MH5200 Advanced Investigations in Linear Algebra I – Revision Notes

Quantitative Research Society @ NTU

Academic Year 2025–2026

Abstract

Comprehensive revision notes for MH5200 Advanced Investigations in Linear Algebra I, integrating theoretical concepts, key formulas, worked examples from problem sheets, and exam-focused applications. The emphasis is on building intuition for each result, understanding how the pieces fit together, and recognising patterns that frequently appear in exam questions and in later modules (e.g. numerical analysis, optimisation, PDEs and graph-based learning).

Contents

How to Use This Document 3

I Foundational Topics 5

1 Linear Systems and Matrix Operations 5
- 1.1 Tridiagonal Systems . . . 5
- 1.2 Matrix Products and Transpose . . . 6
- 1.3 Symmetric and Skew-Symmetric Matrices . . . 7
- 1.4 Orthogonal Matrices . . . 8
- 1.5 Rank-One Updates and Quadratic Forms . . . 8

2 Linear Independence and Special Matrices 9
- 2.1 Linear Dependence via Determinant . . . 9
- 2.2 Orthonormal Expansions . . . 10
- 2.3 Skew-Symmetric Characterisation . . . 11

3 Trace and Matrix Interpretations **11**
3.1 Trace Identities 11
3.2 Matrix Interpretation of Interpolation Systems 12

4 LU Factorisations and Block Inverses **13**
4.1 Lower Triangular Elimination Matrices 13
4.2 Schur Complement Example 13

II Determinants, Eigenvalues and Similarity **15**

5 Determinants: Theory and Techniques **15**
5.1 Basic Properties 15
5.2 Tridiagonal Determinant Recurrences 15

6 Eigenvalues and Eigenvectors **16**
6.1 Definitions 16
6.2 Algebraic vs Geometric Multiplicity 16
6.3 Rank-One Matrices 17

7 Similarity and Diagonalisation **17**
7.1 Similarity 17
7.2 Orthogonal Diagonalisation 18

III Graphs, Power Method, and Cross-Topic Tools **19**

8 Graph Matrices and Laplacians **19**
8.1 Adjacency, Degree and Laplacian 19
8.2 Path Counting 20

9 Power Method for Dominant Eigenvalues **20**

IV Global Tools and Exam Strategy **22**

10 Invertible Matrix Theorem **22**

11 Rank-Nullity Theorem **23**

12 Cayley–Hamilton and Cramer's Rule **23**

A Formula Sheet **25**

How to Use This Document

Structure of the Notes

These notes are designed to be used together with your lectures and tutorials. They are *not* a replacement for the official notes, but a compressed and slightly more opinionated view of the course. Each chapter is organised into:

1. **Core Theory** — formal definitions, results and concise intuition. Here you should make sure you understand every symbol.
2. **Mathematical Framework** — formulas, algebra and derivations that you are expected to reproduce or adapt in exam-style questions.
3. **Worked Examples** — tutorial-style questions with full solutions. These are chosen to be representative of typical exam tricks rather than exhaustive.
4. **Applications** — interpretations or "stories" behind the linear algebra. Many of these connect to numerical methods and data science.
5. **Exam Focus** — high-yield concepts, typical traps and checklist items, often summarised in 1–2 screens worth of material per topic.

A suggested revision flow:

- **First pass (overview)**: skim the section headings, read Core Theory and Key Formula boxes, and draw a mini mind-map of how the concepts link.
- **Second pass (active)**: attempt Worked Examples yourself *before* checking solutions. Time yourself and write full solutions, not just answers.
- **Third pass (consolidation)**: redo 1–2 examples without looking, and summarise each topic in your own words (no formulas) in ≤ 5 sentences.
- **Final pass (exam week)**: review Exam Focus and Common Mistake boxes in the days before the exam. Use the appendix as a quick formula sheet.

Exam Focus

A good mental target is: for each topic you should be able to (i) state the main theorem, (ii) know at least one standard proof idea, (iii) solve a small example *without* notes, and (iv) recognise when the topic is being used inside a more complicated exam question.

Colour and Box System (Legend)

The document uses a colour-coded box system. Each box has a **name** at the top so you can quickly recognise its purpose:

Colour	Box Name	Use This For
Blue	**Key Formula**	Core equations and identities that you should memorise or be able to derive quickly.
Brown	**Worked Example**	Fully worked questions. Try on your own before reading the solution.
Red	**Common Mistake**	Typical errors, misleading shortcuts and conceptual traps to avoid in exams.
Purple	**Exam Focus**	High-yield checklists, "must-know" results and exam strategy tips.
Yellow	**Policy Application**	Interpretations or context where relevant.

When you print in black and white, rely mainly on the **title text** at the top of each box (Key Formula, Worked Example, etc.).

Common Mistake

Many students read worked examples passively and feel they "understand". In linear algebra this is dangerous: the same concept can appear with completely different notation. When you see an example, ask yourself: *If I changed the letters and the dimension, do I still know what to do?*

Part I
Foundational Topics

1 Linear Systems and Matrix Operations

This section covers basic matrix operations, tridiagonal systems, symmetric and skew-symmetric matrices, orthogonal matrices, and rank-one updates — material that appears heavily in Sheet 1 and Sheet 4.

