Lecture Notes

Mathematical Economics Class

August 25, 2025

Lecturer: Ismael Moreno-Martinez
Mail: ismael.moreno@eui.eu

The course is designed to facilitate a smooth transition into a first year PhD program in Economics by providing the the basic mathematical foundation needed by incoming students. By focusing on intuitive understanding and practical applications, the classes are taught as a complement to a background course in Mathematics, such as the one held by Prof. Villanacci at the EUI.

The main references for the course are these lecture notes based on:

- Martin J. Osborne. Mathematical methods for economic theory. https://mjo.osborne.economics.utoronto.ca/index.php/tutorial/index/1/toc
- Cooper Russell. Lecture Notes. Overview of Applied Dynamic Programming.
- Simon, C.P. and Blume, L., 1994. Mathematics for economists (Vol. 7). New York: Norton.
- Villanacci Antonio. Lecture Notes. Background Course on Mathematics.

Contents

1 Calculus

In this lecture, I will provide a few basic definitions and then we will go over some useful calculus tools, namely, (1) the chain rule, (2) total derivatives, and (3) the implicit function theorem. These will be fundamental when working with economic models with multiple interrelated endogenous variables and exogenous parameters. For example, we may want to understand how the equilibrium labor supply changes when we adjust a particular policy parameter (i.e. generosity of government transfers).

1.1 Basic Calculus Definitions

Before we get into those tools, let's define the main subject of this chapter, functions:

Definition 1: Function

A **function** is a rule that assigns to each element in a set called the *domain* exactly one element in a set called the *codomain*.
Formally, a function f from set X to set Y is denoted as:

$$f : X \to Y$$

and for every $x \in X$, there exists a unique $y \in Y$ such that $y = f(x)$.

In economics, we often interpret functions as relationships between economic variables. For example, a utility function $u(x)$ assigns a utility value to each consumption bundle x of n goods, and we define each bundle as a point in the set of positive real numbers $\mathbb{R}^n_+$. Usual domains of functions in economics are the real numbers $\mathbb{R}$, or some interval within them (e.g. $\mathbb{R}_+$, $[0, 1]$)

Now, let's define a couple of interesting properties that make functions easy to work with: continuity and differentiability. You can think of a function as **continuous** if you can draw its graph without lifting your pen from the paper. This means there are no sudden jumps, holes, or breaks in the function.

Definition 2: Continuity

Formally, f is continuous at x_0 if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\|x - x_0\| < \delta \quad \Rightarrow \quad |f(x) - f(x_0)| < \varepsilon$$

If this condition holds for all x_0 in the domain of f, we say that f is continuous on its domain.[a]

[a] Here, $\|\cdot\|$ denotes the standard Euclidean distance in $\mathbb{R}^n$:

$$\|y - x\| = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \cdots + (y_n - x_n)^2}.$$

Intuitively, this means that if we "zoom in" close enough to a point x_0, then the function values $f(x)$ will also be very close to $f(x_0)$. In other words, small changes in the input lead to only small changes in the output. This is precisely what rules out sudden jumps or breaks in the function.

Just like continuity means you can draw the function without lifting your pen, **differentiability** means the function is smooth: there are no sharp corners or kinks. At every point, the function has a well-defined slope (in one dimension) or a well-defined tangent plane (in higher dimensions).

Definition 3: Differentiability

Univariate case: A function $f : \mathbb{R} \to \mathbb{R}$ is **differentiable** at a point $x_0 \in \mathbb{R}$ if the following limit exists as a finite and unique real number:[a]

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

In this case, the value of the limit is **the derivative of** f at x_0, denoted by $f'(x_0)$.

Multivariate case: A function$f$ is differentiable at $x_0 \in \mathbb{R}^n$ if there exists a vector $v \in \mathbb{R}^n$ such that

$$\lim_{\|h\| \to 0} \frac{|f(x_0 + h) - f(x_0) - v \cdot h|}{\|h\|} = 0$$

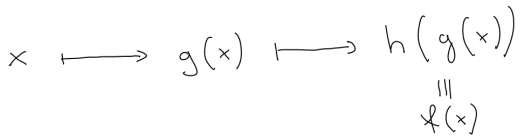
If vector v exists, it is the gradient of derivatives of f at x_0: $\nabla f(x_0)$.

[a] Note that the limit must be the same from both the left and the right. For instance, the absolute value function $f(x) = |x|$ is not differentiable at $x = 0$, because the left and right limits disagree.

Intuitively, this means that near x_0, the function behaves like its tangent plane $\nabla f(x_0) \cdot h$. Differentiability implies continuity, but the converse is not true. In economics, differentiability allows us to compute marginal effects and apply tools like the chain rule, optimization, and comparative statics. Finally, we can say a function is **smooth** if it is differentiable, and all its higher order derivatives are also differentiable.

1.2 The Chain Rule

Given two functions $g(x)$ and $h(x)$ defined from $\mathbb{R}$ to $\mathbb{R}$, consider the function formed by first applying g to any number x and then applying function h to the result $g(x)$. Such a function is called **composite function** and can be denoted as $f(x) = h(g(x))$ or as $f(x) = (h \circ g)(x)$.



Example: The function $f(x) = e^x + 1$ can be seen as $f(x) = h(g(x))$ where $g(x) = e^x$ and $h(x) = x + 1$. □

We are interested in studying the derivative of a composite function. For this purpose, we use the Chain Rule.

Definition 4: Chain Rule

The **Chain Rule** states that for a function $f(x) = h(g(x))$ the derivative is

$$\frac{\partial f(x)}{\partial x} = \frac{\partial h(g(x))}{\partial x} = h'(g(x)) \cdot g'(x)$$

In words, the derivative of a composite function can be thought of as the derivative of the *outside* function (h), evaluated at the inside function, times the derivative of the *inside* function (g). Be aware that to apply the Chain Rule the function h and g have to be differentiable. [1]

For clarity and to get used to the notation, the chain rule can also be written as

$$\frac{\partial f(x)}{\partial x} = \frac{\partial (h \circ g)(x)}{\partial x} = \frac{\partial h(g(x))}{\partial g(x)} \cdot \frac{\partial g(x)}{\partial x}$$

Example: Let's differentiate the function $U(x) = (\alpha x^{\rho} + 1)^{\frac{1}{\rho}}$ with α and ρ as parameters, this function can be seen as $U(x) = h(g(x))$ where $g(x) = (\alpha x^{\rho} + 1)$ is the inside function and $h(x) = x^{\frac{1}{\rho}}$ as outside function. The derivative of the outside function evaluated at the inside function is $h'(g(x)) = \frac{1}{\rho}(g(x))^{\frac{1}{\rho}-1} = \frac{1}{\rho}(g(x))^{\frac{1-\rho}{\rho}} = \frac{1}{\rho}(\alpha x^{\rho} + 1)^{\frac{1-\rho}{\rho}}$, while the derivative of the inside function is $g'(x) = (\alpha \rho x^{\rho-1})$. By applying the Chain Rule we get:

$$\frac{\partial f}{\partial x}(x) = \frac{\partial (h \circ g)}{\partial x}(x) = h'(g(x))g'(x) = \alpha x^{\rho-1}(\alpha x^{\rho} + 1)^{\frac{1-\rho}{\rho}}$$

□

Similarly, the rule can be applied if a function is composed of more than two functions.

Example: Let's derive the function $f(x) = [log(x^{\frac{1}{3}})]^2$. In this case, we have functions

$$\begin{aligned} h(x) &= x^2 \\ g(x) &= log(x) \\ k(x) &= x^{\frac{1}{3}} \end{aligned}$$

and derivative $f'(x) = h'(g(k(x)) \cdot g'(k(x)) \cdot k'(x)$, that is:

$$\frac{\partial f}{\partial x}(x) = 2[log(x^{\frac{1}{3}})]^{2-1} \cdot \frac{1}{x^{\frac{1}{3}}} \cdot \frac{1}{3} x^{\frac{1}{3}-1} = \frac{2}{3x} log(x^{\frac{1}{3}})$$

□

What happens if we have two inside functions? For example, $f(x) = h(g_1(x), g_2(x))$? We apply the chain rule for each component and sum up the results:

$$\frac{\partial f}{\partial x}(x) = h_1'(g_1(x), g_2(x))g_1'(x) + h_2'(g_1(x), g_2(x))g_2'(x)$$

[1] If g and h are both differentiable functions, then the composite function $f(x) = h(g(x))$ is also differentiable.

Notice that for multivariate functions, we can use the subscript to indicate with respect to which input we are taking the derivative (e.g. for function $f(x, y)$, the derivative with respect to x can be written $f_1'(x, y)$).

Example: Let's differentiate the function $V(k) = u(f(k) - k)$

$$\frac{\partial V(k)}{\partial k} = u'(f(k) - k) \cdot f'(k) + u'(f(k) - k) \cdot (-1) = u'(f(k) - k) \cdot (f'(k) - 1)$$

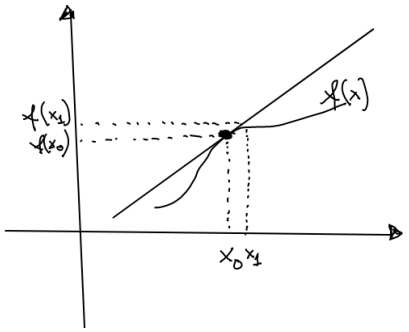
Note that in the above differentiation we are considering $\frac{\partial V}{\partial k}(x) = \frac{dh}{dk}(g_1(k), g_2(k))$ where $g_1(k) = f(k)$ and $g_2(k) = -k$. Alternatively we could have simply considered as outside function $u()$ and as inside function all $f(k) - k$, and applying directly the simple version of the chain rule.

□

1.3 The Total Derivative

In economics we often deal with models that depend on several variables at once: wages, prices, endowments, technology, policy parameters, etc. When more than one of these changes, we want to know how the outcome of interest responds. The **total derivative** gives us a systematic way to approximate such changes, combining all the partial effects into one expression. The total derivative will also serve as a stepping stone for more advanced results such as the Implicit Function Theorem. In the remaining of the lecture, we will apply these concepts to conduct some 'comparative statics' (i.e. we will study how the outcomes of a very simple model change when we change the parameters).

