MATH0047 Advanced Linear Algebra
Revision Note

Ambrose W

May 12, 2026

Contents

1 Chapter 1: Matrices and linear equations

1.1 Core matrix language

Definition 1.1 (Matrix)**.** A matrix is a rectangular array of numbers.

Definition 1.2 (Square matrix)**.** A matrix is *square* if it has the same number of rows and columns, i.e. if it has size $n \times n$ for some n.

Definition 1.3 (Equality of matrices)**.** Two matrices A and B are equal if they have the same size and

$$A_{ij} = B_{ij} \qquad \text{for all relevant } i, j.$$

Definition 1.4 (Matrix addition)**.** If A and B have the same size $m \times n$, then their sum is the $m \times n$ matrix defined by

$$(A + B)_{ij} = A_{ij} + B_{ij} \qquad (1 \le i \le m,\ 1 \le j \le n).$$

Definition 1.5 (Scalar multiplication)**.** If A is an $m \times n$ matrix and $c \in \mathbb{C}$, then the scalar multiple cA is defined entrywise by

$$(cA)_{ij} = c(A_{ij}).$$

Definition 1.6 (Matrix multiplication)**.** If A is an $m \times r$ matrix and B is an $r \times n$ matrix, then the product AB is the $m \times n$ matrix defined by

$$(AB)_{ij} = \sum_{k=1}^{r} A_{ik} B_{kj}.$$

The inner dimensions must agree: the number of columns of A must equal the number of rows of B.

Definition 1.7 (Trace)**.** If A is a square matrix of size $n \times n$, then the trace of A is

$$\operatorname{tr}(A) = A_{11} + A_{22} + \cdots + A_{nn} = \sum_{i=1}^{n} A_{ii}.$$

Definition 1.8 (Transpose)**.** If A is an $m \times n$ matrix, then the transpose of A, denoted by A^T, is the $n \times m$ matrix defined by

$$(A^T)_{ij} = A_{ji}.$$

Exam tip. For entrywise matrix proofs, the standard pattern is: fix i, j, compute the (i, j)-entry of the left-hand side, simplify using the relevant definitions, and identify the (i, j)-entry of the right-hand side.

Proposition 1.9 (Commutativity of matrix addition)**.** *Suppose that A and B are $m \times n$ matrices. Then*

$$A + B = B + A.$$

Proof required

Proof. For all i, j,

$$(A + B)_{ij} = A_{ij} + B_{ij} = B_{ij} + A_{ij} = (B + A)_{ij}.$$

Hence $A + B = B + A$ by Definition 1.3. □

Proposition 1.10 (Associativity of matrix addition). *Suppose that A, B, C are $m \times n$ matrices. Then*

$$A + (B + C) = (A + B) + C.$$

Proof included here

Proof. For all i, j,

$$(A + (B + C))_{ij} = A_{ij} + (B_{ij} + C_{ij}) = (A_{ij} + B_{ij}) + C_{ij} = ((A + B) + C)_{ij}.$$

Therefore $A + (B + C) = (A + B) + C$. □

Proposition 1.11 (Left distributivity over matrix addition). *Suppose that A is an $m \times r$ matrix and that B, C are $r \times n$ matrices. Then*

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

Proof required

Proof. For all i, j,

$$\begin{aligned}(A(B + C))_{ij} &= \sum_{k=1}^{r} A_{ik}(B + C)_{kj} = \sum_{k=1}^{r} A_{ik}(B_{kj} + C_{kj}) \\ &= \sum_{k=1}^{r}(A_{ik}B_{kj} + A_{ik}C_{kj}) = \sum_{k=1}^{r} A_{ik}B_{kj} + \sum_{k=1}^{r} A_{ik}C_{kj} \\ &= (AB)_{ij} + (AC)_{ij} = (AB + AC)_{ij}.\end{aligned}$$

Hence $A(B + C) = AB + AC$. □

Proposition 1.12 (Right distributivity over matrix addition). *Suppose that A is an $r \times n$ matrix and that B, C are $m \times r$ matrices. Then*

$$(B + C)A = BA + CA.$$

Proof included here

Proof. For all i, j,

$$((B + C)A)_{ij} = \sum_{k=1}^{r}(B + C)_{ik}A_{kj} = \sum_{k=1}^{r}(B_{ik} + C_{ik})A_{kj} = (BA + CA)_{ij}.$$

□

Proposition 1.13 (Associativity of matrix multiplication). *Suppose that A is an $m \times r$ matrix, B is an $r \times s$ matrix, and C is an $s \times n$ matrix. Then*

$$A(BC) = (AB)C.$$

Proof required

Proof. For all i, j,

$$\begin{aligned}(A(BC))_{ij} &= \sum_{k=1}^{r} A_{ik}(BC)_{kj} = \sum_{k=1}^{r} A_{ik}\left(\sum_{\ell=1}^{s} B_{k\ell}C_{\ell j}\right)\\ &= \sum_{k=1}^{r}\sum_{\ell=1}^{s} A_{ik}B_{k\ell}C_{\ell j} = \sum_{\ell=1}^{s}\sum_{k=1}^{r} A_{ik}B_{k\ell}C_{\ell j}\\ &= \sum_{\ell=1}^{s}\left(\sum_{k=1}^{r} A_{ik}B_{k\ell}\right)C_{\ell j} = \sum_{\ell=1}^{s}(AB)_{i\ell}C_{\ell j} = ((AB)C)_{ij}.\end{aligned}$$

Hence $A(BC) = (AB)C$. ☐

Remark 1.14 (A warning about multiplication)**.** Matrix multiplication is associative, but it is not commutative in general. Matrix addition is both associative and commutative.

Proposition 1.15 (Distributivity of scalar addition)**.** *Suppose that A is an $m \times n$ matrix and that $c_1, c_2 \in \mathbb{C}$. Then*

$$(c_1 + c_2)A = c_1A + c_2A.$$

Proof required

Proof. For all i, j,

$$((c_1 + c_2)A)_{ij} = (c_1 + c_2)A_{ij} = c_1A_{ij} + c_2A_{ij} = (c_1A + c_2A)_{ij}.$$

☐

Proposition 1.16 (Associativity of scalar multiplication)**.** *Suppose that A is an $m \times n$ matrix and that $c_1, c_2 \in \mathbb{C}$. Then*

$$(c_1c_2)A = c_1(c_2A).$$

Proof included here

Proof. For all i, j,

$$((c_1c_2)A)_{ij} = (c_1c_2)A_{ij} = c_1(c_2A_{ij}) = (c_1(c_2A))_{ij}.$$

☐

Proposition 1.17 (Scalar multiplication distributes over matrix addition)**.** *Suppose that A and B are matrices of the same size and that $c \in \mathbb{C}$. Then*

$$c(A + B) = cA + cB.$$

Proof included here

Proof. For all i, j,

$$(c(A + B))_{ij} = c(A_{ij} + B_{ij}) = cA_{ij} + cB_{ij} = (cA + cB)_{ij}.$$

☐

Proposition 1.18 (Scalar multiplication and matrix products)**.** *Suppose that A is an $m \times r$ matrix, B is an $r \times n$ matrix, and $c \in \mathbb{C}$. Then*

$$c(AB) = (cA)B.$$

Proof included here

Proof. For all i, j,

$$(c(AB))_{ij} = c\sum_{k=1}^{r} A_{ik}B_{kj} = \sum_{k=1}^{r}(cA_{ik})B_{kj} = ((cA)B)_{ij}.$$

$\square$

Editorial note. By the same entrywise argument,

$$c(AB) = A(cB)$$

as well.