A recurring theme is to interpret matrices in *three ways*:

- as **arrays of numbers** on which you perform algebraic manipulations;
- as **linear maps** acting on vectors ($x \mapsto Ax$);
- as **collections of geometric objects** (rows, columns, eigenvectors, etc.).

Switching viewpoint is often the key to solving exam questions efficiently.

1.1 Tridiagonal Systems

Key Formula

A **tridiagonal system** has coefficient matrix T_n of the form

$$T_n = \begin{bmatrix} d_1 & e_1 & 0 & \cdots & 0 \\ f_1 & d_2 & e_2 & \ddots & \vdots \\ 0 & f_2 & d_3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & e_{n-1} \\ 0 & \cdots & 0 & f_{n-1} & d_n \end{bmatrix}.$$

Such systems can be solved in $O(n)$ time using the Thomas algorithm, a specialised Gaussian elimination.

Intuitively, tridiagonal matrices arise whenever each unknown interacts only with its "neighbours" (e.g. finite-difference schemes for 1D PDEs, Markov chains on path graphs). The Thomas algorithm avoids fill-in: you maintain tridiagonal structure instead of creating a dense matrix.

Thomas Algorithm Sketch

For $T_n x = b$ with T_n as above:

1. **Forward sweep:** eliminate f_i by updating

$$\tilde{d}_i = d_i - f_{i-1}\frac{e_{i-1}}{\tilde{d}_{i-1}}, \qquad \tilde{b}_i = b_i - f_{i-1}\frac{\tilde{b}_{i-1}}{\tilde{d}_{i-1}}.$$

   This is just Gaussian elimination specialised to the banded structure.

2. **Backward substitution:** solve

$$x_n = \frac{\tilde{b}_n}{\tilde{d}_n}, \qquad x_i = \frac{\tilde{b}_i - e_i x_{i+1}}{\tilde{d}_i},\ i = n-1, \ldots, 1.$$

Total cost is $\Theta(n)$, much cheaper than the $\Theta(n^3)$ of generic Gaussian elimination.

Worked Example: Boundary Value Tridiagonal System

Consider the finite-difference discretisation of $-u''(x) = f(x)$ on $[0,1]$ with $u(0) = u(1) = 0$. With step $h = 1/5$ and nodes $x_1, \ldots, x_4$, we obtain

$$-u_{i+1} + 2u_i - u_{i-1} = h^2 f(x_i), \quad i = 1, \ldots, 4,$$

with $u_0 = u_5 = 0$. In matrix form $T_4\mathbf{u} = \mathbf{b}$ where

$$T_4 = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}, \qquad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix}.$$

We can solve this by:

1. applying the Thomas algorithm (fast and structured), or
2. performing standard Gaussian elimination (slower but conceptually similar).

The important point for exams is to (i) recognise the tridiagonal structure, and (ii) know that such systems are well-behaved and solvable in linear time in n.

Exam Focus

When you see a second-order finite-difference equation on an interval (discrete Laplacian), *immediately* think “tridiagonal system”. This is a standard pattern and often the exam question is really checking whether you can set up the matrix correctly.

1.2 Matrix Products and Transpose

Key Formula

For conformable matrices A, B:

$$(AB)^T = B^T A^T,$$
$$(A^T)^T = A,$$
$$(AB)^{-1} = B^{-1}A^{-1} \quad \text{(if both are invertible).}$$

Geometric interpretation:

- AB means “apply B first, then A” as linear maps.
- The transpose A^T is the adjoint with respect to the standard inner product:

$$(Ax)^T y = x^T (A^T y).$$

- The formula $(AB)^{-1} = B^{-1}A^{-1}$ just reverses the order of operations.

Worked Example: Coordinates of BAx

Let A be $m \times n$, B be $p \times m$ and $\mathbf{x} \in \mathbb{R}^n$. Then $BA\mathbf{x} \in \mathbb{R}^p$ and its i-th component is

$$(BA\mathbf{x})_i = \sum_{j=1}^{m} b_{ij} \left(\sum_{k=1}^{n} a_{jk} x_k \right) = \sum_{k=1}^{n} \left(\sum_{j=1}^{m} b_{ij} a_{jk} \right) x_k.$$

Algebraically, this shows how the entries of BA are built from those of A and B. Conceptually, it expresses BA as a linear combination of the columns of A with coefficients determined by B.

Common Mistake

When you see ABx, do *not* first multiply B and x if the dimensions do not match. Always check dimensions: if A is $m \times n$ and B is $n \times p$, then AB is $m \times p$, so x must be in $\mathbb{R}^p$ to form ABx.

1.3 Symmetric and Skew-Symmetric Matrices

Definition 1.1: Symmetric and Skew-Symmetric Matrices

A matrix S is **symmetric** if $S^T = S$. A matrix A is **skew-symmetric** if $A^T = -A$.

Symmetric matrices behave like “nice” quadratic forms; skew-symmetric matrices behave like “infinitesimal rotations” in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Key Formula

Every matrix M can be uniquely decomposed as

$$M = \frac{1}{2}(M + M^T) + \frac{1}{2}(M - M^T) = S + A,$$

where S is symmetric and A is skew-symmetric.