We know it is possible to approximate the value of a univariate function around a point x_0 by making use of the derivative. That is, the value of function f around x_0 can be approximated by the line $f(x_0) + f(x_0)'(x - x_0)$. This line can be represented as the tangent to the plot of $f(x)$ at x_0



Similarly, we can use the derivative as a linear approximation of the changes of the function around x_0, so that $f(x) - f(x_0) \approx f(x_0)'\Delta x$. Where $\Delta x = x - x_0$.

The same can happen in a multivariate world. The value of function $f(x_1, ..., x_n)$ around $(\bar{x}_1, ..., \bar{x}_n)$ can be approximated by

$$f(x_1, ..., x_n) \approx f(\bar{x}_1, ..., \bar{x}_n) + \frac{\partial f}{\partial x_1}(\bar{x}_1, ..., \bar{x}_n)\Delta x_1 + \frac{\partial f}{\partial x_2}(\bar{x}_1, ..., \bar{x}_n)\Delta x_2 + ... + \frac{\partial f}{\partial x_n}(\bar{x}_1, ..., \bar{x}_n)\Delta x_n$$

Note that if in $\mathbb{R}^2$ we were graphically interpreting the approximation as a line tangent to the curve, in $\mathbb{R}^3$ we are approximating the function with the plane tangent to the curve.

If one is interested in understanding how the value of a function changes when all its input variables change, one should compute:

$$f(x_1,...,x_n) - f(\bar{x}_1,...,\bar{x}_n) \approx \frac{\partial f}{\partial x_1}(\bar{x}_1,...,\bar{x}_n)\Delta x_1 + \frac{\partial f}{\partial x_2}(\bar{x}_1,...,\bar{x}_n)\Delta x_2 + ... + \frac{\partial f}{\partial x_n}(\bar{x}_1,...,\bar{x}_n)\Delta x_n$$

as the change becomes infinitesimal ($\Delta \to 0$), we get

$$df \approx \frac{\partial f}{\partial x_1}(\bar{x}_1,...,\bar{x}_n)dx_1 + \frac{\partial f}{\partial x_2}(\bar{x}_1,...,\bar{x}_n)dx_2 + ... + \frac{\partial f}{\partial x_n}(\bar{x}_1,...,\bar{x}_n)dx_n$$

the RHS of the equation is called the total differential of a function and is denoted as df. Note that the elements in the RHS are the multiplication of partial derivatives with the marginal changes. If we consider the **gradient** evaluated at $(\bar{x}_1,...,\bar{x}_n)$, that is the vector of partial derivatives evaluated in that point:

$$\nabla f(\bar{x}_1,...,\bar{x}_n) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\bar{x}_1,...,\bar{x}_n) \\ \frac{\partial f}{\partial x_2}(\bar{x}_1,...,\bar{x}_n) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\bar{x}_1,...,\bar{x}_n) \end{pmatrix}$$

and the vector of marginal changes

$$dx = \begin{pmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{pmatrix}$$

we can rewrite the **total differential** in more compact form:

$$df = \nabla f(\bar{x}_1,...,\bar{x}_n) \cdot dx$$

where the $\nabla f(\bar{x}_1,...,\bar{x}_n)$ is the gradient of the function and dx is the vector of marginal changes.

Example: Let's find a linear approximation that describes what happens to the function $F(x_1,x_2) = x_1^{\frac{2}{3}}x_2^{\frac{1}{3}}$ when we slightly change x_1 and x_2, that is we compute the total derivative:

$$\begin{aligned} dF &= \frac{\partial F(x_1,x_2)}{\partial x_1}dx_1 + \frac{\partial F(x_1,x_2)}{\partial x_2}dx_2 \\ &= \frac{2}{3}x_1^{-\frac{1}{3}}x_2^{\frac{1}{3}}dx_1 + \frac{1}{3}x_1^{\frac{2}{3}}x_2^{-\frac{2}{3}}dx_2 \end{aligned}$$

□

Example: The total derivative can also be used to study an optimal point. Let's use the total derivative to do some 'comparative statics', that is, evaluate how the outcome of a model changes as we tweak the exogenous parameters of that model. Consider the problem

$$\max_n u(wn + A) - g(n) \tag{1}$$

where

- u is the utility function with $u' > 0$, $u'' < 0$ (strictly increasing and concave)
- g is the work-disutility function with $g' > 0$, $g'' > 0$ (strictly increasing and convex)
- $u'(0) > g'(0)$ and $u'(1) < g'(1)$
- w is the salary, $0 \leq n \leq 1$ is the amount of hours worked and A is a fix income

We want to find the optimal amount of hours worked and determine how this is influenced by changes in A.
To find the optimal point, we take the derivative of the function w.r.t. n and set it equal to zero. The first order condition (F.O.C.) is

$$wu'(wn + A) - g'(n) = 0 \tag{2}$$

The optimal value of working hours must satisfy the above equation. In other words, every value of n that satisfies the above equation is an optimal value of n (n^*).

We are now interested in understanding what happens to the optimal value once we change the value of A. To do so, we will redefine the optimality condition above as a function $f(n, w, A) = wu'(wn + A) - g'(n)$ and use the total derivative. Notice that we have conceptually changed how we see the elements of the model, now we see w and A as variable inputs of a function, not as fixed parameters. This is because we have put our "modeler hats", while before we were approaching the problem from the point of view of a household living inside our model.
We compute the total derivative:

$$df = (w^2u'')dn + (u' + wnu'')dw + [wu'']dA - g''dn$$

We can interpret the total derivative above as the changes in the LHS of our FOC (see equation 2) with respect to small changes in (n, w, A). This implies that any change in one of the variables, say dA, has to be 'balanced' by an adequate change in another variable, say dn, in order for the FOC to still hold. Put differently, if we want to move to a different point is still a solution to the problem, we need to choose (dn, dw, dA) so that $df = 0$. That is

$$(w^2u'')dn + (u' + wnu'')dw + (wu'')dA - (g'')dn = 0$$

Now, since we are interested in the relation between n and A, we can set $dw = 0$, and solve for $\frac{dn}{dA}$:

$$\frac{dn}{dA} = -\frac{wu''}{[w^2u'' - g'']} < 0$$

We now this relation is negative because the numerator is negative and the denominator is positive, following the model assumptions. Therefore we can say that when the endowment A increases, optimal labor supply n decreases. □

1.4 The Implicit Function Theorem

In the last subsection we used the total derivative to approximate how solutions react to parameter changes. The Implicit Function Theorem (IFT) provides the rigorous justification for why such comparative statics work: it guarantees that under certain conditions, the solution can indeed be expressed as a differentiable function of the parameters, and it gives us the formula for its derivative. In other words, the IFT formalizes and extends the comparative statics

logic we developed using total derivatives.

In economics, equilibrium conditions are usually represented by equations that include both endogenous, x, and exogenous variables (or parameters), p:

$$f(x, p) = 0 \tag{3}$$

These equations, assuming they have a solution, define an implicit function of x with respect to p. That is the mapping between the potential values parameter p could take, and the corresponding values x should take to solve the model. This function, $g(p)$, satisfies $f(g(p), p) = 0$ for all p. Now, say we found a combination (x_0, p_0) so that $f(x_0, p_0) = 0$. We may want to understand the behavior of implicit function g around p_0. This is equivalent to understanding how the outcome of the model, x, changes with parametrization. The IFT help us do exactly this:

Definition 5: Implicit Function Theorem

Let f be a continuously differentiable function of two variables, defined on an open set $S \subseteq \mathbb{R}^2$. Suppose that

$$\frac{\partial f}{\partial x}(x_0, p_0) \neq 0.$$

Then there exists an open interval I containing p_0 and a continuously differentiable function $g : I \to \mathbb{R}$ such that

$$f(g(p), p) = f(x_0, p_0) \quad \text{for all } p \in I,$$

and

$$g'(p_0) = -\frac{\partial f/\partial p(x_0, p_0)}{\partial f/\partial x(x_0, p_0)}.$$

The corolary in the definition above can be easily derived from the expression $f(g(p), p) = 0$. Applying the chain rule:

$$\frac{\partial f}{\partial x}(g(p), p) \cdot g'(p) + \frac{\partial f}{\partial p}(g(p), p) = 0,$$

so that

$$g'(p) = -\frac{\frac{\partial f}{\partial p}(g(p), p)}{\frac{\partial f}{\partial x}(g(p), p)}.$$

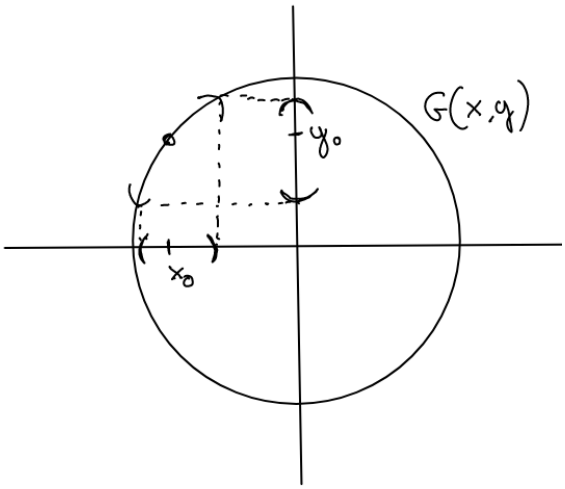
IFT Recipe

1. Write down the equation as a function such that $f(x, p) = 0$ and identify the exogenous (p) and the endogenous (x) variables.
2. Compute the partial derivatives wrt x and p
3. Check if the assumptions hold around a point of interest (x_0, p_0)
   - The function is defined in an open set around (x_0, p_0)
   - The function and its derivatives are continuous in this set
   - The partial derivative of x is different from zero ($\frac{\partial f(x_0, p_0)}{\partial x} \neq 0$)

4. Apply the theorem and compute the derivative

$$g'(x) = -\frac{\frac{\partial f}{\partial p}(x,p)}{\frac{\partial f}{\partial x}(x,p)}$$

Example: Consider the formula of the circle of radius one and centered at the origin, characterized by the equation $x^2 + y^2 = 1$ (or $x^2 + y^2 - 1 = 0$). Let's consider an arbitrary point in the circle, $(-\sqrt{1/2}, \sqrt{1/2})$. We want to see if, in an interval around this point, we can find the relation $g'(-\sqrt{1/2}, \sqrt{1/2})$.



□

Following the recipe presented above...

