Induction Course

In

Quantitative Methods

M.Sc. in Financial Management

Department of Accounting and Finance

Athens University of Economics and Business

Leonidas S. Rompolis

September, 2023

LINEAR ALGEBRA

1. Introduction - Notation

A **matrix** is a rectangular array of numbers, such as:

$$\begin{pmatrix} 2 & -5 & 6 \\ 20 & 1 & 0 \end{pmatrix}$$

An item in a matrix is called **entry** or **element**. The example has entries: 2, -5, 6, 20, 1 and 0.

The horizontal and vertical lines in a matrix are called **rows** and **columns**, respectively. To specify a matrix size, a matrix with m rows and n columns is called a $m \times n$ matrix, while m and n are its **dimensions**. The example is 2×3 matrix.

A matrix with one row (a $1 \times n$ matrix) is called a **row vector** and a matrix with one column (a $m \times 1$ matrix) a **column vector**. For example, the first row of the above example

$$\begin{pmatrix} 2 & -5 & 6 \end{pmatrix}$$

is a row vector, while its second column

$$\begin{pmatrix} -5 \\ 1 \end{pmatrix}$$

is a column vector.

In general, a $m \times n$ matrix is represented as:

$$A_{mn} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

The elements of the matrix are represented as:

$$a_{ij}, i = 1, 2, ..., m \text{ and } j = 1, 2, ..., n$$

The subscripts of each element show its position in the matrix. Thus, the element a_{12} lies on the first row and the second column of the matrix.

The **transpose** of a matrix A is another matrix A' created by any one of the following equivalent actions:

- Write the rows of A as the columns of A'
- Write the columns of A as the rows of A'

Formally, the (i,j) element of A is the (j,i) element of A'. Thus, if A is a $m \times n$ matrix then A' is a $n \times m$ matrix.

A **square matrix** has an equal number of columns and rows.

A **zero matrix** has all its elements equal to zero.

A **diagonal matrix** is a square matrix with all the elements equal to zero except those of the main diagonal.

An identity matrix is a diagonal matrix with all the elements of the main diagonal equal to 1, i.e,

$$I_{nn} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

In a **symmetric matrix** the elements are symmetric with respect to the main diagonal. Thus a symmetric matrix is a square matrix with: $a_{ij} = a_{ji}, \forall\, i, j \text{ with } i \neq j$.

MATRICES ARE USEFUL FOR:

- Solving a linear system of equations, or finding whether or not this system has a (unique) solution
- An input-output analysis which describes a production process in macroeconomics
- A multiple regression analysis in econometrics
- A portfolio optimization problem in investments.

2. Matrix Operations

When two or more matrices are **added** or **subtracted**, these matrices should be **compatible**. This means they should have the same dimensions.

The rule of addition and subtraction of a matrix states that the respective elements of the matrices are added or subtracted.

The new matrix that will come up after the addition or subtraction of two or more matrices will have the same dimensions as them.

Properties of matrix addition:

- Commutative law: A + B = B + A
- Associative law: (A + B) + C = A + (B + C)

The multiplication of a matrix with a real number is called **scalar multiplication**.

The rule of scalar multiplication states that the real number should be multiplied with each element of the matrix.

The **inner product** of a $1 \times n$ row vector $a' = (a_1 \quad a_2 \quad \dots \quad a_n)$ and a $n \times 1$ column vector $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ is equal to the number:

$$a'b = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n = \sum_{i=1}^{n} a_i b_i$$

To define the inner product of two vectors the number of the rows of the first vector should be equal to the number of the columns of the second.

Example 1: In a linear regression model the dependent variable is linearly related to several other variables, called the independent variables. Let y_i be the i-th observation of the dependent variable in question and let $x_i' = \begin{pmatrix} x_{i1} & x_{i2} & \dots & x_{iK} \end{pmatrix}$ be the i-th observation of the K independent variables. Then,

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_K x_{iK} + \varepsilon_i$$

where $\beta' = \begin{pmatrix} \beta_1 & \beta_2 & \dots & \beta_K \end{pmatrix}$ are unknown parameters to be estimated and ε_i is an unobserved error term. The above linear regression model can be written as an inner product as:

$$y_i = x_i'\beta_i + \varepsilon_i$$

Example 2: Assume an investor that invested in a portfolio of n stocks. The weights of the stocks in the portfolio are represented by the row vector $w' = \begin{pmatrix} w_1 & w_2 & \dots & w_n \end{pmatrix}$ with $\sum_{i=1}^{n} w_i = 1$. The expected annual returns of the n stocks are represented by the column vector

$r = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix}$. Then, the expected annual return of the portfolio is given by the inner product of vectors w' and r, i.e.,

$$w'r = \sum_{i=1}^{n} w_i r_i$$

If for example the portfolio contains 4 stocks with weights: $w' = \begin{pmatrix} 0.2 & 0.3 & 0.1 & 0.4 \end{pmatrix}$ and expected annual returns: $r = \begin{pmatrix} 0.1 \\ 0.2 \\ 0.15 \\ 0.05 \end{pmatrix}$, then the expected annual return of the portfolio is equal to:

$$w'r = 0.2 \times 0.1 + 0.3 \times 0.2 + 0.1 \times 0.15 + 0.4 \times 0.05 = 0.115$$

Multiplication of two matrices is defined only if the number of columns of the left matrix is the same as the number of rows of the right matrix. If A is a $m \times n$ matrix and B is a $n \times p$ matrix, then their **matrix product** C = AB is the $m \times p$ matrix whose entries are given by the inner product of the corresponding row of A and the corresponding column of B:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + ... + a_{in} b_{nj}$$

For example, the multiplication of two 2×2 matrices A and B yields:

$$AB = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} =$$

$$= \begin{pmatrix} \begin{pmatrix} a_{11} & a_{12} \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix} & \begin{pmatrix} a_{11} & a_{12} \end{pmatrix} \begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix} \\ \begin{pmatrix} a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix} & \begin{pmatrix} a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix} \end{pmatrix} =$$

$$=\begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix}$$

Example 3: Assume an investor that invested in a portfolio with n stocks. The weights of the stocks in the portfolio are represented by the row vector $w'=\begin{pmatrix} w_1 & w_2 & \dots & w_n \end{pmatrix}$ with $\sum_{i=1}^{n} w_i = 1$. The variance-covariance matrix V of the n stocks is a $n \times n$ matrix with the (i,i) element being the annual return variance of the i-th stock and the (i,j) element being the covariance between the i-th and the j-th stock. Then, the variance of the portfolio annual return is given as:

$$w'Vw$$

If $w'=\begin{pmatrix} 0.2 & 0.1 & 0.1 & 0.35 & 0.15 & 0.05 & 0.05 \end{pmatrix}$ and the variance-covariance matrix is equal to:

	ΕΜΠ	ΑΕΓΕΚ	ΑΛΦΑ	ΒΙΟΧΚ	ΕΛΑΙΣ	ΕΤΕ	ΟΤΕ
ΕΜΠ	0.00473	0.00199	0.00262	0.00132	0.00196	0.00362	0.00213
ΑΕΓΕΚ	0.00199	0.00498	0.00142	0.00173	0.00201	0.00144	0.00143
ΑΛΦΑ	0.00262	0.00142	0.00239	0.00152	0.0015	0.00239	0.00159
ΒΙΟΧΚ	0.00132	0.00173	0.00152	0.0036	0.00176	0.00174	0.00141
ΕΛΑΙΣ	0.00196	0.00201	0.0015	0.00176	0.00438	0.00161	0.00117
ΕΤΕ	0.00362	0.00144	0.00239	0.00174	0.00161	0.00351	0.00198
ΟΤΕ	0.00213	0.00143	0.00159	0.00141	0.00117	0.00198	0.00316

Then, the variance of the portfolio return is equal to 0.0022.

Properties of matrix multiplication:

- The commutative law does not hold: $AB \neq BA$

- Associative law: (AB)C = A(BC)
- Distributive law: (A + B) C = AC + BC
- Multiply with the identity matrix: AI = IA = A

Properties of transpose matrices:

- $(A')' = A$
- $(A + B)' = A' + B'$
- $(AB)' = B'A'$
- $(cA)' = cA'$

3. The determinant of a matrix

The determinant is a special number associated with any **square** matrix. The fundamental geometric meaning of a determinant is a scale factor for measure when the matrix is regarded as a linear transformation.

The **Laplace expansion** of a $n \times n$ square matrix A gives the determinant, denoted as $|A|$, of the matrix. This is the sum of n determinants of $(n-1)\times(n-1)$ sub-matrices of A. There are n^2 such expressions, one for each row and column of A.

Define the (i,j) **minor matrix** M_{ij} of A as the $(n-1)\times(n-1)$ matrix that results from deleting the i-th row and the j-th column of A, and C_{ij} the **cofactor** of A as $C_{ij} = (-1)^{i+j}|M_{ij}|$.

Then the Laplace expansion gives the determinant of A as:

$$|A| = a_{i1}C_{i1} + a_{i2}C_{i2} + \ldots + a_{in}C_{in} =$$
$$= a_{1j}C_{1j} + a_{2j}C_{2j} + \ldots + a_{nj}C_{nj}$$

Example 4: The determinant of a 2×2 matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is equal to: $|A| = a_{11}a_{22} - a_{12}a_{21}$.