This is analogous to decomposing a complex number into real and imaginary parts. Many proofs with quadratic forms and energy functionals use this decomposition to separate “even” and “odd” parts.

Example 1.1. *All 2×2 skew-symmetric matrices have the form*

$$A = \omega \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \omega \in \mathbb{R}.$$

The corresponding linear map is a rotation by 90° scaled by ω. In higher dimensions, skew-symmetric matrices are closely related to cross products and angular velocities.

Exam Focus

In inner-product expressions like $x^T Ax$, the skew-symmetric part of A does not contribute:

$$x^T Ax = x^T \left(\frac{A + A^T}{2} \right) x.$$

So for quadratic forms you may safely replace A by its symmetric part. This is very useful for simplifying proofs.

1.4 Orthogonal Matrices

Definition 1.2: Orthogonal Matrix

A square matrix Q is **orthogonal** if $Q^TQ = I$, i.e. $Q^{-1} = Q^T$.

Key Formula

Properties of orthogonal matrices:

- Columns (and rows) form an orthonormal set.
- Q preserves Euclidean norm: $\|Q\mathbf{x}\| = \|\mathbf{x}\|$.
- Q preserves inner products and angles.
- $\det(Q) = \pm 1$; $\det(Q) = 1$ usually corresponds to a "pure rotation", $\det(Q) = -1$ to a rotation+reflection.

Numerically, orthogonal matrices are "perfectly conditioned": they do not magnify errors. This is why QR factorisation and orthogonal similarity transforms are central in numerical linear algebra.

Orthogonal Matrices in $\mathbb{R}^2$

In $\mathbb{R}^2$, every orthogonal matrix has one of the forms

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix},$$

corresponding respectively to a rotation and a reflection across a line through the origin.

Common Mistake

Do not confuse "orthogonal matrix" with "matrix with orthogonal columns but not unit length". For orthogonality, you need *orthonormal* columns. If the columns are merely orthogonal, you get Q^TQ diagonal but not I.

1.5 Rank-One Updates and Quadratic Forms

Proposition 1.1. *If the i-th row of A is a_i^T, then*

$$A^TA = \sum_i a_i a_i^T.$$

More generally, for a diagonal $C = \operatorname{diag}(c_1, \dots, c_m)$,

$$A^TCA = \sum_{i=1}^m c_i a_i a_i^T.$$

Each outer product $a_i a_i^T$ is rank-one. Thus $A^T A$ is a sum of rank-one positive semi-definite matrices. This "sum-of-squares" viewpoint explains why $A^T A$ is always symmetric positive semi-definite.

Worked Example: Quadratic Form as Sum of Squares

If A has rows $a_1^T, \dots, a_m^T$, then

$$\mathbf{x}^T A^T A\mathbf{x} = (A\mathbf{x})^T (A\mathbf{x}) = \sum_{i=1}^{m} (a_i^T \mathbf{x})^2 \geq 0.$$

Thus $A^T A$ is always symmetric positive semi-definite. If the rows of A are linearly independent, then $A^T A$ is positive definite and hence invertible.

Exam Focus

Whenever you see an expression like $x^T A^T Ax$, immediately rewrite it as $\|Ax\|^2$. This makes it obvious that the quantity is non-negative and often gives a direct geometric interpretation (distance, energy, variance, etc.).

2 Linear Independence and Special Matrices

This section corresponds closely to Problem Sheet 2, including linear dependence via determinants, orthonormal expansions, skew-symmetric characterisations and magic matrices.

2.1 Linear Dependence via Determinant

Technique

To test whether vectors $v_1, \dots, v_n \in \mathbb{R}^n$ are linearly independent, form the matrix $V = [v_1 \ \cdots \ v_n]$ and compute $\det(V)$. If $\det(V) = 0$, they are dependent; otherwise they are independent.
In higher dimensions or parametric situations, it is often easier to:

- compute $\det(V)$ symbolically in terms of parameters;
- find the parameter values where the determinant vanishes.

Worked Example: Parametric Linear Dependence

Consider the vectors

$$v_1 = (\mu, 1, 0)^T, \quad v_2 = (1, \mu, 1)^T, \quad v_3 = (0, 1, \mu)^T.$$

They are dependent when

$$\det \begin{bmatrix} \mu & 1 & 0 \\ 1 & \mu & 1 \\ 0 & 1 & \mu \end{bmatrix} = 0.$$

Computing the determinant:

$$\mu(\mu^2 - 1) - 1(\mu - 0) = \mu^3 - \mu - \mu = \mu^3 - 2\mu = \mu(\mu^2 - 2).$$

Thus dependence occurs for $\mu = 0, \pm\sqrt{2}$.

Interpretation: for all other values of μ, the three vectors form a basis of $\mathbb{R}^3$, while at $\mu = 0, \pm\sqrt{2}$ the dimension of the span drops to 2.

Exam Focus

Parametric determinant questions often appear in the form "for which values of μ does A fail to be invertible?". Mechanically, this is just "solve $\det(A(\mu)) = 0$", but conceptually you should link this to changes in rank, dimension of eigenspaces, and geometry of the column space.