1. We write down the equation as a function $f(x, y)$...

$$f(x,y) = x^2 + y^2 - 1$$

   and identify the exogenous and the endogenous variables. In this case it won't make a difference which one, x or y we choose as endogenous, so we will choose x arbitrarily.

2. We compute the partial derivatives $f_x(x,y) = 2x$ and $f_y(x,y) = 2y$

3. We check if the IFT assumptions hold around $(-\sqrt{1/2}, \sqrt{1/2})$

   - Is the function defined in an open set around $(-\sqrt{1/2}, \sqrt{1/2})$ → Yes, f is in defined in all $\mathbb{R}^2$.
   - Are the function and its derivatives continuous in this set? → Yes (polynomials are continuous an differentiable in $\mathbb{R}$.
   - Is the partial derivative of x different from zero? Yes, $\frac{\partial f(-\sqrt{1/2},\sqrt{1/2})}{\partial x} = -2\sqrt{1/2} \neq 0$

4. We apply the IFT at $(-\sqrt{1/2}, \sqrt{1/2})$ and compute the derivative

$$g'(y) = -\frac{\frac{\partial f(-\sqrt{1/2},\sqrt{1/2})}{\partial y}}{\frac{\partial f(-\sqrt{1/2},\sqrt{1/2})}{\partial x}} = -\frac{2\sqrt{1/2}}{-2\sqrt{1/2}} = 1$$

Done! Now, does the IFT apply for all values of x and y such that $f(x,y)=0$? Why? □

Example: We can re-approach the problem in the last chapter through the IFT. Again, we want to study how the choice of n changes when A changes. We start with the optimality condition (FOC):

$$wu'(wn+A) = g'(n)$$

We follow the recipe:

1. We write down the FOC as an implicit function h (everything on the same side!) and identify the endogenous (n) and exogenous variables (A):

$$h(n,A) = wu'(wn+A) - g'(n) = 0$$

2. We compute the partial derivatives:
   $\frac{\partial h}{\partial n} = w^2u''(wn+A) - g''(n)$
   $\frac{\partial h}{\partial A} = wu''(wn+A)$

3. We check the assumptions
   - By assumption, h is defined at the optimal point that solves $h(n,A)=0$
   - Also by assumption, the function and derivates are continuous in the domain (and around the solution point)
   - Again, from the problem, we know $u''<0$ and $g''>0$. Therefore, the derivative wrt the endogenous variable must be different from zero: $\frac{\partial h}{\partial n} = w^2u''(wn+A) - g''(n) < 0$

4. We can apply the IFT:

$$\frac{\partial n}{\partial A} = -\frac{\frac{\partial h}{\partial A}}{\frac{\partial h}{\partial n}} = -\frac{wu''(wn+A)}{w^2u''(wn+A) - g''(n)}$$

Which is the same result we got in the Total Derivative Example. □

2 Optimization

In economics, we are often interested in finding the maxima or minima of functions subject to a set of restrictions. This is common when we model the decision of economic agents in a model. For example, we can understand households' consumption decision as the maximization problem of a utility function that depends on the consumption of different goods, which quantities are subject to a budget restriction. We also solve maximization/minimization problems in our regressions (e.g. *maximum* likelihood estimators, ordinary *least* squares) or when doing structural econometric modelling (e.g. simulated method of moments minimizes the distance between model-generated moments and real data).

2.1 Basic Topology Definitions

Definition 6: Interior and Boundary Points

Let S be a subset of $\mathbb{R}^n$.

- A point x is an **interior point** of S if there exists $\varepsilon > 0$ such that all points within distance ε of x are members of S.
- A point x is a **boundary point** of S if for every $\varepsilon > 0$, no matter how small, there exists at least one point within distance ε of x that belongs to S and at least one point within that distance that does not belong to S.

The set of interior points of the disk $S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq c\}$ is $\{(x, y) : x^2 + y^2 < c\}$, and the set of boundary points is $\{(x, y) : x^2 + y^2 = c\}$.

Now, the presence of boundary and interior points in a set will determine whether these are open, or closed. As economists, we will very frequently represent the actions an economic agent can take as a set of points in a set (e.g. set of possible levels of investment, hours to work, crimes to commit, pollution to emit, PhD students to fail, etc.). Whether these sets are open or closed, will determine how to solve the problem of the agent, or wether the problem has a solution, at all.

Definition 7: Open and Closed Sets

Let $S \subseteq \mathbb{R}^n$.

- The set S is **open** if every point in S is an interior point.
- The set S is **closed** if every boundary point of S is also a member of S.

Note that both $\mathbb{R}^n$ and the empty set $\emptyset$ are examples of sets that are **both open and closed**, sometimes called *clopen* sets. This is because $\mathbb{R}^n$ has no boundary points, and $\emptyset$ vacuously satisfies the definitions of openness and closedness.

Now, a useful concept in topology is that of a **ball**. A ball can be seen as the multidimensional version of an interval. Let me provide a formal definition:

Definition 8: Ball

Let $x \in \mathbb{R}^n$ and let $\varepsilon > 0$. The **open ball** of radius ε centered at x is the set

$$\{y \in \mathbb{R}^n : \|y - x\| < \varepsilon\}.$$

Sometimes, when we want to check if some property of a function holds around a point, we will check this property in a ball around that point.

Definition 9: Convex Set

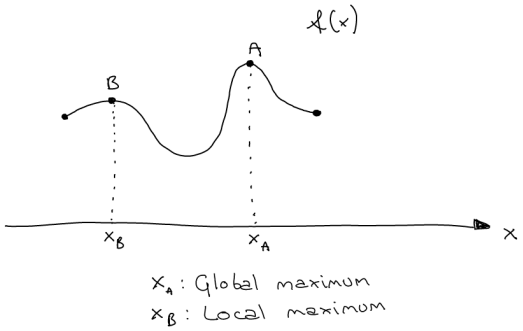
A set $S \subseteq \mathbb{R}^n$ is called **convex** if, for every pair of points $x, y \in S$, the line segment connecting x and y is entirely contained in S. That is, for all $\lambda \in [0, 1]$,

$$\lambda x + (1 - \lambda)y \in S.$$

For example, the set $S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ is convex, because the straight line between any two points inside the disk stays entirely within the disk. Notice commonly used sets, such as $\mathbb{R}^n$, $\mathbb{R}_+$ or $[0, 1]$, are convex.

2.2 Optimization Problems

A function f has a *local maximum* at a point x^* if $f(x) \leq f(x^*)$ for all x in some open interval containing x^*. The maximum is instead *fglobal or absolute* if the unequality holds for all x in the domain of f. A minimum can be similarly defined.



In functions with several variables, such as $F : U \in \mathbb{R}^n \to \mathbb{R}^1$, the maximum is similarly defined. f has a local maximum at a point $\mathbf{x}^* \in U$ if $F(\mathbf{x}^*) \geq F(\mathbf{x})$ for all $\mathbf{x}$ in some open ball containing $\mathbf{x}^*$. The maximum is instead global if the unequality holds for all the $\mathbf{x}$ in the domain, U.

A general optimization problem can be written as

$$\begin{aligned} \text{maximize} \quad & f(x, y) \\ \text{subject to} \quad & h(x, y) = 0 \\ & g(x, y) \leq 0 \\ & x \in \Omega \subset \mathbb{R}, \\ & y \in \hat{\Omega} \subset \mathbb{R}, \end{aligned}$$

where f is the **objective function**, while x and y are the **choice variables**. Here, Ω and $\hat{\Omega}$ are the domains of the variables of interest, x and y respectively. The condition $h(x, y) = 0$ represents equality constraints, and $g(x, y) \leq 0$ inequality constraints. For simplicity, we consider only one constraint of each type.

How do we solve the above optimization problem? The method depends on the presence and type of constraints:

- if we have a maximization problem without any constraint, we can use the derivative of the function to find the optimal points.

- If there is an equality constraint, we can either use substitution or the Lagrange multiplier method.
- If there is an inequality constraint, we need to use the Kuhn-Tucker conditions (also known as the Karush-Kuhn-Tucker (KKT) conditions).

Why have we only discussed the maximization of a function and not the minimization?

Because maximization and minimization are related by the following condition:

$$min(f(x)) \text{ s.t. } g(x) \leq 0 = max(-f(x)) \text{ s.t. } g(x) \leq 0$$

Thus, it is always possible to transform a minimization problem into a maximization one by multiplying the function by -1.

2.3 Unconstrained Optimization

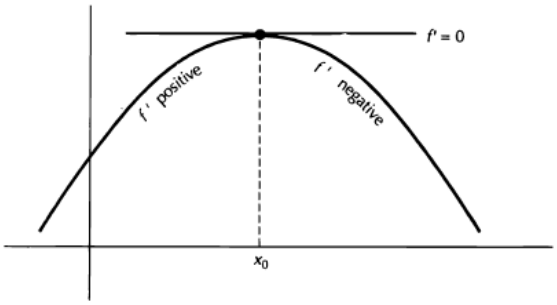
In the absence of constraints, candidates to be a functions' maximum are encountered:

1. at points where the derivative is zero (*critical points*, denoted as x^*).
2. at the points where the derivative cannot be computed.[2]
3. at the boundaries of the domain of f.

In economics we often work with functions that are continuous and differentiable, so frequently we don't need to worry about type 2 cases. It is also common to focus on problems with "interior" solutions, or with open domains (that don't contain their boundaries), so type 3 solutions can be less frequent. Let's focus for now in the most standard case, 1. How do we find such a maximum point? We have a two steps approach:

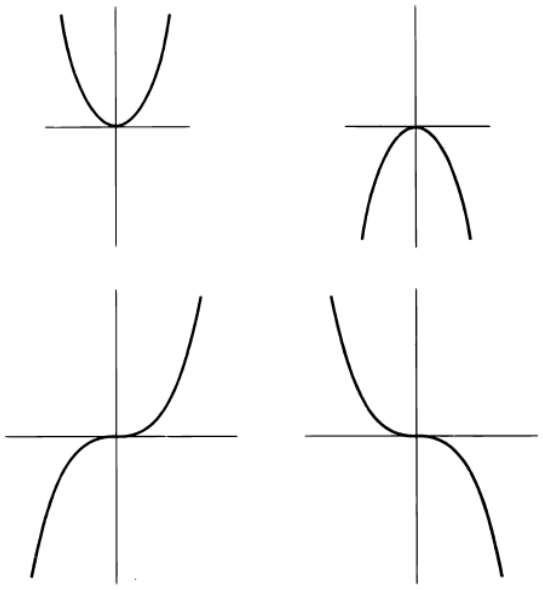
1. **First Order Condition (FOC)**: take the derivative of a function and find the point/s where it is equal to zero (known as *critical* point/s and denoted as x^*)
2. **Second Order Condition (SOC)**: verify that the function is *concave* at the critical point by checking if $f''(x^*) \leq 0$.

