Induction Course in
Quantitative Methods

MSc in Accounting and Finance

Leonidas S. Rompolis
Department of Accounting and Finance
Athens University of Economics and Business

September 2023

Course outline

"Every kind of science, if it has only reached a certain degree of maturity, automatically becomes a part of mathematics", David Hilbert

Part I — Linear algebra

Matrices, operations, determinants, inverses

Part II — Systems of linear equations

Rank, echelon form, existence and uniqueness

Part III — Eigenvalues

Eigenvectors, positive definite matrices

Part IV — Differential calculus

Limits, derivatives, optimisation, Lagrange multipliers

Part V — Integration

Indefinite and definite integrals

Throughout: applications to portfolios, asset pricing and firm behaviour.

PART I

Linear Algebra

Matrices: notation

A **matrix** is a rectangular array of numbers:

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

- Each number is an **entry** or **element**
- Horizontal lines are **rows**, vertical ones **columns**
- m rows $\times$ n columns $\Rightarrow$ an $m \times n$ matrix; m and n are its **dimensions**
- The example above is 2×3

General form

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

The subscripts give the position: a_{12} sits in row 1, column 2.

Vectors and the transpose

A matrix with one row is a **row vector**; with one column, a **column vector**:

$$\begin{pmatrix} 2 & -5 & 6 \end{pmatrix}, \qquad \begin{pmatrix} -5 \\ 1 \end{pmatrix}$$

Transpose

$A^\top$ is obtained by writing the rows of A as columns (equivalently, the columns as rows):

$$(A^\top)_{ij} = a_{ji}$$

If A is $m \times n$, then $A^\top$ is $n \times m$.

Special matrices

- **Square**: equal number of rows and columns
- **Zero**: all elements equal zero
- **Diagonal**: square, all off-diagonal elements zero
- **Identity**: diagonal with ones on the main diagonal
- **Symmetric**: square with $a_{ij} = a_{ji}$, i.e. $A = A^\top$

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix}$$

$$S = \begin{pmatrix} 4 & 1 & 3 \\ 1 & 5 & 2 \\ 3 & 2 & 6 \end{pmatrix} \quad \text{(symmetric)}$$

Why matrices matter

Matrices are useful for

- Solving a linear system of equations — and telling whether a (unique) solution exists
- Input–output analysis describing a production process in macroeconomics
- Multiple regression analysis in econometrics
- Portfolio optimisation problems in investments

Addition, subtraction and scalar multiplication

Matrices can be added or subtracted only if they are **compatible** — same dimensions.

$$(A \pm B)_{ij} = a_{ij} \pm b_{ij}$$

The result has the same dimensions.

Scalar multiplication multiplies every element by the same real number:

$$(cA)_{ij} = c\, a_{ij}$$

Properties of addition

$$A + B = B + A \quad \text{(commutative)}$$

$$(A + B) + C = A + (B + C) \quad \text{(associative)}$$

The inner product of a $1 \times n$ row vector $x^\top$ and an $n \times 1$ column vector y is the number

$$x^\top y = \sum_{i=1}^{n} x_i y_i$$

The number of columns of the first must equal the number of rows of the second.

Example 1 — Linear regression

With y_i the i-th observation of the dependent variable and $x_i = (x_{i1}, \ldots, x_{iK})^\top$ the K independent variables,

$$y_i = \beta_1 x_{i1} + \cdots + \beta_K x_{iK} + \varepsilon_i = x_i^\top \beta + \varepsilon_i$$

where β is the vector of unknown parameters and ε_i an unobserved error term.

Inner product: expected portfolio return

Example 2

An investor holds n stocks with weights $w^\top = (w_1, \dots, w_n)$, $\sum_i w_i = 1$, and expected annual returns $r = (r_1, \dots, r_n)^\top$. The expected portfolio return is

$$E(r_p) = w^\top r = \sum_{i=1}^{n} w_i r_i$$

With four stocks, $w^\top = (0.2,\ 0.3,\ 0.1,\ 0.4)$ and $r^\top = (0.08,\ 0.12,\ 0.05,\ 0.10)$:

$$E(r_p) = 0.2(0.08) + 0.3(0.12) + 0.1(0.05) + 0.4(0.10) = 0.097$$

Defined only when the number of columns of the left matrix equals the number of rows of the right one. If A is $m \times n$ and B is $n \times p$, then $C = AB$ is $m \times p$ with

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

i.e. the inner product of row i of A with column j of B.