2.2 Orthonormal Expansions

Key Formula

Let $a_1, \dots, a_k$ be orthonormal vectors in $\mathbb{R}^n$ and suppose

$$x = \beta_1 a_1 + \cdots + \beta_k a_k.$$

Then

$$\|x\|^2 = \sum_{i=1}^{k} \beta_i^2 = \|\beta\|^2.$$

This is just the Pythagorean theorem in disguise. In an orthonormal basis, inner products and norms behave exactly like in $\mathbb{R}^k$ with the standard coordinates.

Worked Example: Coefficients from Inner Products

If $\{a_1, \dots, a_k\}$ is orthonormal and $x = \sum \beta_i a_i$, then

$$\beta_i = a_i^T x.$$

Thus the coordinate vector $(\beta_1, \dots, \beta_k)^T$ is obtained simply by taking inner products with each a_i. In matrix form, if $Q = [a_1 \ \dots \ a_k]$, then

$$\beta = Q^T x, \quad x = Q\beta, \quad Q^T Q = I_k.$$

Exam Focus

Whenever you have an orthonormal basis (e.g. eigenvectors of a symmetric matrix), *never* solve for coefficients by inverting a matrix. Just take inner products. This is faster, numerically more stable, and conceptually clearer.

2.3 Skew-Symmetric Characterisation

Key Formula

For a square matrix A, the following are equivalent:

1. $A^T = -A$ (skew-symmetric),
2. $x^T Ax = 0$ for all $x \in \mathbb{R}^n$,
3. All diagonal entries of A are zero.

Sketch of (1) ⇒ (2): using $A^T = -A$,

$$x^T Ax = (x^T Ax)^T = x^T A^T x = x^T(-A)x = -x^T Ax$$

so $x^T Ax = 0$. Conversely, one can use polarisation identities to recover A from the bilinear form $(x, y) \mapsto x^T Ay$ and show skew-symmetry.

Common Mistake

The implication "all diagonals zero ⇒ A skew-symmetric" is *false*. Zero diagonal is necessary but not sufficient. For skew-symmetry you must also have $a_{ij} = -a_{ji}$ for all $i \neq j$.

3 Trace and Matrix Interpretations

3.1 Trace Identities

Key Formula

For matrices of compatible sizes:

$$\begin{aligned} \mathrm{tr}(A^T B) &= \sum_{i,j} a_{ij} b_{ij} \quad \text{(Frobenius inner product)}, \\ \mathrm{tr}(AB) &= \mathrm{tr}(BA), \\ \mathrm{tr}(A + B) &= \mathrm{tr}(A) + \mathrm{tr}(B). \end{aligned}$$

Moreover, $\mathrm{tr}(A^T A) \geq 0$ with equality iff $A = 0$.

Interpretation:

- $\mathrm{tr}(A^T B)$ is the standard inner product on matrices; $\|A\|_F^2 = \mathrm{tr}(A^T A)$.
- The identity $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ allows cyclic shifts inside traces, which is critical in many derivations (e.g. differentiating trace expressions).
- For a square matrix A, $\mathrm{tr}(A)$ also equals the sum of eigenvalues (with multiplicity).

Worked Example: Trace as Sum of Diagonal Blocks

Let
$$A = \begin{bmatrix} B & * \\ 0 & C \end{bmatrix}$$
be block upper triangular. Then
$$\operatorname{tr}(A) = \operatorname{tr}(B) + \operatorname{tr}(C).$$
This follows both from the definition (sum of diagonal entries) and from the fact that eigenvalues of A are the union of those of B and C.

Common Mistake

It is *not* true in general that $\operatorname{tr}(AB) = \operatorname{tr}(A)\operatorname{tr}(B)$. This only holds in very special situations (e.g. when one of the matrices is a scalar multiple of the identity).

3.2 Matrix Interpretation of Interpolation Systems

Polynomial Interpolation as $Ac = b$

Let $p(x) = c_1 + c_2x + c_3x^2 + c_4x^3 + c_5x^4$ satisfy conditions such as
$$p(0) = 0, \quad p'(0) = 0, \quad p(1) = 1, \quad p'(1) = 0.$$
Each condition is a linear equation in $c_1, \dots, c_5$, so we can write $Ac = b$ for a suitable 4×5 matrix A. For example,
$$p(0) = 0 \Rightarrow c_1 = 0, \quad p'(0) = 0 \Rightarrow c_2 = 0, \quad p(1) = 1 \Rightarrow c_1 + c_2 + c_3 + c_4 + c_5 = 1,$$
$$p'(1) = 0 \Rightarrow c_2 + 2c_3 + 3c_4 + 4c_5 = 0.$$
Collecting these into matrix form gives an underdetermined system with infinitely many solutions. Conceptually, polynomial interpolation is just solving a linear system in the coefficients; the classical Vandermonde matrix is a special case where the system is square and typically invertible.

Exam Focus

Whenever you see "find a polynomial/curve/function with these interpolation and derivative conditions", first think "unknown coefficients + linear system". The exact functional form is secondary; the primary task is setting up $Ac = b$ correctly.