Example 5: Consider the matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$. The determinant of this matrix can be computed using the Laplace expansion along the first row:

$$|A| = 1\times\begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2\times\begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3\times\begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} =$$
$$= 1\times(-3) - 2\times(-6) + 3\times(-3) = 0$$

Properties of determinants:

- If all the elements of a row or column of a matrix are zero, then the determinant of the matrix is equal to zero
- If B results from A by interchanging two rows or two columns, then $|B| = -|A|$
- If B results from A by multiplying one row or one column with a number c, then $|B| = c|A|$
- If B results from A by adding a multiple of one row to another row, or a multiple of one column to another column, then $|B| = |A|$
- $|cA_{nn}| = c^n |A_{nn}|$
- $|AB| = |A||B|$
- $|A'| = |A|$

Consider a $n \times n$ matrix A and consider that there exists a vector $\lambda' = \begin{pmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_n \end{pmatrix} \neq 0$ for which:

$$\lambda_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + \lambda_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \dots + \lambda_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix} = 0$$

Then the vectors $\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}$ are **linearly dependent.** In this case the determinant of A is equal to zero.[1] This matrix is called **singular**.

This is exactly what happens in Example 5 where the sum of the first and third column is twice the second column.

[1] The same holds when the columns of the matrix are linearly dependent.

4. The inverse of a matrix

The **inverse** of a **square** matrix is a new matrix A^{-1} (with the same dimensions with A) such that:

$$AA^{-1} = A^{-1}A = I$$

The calculation of the inverse of a matrix is required when we need to “move” the matrix from one side of the equation to the other in order to solve a linear system of equations. For example, if $AX = B$, then by multiplying in the left both sides of the equation with A^{-1} we obtain:

$$A^{-1}AX = A^{-1}B \Rightarrow IX = A^{-1}B \Rightarrow X = A^{-1}B$$

The inverse of a matrix can be calculated as:

$$A^{-1} = \frac{1}{|A|} \mathrm{adj}(A)$$

where adj(A) is the **adjoint** matrix of A.

If $C_{ij} = (-1)^{i+j} |M_{ij}|$ defines the cofactor matrix of A, then

$$\mathrm{adj}(A) = C'$$

From the above equation it becomes apparent that the inverse of the matrix is defined only when the determinant $|A| \neq 0$. If this property holds then the matrix is called **invertible** or **non-singular**.

Example 6: The inverse of a 2×2 matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is:

$$A^{-1} = \frac{1}{(a_{22}a_{11} - a_{12}a_{21})}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Properties of the inverse of a matrix:

- $\left(A^{-1}\right)^{-1} = A$
- $\left(cA\right)^{-1} = c^{-1}A^{-1}$
- $\left(A'\right)^{-1} = \left(A^{-1}\right)'$
- $\left(AB\right)^{-1} = B^{-1}A^{-1}$
- $\left|A^{-1}\right| = 1/\left|A\right|$

SYSTEM OF LINEAR EQUATIONS

1. INTRODUCTION – NOTATION

A **system of linear equations** is a collection of linear equations involving the same set of variables.

Example 7: Consider an economic model that describes demand and supply of an asset. We can write that the quantity of the asset that the market demand, denoted as Q_d is:

$$Q_d = a - bP, a, b > 0$$

where P is the price of the asset. The above equation implies that as the price decreases the demand increases. The quantity supplied by the market, denoted as Q_s is:

$$Q_s = -c + dP, c, d > 0$$

Now, price and supply are positively related.

For the market to be in equilibrium we must also impose:

$$Q_s = Q_d \equiv Q$$

These three equations constitute a system of two linear equations with two unknowns that describe demand, supply and the asset price in equilibrium. We can therefore write:

$$Q = a - bP$$
$$Q = -c + dP$$

Solving this system with respect to P and Q we determine the market equilibrium.

In general, a system of linear equations is written as:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

Here $x_1, x_2, \ldots, x_n$ are the unknowns (**endogenous variables**), $a_{11}, a_{12}, \ldots, a_{mn}$ are the **coefficients** of the system and $b_1, b_2, \ldots, b_m$ are the constant terms (**exogenous** variables). This is system with m linear equations and n unknowns.

Every system of linear equations can be written in matrix form as follows:

$$AX = B \tag{1}$$

with

$A_{(m\times n)} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, X_{(n\times 1)} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, B_{(m\times 1)} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ A is the **coefficient matrix**, X is the **solution vector** and B is the **vector of constant terms**.

For example the system of equations that determines the market equilibrium can be written as:

$$\underbrace{\begin{pmatrix} 1 & b \\ 1 & -d \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} Q \\ P \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} a \\ -c \end{pmatrix}}_{B}$$

2. The solution of linear system of equations

A linear system may behave in any one of three possible ways:

- The system has an infinite number of solutions.
- The system has a unique solution.
- The system has no solution (**inconsistency**)

The **column rank** of matrix A is the maximal number of linearly independent columns of A. Likewise, the **row rank** of A is the maximal number of linearly independent rows of A. Since the column rank and the row rank are always equal, they are simply called the **rank** of A.

The rank of a $m \times n$ matrix is at most min(m,n), i.e., $r(A_{mn}) \leq \min(m,n)$. When $r(A_{mn}) = \min(m,n)$ the matrix is said to have **full rank**.

Also define the **augmented matrix** as $C_{m(n+1)} = \begin{pmatrix} A & B \end{pmatrix}$.

How can we calculate the rank of a matrix? To do so, we need to transform it to its **row echelon form**. This form should satisfy the following conditions:

- The first non-zero element in each row, called the leading entry, is 1.
- Each leading entry is in a column to the right of the leading entry in the previous row.
- Rows with all zero elements, if any, are below rows having non-zero elements.

To transform a matrix to its echelon form we need to do a series of elementary row operations as follows:

1. Pivot the matrix.
    - Start with the first row of the matrix if the entry of the first column is different than zero. This entry is known as the **pivot**. Otherwise, interchange it with another row.
    - Multiply each element in the pivot row by the inverse of the pivot, so the pivot equals 1.
    - Add multiples of the pivot row to each of the lower rows, so every element in the pivot column is equal to 0.
2. Repeat step 1.
    - Repeat the procedure from step 1 above, ignoring previous pivot rows.
    - Continued until there are no more pivots to be processed.

Once the row echelon form is found, the rank of the matrix is equal to the number of non-zero rows in its echelon form.

Example 8: Consider the following matrix $A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 7 & 8 \end{pmatrix}$. To transform this matrix to its echelon form we do the following series of elementary row operations:

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 7 & 8 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 2 & 7 & 8 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 3 & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

The last matrix above is the echelon form. It has two non-zero rows so the rank of matrix A is equal to 2, r(A) = 2.

Proposition 1: The linear system of equations (1) has a solution if and only if r(A) = r(C).

Example 9: Assume that $A = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 8 \\ 9 \end{pmatrix}$. Then we can show that r(A) = 1 and r(C) = 2, thus the system does not have a solution.

If $B = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$, then r(C) = 1 and the system has a solution (though not unique).

Proposition 2: The linear system of equations (1) has a unique solution if r(A) = r(C) = n (where n is the number of unknowns).

Consider that m < n (more unknowns than equations). In this case $r(A) \le \min(m, n) = m < n$ and the system cannot have a unique solution.

Proposition 3: If n = m and r(A) = n, then the system (1) has a unique solution.[2]

In this case the matrix A is invertible. This comes from the fact that A is a square matrix and $|A| \neq 0$ (since r(A) = n). Thus the solution of the system is:

$$X = A^{-1}B$$

Example 7 (continued): In this case we have that n = m and r(A) = 2, thus the system has a unique solution. The inverse of A is:

$$A^{-1} = -\frac{1}{d+b}\begin{pmatrix} -d & -b \\ -1 & 1 \end{pmatrix} = \frac{1}{d+b}\begin{pmatrix} d & b \\ 1 & -1 \end{pmatrix}$$

and

---

[2] This comes from Proposition 2. When n = m we can easily prove that r(C) = n.

$$X = \frac{1}{d+b}\begin{pmatrix} d & b \\ 1 & -1 \end{pmatrix}\begin{pmatrix} a \\ -c \end{pmatrix} = \frac{1}{d+b}\begin{pmatrix} ad-bc \\ a+c \end{pmatrix}.$$

Thus, the equilibrium price and quantity are: $P = \frac{a+c}{b+d}$ and $Q = \frac{ad-bc}{b+d}$, respectively.

Example 10: Consider an economic model with two assets 1 and 2. The demand for the two assets is:

$$Q_{d1} = a_0 + a_1 P_1 + a_2 P_2$$
$$Q_{d2} = b_0 + b_1 P_1 + b_2 P_2$$

where P_1 and P_2 are the prices of asset 1 and 2, respectively. Similarly, the supply for the two assets is:

$$Q_{s1} = c_0 + c_1 P_1 + c_2 P_2$$
$$Q_{s2} = d_0 + d_1 P_1 + d_2 P_2$$

In equilibrium,

$$Q_{s1} = Q_{d1} \equiv Q_1$$
$$Q_{s2} = Q_{d2} \equiv Q_2$$

We can write the above problem as a system of linear equations as follows:

$$\underbrace{\begin{pmatrix} 1 & 0 & -a_1 & -a_2 \\ 0 & 1 & -b_1 & -b_2 \\ 1 & 0 & -c_1 & -c_2 \\ 0 & 1 & -d_1 & -d_2 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} Q_1 \\ Q_2 \\ P_1 \\ P_2 \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} a_0 \\ b_0 \\ c_0 \\ d_0 \end{pmatrix}}_{B}$$

Thus, we have a 4×4 system of linear equations. Assume that r(A) = 4, then this system has a unique solution which determines the equilibrium in this two-assets market. To solve this system, one should calculate the inverse of a 4×4 matrix which is not an easy task to do without a computer program. However, we can simplify the above system of equations to give an explicit solution without using a computer. To do so, we will use results of matrix algebra presented in the previous section. First, write the above system as:

$$\begin{pmatrix} I & A_1 \\ I & A_2 \end{pmatrix}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$$

which implies:

$$X_1 + A_1X_2 = B_1$$
$$X_1 + A_2X_2 = B_2$$

Substituting X_1 from the first equation to the second we obtain:

$$(B_1 - A_1X_2) + A_2X_2 = B_2 \Rightarrow (A_2 - A_1)X_2 = B_2 - B_1$$

Thus we have a 2×2 system of equations which can be easily solved, with respect to X_2 as in the previous example. Finally, using one of the above equations we can solve for X_1.

