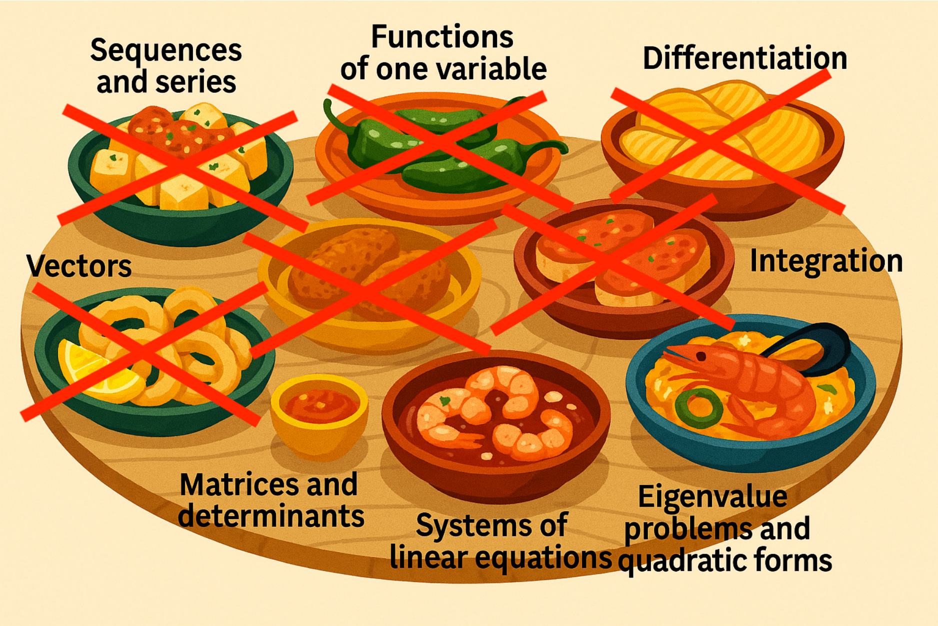
Lecture 4: Calculus and Linear Algebra

Chiara Amorino
Mathematics Brush-up

Barcelona School of Economics

Barcelona School of Economics

Course content (math tapas)



Linear subspaces

Linear subspace

The set $V \subset \mathbb{R}^n$ is a linear subspace of $\mathbb{R}^n$ if

$$\forall \boldsymbol{u}, \boldsymbol{v} \in V \quad \forall a, b \in \mathbb{R} \qquad a\boldsymbol{u} + b\boldsymbol{v} \in V.$$

(closure under taking linear combinations)

There are two ways we can describe a linear subspace V:

- $V = \mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ for some vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m \in V$.
- Solution set of $A\boldsymbol{x} = \mathbf{0}_m$ for some $A \in \mathbb{R}^{m \times n}$.

Some examples:

- kernel of $A \in \mathbb{R}^{m \times n}$
- image of A
- Points in $\mathbb{R}^3$ satisfying $x_1 + x_2 + x_3 = 0$.

Why subspaces must have this form?

Let $V \subseteq \mathbb{R}^n$ be a linear subspace.

Take any nonzero $\boldsymbol{v}_1 \in V$. Then $\mathrm{span}(\boldsymbol{v}_1) \subseteq V$.

If $\mathrm{span}(\boldsymbol{v}_1) \subseteq V$, we are done.

If not, take any nonzero $\boldsymbol{v}_2 \in V \setminus \mathrm{span}(\boldsymbol{v}_1)$.

Then $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ is linearly independent and again $\mathrm{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\} \subseteq V$.

We proceed iteratively until we reach step m such that

$$V \;=\; \mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}.$$

This procedure must stop as we cannot have more than n linearly independent vectors in $V \subseteq \mathbb{R}^n$.

System of linear equations

Systems of linear equations

Let A be an $m \times n$ matrix, $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ unknown, and $\mathbf{b} \in \mathbb{R}^m$ given. The system

$$A\mathbf{x} = \mathbf{b}$$

has the vector form

$$\sum_{j=1}^{n} x_j \, \mathbf{a}_j = \mathbf{b},$$

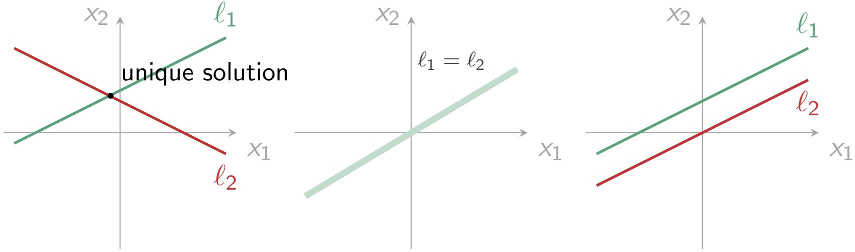
where $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are the columns of A.