4 LU Factorisations and Block Inverses

4.1 Lower Triangular Elimination Matrices

Technique

If L is strictly lower triangular with $L^n = 0$ (nilpotent), then

$$(I - L)^{-1} = I + L + L^2 + \cdots + L^{n-1}.$$

This arises naturally when collecting row operations into a single matrix.

Idea: the Neumann series

$$(I - L)(I + L + L^2 + \cdots + L^{n-1}) = I - L^n = I$$

because all higher powers vanish. This is identical to the geometric series formula for real numbers, but now in matrix form.

Worked Example: Special Lower Triangular Matrix

Let

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -a & 1 & 0 & 0 \\ 0 & -b & 1 & 0 \\ 0 & 0 & -c & 1 \end{bmatrix}.$$

Then

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ ab & b & 1 & 0 \\ abc & bc & c & 1 \end{bmatrix}.$$

You can verify this either by direct multiplication or by writing $A = I - L$ and using the finite Neumann series; L has non-zero entries only on the first subdiagonal, so $L^4 = 0$.

Exam Focus

Special patterns (like constant subdiagonals) often allow you to write the inverse in closed form. Examiners like to test whether you can recognise such patterns and avoid heavy computation.

4.2 Schur Complement Example

Key Formula

For

$$A = \begin{bmatrix} I_n & a \\ a^T & 0 \end{bmatrix}, \quad a \in \mathbb{R}^n,$$

we have A invertible iff $a \neq 0$, and

$$A^{-1} = \begin{bmatrix} I_n - \dfrac{aa^T}{a^T a} & \dfrac{a}{a^T a} \\ \dfrac{a^T}{a^T a} & -\dfrac{1}{a^T a} \end{bmatrix}.$$

Derivation (block inverse via Schur complement):

- Take the partition $A = \begin{bmatrix} I_n & a \\ a^T & 0 \end{bmatrix}$ with upper-left block I_n invertible.
- The Schur complement of I_n is
$$S = 0 - a^T I_n^{-1} a = -a^T a.$$
- The block inverse formula then yields

$$A^{-1} = \begin{bmatrix} I_n + I_n a S^{-1} a^T I_n & -I_n a S^{-1} \\ -S^{-1} a^T I_n & S^{-1} \end{bmatrix} = \begin{bmatrix} I_n - \dfrac{aa^T}{a^T a} & \dfrac{a}{a^T a} \\ \dfrac{a^T}{a^T a} & -\dfrac{1}{a^T a} \end{bmatrix}.$$

Interpretation: Projection onto Orthogonal Complement

The block $I_n - \dfrac{aa^T}{a^T a}$ is the orthogonal projector onto the subspace orthogonal to a. Thus the inverse of A combines a projection onto $a^\perp$ with a one-dimensional correction in the a direction. This viewpoint is useful in optimisation (Lagrange multipliers, constrained least squares) and in numerical linear algebra.

Part II
Determinants, Eigenvalues and Similarity

5 Determinants: Theory and Techniques

5.1 Basic Properties

Key Formula

For $n \times n$ matrices A, B and scalar k:

$$\begin{aligned}
\det(AB) &= \det(A)\det(B),\\
\det(A^T) &= \det(A),\\
\det(kA) &= k^n \det(A),\\
\det(A^{-1}) &= \frac{1}{\det(A)} \quad \text{(if } A \text{ invertible).}
\end{aligned}$$

Conceptually, $\det(A)$ measures how A scales volumes in $\mathbb{R}^n$. Negative determinant corresponds to an orientation-reversing transformation.

Effect of Row Operations

- Swap two rows $\Rightarrow$ determinant changes sign.
- Multiply a row by $k \Rightarrow$ determinant multiplied by k.
- Add multiple of one row to another $\Rightarrow$ determinant unchanged.

These rules are the basis of determinant computation by Gaussian elimination.

5.2 Tridiagonal Determinant Recurrences

Worked Example: Determinant Recurrence

For the $n \times n$ matrix

$$E_n = \begin{vmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \ddots & \vdots \\ 0 & 1 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 & 1 \end{vmatrix},$$

expansion along the first row yields a recurrence

$$E_n = E_{n-1} - E_{n-2}.$$

With initial conditions $E_1 = 1$, $E_2 = 0$, we can compute further values and observe a period-6 pattern:

$$E_1 = 1,\ E_2 = 0,\ E_3 = -1,\ E_4 = -1,\ E_5 = 0,\ E_6 = 1,\ E_7 = 1, \dots$$

This kind of recurrence appears frequently when dealing with Toeplitz or banded matrices arising from discretisations of differential equations.

Exam Focus

For structured matrices (tridiagonal, Toeplitz, banded), *never* expand determinants fully. Instead, look for recurrences by expanding along a row/column that introduces smaller matrices of the same type.

6 Eigenvalues and Eigenvectors

6.1 Definitions

Definition 6.1: Eigenvalues and Eigenvectors

Let A be an $n \times n$ matrix. A non-zero vector v is an **eigenvector** of A with eigenvalue λ if

$$Av = \lambda v.$$

The eigenvalues λ are the roots of the characteristic polynomial

$$p_A(t) = \det(tI - A).$$

Geometric intuition:

- Av is a stretched (and possibly flipped) version of v along the same line.
- Eigenvalues describe the "preferred directions" of the linear map where only scaling occurs.
- In dynamical systems, eigenvalues determine stability: $|\lambda| < 1$ corresponds to contraction, $|\lambda| > 1$ to expansion.