Proposition 1.19 (Double transpose). *For any matrix A,*

$$(A^T)^T = A.$$

Proof required

Proof. For all i, j,

$$((A^T)^T)_{ij} = (A^T)_{ji} = A_{ij}.$$

$\square$

Proposition 1.20 (Transpose of a sum). *If A and B have the same size, then*

$$(A + B)^T = A^T + B^T.$$

Proof included here

Proof. For all i, j,

$$((A + B)^T)_{ij} = (A + B)_{ji} = A_{ji} + B_{ji} = (A^T + B^T)_{ij}.$$

$\square$

Proposition 1.21 (Transpose of a product). *Suppose that A is an $m \times r$ matrix and B is an $r \times n$ matrix. Then*

$$(AB)^T = B^T A^T.$$

Proof required

Proof. For all i, j,

$$((AB)^T)_{ij} = (AB)_{ji} = \sum_{k=1}^{r} A_{jk}B_{ki} = \sum_{k=1}^{r}(B^T)_{ik}(A^T)_{kj} = (B^T A^T)_{ij}.$$

$\square$

Proposition 1.22 (Cyclic trace identity). *Suppose that A is an $m \times n$ matrix and B is an $n \times m$ matrix. Then*

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

Proof included here

Proof. Using the definition of trace and matrix multiplication,

$$\operatorname{tr}(AB) = \sum_{i=1}^{m}(AB)_{ii} = \sum_{i=1}^{m}\sum_{k=1}^{n} A_{ik}B_{ki} = \sum_{k=1}^{n}\sum_{i=1}^{m} B_{ki}A_{ik} = \sum_{k=1}^{n}(BA)_{kk} = \operatorname{tr}(BA).$$

$\square$

Editorial note. A useful immediate identity is $\operatorname{tr}(A^T) = \operatorname{tr}(A)$ for every square matrix A.

1.2 Zero matrices, identity matrices, and inverses

Definition 1.23 (Zero matrix)**.** A matrix, of any size, with every entry equal to 0 is called a *zero matrix*. The same symbol 0 is often used for whichever zero matrix is required by the context.

Definition 1.24 (Identity matrix)**.** An identity matrix is a square matrix with 1 on the main diagonal and 0 everywhere else. The identity matrix of size $n \times n$ is denoted by I_n.

Proposition 1.25 (Identity matrices act as multiplicative identities)**.** *Let I_m and I_n be the $m \times m$ and $n \times n$ identity matrices. Then for every $m \times n$ matrix A,*

$$I_m A = A \qquad \textit{and} \qquad A I_n = A.$$

Proof included here

Proof. For all i, j,

$$(I_m A)_{ij} = \sum_{k=1}^{m} (I_m)_{ik} A_{kj} = A_{ij}$$

because $(I_m)_{ik} = 0$ unless $k = i$, and $(I_m)_{ii} = 1$. The proof of $AI_n = A$ is identical:

$$(AI_n)_{ij} = \sum_{k=1}^{n} A_{ik} (I_n)_{kj} = A_{ij}.$$

□

Definition 1.26 (Invertible matrix)**.** Let A be a square matrix of size $n \times n$. We say that A is *invertible* if there exists an $n \times n$ matrix B such that

$$AB = I_n \qquad \text{and} \qquad BA = I_n.$$

In that case B is called the inverse of A and is denoted by A^{-1}.

Proposition 1.27 (Uniqueness of inverses)**.** *Suppose that B and C are both inverses of a matrix A. Then $B = C$.* *Proof required*

Proof. We assume that B and C are both inverse matrices for A. By the definition of an inverse:

$$\begin{aligned} AB = I \text{ and } BA = I \quad &(\text{since } B \text{ is an inverse of } A) \\ AC = I \text{ and } CA = I \quad &(\text{since } C \text{ is an inverse of } A). \end{aligned}$$

Then, we can show that $B = C$ by using the fact that $(BA)C = B(AC)$ (i.e. by using the fact that matrix multiplication is associative):

We note that:

$$\begin{aligned} (BA)C &= IC = C && \text{using } BA = I, \\ B(AC) &= BI = B && \text{using } AC = I. \end{aligned}$$

So, using the fact that $(BA)C = B(AC)$, we conclude that $B = C$, as required. □

Proposition 1.28 (Inverse of a product of two invertible matrices)**.** *Suppose that A and B are invertible matrices of the same size. Then AB is invertible and*

$$(AB)^{-1} = B^{-1} A^{-1}.$$

Proof required

Proof. We assume that A and B are invertible. Hence, there exists

$$\begin{gathered} A^{-1} \text{ such that } AA^{-1} = I \text{ and } A^{-1}A = I, \\ B^{-1} \text{ such that } BB^{-1} = I \text{ and } B^{-1}B = I. \end{gathered}$$

We compute both products:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = I$$

and

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = I.$$

So $B^{-1}A^{-1}$ is the inverse of AB. ☐

Proposition 1.29 (Inverse of a product of finitely many invertible matrices)**.** *Suppose that* $A_1, \ldots, A_n$ *are invertible matrices of the same size. Then*

$$(A_1 \cdots A_n)^{-1} = A_n^{-1} \cdots A_1^{-1}.$$

Proof included here

Proof. Apply Proposition 1.28 repeatedly. For example,

$$(A_1A_2A_3)^{-1} = A_3^{-1}(A_1A_2)^{-1} = A_3^{-1}A_2^{-1}A_1^{-1},$$

and induction gives the general formula. ☐

Exam workflow: Nilpotent inverses

Key idea. Do *not* try to row-reduce. Questions of this type ask for a *polynomial identity in* A. Because $A^m = 0$, every power $A^m, A^{m+1}, \ldots$ vanishes, so the inverse is a short polynomial in A.

General method (illustrated for $A^3 = 0$, finding $(A + I)^{-1}$).

Step 1. Recognise that A is nilpotent: since $A^3 = 0$, we have $A^3 = A^4 = A^5 = \cdots = 0$. Any inverse should be expressible using only I, A, A^2.

Step 2. Guess the form $(A + I)^{-1} = c_0I + c_1A + c_2A^2$ and expand:

$$(A + I)(c_0I + c_1A + c_2A^2) = c_0I + (c_0 + c_1)A + (c_1 + c_2)A^2 + c_2A^3.$$

Since $A^3 = 0$ this equals $c_0I + (c_0 + c_1)A + (c_1 + c_2)A^2$.

Step 3. Match coefficients to I:

$$c_0 = 1, \qquad c_0 + c_1 = 0, \qquad c_1 + c_2 = 0 \implies c_0 = 1,\ c_1 = -1,\ c_2 = 1.$$

Exam-style verification (always check both sides). Since the definition of inverse requires $BC = CB = I$, verify both products:

$$(A + I)(A^2 - A + I) = A^3 - A^2 + A + A^2 - A + I = A^3 + I = I,$$

$$(A^2 - A + I)(A + I) = A^3 + A^2 - A^2 - A + A + I = A^3 + I = I.$$

Therefore $A + I$ is invertible and

$$\boxed{(A + I)^{-1} = A^2 - A + I.}$$

Finding $(A^2 - I)^{-1}$ when $A^8 = 0$. Treat A^2 as the variable in the finite geometric series and build the product $(A^2 - I)(A^6 + A^4 + A^2 + I)$ row by row:

$$\begin{aligned} (A^2 - I)\,A^6 &= A^8 - A^6, \\ (A^2 - I)\,A^4 &= \quad A^6 - A^4, \\ (A^2 - I)\,A^2 &= \quad A^4 - A^2, \\ (A^2 - I)\,I &= \quad A^2 - I. \end{aligned}$$

Adding the four rows, the interior terms A^6, A^4, A^2 cancel in telescoping pairs, leaving

$$(A^2 - I)(A^6 + A^4 + A^2 + I) = A^8 - I = -I,$$

since $A^8 = 0$. Multiplying both sides by -1 gives

$$(A^2 - I)\big(-A^6 - A^4 - A^2 - I\big) = I,$$

hence $A^2 - I$ is invertible and

$$\boxed{(A^2 - I)^{-1} = -A^6 - A^4 - A^2 - I.}$$

General finite geometric series identities. Whenever $A^m = 0$:

$$(I - A)^{-1} = I + A + A^2 + \cdots + A^{m-1},$$

$$(I + A)^{-1} = I - A + A^2 - \cdots + (-1)^{m-1}A^{m-1}.$$

For $A^3 = 0$ these reduce to $(I - A)^{-1} = I + A + A^2$ and $(I + A)^{-1} = I - A + A^2$.

