2} = \lambda^*,$$

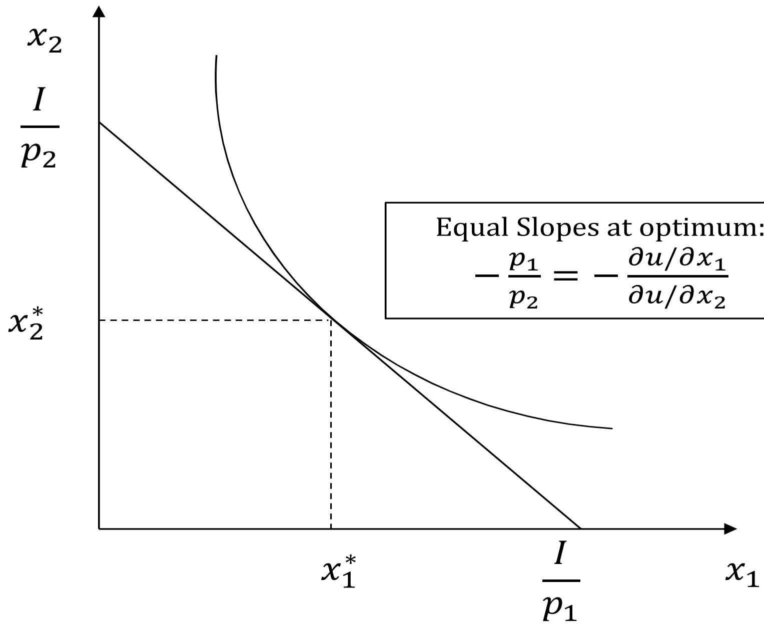

Figure 3: Constrained optimization. At the optimum point the slope of the budget line $-\frac{p_1}{p_2}$ is equal to the slope of the indifference curve $-\frac{\partial u/\partial x_1}{\partial u/\partial x_2}$.

which results in two separate equations that have to hold in the optimum

$$\begin{cases} u_{x_1}(x_1^*, x_2^*) - \lambda^* p_1 = 0, \\ u_{x_2}(x_1^*, x_2^*) - \lambda^* p_2 = 0. \end{cases}$$

Since we have to solve for three unknowns, x_1^*, x_2^*, and λ we need three equations. We therefore add the budget constraint to the two equations which results in the following system:

$$\begin{cases} u_{x_1}(x_1^*, x_2^*) - \lambda^* p_1 = 0, \\ u_{x_2}(x_1^*, x_2^*) - \lambda^* p_2 = 0, \\ I - p_1 x_1^* - p_2 x_2^* = 0. \end{cases}$$

This we can now solve for x_1^*, x_2^*, and λ^*.

There is a convenient way of writing the maximization problem in the form of the Lagrangian function:

$$L(x_1, x_2, \lambda) \equiv u(x_1, x_2) + \lambda (I - p_1 x_1 - p_2 x_2).$$

We can find the critical points of this Lagrangian function by simply deriving the function w.r.t.

its arguments and setting these derivatives equal to zero:

$$\left\{ \begin{array}{l} \frac{\partial L(x_1,x_2,\lambda)}{\partial x_1} = 0 : \frac{\partial u\left(x_1^*,x_2^*\right)}{\partial x_1} - \lambda^* p_1 = 0, \\ \frac{\partial L(x_1,x_2,\lambda)}{\partial x_2} = 0 : \frac{\partial u\left(x_1^*,x_2^*\right)}{\partial x_2} - \lambda^* p_2 = 0, \\ \frac{\partial L(x_1,x_2,\lambda)}{\partial \lambda} = 0 : I - p_1 x_1^* - p_2 x_2^* = 0. \end{array} \right. \tag{5}$$

This system is identical to the one derived above. This system can now be solved for x_1^*, x_2^*, and λ^*. The constant λ is called **Lagrange multiplier.** It measures the sensitivity of the optimal value of the utility function to changes in income. In other words, it measures the marginal utility of wealth at the optimum, i.e. $\frac{du\left(x_1^*,x_2^*\right)}{dI} = \lambda^*$.

Proof. The Lagrangian function at the optimum point is defined as

$$L\left(x_1^*, x_2^*, \lambda^*; I\right) \equiv u\left(x_1^*, x_2^*\right) + \lambda^*\left(I - p_1 x_1 - p_2 x_2\right),$$

where I now enters the Lagrangian function as a parameter so that the optimal allocation is a function of this parameter, i.e. $x_1^*(I), x_2^*(I)$. The first order conditions are therefore

$$\frac{\partial u\left(x_1^*, x_2^*\right)}{\partial x_1} - \lambda^* p_1 = 0, \tag{6}$$

$$\frac{\partial u\left(x_1^*, x_2^*\right)}{\partial x_2} - \lambda^* p_2 = 0, \tag{7}$$

and the total derivative of the budget constraint in the optimium is

$$dI - p_1 dx_1 - p_2 dx_2 = 0 \rightarrow p_1 \frac{dx_1}{dI} + p_2 \frac{dx_2}{dI} = 1 \tag{8}$$

We next derive the utility function w.r.t. I and get

$$\begin{array}{rcl} \frac{\partial u\left(x_1^*(I), x_2^*(I)\right)}{\partial I} & = & \frac{\partial U}{\partial x_1}\frac{dx_1}{dI} + \frac{\partial u}{\partial x_2}\frac{dx_2}{dI}, \\ \text{from FOC } (6),(7) & = & \lambda^* p_1 \frac{dx_1}{dI} + \lambda^* p_2 \frac{dx_2}{dI}, \\ \text{from } (8) & = & \lambda^*\left(p_1 \frac{dx_1}{dI} + p_2 \frac{dx_2}{dI}\right), \\ & = & \lambda^*. \end{array}$$

■

In general, the Lagrangian for a general system

$$\begin{array}{rcl} \max\limits_{x_1,\ldots,x_n} u\left(x_1, ..., x_n\right) & : & \text{s.t.} \\ h\left(x_1, ..., x_n\right) & = & I, \end{array}$$

is

$$L\left(x_1, ..., x_n, \lambda\right) \equiv u\left(x_1, ..., x_n\right) + \lambda\left(I - h\left(x_1, ..., x_n\right)\right)$$

and the interior solution $\left(x_1^*, ..., x_n^*, \lambda^*\right)$, if it exists, can be found by solving the system

$$\begin{aligned}
\frac{\partial L\left(x_1, ..., x_n, \lambda\right)}{\partial x_1} &= 0, \\
&\vdots \\
\frac{\partial L\left(x_1, ..., x_n, \lambda\right)}{\partial x_n} &= 0, \\
\frac{\partial L\left(x_1, ..., x_n, \lambda\right)}{\partial \lambda} &= 0.
\end{aligned}$$

References

Dadkhah, Kamran. 2011. *Foundations of Mathematical and Computational Economics*. 2 ed. Springer.

Hands, Wade D. 1991. *Introductory Mathematical Economics*. Heath and Company.

Simon, Carl P. and Lawrence Blume. 1994. *Mathematics for Economists*. New York: W.W. Norton.

Williamson, Steven D. 2013. *Macroeconomics*. Addison-Wesley.