Key Formula

If A has eigenvalues $\lambda_1, \ldots, \lambda_n$ (counted with algebraic multiplicity), then

$$\operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i, \qquad \det(A) = \prod_{i=1}^{n} \lambda_i.$$

6.2 Algebraic vs Geometric Multiplicity

Definition 6.2: Multiplicities of Eigenvalues

For an eigenvalue λ of A:

- The **algebraic multiplicity** is its multiplicity as a root of $p_A(t)$.
- The **geometric multiplicity** is $\dim\ker(A - \lambda I)$.

Always: geometric multiplicity $\leq$ algebraic multiplicity.

A matrix is diagonalizable iff for each eigenvalue, geometric multiplicity equals algebraic multiplicity and the total number of linearly independent eigenvectors is n.

Common Mistake

Equal eigenvalues do *not* automatically imply non-diagonalizability. The problem arises only when the eigenspaces are too small (geometric multiplicity strictly less than algebraic multiplicity).

6.3 Rank-One Matrices

Eigenvalues of Rank-One Matrix

Let $A = uv^T$ with $u, v \in \mathbb{R}^n$, $u, v \neq 0$. Then

$$Av = u(v^T v) = (v^T u)\, v,$$

so v is an eigenvector with eigenvalue $v^T u$. All other eigenvalues are 0, so A has one possibly non-zero eigenvalue and rank at most 1.
In fact, A acts by "projecting" onto the direction of u and then rescaling. This structure is heavily used in rank-one updates (Sherman–Morrison formula) and in iterative methods.

7 Similarity and Diagonalisation

7.1 Similarity

Definition 7.1: Similarity

Matrices $A, B \in \mathbb{R}^{n\times n}$ are **similar** if there exists an invertible P such that

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

Key Formula

If A and B are similar, then they have the same characteristic polynomial, thus the same eigenvalues (with multiplicities). They also have the same trace and determinant.

Interpretation: similarity means "same linear map, but in a different basis". Many questions ask you to show two matrices are similar by constructing P whose columns are basis vectors relating the two coordinate systems.

Common Mistake

Having the same eigenvalues does *not* guarantee that two matrices are similar. The Jordan structure (or geometric multiplicities) can differ. Counterexample: the identity matrix and a non-trivial Jordan block both have eigenvalue 1 only, but are not similar.

7.2 Orthogonal Diagonalisation

Theorem 7.1: Spectral Theorem (Real Symmetric Case)

If A is a real symmetric $n \times n$ matrix, then there exists an orthogonal matrix Q and a real diagonal matrix Λ such that

$$A = Q\Lambda Q^T.$$

The diagonal entries of Λ are the eigenvalues of A and the columns of Q are orthonormal eigenvectors.

This theorem is one of the central pillars of linear algebra. It underlies:

- principal component analysis (PCA) in data science;
- quadratic forms and optimisation (diagonalising energy functions);
- numerical algorithms (QR algorithm, power method).

Diagonalising a Symmetric 2×2 Matrix

For

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix},$$

we find eigenvalues via $\det(\lambda I - A) = 0$:

$$\det \begin{bmatrix} \lambda - 2 & -1 \\ -1 & \lambda - 2 \end{bmatrix} = (\lambda - 2)^2 - 1 = 0 \Rightarrow \lambda_1 = 1,\ \lambda_2 = 3.$$

Eigenvectors are easily computed and orthogonalised, then normalised to give an orthogonal matrix Q, with $Q^T AQ = \operatorname{diag}(1, 3)$. In this basis, A simply scales the coordinate axes by 1 and 3 respectively.

Exam Focus

For symmetric matrices, always look for an orthonormal eigenbasis and think in the "diagonalised basis". Many inequalities and optimisation problems become trivial once you express everything in terms of eigenvalues.

Part III
Graphs, Power Method, and Cross-Topic Tools

8 Graph Matrices and Laplacians

8.1 Adjacency, Degree and Laplacian

Definition 8.1: Adjacency Matrix

For an undirected simple graph $G = (V, E)$ with $|V| = n$, the **adjacency matrix** $A \in \mathbb{R}^{n\times n}$ is defined by

$$A_{ij} = \begin{cases} 1, & \{v_i, v_j\} \in E, \\ 0, & \text{otherwise.} \end{cases}$$

Definition 8.2: Degree Matrix

The **degree matrix** D is diagonal with $D_{ii} = d(v_i)$ equal to the degree of vertex v_i.

Definition 8.3: Graph Laplacian

The **(combinatorial) Laplacian** of G is

$$L = D - A.$$

Key Formula

If B is an incidence matrix for an oriented version of G, then

$$L = BB^T.$$

Thus L is symmetric and positive semi-definite. Also $L\mathbf{1} = 0$, so 0 is always an eigenvalue of L.