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

Example 3

With weights $w^\top$ and variance–covariance matrix V ($n \times n$; element (i,i) is the return variance of stock i, element (i,j) the covariance between i and j), the portfolio return variance is

$$\sigma_p^2 = w^\top V w$$

	EMP	AEGEK	ALFA	BIOXK	ELAIS	ETE	OTE
EMP	0.00473	0.00199	0.00262	0.00132	0.00196	0.00362	0.00213
AEGEK	0.00199	0.00498	0.00142	0.00173	0.00201	0.00144	0.00143
ALFA	0.00262	0.00142	0.00239	0.00152	0.00150	0.00239	0.00159
BIOXK	0.00132	0.00173	0.00152	0.00360	0.00176	0.00174	0.00141
ELAIS	0.00196	0.00201	0.00150	0.00176	0.00438	0.00161	0.00117
ETE	0.00362	0.00144	0.00239	0.00174	0.00161	0.00351	0.00198
OTE	0.00213	0.00143	0.00159	0.00141	0.00117	0.00198	0.00316

With equal weights across the seven stocks, $\sigma_p^2 = 0.0021$.

Properties of multiplication and transposition

Multiplication

- The commutative law **fails**: $AB \neq BA$ in general
- Associative: $(AB)C = A(BC)$
- Distributive: $(A + B)C = AC + BC$
- Identity: $AI = IA = A$

Transposition

- $(A^\top)^\top = A$
- $(A + B)^\top = A^\top + B^\top$
- $(cA)^\top = cA^\top$
- $(AB)^\top = B^\top A^\top$

The determinant

The determinant is a number attached to any square matrix. Geometrically it is the scale factor for measure when the matrix is read as a linear transformation.

Laplace expansion

Let M_{ij} be the **minor matrix** — the $(n-1) \times (n-1)$ matrix left after deleting row i and column j of A — and let the **cofactor** be $c_{ij} = (-1)^{i+j}|M_{ij}|$. Then

$$|A| = \sum_{j=1}^{n} a_{ij} c_{ij} \qquad \text{(expansion along row } i\text{)}$$

There are $2n$ such expressions, one for each row and each column.

Example 4 — the 2×2 case

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Determinant: a worked expansion

Example 5

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

Expanding along the first row:

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

$$= 1(45-48) - 2(36-42) + 3(32-35) = -3 + 12 - 9 = 0$$

The first and third columns sum to twice the second — the columns are linearly dependent.

Properties of determinants

- If every element of a row or column is zero, then $|A| = 0$
- If B comes from A by interchanging two rows or two columns, $|B| = -|A|$
- If B comes from A by multiplying one row or column by c, then $|B| = c|A|$
- If B comes from A by adding a multiple of one row (column) to another row (column), then $|B| = |A|$

Singular matrices

If there is a non-zero vector x with $Ax = 0$, the columns of A are **linearly dependent** and $|A| = 0$. Such a matrix is called **singular**. The same holds for linearly dependent rows.

The inverse of a matrix

The **inverse** of a square matrix A is a matrix A^{-1} of the same dimensions with

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

We need it to move a matrix from one side of an equation to the other. If $Ax = b$, then

$$A^{-1}Ax = A^{-1}b \implies x = A^{-1}b$$

Calculating the inverse

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

where $\mathrm{adj}(A) = C^\top$ is the **adjoint**, the transpose of the cofactor matrix $C = [c_{ij}]$.

The inverse exists only when $|A| \neq 0$. Such a matrix is **invertible** or **non-singular**.

The inverse: the 2×2 case and properties

Example 6

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Properties of the inverse

- $(A^{-1})^{-1} = A$
- $(A^\top)^{-1} = (A^{-1})^\top$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $|A^{-1}| = 1/|A|$

PART II

Systems of Linear Equations

What is a system of linear equations?

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

Example 7 — demand and supply

Quantity demanded falls with the price P; quantity supplied rises with it:

$$Q_d = a - bP, \qquad Q_s = -c + dP, \qquad a, b, c, d > 0$$

Market equilibrium imposes $Q_d = Q_s = Q$. These three equations give a system of two linear equations in two unknowns:

$$Q + bP = a, \qquad Q - dP = -c$$

Solving for P and Q determines the market equilibrium.

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

The x_j are the unknowns (endogenous variables), the a_{ij} the coefficients, the b_i the constant terms (exogenous variables): m equations, n unknowns.

Matrix form

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

A is the **coefficient matrix**, X the **solution vector**, B the **vector of constant terms**. The equilibrium system above becomes

$$\begin{pmatrix} 1 & b \\ 1 & -d \end{pmatrix} \begin{pmatrix} Q \\ P \end{pmatrix} = \begin{pmatrix} a \\ -c \end{pmatrix}$$

Three possible outcomes

A linear system behaves in exactly one of three ways:

1. infinitely many solutions
2. a unique solution
3. no solution at all (**inconsistency**)

Which case applies is settled by the **rank**.

Rank

The **column rank** is the maximal number of linearly independent columns; the **row rank** the maximal number of linearly independent rows. They are always equal, so we simply speak of the **rank** $r(A)$.