Why are we interested in the point where the derivative is null? Because at a maximum, a function is neither increasing nor decreasing, so its derivative must be null (or undefined or the function must be at an endpoint of the domain). Graphically:



[2] In economics we usually work with differentiable functions, but you might encounter the absolute value function or piecewise functions.

Why check for concavity? To ensure that the critical point is a maximum rather than a minimum or a saddle point. The following plot shows various critical points, but only the concave curve in the top-right has a maximum at the critical point:



Given we find a critical point at x^*, a negative second derivative $f''(x^*)$ will confirm it is a maximum, if $f'' \leq 0$ throughout an interval $I \in \mathbb{R}$, x^* will be a global maximum in I, and it will be unique if $f'' < 0$ throughout. Remember, however, that in cases where functions have several critical points, points at which the function is not differentiable, or when the domain includes boundary points, it may be necessary to check the value of the function throughout all these key points to verify which one is the global maximum. We will see more on these in the next sub-sections.

Example: maximization of

$$\max_{c_1}(-(c-5)^2+3)$$

We take the FOC and the SOC:

$$\text{FOC: } -2(c-5)=0 \Rightarrow c=5$$
$$\text{SOC: } -2<0$$

There is a maximum at $c = 5$. Since the function is defined in all $\mathbb{R}$ and the SOC holds strictly in all the domain, the max is global. □

The process to find extreme points with multivariate functions ($f : \mathbb{R}^n \to \mathbb{R}$) is very similar, intuitively, though with a couple of additional tweaks. Let's first revisit the two-step procedure to find local maxima based on the derivatives of $f(x_1, ..., x_n)$.

1. **First Order Conditions**: take all the partial derivatives of the function and find the interior points in which they are all equal to zero (so-called *critical* points)

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^*) = 0 \text{ for } i = 1, ..., n$$

2. **Second Order Conditions**: At this point, we assume f is twice differentiable at $\mathbf{x}^*$. We need to study the second derivatives of f to ensure that $\mathbf{x}^*$ is a maximum. We can represent the set of second order derivatives of a function in a matrix that we call the Hessian:

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_2^2}(\mathbf{x}) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{x}) & \dots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}) \end{bmatrix}.$$

   Now, while in the univariate case we needed to simply check for the sign of the second derivative, now we will need to verify if the Hessian is *negative definite* or *negative semi-definite*. These are two properties that we will study in later chapters. For now suffices to note that, these properties speak about the concavity of a function. Let's finally state how these properties help us identify if $\mathbf{x}^*$ is a maximum:

   - If the Hessian of the Function evaluated at the critical point, $\nabla^2 f(\mathbf{x}^*)$, is *negative definite*, the critical point is a strict local maximum (sufficient condition).[a]
   - If the Hessian $\nabla^2 f(\mathbf{x}^*)$ is *negative semidefinite*, then the critical point ***may*** be a local maximum. (necessary condition).[b]
   - If the Hessian $\nabla^2 f$ evaluated *in a ball* around a critical point is *negative semidefinite*, then the critical point is a local maximum.

[a] This is sufficient because, if satisfied, guarantee that a point is a maximum, but this proof does not go both ways. For example, it is not true that all local maxima have negative definite Hessian matrices.

[b] This is necessary because it must be satisfied for a point to be a maximum. However, satisfying these conditions alone does not guarantee that the point is an extremum; it only indicates potential candidates, the point could also be a saddle or a flat point.

Again, if the conditions above are fulfilled, we got ourselves a *candidate* for a global maximum. To confirm the golbal max, we need to evaluate the function at all critical points, at points where f is non-differentiable points and at the boundaries. In cases were the function is always differentiable, and the domain open, we can focus on the critical points. One way to confirm if a critical point is a global max in that context is if (1) the domain is convex, and the Hessian is negative semidefinite across the entire domain.

2.4 Constrained optimization: Lagrange approach

We can now move on to optimization under equality constraints. To gain intuition, we will focus in functions with two variables in the real numbers space ($f : S \subset \mathbb{R}^2 \to \mathbb{R}$). We want to find solutions for problems of the type

$$\begin{aligned} \text{maximize} \quad & f(x,y) \\ \text{subject to} \quad & g(x,y) = 0 \end{aligned}$$

One handy tool to approach such maximization problems is the Lagrangian. A lagrangian is a way reformulate the constraint problem into an unconstraint one by including a new "artificial" variable, λ, called the Lagrange multiplier.

Proposition 1: Lagrange multiplier condition

Let $f, g : S \subset \mathbb{R}^2 \to \mathbb{R}$ be continuously differentiable on the interior of S, and suppose the interior point $(x^*, y^*) \in S$ solves

$$\max_{x,y} f(x,y) \quad \text{s.t. } g(x,y) = 0.$$

Assume $g_1'(x^*, y^*) \neq 0$ or $g_2'(x^*, y^*) \neq 0$. Then there exists a unique $\lambda \in \mathbb{R}$ such that

$$\begin{aligned} f_1'(x^*, y^*) - \lambda g_1'(x^*, y^*) &= 0, \\ f_2'(x^*, y^*) - \lambda g_2'(x^*, y^*) &= 0 \\ g(x^*, y^*) &= 0. \end{aligned}$$

Consequently, (x^*, y^*) is a critical point of the Lagrangian

$$\mathcal{L}(x, y, \lambda) = f(x,y) - \lambda\big(g(x,y)\big).$$

We deduce from the result above that the following recipe may be used to solve a maximization problem of the type we are considering:

Recipe: Lagrangian Application

Let f and g be functions of two variables defined on a set S, and suppose they are continuously differentiable on the interior of S. If the problem

$$\max_{x,y} f(x,y) \quad \text{subject to } g(x,y) = 0$$

has a solution, it may be found by the following steps:

1. **State the Lagrangian**:

$$\mathcal{L}(x, y, \lambda) = f(x,y) - \lambda\big(g(x,y)\big).$$

2. **Find interior candidates (regular case)**: Find all values of (x, y, λ) such that
   - (x, y) is an interior point of S, and

- the first-order conditions hold (i.e. the partial derivatives of $\mathcal{L}(x, y, \lambda)$ equal zero):

$$\begin{cases} f_x(x,y) - \lambda g_x(x,y) = 0 \\ f_y(x,y) - \lambda g_y(x,y) = 0 \\ g(x,y) = 0 \end{cases}$$

3. **Find abnormal candidates (rare case)**: Find all points (x, y) such that

$$\begin{aligned} g_x(x,y) &= 0 \\ g_y(x,y) &= 0 \\ g(x,y) &= 0 \end{aligned}$$

(In most problems, there are no such points. In particular, if g is linear, this step can be skipped.)

4. **Find boundary candidates (if S has a boundary)**: Find all points on the boundary of S that satisfy both

$$\begin{aligned} g(x,y) &= 0 \\ (x,y) &\in \text{boundary of } S \end{aligned}$$

and solve the constrained maximization problem on this subset. We can forget about these type of candidates if S is open.

5. **Compare values**: Evaluate $f(x, y)$ at all the candidate points found in steps 2–4. The point(s) where f is largest are the solutions to the original problem.

While sometimes we need to go through all the steps of the recipe above, there is occasions where we can take a shortcut and be sure an interior candidate (x^*, y^*) is the solution to the maximization problem. Say we found a solution to the FOCs (x^*, y^*, λ^*). If, additionally to the previous assumptions, we have that S is convex and the Lagrangian is concave with respect to (x, y), once we fix λ^* (that is, $\mathcal{L}(x, y, \lambda^*)$ is concave), then (x^*, y^*) solves $\max f(x, y)$ subject to $g(x, y) = 0$. This saves us some steps to find *a* solution. Notice that there might be more than one solution to the problem.

Example: Let's solve

$$\begin{aligned} \text{maximize} \quad & -x - 2y^2 \\ \text{subject to} \quad & x + y = 1 \end{aligned}$$

First, note that we are lucky enough that the objective function $f(x, y) = -x-2y^2$ and the constrain $g(x, y) = x+y-1$ are both continuously differentiable in their domain, $\mathbb{R}^2$. Therefore, we can apply the recipe above to try and find a solution:

1. Construct the Lagrangian

$$\mathcal{L}(x, y, \lambda) = -x - 2y^2 - \lambda(x + y - 1)$$

2. Find the interior candidates, by taking FOC:

$$\frac{\partial \mathcal{L}}{\partial x}(x^*, y^*, \lambda^*) = -1 - \lambda = 0$$
$$\frac{\partial \mathcal{L}}{\partial y}(x^*, y^*, \lambda^*) = -4y - \lambda = 0$$
$$\frac{\partial \mathcal{L}}{\partial \lambda}(x^*, y^*, \lambda^*) = -(x + y - 1) = 0$$

   From the first equation we get $\lambda = -1$, from the second $y = \frac{1}{4}$, from the last one $x = \frac{3}{4}$. Therefore the critical point is at $(x, y, \lambda) = (\frac{1}{4}, \frac{3}{4}, -1)$

3. There is not abnormal candidates, since, for example, $g_x(x, y) = 1 \neq 0$, $\forall (x, y) \in \mathbb{R}^2$

4. The domain $\mathbb{R}^2$ has no boundary points, and therefore no boundary candidates.

5. The only candidate for a maximum is $(x, y) = (\frac{1}{4}, \frac{3}{4})$, so this is the maximum point.

On the interpretation of the Lagrange multiplier: Look at the FOC of the lagrangian with respect to x, $f_x(x^*, y^*) - \lambda g_x(x^*, y^*) = 0$. We can rearrange it and obtain

$$\lambda = \frac{f_x(x^*, y^*)}{g_x(x^*, y^*)}$$

How can we interpret this expression? The denominator represents the marginal increase in the restriction function, when we marginally increase x. The numerator represents the marginal increase in the objective function when we increase x. Therefore, we can interpret λ as the marginal increase in the objective function when we loosen a bit the restriction, and let x increase a bit. That is, if we allow g to increase by one (infinitesimal) unit, x can increse by $\frac{1}{g_x}$. In turn, this increase in x leads to an increase in our objective function of $\lambda = \frac{f_x}{g_x}$.