Interpretation of L:

- $(Lx)_i = d(v_i)x_i - \sum_{j\sim i} x_j$ measures how "different" x_i is from its neighbours.
- The quadratic form
$$x^T Lx = \frac{1}{2} \sum_{\{i,j\}\in E} (x_i - x_j)^2$$
rewards smooth vectors on the graph (neighbouring entries close to each other).
- This is a discrete analogue of the continuous Laplacian in PDEs.

Connected Components and Eigenvalues of L

The multiplicity of the eigenvalue 0 of L equals the number of connected components of the graph. For a connected graph, 0 is simple and the corresponding eigenvector is the all-ones vector $\mathbf{1}$.

8.2 Path Counting

Key Formula

For an undirected graph with adjacency matrix A, the (i, j) entry of A^N equals the number of distinct walks of length N from v_i to v_j.

This result follows by induction using matrix multiplication and provides a direct combinatorial meaning to powers of A.

Example: Walks of Length 2

If A is the adjacency matrix of a graph, then

$$(A^2)_{ij} = \sum_k A_{ik} A_{kj}$$

counts all vertices k which are neighbours of both i and j. Thus $(A^2)_{ij}$ equals the number of length-2 walks from i to j, and in particular $(A^2)_{ii}$ counts the number of length-2 closed walks starting and ending at i.

9 Power Method for Dominant Eigenvalues

Algorithm 9.1: Power Method

Let A be a matrix with eigenvalues $|\lambda_1| \leq \cdots \leq |\lambda_n|$ and corresponding orthonormal eigenvectors $u_1, \ldots, u_n$, with λ_n dominant (i.e. $|\lambda_n| > |\lambda_{n-1}|$). Given an initial vector $x^{(0)}$ with $u_n^T x^{(0)} \neq 0$, define

$$x^{(k+1)} = \frac{Ax^{(k)}}{\|Ax^{(k)}\|}, \quad k = 0, 1, 2, \ldots$$

Then $x^{(k)} \to \pm u_n$ as $k \to \infty$, and the Rayleigh quotient

$$\rho(x^{(k)}) = \frac{(x^{(k)})^T A x^{(k)}}{(x^{(k)})^T x^{(k)}}$$

converges to λ_n.

Conditions for convergence:

- The dominant eigenvalue must be unique in magnitude.
- The starting vector must have a non-zero component in the dominant eigenvector direction.

- The matrix should not be defective in a way that spoils convergence (usually not an issue for symmetric matrices).

Worked Example Sketch: Convergence of Power Method

Write $x^{(0)} = \sum_{i=1}^{n} \alpha_i u_i$ with $\alpha_n \neq 0$. Then

$$A^k x^{(0)} = \sum_{i=1}^{n} \alpha_i \lambda_i^k u_i = \lambda_n^k \left(\alpha_n u_n + \sum_{i=1}^{n-1} \alpha_i \left(\frac{\lambda_i}{\lambda_n} \right)^k u_i \right).$$

Since $|\lambda_i/\lambda_n| < 1$ for $i < n$, the bracket tends to $\alpha_n u_n$ as $k \to \infty$. After normalisation, $x^{(k)}$ converges to $\pm u_n$. The Rayleigh quotient then converges to λ_n.

Common Mistake

If the dominant eigenvalue is negative, the direction of $x^{(k)}$ may alternate sign. This does *not* mean the method is failing; eigenvectors are only defined up to sign. Watch the Rayleigh quotient for convergence of the eigenvalue.

Exam Focus

In exam questions about the power method, focus on:

- writing $x^{(0)}$ in the eigenbasis;
- tracking the ratio λ_i^k/λ_n^k;
- describing the limit behaviour both for the vector and for the Rayleigh quotient.

You almost never need to implement the method numerically; the proof idea is the main target.

Part IV
Global Tools and Exam Strategy

10 Invertible Matrix Theorem

Key Formula

For an $n \times n$ matrix A, the following statements are equivalent:

1. A is invertible.
2. $\det(A) \neq 0$.
3. $\operatorname{rank}(A) = n$.
4. $\operatorname{null}(A) = \{\mathbf{0}\}$.
5. The columns of A are linearly independent.
6. A is row-equivalent to the identity matrix I_n.
7. 0 is not an eigenvalue of A.
8. A can be expressed as a product of elementary matrices.

This theorem compresses many different ways of thinking about invertibility into one list. In proofs and exam questions, you frequently show that one condition implies another, or you pick whichever viewpoint is most convenient.

Using Different Faces of Invertibility

Suppose A is $n \times n$ and you know that $\operatorname{rank}(A) = n$. Then

- there is a unique solution to $Ax = b$ for every b;
- A is row-equivalent to I_n (by Gaussian elimination);
- A has non-zero determinant and can be written as a product of elementary matrices;
- 0 is not an eigenvalue of A.

Depending on the question, you may choose any of these as your starting point.

Exam Focus

In proofs, explicitly state which version of the invertible matrix theorem you are using. For instance: "Since $\det(A) \neq 0$, A is invertible and hence has full rank" is clearer than "Clearly A is invertible".

11 Rank-Nullity Theorem

Theorem 11.1: F

r an $m \times n$ matrix A,

$$\operatorname{rank}(A) + \operatorname{nullity}(A) = n,$$

where $\operatorname{nullity}(A) = \dim(\operatorname{null}(A))$.