Example 11: **A pure, primitive, or Arrow-Debreu security** is defined as a security that pays $1 at the end of a period if a given state occurs and nothing if any other state occurs. The concept of pure securities allows the decomposition of market securities into portfolios of pure securities. Thus, every market security may be considered a combination of various pure securities. Of course, these securities are not traded so they should be inferred from market traded securities.

Consider two securities j and k, whose payoff table is given below:

Payoff table for securities j and k

Security	State 1	State 2	
j	$10	$20	p_j = $8
k	$30	$10	p_k = $9

We want to determine the prices of the pure securities (1, 0) and (0 ,1). Note that we have two states of nature and two linearly independent market securities, thus the market is **complete**. This means that the pure securities and uniquely determined.

By buying 10 units of (1, 0) and 20 units of (0, 1) we can construct the end-of-period payoff of security j. Similarly, by buying 30 units of (1, 0) and 10 units of (0, 1) we can construct the end-of-period payoff of security k. Since the payoffs of these portfolios equate the payoffs of the two market securities, we can assume that the prices should do so. Thus, we can write:

$$p_1 Q_{j1} + p_2 Q_{j2} = p_j$$
$$p_1 Q_{k1} + p_2 Q_{k2} = p_k$$

Substituting the values given in the table we obtain:

$$p_1 10 + p_2 20 = 8$$
$$p_1 30 + p_2 10 = 9$$

This is a 2×2 system of equations. Solving for p_1 and p_2 we obtain: p_1 = \$0.2 and p_2 = \$0.3.

EIGENVALUES AND EIGENVECTORS

1. POSITIVE DEFINITE MATRICES

A symmetric $n \times n$ matrix A is said to be **positive definite** if $x'Ax > 0$ for all non zero $n \times 1$ vectors x. Similarly, we can define **positive semi-definite** ($x'Ax \geq 0$), **negative definite** ($x'Ax < 0$) and **negative semi-definite** ($x'Ax \leq 0$).

If the matrix A is positive definite then it is invertible.

Positive definite matrices are important for linear regressions analysis and the maximization or minimization of a multivariate function.

2. Eigenvalues and eigenvectors

Let know write the system of equations:

$$Ax = \lambda x$$

where A is a $n \times n$ square matrix, x is $n \times 1$ vector and λ is real number. This system can be also written as:

$$(A - \lambda I)x = 0$$

This is a **homogenous** system of equations (since the vector of constant terms is zero). A trivial solution to the above system (when $A - \lambda I$ is invertible) is X = 0. When, however,

$$|A - \lambda I| = 0 \qquad (2)$$

the system has also **non-trivial** solutions. The determination of these solutions is known as the **eigenvalue problem**. The values of λ that satisfy equation (2) are known as the **eigenvalues** of matrix A. The non-trivial solutions of x are known as the **eigenvectors** of matrix A.

Example 12: Consider the matrix $A = \begin{pmatrix} -5 & 2 \\ 2 & -2 \end{pmatrix}$. Then, to find the eigenvalues we solve:

$$|A - \lambda I| = 0 \Rightarrow \begin{vmatrix} -5-\lambda & 2 \\ 2 & -2-\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 + 7\lambda + 6 = 0$$

The solutions are λ = -6 and λ = -1. These are the eigenvalues of matrix A.

For each eigenvalue there is a corresponding eigenvector. This vector can be found by substituting one of the eigenvalues back into the original equation.

For λ = -6 we have that:

$$\begin{pmatrix} -5+6 & 2 \\ 2 & -2+6 \end{pmatrix} x = 0 \Rightarrow \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} x = 0 \Rightarrow x_1 = -2x_2$$

So, the eigenvector corresponding to λ = -6 is any vector of the form $\begin{pmatrix} -2c \\ c \end{pmatrix}, c \neq 0$.

For λ = -1 we have that:

$$\begin{pmatrix} -5+1 & 2 \\ 2 & -2+1 \end{pmatrix} x = 0 \Rightarrow \begin{pmatrix} -4 & 2 \\ 2 & -1 \end{pmatrix} x = 0 \Rightarrow x_2 = 2x_1$$

So the eigenvector corresponding to λ = -1 is any vector of the form $\begin{pmatrix} c \\ 2c \end{pmatrix}, c \neq 0$.

Eignevalues are directly related to whether a matrix is positive definite. More specifically,

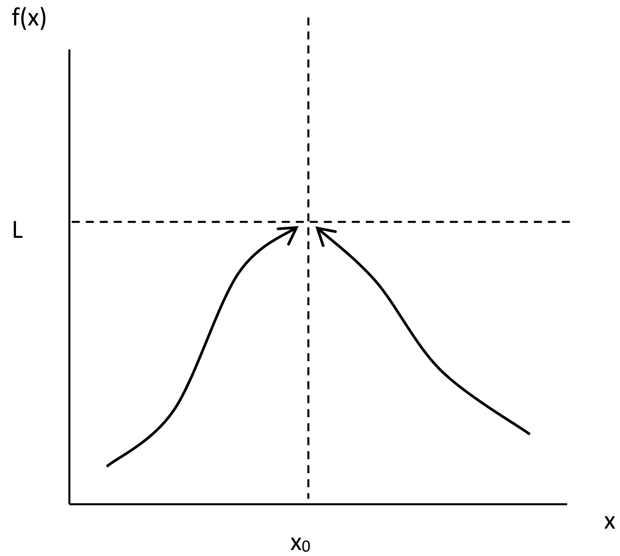
- When all eigenvalues are positive the matrix is positive definite
- When all eigenvalues are negative the matrix is negative definite (as in the Example 11)
- When all eigenvalues are non-negative (for example all are positive and one is equal to zero) the matrix is positive semi-definite
- When all eigenvalues are non-positive (for example all are negative and one is equal to zero) the matrix is negative semi-definite
- When eignevalues are both positive and negative the matrix is **indefinite**.

DIFFERENTIATION – DIFFERENTIAL CALCULUS

LIMITS

We define as the **limit** of a function y = f(x) as x approaches x_0 a number L in which the function converges. Informally the function has a limit L at an input x_0 if f(x) is "close" to L whenever x is "close" to x_0. In other words, f(x) becomes closer and closer to L as x moves closer and closer to x_0. In this case we write,

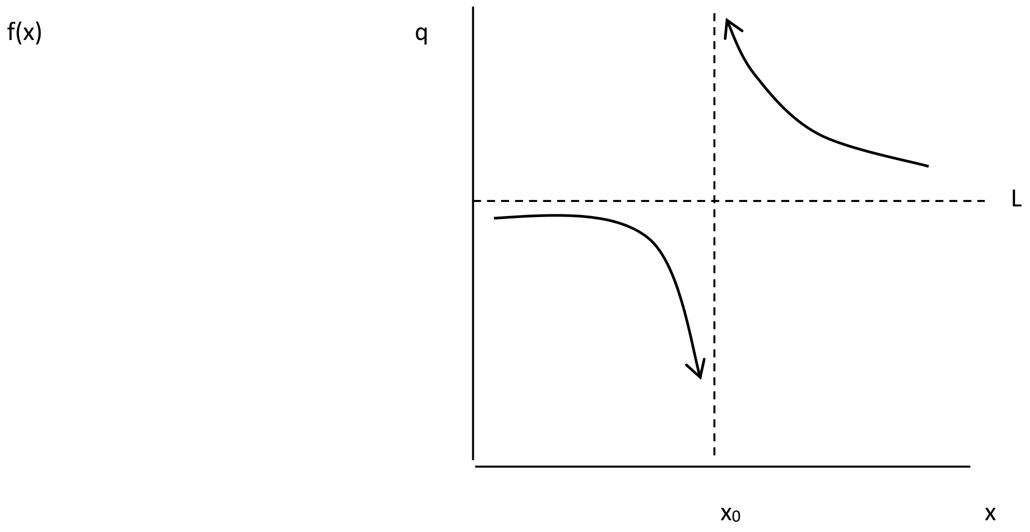
$$\lim_{x \to x_0} f(x) = L$$



In order for the limit to exist the function f(x) should approach L both from the left and the right, that is,

$$\lim_{x \to x_0^-} f(x) = \lim_{x \to x_0^+} f(x) = L$$

In the following graph this is not true. As you observe the function converges to different numbers as x approaches x_0 from the left and the right. In this case the limit of f at x_0 does not exist.