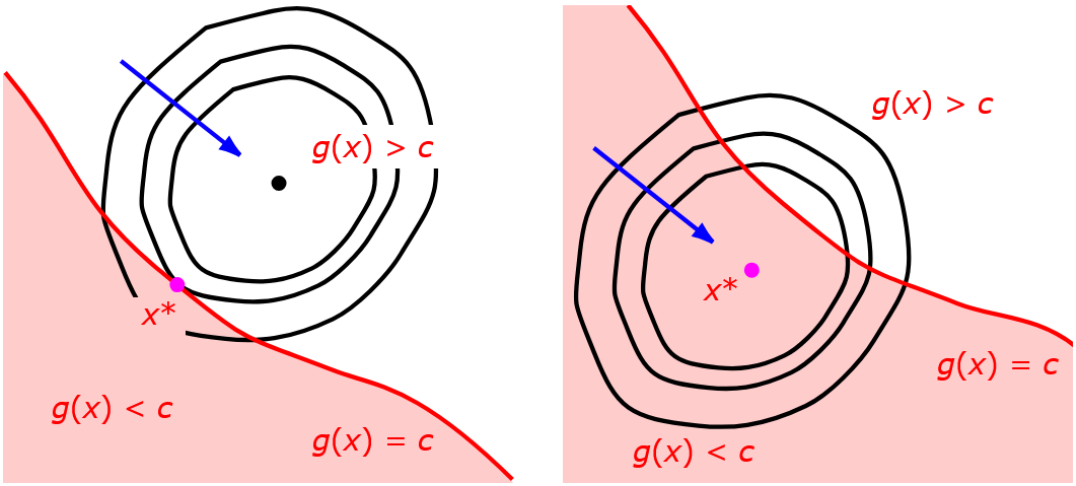
This interpretation of the Lagrange multiplier will be useful in the next subsection, in which we will introduce inequality constraints. Meanwhile, you can try to reason how can we interpret the cases when $f_x = 0 \rightarrow \lambda = 0$, or when $g_x = 0 \rightarrow \lambda = \infty$

2.5 Constrained optimization: Karush-Kuhn-Tucker

In this subsection we will consider optimization problems with inequality constraints. We will approach them through a powerful tool to find solutions to such problems: The Karush-Kuhn-Tucker (KKT) Conditions, named after three mathematicians that established them. First, we will get an intuitive understanding of these conditions. Second, I will briefly discuss the conditions **necessary** for KKT to help us find solutions to a problem, and when KKT is **sufficient** to find a solution. Now, in this lecture we will focus on the bivariate case with one restriction:

$$\max_{x,y} f(x, y) \quad \text{subject to } g(x, y) \leq 0. \tag{4}$$

There are two possibilities for the solution of this problem, both illustrated with one example in the figures below. The black closed curves are contours of f; values of the function increase in the direction shown by the blue arrows. The downward-sloping red line is the set of points x satisfying $g(x) = 0$. The set of points x satisfying $g(x) \leq 0$ is the shaded set below and to the left of the line.



In each figure the solution of the problem is the point x^*. In the first figure the constraint **binds** at the solution: loosening the constraint changes the solution. In the second figure, the constraint is **slack** at the solution: a small loosening of the constraint wouldn't change the solution.

Now, let's look at this maximization problems from the lens of their Lagrangians:

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda\big(g(x, y)\big).$$

In the first figure, we have that $f_i(x^*, y^*) \neq 0$, because the maximum is not a critical point. Therefore, from the first order conditions, we have that $f_x(x, y) = \lambda g_x(x, y) \neq 0$, which implies that both λ and g_x are different from zero. Additionally, from the previous subsection we know we can interpret λ as the marginal change in the objective function of relaxing a binding constraint. Intuitively, this number must be positive in a binding solution, so that $\lambda > 0$. In other words, it makes sense that the objective function was increasing, as it hit the constraint, making the Lagrange multiplier positive. To recap, in a binding solution (x^*, y^*) we can expect:

1. $g(x^*, y^*) = 0$, by definition
2. $f_x(x^*, y^*) > 0$ or $f_y(x^*, y^*) > 0$ or both (objective function must be growing while hitting the constraint)
3. $\lambda > 0$

Following a similar reasoning, we find that in a slack solution, such as in the second figure, we have:

1. $g(x^*, y^*) < 0$, by definition
2. $f_x(x^*, y^*) = f_y(x^*, y^*) = 0$
3. $\lambda = 0$ (Follows from the two points above, plus the FOCs of the Lagrangian)

Notice the interpretation of the Lagrangian multiplier in this case: a marginal shift in the inequality constraint would not change the solution value of the objective function. We may combine the two cases by writing the solution conditions as

$$\begin{aligned} L_i'(x^*) &= 0 && \text{for } i = 1, \ldots, n, \\ \lambda &\geq 0, \\ g(x^*) &\leq 0, \end{aligned}$$

and either $\lambda = 0$ or $g(x, y) = 0$. Now, the product of two numbers is zero if and only if at least one of them is zero, so we can alternatively write the conditions as

$$\begin{aligned} L_i'(x^*) &= 0 && \text{for } i = 1, \ldots, n, \\ \lambda &\geq 0, \\ g(x^*) &\leq 0, \\ \lambda\left[g(x^*) - c\right] &= 0 \end{aligned}$$

Note that the conditions do not rule out the possibility that both $\lambda = 0$ and $g(x^*) = 0$.

The argument above suggests that **if x^* solves the problem** and the constraint satisfies a regularity condition, **then x^* must satisfy these conditions**. In other words, for a point (x^*, y^*) to be a solution to the problem, it must satisfy the above conditions.

Some notes on the applicability of KKT: It is important to realise that the KKT conditions do not help us solve *every* problem. They are guaranteed to hold at a solution only under certain assumptions. Formally, we need a *regularity condition* (a "constraint qualification") and, for sufficiency, some shape restrictions on f and the g's. In this course we will not cover all possible cases, but you should know a few easy ones:

- **For necessity**: if f and the g_j's are differentiable on an open set containing the feasible region, and each active constraint has a nonzero gradient at the optimum, then any local (and hence global) maximum must satisfy KKT.
- **For sufficiency (global optimality)**: if the feasible set is convex (e.g. all g_j are convex) and f is concave *on the feasible set*, then any point satisfying the KKT conditions is a global maximum

These are not the only possibilities, but they are common in economic applications. Outside these cases, KKT can still be applied, but one must check the relevant constraint qualification and be aware that the conditions might identify candidate points that are not actual optima.

Example: Let's solve

$$\begin{aligned} \text{maximize} \quad & -x^2 - y^2 + 4x + 6y \\ \text{subject to} \quad & 0 \geq 2x + y - 3 \end{aligned}$$

First, I can let you know (and in this case, you could even check yourself) the problem satisfies the necessity and sufficiency conditions we just discussed, so we can confidently use the KKT conditions to find its solution.

1. Let's start by setting up the Lagrangian:

$$\mathcal{L}(x, y, \lambda) = -x^2 - y^2 + 4x + 6y - \lambda(2x + y - 3)$$

2. Now we state the KKT conditions:

$$\frac{\partial L}{\partial x} = -2x + 4 - 2\lambda = 0 \tag{5}$$
$$\frac{\partial L}{\partial y} = -2y + 6 - \lambda = 0 \tag{6}$$
$$g(x, y) = 2x + y - 3 \leq 0 \tag{7}$$
$$\lambda \geq 0 \tag{8}$$
$$\lambda(2x + y - 3) \geq 0 \tag{9}$$

3. Let's find points satisfying all the conditions above:

   - Say $\lambda = 0$, from the FOCs we obtain $x = 2$, and $y = 3$. However, this values do not satify the contraint condition $g(x, y) \leq 0$. Not a solution.
   - Say $g(x, y) = 0$, then $y = 3 - 2x$, which together with $\frac{\partial L}{\partial y}$ gives us $\lambda = 4x$. Plug the latter in $\frac{\partial L}{\partial x}$ to obtain $x = \frac{2}{5}$. Then it follows, $y = \frac{11}{5}$ and $\lambda = \frac{8}{5}$. Therefore, $(x, y) = (\frac{2}{5}, \frac{11}{5})$ is the solution to the problem.

3 Dynamic Programming

Dynamic programming is a powerful framework for solving intertemporal decision problems. That is, situations where today's choices shape tomorrow's opportunities. In economics, such problems are everywhere: consumers deciding how to smooth consumption over their lifetimes, firms planning investment and production, or governments designing policies under uncertainty. What makes dynamic programming unique is its ability to break a complex, multi-period problem into simpler, recursive steps, allowing us to focus on one decision at a time while accounting for optimal future behavior. We will illustrate this approach through the "cake-eating problem," a stripped-down yet revealing model in which an agent decides how to consume a finite resource over time. Despite its simplicity, it captures the core trade-off between present enjoyment and future well-being, and serves as a first-step towards understanding richer economic applications.

3.1 Cake Eating Problem: Sequential vs. Recursive Formulation

Consider the following problem. You have a cake of size W_1. At each period t, you can choose how much of the cake to eat c_t. That leaves $W_{t+1} = W_t - c_t$ for next period. Let's say you have T periods to eat the cake, after which, the cake expires, or you die, it's the same. Eating cake today gives you a utility flow $u(c)$, and you discount future utility by β, with $0 < \beta < 1$. We assume u is differentiable, strictly increasing, strictly concave. We also assume that $lim_{c \to 0} u'(c) = \infty$. We can present the problem the following way, sometimes called the **sequence problem**:

$$max \quad \sum_{t=1}^{T} \beta^{t-1} u(c_t) \tag{10}$$

$$\begin{aligned} s.t \quad W_{t+1} &= W_t - c_t \quad \text{for } t = 1, 2, ... \\ c_t &\leq W_t \qquad \text{for } t = 1, 2, ... \\ 0 &\leq W_{T+1} \end{aligned}$$

As we have learned in previous chapters, we could set up a Lagrangian and take first order conditions. These would tell us $\lambda = u'(c_1) = \beta u'(c_2) = \beta^2 u'(c_3) = ...$ From this result, we can take the general expression

$$u'(c_t) = \beta u'(c_{t+1})$$

Which is generally refered to as **Euler equation**. The Euler equation says that at the optimum, the marginal utility you give up by reducing consumption by a tiny amount today is exactly balanced by the discounted marginal utility you gain from consuming that amount tomorrow. If marginal utility today were higher, you would want to shift some consumption from tomorrow to today; if it were lower, you would postpone consumption. This balance ensures that no reallocation of cake across periods can make you better off.