1.3 Systems of linear equations and row reduction

Definition 1.31 (Complex linear equation)**.** A complex linear equation is an equation of the form

$$a_1x_1 + \cdots + a_nx_n = b,$$

where $a_1, \ldots, a_n, b \in \mathbb{C}$.

Definition 1.32 (Real linear equation)**.** A real linear equation is an equation of the form

$$a_1x_1 + \cdots + a_nx_n = b,$$

where $a_1, \ldots, a_n, b \in \mathbb{R}$.

Definition 1.33 (Solution of a real linear equation)**.** A solution of the real linear equation $a_1x_1 + \cdots + a_nx_n = b$ is a choice of real numbers $s_1, \ldots, s_n$ such that

$$a_1s_1 + \cdots + a_ns_n = b.$$

Definition 1.34 (System of real linear equations)**.** A system of real linear equations is a finite list of m real linear equations in n unknowns:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots &= \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

It can be written as the matrix equation

$$Ax = b.$$

Definition 1.35 (Real elementary row operation)**.** A real elementary row operation is one of the following:

1. multiply a row by a non-zero real number;
2. add a real multiple of one row to another row;
3. interchange (i.e. permutate) two rows.

The course notation is

$$R_i \mapsto \lambda R_i, \qquad R_i \mapsto R_i + \lambda R_j, \qquad R_i \leftrightarrow R_j.$$

Definition 1.36 (Reduced row echelon form)**.** A matrix is in reduced row echelon form if it satisfies all four conditions below:

1. any zero rows appear at the bottom;
2. each non-zero row begins with a leading 1;
3. leading 1 of the lower row is strictly to the right of the leading 1 of the row above;
4. each column containing a leading 1 has zeros everywhere else.

Definition 1.37 (Row echelon form)**.** A matrix is in row echelon form if it satisfies the first three conditions in Definition 1.36.

Proposition 1.38 (Uniqueness of reduced row echelon form)**.** *A matrix has a unique reduced row echelon form.* Statement only

Exam workflow: Gauss–Jordan elimination

To solve $Ax = b$, row reduce the augmented matrix $(A \,|\, b)$ to reduced row echelon form.

- A row of the form $[0 \ \cdots \ 0 \,|\, c]$ with $c \neq 0$ represents the impossible equation $0 = c$, so the system is *inconsistent* (a system of real linear equations is inconsistent if no real solution exists).
- If every variable column contains a leading 1, the solution is unique.
- If at least one variable is free, the system has infinitely many solutions. Express the pivot variables in terms of the free variables and then parameterise the full vector solution.

Notice that you MUST reduce COMPLETELY when you are asked to *'Use the Gauss-Jordan process'* in the exam.

Proposition 1.43 (Possible numbers of solutions)**.** *A system of real linear equations has no solutions, one unique solution, or infinitely many solutions.* Statement only

> If a system of real linear equations has two distinct solutions, then any affine combination of them is also a solution. That is, if u and v are solutions, then for any $\lambda \in \mathbb{R}$,
>
> $$\lambda u + (1 - \lambda) v$$
>
> is also a solution.

Proposition 1.44 (Homogeneous systems)**.** *A homogeneous system of real linear equations either has the unique solution (the zero solution) or infinitely many solutions.* Statement only

Exam tip. A homogeneous system always has the zero solution, so "no solutions" is impossible in the homogeneous case.

1.4 Elementary matrices and the invertibility theorem

Definition 1.45 (Elementary matrix)**.** A real square matrix of size $n \times n$ is *elementary* if it is one of the following:

1. $D_n(i;\lambda)$: the same as I_n, except that the entry in row i, column i is $\lambda \neq 0$;
2. $E_n(i,j;\lambda)$: the same as I_n, except that the entry in row i, column j is λ with $i \neq j$;
3. $P_n(i,j)$: the same as I_n, except that rows i and j are interchanged.

They correspond respectively to the row operations

$$R_i \mapsto \lambda R_i, \qquad R_i \mapsto R_i + \lambda R_j, \qquad R_i \leftrightarrow R_j.$$

Remark 1.46 (Dropping the size subscript)**.** When the matrix size is clear from the context, the notes often write $D(i;\lambda)$, $E(i,j;\lambda)$, and $P(i,j)$ instead of $D_n(i;\lambda)$, $E_n(i,j;\lambda)$, and $P_n(i,j)$.

Remark 1.47 (Elementary matrices act on the left)**.** Left multiplication by an elementary matrix performs the corresponding row operation. Multiplying on the right performs the analogous operation on columns, not on rows.

Proposition 1.49 (Inverses of elementary matrices)**.** *Each elementary matrix is invertible, and its inverse is elementary of the same type:*

$$D_n(i;\lambda)^{-1} = D_n(i;\lambda^{-1}), \qquad E_n(i,j;\lambda)^{-1} = E_n(i,j;-\lambda), \qquad P_n(i,j)^{-1} = P_n(i,j).$$

Proof required

Proof. Each inverse simply undoes the corresponding row operation. Multiplying row i by λ is undone by multiplying by λ^{-1}; adding λ times row j to row i is undone by adding $-\lambda$ times row j to row i; and swapping the same two rows twice returns the original matrix. ☐

Theorem 1.52 (Equivalent characterisations of invertibility)**.** *Suppose that A is a real square $n \times n$ matrix, and let $x, b \in \mathbb{R}^n$. The following statements are equivalent:*

1. *A is invertible.*
2. *The matrix equation $Ax = b$ has a unique solution for every $b \in \mathbb{R}^n$.*
3. *The reduced row echelon form of A is I_n.*
4. *A can be expressed as a product of elementary matrices.*

Proof required

Proof. It suffices to establish the cyclic chain (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (1), since this connects any two statements in both directions.

(1) $\Rightarrow$ (2): Since A is invertible, A^{-1} exists. Setting $x = A^{-1}b$ gives

$$Ax = A(A^{-1}b) = (AA^{-1})b = I_n b = b,$$

so $A^{-1}b$ is a solution. For uniqueness, suppose x and y both satisfy $Ax = Ay = b$. Left-multiplying by A^{-1} produces $(A^{-1}A)x = (A^{-1}A)y$, that is, $x = y$. The solution therefore exists and is unique.

(2) $\Rightarrow$ (3): Let $A' = \mathrm{RREF}(A)$, and let $(A' \,|\, b')$ denote the reduced row echelon form of the augmented matrix $(A \,|\, b)$. Argue by contradiction: suppose A' contained a zero row. For the corresponding row of the augmented system, exactly one of two situations would arise:

- the zero row of A' is paired with a non-zero entry in b'. The resulting equation reads $0 = c$ with $c \neq 0$, so $Ax = b$ is inconsistent and admits no solution—contradicting unique solvability;
- the zero row of A' is paired with zero in b'. Then A' contains fewer than n leading 1s, so at least one column lacks a leading 1 and produces a free variable, hence infinitely many solutions—again contradicting uniqueness.

Either alternative violates the hypothesis, so A' has no zero row, and each of its n rows carries a leading 1. The only $n \times n$ matrix in reduced row echelon form with this property is I_n; hence $A' = I_n$.

(3) $\Rightarrow$ (4): If $\mathrm{RREF}(A) = I_n$, a sequence of elementary row operations transforms A into I_n. Recording each operation as left multiplication by an elementary matrix produces $E_1, \ldots, E_k$ such that

$$E_k \cdots E_1 A = I_n.$$

Each E_i is invertible by Proposition 1.49, so the product $E_k \cdots E_1$ is invertible. Left-multiplying both sides by $(E_k \cdots E_1)^{-1}$ and applying Proposition 1.29 gives

$$A = (E_k \cdots E_1)^{-1} = E_1^{-1} \cdots E_k^{-1}.$$

By Proposition 1.49 again, the inverse of an elementary matrix is itself elementary, so every factor E_i^{-1} is elementary and A is expressed as a product of elementary matrices.