For an $m \times n$ matrix, $r(A) \leq \min(m, n)$. When equality holds the matrix has **full rank**.

The **augmented matrix** is $C = [A \mid B]$.

Row echelon form

To find the rank, transform the matrix into **row echelon form**, which satisfies:

- the first non-zero element of each row (the **leading entry**) is 1
- each leading entry lies in a column to the right of the leading entry of the row above
- rows of all zeros, if any, sit below rows with non-zero elements

Elementary row operations

1. Start with the first row if its first-column entry is non-zero — this entry is the **pivot**. Otherwise interchange it with another row.
2. Multiply every element of the pivot row by the inverse of the pivot, so the pivot becomes 1.
3. Add multiples of the pivot row to the lower rows so that every other element of the pivot column becomes 0.
4. Repeat, ignoring the rows already used as pivots, until no pivots remain.

The rank is the number of non-zero rows in the echelon form.

Echelon form: worked example

Example 8

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 3 & 5 \end{pmatrix} \xrightarrow{R_2 - 2R_1\,,\ R_3 - R_1} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

The last matrix is in echelon form. It has two non-zero rows, so $r(A) = 2$.

When does a solution exist?

Proposition 1

System (1) has a solution if and only if $r(A) = r(C)$.

Proposition 2

System (1) has a *unique* solution if $r(A) = r(C) = n$, where n is the number of unknowns.

Proposition 3

If $n = m$ and $r(A) = n$, then system (1) has a unique solution.

If $m < n$ (more unknowns than equations) then $r(A) \leq m < n$, so the system cannot have a unique solution.

Example 9

Take

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

Here $r(A) = 1$ but $r(C) = 2$, so by Proposition 1 the system has **no solution**.

If instead $B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, then $r(C) = 1 = r(A)$ and the system **does** have a solution — though not a unique one.

When $n = m$ and $r(A) = n$, the matrix A is square with full rank, hence invertible, and

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

Example 7 (continued)

For the demand–supply system $n = m = 2$ and $r(A) = 2$, so the solution is unique:

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

$$\begin{pmatrix} Q^* \\ P^* \end{pmatrix} = \frac{1}{b+d}\begin{pmatrix} d & b \\ 1 & -1 \end{pmatrix}\begin{pmatrix} a \\ -c \end{pmatrix} = \begin{pmatrix} \dfrac{ad-bc}{b+d} \\ \dfrac{a+c}{b+d} \end{pmatrix}$$

Example 10: a two-asset market

Two assets with prices P_1 and P_2. Demand and supply are

$$\begin{aligned} Q_{d1} &= a_0 + a_1 P_1 + a_2 P_2, & Q_{s1} &= b_0 + b_1 P_1 + b_2 P_2 \\ Q_{d2} &= \alpha_0 + \alpha_1 P_1 + \alpha_2 P_2, & Q_{s2} &= \beta_0 + \beta_1 P_1 + \beta_2 P_2 \end{aligned}$$

with equilibrium $Q_{d1} = Q_{s1}$ and $Q_{d2} = Q_{s2}$: a 4×4 system.

The practical point

Inverting a 4×4 matrix by hand is painful. Instead substitute the equilibrium conditions to eliminate the quantities, leaving a 2×2 system in P_1 and P_2 which solves as in the previous example. The quantities then follow from either original equation.

Example 11: Arrow–Debreu securities

A **pure** (**primitive**, **Arrow–Debreu**) **security** pays \$1 at the end of the period if a given state occurs and nothing otherwise. Every market security can be decomposed into a portfolio of pure securities. Pure securities are not traded, so their prices must be inferred from traded securities.

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

Two states of nature and two linearly independent market securities: the market is **complete**, so the pure security prices are uniquely determined.

Arrow–Debreu: solving for the state prices

Buying 10 units of $(1, 0)$ and 20 units of $(0, 1)$ replicates security j; 30 and 10 units replicate security k. Replicating payoffs must have replicating prices:

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

Solution

$$p_1 = \$0.20, \qquad p_2 = \$0.30$$

A 2×2 linear system delivers the state-price vector — the foundation of asset pricing.

PART III

Eigenvalues and Eigenvectors

Positive definite matrices

Positive definite matrices

A symmetric $n \times n$ matrix A is **positive definite** if

$$x^\top Ax > 0 \quad \text{for all non-zero } n \times 1 \text{ vectors } x$$

Replacing $>$ accordingly gives **positive semi-definite** (≥ 0), **negative definite** (< 0) and **negative semi-definite** (≤ 0).