The Euler equation tells us how to spread the cake across periods so that the marginal utility trade-off is balanced at each step. The resource constraint $\sum_{t=1}^{T} c_t \leq W_1$ adds a simple but powerful condition: in the very last period, you eat whatever cake is left, because there is no point saving for period $T + 1$. Working backwards, this means the amount available in period T must be exactly what you plan to consume then, and similarly for earlier periods. Combining this "eat it all in the end" logic with the Euler equations means the initial cake size W_1 pins down the exact consumption path for every period.

We could summarize the solution to the (finite horizon) cake eating problem in the **value function** $V_T(W_1)$. That is, the solution value of problem (12) when the initial cake has size W_1, and the horizon is of T periods. Notice that, from the first order conditions of the sequence problem, we know that an increase in the size of the cake leads to an increase in lifetime utility equal to the marginal utility in any period:

$$V_T'(W_1) = \lambda = \beta^{t-1} u'(c_t)$$

Given that, in the optimum, marginal utility is equalized across periods, we know increasing the cake size a little bit should increase utility by that amount. Therefore this value function has the power to summarize the lifetime value of the cake, without the need of knowing exactly when the cake is eaten, we just need to know that it will be done optimally.

Now, say we want to add a period $t = 0$ to the problem, with an initial cake of size W_0. One way to solve this augmented problem is to set up again the sequence problem as we did before in (12), and solve the problem for $T+1$ periods. This sounds pretty boring to me. The dynamic programming approach provides a way to make our life easier. We can turn the T+1 period problem into a two-period problem by using the value function V_T, and a small rewrite of the objective function:

$$max \quad u(c_0) + V_T(W_1) \tag{11}$$

$$s.t \quad W_1 = W_0 - c_0$$

In this formulation, the choice of consumption in period 0 determines the size of the cake that will be available starting in period 1, W_1. Now, instead of choosing a sequence of consumption levels, we just find c_0. Once c_0 and thus W_1 are determined, the value of the problem from then on is given by $V_T(W_1)$. This function completely summarizes optimal behavior from period 1 onward. For the purposes of the dynamic programming problem, it does not matter how the cake will be consumed after the initial period. All that is important is that the agent will be acting optimally and thus generating utility given by $V_T(W_1)$. This is the **principle of optimality**, due to Richard Bellman, at work. With this knowledge, an optimal decision can be made regarding consumption in period 0.

3.2 Infinite Cake Eating

Now, consider we expand the horizon of the cake eating problem until infinity, so that the cake never expires, and we never die.

$$max \quad \sum_{t=1}^{\infty} \beta^{t-1} u(c_t) \tag{12}$$

$$\begin{aligned} s.t \quad W_{t+1} &= W_t - c_t \quad \text{for } t = 1, 2, ... \\ c_t &\leq W_t \qquad\quad \text{for } t = 1, 2, ... \end{aligned} \tag{13}$$

While an infinite live may sound good at first, solving our problem now involves finding an infinite consumption path $\{c_t\}_{t=1}^{\infty}$ (and eating a lot less cake everyday). To make this less overwhelming, let's formulate this as a dynamic programing problem:

$$V(W) = max \quad u(c) + \beta V(W - c) \tag{14}$$

Here $u(c)$ is the utility from eating cake in the current period, and $V(W)$ is the value of the infinite horizon problem starting with a cake of size W. In this problem, we call W the **state variable**. The state completely summarizes all information from the past or present needed for the forward-looking problem. The **control variable** is the one being chosen, this is current consumption c in our case. Finally, the **transition equation** $W' = W - c$ tells us how the state today and tomorrow are related (We will use primes, such as W' to denote the future values of a variable).

In the finite-horizon cake-eating problem, we can solve the value function by working backwards from the last period. This is because we know exactly what happens in the final step: the agent eats whatever cake is left. This "terminal condition" lets us pin down the value function in the previous period, then the one before that, and so on until we reach the first period.
The loss of the terminal condition puts us at a disadvantage to solve the problem. However, the infinite horizon has

an important upside: Now the problem is **stationary**. That is, because the horizon of the problem doesn't change from one period to the other, the value function V is the same on both sides of (14). That is, while before we had one value function for each horizon $V_1, ..., V_T$, now we have only one. This makes (14) a **functional equation**, frequently named the Bellman Equation, after one of the originators of dynamic programming. The unknown in the Bellman equation is the value function itself: the idea is to find a function $V(W)$ that satisfies this condition for all W.

A very important property of this problem is that all information about the past that bears on current and future decisions is summarized by W, the size of the cake at the start of the period. Whether the cake is of this size because we initially have a large cake and can eat a lot of it or a small cake and are frugal eaters is not relevant. All that matters is that we have a cake of a given size.

Now, say we find a solution to the functional equation. In other words, we find V. Then we can propose there is an implicit function $c = \phi(W)$ that solves for $max \;\; u(c) + \beta V(W - c)$. This function, provides a mapping between states to actions, and is a solution to the infinite cake eating problem, and will satisfy the Euler equation:

$$u'(\phi(W)) = \beta u'(\phi(W - \phi(W))) \quad \text{for all} \;\; W$$

4 Matrix Algebra

Matrices are fundamental objects in mathematics and economics. They are simply rectangular arrays of numbers arranged in rows and columns, and they provide a compact way to represent and manipulate systems of equations, linear transformations, and data. Just like numbers, matrices can be added, subtracted, and multiplied, though some of their operations follow different rules.

Given $m, n \in \mathbb{N}$ a matrix of order $m \times n$ is a table of numbers with m rows and n columns as displayed below

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \ddots & & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & & \ddots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} \tag{15}$$

where each element a_{ij} is a number and is called the (i, j) entry of the matrix.

The standard is to use the first subscript of the entries of a matrix to the row, and the second subscript j to denote the column. Usually we use capital letters to refer to matrices and we write $A_{m\times n}$ to denote a matrix of order $m \times n$. The **main diagonal** of a matrix is made up by the entries a_{ii} with $i \in \{1, ..., m\}$.

Example: An example of a matrix of order 3×4 is

$$\begin{bmatrix} 1 & 4 & 5 & 8 \\ 34 & 2.3 & 9 & 10 \\ \frac{2}{3} & 0 & 1 & 0 \end{bmatrix}$$

where the main diagonal is made up by $\{1, 2.3, 1\}$ □

We will know study specific types of matrices:

- A **vector** is a matrix of the form

$$\begin{bmatrix} a_1 \\ \cdots \\ a_m \end{bmatrix} \text{ or } \begin{bmatrix} a_1, & \cdots & a_n \end{bmatrix} \tag{16}$$

  where the first is called column vector and the second row vector.

- A **square matrix** is a matrix where the number of rows is equal to the number of columns. That is a matrix $A_{m\times n}$ is square if $m = n$.

  Example:

$$B_{2\times 2} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ and } C_{3\times 3} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 4 & 5 & 7 \end{bmatrix}$$

□

- A **diagonal matrix** is a matrix in which the entries outside the main diagonal are all zero

  Example:

$$A_{2\times 2} \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \text{ and } B_{3\times 3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

□

- The **zero matrix** is a matrix of zeros

  Example:

$$\mathbf{0}_{2\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \mathbf{0}_{3\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

□

  Note that the zero matrix is usually denoted by a 0 in bold. If the subscpit is of the type $m \times n$ you have a zero matrix of order $m \times n$, otherwise the subscript could simply be a single natural number ($\mathbf{0}_m$) to indicate a zero vector.

- The **identity matrix** is a square matrix in which all the elements of principal diagonals are one, and all other elements are zeros

  Example:

$$\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

□

  Note that it is denoted by the notation $\mathbf{I}_n$ or simply $\mathbf{I}$.

- A **upper-triangular matrix** is a matrix in which all entries below the diagonal are 0.

Example:

$$B_{2\times 2} = \begin{bmatrix} 1 & 2 \\ 0 & 4 \end{bmatrix} \text{ and } C_{3\times 3} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 7 \end{bmatrix}$$

□

- A **lower-triangular matrix** is a matrix in which all entries above the diagonal are 0.

  Example:

$$B_{2\times 2} = \begin{bmatrix} 1 & 0 \\ 3 & 4 \end{bmatrix} \text{ and } C_{3\times 3} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 4 & 5 & 7 \end{bmatrix}$$

□

4.1 Fundamentals of Matrix algebra

Addition: It is possible to add two matrices if they are of the same size, that is if they have the same number of rows and columns. Their sum is a new matrix of the same size as the two matrices being added.

How does addition work? The (i, j)th entry of the sum matrix is simply the sum of the $(i, j)th$ entried of the two matrices being added:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & b_{ij} & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & a_{ij}+b_{ij} & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix}$$

Example:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix} = \begin{pmatrix} 1+1 & 2+1 & 3+1 \\ 4+2 & 5+2 & 6+2 \\ 7+3 & 8+3 & 9+3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 4 \\ 6 & 7 & 8 \\ 10 & 11 & 12 \end{pmatrix}$$

□

Subtraction: It is possible to subtract two matrices if they are of the same size, that is if they have the same number of rows and columns. Their subtraction is a new matrix of the same size as the two matrices being added.

How does subtraction work? The (i, j)th entry of the sum matrix is simply the subtraction of the $(i, j)th$ entries of the two matrices:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} - \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & b_{ij} & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}-b_{11} & \cdots & a_{1n}-b_{1n} \\ \vdots & a_{ij}-b_{ij} & \vdots \\ a_{m1}-b_{m1} & \cdots & a_{mn}-b_{mn} \end{pmatrix}$$

Example:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} - \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix} = \begin{pmatrix} 1-1 & 2-1 & 3-1 \\ 4-2 & 5-2 & 6-2 \\ 7-3 & 8-3 & 9-3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 3 & 4 \\ 4 & 5 & 6 \end{pmatrix}$$

□

Scalar multiplication: It is possible to multiply a scalar by a ordinary number. The product of a matrix A and a scalar c, denoted by cA, is the matrix created by multiplying each entry of A by c.

$$cA = c\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} ca_{11} & \cdots & ca_{1n} \\ \vdots & ca_{ij} & \vdots \\ ca_{m1} & \cdots & ca_{mn} \end{pmatrix}$$

Example:

$$2A = 2\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \\ 14 & 16 & 18 \end{pmatrix}$$

□

Matrix Multiplication: It is also possible to multiply two matrices, but not all matrices can be multiplied and the order of multiplication does matter. It is possible to multiply two matrices A and B, only if

$$\text{the number of columns of } A = \text{to the number of rows of } B$$

In other words if, given two matrices $A_{m\times n}$ and $B_{p\times k}$, it is possible to have the product AB only if $n = p$ or to have the product BA only if $k = m$.