Properties of limits:

Assume that $\lim_{x \to p} f(x) = a, \lim_{x \to p} g(x) = b$, then

- $\lim_{x \to p} \lambda f(x) = \lambda a, \text{ for } \lambda \in \mathbb{R}$
- $\lim_{x \to p} (f(x) \pm g(x)) = a \pm b$

- $\lim_{x \to p} f(x)g(x) = ab$
- $\lim_{x \to p} \frac{f(x)}{g(x)} = \frac{a}{b}$, if $b \neq 0$
- $\lim_{x \to p} |f(x)| = |a|$
- $\lim_{x \to p} f(x)^n = a^n$

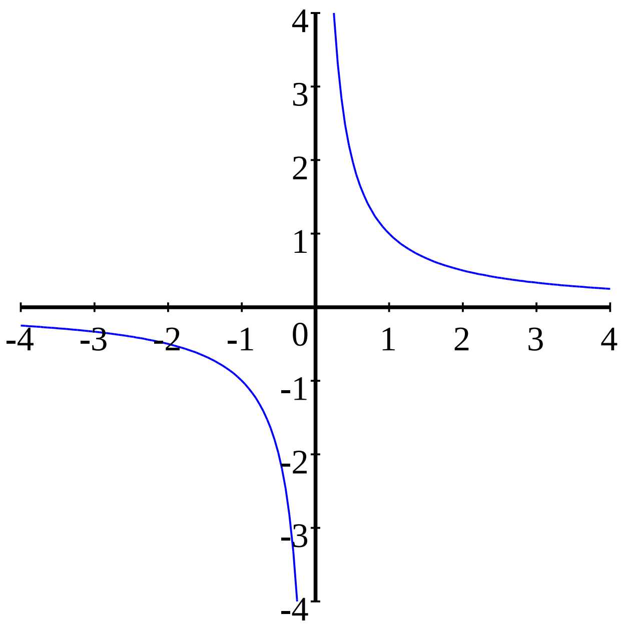
Limits can be used to determine the **asymptotes** of a function. The line y = a is the **horizontal asymptote** of a function f if:

$$\lim_{x \to +\infty} f(x) = a \text{ or } \lim_{x \to -\infty} f(x) = a$$

The line x = a is the **vertical asymptote** of a function f if:

$$\lim_{x \to a^-} f(x) = \pm\infty \text{ or } \lim_{x \to a^+} f(x) = \pm\infty$$

Example 13: Consider the function f(x) = 1/x.



Then we can prove that:

- $\lim_{x\to+\infty} f(x) = 0$ and $\lim_{x\to-\infty} f(x) = 0$, so 0 is a horizontal asymptote of the function.
- $\lim_{x\to 0^+} f(x) = +\infty$ and $\lim_{x\to 0^-} f(x) = -\infty$, so the function f is discontinuous at 0 and 0 is a vertical asymptote of the function.

2. CONTINUITY

A function is **continuous** at x_0 if:

$$\lim_{x \to x_0} f(x) = f(x_0)$$

The above definition implies that for a function to be continuous at a specific point then the following conditions should be satisfied:

- The limit of f(x) at x_0 exists.
- This limit is equal to $f(x_0)$

The function f is continuous at A if it continuous at every $x \in A$.

Using the properties of limits one can easily show that if the functions f and g are continuous at x_0 then the functions: λf, $f \pm g$, fg, $\frac{f}{g}$ if $g(x_0) \neq 0$, $|f|$ and f^n are also continuous at x_0.

We can also define continuity using the left and right limits.

If f is defined on x_0 and $\lim_{x \to x_0^+} f(x) = f(x_0)$ the function is **right continuous** at x_0.

If f is defined on x_0 and $\lim_{x \to x_0^-} f(x) = f(x_0)$ the function is **left continuous** at x_0.

A function f is continuous at x_0 if it is both right and left continuous at this point, i.e,

$$\lim_{x \to x_0^+} f(x) = \lim_{x \to x_0^-} f(x) = f(x_0)$$

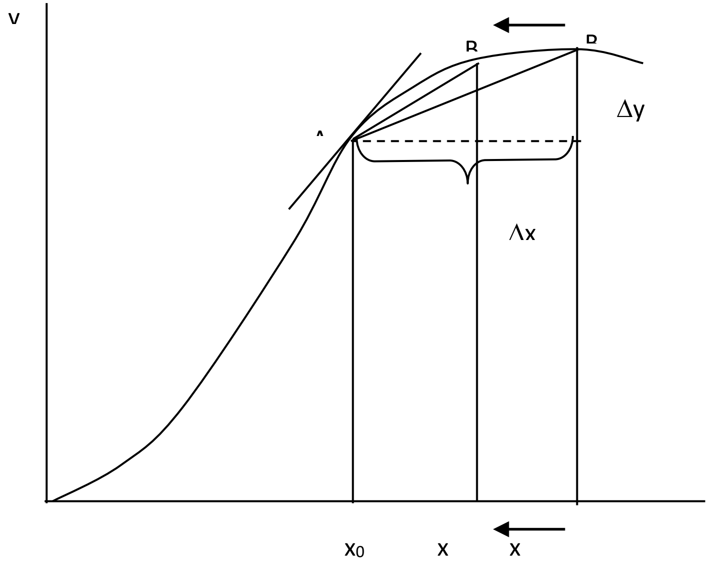
If a function is not continuous at point x_0 it is called **discontinuous** at this point. From the above definition a function is discontinuous if one of the following conditions holds:

- $\lim_{x \to x_0^+} f(x), \lim_{x \to x_0^-} f(x)$ do not exist.
- $\lim_{x \to x_0^+} f(x) \neq \lim_{x \to x_0^-} f(x)$
- $\lim_{x \to x_0^+} f(x) = \lim_{x \to x_0^-} f(x) = L \neq f(x_0)$

3. Definition of the first derivative

Assume a function y = f(x). Consider point A with coordinates $(x_0, f(x_0))$ and B another point on the curve with coordinates (x,f(x)). The slope of the line AB is: $\frac{f(x)-f(x_0)}{x-x_0}$, that is the rate of change of y with respect to the change of x. As x approaches x_0 (or B approaches A) line AB approaches the tangent line at the point A. The slope of this tangent line at A is the **first derivative** of the function f at x_0. It is also called the **instantaneous rate of change**. Formally this is defined as:

$$\lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} = f'(x_0)$$



Alternatively, one could consider that the independent variable x changes by Δx, that is, from x_0 to $x_0 + \Delta x$. In this case the dependent variable would change from $f(x_0)$ to $f(x_0 + \Delta x)$. This change is equal to: $\Delta y = f(x_0 + \Delta x) - f(x_0)$. The mean change of y with respect to x in the interval $(x_0, x_0 + \Delta x)$ is:

$$\frac{f(x_0+\Delta x)-f(x_0)}{x+\Delta x-x_0}=\frac{\Delta y}{\Delta x}$$

If the limit

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}$$

is finite then we say that the function f is **differentiable** at x_0. This limit is the first derivative of f at x_0. Setting $x = x_0 + \Delta x$ we obtain the previous definition of the first derivative.

If the function f is differentiable at a set A of its domain, then we define a new function on A that assigns to every point $x \in A$ the first derivative of *f* at x. This function is called **first derivative** or simply **derivative** of f and it is denoted as $f'(x)$ or $\frac{df(x)}{dx} \equiv \frac{dy}{dx}$.

Example 14: Consider the function $f(x) = 3x^2 - 4$. Then we have that:

$$\frac{\Delta y}{\Delta x} = \frac{3(x_0 + \Delta x)^2 - 4 - (3x_0^2 - 4)}{\Delta x} =$$
$$= \frac{3(x_0^2 + 2x_0\Delta x + \Delta x^2) - 3x_0^2}{\Delta x} =$$
$$= \frac{6x_0\Delta x + 3\Delta x^2}{\Delta x} = 6x_0 + 3\Delta x$$

So, $\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = 6x_0$.

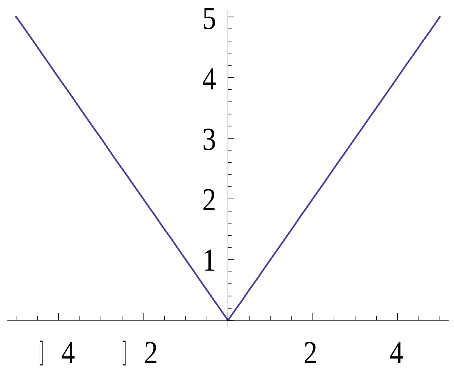
Since x_0 is an arbitrarily chosen point of the domain of f the last result implies that the derivative of f is:

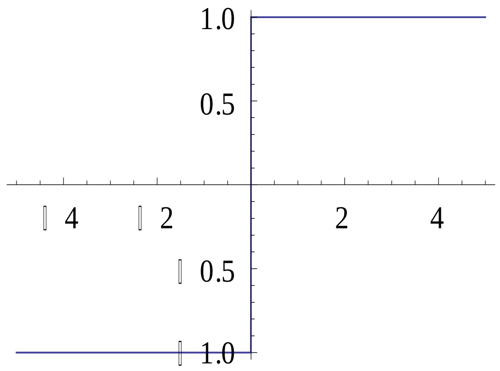
$$f'(x) = 6x$$

The above result implies that the instantaneous rate of change of y with respect to x is different at different points of the domain. In other words, how much would y change with respect to x depends on the value of x.

Theorem 1: If a function f is differentiable at a point x_0 of its domain, then f must also be continuous at x_0.

The opposite does not hold. Thus, continuity of f is necessary but not sufficient for the differentiability of f at x_0. In other words, if f is discontinuous at x_0 then the derivative of f at x_0 does not exist. On the other hand, if f is continuous at x_0 the derivative may or may not exist.