If $\mathbf{b} = \mathbf{0}$ the system is homogeneous; otherwise it is inhomogeneous.

An inhomogeneous system can be inconsistent (no solution) or consistent (at least one solution).

A homogeneous system always has $\mathbf{x} = \mathbf{0}$ as a solution.

Geometry of 2×2 linear systems



Unique solution | Infinitely many solutions | No solution (inconsistent)

Key: Lines are the solution sets of single linear equations. Intersecting $\Rightarrow$ one point. Coincident $\Rightarrow$ all points. Parallel $\Rightarrow$ none.

Augmented matrix; consistency

The augmented matrix of the linear system $A\boldsymbol{x} = \boldsymbol{b}$ is

$$A_{\mathbf{b}} = \begin{bmatrix} A & \mathbf{b} \end{bmatrix} \in \mathbb{R}^{m \times (n+1)}.$$

Remark: Either $\operatorname{rank}(A_{\mathbf{b}}) = \operatorname{rank}(A)$ or $\operatorname{rank}(A_{\mathbf{b}}) = \operatorname{rank}(A) + 1$.

Theorem (Rouché–Capelli)

$A\mathbf{x} = \mathbf{b}$ is consistent $\iff$ $\operatorname{rank}(A_{\mathbf{b}}) = \operatorname{rank}(A)$ (equivalently $\mathbf{b} \in \operatorname{span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$).

Moreover:

1. If $\operatorname{rank}(A) = \operatorname{rank}(A_{\mathbf{b}}) = n$, the solution is unique. Q: What is it?
2. If $\operatorname{rank}(A) = \operatorname{rank}(A_{\mathbf{b}}) < n$, there are infinitely many solutions.

The rank tells us how many variables are **pivots** (determined by equations).

The remaining variables can be chosen freely, hence called **free variables**.

Solving with Gaussian elimination

Row echelon forms

A matrix is in row echelon form (REF) if:

1. All nonzero rows appear above any zero rows.
2. Each pivot (leading entry in a row) is to the right of the pivot in the row above it.
3. All entries below a pivot are zero.

It is in reduced row echelon form (RREF) if, in addition:

1. Each pivot is equal to 1.
2. Each pivot is the only nonzero entry in its column.

Example:

$$\begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 1 & 4 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{(REF, not RREF).} \qquad \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 5 \end{pmatrix} \quad \text{(RREF).}$$

Gaussian elimination and RREF

Theorem

By applying elementary row operations to the augmented matrix $A_{\mathbf{b}}$, any system $A\mathbf{x} = \mathbf{b}$ can be transformed into an equivalent system in reduced row echelon form (RREF).

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ x_1 - x_2 + 2x_3 &= 2 \\ 4x_1 + 6x_2 - x_3 &= 9 \end{aligned} \quad \iff \quad \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

So the unique solution is $\mathbf{x} = (1, 1, 1)$.

Why equivalence holds:

$A\mathbf{x} = \mathbf{b} \;\Leftrightarrow\; A_{\mathbf{b}} \begin{bmatrix} \mathbf{x} \\ -1 \end{bmatrix} = \mathbf{0}$. Row operations correspond to multiplying $A_{\mathbf{b}}$ on the left by an invertible matrix C, so the solution set is unchanged.

A homogeneous example

Consider

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 + x_6 &= 0, \\ 2x_1 + 2x_2 + 2x_3 + 2x_4 - x_5 - x_6 &= 0, \\ 3x_1 + 3x_2 - x_3 - x_4 - x_5 - x_6 &= 0. \end{aligned}$$

Row-reducing the coefficient matrix gives

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

So there are 3 pivots, $\operatorname{rank}(A) = 3$, and the solution space has dimension $6 - \operatorname{rank}(A) = 3$:

$$(x_1, -x_1,\, x_3, -x_3,\, x_5, -x_5).$$

A basis is $\{(1, -1, 0, 0, 0, 0), (0, 0, 1, -1, 0, 0), (0, 0, 0, 0, 1, -1)\}$.

General solution structure

Theorem

Every solution of $A\mathbf{x} = \mathbf{b}$ can be written as $\mathbf{x} = \mathbf{x}_h + \mathbf{x}_p$, where $\mathbf{x}_h$ is the general solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$ and $\mathbf{x}_p$ is any particular solution of $A\mathbf{x} = \mathbf{b}$.

Corollary: If $A\mathbf{x} = \mathbf{b}$ is consistent, then the set of solutions is an affine subspace parallel to $\ker(A)$.