(4) $\Rightarrow$ (1): Suppose $A = F_1 \cdots F_m$ with each F_i elementary. Each F_i is invertible by Proposition 1.49, and Proposition 1.28 applied inductively shows that a product of invertible matrices is invertible. Hence A is invertible, with inverse $A^{-1} = F_m^{-1} \cdots F_1^{-1}$. □

2 Chapter 2: Determinants

2.1 Definition of the determinant

Definition 2.1 (Permutation)**.** A permutation of $\{1, \ldots, n\}$ is an arrangement of the numbers $1, \ldots, n$ in some order, without omissions or repetitions.

Proposition 2.2 (Number of permutations)**.** *There are $n!$ permutations of n numbers.* Proof required

Proof. There are n choices for the first position, then $n-1$ for the second, then $n-2$ for the third, and so on down to 1 choice for the last. Multiplying gives $n!$. □

Definition 2.3 (Inversion)**.** An inversion in a permutation is a pair that appears in the wrong order, in the sense that a larger number comes before a smaller one.

Definition 2.4 (Sign of a permutation)**.** If a permutation ϑ has k inversions, then its sign is

$$\operatorname{sgn}(\vartheta) = (-1)^k.$$

So an even number of inversions gives sign $+1$, and an odd number of inversions gives sign -1.

Definition 2.5 (Elementary product)**.** An elementary product from a $n \times n$ matrix A is a product of n entries of A, such that there is exactly one entry from each row and each column.

Definition 2.6 (Signed elementary product)**.** A signed elementary product is an elementary product multiplied by the sign of the permutation that selects its columns.

Definition 2.7 (Determinant)**.** If A is a square $n \times n$ matrix, then the determinant of A is the sum of all signed elementary product from A,

$$\det(A) = \sum_{\vartheta \in S_n} \operatorname{sgn}(\vartheta) A_{1\vartheta(1)} A_{2\vartheta(2)} \cdots A_{n\vartheta(n)}.$$

Definition 2.8 (Minor)**.** Let A be a square $n \times n$ matrix. For each entry A_{ij}, let $A(\hat{\imath}, \hat{\jmath})$ denote the $(n-1) \times (n-1)$ matrix obtained by deleting row i and column j. The determinant

$$\det\big(A(\hat{\imath}, \hat{\jmath})\big)$$

is the *minor* of A_{ij}.

Definition 2.9 (Cofactor)**.** Let A be a square $n \times n$ matrix. The *cofactor* of the entry A_{ij} is

$$(-1)^{i+j} \det\big(A(\hat{\imath}, \hat{\jmath})\big).$$

Exam workflow: Cofactor expansion

Expanding along row i gives

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det\big(A(\hat{\imath}, \hat{\jmath})\big),$$

and expanding along column j gives

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det\big(A(\hat{\imath}, \hat{\jmath})\big).$$

For a 2×2 matrix,

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

In exam work, cofactor expansion is usually best along a row or column containing many zeros.

2.2 Basic determinant rules

Proposition 2.11 (Zero row or zero column)**.** *Let A be a square matrix. If A has a zero row or a zero column, then*

$$\det(A) = 0.$$

Proof required

Proof. Every signed elementary product contains one entry from each row and one entry from each column. If one whole row or column is zero, every signed elementary product contains a zero factor, so every signed elementary product is zero. Their sum is therefore zero. ☐

Proposition 2.12 (Determinant of a transpose)**.** *Let A be a square $n \times n$ matrix. Then*

$$\det(A) = \det(A^T).$$

Proof included here

Proof. From the definition,

$$\det(A^T) = \sum_{\vartheta \in S_n} \operatorname{sgn}(\vartheta)(A^T)_{1\vartheta(1)} \cdots (A^T)_{n\vartheta(n)} = \sum_{\vartheta \in S_n} \operatorname{sgn}(\vartheta) A_{\vartheta(1)1} \cdots A_{\vartheta(n)n}.$$

Now let $\sigma = \vartheta^{-1}$. Then σ also runs through S_n and $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\vartheta)$. The product above becomes

$$A_{1\sigma(1)} \cdots A_{n\sigma(n)},$$

so the sum is exactly $\det(A)$. ☐

Definition 2.13 (Upper triangular matrix)**.** A square matrix A is *upper triangular* if $A_{ij} = 0$ whenever $i > j$.

Definition 2.14 (Lower triangular matrix)**.** A square matrix A is *lower triangular* if $A_{ij} = 0$ whenever $j > i$.

Proposition 2.15 (Determinant of an upper triangular matrix)**.** *If A is an upper triangular $n \times n$ matrix, then*

$$\det(A) = A_{11}A_{22} \cdots A_{nn}.$$

Proof required

Proof. We construct a possibly non-zero elementary product column by column. In column j, the entries in rows $j+1, \dots, n$ are all zero (upper triangular), so we must choose from rows $1, \dots, j$. But rows $1, \dots, j-1$ have already been used by columns $1, \dots, j-1$, so the only available choice is A_{jj}.

Hence the only possibly non-zero elementary product is $A_{11}A_{22} \cdots A_{nn}$, corresponding to the identity permutation $(1, 2, \dots, n)$, which has 0 inversions and sign $+1$. Therefore $\det(A) = A_{11}A_{22} \cdots A_{nn}$. ☐

Proposition 2.16 (Determinant of a lower triangular matrix)**.** *If A is a lower triangular $n \times n$ matrix, then*

$$\det(A) = A_{11}A_{22} \cdots A_{nn}.$$

Proof included here

Proof. Since A^T is upper triangular, Proposition 2.15 and Proposition 2.12 give

$$\det(A) = \det(A^T) = A_{11}A_{22}\cdots A_{nn}.$$

☐

Definition 2.17 (Diagonal matrix)**.** A square matrix A is *diagonal* if it has zero entries both above and below the main diagonal.

Editorial note. A diagonal matrix is both upper triangular and lower triangular, so its determinant is the product of its diagonal entries.

Exam workflow: Choosing a determinant method

Use the quickest valid route.

1. If the matrix is triangular, use the diagonal product immediately.
2. If a row or column has many zeros, use cofactor expansion there.
3. If the matrix is dense but easy to triangularise, use row operations and track their effect on the determinant.
4. If the question explicitly says "use the definition", write the signed elementary products.

2.3 Elementary row operations and determinants

Proposition 2.19 (Effect of elementary row operations on determinants)**.** *For the elementary matrices $D_n(i;\lambda)$, $E_n(i,j;\lambda)$, and $P_n(i,j)$, and every square $n \times n$ matrix A,*

$$\begin{aligned}\det\big(D_n(i;\lambda)A\big) &= \lambda\det(A),\\ \det\big(E_n(i,j;\lambda)A\big) &= \det(A),\\ \det\big(P_n(i,j)A\big) &= -\det(A).\end{aligned}$$

Equivalently,

$$\begin{aligned}\det(A) &= \tfrac{1}{\lambda}\det\big(D_n(i;\lambda)A\big),\\ \det(A) &= \det\big(E_n(i,j;\lambda)A\big),\\ \det(A) &= -\det\big(P_n(i,j)A\big).\end{aligned}$$

Proof required

Proof. For $D_n(i;\lambda)$, left multiplication multiplies row i of A by λ. Every signed elementary product contains exactly one entry from row i, so every signed elementary product is multiplied by λ. Hence the whole determinant is multiplied by λ.

For $P_n(i,j)$, swapping rows i and j changes the sign of every signed elementary product, so the determinant changes sign.