- A positive definite matrix is invertible
- Positive definiteness underpins linear regression analysis and the maximisation or minimisation of a multivariate function
- A variance–covariance matrix is positive semi-definite by construction — recall $\sigma_p^2 = w^\top V w \geq 0$

The eigenvalue problem

Consider

$$Ax = \lambda x \quad \Longleftrightarrow \quad (A - \lambda I)x = 0$$

where A is $n \times n$, x is $n \times 1$ and λ is a real number.

This is a **homogeneous** system: the constant terms are zero. When $A - \lambda I$ is invertible the only solution is the trivial $x = 0$. Non-trivial solutions require

$$|A - \lambda I| = 0 \tag{2}$$

The values of λ satisfying (2) are the **eigenvalues** of A. The corresponding non-trivial x are the **eigenvectors**.

Example 12: finding eigenvalues and eigenvectors

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

so $\lambda_1 = -6$ and $\lambda_2 = -1$.

Substituting each eigenvalue back into $(A - \lambda I)x = 0$:

$$\lambda = -6: \quad x_1 + 2x_2 = 0 \qquad\qquad \lambda = -1: \quad -2x_1 + x_2 = 0$$

$$\text{eigenvector} \propto \begin{pmatrix} 2 \\ -1 \end{pmatrix} \qquad\qquad \text{eigenvector} \propto \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Eigenvectors are determined only up to a scalar multiple.

The link

- all eigenvalues positive $\Rightarrow$ **positive definite**
- all eigenvalues negative $\Rightarrow$ **negative definite** (as in Example 12)
- all eigenvalues non-negative, at least one zero $\Rightarrow$ **positive semi-definite**
- all eigenvalues non-positive, at least one zero $\Rightarrow$ **negative semi-definite**
- eigenvalues of both signs $\Rightarrow$ the matrix is **indefinite**

This is the test we will use for second-order conditions in multivariate optimisation.

PART IV

Differentiation

Differential calculus

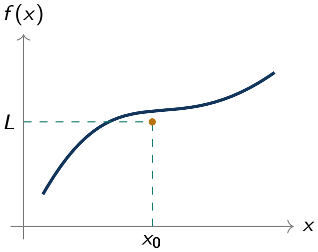
The **limit** of $y = f(x)$ as x approaches x_0 is the number L to which the function converges: $f(x)$ gets closer and closer to L as x gets closer and closer to x_0,

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

For the limit to exist the function must approach L from both sides:

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

If the left and right limits differ, the limit at x_0 does not exist.



Properties of limits and asymptotes

Assume $\lim_{x\to x_0} f(x) = L$ and $\lim_{x\to x_0} g(x) = M$. Then

$$\lim(f \pm g) = L \pm M, \qquad \lim(fg) = LM, \qquad \lim \frac{f}{g} = \frac{L}{M} \ (M \neq 0), \qquad \lim cf = cL$$

Asymptotes

The line $y = a$ is a **horizontal asymptote** of f if $\lim\limits_{x\to\pm\infty} f(x) = a$.

The line $x = a$ is a **vertical asymptote** of f if $\lim\limits_{x\to a^\pm} f(x) = \pm\infty$.

Example 13

For $f(x) = 1/x$:

$$\lim_{x\to\pm\infty} \frac{1}{x} = 0 \ \Rightarrow \ y = 0 \text{ is a horizontal asymptote}$$

$$\lim_{x\to 0^\pm} \frac{1}{x} = \pm\infty \ \Rightarrow \ x = 0 \text{ is a vertical asymptote, and } f \text{ is discontinuous at } 0$$

A function is **continuous** at x_0 if

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

which requires both that the limit exists and that it equals $f(x_0)$. The function f is continuous on A if it is continuous at every point of A.

- If f and g are continuous at x_0, so are $f \pm g$, fg and f/g (where $g(x_0) \neq 0$)
- f is **right continuous** at x_0 if $\lim_{x \to x_0^+} f(x) = f(x_0)$; **left continuous** if the left limit does the same
- f is continuous at x_0 exactly when it is both right and left continuous there
- Otherwise f is **discontinuous** at x_0 — either a one-sided limit fails to exist, or they disagree, or they differ from $f(x_0)$

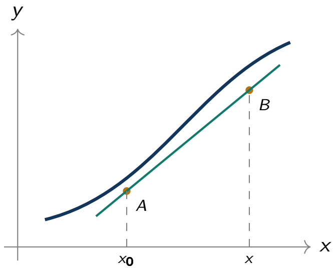
The first derivative

Take $A = (x_0, f(x_0))$ and $B = (x, f(x))$ on the curve. The slope of the chord AB is

$$\frac{f(x) - f(x_0)}{x - x_0}$$

the rate of change of y with respect to x. As $x \to x_0$, the chord approaches the tangent line at A, whose slope is the **first derivative**, also called the **instantaneous rate of change**:

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



The derivative: the Δx formulation

Let x change by Δx, from x_0 to $x_0 + \Delta x$. Then y changes by

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

and the mean rate of change over $(x_0, x_0 + \Delta x)$ is $\Delta y / \Delta x$. If

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

is finite, f is **differentiable** at x_0 and the limit is $f'(x_0)$. Setting $x = x_0 + \Delta x$ recovers the previous definition.