What is the matrix product? Each entry of the matrix product $A_{m\times n}B_{n\times k}$ can be obtained by multiplying the ith row of A and the jth column of B as follows:

$$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \cdot \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

Example:

$$\begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{pmatrix}\begin{pmatrix} 0 & 3 \\ 6 & 9 \end{pmatrix} = \begin{pmatrix} 1\cdot 0 + 2\cdot 6 & 1\cdot 3 + 2\cdot 9 \\ 4\cdot 0 + 5\cdot 6 & 4\cdot 3 + 5\cdot 9 \\ 7\cdot 0 + 8\cdot 6 & 7\cdot 3 + 8\cdot 9 \end{pmatrix} = \begin{pmatrix} 0 + 12 & 3 + 18 \\ 0 + 30 & 12 + 45 \\ 0 + 48 & 21 + 72 \end{pmatrix} = \begin{pmatrix} 12 & 21 \\ 30 & 57 \\ 48 & 93 \end{pmatrix}$$

□

The identity matrix has the following property: if you multiply a matrix with the identity matrix you get the original matrix. That is

$$AI = A \text{ and } IA = A$$

An idempotent matrix is a square matrix that multiplied by itself gives itself ($B \cdot B = B$ such as the identity matrix

Example: an idempotent matrix is

$$\begin{pmatrix} 5 & -5 \\ 4 & -4 \end{pmatrix}$$

□

Laws of Matrix Algebra The following are some of the fundamental properties of matrix operations that hold for all matrices of compatible dimensions:

- **Associative Laws**:

$$(A+B)+C = A+(B+C)$$
$$(AB)C = A(BC)$$

- **Commutative Law for Addition**:

$$A+B = B+A$$

- **Distributive Laws**:

$$A(B+C) = AB+AC$$
$$(A+B)C = AC+BC$$

Example: Given A,B,C square matrices with the same number of rows, simplify the following expression using matrix algebra rules:

$$A(BC+\alpha I) - B(2AC - 3\beta I)$$

and arrive at

$$(AB-2BA)C + \alpha A + 3\beta B$$

Follow the following steps:

1. Distribute A and B:
$$ABC + \alpha A - 2BAC + 3\beta B$$
2. Combine like terms involving AC:
$$ABC - 2BAC + \alpha A + 3\beta B$$
3. Factor out common terms:
$$(AB-2BA)C + \alpha A + 3\beta B$$
4. Final simplification:
$$(AB-2BA)C + \alpha A + 3\beta B$$

□

Transposition: The transpose of a $m \times n$ matrix A is the $n \times m$ matrix obtained by interchanging the rows and columns of A: the first row of A becomes the first column of the transpose of A and so on.

Example:

$$A = \begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{pmatrix}, \; A' = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \end{pmatrix}$$

□

Note that the transpose of a matrix A is usually denoted as A' or A^T.

The following ***rules*** apply to the transpose:

$$\begin{aligned}(A+B)' &= A'+B'\\ (A-B)' &= A'-B'\\ (A')' &= A\\ (cA)' &= cA'\\ (AB)' &= B'A'\end{aligned}$$

Example: Given A,B,C square matrices with the same number of rows, simplify the following expression using matrix algebra rules:

$$\left((AB)'C + 2A - 3(B'C+\alpha A)\right)B' - A(C-\beta I)B'$$

and arrive at

$$(2-3\alpha+\beta)AB' - (A + B'A' - 3B')CB'$$

Follow the following steps:

1. Distribute and apply transpose properties:

$$(AB)'C = B'A'C$$

2. Expand and combine like terms:

$$(B'A'C + 2A - 3B'C - 3\alpha A)B' - ACB' + \beta AB'$$

3. Distribute B':

$$B'A'CB' + 2AB' - 3B'CB' - 3\alpha AB' - ACB' + \beta AB'$$

4. Combine like terms:

$$2AB' - ACB' - 3\alpha AB' + \beta AB' + B'A'CB' - 3B'CB'$$

5. Final simplification:

$$(2-3\alpha+\beta)AB' - (A + B'A' - 3B')CB'$$

□

4.2 The determinant of a matrix

Each square matrix has an associated number called the determinant. This number is frequently useful in economics. For example, it is a quick test of whether a system of equations has a unique solution. Also, a nonzero determinant means the matrix is invertible. This matters in many settings: when estimating OLS, the data matrix must be invertible to compute the coefficients (no perfect collinearity); when solving or forecasting with vector autoregressions (VARs), singular matrices make it impossible to distinguish shocks or trace dynamics.

The determinant of a 1×1 matrix is simply the entry of the matrix:

$$det(a) = a$$

Note that the determinant of a matrix A is usually denoted as $|A|$ or $det(A)$. The determinant of a 2×2 matrix is

$$\text{given } A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, |A| = a_{11}a_{22} - a_{12}a_{21}$$

This is equivalent to write

$$|A| = a_{11}|a_{22}| - a_{12}|a_{21}|$$

where the first term on the RHS is the $(1,1)$th entry of A times the determinant of the submatrix obtained by deleting from A the row and columns which containt that entry; the second term is the $(1,2)$th entry times the determinant of th esubmatrix obtained by deleting from A the row and column which contain the entry. The terms alternate in sign.

The determinant of a 3×3 matrix is a bit more complicated to get, but it exploits the same logic mentioned in the paragraph above:

$$\text{given } A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, |A| = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

where the first term on the RHS is the $(1,1)$th entry of A times the determinant of the submatrix obtained by deleting from A the row and columns which containt that entry; the second term is the $(1,2)$th entry times the determinant of the submatrix obtained by deleting from A the row and column which contain the entry;the third term is te $(1,3)$th entry times te determinant of the submatrix obtained y deleting from A the row and column which contain the entry. The terms alternate in sign (or alternatively if $i+j$ of the specific entry is even a plus must be added before the entry, while if it is odd a minus must be added).

The above formula is only one of the possible formula to calculate the determinant of a 3×3 matrix. The general recipe is

1. Select a row or a column of the matrix $A_{i \times j}$
2. Take the first entry of the row/column considered and multiply it by
   - the determinant of the submatrix obtained by deleting from A the row and columns which containt that entry (**such a determinant is called the** (i,j)**th minor of** A)
   - $+1$ if, given the (i,j)th entry, the sum $i+j$ is even or by -1 if the sum is odd (such a $+1/-1$ is called the (i,j)th cofactor of A)
3. Do the same of all the entries of the row/column
4. Sum the above quantities

Again, in econometrics we will be often interested in making sure a matrix is **nonsingular**, that is a matrix has a determinant different from zero.

Properties of the determinant:

- $Det(\mathbf{0}) = 0$
- $Det(I) = 1$
- $Det(A) = Det(A')$
- If A has one row/column of zeros $\rightarrow Det(A) = 0$
- If A has rows/columns linearly dependent $\rightarrow Det(A) = 0$
- If A is diagonal (or triangular), the determinant is the product of the elements in the diagonal
- If A and B are both $n \times n$, then $det(AB) = Det(A) \cdot Det(B) = Det(BA)$
- If A and B are both $n \times n$, then $det(AB) \neq 0 \Leftrightarrow Det(A) \neq 0$ and $Det(B) \neq 0$

4.3 Minors of a Matrix

Minors of order k vs. minors of element a_{ij}

Let B be an $m \times n$ matrix. Let $\tilde{B}$ be any $k \times k$ *square* submatrix obtained by selecting k rows and k columns of B. Then $\det(\tilde{B})$ is called a **minor of B of order k**.

Thus, the order of a minor is the number of rows (or columns) it contains.

Example: Consider the matrix

$$A = \begin{pmatrix} 1 & -2 & -1 & 0 \\ 3 & -6 & -3 & 1 \end{pmatrix}.$$

Possible minors of order 2:

$$\det\begin{pmatrix} 1 & -2 \\ 3 & -6 \end{pmatrix},\ \det\begin{pmatrix} -2 & -1 \\ -6 & -3 \end{pmatrix},\ \det\begin{pmatrix} -1 & 0 \\ -3 & 1 \end{pmatrix},\ \det\begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix},\ \det\begin{pmatrix} 1 & -1 \\ 3 & -3 \end{pmatrix},\ \det\begin{pmatrix} -2 & 0 \\ -6 & 1 \end{pmatrix}.$$

Note that we only consider square submatrices. Thus, for any matrix $A_{m\times n}$, the maximum order of any minor is $k \leq \min(m, n)$. □

Let A a $n \times n$ square matrix with $n \geq 2$ and let an element in it be denoted as a_{ij}, where i denotes the row and j the column. The minor M_{ij} of A is the determinant of the $(n-1) \times (n-1)$ submatrix formed by removing the i-th row and the j-th column. To be more specific, this is called **minors of element a_{ij}**

Take some time to focus on the difference in the definition of the minors of order k and of the minors of element a_{ij}, they are two very similar concepts, but the first applies to any matrix, the second one only to square matrices.

Example: Consider the following matrix

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 3 & 1 \\ 0 & -2 & 2 \end{pmatrix}$$

- $M_{2,1} = Det \begin{pmatrix} 0 & 3 \\ -2 & 2 \end{pmatrix} = 6$
- $M_{3,3} = Det \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} = 3$

□

The product $A_{ij} = (-1)^{i+j} M_{ij}$ is called the **cofactor of the element** a_{ij}.