For $E_n(i,j;\lambda)$, let A' be the matrix obtained from A by replacing row i by row $i + \lambda$ row j. Expand $\det(A')$ along row i. The cofactors are the same as for A, because deleting row i removes the only changed row. Thus

$$\det(A') = \det(A) + \lambda\det(B),$$

where B is obtained from A by replacing row i by row j. The matrix B has two equal rows, so $\det(B) = 0$. Therefore $\det(A') = \det(A)$. ☐

Proposition 2.20 (Determinants of elementary matrices)**.** *For the elementary matrices* $D_n(i;\lambda)$*,* $E_n(i,j;\lambda)$*, and* $P_n(i,j)$*,*

$$\det\bigl(D_n(i;\lambda)\bigr) = \lambda, \qquad \det\bigl(E_n(i,j;\lambda)\bigr) = 1, \qquad \det\bigl(P_n(i,j)\bigr) = -1.$$

Proof required

Proof. Apply Proposition 2.19 with $A = I_n$ and use $\det(I_n) = 1$. ☐

Proposition 2.21 (Multiplicativity with an elementary matrix)**.** *For any elementary* $n \times n$ *matrix* E *and any* $n \times n$ *matrix* A*,*

$$\det(EA) = \det(E)\det(A).$$

Proof included here

Proof. Combine Proposition 2.19 with Proposition 2.20 case by case. Each of the three elementary-matrix types satisfies the stated identity. ☐

Remark 2.22 (Determinant computation by row operations)**.** To compute $\det(A)$ efficiently, row reduce A to an upper or lower triangular matrix, keep track of how each row operation changes the determinant, and then use Propositions 2.15 or 2.16.

Theorem 2.23 (Determinant criterion for invertibility)**.** *A real square* $n \times n$ *matrix* A *is invertible if and only if*

$$\det(A) \neq 0.$$

Statement only

Exam tip. This is one of the main conversion rules of the course:

$$\det(A) \neq 0 \iff A \text{ invertible} \iff \mathrm{RREF}(A) = I_n \iff Ax = b \text{ has a unique solution for every } b.$$

Theorem 2.24 (Multiplicativity of the determinant)**.** *For any square matrices* A *and* B *of the same size,*

$$\det(AB) = \det(A)\det(B).$$

Proof required

Proof. The lecture-note proof breaks into four cases.

Case 1&2: A is invertible. Since A is an invertible matrix, we may use Theorem 1.52 in order to deduce that it is possible to express A as a product of elementary matrices, $A = E_1 \cdots E_k$ say. In that case:

$$\begin{aligned}
\det(AB) &= \det(E_1 \cdots E_k B) && \\
&= \det(E_1)\det(E_2 \cdots E_{k-1}E_k B) && \text{using Proposition 2.21} \\
&\ \ \vdots && \quad\vdots \\
&= \det(E_1)\det(E_2)\cdots\det(E_{k-1})\det(E_k)\det(B) && \text{using Proposition 2.21} \\
&= \det(E_1)\det(E_2)\cdots\det(E_{k-1}E_k)\det(B) && \text{using Proposition 2.21} \\
&\ \ \vdots && \quad\vdots \\
&= \det(E_1 \cdots E_k)\det(B) && \text{using Proposition 2.21} \\
&= \det(A)\det(B).
\end{aligned}$$

Case 3: A is not invertible and B is invertible. By Theorem 2.23, $\det(A) = 0$. Because A is not invertible, Theorem 1.52 gives a non-zero vector x^* with $Ax^* = 0$. Since B is invertible, there exists $y^* \neq 0$ with $By^* = x^*$. Then

$$(AB)y^* = A(By^*) = Ax^* = 0,$$

so AB is not invertible. Therefore $\det(AB) = 0 = \det(A)\det(B)$.

Case 4: A and B are both not invertible. Then $\det(B) = 0$ by Theorem 2.23. Again Theorem 1.52 gives a non-zero vector x^* with $Bx^* = 0$. Hence

$$(AB)x^* = A(Bx^*) = A0 = 0,$$

so AB is not invertible and $\det(AB) = 0 = \det(A)\det(B)$. ☐

Exam workflow: Standard determinant consequences

1. Repeated use of Theorem 2.24 gives

   $$\det(A^k) = \det(A)^k.$$

2. If A is invertible, then

   $$1 = \det(I) = \det(AA^{-1}) = \det(A)\det(A^{-1}),$$

   so

   $$\det(A^{-1}) = \frac{1}{\det(A)}.$$

   Notice that the above result is derived from Theorem 2.24, meaning that you MUST include this derivation when quoting this result in the exam.

3. If $B = P^{-1}AP$, then

   $$\det(B) = \det(P^{-1})\det(A)\det(P) = \det(A),$$

   so similar matrices have the same determinant.

3 Chapter 3: Vector spaces

3.1 Real and complex vector spaces

Definition 3.1 (The real vector space $\mathbb{R}^n$)**.** The real vector space $\mathbb{R}^n$ consists of all real column vectors of length n.

Definition 3.2 (The complex vector space $\mathbb{C}^n$)**.** The complex vector space $\mathbb{C}^n$ consists of all complex column vectors of length n.

Definition 3.3 (Addition of vectors)**.** If a, b are vectors in $\mathbb{R}^n$ or $\mathbb{C}^n$, their sum is defined entrywise by

$$(a+b)_i = a_i + b_i \qquad (1 \leq i \leq n).$$

Definition 3.4 (Scalar multiplication of vectors)**.** If $a \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$, or $a \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$, then the scalar multiple λa is defined entrywise by

$$(\lambda a)_i = \lambda a_i.$$

Exam tip. Throughout the course, vectors are columns. In matrix questions, the columns of a matrix are usually the relevant vectors unless the problem explicitly says otherwise.

3.2 Real inner products, norms, and orthogonality

Definition 3.5 (Standard real inner product)**.** For $a, b \in \mathbb{R}^n$, the standard real inner product is

$$\langle a, b\rangle = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + \cdots + a_n b_n.$$

The standard real inner product satisfies the following identities for all $a, b, c \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$:

- $\langle a, b\rangle = \langle b, a\rangle$
- $\langle a+b, c\rangle = \langle a, c\rangle + \langle b, c\rangle$
- $\lambda\langle a, b\rangle = \langle \lambda a, b\rangle = \langle a, \lambda b\rangle$
- $\langle a, a\rangle = a_1^2 + \cdots + a_n^2 \geq 0$

Definition 3.6 ((Associated) Norm in $\mathbb{R}^n$)**.** For $a \in \mathbb{R}^n$, the norm of a is

$$\|a\| = \sqrt{\langle a, a\rangle} = \sqrt{\sum_{i=1}^{n} (a_i)^2}.$$

Definition 3.7 (Orthogonality in $\mathbb{R}^n$)**.** If $a, b \in \mathbb{R}^n$ are non-zero, then a is orthogonal to b if

$$\langle a, b\rangle = 0.$$

Remark 3.8 (Angle between non-zero vectors)**.** If $a, b \in \mathbb{R}^n$ are non-zero, then the angle θ between them is defined by

$$\cos\theta = \frac{\langle a, b\rangle}{\|a\|\,\|b\|}.$$

This makes sense because Theorem 3.9 shows that the right-hand side lies in $[-1, 1]$.

Theorem 3.9 (Cauchy–Schwarz inequality)**.** *Suppose that $a, b \in \mathbb{R}^n$ are non-zero. Then*

$$|\langle a, b\rangle| \leq \|a\| \, \|b\|.$$

Statement only

Theorem 3.10 (Triangle inequality)**.** *For all $a, b \in \mathbb{R}^n$,*

$$\|a + b\| \leq \|a\| + \|b\|.$$

Proof required

Proof. Expand the square of the norm and use Theorem 3.9:

$$\begin{aligned}\|a+b\|^2 &= \langle a+b, a+b\rangle = \langle a, a\rangle + \langle a, b\rangle + \langle b, a\rangle + \langle b, b\rangle \\ &= \|a\|^2 + 2\langle a, b\rangle + \|b\|^2 \leq \|a\|^2 + 2\|a\| \, \|b\| + \|b\|^2 = (\|a\| + \|b\|)^2.\end{aligned}$$

Both sides are non-negative, so taking square roots preserves the inequality. ☐

Theorem 3.11 (Pythagoras for orthogonal vectors)**.** *Suppose that $a, b \in \mathbb{R}^n$ are non-zero and orthogonal. Then*

$$\|a + b\|^2 = \|a\|^2 + \|b\|^2.$$

Proof required

Proof. As above,

$$\|a + b\|^2 = \|a\|^2 + 2\langle a, b\rangle + \|b\|^2.$$

Since a and b are orthogonal, $\langle a, b\rangle = 0$, so the middle term vanishes. ☐

Exam workflow: Norm and angle questions

1. Convert norms into inner products by squaring.
2. Expand using bilinearity and symmetry.
3. Use orthogonality by setting $\langle a, b\rangle = 0$, or bound $\langle a, b\rangle$ with Cauchy–Schwarz.
4. Take square roots only at the end.