If f is differentiable on a set A, the map assigning $f'(x)$ to each $x \in A$ is itself a function — the **derivative** of f, written $f'(x)$ or $\mathrm{d}y/\mathrm{d}x$.

Example 14

For $f(x) = x^2$:

$$\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \frac{(x_0 + \Delta x)^2 - x_0^2}{\Delta x} = 2x_0 + \Delta x \xrightarrow[\Delta x \to 0]{} 2x_0$$

Since x_0 was arbitrary, $f'(x) = 2x$. The instantaneous rate of change differs at different points: how much y moves with x depends on where you are.

Theorem 1

If f is differentiable at x_0, then f is continuous at x_0.

The converse fails. Continuity is necessary but not sufficient for differentiability: if f is discontinuous at x_0 the derivative does not exist there, but continuity alone does not guarantee it does.

Rules of differentiation

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

Properties

$$(f \pm g)' = f' \pm g'$$
$$(cf)' = cf'$$
$$(fg)' = f'g + fg'$$
$$\left(\tfrac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$$
$$(f(g(x)))' = f'(g(x))\,g'(x) \quad \text{(chain rule)}$$

Combining the table with the chain rule gives, for example,

$$\left(e^{g(x)}\right)' = g'(x)e^{g(x)}, \qquad \left(\ln g(x)\right)' = \frac{g'(x)}{g(x)}$$

The instantaneous rate of growth

Let a variable y be a function of time, $y = f(t)$ — for instance an amount of money being compounded.

$$\text{instantaneous rate of growth} = \frac{f'(t)}{f(t)} = \frac{\mathrm{d} \ln f(t)}{\mathrm{d}t}$$

Example 15

With $y(t) = y_0 e^{rt}$:

$$\ln y(t) = \ln y_0 + rt \quad \Longrightarrow \quad \frac{\mathrm{d} \ln y(t)}{\mathrm{d}t} = r$$

The growth rate is constant. If $r > 0$ the variable grows exponentially in t; if $r < 0$ it decays exponentially.

Continuous compounding

If r is an interest rate, $y(t) = y_0 e^{rt}$ is what an investment of y_0 at time 0 is worth at time t under **continuous compounding**: in every instant a flow of cash increases the balance by r, the percentage change over a very short period.

> The *absolute* change is not constant — it grows with t, because the larger the balance the more it earns.

The link to discrete compounding:

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

As the number of compounding periods per year goes to infinity, the discrete formula converges to the exponential one.

f is **increasing** (**decreasing**) on its domain if for every $x_1 < x_2$ we have $f(x_1) \leq f(x_2)$ ($f(x_1) \geq f(x_2)$). Replacing $\leq$ with $<$ gives **strictly increasing**; $\geq$ with $>$ gives **strictly decreasing**.

Theorem 2

For f differentiable on its domain A:

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

The last two do not reverse: $f(x) = x^3$ is strictly increasing yet $f'(0) = 0$.

f is **convex** if for any x_1, x_2 in its domain and any $\lambda \in [0,1]$

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

and **concave** if the inequality is reversed. Replacing $\leq$ by the strict $<$ gives a **strictly convex** function; reversing again gives **strictly concave**.

Theorem 3

For f twice differentiable on its domain A:

- f is convex on A *if and only if* $f''(x) \geq 0$ for every $x \in A$
- f is concave on A *if and only if* $f''(x) \leq 0$ for every $x \in A$
- $f''(x) > 0$ on A implies f is strictly convex; $f''(x) < 0$ implies strictly concave

Extrema: why we care

A great many problems in economics, finance and business come down to locating the maximum or minimum of a function — often subject to constraints.

Examples we will meet

- find the optimal portfolio maximising return, or minimising risk
- determine asset prices and returns for a representative investor maximising utility
- determine the quantity a firm should produce to maximise profit
- determine the optimal timing of investing in, or selling, an asset

Local and global extrema

Local

f has a **local maximum** at x^* if there is some $\varepsilon > 0$ with

$$f(x) \leq f(x^*) \quad \text{when } |x - x^*| < \varepsilon$$

and a **local minimum** if $f(x) \geq f(x^*)$ on such a neighbourhood.

Global

f has a **global maximum** at x^* if $f(x) \leq f(x^*)$ for *all* x, and a **global minimum** if $f(x) \geq f(x^*)$ for all x.

Local and global minima and maxima together are the **local (global) extrema**. If $\leq$ is replaced by the strict $<$ (or $\geq$ by $>$), the maxima and minima are unique.