Interpretation: the domain $\mathbb{R}^n$ is decomposed into

$$\underbrace{\operatorname{null}(A)}_{\text{“invisible directions”}} \oplus \underbrace{\text{complementary subspace}}_{\text{mapped injectively into range}}.$$

Rank-nullity is a dimension bookkeeping tool: if you know the rank, you immediately know the number of degrees of freedom in the homogeneous solution.

Homogeneous Solution Space Dimension

If A is 4×6 and $\operatorname{rank}(A) = 3$, then any homogeneous system $Ax = 0$ has a solution space of dimension $6 - 3 = 3$. So you expect to find 3 free variables in row-reduced echelon form.

12 Cayley–Hamilton and Cramer's Rule

Theorem 12.1: Cayley–Hamilton

very square matrix A satisfies its own characteristic polynomial $p_A(t) = \det(tI - A)$:

$$p_A(A) = 0.$$

Typical applications:

- expressing A^k as a linear combination of lower powers of A;
- computing A^{-1} when A is 2×2 or 3×3 using the polynomial identity;
- deriving recurrence relations satisfied by sequences like $x_k = A^k x_0$.

Theorem 12.2: Cramer's Rule

f A is invertible and $Ax = b$, then

$$x_i = \frac{\det(A_i)}{\det(A)},$$

where A_i is obtained from A by replacing the i-th column by b.

Cramer's rule is conceptually elegant but numerically inefficient for large systems. In exam settings, it is mainly used to:

- prove existence and uniqueness of solutions for small systems;
- derive explicit formulas for x_i in terms of determinants.

Common Mistake

Do not use Cramer's rule on high-dimensional systems unless the question *explicitly* asks for it. Gaussian elimination or matrix inverses are usually preferred; Cramer is mainly for theoretical arguments and low dimensions.

A Formula Sheet

Matrix Algebra

Key Formula

$$
\begin{aligned}
(AB)^T &= B^T A^T, & (A^T)^T &= A,\\
(AB)^{-1} &= B^{-1}A^{-1}, & (A^T)^{-1} &= (A^{-1})^T,\\
\operatorname{tr}(AB) &= \operatorname{tr}(BA), & \operatorname{tr}(A+B) &= \operatorname{tr}(A)+\operatorname{tr}(B).
\end{aligned}
$$

Determinants

Key Formula

For $A \in \mathbb{R}^{n\times n}$ and scalar k:

$$
\begin{aligned}
\det(AB) &= \det(A)\det(B),\\
\det(A^T) &= \det(A),\\
\det(A^{-1}) &= \frac{1}{\det(A)} \ (\det A \neq 0),\\
\det(kA) &= k^n \det(A).
\end{aligned}
$$

Row operations:

$$
\begin{aligned}
\text{row swap}: &\ \det \to -\det,\\
\text{row} \times k: &\ \det \to k \det,\\
\text{row} + \alpha(\text{other row}): &\ \det \ \text{unchanged}.
\end{aligned}
$$

Eigenvalues & Diagonalisation

Key Formula

Let A have eigenvalues $\lambda_1, \dots, \lambda_n$:

$$\operatorname{tr}(A) = \sum_{i=1}^n \lambda_i, \qquad \det(A) = \prod_{i=1}^n \lambda_i.$$

If A is diagonalizable with eigenbasis columns of P:

$$P^{-1}AP = \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n), \qquad A^k = P\Lambda^k P^{-1}.$$

Orthogonality & Projections

Key Formula

If $\{q_1, \dots, q_k\}$ is orthonormal and

$$x = \sum_{i=1}^k \alpha_i q_i,$$

then

$$\alpha_i = q_i^T x, \qquad \|x\|^2 = \sum_{i=1}^k \alpha_i^2.$$

Orthogonal projector onto $\operatorname{span}\{q_1, \dots, q_k\}$:

$$P = QQ^T, \quad Q = [q_1 \ \cdots \ q_k].$$

Graph Laplacian & Paths

Key Formula

For Laplacian $L = D - A$ of a graph:

$$L\mathbf{1} = 0, \qquad L = BB^T \ \Rightarrow \ L \succeq 0.$$

For adjacency matrix A:

$$(A^N)_{ij} = \text{number of length-}N \text{ walks from } v_i \text{ to } v_j.$$

Power Method

Key Formula

Power iteration:

$$x^{(k+1)} = \frac{Ax^{(k)}}{\|Ax^{(k)}\|},$$

converges (under standard assumptions) to an eigenvector of the dominant eigenvalue. The Rayleigh quotient

$$\rho(x) = \frac{x^T A x}{x^T x}$$

approximates that eigenvalue.

Summary Identities

Rank–Nullity:

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

Cayley–Hamilton:

$$p_A(A) = 0, \quad p_A(t) = \det(tI - A).$$

Invertible Matrix Theorem (short):

$$A^{-1} \text{ exists} \iff \operatorname{rank}(A) = n \iff 0 \notin \{\lambda_i\}.$$

Gram–Schmidt (one step):

$$u_i = v_i - \sum_{j<i} (v_i^T q_j) q_j, \quad q_i = \frac{u_i}{\|u_i\|}.$$