Example: Three identical equations

$$x_1 + x_2 + x_3 = 1$$

reduce to

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \rightsquigarrow \begin{array}{r} x_1 = 1 - x_2 - x_3, \\ x_2, x_3 \text{ free.} \end{array}$$

So

$$\mathbf{x} = (1, 0, 0) + \lambda_1(-1, 1, 0) + \lambda_2(-1, 0, 1).$$

Recipe: solving $A\mathbf{x} = b$ with Gaussian elimination

1. Write down the augmented matrix $\left[A \mid b\right]$.
2. Use row operations to reduce to RREF.
3. Identify pivot and free variables.
4. Express the solution: $\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h$.

Always

- The number of pivot variables is the rank of A,
- the number of free variables $= n - \text{rank}(A)$,
- the free variables parametrize the solution set.

Two quick examples

1. For $A = \begin{pmatrix} 1 & 2 & 0 \\ 4 & 6 & 2 \\ 3 & 2 & 4 \end{pmatrix}$, RREF gives $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$. So $\operatorname{rank}(A) = 2$, $\dim\ker(A) = 1$, and

$$\ker(A) = \{\lambda(-2, 1, 1) : \ \lambda \in \mathbb{R}\}.$$

2. For $\begin{cases} x_1 + x_2 + x_3 = -1 \\ x_1 + 2x_2 + 4x_3 = 2 \end{cases}$, the augmented matrix reduces to $\left(\begin{array}{ccc|c} 1 & 0 & -2 & -4 \\ 0 & 1 & 3 & 3 \end{array}\right)$. So $\operatorname{rank}(A) = 2$, one free variable $x_3 = \lambda$, and

$$\mathbf{x} = (-4, 3, 0) + \lambda(2, -3, 1).$$

Finding a basis for a span

Let

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ -1 \\ -2 \end{pmatrix}.$$

Put them as columns of a matrix:

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

Row-reduction gives $M \sim \begin{pmatrix} 1 & 0 & -1/3 \\ 0 & 1 & -2/3 \\ 0 & 0 & 0 \end{pmatrix}$.

Thus $\operatorname{rank}(M) = 2$ and there is one linear dependence: $3\mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3 = 0$.
A basis for $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ can be taken as the first two pivot columns, e.g.

$$\{\mathbf{v}_1,\ \mathbf{v}_2\}.$$

Quick recap

- **Homogeneous system:** $A\mathbf{x} = 0$. Always has at least the trivial solution.
- **Inhomogeneous system:** $A\mathbf{x} = b$. May or may not be solvable.
- **Augmented matrix:** $\left[A \mid b\right]$, i.e. stick b as a new column.
- **Rank:** Number of pivot columns after row reduction = dimension of the column space.

Standing on the shoulders of giants

“If I have seen further, it is by standing on the shoulders of giants.” **Isaac Newton (1675)**.

Some of the giants of linear/matrix algebra:

- Gauss (Gaussian elimination, least squares, determinants)
- Cauchy (Cauchy–Schwarz, eigenvalues, determinants)
- Laplace (Laplace expansion of determinants)
- Cayley (Cayley–Hamilton theorem, matrix theory)
- Sylvester (matrix rank, invariant theory)
- Jordan (Jordan canonical form)
- Frobenius (matrix factorizations, linear algebra foundations)

(pictured here as the base of a *castell*)

Chapter 8: Eigenvalues and Quadratic Forms

Eigenvalues/eigenvectors appear in growth models, Markov chains, PCA, and stability of dynamical systems.

Quadratic forms drive optimization via second-order conditions.

Read Chapter 10 of Werner–Sotskov (optional: Simon–Blume, Ch. 7).

Exercises 10.2; 10.5 A,B,C; 10.6(a,b).

Eigenvalues and eigenvectors

Eigenvalues and eigenvectors: definitions

For $A \in \mathbb{R}^{n\times n}$, a scalar λ is an eigenvalue if there exists $\mathbf{x} \neq \mathbf{0}$ with

$$A\mathbf{x} = \lambda\mathbf{x} \quad \Longleftrightarrow \quad (A - \lambda I_n)\mathbf{x} = \mathbf{0}.$$

Then $\mathbf{x}$ is an eigenvector for λ.

Characterization:

$$\lambda \text{ eigenvalue} \iff \operatorname{rank}(A - \lambda I_n) < n \iff \det(A - \lambda I_n) = 0.$$

The characteristic polynomial $P(\lambda) = \det(A - \lambda I_n)$ has degree n; its real roots (at most n) are the real eigenvalues.

Procedure:

1. Solve $\det(A - \lambda I_n) = 0$ for eigenvalues.
2. For each λ, solve $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ for eigenvectors.

e,g. $\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$, $\begin{bmatrix}2 & 3\\ 1 & 4\end{bmatrix}$, $\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}$

Why eigenvalues matter (at a glance)

If $A\boldsymbol{x} = \lambda\boldsymbol{x}$ then $A^k\boldsymbol{x} = A^{k-1}(A\boldsymbol{x}) = \lambda A^{k-1}\boldsymbol{x} = \cdots = \lambda^k\boldsymbol{x}$.