First order condition

Theorem 4

Let f be differentiable on the interior of its domain. If x^* is a local extremum, then

$$f'(x^*) = 0$$

Remarks

- A point with $f'(x) \neq 0$ cannot be a local extremum: the condition is *necessary*, and is called the **first order condition**
- It is not sufficient. $f(x) = x^3$ has $f'(0) = 0$ but 0 is neither a maximum nor a minimum. Points with $f'(x) = 0$ are **stationary points**
- Interiority matters: $f(x) = x^2 + 1$ on $[1, 4]$ has its minimum at $x = 1$ and maximum at $x = 4$, yet neither derivative is zero
- A function can have a local extremum where it is not differentiable: $f(x) = |x|$ has a local minimum at 0

Critical points and second order conditions

Critical points

On a closed and bounded domain A, an extremum can occur only at

- stationary points
- bounds of the domain A
- points where f is not differentiable
- points where f is not continuous

Theorem 5 — second order conditions

Let f be twice differentiable on the interior of its domain, at a point x^* with $f'(x^*) = 0$:

- if $f''(x^*) > 0$, then f has a local minimum at x^*
- if $f''(x^*) < 0$, then f has a local maximum at x^*

These are the **sufficient conditions**, or **second order conditions**.

When the second derivative vanishes

Theorem 5 is silent when $f''(x^*) = 0$: $f(x) = x^4$ has a local minimum at 0, while $f(x) = x^5$ has only a point of inflexion there.

Theorem 6

Let f be n-times differentiable on the interior of its domain, with

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

Then:

- if n is **even**: $f^{(n)}(x^*) > 0$ gives a local minimum; $f^{(n)}(x^*) < 0$ gives a local maximum
- if n is **odd**: f has a point of inflexion at x^*

$f(x) = x^6$: the first non-zero derivative at 0 is $f^{(6)}(0) = 720 > 0$, $n = 6$ even $\Rightarrow$ local minimum.
$f(x) = x^5$: the first non-zero derivative at 0 is $f^{(5)}(0) = 120$, $n = 5$ odd $\Rightarrow$ point of inflexion.

From local to global

Theorem 7

Let f be 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), the global maximum (minimum) is unique

This is why concavity of the utility function and convexity of the variance matter so much in portfolio theory: they turn a local answer into the answer.

$f(x) = x^3 - 12x^2 + 5$, continuous and differentiable on $\mathbb{R}$. Find (a) local extrema, (b) intervals of monotonicity, (c) intervals of convexity.

First order condition: $f'(x) = 3x^2 - 24x = 3x(x-8) = 0$, so the stationary points are $x = 0$ and $x = 8$.

Second order condition: $f''(x) = 6x - 24$, so $f''(0) = -24 < 0$ and $f''(8) = 24 > 0$.

$$\Rightarrow \; x = 0 \text{ is a local maximum,} \qquad x = 8 \text{ is a local minimum}$$

Monotonicity: f' changes sign at its roots. On $(-\infty, 0)$, $f' > 0$ (strictly increasing); on $(0, 8)$, $f' < 0$ (strictly decreasing); on $(8, \infty)$, $f' > 0$ (strictly increasing).

Convexity: f'' changes sign at $x = 4$. On $(-\infty, 4)$, $f'' < 0$ (strictly concave); on $(4, \infty)$, $f'' > 0$ (strictly convex).

Profit maximisation: cost and revenue

The **cost function** $C = C(Q)$ gives total cost of producing quantity Q. Marginal and average cost are

$$\text{MC} = C'(Q), \qquad \text{AC} = \frac{C(Q)}{Q}$$

Shape implied by theory

- strictly increasing, $C'(Q) > 0$
- for small Q cost rises at a decreasing rate, $C''(Q) < 0$ for $Q < Q_1$; thereafter at an increasing rate, $C''(Q) > 0$ for $Q > Q_1$
- Q_1 is a point of inflexion — the law of diminishing returns

The **revenue function** $R = R(Q)$ gives total income from selling Q. Marginal and average revenue are

$$\text{MR} = R'(Q), \qquad \text{AR} = \frac{R(Q)}{Q}$$

Profit

$$\Pi(Q) = R(Q) - C(Q)$$

Profit maximisation: the conditions

First order condition

$$\Pi'(Q^*) = R'(Q^*) - C'(Q^*) = 0 \quad \Longleftrightarrow \quad R'(Q^*) = C'(Q^*)$$

Marginal revenue equals marginal cost.

Second order condition

$$\Pi''(Q^*) = R''(Q^*) - C''(Q^*) < 0 \quad \Longleftrightarrow \quad R''(Q^*) < C''(Q^*)$$

At the profit-maximising point the slope of marginal revenue is smaller than the slope of marginal cost.