5
4
3
2
1
4
2
2
4


1.0
0.5
4
2
2
4
0.5
1.0

4. Rules of differentiation

The following table gives the derivative of some well-known functions.

$f(x)$	$f'(x)$
c	0
x	1
cx^n	cnx^{n-1}
e^x	e^x
lnx	1/x

The derivative has also the following properties:

- $\frac{d}{dx}(f(x) \pm g(x)) = f'(x) \pm g'(x)$
- $\frac{d}{dx}(f(x)g(x)) = f'(x)g(x) + f(x)g'(x)$
- $\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}, \text{if } g(x) \neq 0$
- $\frac{d}{dx}f^{-1}(x) = \frac{1}{f'(x)}$
- $\frac{d}{dx}f(g(x)) = f'(g(x))g'(x)$ (Chain rule)

The last property and the results of the previous table imply:

- $\frac{d}{dx}f(x)^n = nf(x)^{n-1}f'(x)$

- $\frac{d}{dx} e^{f(x)} = e^{f(x)} f'(x)$
- $\frac{d}{dx} \ln f(x) = \frac{f'(x)}{f(x)}$

5. The instantaneous rate of growth

Assume that a variable y is a function of time t. For example, the final amount of money which is compounded is a function of time, y = f(t).

The **instantaneous rate of growth** of y is defined as:

$$\frac{dy/dt}{y} = \frac{f'(t)}{f(t)}$$

which is also equivalent (see the last bullet) to:

$$\frac{d}{dt}\ln y$$

This last property enables us to determine the instantaneous rate of growth in many economic problems.

Example 15: Assume that $y(t) = y_0 e^{rt}$. Then we can write,

$$\ln y(t) = \ln y_0 + rt$$

and the instantaneous rate of growth equals:

$$\frac{d}{dt}\ln y = r$$

which is constant. If r > 0 the variable y exponentially increases with respect to t, while when r < 0 the variable exponentially decreases.

The notion of constant instantaneous growth rate is very familiar in finance. Assume that r is an interest rate. Then y(t) gives the amount of money one would collect from time 0 (when he/she invests y_0) to time t if this amount of money is continuously compounded. That is in every time period there is a flow of cash that increases your money by r. This r is equal to the percentage change of the variable in a very small period of time. Notice that the absolute change is not constant, actually it increases with respect to t, because the more money you invest the more you gain from them.

The exponential function that appears in the continuous compounding formula is related to the well-known discrete compounding formula

$$y(t) = y_0 \left(1 + \frac{r}{n}\right)^{nt}$$

This comes from the fact that:

$$\lim_{n \to \infty} \left(1 + \frac{r}{n}\right)^{nt} = e^{rt}$$

So, when the number of compounding during the year (given by n) approaches infinity the discrete compounding formula approaches the exponential function.

5. Monotonicity

Consider a function y = f(x). This function is called **increasing** (**decreasing**) over its domain if for every x_1, x_2 of the domain with $x_1 < x_2$ we have $f(x_1) \leq f(x_2)$ ($f(x_1) \geq f(x_2)$). This is equivalent to:

$$\frac{\Delta y}{\Delta x} = \frac{f(x_1) - f(x_2)}{x_1 - x_2} \geq 0 (\leq 0)$$

The function is called **strictly increasing** (**strictly decreasing**) if for $x_1 < x_2$ we have $f(x_1) < f(x_2)$ ($f(x_1) > f(x_2)$).

Theorem 2: Consider a function f which is differentiable over its domain A. Then the following arguments hold:

- f is increasing on A *if and only if* $f'(x_0) \geq 0$ for every $x_0 \in A$.
- f is decreasing on A *if and only if* $f'(x_0) \leq 0$ for every $x_0 \in A$.
- If $f'(x_0) > 0$ for every $x_0 \in A$ then f is strictly increasing on A.
- If $f'(x_0) < 0$ for every $x_0 \in A$ then f is strictly decreasing on A.

From the theorem it is obvious that if f is strictly increasing (decreasing) this does not imply that $f'(x_0) > 0$ ($f'(x_0) < 0$). For example the function $f(x) = x^3$ is strictly increasing but $f'(0) = 0$.

6. Convexity

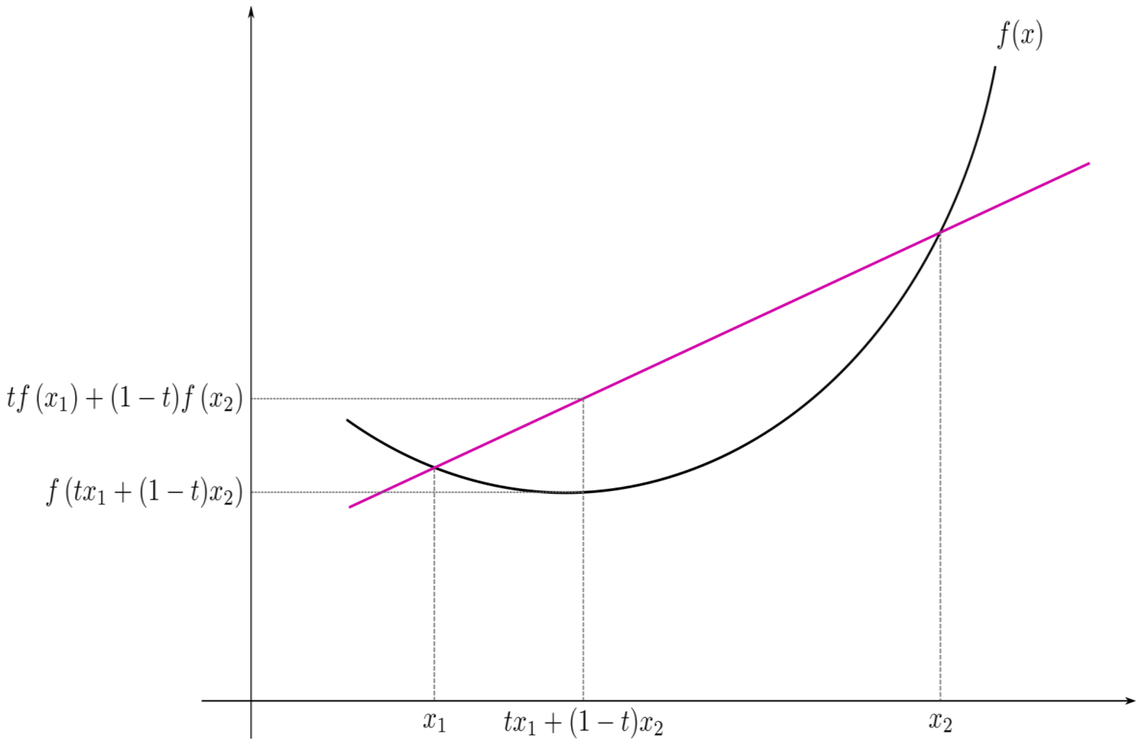
Consider the function y = f(x). This function is called **convex** if for any two points x_1 and x_2 in its domain and any $t \in [0,1]$,

$$f(tx_1 + (1-t)x_2) \le tf(x_1) + (1-t)f(x_2)$$

The function is called **concave** if

$$f(tx_1 + (1-t)x_2) \ge tf(x_1) + (1-t)f(x_2)$$

If the order $\le$ in the definition of convexity is replaced by the strict order < then one obtains a **strictly convex** function. Again, by inverting the order symbol, one finds a **strictly concave** function.


$f(x)$
$tf(x_1) + (1-t)f(x_2)$
$f(tx_1 + (1-t)x_2)$
x_1
$tx_1 + (1-t)x_2$
x_2

Theorem 3: Consider a function f twice differentiable on its domain A. Then the following arguments hold:

- f is convex on A *if and only if* $f''(x_0) \geq 0$ for every $x_0 \in A$.[3]
- f is concave on A *if and only if* $f''(x_0) \leq 0$ for every $x_0 \in A$.
- If $f''(x_0) > 0$ for every $x_0 \in A$ then f is strictly convex on A.
- If $f''(x_0) < 0$ for every $x_0 \in A$ then f is strictly concave on A.

[3] The function $y = f''(x)$ is the second derivative of the function f. It can be defined as the first derivative of the first derivative function $y = f'(x)$.

7. Extrema

A large number of problems in economics, finance and business administration are related to determine the maximum or the minimum point of a function. In some of these problems additional constraints should be imposed, leading to the constrained optimization of the function. We will examine these problems in the reminder of the course. Examples of such problems are:

- Find the optimal portfolio which maximizes return (or minimizes the risk)
- Determine asset prices and returns assuming a representative investor maximizing his/her utility function.
- Determine the quantity of products that a firm should produce to maximize its profits.
- Determine the optimal timing of investing or selling an asset.