3.3 Complex inner products

Definition 3.14 (Complex conjugate)**.** If $a = x + iy$ with $x, y \in \mathbb{R}$, then the complex conjugate of a is

$$\overline{a} = x - iy.$$

Definition 3.15 (Standard complex inner product)**.** For $a, b \in \mathbb{C}^n$, the standard complex inner product is

$$\langle a, b\rangle = a_1\overline{b_1} + \cdots + a_n\overline{b_n}.$$

Definition 3.16 (Norm in $\mathbb{C}^n$)**.** For $a \in \mathbb{C}^n$, the norm of a is

$$\|a\| = \sqrt{\langle a, a\rangle}.$$

Definition 3.17 (Orthogonality in $\mathbb{C}^n$)**.** If $a, b \in \mathbb{C}^n$ are non-zero, then a is orthogonal to b if

$$\langle a, b\rangle = 0.$$

Editorial note. In the complex setting,

$$\langle a, b\rangle = \overline{\langle b, a\rangle}, \qquad \langle \lambda a, b\rangle = \lambda\langle a, b\rangle, \qquad \langle a, \lambda b\rangle = \overline{\lambda}\,\langle a, b\rangle.$$

The conjugate appears in the second slot.

3.4 Subspaces, span, independence, basis, and dimension

Definition 3.21 (Subspace of $\mathbb{R}^n$)**.** A subset $S \subseteq \mathbb{R}^n$ is a subspace of $\mathbb{R}^n$ if:

1. $0 \in S$;
2. Closed under addition: $a, b \in S \Rightarrow a + b \in S$;
3. Closed under scalar multiplication: $a \in S$ and $\lambda \in \mathbb{R} \Rightarrow \lambda a \in S$.

Note. The set consisting only of the relevant zero vector is always a subspace of whichever real vector space $\mathbb{R}^n$ we are considering. It is often referred to as a *trivial subspace.*

Definition 3.22 (Subspace of $\mathbb{C}^n$)**.** A subset $S \subseteq \mathbb{C}^n$ is a subspace of $\mathbb{C}^n$ if the three conditions from Definition 3.21 hold with complex scalars.

Exam tip. The fastest way to show that a set is *not* a subspace is usually to show that $0 \notin S$, or that the set fails closure under addition or scalar multiplication.

Definition 3.28 (Linear combination)**.** A linear combination of vectors $v_1, \dots, v_m \in \mathbb{R}^n$ is an expression of the form

$$\lambda_1 v_1 + \cdots + \lambda_m v_m.$$

Definition 3.29 (Span)**.** A set of vectors $\{v_1, \dots, v_m\}$ in $\mathbb{R}^n$ spans a subspace V of $\mathbb{R}^n$ if

1. each v_i are vectors in V;
2. every vector in V can be written as a linear combination of $v_1, \dots, v_m$.

Equivalently,

$$V = \operatorname{span}\{v_1, \dots, v_m\}.$$

Definition 3.30 (Linear independence and dependence)**.** Vectors $v_1, \dots, v_m \in \mathbb{R}^n$ are linearly independent if for $\lambda_1, \dots, \lambda_m \in \mathbb{R}$

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$$

only have the zero solution ($\lambda_1 = \cdots = \lambda_m = 0$). Otherwise they are linearly dependent.

Definition 3.34 (Basis)**.** Suppose that V is a subspace of $\mathbb{R}^n$. A set $\{v_1, \dots, v_m\} \subseteq V$ is a basis for V if it spans V and vectors $v_1, \dots, v_m$ are linearly independent.

Proposition 3.36 (The span of vectors is a subspace)**.** *Suppose that* $v_1, \dots, v_m \in \mathbb{R}^n$*. Then* $\operatorname{span}\{v_1, \dots, v_m\}$ *is a subspace of* $\mathbb{R}^n$*.* Statement only

Proposition 3.37 (Redundant vectors in a dependent set)**.** *Suppose that* $v_1, \dots, v_m \in \mathbb{R}^n$ *are linearly dependent. Then at least one of them can be written as a linear combination of the others.* Statement only

Theorem 3.38 (Any two finite bases have the same size)**.** *Consider a subspace* V *of a real vector space. Any two finite bases for* V *contain the same number of elements.* Statement only

Definition 3.39 (Dimension)**.** The dimension of a subspace V of a real vector space, written $\dim(V)$, is the number of vectors in any basis of V.

Exam workflow: Span, independence, basis, and dimension by row reduction

Suppose $v_1, \dots, v_m \in \mathbb{R}^n$ and form the matrix

$$A = [v_1 \ \cdots \ v_m].$$

- To test whether $b \in \mathbb{R}^n$ lies in $\operatorname{span}\{v_1, \dots, v_m\}$, solve $A\lambda = b$.
- To test linear independence, solve $A\lambda = 0$.
- To find a basis of a subspace given by equations, solve the system, write the general vector in parametric form, and read off the vectors multiplying the free variables.
- If there are exactly n vectors in $\mathbb{R}^n$, then

$$v_1, \dots, v_n \text{ are independent} \iff \det([v_1 \ \cdots \ v_n]) \neq 0.$$

3.5 Orthonormal sets, Gram–Schmidt, and orthogonal complements

Definition 3.40 (Orthonormal set)**.** A set $\{v_1, \dots, v_m\} \subseteq \mathbb{R}^n$ is orthonormal if

$$\|v_i\| = 1 \quad \text{for all } i, \qquad \langle v_i, v_j \rangle = 0 \quad \text{whenever } i \neq j.$$

Exam workflow: Gram–Schmidt orthonormalisation

Given linearly independent vectors $u_1, \dots, u_m \in \mathbb{R}^n$, define

$$u_1' = u_1, \qquad u_k' = u_k - \sum_{j=1}^{k-1} \frac{\langle u_k, u_j' \rangle}{\langle u_j', u_j' \rangle} u_j' \quad (k \geq 2).$$

Then $u_1', \dots, u_m'$ are orthogonal and span the same subspace as $u_1, \dots, u_m$. Finally normalise:

$$e_k = \frac{u_k'}{\|u_k'\|}.$$

The vectors $e_1, \dots, e_m$ are orthonormal.

Exam tip. In calculations, first produce the orthogonal vectors $u_1', u_2', \dots$ and only then normalise them. If the question asks only for an orthogonal basis, stop before dividing by norms.

Definition 3.43 (Orthogonal complement)**.** Suppose that V is a subspace of $\mathbb{R}^n$. The orthogonal complement of V is

$$V^\perp = \{w \in \mathbb{R}^n : \langle w, v \rangle = 0 \text{ for every } v \in V\}.$$

Proposition 3.44 (The orthogonal complement is a subspace)**.** *Suppose that V is a subspace of $\mathbb{R}^n$. Then $V^\perp$ is a subspace of $\mathbb{R}^n$.* Statement only

Exam workflow: Finding $V^\perp$

Let $w = (w_1, \dots, w_n)^T$ be a general vector and impose

$$\langle w, v \rangle = 0$$

for each basis vector v of V. Solve the resulting homogeneous system and read off a basis of the solution space.

3.6 Polynomial inner products and Fourier series

Editorial note. The lecture notes treat spaces such as

$$\mathbb{R}[x]_2 = \{a_2x^2 + a_1x + a_0 : a_0, a_1, a_2 \in \mathbb{R}\}$$

with inner products defined by integration, for example

$$\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx \qquad \text{or} \qquad \langle f, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)\,dx.$$

All the same ideas continue to hold: orthogonality means the integral is zero, norms are $\|f\| = \sqrt{\langle f, f\rangle}$, and Gram–Schmidt uses the same projection formula with integrals replacing finite sums.