A **monopolist** sets both quantity and price — but cannot set the price arbitrarily high, since demand would fall and profit with it. Let $P = f(Q)$ with $f'(Q) < 0$; take

$$f(Q) = 250 - 20Q \quad \Longrightarrow \quad R(Q) = PQ = 250Q - 20Q^2, \qquad R'(Q) = 250 - 40Q$$

Take the cost function $C(Q) = \frac{1}{3}Q^3 - 3.5Q^2 + 15Q + k$, so that $C'(Q) = Q^2 - 7Q + 15$.

First order condition: $250 - 40Q = Q^2 - 7Q + 15 \;\Rightarrow\; Q^2 + 33Q - 235 = 0$

$$Q_1 = -39.02 \;\; \text{(rejected)}, \qquad Q_2 = 6.02$$

Second order condition: $R''(Q) = -40$ and $C''(Q) = 2Q - 7 = 5.04$ at $Q = 6.02$, so $R'' < C''$ and profit is maximised there.

At that quantity the price is set at $P = 250 - 20(6.02) \approx \$129.6$.

In a **perfectly competitive market** the firm cannot influence the price: P is constant for buyers and sellers alike. Hence

$$R(Q) = PQ, \qquad R'(Q) = P \quad \text{(constant marginal revenue)}$$

Take the same firm as Example 17, now facing a market price $P = \$50$.

First order condition: $50 = Q^2 - 7Q + 15 \;\Rightarrow\; Q^2 - 7Q - 35 = 0$

$$Q_1 = -3.37 \;\text{(rejected)}, \qquad Q_2 = 10.37$$

Second order condition: $R''(Q) = 0 < C''(Q) = 2Q - 7 = 13.75$ at $Q = 10.37$, so profit is maximised at that quantity.

Partial derivatives

In most real problems a variable depends on several others — the demand for a car depends on its price, the prices of competing cars, consumer income, the age of the model, and so on. We therefore need **multivariate functions**:

$$y = f(x_1, x_2, \ldots, x_n)$$

A **partial derivative** measures the instantaneous rate of change of y with respect to one independent variable, holding all the others constant:

$$\frac{\partial f}{\partial x_i}, \qquad i = 1, \ldots, n$$

The rules are exactly those of ordinary differentiation, applied one variable at a time:

- $z = ax + by \;\Rightarrow\; \partial z/\partial x = a$ and $\partial z/\partial y = b$
- $z = xy \;\Rightarrow\; \partial z/\partial x = y$ and $\partial z/\partial y = x$
- $z = f(x, y)$ with $x = g(u, v)$ and $y = h(u, v) \Rightarrow$

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}, \qquad \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}$$

Constrained optimisation

Many problems require maximising or minimising a multivariate function subject to constraints:

- maximise a production function subject to a budget constraint
- maximise an investor's expected utility subject to a budget constraint
- minimise portfolio risk subject to a required expected return

The problem

$$\max_x \; f(x_1, \ldots, x_n) \quad \text{subject to} \quad g_j(x_1, \ldots, x_n) = 0, \quad j = 1, \ldots, m \tag{3}$$

Theorem 8

If x^* is a local extremum of problem (3), then there exists a vector $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)$ such that

$$\nabla f(x^*) = \sum_{j=1}^{m} \lambda_j^* \, \nabla g_j(x^*)$$

To find x^* and λ^*, write the **Lagrangian function**

$$L(x, \lambda) = f(x_1, \ldots, x_n) - \sum_{j=1}^{m} \lambda_j \, g_j(x_1, \ldots, x_n)$$

where λ_j is the **Lagrange multiplier**, and solve

$$\frac{\partial L}{\partial x_i} = 0 \;\; (i = 1, \ldots, n), \qquad \frac{\partial L}{\partial \lambda_j} = 0 \;\; (j = 1, \ldots, m)$$

These are the **first order conditions** for x^* to be a local extremum of (3). The conditions on the multipliers simply return the constraints.

Second order conditions: the bordered Hessian

Take $n = 2$ and $m = 1$ for simplicity.

Bordered Hessian

$$\bar{H} = \begin{pmatrix} 0 & g_1 & g_2 \\ g_1 & L_{11} & L_{12} \\ g_2 & L_{21} & L_{22} \end{pmatrix} \qquad g_i = \frac{\partial g}{\partial x_i}, \quad L_{ij} = \frac{\partial^2 L}{\partial x_i \partial x_j}$$

- if $|\bar{H}| > 0$ then x^* is a local **maximum**
- if $|\bar{H}| < 0$ then x^* is a local **minimum**

In general the bordered Hessian is $(n+m) \times (n+m)$. Then x^* is a local maximum if the matrix is negative definite, and a local minimum if it is positive definite — which is where Part III pays off.