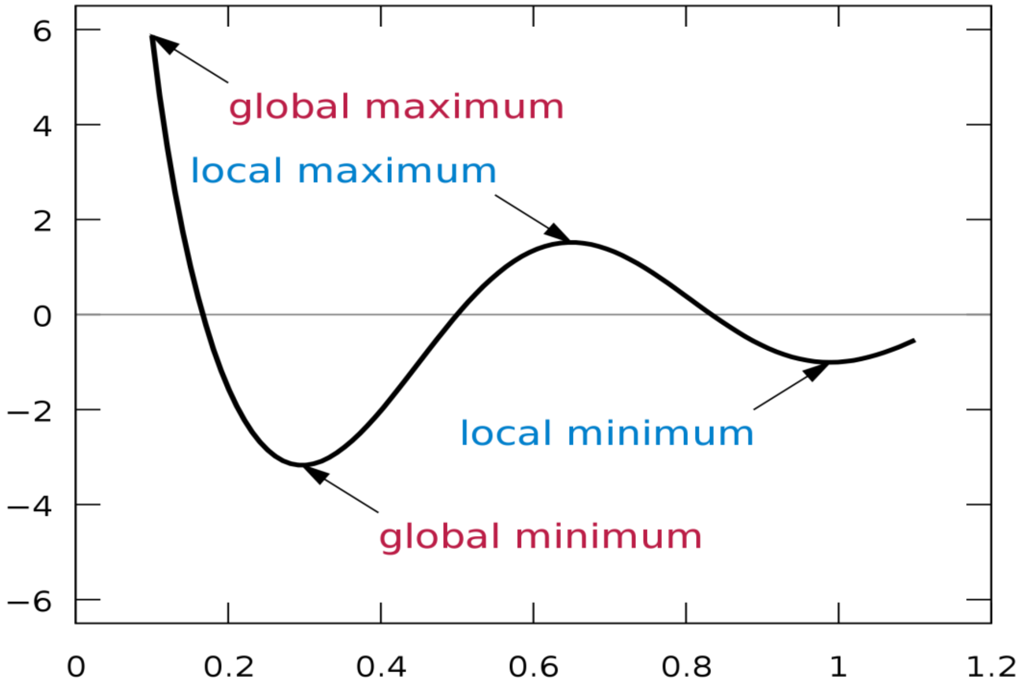
A function f is said to have a **local maximum** at x^*, if there exists some ε > 0 such that $f(x^*) \geq f(x)$ when $|x - x^*| < \varepsilon$.

A function f is said to have a **local minimum** at x^*, if there exists some ε > 0 such that $f(x^*) \leq f(x)$ when $|x - x^*| < \varepsilon$.

A function is said to have a **global maximum** at x^*, if $f(x^*) \geq f(x)$ for all x.

A function is said to have a **global minimum** at x^*, if $f(x^*) \leq f(x)$ for all x.

The local (global) minima and maxima are known as the **local (global) extrema**.



If the order $\leq \text{or} \geq$ is replaced by the strict order < or > for all $x \neq x^*$ then the global and local maxima are unique.

Theorem 4: Consider a function f which is differentiable on the interior of its domain. If the point x^* is a local extremum of the function, then $f'(x^*) = 0$.

Remarks:

- The above theorem implies that a point x^* with $f'(x^*) \neq 0$ cannot be a local extremum of f. To this end, the above condition is a *necessary condition for the existence of a local extremum*. It is also called **first order condition**.
- The opposite argument does not hold. For example, the function $f(x) = x^3$ has $f'(0) = 0$ but 0 is not a local maximum or minimum. These points where the first derivative is equal to zero are called **stationary points**.
- The requirement that x^* belongs to the interior of the function domain is necessary. A function which is defined on a closed set could have local extrema on the bounds of this domain. For example, the function $f(x) = x^2 + 1$ defined on the closed set [1,4] has minimum on x = 1 and maximum on x = 4, but $f'(1) = 2 \neq 0$ and $f'(4) = 8 \neq 0$.

- A function could have local extrema in the interior of its domain without being differentiable at this point. For example, the function $f(x) = |x|$ is not differentiable on 0 but it has a local minimum on this point.

In general, a function f defined over a closed and bounded domain A can have an extremum in the following points:

- Stationary points
- Bounds of the domain A
- Points where f is not differentiable
- Points where f is not continuous

The points where f can have an extremum are called **critical points** of f.

Consider that the function f has an extremum on the interior of its domain. Then the following theorem characterizes this extremum.

Theorem 5: Consider a function f twice differentiable on the interior of its domain and a point x^* of this interior with $f'(x^*) = 0$.

- If $f''(x^*) > 0$, then f has a local minimum on x^*.
- If $f''(x^*) < 0$, then f has a local maximum on x^*.

These conditions are known as the sufficient conditions for determining a local extremum. They are also called **second order conditions**.

The sufficient conditions of the above theorem do not cover the case where $f''(x^*) = 0$. For example, the function $f(x) = x^4$ has $f'(0) = f''(0) = 0$ and 0 is a local minimum. On the other hand, the function f(x) = x^5 has also $f'(0) = f''(0) = 0$ but 0 is just a point of inflexion. Thus, we need a more general statement than theorem 5 to cover all the possible cases. This is given by the following theorem.

Theorem 6: Consider a function f n-differentiable on the interior of its domain and consider a point x^* of this interior for which:

$$f'(x^*) = f''(x^*) = \dots = f^{(n-1)}(x^*) = 0, f^{(n)}(x^*) \neq 0$$

Then,

- If n is even then,
  - If $f^{(n)}(x^*) > 0$, then f has a local minimum on x^*
  - If $f^{(n)}(x^*) < 0$, then f has a local maximum on x^*
- If n is odd, then f has a point of inflexion on x^*.

For example, the function f(x) = x^6 has:

$$f'(0) = f''(0) = \ldots = f^{(5)}(0) = 0$$

and $f^{(6)}(0) = 720 > 0$. Thus, it has a local minimum on 0. On the other hand, the function f(x) = x^5 has:

$$f'(0) = f''(0) = \ldots = f^{(4)}(0) = 0$$

and $f^{(5)}(0) = 120 > 0$. Thus, it has a point of inflexion on 0.

The above theorems characterize local minima and maxima. Under certain conditions these local extrema are also the global extrema of a function. These are given in the following theorem.

Theorem 7: Consider a function f defined on a closed and bounded domain A.

- If f is continuous and concave (convex) on A, then every local maximum (minimum) is also a global maximum (minimum).
- If f is strictly concave (convex), then the global maximum (minimum) is unique.

Example 16: Consider the function $f(x) = \frac{1}{3}x^3 - 4x^2 + 2$ which is continuous and differentiable in the domain $\mathbb{R}$.

1. Find the local maxima and minima of f.
2. Find the intervals where f is increasing and decreasing.
3. Find the intervals where f is convex and concave.

Using the first order condition we specify the points which we may observe local extrema. Thus,

$$f'\left(x^*\right)=0 \Rightarrow \left(x^*\right)^2 - 8x^* = 0 \Rightarrow x^*\left(x^* - 8\right) = 0$$

Thus, the stationary points of this function are $x^* = 0$ and $x^* = 8$. Using the second order condition we can specify if these points are local extrema or points of inflexion. We have that,

$$f''(x) = 2x - 8$$

Thus, $f''(0) = -8 < 0$ and $f''(8) = 8 > 0$. So, in x = 0 the function has a local maximum and in x = 8 it has a local minimum.

To determine the intervals where f is increasing or decreasing, we should examine the sign of the first derivative. But we know that a function changes signs around its root. Solving for the first order condition we find that the roots of f' are 0 and 8. At the interval $\left(-\infty, 0\right)$, $f'(x) > 0$, so the function f is (strictly) increasing. At the interval $\left(0,8\right)$, $f'(x) < 0$, and the function f is (strictly) decreasing. Finally, at the interval $\left(8, +\infty\right)$, $f'(x) > 0$, so f is (strictly) increasing.

To determine the intervals where f is concave or convex, we should examine the sign of the second derivative function. If we solve $f''(x) = 0 \Rightarrow 2x - 8 = 0 \Rightarrow x = 4$, we find that f'' changes signs around 4. At the interval $(-\infty, 4)$, $f''(x) < 0$, so f is (strictly) concave, whereas at the interval $(4, +\infty)$, $f''(x) > 0$, so f is (strictly) convex.

8. Profit maximization

The *cost function* C = C(Q) specifies the total cost C for the production of a quantity Q. We also define the *marginal cost* as

$$MC = \frac{dC}{dQ}$$

and the *average cost* as

$$AC = \frac{C}{Q}$$

According to economic theory the function C has the following characteristics:

- It is strictly increasing $(C'(Q) > 0)$ because as the quantity increases the cost also increases.
- For small values of Q the cost increases with a decreasing rate, so $C''(Q) < 0$ for $Q < Q_1$, and then it increases with increasing rate, so $C''(Q) > 0$ for $Q > Q_1$. Thus Q_1 is a point of inflexion of C. This is due to the law of diminishing returns.

The *revenue function* R = R(Q) specifies the total income R earned by selling a quantity Q. We define the *marginal revenue* as

$$\text{MR} = \frac{\text{dR}}{\text{dQ}}$$

and the *average revenue* as

$$\text{AR} = \frac{\text{R}}{\text{Q}}$$

Consider a firm which wants to determine the quantity of products Q that maximizes its profits.

The *profit function* is defined as:

$$\pi(\text{Q}) = \text{R}(\text{Q}) - \text{C}(\text{Q})$$

The first order condition determines that a local extremum Q^* exists when:

$$\pi'(\text{Q}^*) = 0 \Rightarrow \text{R}'(\text{Q}^*) = \text{C}'(\text{Q}^*) \Rightarrow \text{MR} = \text{MC}$$

that is, when marginal revenue is equal to marginal cost.

The second order condition determines that this point correspond to a maximum if

$$\pi''(Q^*) < 0 \Rightarrow R''(Q^*) < C''(Q^*) \Rightarrow \frac{dMR}{dQ} < \frac{dMC}{dQ}$$

Thus at the point where the profit function is maximized the slope of the marginal revenue is smaller to the slope of the marginal cost.