Exam workflow: Fourier coefficients on $[-\pi, \pi)$

For the lecture-note convention

$$f(x) = c + \sum_{n=1}^{\infty} \big(a_n \cos(nx) + b_n \sin(nx)\big),$$

the coefficients are

$$c = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,dx, \qquad a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\,dx, \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,dx.$$

Useful parity shortcuts:

- if f is even, then every $b_n = 0$;
- if f is odd, then $c = 0$ and every $a_n = 0$.

Example 3.45 (Fourier series)**.** For the function

$$f(x) = \begin{cases} 0, & -\pi \le x < -\frac{\pi}{2}, \\ 1, & -\frac{\pi}{2} \le x < \frac{\pi}{2}, \\ 0, & \frac{\pi}{2} \le x < \pi, \end{cases}$$

Example 3.45 in the lecture notes obtains the Fourier expansion

$$f(x) = \frac{1}{2} + \frac{2}{\pi}\left(\cos x - \frac{1}{3}\cos 3x + \frac{1}{5}\cos 5x - \frac{1}{7}\cos 7x + \cdots\right).$$

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{3\pi}{4}$	$\frac{5\pi}{6}$	π
$\sin\theta$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0
$\cos\theta$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0	$-\frac{1}{2}$	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{3}}{2}$	-1

4 Chapter 4: Linear maps

4.1 Linear maps, matrices, kernels, and images

Definition 4.1 (Real linear map)**.** A real linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ is a function satisfying:

1. $T(0) = 0$;
2. $T(a + b) = T(a) + T(b)$ for all $a, b \in \mathbb{R}^n$;
3. $T(\lambda a) = \lambda T(a)$ for all $\lambda \in \mathbb{R}$ and $a \in \mathbb{R}^n$.

Definition 4.6 (Kernel)**.** Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be linear. The kernel of T, written $\mathrm{Ker}(T)$, is the set of vectors in $\mathbb{R}^n$ that T sends to the zero vector in $\mathbb{R}^m$:

$$\mathrm{Ker}(T) = \{x \in \mathbb{R}^n : T(x) = 0\}.$$

Definition 4.7 (Image)**.** Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be linear. The image of T, written $\mathrm{Im}(T)$, is the set of vectors in $\mathbb{R}^m$ that are attained by the map T:

$$\mathrm{Im}(T) = \{y \in \mathbb{R}^m : T(x) = y \text{ for some } x \in \mathbb{R}^n\}.$$

Editorial note. If $T(x) = M_T x$ and the columns of M_T are $c_1, \dots, c_n$, then

$$\mathrm{Im}(T) = \mathrm{span}\{c_1, \dots, c_n\},$$

because every output vector has the form $x_1 c_1 + \cdots + x_n c_n$.

Proposition 4.8 (The kernel is a subspace)**.** *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be linear. Then $\mathrm{Ker}(T)$ is a subspace of $\mathbb{R}^n$.* Proof required

Proof. Because $T(0) = 0$, the zero vector lies in $\mathrm{Ker}(T)$. If $a, b \in \mathrm{Ker}(T)$, then

$$T(a + b) = T(a) + T(b) = 0 + 0 = 0,$$

so $a + b \in \mathrm{Ker}(T)$. If $a \in \mathrm{Ker}(T)$ and $\lambda \in \mathbb{R}$, then

$$T(\lambda a) = \lambda T(a) = \lambda \cdot 0 = 0,$$

so $\lambda a \in \mathrm{Ker}(T)$. Hence $\mathrm{Ker}(T)$ is a subspace. □

Proposition 4.9 (The image is a subspace)**.** *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be linear. Then $\mathrm{Im}(T)$ is a subspace of $\mathbb{R}^m$.* Proof required

Proof. Since $T(0) = 0$, the zero vector lies in $\mathrm{Im}(T)$. If $a, b \in \mathrm{Im}(T)$, choose $x, y \in \mathbb{R}^n$ with $T(x) = a$ and $T(y) = b$. Then

$$T(x + y) = T(x) + T(y) = a + b,$$

so $a + b \in \mathrm{Im}(T)$. If $a \in \mathrm{Im}(T)$ and $T(x) = a$, then

$$T(\lambda x) = \lambda T(x) = \lambda a,$$

so $\lambda a \in \mathrm{Im}(T)$. □

Definition 4.10 (Nullity)**.** For a linear map $T : \mathbb{R}^n \to \mathbb{R}^m$, the nullity of T is

$$\mathrm{nullity}(T) = \dim\bigl(\mathrm{Ker}(T)\bigr).$$

Definition 4.11 (Rank)**.** For a linear map $T : \mathbb{R}^n \to \mathbb{R}^m$, the rank of T is

$$\operatorname{rank}(T) = \dim\big(\operatorname{Im}(T)\big).$$

Theorem 4.13 (Rank–nullity theorem)**.** *If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a real linear map, then*

$$\operatorname{rank}(T) + \operatorname{nullity}(T) = n.$$

Proof included here

Proof. Let M_T be the matrix of T and row reduce it. Suppose the row-reduced matrix has r pivot columns. Then the homogeneous system $M_T x = 0$ has $n - r$ free variables, so

$$\operatorname{nullity}(T) = n - r.$$

Also, the pivot columns of the original matrix form a basis of the column space, which is $\operatorname{Im}(T)$, so

$$\operatorname{rank}(T) = r.$$

Adding gives $\operatorname{rank}(T) + \operatorname{nullity}(T) = r + (n - r) = n$. ☐

Exam workflow: Kernel, image, rank, and nullity from a matrix

1. Solve $M_T x = 0$ to obtain a basis of $\operatorname{Ker}(T)$ and hence $\operatorname{nullity}(T)$.
2. Row reduce M_T.
3. Identify the pivot columns.
4. Use the corresponding columns of the *original* matrix for a basis of $\operatorname{Im}(T)$.
5. Count dimensions and check the rank–nullity formula.

Exam tip. Row operations preserve the solution set of $M_T x = 0$, so they preserve the kernel. They do *not* preserve the image as a subspace of $\mathbb{R}^m$, so image bases must be taken from the pivot columns of the original matrix.

4.2 Eigenvalues, eigenvectors, and characteristic equations

Definition 4.14 (Eigenvalue and eigenvector: real case)**.** Let A be a real square matrix of size $n \times n$. A non-zero vector $v \in \mathbb{R}^n$ is an eigenvector of A if there exists a real number λ such that

$$Av = \lambda v.$$

The number λ is then an eigenvalue of A.

Definition 4.15 (Eigenvalue and eigenvector: complex case)**.** Let A be a complex square matrix of size $n \times n$. A non-zero vector $v \in \mathbb{C}^n$ is an eigenvector of A if there exists a complex number λ such that

$$Av = \lambda v.$$

The complex number λ is then an eigenvalue of A.

Definition 4.16 (Characteristic equation)**.** Let A be a complex square matrix of size $n \times n$. The characteristic equation of A is

$$\det(A - \lambda I) = 0.$$

Proposition 4.17 (Eigenvalue criterion via determinants)**.** *Let A be a complex square matrix of size $n \times n$. A complex number λ is an eigenvalue of A if and only if λ is a solution to the equation*

$$\det(A - \lambda I) = 0.$$

Proof required

Proof. By definition, a complex number λ is an eigenvalue of A if and only if there exists a non-zero vector $v \in \mathbb{C}^n$ such that $Av = \lambda v$. But, then:

$$\begin{aligned} Av = \lambda v \quad &\Leftrightarrow \quad Av - \lambda v = 0 \\ &\Leftrightarrow \quad Av - \lambda I v = 0 \\ &\Leftrightarrow \quad (A - \lambda I)v = 0 \end{aligned}$$

Since $v \neq 0$ and $(A - \lambda I)v = 0$, while also $(A - \lambda I)0 = 0$, the equation $(A - \lambda I)x = 0$ has at least two solutions. Hence it does not have a unique solution. By Theorem 1.52, the matrix $A - \lambda I$ is therefore not invertible.