- **Markov chains:** steady state $=$ eigenvector for $\lambda = 1$.
- **Growth/dynamics:** long-run rate $\approx$ spectral radius $\rho(A)$.
- **PCA:** principal directions $=$ eigenvectors of covariance; explained variances $=$ eigenvalues.
- **Quadratic forms:** curvature of $x^\top Ax$ aligned with eigenvectors; definiteness from eigenvalue signs.

Eigenvalues and eigenvectors: migration example

Two regions: city (x_t^C) and countryside (x_t^R). Let $\boldsymbol{x}_t = (x_t^C, x_t^R)^\top$. Assume next-year proportions follow a **column-stochastic** transition:

$$\boldsymbol{x}_{t+1} = A\,\boldsymbol{x}_t, \qquad A = \begin{pmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{pmatrix},$$

so each column sums to 1 (fractions of people who *arrive* from each region).

A stable composition $\boldsymbol{u}$ satisfies $A\boldsymbol{u} = \boldsymbol{u}$ and $u_1 + u_2 = 1$:

$$(A - I)\boldsymbol{u} = 0 \;\Rightarrow\; \boldsymbol{u} = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}.$$

Spectrum: $\lambda_1 = 1$, $\lambda_2 = 0.4$.

Any initial x_0 with positive entries converges to u.

- Check $\boldsymbol{v} = (1, -1)$ is the 0.4-eigenvector.
- $\boldsymbol{x}_0$ can be written as $\boldsymbol{x}_0 = \boldsymbol{u} + \lambda\boldsymbol{v}$
- $\boldsymbol{x}_t = A^t(\boldsymbol{u} + \lambda\boldsymbol{v}) = \boldsymbol{u} + \lambda \cdot 0.4^t \cdot \boldsymbol{v}$.
- $\|\boldsymbol{x}_t - \boldsymbol{u}\| = |\lambda| \cdot 0.4^t\sqrt{2} \to 0$.

Basic spectral facts

For $A \in \mathbb{R}^{n \times n}$

1. Eigenvectors associated with distinct eigenvalues are linearly independent.
2. If A is symmetric, then all eigenvalues are real and eigenvectors for distinct eigenvalues are orthogonal (Spectral Theorem).

Example:

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

Then $P(\lambda) = \lambda^2 - 2\lambda - 3$, so $\lambda_1 = 3$, $\lambda_2 = -1$ with eigenvectors along $(1, 1)$ and $(-1, 1)$, respectively.

Matrix diagonalization

Diagonalizable or not?

Definitions:

- $A \in \mathbb{R}^{n \times n}$ is **diagonalizable** if there exists an invertible P such that

$$P^{-1}AP = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

  Equivalently, A has a basis of eigenvectors.
- For each eigenvalue λ, the **eigenspace** is

$$E_\lambda = \ker(A - \lambda I).$$

  Its dimension is the **geometric multiplicity**.
- The number of times λ appears as a root of the characteristic polynomial is its **algebraic multiplicity**.

Why it matters: diagonalization simplifies A^k, e^A.

Fact:

A is diagonalizable $\iff$ $\sum_\lambda \dim E_\lambda = n$.

Example 1 (diagonalizable).

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

Characteristic polynomial: $(2-\lambda)^2 - 1$. Distinct eigenvalues $\Rightarrow$ eigenvectors span $\mathbb{R}^2$. So A is diagonalizable: $A = PDP^{-1}$ with $D = \mathrm{diag}(3,1)$ and $P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Example 2 (defective).

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Characteristic polynomial: $(\lambda - 1)^2$ (algebraic multiplicity 2). Eigenspace: $\ker(B - I) = \mathrm{span}\{(1,0)\}$ has dimension 1. So B is **not** diagonalizable.

Repeated eigenvalues

Consider

$$A = \begin{pmatrix} 0 & -1 & 1 \\ -7 & 0 & 5 \\ -5 & -2 & 5 \end{pmatrix}.$$

The characteristic polynomial is

$$p_A(\lambda) = \lambda^3 - 5\lambda^2 + 8\lambda - 4 = (\lambda - 2)^2(\lambda - 1).$$

Hence the eigenvalues are $\lambda = 2$ (algebraic mult. 2) and $\lambda = 1$ (algebraic mult. 1).

Eigenspaces:

$$E_2 = \ker(A - 2I) = \operatorname{span}\{(1, -1, 1)\}, \qquad E_1 = \ker(A - I) = \operatorname{span}\{(2, 1, 3)\}.$$

Since $\dim E_2 = 1 < 2$, A is defective.