Example: Consider the following matrix

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 3 & 1 \\ 0 & -2 & 2 \end{pmatrix}$$

- $A_{2,1} = (-1)^3 \cdot Det \begin{pmatrix} 0 & 3 \\ -2 & 2 \end{pmatrix} = -6$
- $A_{3,3} = (-1)^6 \cdot Det \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} = 3$

□

Principal Minors and Leading Principal Minors of a Matrix

Let A a $n \times n$ square matrix. The minors obtained after considering k columns and the *same* k rows from A is called **principal minor** of A of order k.

Example: Consider the following matrix

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 2 \\ 4 & 3 & 6 \end{pmatrix}$$

- The principal minors of order 1: 1,1,6;
- The principal minors of order 2:

$$det \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} = -5$$
$$det \begin{pmatrix} 1 & 4 \\ 4 & 6 \end{pmatrix} = -10$$
$$det \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} = 0$$

- The principal minor of order 3: $det(A) = 0$;

□

Let A be an $n \times n$ square matrix. The **leading principal minor of order** k is the determinant of the $k \times k$ submatrix formed by the first k rows and the first k columns of A (i.e., $A_{1:k,\,1:k}$).

Example: Consider

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 3 & 1 \\ 0 & -2 & 2 \end{pmatrix}.$$

- Leading principal minor of order 1: 1.
- Leading principal minor of order 2:

$$\det \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} = 1 \cdot 3 - 0 \cdot 2 = 3.$$

- Leading principal minor of order 3: $\det(A) = -4$.

□

4.4 Rank

The rank of a matrix A is the maximum number of linearly independent rows (or equivalently of linearly independent columns).

Therefore, the rank (r) is a natural number (0,1,2,3...) and is smaller or equal than the number of rows and columns. That is, the rank of a matrix $A_{m\times n}$ can be $min\{m, n\}$ at most.

Example:
Consider the following matrix

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

note that here the third column is a linear combination of the first two (because if you sum the first two columns you get the third one). We will see that the rank of this matrix is 2. □

We can also use the minors to find the rank of a matrix: Given $A_{m\times n}$, the rank of A is the maximum order of its non null minors.

Example:

$$A = \begin{pmatrix} 1 & 2 & -1 & 4 \\ 3 & 1 & 2 & 2 \\ 4 & 3 & 1 & 6 \end{pmatrix}$$

let's check if there are minors different from zero.

- minors of order 1: yes because at least one element is different from 0
- minors of order 2: yes, for example $det\begin{pmatrix}1 & 2\\ 3 & 1\end{pmatrix} = -5$
- minors of order 3: no, in fact

$$Det\begin{pmatrix}1 & 2 & -1\\ 3 & 1 & 2\\ 4 & 3 & 1\end{pmatrix} = 0, Det\begin{pmatrix}1 & 2 & 4\\ 3 & 1 & 2\\ 4 & 3 & 6\end{pmatrix} = 0, Det\begin{pmatrix}1 & -1 & 4\\ 3 & 2 & 2\\ 4 & 1 & 6\end{pmatrix} = 0, Det\begin{pmatrix}2 & -1 & 4\\ 1 & 2 & 2\\ 3 & 1 & 6\end{pmatrix} = 0,$$

The rank of A is 2. □

4.5 Gradient, Jacobian and Hessian Matrices

Now, we will introduce several matrices that characterice the derivatives of multivariate functions: The gradient, the Jacobian, and the Hessian. The **gradient** is the vector of first-order partial derivatives of a **scalar-valued function**, that is, a function $f : \mathbb{R}^n \to \mathbb{R}$ that returns a real number.

The **gradient vector** of f at a point $p = (x_1, \ldots, x_n)$ is

$$\nabla f(p) = \begin{pmatrix}\frac{\partial f}{\partial x_1}\\ \frac{\partial f}{\partial x_2}\\ \vdots\\ \frac{\partial f}{\partial x_n}\end{pmatrix}.$$

Geometrically, the gradient points in the direction of the fastest increase of the function.

The Jacobian is a matrix that holds all first-order partial derivatives of a **vector-valued function**, that is a function f that returns a vector ($f : \mathbb{R}^n \to \mathbb{R}^m$). To understand better, consider the case of $f : \mathbb{R}^2 \to \mathbb{R}^2$

The **Jacobian** matrix vector of a function $f = \begin{pmatrix}h(x,y)\\ g(x,y)\end{pmatrix}$ at a point (p_x, p_y) is

$$J(p_x, p_y) = \begin{pmatrix}h_x(p_x,p_y) & h_y(p_x,p_y)\\ g_x(p_x,p_y) & g_y(p_x,p_y)\end{pmatrix}$$

Note that the Jacobian matrix can also be written as a vector of gradients:

$$Df = J(p_x, p_y) = \begin{pmatrix}h_x & h_y\\ g_x & g_y\end{pmatrix} = \begin{pmatrix}\nabla h(p)^T\\ \nabla g(p)^T\end{pmatrix}$$

Example: Consider the function

$$f(x,y) = \begin{pmatrix}e^x + y\\ x + y^2\end{pmatrix}$$

The Jacobian matrix is

$$J(x,y) = \begin{pmatrix} e^x & 1 \\ 1 & 2y \end{pmatrix}$$

□

4.6 Hessian matrix

The Hessian matrix is the matrix of the second-order partial derivatives of a **scalar-valued function**, that is a function f that returns a scalar value ($f : \mathbb{R}^n \to \mathbb{R}$).

The Hessian matrix of a function f at a point $p = (x_1, \ldots, x_n)$ is

$$Hessian = \nabla^2 f(p) = \begin{pmatrix} \frac{\partial^2 f(p)}{\partial x_1^2} & \frac{\partial^2 f(p)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(p)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(p)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(p)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(p)}{\partial x_2 \partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 f(p)}{\partial x_n \partial x_1} & \frac{\partial^2 f(p)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(p)}{\partial x_n^2} \end{pmatrix}$$

Note that the Hessian matrix is **symmetric** and that sometimes is also denoted as $D^2 f(p)$.

Example: Compute the Hessian of the following function: $f(x,y) = y\,x^3$.

$$\nabla^2 f(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 6xy & 3x^2 \\ 3x^2 & 0 \end{pmatrix}.$$

□

4.7 Matrix Definiteness

Positive and negative definiteness are useful concepts in multivariate optimization when applied to Hessian matrices.

Given a function $f(x)$ and its Hessian matrix H, the matrix H is said to be **negative definite** if $xHx \leq 0$ for all $x \neq 0$, where x is the vector of variables ($x \in \mathbb{R}^n$), and if $x'Hx = 0$ only if $x = 0$, that is $x'Hx$ is zero only in zero. Similarly, it is possible to define positive matrices.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and let H be its Hessian matrix. We say that H is **negative definite** if, for every nonzero vector $x \in \mathbb{R}^n$,

$$x^T H x < 0.$$

In other words, the quadratic form $x^\top Hx$ is strictly negative for all $x \neq 0$, and equals zero only at $x = 0$. Similarly, H is **positive definite** if $x^\top Hx > 0$ for all $x \neq 0$.

In some cases, the condition $xHx = 0$ may occur for some non-zero vectors x. In such situations, the matrix is referred to as **semidefinite**:

- H is **positive semidefinite** if $x'Hx \geq 0$ for all $x \in \mathbb{R}^n$, with $x'Hx = 0$ for $x = 0$ and some $x \neq 0$.
- H is **negative semidefinite** if $x'Hx \leq 0$ for all $x \in \mathbb{R}^n$, with $x'Hx = 0$ for $x = 0$ and some $x \neq 0$.

Why do we need these definitions? Because a function f is

- Strictly concave if and only if its Hessian matrix is negative definite
- Strictly convex if and only if its Hessian matrix is positive definite
- Convex if and only if its Hessian matrix is positive semidefinite
- Concave if and only if its Hessian matrix is negative semidefinite

More concretely, how to determine the definiteness of a function?
Determine matrix definiteness

- The Hessian Matrix is **negative definite** if and only if all *leading principal minors* of the Hessian are *strictly* positive if their order is even and *strictly* negative if their order is odd (or equivalently all eigenvalues of the Hessian are strictly negative)
- The Hessian Matrix is **positive definite** if and only if all *leading principal minors* of the Hessian are *strictly* positive (or equivalently all eigenvalues of the Hessian are strictly positive)
- The Hessian Matrix is **negative semidefinite** if and only if all the *principal minors* of the Hessian are *nonnegative* (≥ 0) if their order is even and *nonpositive* (≤ 0) if their order is odd (or equivalently all eigenvalues of the Hessian are nonpositive and exist at least one eigenvalue which is equal to zero)
- The Hessian Matrix is **positive semidefinite** if and only if all the *principal minors* of the Hessian are *non negative* (≥ 0) (or equivalently all eigenvalues of the Hessian are nonnegative and exist at least one eigenvalue which is equal to zero)

Example: Let's check the definiteness of the following function

$$-x - 2y^2$$

We compute the Hessian

$$\begin{pmatrix} 0 & 0 \\ 0 & -4 \end{pmatrix}$$

- We compute the leading principal minors: the leading principal minor of order 1 is 0, the leading principal minor of order 2 is 0. The matrix is not negative definite or positive definite.
- Let's now compute the other principal minors, we computed all the principal minors of order 2 (equal to 0), and only the principal minors of order 1 are left. They are all 0 except for one which is -4, given that all the even principal minors of the Hessian are nonnegative and all the odd are nonpositive, the Hessian matrix is negative semidefinite.

Given that the hessian is negative semidefinite the function is concave. Note that we could also have equivalently calculated the eigenvalues to determine the definitiveness of a matrix. □
We summarize the above information in the following conditions:

Negative Hessian matrix	$\iff$	All leading principal minors of the Hessian are strictly positive if their order is even and strictly negative if their order is odd	or	All eigenvalues of the Hessian are strictly negative	$\iff$	Function is strictly concave
Positive Definite Hessian matrix	$\iff$	All leading principal minors of the Hessian are strictly positive	or	All eigenvalues of the Hessian are strictly positive	$\iff$	Function is strictly convex
Negative Semidefinite Hessian matrix	$\iff$	All principal minors of the Hessian are nonnegative if their order is even and nonpositive if their order is odd	or	All eigenvalues of the Hessian are nonpositive and exist at least one eigenvalue which is equal to zero	$\iff$	Function is concave
Positive Semidefinite Hessian matrix	$\iff$	All principal minors of the Hessian are nonnegative	or	All eigenvalues of the Hessian are nonnegative and exist at least one eigenvalue which is equal to zero	$\iff$	Function is convex