Example 17: Assume that the firm is a *monopoly*. In this case the firm can determine for itself the quantity and the price of the product. Even in this case the firm cannot impose an extremely high price because demand would decrease, causing the decrease of profits. Assume that the relation between demand quantity Q and price P is determined by the function P = f(Q) for which $f'(Q) < 0$.

Assume that f(Q) = 250 – 20Q.

In this case the revenue function is equal to:

$$R(Q) = QP = Q(250 - 20Q) = 250Q - 20Q^2$$

and the marginal revenue equals:

$$MR = \frac{dR}{dQ} = 250 - 40Q$$

Assume that $C(Q) = \frac{1}{3}Q^3 - \frac{7}{2}Q^2 + 15Q + 100$

Then,

$$MC = \frac{dC}{dQ} = Q^2 - 7Q + 15$$

The first order condition gives:

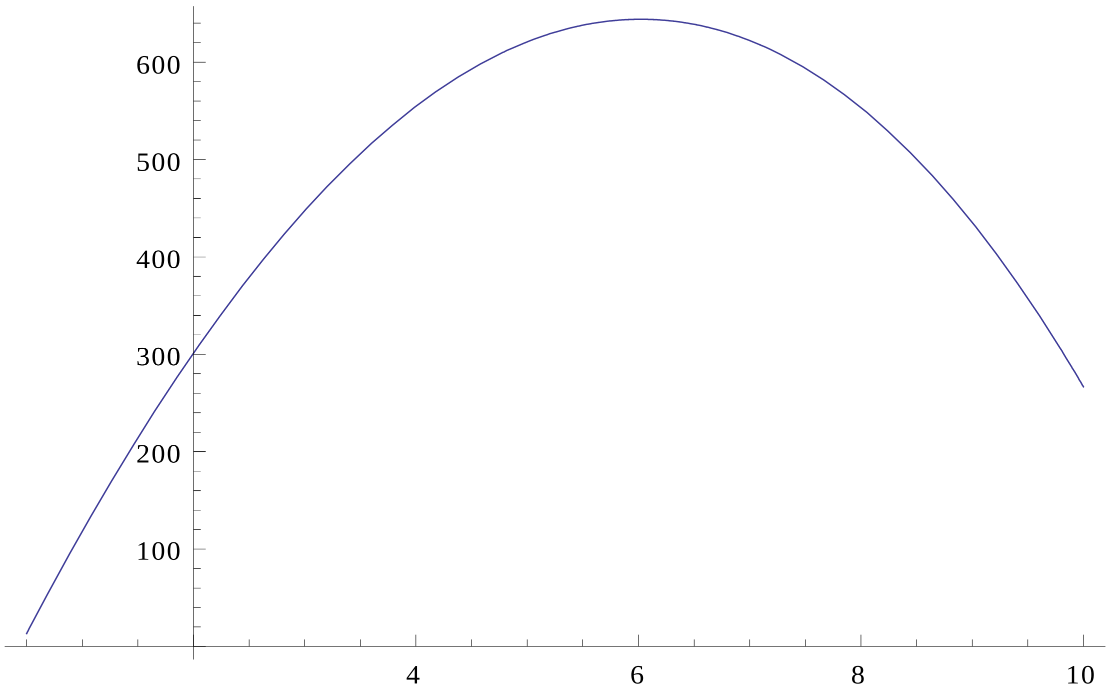
$$MR = MC \Rightarrow 250 - 40Q = Q^2 - 7Q + 15 \Rightarrow Q^2 + 33Q - 235 = 0$$

The solutions of the last equation are Q_1 = -39.02 which is rejected and Q_2 = 6.02.

The second order condition is:

$$\frac{dMR}{dQ} < \frac{dMC}{dQ} \Rightarrow -40 < 2Q - 7$$

which holds for Q = 6.02, so at this point the profit of the firm is maximized. For this quantity demanded the price should be set at P = 129.6.



Example 18: Assume that the firm operates in *a perfectly competitive market*. In this case it cannot influence the price of the product and the price is considered constant both for demanders and suppliers.

The revenue function is equal to: $R(Q) = \overline{P}Q$ and the marginal revenue is constant, $MR = \overline{P}$.

Assume now that the firm of the previous example operates in a perfectly competitive market with $\overline{P} = 50$.

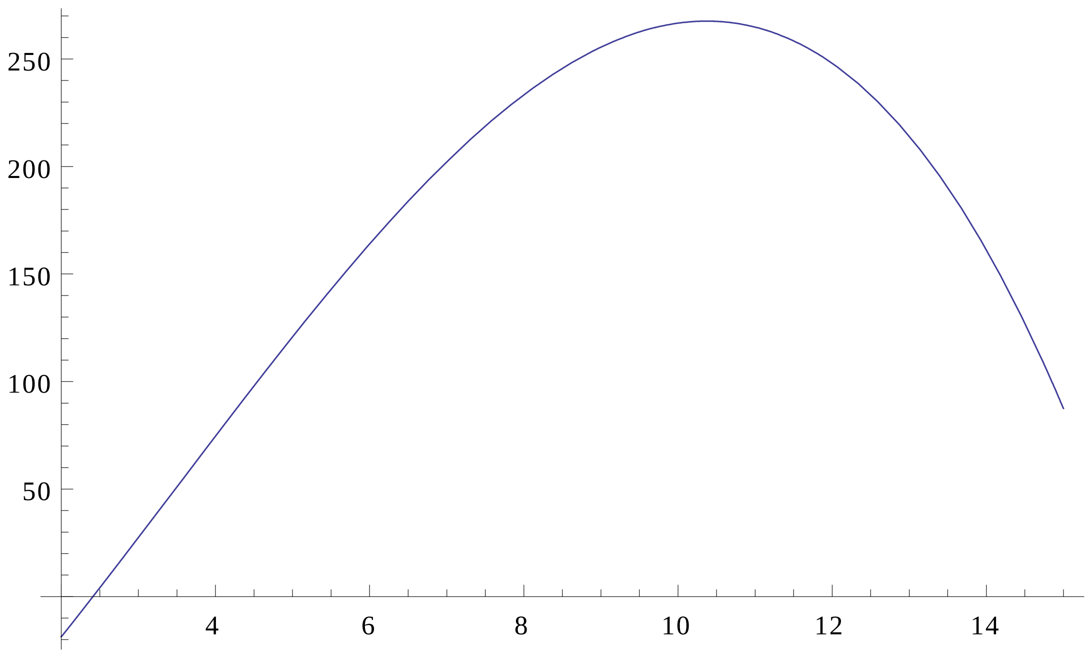
The first order condition is:

$$MR = MC \Rightarrow 50 = Q^2 - 7Q + 15 \Rightarrow Q^2 - 7Q - 35 = 0$$

The solutions of the last equation are $Q_1 = -3.37$ which is rejected and $Q_2 = 10.37$.

$$\frac{dMR}{dQ} < \frac{dMC}{dQ} \Rightarrow 0 < 2Q - 7$$

For $Q_2 = 10.37$ this condition is satisfied, so for this quantity produced the firm maximizes its profits.


250
200
150
100
50
4
6
8
10
12
14

9. Partial derivatives

In many real problems that we face in economics and finance a variable depends on a number of other variables. For example, the demand for a specific car depends on its price, the prices of other cars in the same category, the income of consumers, the age of the car e.t.c. Thus, we need to introduce **multivariate functions** in order to describe these types of relations. Their general form is:

$$y = f(x_1, x_2, \dots, x_n) = f(x)$$

where y is the dependent variable and $x' = (x_1, x_2, \dots, x_n)$ is the vector of independent variables.

A **partial derivative** measures the instantaneous rate of change of the dependent variable with respect to one independent variable assuming that all the others are constant. For the previous function f we can define n partial derivatives,

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}$$

The rules for calculating a partial derivative are the same to the rules of calculating a derivative and are given below:

- If $z = f(x)g(x)$, then $\frac{\partial z}{\partial x_j} = f(x)\frac{\partial g}{\partial x_j} + g(x)\frac{\partial g}{\partial x_j}$
- If $z = \frac{f(x)}{g(x)}, g(x) \neq 0$ then $\frac{\partial z}{\partial x_j} = \frac{f(x)\frac{\partial g}{\partial x_j} - g(x)\frac{\partial g}{\partial x_j}}{g(x)^2}$
- If $z = f(x)^n$ then $\frac{\partial z}{\partial x_j} = nf(x)^{n-1}\frac{\partial f}{\partial x_j}$
- If z = f(x,y) with x = g(u,v) and y = h(u,v) then, $\frac{\partial z}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}$ and $\frac{\partial z}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}$.

10. Constrained optimization

In several problems that we face in economics and finance we must maximize or minimize a multivariate function with respect to several constraints. For example,

- To maximize the production function with respect to budget constraints.
- To maximize the expected utility of an investor with respect to budget constraints.
- To minimize the risk of a portfolio with respect to the return expected by the investor.

The problem can be written as follows:

$$\begin{aligned} &\max_{x_1,\ldots,x_n} f(x_1,\ldots,x_n) \\ &\text{s.t.c } g_j(x_1,\ldots,x_n) = b_j,\ j = 1,2,\ldots,m. \end{aligned} \qquad (3)$$

The following theorem helps us solve this problem.