Finally, we may now use Theorem 2.23 to conclude that the matrix $A - \lambda I$ is not invertible if and only if $\det(A - \lambda I) = 0$, i.e. if and only if $\det(A - \lambda I) = 0$.

Therefore, overall, we deduce that a complex number is an eigenvalue of a complex square matrix A if and only if it is a root of the characteristic equation of A, as required. ☐

Definition 4.19 (Diagonalisable matrix)**.** Let A be a real square matrix of size $n \times n$. The matrix A is diagonalisable if there exists an invertible matrix P such that

$$P^{-1}AP$$

is diagonal.

Definition 4.21 (Algebraic multiplicity)**.** Suppose that λ is an eigenvalue of a real or complex square matrix. Its algebraic multiplicity is its multiplicity as a root of the characteristic equation.

Definition 4.22 (Geometric multiplicity)**.** Suppose that λ is an eigenvalue of a real or complex square matrix. Its geometric multiplicity is the largest number of linearly independent eigenvectors corresponding to λ, equivalently the dimension of the eigenspace corresponding to λ.

Exam workflow: Diagonalisation

To diagonalise a square matrix A:

1. solve $\det(A - \lambda I) = 0$ for the eigenvalues;
2. for each eigenvalue λ, solve $(A - \lambda I)x = 0$;
3. collect enough linearly independent eigenvectors to form an invertible matrix P;
4. put those eigenvectors into the columns of P;
5. place the corresponding eigenvalues on the diagonal of D in the same order.

If you obtain n linearly independent eigenvectors, then $P^{-1}AP = D$.

Exam tip. A matrix A of size $n \times n$ is diagonalisable if and only if any of the following equivalent criteria hold:

- **Eigenvector basis criterion:** A has n linearly independent eigenvectors.

- **Algebraic vs geometric multiplicity criterion:** For every eigenvalue λ,

$$\text{geometric multiplicity of } \lambda = \text{algebraic multiplicity of } \lambda,$$

  assuming the characteristic polynomial splits over the field being used.
- **Distinct eigenvalues sufficient condition:** If A has n distinct eigenvalues, then it is diagonalisable.
- **Repeated eigenvalues warning:** Repeated eigenvalues do **not** automatically mean non-diagonalisable. You must check whether there are enough independent eigenvectors.

4.3 Powers, inverses, and similarity

Proposition 4.24 (Eigenvectors of powers)**.** *Suppose that A is a square matrix. If v is an eigenvector of A corresponding to the eigenvalue λ, then v is an eigenvector of A^n corresponding to the eigenvalue λ^n.* Proof required

Proof. We assume that v is an eigenvector of A, corresponding to an eigenvalue λ, so that $Av = \lambda v$. Consider $A^n v$:

$$\begin{aligned}
A^n v &= A^{n-1}(Av) && \\
&= A^{n-1}(\lambda v) && \text{using } Av = \lambda v \\
&= \lambda(A^{n-1}v) && \text{since } A \text{ corresponds to a linear map} \\
&= \lambda\big(A^{n-2}(Av)\big) && \\
&= \lambda\big(A^{n-2}(\lambda v)\big) && \text{using } Av = \lambda v \\
&= \lambda^2(A^{n-2}v) && \text{since } A \text{ corresponds to a linear map} \\
&\vdots && \\
&= \lambda^n v
\end{aligned}$$

Therefore, we have shown that $A^n v = \lambda^n v$, i.e. that v is an eigenvector of A^n, corresponding to the eigenvalue λ^n, as required. ☐

Proposition 4.25 (Eigenvectors of inverses)**.** *Suppose that A is an invertible square matrix. If v is an eigenvector of A corresponding to the eigenvalue λ, then $\lambda \neq 0$ and v is an eigenvector of A^{-1} corresponding to the eigenvalue λ^{-1}.* Proof required

Proof. If A is invertible, then there is an unique solution to $Ax = 0$, which is $x = 0$ (by Theorem 1.52). Hence, there is NO **non-zero** eigenvector corresponding to the eigenvalue 0 (s.t. $Av = 0v = 0$). So $\lambda \neq 0$. Multiplying $Av = \lambda v$ by A^{-1} gives

$$\begin{aligned}
Av = \lambda v &\iff A^{-1}(Av) = A^{-1}(\lambda v) \\
&\iff (A^{-1}A)v = \lambda(A^{-1}v) \\
&\iff v = \lambda(A^{-1}v) \\
&\iff A^{-1}v = \tfrac{1}{\lambda}v.
\end{aligned}$$

Hence, v is an eigenvector of A^{-1} corresponding to the eigenvalue λ^{-1}, as required. ☐

Proposition 4.26 (Powers from a diagonalisation)**.** *Suppose that A and D are square matrices of the same size and that there exists an invertible matrix P such that*

$$D = P^{-1}AP.$$

Then, for every positive integer n,

$$A^n = PD^nP^{-1}.$$

Proof required

Proof. We assume that $P^{-1}AP = D$, and may "cancel out" P and P^{-1} from this equation in order to obtain an expression for A:

$$\begin{aligned} P^{-1}AP = D \quad &\Leftrightarrow \quad P(P^{-1}AP) = PD \\ &\Leftrightarrow \quad AP = PD \\ &\Leftrightarrow \quad (AP)P^{-1} = (PD)P^{-1} \\ &\Leftrightarrow \quad A = PDP^{-1} \end{aligned}$$

Let us now substitute this expression for A into A^n and simplify:

$$\begin{aligned} A^n &= (PDP^{-1})^n && \text{by substituting } A = PDP^{-1} \\ &= (PDP^{-1})(PDP^{-1})\cdots(PDP^{-1}) \\ &= PD(P^{-1}P)D(P^{-1}P)D\cdots D(P^{-1}P)DP^{-1} \\ &= PDIDID\cdots DIDP^{-1} && \text{by using } P^{-1}P = I \\ &= PD^nP^{-1} \end{aligned}$$

So, we have shown that, if $P^{-1}AP = D$, then $A^n = PD^nP^{-1}$, as required. □

Exam workflow: Computing powers of a diagonalisable matrix

If $P^{-1}AP = D$, then

$$A^n = PD^nP^{-1}, \qquad D^n = \operatorname{diag}(\lambda_1^n, \dots, \lambda_n^n).$$

So once a diagonalisation is known, powers of A reduce to powers of diagonal entries.

4.4 Hermitian and symmetric matrices

Definition 4.27 (Conjugate matrix)**.** If $A = (A_{ij})$ is a complex matrix, the conjugate of A, denoted by $\overline{A}$, is defined entrywise by

$$(\overline{A})_{ij} = \overline{A_{ij}}.$$

Definition 4.28 (Hermitian matrix)**.** A complex square matrix A is Hermitian if

$$\overline{A}^T = A.$$

Definition 4.29 (Symmetric matrix)**.** A real square matrix A is symmetric if

$$A^T = A.$$

Equivalently, $A_{ij} = A_{ji}$ for all i, j.

Theorem 4.30 (Spectral theorem for Hermitian matrices)**.** *Suppose that M is a Hermitian matrix of size $n \times n$. Then:*

- *M is diagonalisable;*
- *every eigenvalue of M is real;*
- *eigenvectors corresponding to distinct eigenvalues are orthogonal with respect to the standard complex inner product;*
- *there exists an orthonormal set of n eigenvectors for M.*

Statement only

Exam workflow: Hermitian and symmetric matrix questions

1. Compute the eigenvalues from $\det(A - \lambda I) = 0$.
2. Use Theorem 4.30 to justify that, in the Hermitian case, all eigenvalues are real and eigenvectors from distinct eigenvalues are orthogonal.
3. Normalise the eigenvectors if an orthonormal basis is required.
4. Build P from eigenvectors and D from eigenvalues.
5. Use diagonalisation for powers if needed.

For real symmetric matrices, the same conclusions hold with real eigenvectors.