Example 19: maximising expected utility

An individual's expected utility of end-of-period wealth W_1 is

$$E\big[U(W_1)\big] = \sum_{s=1}^{S} \pi_s\, U(Q_s)$$

where Q_s is the number of state-s pure securities bought — which is also end-of-period wealth if state s occurs, and π_s the probability of state s.

With initial wealth W_0 and pure security prices p_s, the problem is

$$\max_{Q_1,\dots,Q_S} \sum_s \pi_s U(Q_s) \qquad \text{subject to} \qquad \sum_s p_s Q_s = W_0$$

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

Setting the partial derivatives to zero:

$$\frac{\partial L}{\partial Q_s} = \pi_s U'(Q_s) - \lambda p_s = 0, \qquad \frac{\partial L}{\partial \lambda} = W_0 - \sum_s p_s Q_s = 0$$

Logarithmic utility $U(C) = \ln C$, $W_0 = \$5{,}000$, two states with $p_1 = \$0.4$, $p_2 = \$0.6$ and $\pi_1 = 1/3$, $\pi_2 = 2/3$. The first two conditions give

$$Q_1 = \frac{\pi_1}{\lambda p_1}, \qquad Q_2 = \frac{\pi_2}{\lambda p_2}$$

Substituting into the budget constraint: $\pi_1/\lambda + \pi_2/\lambda = W_0$, so $\lambda = 1/W_0$ and

$$Q_s = \frac{\pi_s W_0}{p_s} \quad \Longrightarrow \quad Q_1 = \frac{(1/3)(5000)}{0.4} = 4{,}166.7, \qquad Q_2 = \frac{(2/3)(5000)}{0.6} = 5{,}555.5$$

Example 19: checking it is a maximum

With $U(C) = \ln C$ we have $L_{ss} = -\pi_s / Q_s^2$ and $L_{12} = L_{21} = 0$, while $g_s = p_s$. The bordered Hessian is

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

Expanding,

$$|\bar{H}| = p_1^2 \frac{\pi_2}{Q_2^2} + p_2^2 \frac{\pi_1}{Q_1^2} > 0$$

The determinant is positive, so the solution $Q_1 = 4{,}166.7$, $Q_2 = 5{,}555.5$ is a local maximum — and since ln is strictly concave, by Theorem 7 it is the global maximum.

PART V

Integration

Differentiating f gives the instantaneous rate of change. When the rate of change is known and we want the function itself, we take the opposite path: **integration** is the inverse of differentiation.

If $F'(x) = f(x)$, then the **indefinite integral** of f is the family of functions

$$\int f(x)\,\mathrm{d}x = F(x) + c$$

Example 20

Both $f(x) = x^3$ and $f(x) = x^3 + 2$ have derivative $3x^2$. Starting from $3x^2$ we cannot pin down f — but every candidate differs only by a constant c.

A **boundary condition** fixes c. If $f(1) = 3$, then $1 + c = 3$, so $c = 2$ and $f(x) = x^3 + 2$.

Basic integrals

$$\int c\,\mathrm{d}x = cx + k$$

$$\int x^n\,\mathrm{d}x = \frac{x^{n+1}}{n+1} + k, \quad n \neq -1$$

$$\int \frac{1}{x}\,\mathrm{d}x = \ln|x| + k$$

$$\int e^x\,\mathrm{d}x = e^x + k$$

$$\int e^{ax}\,\mathrm{d}x = \frac{1}{a}e^{ax} + k$$

General rules

$$\int (f \pm g)\,\mathrm{d}x = \int f\,\mathrm{d}x \pm \int g\,\mathrm{d}x$$

$$\int cf(x)\,\mathrm{d}x = c\int f(x)\,\mathrm{d}x$$

$$\int fg'\,\mathrm{d}x = fg - \int f'g\,\mathrm{d}x$$

(integration by parts)

$$\int f(g(x))g'(x)\,\mathrm{d}x = \int f(u)\,\mathrm{d}u$$

(integration by substitution)

Indefinite integrals let us compute **definite integrals**. If $F'(x) = f(x)$, then

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$$

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

Properties

$$\int_a^a f\,\mathrm{d}x = 0, \qquad \int_a^b f\,\mathrm{d}x = -\int_b^a f\,\mathrm{d}x, \qquad \int_a^b f\,\mathrm{d}x = \int_a^c f\,\mathrm{d}x + \int_c^b f\,\mathrm{d}x$$

$$\int_a^b (f \pm g)\,\mathrm{d}x = \int_a^b f\,\mathrm{d}x \pm \int_a^b g\,\mathrm{d}x, \qquad \int_a^b cf\,\mathrm{d}x = c\int_a^b f\,\mathrm{d}x$$

End of induction course

Linear algebra · Linear systems · Eigenvalues

Differential calculus · Integration