Theorem 8: If $x^* = (x_1^*, x_2^*, ..., x_n^*)'$ is a local extremum of the problem (3) then it exists a vector $\lambda^* = (\lambda_1^*, \lambda_2^*, ..., \lambda_n^*)'$ such as

$$\frac{\partial f(x^*)}{\partial x_i} - \sum_{j=1}^{m} \lambda_j * \frac{\partial g_j(x^*)}{\partial x_i} = 0, i = 1, 2, ..., n$$

The above theorem implies that in order to find x^* and λ^* we must first write the **Lagrangian function**

$$L(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j (b_j - g_j(x))$$

where λ is the **Lagrange multiplier**. Then we must solve the following system of equations:

$$\frac{\partial L}{\partial x_i} = 0, i = 1, 2, ..., n$$

$$\frac{\partial L}{\partial \lambda_j} = 0, \ j = 1, 2, ..., m$$

The above system of equations determines the **sufficient first order conditions** for x^* to be a local extremum of (3).

We need to determine if this extremum is a maximum or a minimum. For ease of simplicity, we will solve the problem for n = 2 and m = 1.

We define the **bordered Hessian matrix** as:

$$\overline{H}_L(x_1, x_2) = \begin{pmatrix} 0 & \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} \\ \frac{\partial g}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} \\ \frac{\partial g}{\partial x_2} & \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} \end{pmatrix}$$

If $\left|\overline{H}_L(x^*)\right| > 0$ then x* is a local maximum If $\left|\overline{H}_L(x^*)\right| < 0$ then x* is a local minimum. These are the **second order conditions**.[4]

Example 19: An individual's expected utility of end-of-period wealth W_1 can be written as

$$E\left[U\left(W_1\right)\right] = \sum \pi_s U\left(Q_s\right)$$

where Q_s = the number of pure securities paying a dollar if state s occurs. In this context, Q_s represents the number of state s pure securities the individual buys as well as his end-of-period wealth if state s occurs.

Now consider the problem we face when we must decide how to invest our initial wealth W_0 (how much of each pure security we should buy) in order to maximize the expected utility. We

[4] In general, the bordered Hessian matrix is of dimension $(n + m \times n + n)$. In this case x* is a local maximum if the matrix is negative definite. If the matrix is positive definite x* is a local minimum.

therefore face a problem of maximizing the expected utility subject to a wealth constraint. The problem can be written as:

$$\max_{Q_s} \sum_s \pi_s U(Q_s)$$

subject to

$$\sum_s p_s Q_s = W_0$$

Our portfolio decision consists of choices we make for Q_s. To solve the above problem, we construct the Lagrange function:

$$L = \sum_s \pi_s U(Q_s) - \lambda \left[\sum_s p_s Q_s - W_0 \right]$$

where λ is a Lagrange multiplier.

We take the first partial derivatives with respect to the unknown variables Q_s and the Lagrange multiplier, and we set them equal to zero.

$$\frac{\partial L}{\partial Q_s} = \pi_s U'(Q_s) - \lambda p_s = 0 \text{ for every s}$$

$$\frac{\partial L}{\partial \lambda} = \sum_s p_s Q_s - W_0 = 0$$

Solving the above system of equations, we obtain the solution to our problem.

Consider an investor with a logarithmic utility function of wealth, i.e., U(C) = lnC, and initial wealth $5,000. Assume a two-state world where the pure security prices are $0.4 and $0.6 and the state probabilities are 1/3 and 2/3. The system of equations can be written as:

$$\frac{\partial L}{\partial Q_1} = \frac{1}{3Q_1} - 0.4\lambda = 0$$

$$\frac{\partial L}{\partial Q_2} = \frac{2}{3Q_2} - 0.6\lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = 5,000 - 0.4Q_1 - 0.6Q_2 = 0$$

Solving the first two equations we obtain:

$$Q_1 = \frac{1}{1.2\lambda}$$

$$Q_2 = \frac{1}{0.9\lambda}$$

Substituting these equations to the last equation we obtain:

$$\frac{0.4}{1.2\lambda} + \frac{0.6}{0.9\lambda} = 5,000 \Leftrightarrow \lambda = \frac{1}{5,000}$$

Therefore, the optimal investment choices are: Q_1 = 4,166.7 and Q_2 = 5,555.5.

To ensure that the above solutions define a local maximum we must calculate the bordered Hessian matrix. We have that,

$$\frac{\partial g}{\partial Q_1} = p_1, \frac{\partial g}{\partial Q_2} = p_2$$

and

$$\frac{\partial^2 L}{\partial Q_1^{\ 2}} = -\frac{\pi_1}{Q_1^{\ 2}}, \frac{\partial^2 L}{\partial Q_2^{\ 2}} = -\frac{\pi_2}{Q_2^{\ 2}}, \frac{\partial^2 L}{\partial Q_1 \partial Q_2} = 0$$

So,

$$\overline{H}_L = \begin{pmatrix} 0 & p_1 & p_2 \\ p_1 & -\pi_1 / Q_1^{\ 2} & 0 \\ p_2 & 0 & -\pi_2 / Q_2^{\ 2} \end{pmatrix}$$

and we can prove that $\left|\overline{H}_L\right| > 0$, so above solutions define a local maximum.

V. Integration

1. Indefinite integral

When we differentiate a function f we calculate the instantaneous rate of change at a point of its domain. When this rate of change is known, and we want to find the function itself we take the opposite path, and we must apply **integration.** Thus, integration is the opposite of differentiation. If y = f(x) then $\frac{dy}{dx} = f'(x)$ and $\int f'(x)dx = f(x) + c$.

In general we define the **indefinite integral** of a function f, denoted as $\int f(x)dx$, as a family of functions described by the formula

$$\int f(x)dx = F(x) + c, \text{ where } F'(x) = f(x)$$

Example 20: Consider the function f(x) = x^3. The first derivative function is $f'(x) = 3x^2$. So one could assume that the indefinite integral of f' is $\int f'(x)dx = x^3$. However, the first derivative function of f(x) = x^3 + 2 is also equal to $f'(x) = 3x^2$. So starting with f' we are not exactly sure about f. However, one thing is certain that the family of functions that have the same f' differs by a constant c.

To specify c a **boundary condition** should be given. If for example in the previous example we know that f(1) = 3 then,

$$f(x) = x^3 + c \Rightarrow f(1) = 1 + c \Rightarrow 3 = 1 + c \Rightarrow c = 2$$

So, f(x) = x^3 + 2.

2. Rules of integration

We now give the indefinite integrals of some basic functions:

- $\int a\,dx = ax + c$
- $\int ax^n dx = a\frac{1}{n+1}x^{n+1} + c,\ n \neq -1$
- $\int ax^{-1}dx = a\ln|x| + c$
- $\int a^{bx}dx = \frac{a^{bx}}{b\ln a} + c$

- $\int e^{ax}dx = \frac{e^{ax}}{a} + c$

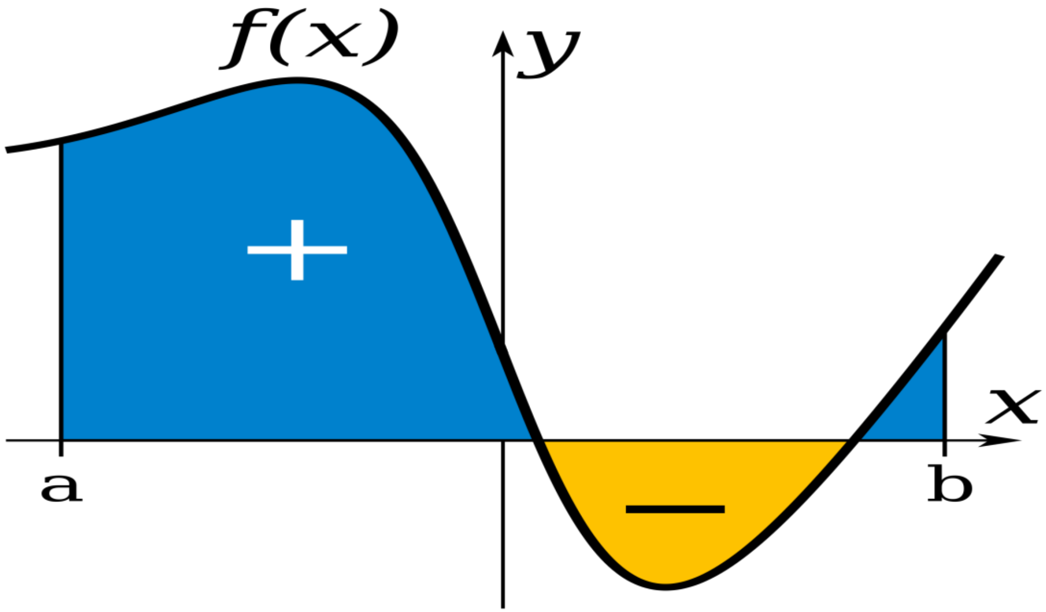
The following rules also hold:

- $\int af(x)dx = a\int f(x)dx$
- $\int (f(x) \pm g(x))dx = \int f(x)dx \pm \int g(x)dx$
- $\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx$ (integration by parts)
- $\int f(g(x))g'(x)dx = \int f(u)du, \text{ with } u = g(x)$ (integration by substitution)

3. Definite integral

Indefinite integrals can be used to compute definite integrals. If $F'(x) = f(x)$ then

$$\int_a^b f(x)dx = F(b) - F(a)$$



This number defines the net signed area of the region bounded by the graph of the function f, the x-axis and the vertical lines x = a and x = b.

The definite integrals have the following properties:

- $\int_a^b f(x)dx = -\int_b^a f(x)dx$
- $\int_a^a f(x)dx = 0$
- $\int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx$