MH1201 Linear Algebra II – Revision Notes

Quantitative Research Society @ NTU

Academic Year 2025–2026

Abstract

Comprehensive revision notes for MH1201 Linear Algebra II, integrating theoretical concepts, key formulas, worked examples, and exam-focused applications. These notes cover abstract vector spaces, linear transformations, matrix representations, eigenvalues, Jordan Canonical Form, and inner-product spaces.

Topic Area	Study Weeks	Key Concepts	ILOs
Vector Spaces: Definition & Subspaces	1-2	Definition of vector space; examples	1
Finite-Dimensional Vector Spaces	3–4	Linear independence, span, bases, dimension	2, 3
Linear Transformations	5–8	Vector spaces of linear transformations; null space & range; injectivity & surjectivity; Rank-Nullity Theorem; matrices; matrix multiplication; invertible transformations; isomorphic spaces; composition	4, 5
Eigenvalues & Eigenvectors	9–10	Eigenvalues, eigenvectors, characteristic polynomials, diagonalization, minimal polynomials, Jordan Canonical Form	6, 7
Inner-Product Spaces	11–12	Inner products, norms, Cauchy–Schwarz Inequality, orthogonality, orthonormal bases, Gram–Schmidt Orthogonalization	8, 9
Course Summary & Review	13	Review of all ILOs (1–9)	1–9

For academic reference only. This document is provided for educational purposes and does not constitute an official question paper, answer key, or any formally authorised academic material. Its content is not guaranteed to be complete or accurate.

Contents

How to Use This Document 4

1 Vector Spaces and Subspaces 5
- 1.1 Definition of a Vector Space . . . 5
- 1.2 Examples of Vector Spaces . . . 7
- 1.3 Vector Subspaces . . . 7
- 1.4 Span, Generating Sets, and Linear Combinations . . . 12
- 1.5 Subspace Operations . . . 15

2 Finite-Dimensional Vector Spaces 23
- 2.1 Span and Generating Sets . . . 23
- 2.2 Linear Independence and Dependence . . . 24
- 2.3 Bases and Coordinate Representations . . . 27
- 2.4 Dimension and Core Theorems in Finite Dimensions . . . 31
- 2.5 Adding Vectors, Extending Bases, and Extracting Bases . . . 34
- 2.6 Dimension Computations and Dimension Formulae . . . 36
- 2.7 High-Yield Worked Problems from Tutorial 3 . . . 41

3 Linear Transformations 45
- 3.1 Definitions, linearity tests, and core examples (Week 5) . . . 45
- 3.2 Kernel (null space) and range (image) (Week 5) . . . 47
- 3.3 Rank–Nullity theorem and injective/surjective criteria (Weeks 5–6) . . . 48
- 3.4 Matrix representations and composition (Week 6) . . . 51
- 3.5 Change of coordinates and change of matrix representations (Week 7) . . . 53
- 3.6 Similarity of square matrices (Week 8) . . . 57
- 3.7 Eigenvalues, eigenvectors, eigenspaces (Week 8) . . . 59

4 Matrix Representations and Change of Basis 68
- 4.1 Ordered bases and coordinate vectors . . . 68
- 4.2 Matrix representation of a linear map . . . 68
- 4.3 Change of coordinates and transition matrices . . . 70
- 4.4 Representation matrices under basis change . . . 73
- 4.5 Composition, identity, and inverse maps . . . 74
- 4.6 Similarity viewpoint and basis-independence statements . . . 77
- 4.7 Common workflows (what to do on exams) . . . 79

5 Eigenvalues, Diagonalization, and Jordan Canonical Form 80
- 5.1 Eigenvalues, eigenvectors, and eigenspaces . . . 80
- 5.2 Characteristic polynomial and computing eigenvalues . . . 82
- 5.3 Eigenspaces: direct sum structure and dimension bounds . . . 83
- 5.4 Diagonalization: definitions, criteria, and multiplicities . . . 84
- 5.5 Core applications: powers, sums of powers, and conjugacy . . . 87
- 5.6 Edge cases and diagnostic traps . . . 90

6 Inner Product Spaces **93**
6.1 Inner products and norms . . . 93
6.2 Cauchy–Schwarz and the triangle inequality . . . 94
6.3 Orthogonality, orthonormal bases, and Fourier coefficients . . . 96
6.4 Orthogonal projection onto a vector . . . 97
6.5 Gram–Schmidt orthogonalization . . . 98
6.6 Orthogonal complements . . . 99
6.7 Orthogonal projection onto a subspace . . . 101
6.8 Frobenius inner product and orthogonal complements in matrix spaces . . . 102
6.9 Least squares viewpoint (if assessed in the module) . . . 103

7 Comprehensive Course Summary and Final Review **104**
7.1 The logical architecture of MH1201 . . . 104
7.2 Vector spaces, subspaces, span, and direct sums . . . 105
7.3 Linear independence, bases, coordinates, and dimension . . . 107
7.4 Linear transformations, kernel, range, rank, and invertibility . . . 110
7.5 Matrices, coordinates, change of basis, and similarity . . . 111
7.6 Eigenvalues, eigenspaces, diagonalization, and operator structure . . . 113
7.7 Inner products, orthogonality, orthonormal bases, and projections . . . 117
7.8 Standard problem-solving workflows . . . 122
7.9 Common mistakes and hidden assumptions . . . 123
7.10 Representative worked illustrations . . . 124
7.11 Final comprehensive revision checklist . . . 127

A Formula & Key Results Reference **128**

B Notation Reference **140**

C Examination and Administrative Information **142**
C.1 Assessment Breakdown . . . 142
C.2 Important Dates . . . 142
C.3 Help-Sheet Policy for Exams . . . 142
C.4 Study Resources . . . 142

How to Use This Document

Structure of the Notes

These notes are designed to be used together with your lectures and tutorials. Each chapter is organised into:

1. **Core Theory** — formal definitions, theorems, and mathematical intuition.
2. **Mathematical Framework** — key formulas, identities, and derivations.
3. **Worked Examples** — tutorial-style questions with full solutions.
4. **Exam Focus** — high-yield concepts, typical traps, and checklist items.

A good way to revise:

- First pass: read Core Theory + Key Formula boxes.
- Second pass: attempt Worked Examples yourself *before* checking solutions.
- Final pass: review Exam Focus and Common Mistake boxes the week before the exam.

Colour and Box System (Legend)

The document uses a colour-coded 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.
Yellow	**Policy Application**	Real-world or mathematical interpretations to help build intuition.
Purple	**Exam Focus**	High-yield checklists, "must-know" results and exam strategy tips.

When you print in black and white, you can still identify boxes by their **title text**:

- Look for the name at the top: *Key Formula*, *Worked Example*, *Common Mistake*, *Policy Application*, *Exam Focus*.
- In digital version, use the colour + name; in printed version, rely mainly on the name.

1 Vector Spaces and Subspaces

This chapter collects the core definitions and proof templates for *vector spaces*, *subspaces*, *span*, and basic subspace operations. The material corresponds to Study Weeks 1–2 and addresses ILO 1.

1.1 Definition of a Vector Space

Definition 1.1: Vector Space

Let $\mathbb{F}$ be a field (in MH1201, typically $\mathbb{R}$ or $\mathbb{C}$).
A *vector space over* $\mathbb{F}$ is a set V equipped with

- a binary operation $+_V : V \times V \to V$ (vector addition), and
- a scalar multiplication $\cdot_V : \mathbb{F} \times V \to V$,

such that for all $u, v, w \in V$ and all $c, d \in \mathbb{F}$, the following axioms hold:

(VS1) $(u +_V v) +_V w = u +_V (v +_V w)$ (associativity of $+_V$)

(VS2) $u +_V v = v +_V u$ (commutativity of $+_V$)

(VS3) $c \cdot_V (u +_V v) = (c \cdot_V u) +_V (c \cdot_V v)$ (left distributivity)

(VS4) $(c + d) \cdot_V u = (c \cdot_V u) +_V (d \cdot_V u)$ (right distributivity)

(VS5) $c \cdot_V (d \cdot_V u) = (cd) \cdot_V u$ (compatibility with field multiplication)

(VS6) $1_{\mathbb{F}} \cdot_V u = u$ (scalar identity)

(VS7) There exists $0_V \in V$ such that $u +_V 0_V = u$ for all $u \in V$ (additive identity)

(VS8) For each $u \in V$ there exists $(-u) \in V$ such that $u +_V (-u) = 0_V$. (additive inverse)

The definition of a vector space distills two fundamental ideas: (i) you can *add* objects together, and (ii) you can *scale* objects by elements of a field. The eight axioms guarantee that these two operations interact in a "reasonable" way—exactly the way they do for arrows in $\mathbb{R}^n$. In particular, axioms (VS1)–(VS2) and (VS7)–(VS8) ensure that $(V, +_V)$ is an *abelian group*, while axioms (VS3)–(VS6) tie scalar multiplication to the field structure.

Exam Focus

Exam strategy for verifying a vector space from scratch. In most exam situations you will *not* verify all eight axioms. Instead, you verify that V is a subspace of a known vector space (see Theorem 1.1) or recognise V as one of the standard examples ($\mathbb{F}^n$, $\mathbb{F}^{m,n}$, $P_n(\mathbb{F})$, $\mathbb{F}^S$). The only time you need all eight axioms is when the operations are *non-standard* (e.g. a "weird" addition or scalar multiplication).

Policy Application

Terminology and notation.
When the ambient space V is clear, we usually write $u + v$ and cu instead of $u +_V v$ and $c \cdot_V u$. Likewise, we write 0 for 0_V when there is no ambiguity.

Key Formula

Standard derived identities in any vector space.
Let V be a vector space over $\mathbb{F}$. For all $u \in V$ and $c \in \mathbb{F}$,

$$-(-u) = u, \qquad 0_{\mathbb{F}}\, u = 0_V, \qquad (-1_{\mathbb{F}})\, u = -u, \qquad c\, 0_V = 0_V.$$

How to prove: each follows from the axioms via cancellation (add the same vector to both sides) and distributivity.

Worked Example

Worked Example: Proving $0_{\mathbb{F}}\, u = 0_V$ from the axioms.
We want to show that multiplying any vector by the field zero gives the vector-space zero.
Step 1. Observe that $0_{\mathbb{F}} = 0_{\mathbb{F}} + 0_{\mathbb{F}}$ in $\mathbb{F}$.
Step 2. Using (VS4) (right distributivity):

$$0_{\mathbb{F}}\, u = (0_{\mathbb{F}} + 0_{\mathbb{F}})\, u = 0_{\mathbb{F}}\, u + 0_{\mathbb{F}}\, u.$$

Step 3. Add the additive inverse $-(0_{\mathbb{F}}\, u)$ to both sides:

$$0_V = 0_{\mathbb{F}}\, u.$$

The proofs of the other three identities follow a similar pattern: exploit a trivial identity in $\mathbb{F}$ (or in V), distribute, and cancel.

Definition 1.2: Vector Subtraction

Let V be a vector space. The *vector subtraction* operation is defined by

$$u - v := u + (-v).$$

Worked Example

Worked Example: A "shifted" subset can fail to be a vector space.
Let $V = \mathbb{R}^3$ with the usual operations, and let

$$W = \{(a, b, c) \in \mathbb{R}^3 \mid a + 2b + 3c = 4\}.$$

Then W is *not* a vector subspace of $\mathbb{R}^3$ because the zero vector is not in W:

$$(0, 0, 0) \notin W \quad \text{since} \quad 0 \neq 4.$$

(Geometrically, W is an affine plane not passing through the origin.)

Common Mistake

"I checked closure, so it is a subspace."
Closure under addition and scalar multiplication *must* be accompanied by non-emptiness (equivalently, $0 \in W$).
A fast disproof for subspaces is: check whether 0 belongs.

1.2 Examples of Vector Spaces

Unless stated otherwise, the operations below are the usual entry-wise addition and scalar multiplication.

- $\mathbb{F}^n$ for $n \in \mathbb{N}$.
- $\mathbb{F}^{m,n}$: the set of all $m \times n$ matrices with entries in $\mathbb{F}$.
- $P_n(\mathbb{F})$: polynomials with coefficients in $\mathbb{F}$ of degree $\leq n$.
- $\mathbb{F}^{\mathbb{F}}$ (or $\mathbb{R}^{\mathbb{R}}$): functions $\mathbb{F} \to \mathbb{F}$ under pointwise operations.
- Non-example: view $\mathbb{R}$ as a vector space over $\mathbb{R}$. Then $\mathbb{Q} \subseteq \mathbb{R}$ is *not* a subspace since it is not closed under scalar multiplication by irrational scalars (e.g. $\sqrt{2} \cdot 1 \notin \mathbb{Q}$).

Policy Application

The field matters. Whether a set is a vector space (or subspace) can depend on *which field* you use. For instance, $\mathbb{Q}$ is *not* a subspace of $\mathbb{R}$ over $\mathbb{R}$ (scalar closure fails), but $\mathbb{Q}$ *is* a vector space over $\mathbb{Q}$ (with usual addition and multiplication). Always check what field the problem specifies.

Worked Example

Worked Example: $P_3(\mathbb{R})$ is a vector space.
Let $P_3(\mathbb{R}) = \{a_0 + a_1X + a_2X^2 + a_3X^3 : a_i \in \mathbb{R}\}$ with the usual polynomial addition and scalar multiplication.

- *Closure:* if $p, q \in P_3(\mathbb{R})$, then $p + q$ has degree ≤ 3, and if $c \in \mathbb{R}$ then cp has degree ≤ 3.
- The remaining axioms reduce to the corresponding properties of real numbers, verified coefficient-wise.

Hence $P_3(\mathbb{R})$ is a vector space over $\mathbb{R}$.

Worked Example

Worked Example (Lecture Week 2): A differential-equation solution set as a subspace.
Consider the set

$$S = \{P \in P(\mathbb{R}) : P'' + 7P' - 3P = 0\}.$$

Define $T : P(\mathbb{R}) \to P(\mathbb{R})$ by $T(P) = P'' + 7P' - 3P$. Then T is linear (differentiation and scalar multiplication are linear operations), and $S = \ker(T)$. Since the kernel of a linear map is always a subspace, S is a subspace of $P(\mathbb{R})$.

Key idea (kernel viewpoint): whenever a subset is defined by setting a *linear* expression equal to zero, recognise it as a kernel and conclude it is a subspace immediately.

1.3 Vector Subspaces

Definition 1.3: Vector Subspace

Let V be a vector space over $\mathbb{F}$, and let $W \subseteq V$.
We say that W is a *vector subspace of* V if W becomes a vector space over $\mathbb{F}$ when equipped with the *restricted* operations inherited from V.

Intuitively, a subspace is a "flat" subset of V that passes through the origin and is "closed" under the ambient operations. In $\mathbb{R}^3$, subspaces are: $\{0\}$, lines through the origin, planes through the origin, and $\mathbb{R}^3$ itself. A line or plane that does *not* pass through the origin is *not* a subspace (it is an *affine* subspace).

Theorem 1.1: Subspace Test

Let V be a vector space over $\mathbb{F}$ and let $W \subseteq V$.
Then W is a vector subspace of V if and only if:

(1) $W \neq \emptyset$;

(2) for all $u, v \in W$, we have $u + v \in W$;

(3) for all $c \in \mathbb{F}$ and $u \in W$, we have $cu \in W$.

Proof. ($\Rightarrow$) If W is a vector space under the inherited operations, then it contains its additive identity (so $W \neq \emptyset$) and is closed under its own addition and scalar multiplication; since these coincide with those of V, (2) and (3) hold.

($\Leftarrow$) If (1)–(3) hold, then the restricted operations map $W \times W \to W$ and $\mathbb{F} \times W \to W$.

All vector space axioms for W follow because they hold in V and all terms appearing remain in W by (2)–(3).

□

Policy Application

Why only three conditions? The Subspace Test works because six of the eight vector space axioms (associativity, commutativity, distributivities, compatibility, scalar identity) are *inherited* from the ambient space V—they hold for *all* vectors in V, so they certainly hold for vectors restricted to W. The only conditions that can fail are: (i) the operations might not map back into W (closure), and (ii) the identity/inverse elements might not lie in W (non-emptiness). But once $W \neq \emptyset$ and closure holds, picking any $w \in W$ gives $0 = 0_{\mathbb{F}}\, w \in W$ (by scalar closure), and $-w = (-1)w \in W$ (by scalar closure), so the identity and inverses are automatic.

Key Formula

Combined one-line subspace test. W is a subspace of V if and only if $W \neq \emptyset$ and for all $u, v \in W$ and $c \in \mathbb{F}$:

$$cu + v \in W.$$

This single condition combines closure under scalar multiplication and addition. It is sometimes faster in proofs.

Key Formula

Non-subspace test (fast disproof).
If $0_V \notin W$, then W is *not* a vector subspace of V.

Useful "kernel viewpoint".
If $T : V \to U$ is a linear map, then $\ker(T) = \{v \in V : T(v) = 0\}$ is a vector subspace of V.
In particular, solution sets of *homogeneous* linear equations are subspaces.

Worked Example

Worked Example (Tutorial 1, Question 1): Subspace test + describing the set.
Consider

$$U = \{(a, b, c) \in \mathbb{R}^3 \mid a = 5c\}.$$

Subspace test.
The set is non-empty because $(0, 0, 0) \in U$.
If $(a_1, b_1, c_1), (a_2, b_2, c_2) \in U$, then $a_1 = 5c_1$ and $a_2 = 5c_2$, so

$$a_1 + a_2 = 5(c_1 + c_2),$$

hence $(a_1 + a_2, b_1 + b_2, c_1 + c_2) \in U$ (closure under addition).
If $(a, b, c) \in U$ and $\lambda \in \mathbb{R}$, then $\lambda a = 5(\lambda c)$, so $(\lambda a, \lambda b, \lambda c) \in U$ (closure under scalar multiplication).
Thus U is a subspace.

Describe U as a span.
Any $(a, b, c) \in U$ satisfies $a = 5c$, so

$$(a, b, c) = (5c, b, c) = c(5, 0, 1) + b(0, 1, 0) \in \text{Span}\{(5, 0, 1), (0, 1, 0)\}.$$

Conversely, any linear combination $c(5, 0, 1) + b(0, 1, 0) = (5c, b, c)$ satisfies $a = 5c$, so it lies in U.
Therefore

$$U = \text{Span}\{(5, 0, 1), (0, 1, 0)\}.$$

(This is a typical *double-inclusion* proof of subspace equality.)

Extension (Tutorial 1, Question 1, Method 2: Kernel viewpoint)

Define $L : \mathbb{R}^3 \to \mathbb{R}$ by $L(a, b, c) = a - 5c$. Then L is linear because:

$$L\big((a_1, b_1, c_1) + (a_2, b_2, c_2)\big) = (a_1 + a_2) - 5(c_1 + c_2) = L(a_1, b_1, c_1) + L(a_2, b_2, c_2),$$

and $L(\lambda(a, b, c)) = \lambda a - 5\lambda c = \lambda L(a, b, c)$. Since $U = \{(a, b, c) : L(a, b, c) = 0\} = \ker(L)$, and the kernel of a linear map is always a subspace, U is a subspace.

When to use the kernel viewpoint: whenever a subset is defined by setting a linear expression (in the coordinates/entries) to zero, you can define the corresponding linear map and immediately conclude the set is a subspace as its kernel. This avoids checking closure conditions manually.

Key Formula

Immediate consequences for subspaces.
If W is a subspace of V, then:

- $0 \in W$;
- if $u \in W$, then $-u \in W$ (take scalar $-1_{\mathbb{F}}$);
- if $u, v \in W$, then $u - v \in W$ (since $u - v = u + (-v)$).

These are often used to simplify subspace proofs.

Worked Example

Worked Example (Tutorial 1, Question 2): A subspace inside a function space.
Let $\mathcal{D}$ be the set of differentiable functions $f : \mathbb{R} \to \mathbb{R}$, viewed as a subspace of $\mathbb{R}^{\mathbb{R}}$ under pointwise operations.
Consider

$$W = \{f \in \mathcal{D} : f'(1) = 3f(2)\}.$$

Non-empty: the zero function $f \equiv 0$ lies in W because $0 = 3 \cdot 0$.
Closure under addition: if $f, g \in W$, then using linearity of differentiation and evaluation,

$$(f+g)'(1) = f'(1) + g'(1) = 3f(2) + 3g(2) = 3(f+g)(2),$$

so $f + g \in W$.
Closure under scalar multiplication: if $c \in \mathbb{R}$ and $f \in W$, then

$$(cf)'(1) = cf'(1) = c \cdot 3f(2) = 3(cf)(2),$$

so $cf \in W$. Hence W is a subspace of $\mathbb{R}^{\mathbb{R}}$ (and of $\mathcal{D}$).

Extension (Tutorial 1, Question 2, Method 2: Kernel of a linear functional)

Define $L : \mathcal{D} \to \mathbb{R}$ by $L(f) = f'(1) - 3f(2)$. We verify linearity:

$$L(f+g) = (f+g)'(1) - 3(f+g)(2) = (f'(1) - 3f(2)) + (g'(1) - 3g(2)) = L(f) + L(g),$$
$$L(\lambda f) = (\lambda f)'(1) - 3(\lambda f)(2) = \lambda(f'(1) - 3f(2)) = \lambda L(f).$$

Since $W = \ker(L)$ and the kernel of a linear map is a subspace, W is a subspace.

Takeaway: the kernel viewpoint generalises to *any* subset defined by a condition of the form "*linear expression* $= 0$." For function spaces, common linear operations include evaluation ($f \mapsto f(a)$), differentiation ($f \mapsto f'(a)$), and integration.

Worked Example

Worked Example (Tutorial 1, Question 6): Disproving a subspace via the zero element.
Let

$$S = \left\{ \begin{pmatrix} a & b \\ 3b - 4c & a + 2 \end{pmatrix} \in \mathbb{R}^{2,2} \;\middle|\; a, b, c \in \mathbb{R} \right\}.$$

If S were a subspace, it must contain the zero matrix.
But the $(2,2)$-entry is always $a + 2$, so for any matrix in S we have $(2,2) = a + 2$.
To get the zero matrix, we would need $a + 2 = 0$ and simultaneously the $(1,1)$-entry $a = 0$, which is impossible.
Therefore $0 \notin S$ and S is *not* a subspace.

Worked Example

Worked Example: Homogeneous vs. non-homogeneous constraints.
Let

$$W_0 = \{(a,b,c) \in \mathbb{R}^3 \mid a+2b+3c = 0\}, \qquad W_4 = \{(a,b,c) \in \mathbb{R}^3 \mid a+2b+3c = 4\}.$$

Define $L : \mathbb{R}^3 \to \mathbb{R}$ by $L(a,b,c) = a+2b+3c$.
Then $W_0 = \ker(L)$ is a subspace (kernel viewpoint).
In contrast, W_4 is *not* a subspace since $0 \notin W_4$.

Key Formula

When is an affine subset a subspace? (Lecture Week 1)
Let $a_1, \ldots, a_n, b \in V$. Then

$$\{\lambda_1 a_1 + \cdots + \lambda_n a_n + b : \lambda_i \in \mathbb{F}\}$$

is a subspace of V if and only if $b \in \text{Span}\{a_1, \ldots, a_n\}$.

Intuition: if b already lies in the span of the a_i's, then the "$+b$" shift does not move the set away from the origin—the set is the same as $\text{Span}\{a_1, \ldots, a_n\}$. Otherwise, the set is a genuine affine translate and misses the origin.

Worked Example

Worked Example (Lecture Week 1): Affine subset test.
Is the set $S = \left\{ \lambda \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \;\middle|\; \lambda, \mu \in \mathbb{R} \right\}$ a subspace of $\mathbb{R}^{3,1}$?
We need to check if $b = (0,1,0)^T \in \text{Span}\{(1,0,-1)^T, (1,2,3)^T\}$. Solving $\lambda(1,0,-1)^T + \mu(1,2,3)^T = (0,1,0)^T$ gives the system $\lambda+\mu = 0$, $2\mu = 1$, $-\lambda+3\mu = 0$. From the second equation $\mu = \frac{1}{2}$, so $\lambda = -\frac{1}{2}$.
Check the third: $\frac{1}{2} + \frac{3}{2} = 2 \neq 0$. Contradiction.
Hence $b \notin \text{Span}\{a_1, a_2\}$, so S is *not* a subspace.

Common Mistake

Confusing "subset" with "subspace".
Even if $W \subseteq V$, W need not inherit the vector space structure unless it is closed under *both* addition and scalar multiplication and contains 0.

Example (Tutorial 1, Question 3): a set may be closed under addition and additive inverses, yet fail scalar closure (e.g. integer lattices in $\mathbb{R}^2$).

Policy Application

Detail on the $\mathbb{Z}^2$ non-example (Tutorial 1, Question 3). $\mathbb{Z}^2 = \{(m,n) : m,n \in \mathbb{Z}\}$ is closed under addition (sums of integers are integers) and additive inverses (negatives of integers are integers). However, $\frac{1}{2}(1,0) = (\frac{1}{2}, 0) \notin \mathbb{Z}^2$, so it fails scalar closure over $\mathbb{R}$. This shows that closure under addition and inverses alone is *not sufficient* for the subspace property—you must also verify scalar closure.

Exam Focus

- **Fast subspace proof templates:**
  (i) use the Subspace Test; or
  (ii) recognize W as a kernel / solution set to homogeneous linear equations; or
  (iii) write $W = \text{Span}(S)$ for some set S.
- **Fast disproof:** check whether $0 \in W$. If not, you are done.
- **Do not claim "closed under subtraction" in place of scalar closure:**
  closure under subtraction implies closure under addition *only if* you already know closure under additive inverses, but it still does not give scalar closure.

1.4 Span, Generating Sets, and Linear Combinations

Definition 1.4: Linear Combination and Span

Let V be a vector space over $\mathbb{F}$.

- Given vectors $v_1, \ldots, v_n \in V$ ($n \in \mathbb{N}_0$), a *linear combination* of $v_1, \ldots, v_n$ is a vector of the form
$$\lambda_1 v_1 + \cdots + \lambda_n v_n \qquad (\lambda_i \in \mathbb{F}).$$
By convention, the "empty" linear combination (when $n = 0$) equals 0_V.
- Given a subset $S \subseteq V$ (not necessarily finite), the *span* of S is
$$\text{Span}(S) = \{\text{all linear combinations of finitely many vectors in } S\}.$$

We say that S *spans* V if $\text{Span}(S) = V$, and call S a *generating set* of V.

The span of a set S captures *all vectors reachable* from S using only addition and scaling. It is the "smallest sandbox" containing S that is closed under the vector space operations. A key subtlety: even if S is infinite, each individual element of $\text{Span}(S)$ uses only *finitely many* vectors from S (infinite sums are not defined in a general vector space without topology).

Key Formula

What is $\text{Span}(\emptyset)$? $\text{Span}(\emptyset) = \{0_V\}$.

Reason: the only linear combination using no vectors is 0_V.

Theorem 1.2: The span of a set is a vector subspace

Let V be a vector space over $\mathbb{F}$ and let $S \subseteq V$.
Then $\text{Span}(S)$ is a vector subspace of V.

Proof. *Non-empty:* $0_V \in \text{Span}(S)$ as the empty linear combination.
Closed under addition: if $u = \sum_{i=1}^{m} \alpha_i s_i$ and $v = \sum_{j=1}^{n} \beta_j t_j$

with $s_i, t_j \in S$, then

$$u + v = \sum_{i=1}^{m} \alpha_i s_i + \sum_{j=1}^{n} \beta_j t_j$$

is a linear combination of finitely many vectors in S, hence is in $\mathrm{Span}(S)$.
Closed under scalar multiplication: for $c \in \mathbb{F}$,

$$c\,u = c \sum_{i=1}^{m} \alpha_i s_i = \sum_{i=1}^{m} (c\alpha_i) s_i \in \mathrm{Span}(S).$$

Thus $\mathrm{Span}(S)$ is a subspace by the Subspace Test.

□

Theorem 1.3: Smallest subspace containing a set

Let V be a vector space over $\mathbb{F}$ and $S \subseteq V$.
Then $\mathrm{Span}(S)$ is the *smallest* vector subspace of V that contains S:

(1) $S \subseteq \mathrm{Span}(S)$;

(2) if W is a subspace of V with $S \subseteq W$, then $\mathrm{Span}(S) \subseteq W$.

Proof. (1) For any $s \in S$, we have $s = 1_{\mathbb{F}}\, s$, hence $s \in \mathrm{Span}(S)$.
(2) Let W be a subspace with $S \subseteq W$. Any element of $\mathrm{Span}(S)$ is a finite linear combination of elements of S, hence of elements of W; since W is closed under addition and scalar multiplication, that combination lies in W.
Thus $\mathrm{Span}(S) \subseteq W$.

□

Policy Application

Two equivalent ways to think about $\mathrm{Span}(S)$**.**

- *Bottom-up (constructive):* $\mathrm{Span}(S)$ consists of all *finite* linear combinations of vectors in S.
- *Top-down (minimality):* $\mathrm{Span}(S)$ is the intersection of *all* subspaces that contain S.

The bottom-up view is useful for computations; the top-down view is useful for abstract proofs (e.g. proving that a subspace containing S must contain $\mathrm{Span}(S)$).

Extension (Lecture Week 2: Span as an intersection)

Let $\mathcal{W} = \{\, W \subseteq V : \ W \text{ is a subspace and } S \subseteq W \,\}$.
Then

$$\mathrm{span}(S) = \bigcap_{W \in \mathcal{W}} W.$$

Indeed, $\mathrm{span}(S) \subseteq W$ for every such W (smallest-subspace property), so $\mathrm{span}(S)$ is contained in the intersection; conversely, $\mathrm{span}(S) \in \mathcal{W}$, so the intersection is contained in $\mathrm{span}(S)$.

Theorem 1.4: Basic properties of spans

Let V be a vector space, and let $A, B \subseteq V$.

(a) $A \subseteq \mathrm{Span}(A)$.

(b) If $A \subseteq B$, then $\mathrm{Span}(A) \subseteq \mathrm{Span}(B)$.

(c) $\mathrm{Span}(A \cup B) = \mathrm{Span}(A) + \mathrm{Span}(B)$.

(d) $\mathrm{Span}(\mathrm{Span}(A)) = \mathrm{Span}(A)$.

Proof. (a) is Theorem 1.3(1).

(b) If $A \subseteq B$, then $\mathrm{Span}(B)$ is a subspace containing A; apply Theorem 1.3(2).

(c) ($\subseteq$) Any element of $\mathrm{Span}(A \cup B)$ is a finite linear combination of vectors from $A \cup B$. Group the terms coming from A and from B to write it as $u + v$ with $u \in \mathrm{Span}(A)$ and $v \in \mathrm{Span}(B)$. Hence it lies in $\mathrm{Span}(A) + \mathrm{Span}(B)$.

($\supseteq$) If $u \in \mathrm{Span}(A)$ and $v \in \mathrm{Span}(B)$, then both u and v are linear combinations of vectors in $A \cup B$, so $u + v$ is also such a linear combination; hence $u + v \in \mathrm{Span}(A \cup B)$.

(d) Since $A \subseteq \mathrm{Span}(A)$, (b) gives $\mathrm{Span}(A) \subseteq \mathrm{Span}(\mathrm{Span}(A))$.

Conversely, $\mathrm{Span}(A)$ is already a subspace containing $\mathrm{Span}(A)$, so Theorem 1.3(2) gives $\mathrm{Span}(\mathrm{Span}(A)) \subseteq \mathrm{Span}(A)$.

□

Policy Application

Why does property (d) (idempotence) matter? It says "spanning an already-spanned set gives nothing new." Formally, Span is a *closure operator* on subsets of V: it is extensive ($A \subseteq \mathrm{Span}(A)$), monotone ($A \subseteq B \Rightarrow \mathrm{Span}(A) \subseteq \mathrm{Span}(B)$), and idempotent ($\mathrm{Span}(\mathrm{Span}(A)) = \mathrm{Span}(A)$). The fixed points of this operator are precisely the subspaces of V.

Worked Example

Worked Example: Describing Span geometrically in $\mathbb{R}^2$.
Let $S = \{(1,2), (2,4)\} \subseteq \mathbb{R}^2$.
Since $(2,4) = 2(1,2)$, any linear combination of vectors in S is a scalar multiple of $(1,2)$:

$$\lambda(1,2) + \mu(2,4) = (\lambda + 2\mu)(1,2).$$

Thus

$$\mathrm{Span}(S) = \{t(1,2) : t \in \mathbb{R}\},$$

the line through the origin in direction $(1,2)$.

Worked Example

Worked Example (Lecture Week 2): Does a set span $P_3(\mathbb{R})$?
Determine whether $S = \{4 - X + X^3,\ 1 + 2X - 3X^2,\ 3 - 3X + 3X^2 + X^3,\ X - X^3\}$ spans $P_3(\mathbb{R})$.
Strategy: express each element of S in the standard basis $\{1, X, X^2, X^3\}$ and check if the coefficient vectors span $\mathbb{R}^4$.
The coefficient vectors (reading off powers $1, X, X^2, X^3$) are:

$$(4, -1, 0, 1), \quad (1, 2, -3, 0), \quad (3, -3, 3, 1), \quad (0, 1, 0, -1).$$

Row-reducing the 4×4 matrix formed by these rows determines whether they span all of $\mathbb{R}^4$. If the rank is 4, then S spans $P_3(\mathbb{R})$; otherwise it does not.
(The full row reduction is left as an exercise; the key point is the *method*: reduce a spanning question to a rank computation.)

Common Mistake

Forgetting "finitely many" in the definition of span.
Even if S is infinite, each element of $\mathrm{Span}(S)$ must be a linear combination of *finitely many* vectors from S.

1.5 Subspace Operations

Definition 1.5: Intersection and Sum of Subspaces

Let V be a vector space over $\mathbb{F}$, and let $V_1, \ldots, V_n$ be subspaces of V.

- The *intersection* $\bigcap_{i=1}^{n} V_i$ is the set of vectors lying in every V_i.
- The *sum* $V_1 + \cdots + V_n$ is defined by

$$V_1 + \cdots + V_n = \{v_1 + \cdots + v_n \in V : v_i \in V_i \text{ for each } i\}.$$

Policy Application

Geometric intuition for intersection and sum.

- The *intersection* $V_1 \cap V_2$ is the largest subspace contained in *both* V_1 and V_2. In $\mathbb{R}^3$, the intersection of two distinct planes through the origin is typically a line through the origin.
- The *sum* $V_1 + V_2$ is the smallest subspace containing *both* V_1 and V_2. If V_1 is the xy-plane and V_2 is the z-axis, then $V_1 + V_2 = \mathbb{R}^3$.

Note the duality: intersection makes spaces *smaller*, sums make spaces *larger*.

Theorem 1.5: Intersection of subspaces is a subspace

Let V be a vector space over $\mathbb{F}$ and let $V_1, \ldots, V_n$ be subspaces of V.
Then $\bigcap_{i=1}^{n} V_i$ is a vector subspace of V.

Proof. Non-empty: $0 \in V_i$ for all i, hence $0 \in \bigcap_i V_i$.
If u, v lie in every V_i, then $u+v$ lies in every V_i by closure, so $u+v \in \bigcap_i V_i$.
Similarly, for $c \in \mathbb{F}$, $cu \in V_i$ for all i, so $cu \in \bigcap_i V_i$.

□

Common Mistake

Intersection is a subspace, but union generally is not. The intersection of any collection of subspaces is always a subspace (the proof above works for *arbitrary* intersections, even infinite). In contrast, the *union* of two subspaces is almost never a subspace (see Theorem 1.7 below). The correct "additive" counterpart of intersection is the *sum*, not the union.

Theorem 1.6: Sum of subspaces is a subspace

Let V be a vector space over $\mathbb{F}$ and let $V_1, \ldots, V_n$ be subspaces of V.
Then $V_1 + \cdots + V_n$ is a vector subspace of V.

Proof. Non-empty: $0 = 0 + \cdots + 0 \in V_1 + \cdots + V_n$.
If $x = v_1 + \cdots + v_n$ and $y = w_1 + \cdots + w_n$ with $v_i, w_i \in V_i$, then

$$x + y = (v_1 + w_1) + \cdots + (v_n + w_n) \in V_1 + \cdots + V_n$$

since each V_i is closed under addition.
For $c \in \mathbb{F}$,

$$cx = (cv_1) + \cdots + (cv_n) \in V_1 + \cdots + V_n$$

since each V_i is closed under scalar multiplication.

□

Worked Example

Worked Example (Lecture Week 2): Sum of subspaces in $\mathbb{R}^4$.
Let

$$U = \{(x, x, y, y) \in \mathbb{R}^4 : x, y \in \mathbb{R}\}, \qquad V = \{(x, x, x, y) \in \mathbb{R}^4 : x, y \in \mathbb{R}\}.$$

Find $U + V$. A generic element of $U + V$ is

$$(a, a, b, b) + (c, c, c, d) = (a + c,\ a + c,\ b + c,\ b + d).$$

Set $s = a + c$ (the first coordinate). Then the second coordinate is also s, the third is $b + c$, and the fourth is $b + d$. Since a, b, c, d range freely over $\mathbb{R}$, the third and fourth coordinates can be any real numbers independently, while the first two are always equal.
Hence

$$U + V = \{(s, s, t, u) \in \mathbb{R}^4 : s, t, u \in \mathbb{R}\}.$$

Key Formula

Relationship between sum and span.
For subspaces $W_1, W_2 \leq V$,

$$W_1 + W_2 = \mathrm{Span}(W_1 \cup W_2).$$

(This is Theorem 1.4(c) specialized to subspaces.)

Worked Example

Worked Example: Computing a sum of subspaces in $\mathbb{R}^3$.
Let

$$U = \{(x, 0, 0) \in \mathbb{R}^3 : x \in \mathbb{R}\}, \qquad V = \{(0, y, 0) \in \mathbb{R}^3 : y \in \mathbb{R}\}.$$

Then any element of $U + V$ has the form $(x, 0, 0) + (0, y, 0) = (x, y, 0)$, hence

$$U + V = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\},$$

the xy-plane.

Worked Example

Worked Example (Tutorial 1, Question 4): Union of subspaces need not be a subspace.
Let

$$W_1 = \{(x, 0) \in \mathbb{R}^2 : x \in \mathbb{R}\}, \qquad W_2 = \{(0, y) \in \mathbb{R}^2 : y \in \mathbb{R}\}.$$

Both are subspaces, but $W_1 \cup W_2$ is not: $(1, 0) \in W_1$ and $(0, 1) \in W_2$, yet

$$(1, 0) + (0, 1) = (1, 1) \notin W_1 \cup W_2.$$

Theorem 1.7: When is $W_1 \cup W_2$ a subspace?

Let V be a vector space and let W_1, W_2 be subspaces of V.
Then $W_1 \cup W_2$ is a subspace of V if and only if $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

Proof. ($\Leftarrow$) If $W_1 \subseteq W_2$, then $W_1 \cup W_2 = W_2$ is a subspace (and similarly if $W_2 \subseteq W_1$).
($\Rightarrow$) Suppose $W_1 \cup W_2$ is a subspace but neither contains the other.
Pick $u \in W_1 \setminus W_2$ and $v \in W_2 \setminus W_1$.
Then $u, v \in W_1 \cup W_2$ but $u + v \in W_1 \cup W_2$ by closure.
If $u + v \in W_1$, then $v = (u + v) - u \in W_1$ (since W_1 is closed under subtraction), contradicting $v \notin W_1$.
Similarly $u + v \notin W_2$. Contradiction. Hence one subspace must contain the other.
□

Extension (Tutorial 1, Question 5, Method 2)

Assume $W_1 \cup W_2$ is a subspace and pick $u \in W_1$, $v \in W_2$.
Since a subspace containing u and v must contain $\text{span}\{u, v\}$, we have $\text{span}\{u, v\} \subseteq W_1 \cup W_2$.
If $u \notin W_2$ and $v \notin W_1$, then $\{u, v\}$ is linearly independent, so $u + v \in \text{span}\{u, v\}$ but $u + v \notin W_1$ and $u + v \notin W_2$ (otherwise subtract to force $v \in W_1$ or $u \in W_2$), contradiction.
Therefore one of W_1, W_2 must contain the other.

Definition 1.6: Direct Sum of Subspaces

Let V be a vector space and let $V_1, \ldots, V_n$ be subspaces of V.
A subspace W is said to be the *direct sum* of $V_1, \ldots, V_n$, written

$$W = V_1 \oplus \cdots \oplus V_n,$$

if and only if:

(1) $W = V_1 + \cdots + V_n$;

(2) every $w \in W$ can be written in *precisely one* way as $w = v_1 + \cdots + v_n$ with $v_i \in V_i$.

A direct sum is a sum with "no redundancy." Condition (1) says the summands V_i together generate W; condition (2) says they do so *without overlap* (each vector has a unique "address" given by its components). Think of it as a coordinate system for W built from the V_i's.

Theorem 1.8: Necessary and sufficient condition for a sum to be a direct sum

Let V be a vector space and let $V_1, \ldots, V_n$ be subspaces of V.
Then

$$V_1 + \cdots + V_n = V_1 \oplus \cdots \oplus V_n$$

if and only if whenever $v_i \in V_i$ satisfy $v_1 + \cdots + v_n = 0$, we must have $v_1 = \cdots = v_n = 0$.

Proof. ($\Rightarrow$) If the sum is direct, then 0 has a unique representation as $0 = 0 + \cdots + 0$ with $0 \in V_i$.
So if $0 = v_1 + \cdots + v_n$, uniqueness forces $v_i = 0$ for all i.
($\Leftarrow$) Suppose the stated condition holds. Let $w \in V_1 + \cdots + V_n$ and assume

$$w = v_1 + \cdots + v_n = \tilde{v}_1 + \cdots + \tilde{v}_n \qquad (v_i, \tilde{v}_i \in V_i).$$

Then $0 = (v_1 - \tilde{v}_1) + \cdots + (v_n - \tilde{v}_n)$ with $v_i - \tilde{v}_i \in V_i$.
By the condition, each $v_i - \tilde{v}_i = 0$, hence $v_i = \tilde{v}_i$ for all i.
Thus the representation is unique, so the sum is direct.

□

Theorem 1.9: The Direct-Sum Criterion (intersection form)

Let V be a vector space and let $V_1, \ldots, V_n$ be subspaces of V. Let W be a subspace of V.
Then $W = V_1 \oplus \cdots \oplus V_n$ if and only if both of the following hold:

(1) $W = V_1 + \cdots + V_n$.

(2) For every $i \in \{1, \ldots, n\}$: $V_i \cap (V_1 + \cdots + V_{i-1} + V_{i+1} + \cdots + V_n) = \{0\}$.

Proof. ($\Rightarrow$) Suppose $W = V_1 \oplus \cdots \oplus V_n$. Then $W = V_1 + \cdots + V_n$ by definition. Let $v \in V_i \cap (V_1 + \cdots + V_{i-1} + V_{i+1} + \cdots + V_n)$. Then $v = v_1 + \cdots + v_{i-1} + v_{i+1} + \cdots + v_n$ for some $v_j \in V_j$. Rearranging: $0 = v_1 + \cdots + v_{i-1} + (-v) + v_{i+1} + \cdots + v_n$, where each summand lies in the corresponding V_j (and $-v \in V_i$). By uniqueness of decomposition (Theorem 1.8), all summands are zero. In particular $v = 0$.
($\Leftarrow$) Suppose (1) and (2) hold. Let $v_1 + \cdots + v_n = \tilde{v}_1 + \cdots + \tilde{v}_n$ with $v_i, \tilde{v}_i \in V_i$. For each i: $v_i - \tilde{v}_i = \sum_{j \neq i} (\tilde{v}_j - v_j) \in V_i \cap (\sum_{j \neq i} V_j) = \{0\}$. Hence $v_i = \tilde{v}_i$ for all i, so the decomposition is unique. □

Exam Focus

Which direct-sum test to use?

- **Theorem 1.8** ("zero-sum test"): best for *disproving* directness (exhibit a non-trivial $0 = v_1 + \cdots + v_n$) or for *abstract proofs*.
- **Theorem 1.9** ("intersection test"): best for *two* subspaces, where it reduces to checking $V_1 \cap V_2 = \{0\}$. For $n \geq 3$, the intersection conditions are more involved (you check each V_i against the sum of the others).

Key Formula

Direct sum of two subspaces (special case).
$W = V_1 \oplus V_2$ if and only if $W = V_1 + V_2$ and $V_1 \cap V_2 = \{0\}$.

This is the most commonly tested scenario. The intersection condition $V_1 \cap V_2 = \{0\}$ is equivalent to: every $w \in W$ has a *unique* decomposition $w = v_1 + v_2$ with $v_1 \in V_1$, $v_2 \in V_2$.

Extension (Lecture Week 2: Direct sums via the sum map)

Define the linear map $\phi : U \times W \to V$ by $\phi(u, w) = u + w$.
Then $\mathrm{Im}(\phi) = U + W$ and

$$\ker(\phi) = \{(u, w) : u = -w\} \cong U \cap W.$$

In particular, $U \cap W = \{0\}$ iff $\ker(\phi) = \{(0, 0)\}$ iff ϕ is injective.
If additionally $V = U + W$, then ϕ is bijective, so every $v \in V$ has a *unique* representation $v = u + w$.

Extension (Tutorial 2, Question 1, Method 2: Sum map for n subspaces)

For the general case with n subspaces, define

$$T : V_1 \times \cdots \times V_n \to V, \qquad T(v_1, \ldots, v_n) = v_1 + \cdots + v_n.$$

Then T is linear, $\mathrm{Im}(T) = V_1 + \cdots + V_n$, and

$$\ker(T) = \{(v_1, \ldots, v_n) \in V_1 \times \cdots \times V_n : v_1 + \cdots + v_n = 0\}.$$

The sum is direct if and only if T is injective, which holds if and only if $\ker(T) = \{(0, \ldots, 0)\}$. This gives a clean reformulation: *direct sum* $\iff$ *the canonical sum map is injective.*

Worked Example

Worked Example: Sum equals $\mathbb{R}^3$ but is not direct.
Let

$$U = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\}, \quad V = \{(0, 0, z) \in \mathbb{R}^3 : z \in \mathbb{R}\}, \quad W = \{(0, y, y) \in \mathbb{R}^3 : y \in \mathbb{R}\}.$$

(1) Show $\mathbb{R}^3 = U + V + W$.
Given $(a, b, c) \in \mathbb{R}^3$, write

$$(a, b, c) = (a, b - c, 0) + (0, 0, c - b) + (0, b, b),$$

where the three summands lie in U, V, and W respectively. Hence $\mathbb{R}^3 \subseteq U + V + W$, and the reverse inclusion is obvious.

(2) Show $\mathbb{R}^3 \neq U \oplus V \oplus W$.
We have a non-trivial zero decomposition:

$$0 = (0, 1, 0) + (0, 0, 1) + (0, -1, -1),$$

where $(0, 1, 0) \in U$, $(0, 0, 1) \in V$, and $(0, -1, -1) \in W$.
By Theorem 1.8, the sum cannot be direct.

Policy Application

How to find a non-trivial zero decomposition (disproof strategy). Look for vectors that live in the "overlap" between the subspaces. For the example above, V and W both involve the second and third coordinates. In particular, $(0, 1, 1) \in W$ and $(0, 0, 1) \in V$ share the third-coordinate direction, so their interaction with U (which controls the second coordinate) creates redundancy.
A systematic approach: check whether $V_i \cap (\sum_{j \neq i} V_j) \neq \{0\}$ for some i. If so, any non-zero vector in that intersection yields a non-trivial zero decomposition.

Common Mistake

"$\oplus$ just means $+$."
Direct sum is *stronger* than sum: you must also have uniqueness of decomposition (equivalently, no non-trivial way to sum to 0 across the summands).

Common Mistake

"I checked $V_i \cap V_j = \{0\}$ for all pairs, so the sum is direct."
For $n \geq 3$ subspaces, pairwise trivial intersections do *not* guarantee a direct sum. The correct condition (Theorem 1.9) requires each V_i to intersect the *sum of all the others* trivially.

Example: in $\mathbb{R}^2$, let $V_1 = \text{Span}\{(1, 0)\}$, $V_2 = \text{Span}\{(0, 1)\}$, $V_3 = \text{Span}\{(1, 1)\}$. Then $V_i \cap V_j = \{0\}$ for all pairs $i \neq j$, but $V_1 + V_2 = \mathbb{R}^2 \supseteq V_3$, so $(1, 1) = (1, 0) + (0, 1) + (0, 0) = (0, 0) + (0, 0) + (1, 1)$ gives two different decompositions of $(1, 1)$. The sum is *not* direct.

Worked Example

Worked Example (Tutorial 1, Question 7): Finding a direct-sum complement in $\mathbb{R}^5$.
Let

$$V = \{(a, b, a+b, a-b, 2a) \in \mathbb{R}^5 : a, b \in \mathbb{R}\} = \text{Span}\{(1,0,1,1,2), (0,1,1,-1,0)\}.$$

Goal: find W such that $\mathbb{R}^5 = V \oplus W$.
Step 1 (choose W): Since V is determined by its first two coordinates, a natural complement is $W = \{(0,0,s,t,u) : s,t,u \in \mathbb{R}\} = \text{Span}\{e_3, e_4, e_5\}$.
Step 2 (check $V \cap W = \{0\}$): If $x \in V \cap W$, then $x = (a, b, a+b, a-b, 2a)$ with $a = 0$ and $b = 0$ (since $x \in W$ forces the first two coordinates to be zero). So $x = 0$.
Step 3 (check $V + W = \mathbb{R}^5$): For any $(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5$, set $v = (x_1, x_2, x_1 + x_2, x_1 - x_2, 2x_1) \in V$ and $w = (0, 0, x_3 - x_1 - x_2, x_4 - x_1 + x_2, x_5 - 2x_1) \in W$. Then $v + w = (x_1, x_2, x_3, x_4, x_5)$.
Hence $\mathbb{R}^5 = V \oplus W$.

General strategy: if V has dimension k in $\mathbb{R}^n$, any complementary subspace W must have dimension $n - k$. A common approach is to extend a basis for V to a basis for $\mathbb{R}^n$; the span of the added vectors gives W.

Exam Focus

- For **sum** questions, start from the definition: take a generic element $v_1 + \cdots + v_n$ and simplify the description.
- For **direct sum** questions, the quickest test is often Theorem 1.8:
  - to *disprove* directness, exhibit a non-trivial decomposition $0 = v_1 + \cdots + v_n$;
  - to *prove* directness, assume $v_1 + \cdots + v_n = 0$ and force each $v_i = 0$.

Theorem 1.10: Three-Concepts-in-One Theorem (Lecture Week 2)

Let V be a vector space and let S be a linearly independent subset of V. Let $\{A, B\}$ be a partition of S (i.e. $A \cup B = S$ and $A \cap B = \emptyset$).
Then $\text{Span}(S) = \text{Span}(A) \oplus \text{Span}(B)$.

Proof. By Theorem 1.4(c), $\text{Span}(S) = \text{Span}(A \cup B) = \text{Span}(A) + \text{Span}(B)$.
It remains to show $\text{Span}(A) \cap \text{Span}(B) = \{0\}$. Let $v \in \text{Span}(A) \cap \text{Span}(B)$. Write

$$v = \sum_{i=1}^{m} \lambda_i a_i = \sum_{j=1}^{n} \mu_j b_j$$

with distinct $a_i \in A$ and distinct $b_j \in B$. Then

$$0 = \sum_{i=1}^{m} \lambda_i a_i + \sum_{j=1}^{n} (-\mu_j) b_j.$$

Since $A \cap B = \emptyset$, the vectors $a_1, \ldots, a_m, b_1, \ldots, b_n$ are all distinct elements of S. By linear independence of S, all coefficients are zero: $\lambda_i = 0$ and $\mu_j = 0$. Hence $v = 0$.
By the Direct-Sum Criterion (two subspaces), $\text{Span}(S) = \text{Span}(A) \oplus \text{Span}(B)$. □

Policy Application

Why the name "Three-Concepts-in-One"? This single result connects three core notions:

- *Linear independence* (the hypothesis on S),
- *Span* (the building blocks Span(A) and Span(B)),
- *Direct sum* (the conclusion).

It is especially useful for constructing direct-sum decompositions in practice: take a linearly independent set that spans your space, partition it, and the spans of the parts give a direct sum.

QRS@NTU — qrsntu.org

2 Finite-Dimensional Vector Spaces

This chapter corresponds to Study Weeks 3–4 (ILOs 2–3). The focus is on *finite-dimensional* structure: linear independence, bases, coordinates, and dimension arguments.

2.1 Span and Generating Sets

Throughout, let V be a vector space over a field $\mathbb{F}$ (typically $\mathbb{R}$ or $\mathbb{C}$).

Definition 2.1: Linear combination and span

Let $S \subseteq V$.

- A **linear combination** of vectors in S means an expression

$$\lambda_1 v_1 + \cdots + \lambda_n v_n$$

  where $n \in \mathbb{N}$, $v_1, \ldots, v_n \in S$ are *distinct*, and $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$.

- The **span** of S is

$$\operatorname{Span}(S) \;=\; \{\lambda_1 v_1 + \cdots + \lambda_n v_n \,|\, n \in \mathbb{N},\; v_1, \ldots, v_n \in S \text{ distinct},\; \lambda_i \in \mathbb{F}\}\,.$$

- We say S **generates** (or **spans**) V if $\operatorname{Span}(S) = V$.

Theorem 2.1: Span is the smallest subspace containing a set

Let $S \subseteq V$.

1. $\operatorname{Span}(S)$ is a vector subspace of V and contains S.
2. If W is a subspace of V and $S \subseteq W$, then $\operatorname{Span}(S) \subseteq W$.

Hence $\operatorname{Span}(S)$ is the *smallest* subspace of V containing S.

Proof. (1) By definition, $0 \in \operatorname{Span}(S)$ (take all coefficients 0), and the set is closed under addition and scalar multiplication because sums/scalar multiples of finite linear combinations are again finite linear combinations of elements of S. Also $S \subseteq \operatorname{Span}(S)$ since each $s \in S$ equals $1 \cdot s$.
(2) If $S \subseteq W$ and W is a subspace, then every linear combination of vectors in S lies in W by closure. Therefore $\operatorname{Span}(S) \subseteq W$.

□

Common Mistake

- Confusing S with $\operatorname{Span}(S)$: the span is usually *much larger* than the set.
- Forgetting that $\operatorname{Span}(\varnothing) = \{0\}$ (the empty linear combination is 0).

2.2 Linear Independence and Dependence

Definition 2.2: Linear independence

Let $S \subseteq V$ (not necessarily finite).

- S is **linearly independent** if whenever

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$$

for some $n \in \mathbb{N}$, distinct $v_1, \ldots, v_n \in S$, and scalars $\lambda_i \in \mathbb{F}$, then

$$\lambda_1 = \cdots = \lambda_n = 0.$$

- Otherwise, S is **linearly dependent**.

Intuitively, a set S is linearly independent when *no vector in S is redundant*: you cannot "build" any element of S from the others. Equivalently, the only way to combine vectors from S to get 0 is by using all-zero coefficients—there is no "shortcut" or "hidden relation." The negation (linear dependence) means some vector in S *can* be expressed using the others; this is made precise in Theorem 2.2.

Policy Application

Useful quick facts about linear independence.

- *Any subset of a linearly independent set is linearly independent.* (If a finite sub-selection could give a nontrivial relation, that same relation would work in the larger set.)
- *Any superset of a linearly dependent set is linearly dependent.* (A nontrivial relation among a subset extends to the whole set by assigning coefficient 0 to the extra vectors.)
- $\varnothing$ is linearly independent (vacuously: there are no finite selections to test).
- $\{v\}$ is linearly independent if and only if $v \neq 0$.
- $\{u, v\}$ is linearly dependent if and only if one is a scalar multiple of the other.

Key Formula

Finite-set test (most used in computations).
For distinct $v_1, \ldots, v_n \in V$,

$$\{v_1, \ldots, v_n\} \text{ is linearly independent} \iff \lambda_1 v_1 + \cdots + \lambda_n v_n = 0 \Rightarrow \lambda_1 = \cdots = \lambda_n = 0.$$

In $\mathbb{F}^m$, place $v_1, \ldots, v_n$ as columns of a matrix A:

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

Then $\{v_1, \ldots, v_n\}$ is linearly independent $\iff$ the homogeneous system $A\lambda = 0$ has only the trivial solution $\iff$ every column is a pivot column in $\operatorname{rref}(A)$.

Exam Focus

Quick dimension sanity check for independence. In $\mathbb{F}^m$, any set of more than m vectors is automatically linearly dependent (you have more unknowns than equations in the homogeneous system, so a nontrivial solution must exist). In particular, in $\mathbb{R}^{2,2} \cong \mathbb{R}^4$, at most 4 matrices can be independent; in $P_n(\mathbb{F})$, at most $n+1$ polynomials can be independent.

Theorem 2.2: Dependence iff one vector lies in the span of the others (Tutorial 3, Question 1)

Let $S \subseteq V$. The following are equivalent:

1. S is linearly dependent.
2. There exist $n \in \mathbb{N} \cup \{0\}$ and $n+1$ distinct vectors $v, w_1, \ldots, w_n \in S$ such that

$$v \in \text{Span}(\{w_1, \ldots, w_n\}).$$

(For $n = 0$, this means $v \in \text{Span}(\varnothing) = \{0\}$, i.e. $v = 0$.)

Proof. (1) $\Rightarrow$ (2): If S is dependent, then there exist distinct $s_0, \ldots, s_m \in S$ and scalars $\alpha_0, \ldots, \alpha_m$, not all zero, such that

$$\alpha_0 s_0 + \cdots + \alpha_m s_m = 0.$$

Choose k with $\alpha_k \neq 0$ and solve for s_k:

$$s_k = -\sum_{i \neq k} \frac{\alpha_i}{\alpha_k} s_i \in \text{Span}(\{s_0, \ldots, s_m\} \setminus \{s_k\}).$$

Set $v = s_k$ and list the remaining vectors as $w_1, \ldots, w_n$.
(2) $\Rightarrow$ (1): If $v = \beta_1 w_1 + \cdots + \beta_n w_n$, then

$$v - \beta_1 w_1 - \cdots - \beta_n w_n = 0$$

is a nontrivial linear dependence relation among distinct vectors from S (the coefficient of v is $1 \neq 0$). Hence S is dependent.

□

Policy Application

Refinement via minimal dependent subsets (Tutorial 3, Question 1, Method 2). One can strengthen the (1) $\Rightarrow$ (2) direction by choosing a *minimal* finite dependent subset $T \subseteq S$ (one of smallest cardinality). In such a minimal set, *every* coefficient in the dependence relation is nonzero (if some $\alpha_k = 0$, the smaller set $T \setminus \{s_k\}$ would still be dependent, contradicting minimality). This is useful in proofs where you need to remove a *specific* vector (e.g. the one with the largest index, or one satisfying an extra condition) and be certain its coefficient is nonzero.

Worked Example

Worked Example (Week 3 lecture): linear independence in $\mathbb{R}^{2,2}$.
Determine whether

$$M_1 = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix},\ M_2 = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix},\ M_3 = \begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix},\ M_4 = \begin{pmatrix} 2 & 1 \\ 2 & -2 \end{pmatrix}$$

is a linearly independent subset of $\mathbb{R}^{2,2}$.

Solution.
Consider

$$c_1M_1 + c_2M_2 + c_3M_3 + c_4M_4 = 0.$$

Equating entries gives a homogeneous system in (c_1, c_2, c_3, c_4):

$$\begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & -1 & 2 & 1 \\ -2 & 1 & 1 & 2 \\ 1 & 1 & 0 & -2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Row reduction yields rref $= I_4$, so the only solution is $c_1 = c_2 = c_3 = c_4 = 0$.
Hence $\{M_1, M_2, M_3, M_4\}$ is linearly independent.

Policy Application

Strategy for non-$\mathbb{R}^n$ spaces (matrices, polynomials, functions). The worked example above illustrates the general recipe:

1. *Flatten* the vectors into coordinate vectors (read off entries of matrices, or coefficients of polynomials).
2. *Form a matrix* with those coordinate vectors as columns.
3. *Row reduce* to determine whether every column is a pivot column.

For function spaces (e.g. $\{e^x, e^{-x}\}$ in $\mathbb{R}^{\mathbb{R}}$), there is no natural finite coordinate system, so instead you *evaluate at specific points* to generate equations, as in the next example.

Worked Example

Worked Example (Week 3 lecture): linear independence of functions.
Let $f, g : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = e^x$ and $g(x) = e^{-x}$. Show that $\{f, g\}$ is linearly independent.

Solution.
Assume $\alpha f + \beta g = 0$ as functions. Then for all $x \in \mathbb{R}$,

$$\alpha e^x + \beta e^{-x} = 0.$$

Multiply by e^x:

$$\alpha e^{2x} + \beta = 0 \quad \text{for all } x.$$

Setting $x = 0$ gives $\alpha + \beta = 0$, and setting $x = 1$ gives $\alpha e^2 + \beta = 0$.
Subtracting the two equations yields $\alpha(e^2 - 1) = 0$, so $\alpha = 0$, hence $\beta = 0$.
Therefore $\{f, g\}$ is linearly independent.

Worked Example

Worked Example (Week 3 lecture): independence of differences.
Let v_1, v_2, v_3, v_4 be distinct vectors such that $\{v_1, v_2, v_3, v_4\}$ is linearly independent. Show that $\{v_1 - v_2,\ v_2 - v_3,\ v_3 - v_4,\ v_4\}$ is linearly independent.

Solution. Suppose $\alpha_1(v_1 - v_2) + \alpha_2(v_2 - v_3) + \alpha_3(v_3 - v_4) + \alpha_4 v_4 = 0$. Expanding:

$$\alpha_1 v_1 + (\alpha_2 - \alpha_1)v_2 + (\alpha_3 - \alpha_2)v_3 + (\alpha_4 - \alpha_3)v_4 = 0.$$

By independence of $\{v_1, v_2, v_3, v_4\}$, all coefficients are zero:

$$\alpha_1 = 0, \quad \alpha_2 - \alpha_1 = 0, \quad \alpha_3 - \alpha_2 = 0, \quad \alpha_4 - \alpha_3 = 0.$$

This gives $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$. Hence $\{v_1 - v_2,\ v_2 - v_3,\ v_3 - v_4,\ v_4\}$ is linearly independent.

Key idea: express the new vectors in terms of the old basis and check that the "change-of-basis" matrix is invertible. Here the matrix is lower-triangular with 1's on the diagonal, hence invertible.

Common Mistake

- Using the finite-set definition incorrectly for infinite sets: linear dependence is witnessed by a *finite* nontrivial linear combination.
- Forgetting "distinct vectors" in the definition: repeated vectors can hide dependence in notation.

2.3 Bases and Coordinate Representations

Definition 2.3: Basis

A **basis** for V is a subset $B \subseteq V$ such that

1. B is linearly independent;
2. $\mathrm{Span}(B) = V$.

If B is finite, we often write it as an ordered list $(b_1, \ldots, b_n)$ when computing coordinates.

A basis is a "coordinate system" for V: it provides a minimal set of building blocks from which every vector can be uniquely assembled. Condition (1) (independence) ensures no building block is redundant; condition (2) (spanning) ensures every vector is reachable. Together they give the key consequence: every $v \in V$ has a *unique* representation as a linear combination of basis vectors.

Key Formula

Standard bases (know these by heart).

- $\mathbb{F}^n$: the standard basis is $\{e_1, \ldots, e_n\}$ where e_i has a 1 in position i and 0's elsewhere.
- $P_n(\mathbb{F})$: the monomial basis is $\{1, X, X^2, \ldots, X^n\}$; note $\dim(P_n(\mathbb{F})) = n + 1$.
- $\mathbb{F}^{m,n}$: the matrix-unit basis is $\{E_{ij} \mid 1 \leq i \leq m,\ 1 \leq j \leq n\}$; note $\dim(\mathbb{F}^{m,n}) = mn$.
- $P(\mathbb{F})$ (all polynomials): the infinite monomial basis $\{1, X, X^2, \ldots\}$; $\dim(P(\mathbb{F})) = \infty$.

Theorem 2.3: Unique linear combination in a basis

Let B be a basis of V. Then each $v \in V$ can be written as a linear combination of vectors of B in *exactly one* way (as a finite sum):

$$v = \sum_{b \in B} \lambda_b \, b,$$

where only finitely many $\lambda_b \in \mathbb{F}$ are nonzero.

Proof. Existence: since $\mathrm{Span}(B) = V$, every v is a finite linear combination of elements of B.
Uniqueness: if also $v = \sum_{b \in B} \mu_b \, b$ (finite sums), subtracting gives

$$\sum_{b \in B} (\lambda_b - \mu_b) b = 0$$

as a finite linear combination of distinct vectors in B. Linear independence forces $\lambda_b - \mu_b = 0$ for all b, so $\lambda_b = \mu_b$.

□

Theorem 2.4: Basis Existence Theorem (Lecture Week 3)

Every vector space has a basis.

This theorem is equivalent to the Axiom of Choice and is a pure existence result. In practice, for the spaces encountered in MH1201 ($\mathbb{F}^n$, $P_n(\mathbb{F})$, $\mathbb{F}^{m,n}$, etc.), explicit bases are readily constructed. There exist vector spaces (e.g. $\mathbb{R}^{\mathbb{R}}$) for which no explicit basis can be written down, even though one must exist.

Definition 2.4: Coordinate vector (finite basis)

Let V be a vector space with an *ordered* basis $\mathcal{B} = (b_1, \ldots, b_n)$.
For $v \in V$, write uniquely

$$v = \lambda_1 b_1 + \cdots + \lambda_n b_n.$$

The column vector

$$[v]_{\mathcal{B}} \overset{\text{def}}{=} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{F}^{n,1}$$

is called the **coordinate vector** of v with respect to $\mathcal{B}$.

Common Mistake

Ordering matters for coordinate vectors. The coordinate vector depends on the *order* of the basis vectors. If $\mathcal{B} = (b_1, b_2)$ and $\mathcal{B}' = (b_2, b_1)$, and $v = 3b_1 + 5b_2$, then

$$[v]_{\mathcal{B}} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \qquad [v]_{\mathcal{B}'} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}.$$

Always check which ordering the problem specifies.

QRS@NTU

Worked Example

Worked Example: compute coordinates in a nonstandard basis.
Let $\mathcal{B} = (b_1, b_2, b_3)$ be a basis of $\mathbb{R}^3$ where

$$b_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}, \quad b_2 = \begin{pmatrix}0\\1\\1\end{pmatrix}, \quad b_3 = \begin{pmatrix}1\\1\\0\end{pmatrix}.$$

Find $[v]_{\mathcal{B}}$ for $v = \begin{pmatrix}2\\1\\3\end{pmatrix}$.

Solution.
Solve $v = \lambda_1 b_1 + \lambda_2 b_2 + \lambda_3 b_3$:

$$\begin{pmatrix}2\\1\\3\end{pmatrix} = \lambda_1 \begin{pmatrix}1\\0\\1\end{pmatrix} + \lambda_2 \begin{pmatrix}0\\1\\1\end{pmatrix} + \lambda_3 \begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}\lambda_1 + \lambda_3\\ \lambda_2 + \lambda_3\\ \lambda_1 + \lambda_2\end{pmatrix}.$$

Thus

$$\lambda_1 + \lambda_3 = 2, \quad \lambda_2 + \lambda_3 = 1, \quad \lambda_1 + \lambda_2 = 3.$$

Subtract the first two equations from the third:

$$(\lambda_1 + \lambda_2) - (\lambda_1 + \lambda_3) - (\lambda_2 + \lambda_3) = 3 - 2 - 1 \Rightarrow -2\lambda_3 = 0 \Rightarrow \lambda_3 = 0.$$

Then $\lambda_1 = 2$ and $\lambda_2 = 1$. Hence

$$[v]_{\mathcal{B}} = \begin{pmatrix}2\\1\\0\end{pmatrix}.$$

Policy Application

General method for finding $[v]_{\mathcal{B}}$ in $\mathbb{R}^n$. Form the augmented matrix $[b_1\ b_2\ \cdots\ b_n \mid v]$ and row reduce. The solution gives the coordinates $\lambda_1, \ldots, \lambda_n$. Equivalently, if P is the matrix with columns $b_1, \ldots, b_n$, then $[v]_{\mathcal{B}} = P^{-1}v$.

Exam Focus

When asked to show a set B is a basis, do *not* argue informally. Use one of:

- show **spanning** + **independence**; or
- in finite dimensions, show $|B| = \dim(V)$ and either spanning or independence (then the other follows by theorem).

2.4 Dimension and Core Theorems in Finite Dimensions

Theorem 2.5: Size-Limitation Lemma

Let V be a vector space that has a *finite* basis of size n.
Then every linearly independent subset of V is finite and has size $\leq n$.

Proof. Let $B = \{b_1, \ldots, b_n\}$ be a basis of V. Suppose for contradiction that V has a linearly independent set containing $n+1$ distinct vectors $v_1, \ldots, v_{n+1}$.
Each v_i is a linear combination of $b_1, \ldots, b_n$, so there exist scalars $\lambda_{i,1}, \ldots, \lambda_{i,n} \in \mathbb{F}$ with

$$v_i = \lambda_{i,1} b_1 + \cdots + \lambda_{i,n} b_n \qquad (i = 1, \ldots, n+1).$$

For scalars $\mu_1, \ldots, \mu_{n+1}$, we have

$$\mu_1 v_1 + \cdots + \mu_{n+1} v_{n+1} = \sum_{j=1}^{n} (\lambda_{1,j}\mu_1 + \cdots + \lambda_{n+1,j}\mu_{n+1})\, b_j.$$

Thus $\mu_1 v_1 + \cdots + \mu_{n+1} v_{n+1} = 0$ holds whenever

$$\lambda_{1,j}\mu_1 + \cdots + \lambda_{n+1,j}\mu_{n+1} = 0 \qquad (j = 1, \ldots, n).$$

This is a homogeneous system of n equations in $n+1$ unknowns, so it has a nontrivial solution $(\mu_1, \ldots, \mu_{n+1}) \neq 0$.
Hence there is a nontrivial linear combination of $v_1, \ldots, v_{n+1}$ equal to 0, contradicting linear independence.

□

Policy Application

Why does the Size-Limitation Lemma work? The key idea is a *counting argument*: $n+1$ vectors, each expressed in terms of n basis vectors, give an under-determined homogeneous system (n equations, $n+1$ unknowns). Under-determined homogeneous systems always have nontrivial solutions (this is a fundamental fact from MH1200 / linear systems theory). This lemma is the engine behind almost every dimension result in the course.

Theorem 2.6: Fundamental Theorem of MH1201 (basis size is well-defined for finite bases)

If V has a finite basis of size n, then every basis of V is finite and has size n.

Proof. Let B be a basis with $|B| = n$, and let $\widetilde{B}$ be any other basis.
Since $\widetilde{B}$ is linearly independent, the Size-Limitation Lemma gives $|\widetilde{B}| \leq n$.
Applying the same lemma with $\widetilde{B}$ as the finite basis yields $n \leq |\widetilde{B}|$.
Therefore $|\widetilde{B}| = n$.

□

Policy Application

Why "Fundamental"? This theorem guarantees that dimension is an intrinsic property of the vector space, not an artefact of a particular basis choice. Without it, we could not define $\dim(V)$ unambiguously. The proof uses the Size-Limitation Lemma *twice*, with the two bases swapping roles—this "ping-pong" argument is a standard proof technique.

Definition 2.5: Dimension and finite-dimensionality

Let V be a vector space. If V has a finite basis, define

$$\dim(V) \stackrel{\text{def}}{=} \text{the number of vectors in any (hence every) basis of } V.$$

Then V is **finite-dimensional** iff $\dim(V)$ is finite.

Key Formula

Quick reference: dimensions of common spaces.

Space	Dimension
$\mathbb{F}^n$	n
$P_n(\mathbb{F})$	$n+1$
$\mathbb{F}^{m,n}$	mn
$\{M \in \mathbb{F}^{n,n} \mid \text{trace}(M) = 0\}$	$n^2 - 1$
$\{M \in \mathbb{R}^{n,n} \mid M^T = M\}$ (symmetric)	$n(n+1)/2$
$\{M \in \mathbb{R}^{n,n} \mid M^T = -M\}$ (skew-symmetric)	$n(n-1)/2$
$\{M \in \mathbb{R}^{n,n} \mid M$ upper triangular$\}$	$n(n+1)/2$

Key Formula

Equivalences in finite dimensions.
Let $\dim(V) = n$ and let $S \subseteq V$ be a set with exactly n vectors. Then the following are equivalent:

$$S \text{ is a basis} \iff S \text{ spans } V \iff S \text{ is linearly independent.}$$

Also:

- any linearly independent set has size $\leq n$;
- any spanning set has size $\geq n$.

Policy Application

Why is this equivalence so useful in practice? It halves your work. If you know $|S| = \dim(V)$, you only need to prove *one* of the two conditions (independence or spanning); the other is then automatic. This is the go-to shortcut on exams: count vectors, compare with the known dimension, and check whichever condition is easier.

Extension (Tutorial 3, Question 2, Method 2)

Let $S = \{s_1, \ldots, s_n\} \subseteq V$ and define $T : \mathbb{F}^n \to V$ by
$T(\alpha_1, \ldots, \alpha_n) = \sum_{i=1}^{n} \alpha_i s_i$.
Then $\mathrm{Im}(T) = \mathrm{Span}(S)$ and $\ker(T)$ is the space of linear relations among the s_i.
If $n = \dim(V)$ and S is linearly independent, then $\ker(T) = \{0\}$ so T is injective; by rank–nullity, $\dim(\mathrm{Im}(T)) = n = \dim(V)$, hence $\mathrm{Im}(T) = V$ and S spans V.

Worked Example

Worked Example (Week 3 lecture): basis for a trace-zero subspace.
Let

$$W = \left\{ M \in \mathbb{R}^{2,2} \;\middle|\; \mathrm{trace}(M) = 0 \right\}.$$

Find a basis of W and compute $\dim(W)$.

Solution.
Write $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. The constraint $\mathrm{trace}(M) = a + d = 0$ implies $d = -a$, so

$$M = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

The three displayed matrices are linearly independent (distinct support patterns in entries), hence they form a basis.
Therefore $\dim(W) = 3$.

Policy Application

Kernel viewpoint for the trace-zero subspace. The trace map $\mathrm{tr} : \mathbb{R}^{2,2} \to \mathbb{R}$ is a linear functional (a linear map to the scalar field). Since $\mathrm{tr} \neq 0$ (e.g. $\mathrm{tr}(I) = 2$), it is surjective, so $\mathrm{rank}(\mathrm{tr}) = 1$. By rank–nullity, $\dim(\ker(\mathrm{tr})) = \dim(\mathbb{R}^{2,2}) - 1 = 4 - 1 = 3$. This gives $\dim(W) = 3$ without explicitly finding a basis—useful as a sanity check.

2.5 Adding Vectors, Extending Bases, and Extracting Bases

Theorem 2.7: Add-One-More Lemma

Let $S \subseteq V$ be linearly independent and let $v \in V$.
Then the following are equivalent:

1. $v \notin S$ and $S \cup \{v\}$ is linearly independent.
2. $v \notin \text{Span}(S)$.

Proof. (1) $\Rightarrow$ (2): If $v \in \text{Span}(S)$, then $v = \sum_{i=1}^{n} \lambda_i s_i$ for distinct $s_i \in S$, so

$$1 \cdot v - \sum_{i=1}^{n} \lambda_i s_i = 0$$

is a nontrivial dependence among distinct vectors in $S \cup \{v\}$, contradiction.
(2) $\Rightarrow$ (1): If $v \notin \text{Span}(S)$, then $v \notin S$ since $S \subseteq \text{Span}(S)$. If a nontrivial linear combination of distinct vectors in $S \cup \{v\}$ equals 0, the coefficient of v must be nonzero (otherwise we contradict independence of S), and then we can solve for v as a linear combination of vectors in S, contradicting $v \notin \text{Span}(S)$.
□

Policy Application

The Add-One-More Lemma in action. This lemma is the workhorse behind *both* the Basis Extension Theorem (keep adding vectors outside the span until you reach a basis) and the hereditary dimension theorem (a maximal independent set in a subspace must span it, because otherwise we could add one more). It encodes the precise boundary between "can still grow" and "already spans."

Theorem 2.8: Finite-dimensionality is hereditary

Let V be finite-dimensional and let W be a subspace of V. Then W is finite-dimensional.

Proof. Let $n = \dim(V)$. Any linearly independent subset of W is also linearly independent in V, hence has size $\leq n$ by the Size-Limitation Lemma. Take a linearly independent subset $B \subseteq W$ of maximal size (existence follows since sizes are bounded by n). We claim $\text{Span}(B) = W$. If not, pick $v \in W \setminus \text{Span}(B)$; by the Add-One-More Lemma, $B \cup \{v\}$ is linearly independent in W, contradicting maximality. Hence B is a finite basis of W.
□

Key Formula

Dimension inequality for subspaces. If W is a subspace of a finite-dimensional space V, then:

$$\dim(W) \leq \dim(V),$$

with equality if and only if $W = V$.
Proof of equality case: if $\dim(W) = \dim(V) = n$, any basis of W is an independent set of n vectors in V, hence a basis of V by the equivalence theorem, so $W = V$.

Theorem 2.9: Basis Extension Theorem

Let V be finite-dimensional and let $S \subseteq V$ be linearly independent. Then S can be extended to a basis of V.

Proof. Let $n = \dim(V)$. Among all linearly independent subsets of V that contain S, pick one of maximal size, call it B (sizes are bounded by n). If $\text{Span}(B) \neq V$, choose $v \in V \setminus \text{Span}(B)$; then $B \cup \{v\}$ is linearly independent by the Add-One-More Lemma, contradicting maximality. Hence $\text{Span}(B) = V$, so B is a basis extending S.

□

Policy Application

Practical algorithm for basis extension (Lecture Week 4). To extend an independent set $S = \{v_1, \ldots, v_m\}$ to a basis of V with $\dim(V) = n$:

1. Append a known basis $\{b_1, \ldots, b_n\}$ of V to form the list $v_1, \ldots, v_m, b_1, \ldots, b_n$.
2. Convert to coordinate vectors (columns in $\mathbb{F}^n$) if necessary.
3. Row reduce the combined matrix. Identify which of the appended basis vectors b_j correspond to non-pivot columns—these are the *redundant* ones.
4. Remove the redundant b_j's. The remaining list $v_1, \ldots, v_m, b_{i_1}, \ldots, b_{i_{n-m}}$ is a basis of V.

As a bonus, $V = \text{Span}(S) \oplus \text{Span}\{b_{i_1}, \ldots, b_{i_{n-m}}\}$, giving a direct-sum complement "for free."

Theorem 2.10: Spanning sets contain a basis (Tutorial 3, Question 3)

Let V be finite-dimensional and let $S \subseteq V$ be a spanning set: $\text{Span}(S) = V$.
Then S contains a basis of V.

Proof. Start with a finite spanning subset of S (in finite dimensions, spanning implies there is a finite spanning subset).
List it as $(u_1, \ldots, u_m)$.
If $\{u_1, \ldots, u_m\}$ is linearly independent, it is already a basis.
Otherwise it is dependent; by Theorem 2.2, one of the u_k lies in the span of the others, so removing u_k does not change the span.
Repeat until the remaining list is linearly independent. The process terminates because the list size strictly decreases.
The remaining set is linearly independent and still spans V, hence is a basis contained in S.

□

Key Formula

Two directions for building a basis.

- *Bottom-up (extension):* start with an independent set, keep adding vectors outside the span. Terminates when the set spans V.
- *Top-down (extraction):* start with a spanning set, keep removing vectors that lie in the span of the remaining ones. Terminates when the set is independent.

Both produce a basis. The choice depends on what you start with.

Exam Focus

- **To prove a set is linearly independent:** start from $\sum \alpha_i v_i = 0$ and force $\alpha_i = 0$.
- **To prove a set spans:** take an arbitrary v and explicitly solve for coefficients.
- **Fast finite-dimensional shortcut:** if you already know $\dim(V) = n$, then for n vectors, independence $\Longleftrightarrow$ spanning $\Longleftrightarrow$ basis.

2.6 Dimension Computations and Dimension Formulae

Worked Example

Worked Example (Week 4 lecture): dimension of a plane in $\mathbb{R}^3$.
Let

$$\Pi = \left\{(x, y, z) \in \mathbb{R}^3 \;\middle|\; x + 2y + 3z = 0\right\}.$$

Find a basis and $\dim(\Pi)$.

Solution.
Solve $x = -2y - 3z$ and parametrize by $y = s$, $z = t$:

$$(x, y, z) = (-2s - 3t,\ s,\ t) = s(-2, 1, 0) + t(-3, 0, 1).$$

Thus

$$\Pi = \mathrm{Span}\{(-2, 1, 0),\ (-3, 0, 1)\}.$$

The two spanning vectors are linearly independent, hence form a basis. Therefore $\dim(\Pi) = 2$.

Policy Application

General recipe for constraint-defined subspaces. If $W = \{v \in V \mid \text{linear conditions on } v\}$:

1. Write out the constraints as linear equations in the coordinates/entries.
2. Solve the system: express the constrained variables in terms of the free variables.
3. Parametrize: write a generic element of W as a linear combination of fixed vectors (one per free parameter).
4. The fixed vectors form a spanning set for W; check independence (usually immediate from the parametrisation).

The number of free parameters equals $\dim(W)$. As a sanity check: each *independent* linear constraint reduces the dimension by 1 (when the constraints are independent), so $\dim(W) \geq \dim(V) -$ (number of constraints).

Worked Example

Worked Example (Week 4 lecture): dimension of upper-triangular matrices.
Let $U_n = \{M \in \mathbb{R}^{n,n} \mid M \text{ is upper triangular}\}$.
Find $\dim(U_n)$.

Solution.
A matrix in U_n has free entries in positions (i, j) with $1 \leq i \leq j \leq n$.
There are

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

such positions. A natural basis is $\{E_{ij} \mid 1 \leq i \leq j \leq n\}$ (matrix units), so

$$\dim(U_n) = \frac{n(n+1)}{2}.$$

Theorem 2.11: Dimension of a sum

Let V be a finite-dimensional vector space and let V_1, V_2 be subspaces of V.
Then $V_1 + V_2$ and $V_1 \cap V_2$ are finite-dimensional and

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2).$$

Proof. Let A be a basis of $V_1 \cap V_2$.
Extend A to a basis of V_1 by adding a set B (so $A \cup B$ is a basis of V_1),
and extend A to a basis of V_2 by adding a set C (so $A \cup C$ is a basis of V_2).
One shows that $A \cup B \cup C$ is linearly independent and spans $V_1 + V_2$, hence is a basis of $V_1 + V_2$.
Therefore

$$\dim(V_1 + V_2) = |A| + |B| + |C| = \big(|A| + |B|\big) + \big(|A| + |C|\big) - |A| = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2).$$

□

Policy Application

Why is $A \cup B \cup C$ linearly independent? (Lecture Week 4, full detail).
Suppose $\sum_{a\in A} \lambda_a\, a + \sum_{b\in B} \mu_b\, b + \sum_{c\in C} \kappa_c\, c = 0$. Then $\sum_{b\in B} \mu_b\, b = -\sum_{a\in A} \lambda_a\, a - \sum_{c\in C} \kappa_c\, c$. The left side lies in V_1 (since $B \subseteq V_1$), and the right side can be rewritten as $\sum(-\lambda_a)\, a + \sum(-\kappa_c)\, c \in V_2$ (since $A \subseteq V_1 \cap V_2 \subseteq V_2$ and $C \subseteq V_2$).
Hence $\sum \mu_b\, b \in V_1 \cap V_2 = \mathrm{Span}(A)$. But $\sum \mu_b\, b \in \mathrm{Span}(B)$, and $\mathrm{Span}(A) \cap \mathrm{Span}(B) = \{0\}$ (since $A \cup B$ is a basis of V_1, and by the Three-Concepts-in-One Theorem from Chapter 1). So $\sum \mu_b\, b = 0$, and by independence of B, all $\mu_b = 0$.
Similarly, $\sum \kappa_c\, c \in V_1 \cap V_2 = \mathrm{Span}(A)$, and $\mathrm{Span}(A) \cap \mathrm{Span}(C) = \{0\}$, so all $\kappa_c = 0$.
Substituting back: $\sum \lambda_a\, a = 0$, so all $\lambda_a = 0$ by independence of A. Hence $A \cup B \cup C$ is linearly independent.

Policy Application

Analogy with inclusion-exclusion (Lecture Week 4). The dimension formula $\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2)$ is the vector-space analogue of the set-theoretic inclusion-exclusion formula $|A \cup B| = |A| + |B| - |A \cap B|$.
Caution (also from Lecture Week 4): this analogy does *not* extend to three subspaces. The "three-subspace inclusion-exclusion" formula $\dim(V_1 + V_2 + V_3) = \dim(V_1) + \dim(V_2) + \dim(V_3) - \dim(V_1 \cap V_2) - \dim(V_2 \cap V_3) - \dim(V_1 \cap V_3) + \dim(V_1 \cap V_2 \cap V_3)$ is *false* in general for vector subspaces.

Key Formula

Dimension identity and immediate inequalities.
For subspaces $U, V \subseteq \mathbb{F}^n$,

$$\dim(U \cap V) = \dim(U) + \dim(V) - \dim(U + V).$$

Since $\dim(U + V) \leq n$, we get the useful bound

$$\dim(U \cap V) \;\geq\; \dim(U) + \dim(V) - n.$$

This is often the fastest way to prove intersections are nontrivial.

Worked Example

Worked Example (Tutorial 3, Question 7): two 5D subspaces in $\mathbb{R}^9$ intersect nontrivially.
Let U and V be 5-dimensional subspaces of $\mathbb{R}^9$. Show $U \cap V \neq \{0\}$.

Solution.
By Theorem 2.11,

$$\dim(U \cap V) = \dim(U) + \dim(V) - \dim(U + V) \geq 5 + 5 - 9 = 1,$$

since $U + V \subseteq \mathbb{R}^9$ implies $\dim(U + V) \leq 9$.
Thus $\dim(U \cap V) \geq 1$, so $U \cap V$ contains a nonzero vector.

Extension (Tutorial 3, Question 7, Method 2)

If $U \cap V = \{0\}$, then $\mathbb{R}^9 \supseteq U + V = U \oplus V$, so

$$\dim(U + V) = \dim(U) + \dim(V) = 10.$$

But any subspace of $\mathbb{R}^9$ has dimension at most 9, contradiction.
Hence $U \cap V \neq \{0\}$.

Worked Example

Worked Example (Week 4 lecture): intersection of two planes in $\mathbb{R}^3$.
Let Π_1 and Π_2 be linear planes (i.e. 2-dimensional subspaces) in $\mathbb{R}^3$. What are the possible dimensions for $\Pi_1 \cap \Pi_2$?

Solution. By the dimension formula,

$$\dim(\Pi_1 \cap \Pi_2) = \dim(\Pi_1) + \dim(\Pi_2) - \dim(\Pi_1 + \Pi_2) = 4 - \dim(\Pi_1 + \Pi_2).$$

Since $\Pi_1 + \Pi_2$ is a subspace of $\mathbb{R}^3$, we have $2 \leq \dim(\Pi_1 + \Pi_2) \leq 3$ (it must contain Π_1, so dim ≥ 2).

- If $\dim(\Pi_1 + \Pi_2) = 3$: then $\dim(\Pi_1 \cap \Pi_2) = 1$ (a line through the origin).
- If $\dim(\Pi_1 + \Pi_2) = 2$: then $\dim(\Pi_1 \cap \Pi_2) = 2$, which forces $\Pi_1 = \Pi_2$ (same plane).

Both cases are realisable, so the possible dimensions are $\{1, 2\}$.

Definition 2.6: Direct sum (as used in Weeks 3–4)

Let U, W be subspaces of V.
We write $V = U \oplus W$ if

$$V = U + W \quad \text{and} \quad U \cap W = \{0\}.$$

In this case, each $v \in V$ has a *unique* decomposition $v = u + w$ with $u \in U$, $w \in W$.

Theorem 2.12: Dimension of a direct sum

If $V = U \oplus W$ with V finite-dimensional, then

$$\dim(V) = \dim(U) + \dim(W).$$

Proof. Since $U \cap W = \{0\}$, Theorem 2.11 gives

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W) = \dim(U) + \dim(W).$$

But $U + W = V$, so $\dim(V) = \dim(U) + \dim(W)$.

□

Key Formula

Checking a direct sum via dimensions (very common on exams). Let $U, W \leq V$ with V finite-dimensional. Then $V = U \oplus W$ if and only if *any two* of the following three conditions hold:

1. $V = U + W$;
2. $U \cap W = \{0\}$;
3. $\dim(U) + \dim(W) = \dim(V)$.

Why any two suffice: conditions (1) and (2) are the definition; if (1) and (3) hold, the dimension formula gives $\dim(U \cap W) = 0$, so (2) follows; if (2) and (3) hold, the dimension formula gives $\dim(U + W) = \dim(V)$, so (1) follows.
This shortcut often saves a full spanning argument: just verify (2) and (3).

Worked Example

Worked Example (Week 4 lecture): a direct-sum conclusion from dimensions.
Let $U, V \subseteq \mathbb{R}^8$ be subspaces such that $\dim(U) = 3$, $\dim(V) = 5$, and $\dim(U + V) = 8$.
Show that $\mathbb{R}^8 = U \oplus V$.

Solution.
By Theorem 2.11,

$$\dim(U \cap V) = \dim(U) + \dim(V) - \dim(U + V) = 3 + 5 - 8 = 0.$$

Hence $U \cap V = \{0\}$. Since $U + V = \mathbb{R}^8$ by the dimension assumption, we conclude $\mathbb{R}^8 = U \oplus V$.

Worked Example

Worked Example (Week 4 lecture): basis, extension, and direct-sum complement for $\{P \in P_4(\mathbb{R}) \mid P(6) = 0\}$.
Define $V = \{P \in P_4(\mathbb{R}) \mid P(6) = 0\}$.

(a) Basis and dimension. The evaluation map $\mathrm{ev}_6 : P_4(\mathbb{R}) \to \mathbb{R}$ given by $\mathrm{ev}_6(P) = P(6)$ is a nonzero linear functional, so $\dim(\ker(\mathrm{ev}_6)) = 5 - 1 = 4$. Every polynomial vanishing at 6 has $(t-6)$ as a factor (up to degree consideration): $P(t) = (t-6)Q(t)$ with $\deg(Q) \leq 3$. A basis is $\{t-6,\ t(t-6),\ t^2(t-6),\ t^3(t-6)\}$, confirming $\dim(V) = 4$.

(b) Extension and complement. Choose any $p \notin V$, e.g. $p = 1$ since $1(6) = 6 \neq 0$. Then $\{t-6,\ t(t-6),\ t^2(t-6),\ t^3(t-6),\ 1\}$ is a basis of $P_4(\mathbb{R})$, and $W = \mathrm{Span}\{1\}$ satisfies $P_4(\mathbb{R}) = V \oplus W$.

2.7 High-Yield Worked Problems from Tutorial 3

Worked Example

Worked Example (Tutorial 3, Question 4): commuting matrices form a subspace and its dimension.
Let

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \qquad \mathcal{V} = \{B \in \mathbb{R}^{2,2} \mid AB = BA\}.$$

Find $\dim(\mathcal{V})$.

Solution.
Write $B = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$.
Compute

$$AB = \begin{pmatrix} x+2z & y+2t \\ z & t \end{pmatrix}, \qquad BA = \begin{pmatrix} x & 2x+y \\ z & 2z+t \end{pmatrix}.$$

Equating entries $AB = BA$ gives

$$x + 2z = x \Rightarrow z = 0, \qquad y + 2t = 2x + y \Rightarrow t = x,$$

and the remaining equations are automatic once $z = 0$.
Thus

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

Therefore $\mathcal{V} = \text{Span}\left\{I, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right\}$ and $\dim(\mathcal{V}) = 2$.

Extension (Tutorial 3, Question 4, Method 2)

Write $A = I + 2N$ with $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $N^2 = 0$.
Since I commutes with everything, $AB = BA$ iff $NB = BN$.
For $B = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$,

$$NB = \begin{pmatrix} z & t \\ 0 & 0 \end{pmatrix}, \qquad BN = \begin{pmatrix} 0 & x \\ 0 & z \end{pmatrix}.$$

Thus $NB = BN$ forces $z = 0$ and $t = x$, giving $B = xI + yN$ and hence $\dim(\mathcal{V}) = 2$ with fewer computations.

Policy Application

Why the nilpotent decomposition helps (general principle). If $A = \alpha I + N$ where N is nilpotent, then $AB = BA \iff NB = BN$ because I commutes with everything. This reduces the problem to commuting with a simpler (nilpotent) matrix. For 2×2 matrices, the commutant of a non-scalar matrix always has dimension 2; the commutant of a scalar matrix ($A = \alpha I$) is all of $\mathbb{R}^{2,2}$, with dimension 4.

Worked Example

Worked Example (Tutorial 3, Question 5): a polynomial subspace, basis extension, and direct sum.
Let

$$V = \{P \in P_4(\mathbb{R}) \mid P(2) = P(5)\}.$$

1. Find a basis of V and $\dim(V)$.
2. Extend your basis to a basis of $P_4(\mathbb{R})$.
3. Find a subspace W such that $P_4(\mathbb{R}) = V \oplus W$.

Solution.
Let $q(t) = (t-2)(t-5)$ (degree 2). For any Q, $(qQ)(2) = (qQ)(5) = 0$, so $qQ \in V$.
Every $P \in P_4(\mathbb{R})$ can be written uniquely as

$$P(t) = q(t)Q(t) + L(t),$$

where $\deg(Q) \leq 2$ and $\deg(L) \leq 1$ (polynomial division by q). Write $L(t) = at + b$.
Then

$$P(2) = L(2) = 2a + b, \qquad P(5) = L(5) = 5a + b,$$

so the constraint $P(2) = P(5)$ forces $2a + b = 5a + b$, hence $a = 0$.
Thus L must be constant, and we obtain the description

$$V = \{q(t)Q(t) + c \mid \deg(Q) \leq 2,\ c \in \mathbb{R}\}.$$

A convenient basis is

$$\mathcal{B}_V = \{1,\ q(t),\ t\,q(t),\ t^2 q(t)\},$$

so $\dim(V) = 4$.
Since $\dim(P_4(\mathbb{R})) = 5$, we extend by adding any polynomial not in V; for instance $t \notin V$ because $t(2) \neq t(5)$.
Hence

$$\mathcal{B} = \{1,\ q(t),\ t\,q(t),\ t^2 q(t),\ t\}$$

is a basis of $P_4(\mathbb{R})$.
Let $W = \mathrm{Span}\{t\}$. Then $V + W = P_4(\mathbb{R})$ by the basis above.
Also $V \cap W = \{0\}$: if $ct \in V$, then $(ct)(2) = (ct)(5)$ implies $2c = 5c$, so $c = 0$.
Therefore $P_4(\mathbb{R}) = V \oplus W$.

Extension (Tutorial 3, Question 5, Method 2)

Define the linear functional $L : P_4(\mathbb{R}) \to \mathbb{R}$ by $L(P) = P(2) - P(5)$.
Then $V = \ker(L)$. Since $L \neq 0$, $\dim(\ker(L)) = \dim(P_4) - 1 = 4$.
To build a complement, choose any p with $L(p) \neq 0$; for example $p(t) = t$ gives $L(t) = 2 - 5 = -3$.
Then $P_4(\mathbb{R}) = \ker(L) \oplus \mathrm{Span}\{t\}$, with decomposition
$P = \left(P - \frac{L(P)}{L(t)}t\right) + \frac{L(P)}{L(t)}t$.

Exam Focus

General recipe for complementing a codimension-1 subspace. If $V = \ker(L)$ for a nonzero linear functional $L : W \to \mathbb{F}$, then:

- $\dim(V) = \dim(W) - 1$.
- Pick *any* $p \in W$ with $L(p) \neq 0$.
- Then $W = V \oplus \mathrm{Span}\{p\}$.
- The explicit decomposition is: $w = \left(w - \frac{L(w)}{L(p)}p\right) + \frac{L(w)}{L(p)}p$.

This pattern appears repeatedly in MH1201 (trace-zero matrices, evaluation constraints, etc.).

Worked Example

Worked Example (Week 4 lecture / dimension application): translate a large independent set.
Let $v_1, \ldots, v_{2024} \in V$ be linearly independent. Show that for any $w \in V$,

$$\dim\Big(\mathrm{Span}\{v_1 + w, \ldots, v_{2024} + w\}\Big) \geq 2023.$$

Solution.
Consider the 2023 vectors

$$(v_1+w)-(v_2+w) = v_1-v_2, \quad (v_2+w)-(v_3+w) = v_2-v_3, \quad \ldots, \quad (v_{2023}+w)-(v_{2024}+w) = v_{2023}-v_{2024}.$$

Each difference lies in $\mathrm{Span}\{v_1 + w, \ldots, v_{2024} + w\}$, so it suffices to show that

$$\{v_1 - v_2,\ v_2 - v_3,\ \ldots,\ v_{2023} - v_{2024}\}$$

is linearly independent.
Suppose $\sum_{i=1}^{2023} \alpha_i(v_i - v_{i+1}) = 0$. Expanding gives

$$\alpha_1 v_1 + (\alpha_2 - \alpha_1)v_2 + \cdots + (\alpha_{2023} - \alpha_{2022})v_{2023} - \alpha_{2023}v_{2024} = 0.$$

Since $\{v_1, \ldots, v_{2024}\}$ is linearly independent, all coefficients must be zero:

$$\alpha_1 = 0, \quad \alpha_2 - \alpha_1 = 0, \ \ldots, \ \alpha_{2023} - \alpha_{2022} = 0, \quad -\alpha_{2023} = 0.$$

Hence $\alpha_1 = \cdots = \alpha_{2023} = 0$, proving the differences are independent.
Therefore the span of $\{v_1 + w, \ldots, v_{2024} + w\}$ contains 2023 independent vectors, so its dimension is at least 2023.

Extension (Tutorial 3, Question 6, Method 2)

Stronger fact (and a structural proof).
If $\{v_1, \ldots, v_n\}$ is linearly independent and $w \notin \mathrm{Span}\{v_1, \ldots, v_n\}$, then $\{v_1 + w, \ldots, v_n + w\}$ is linearly independent.
Let $U = \mathrm{Span}\{v_1, \ldots, v_n, w\}$ and define $T : U \to U$ by $T(v_i) = v_i + w$ and $T(w) = w$, extended linearly. Then T is invertible with inverse $S(v_i) = v_i - w$, $S(w) = w$, so T is injective; injective linear maps send independent sets to independent sets, hence $\{v_i + w\}$ is independent.

Worked Example

Worked Example (Tutorial 3, Question 8): $\mathbb{R}$ as a $\mathbb{Q}$-vector space.
Let $\mathbb{R}_{\mathbb{Q}}$ denote $\mathbb{R}$ viewed as a vector space over $\mathbb{Q}$ (with usual addition and $\mathbb{Q}$-scalar multiplication). For a fixed $r \in \mathbb{R}$, define $V = \{a + br \mid a, b \in \mathbb{Q}\}$.

1. V is a $\mathbb{Q}$-subspace of $\mathbb{R}_{\mathbb{Q}}$: it is non-empty ($0 \in V$), closed under addition ($\mathbb{Q}$-sums of rationals are rational), and closed under $\mathbb{Q}$-scalar multiplication ($\mathbb{Q} \cdot \mathbb{Q} \subseteq \mathbb{Q}$).

2. If $r \in \mathbb{Q}$: then $V = \{a + br \mid a, b \in \mathbb{Q}\} = \mathbb{Q}$ (every element is rational), so $\dim_{\mathbb{Q}}(V) = 1$ with basis $\{1\}$.

3. If $r \notin \mathbb{Q}$: then $\{1, r\}$ spans V by definition. For independence: if $\alpha \cdot 1 + \beta \cdot r = 0$ with $\alpha, \beta \in \mathbb{Q}$ and $\beta \neq 0$, then $r = -\alpha/\beta \in \mathbb{Q}$, contradiction. Hence $\{1, r\}$ is a $\mathbb{Q}$-basis and $\dim_{\mathbb{Q}}(V) = 2$.

Key takeaway: dimension depends on the field. Over $\mathbb{R}$, $\dim_{\mathbb{R}}(\mathbb{R}) = 1$. Over $\mathbb{Q}$, $\dim_{\mathbb{Q}}(\mathbb{R})$ is infinite (uncountable, in fact).

Common Mistake

- Forgetting that **dimension depends on the field**. For example, $\mathbb{R}$ as a vector space over $\mathbb{Q}$ has very different dimension properties than over $\mathbb{R}$.
- Misusing the dimension formula: it requires *subspaces* of a *finite-dimensional* vector space.

Exam Focus

Dimension-proof checklist (very common in MH1201).

- If asked to show $U \cap V \neq \{0\}$, try $\dim(U \cap V) \geq \dim(U) + \dim(V) - \dim(\text{ambient})$.
- If given $\dim(U) + \dim(V) = \dim(U + V)$, then $\dim(U \cap V) = 0$ and (if $U + V$ is the whole space) you get a direct sum.
- If asked to build a basis for a constraint-defined subspace (e.g. $P(2) = P(5)$, $\text{trace}(M) = 0$, $AB = BA$):

  translate the condition into *linear equations in coefficients/entries*, then parametrize and read off a spanning set; finally check independence (often immediate).

3 Linear Transformations

3.1 Definitions, linearity tests, and core examples (Week 5)

Definition 3.1: Linear transformation

Let V and W be vector spaces over the same field $\mathbb{F}$. A function $T : V \to W$ is a *linear transformation* if for all $u, v \in V$ and all $\lambda \in \mathbb{F}$,

$$T(u+v) = T(u) + T(v) \quad \text{and} \quad T(\lambda u) = \lambda T(u).$$

Equivalently, T is linear if and only if for all $u, v \in V$ and all $\alpha, \beta \in \mathbb{F}$,

$$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v).$$

Intuitively, a linear transformation is a "structure-preserving" map: it respects the two operations that define a vector space (addition and scalar multiplication). An equivalent one-shot test is to check

$$\forall v_1, \ldots, v_k \in V,\ \forall \lambda_1, \ldots, \lambda_k \in \mathbb{F} : \qquad T\left(\sum_{i=1}^{k} \lambda_i v_i\right) = \sum_{i=1}^{k} \lambda_i T(v_i),$$

i.e. T commutes with *all* finite linear combinations. This general form follows by induction from the two-vector case and is often what you actually use in proofs. Two immediate consequences that serve as quick sanity checks:

- $T(0_V) = 0_W$ always (set $\lambda = 0$ in homogeneity). If $T(0) \neq 0$, then T is *not* linear.
- $T(-v) = -T(v)$ for all $v \in V$ (set $\lambda = -1$).

Key Formula

Linearity test (one line). To prove $T : V \to W$ is linear, show

$$\forall u, v \in V,\ \forall \alpha, \beta \in \mathbb{F} : \qquad T(\alpha u + \beta v) = \alpha T(u) + \beta T(v).$$

To disprove linearity, it suffices to find one counterexample to additivity or homogeneity.

Definition 3.2: The vector space of linear maps

Let V and W be vector spaces over $\mathbb{F}$. Define

$$\mathcal{L}(V, W) \coloneqq \{T : V \to W \mid T \text{ is linear}\}, \qquad \mathcal{L}(V) \coloneqq \mathcal{L}(V, V).$$

For $S, T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbb{F}$, define pointwise operations:

$$(S+T)(v) \coloneqq S(v) + T(v), \qquad (\lambda T)(v) \coloneqq \lambda T(v) \qquad (\forall v \in V).$$

Theorem 3.1: $\mathcal{L}(V, W)$ is a vector space

Let V, W be vector spaces over $\mathbb{F}$. Then $\mathcal{L}(V, W)$ is a vector space over $\mathbb{F}$ under pointwise addition and scalar multiplication.

Proof. We verify closure; the remaining axioms follow pointwise from those in W.
Let $S, T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbb{F}$. For all $u, v \in V$,

$$(S+T)(u+v) = S(u+v) + T(u+v) = S(u) + S(v) + T(u) + T(v) = (S+T)(u) + (S+T)(v),$$

and similarly $(S+T)(\lambda u) = \lambda(S+T)(u)$, hence $S+T \in \mathcal{L}(V, W)$. Also,

$$(\lambda T)(u+v) = \lambda T(u+v) = \lambda T(u) + \lambda T(v) = (\lambda T)(u) + (\lambda T)(v), \qquad (\lambda T)(\mu u) = \mu(\lambda T)(u),$$

so $\lambda T \in \mathcal{L}(V, W)$. □

When both V and W are finite-dimensional with $\dim(V) = m$ and $\dim(W) = n$, we will later show that $\mathcal{L}(V, W) \cong \mathbb{F}^{n \times m}$ as vector spaces (see the extension box in Section 3.8ff.), hence

$$\dim(\mathcal{L}(V, W)) = mn.$$

This matches the dimension of $\mathbb{F}^{n \times m}$ and reinforces the idea that choosing bases turns abstract linear maps into concrete matrices.

Exam Focus

Tutorial 5, Question 1 note. When asked to show $\mathcal{L}(V, W)$ is a vector space, you can either (i) verify all eight axioms directly (pointwise, borrowing from W), or (ii) show $\mathcal{L}(V, W)$ is a subspace of the function space W^V by checking closure under addition and scalar multiplication plus containing the zero map. Method (ii) is usually faster on exams because the axioms are inherited automatically.

Worked Example

Worked Example: linear vs. non-linear maps on $\mathbb{R}^2$. Let $T_1, T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T_1(x, y) = (x + 2y,\ 3x - y), \qquad T_2(x, y) = (x^2,\ y).$$

- T_1 is linear: both additivity and homogeneity follow by direct expansion.
- T_2 is not linear: homogeneity fails since $T_2(2, 0) = (4, 0) \neq 2(1, 0) = 2T_2(1, 0)$.

Worked Example

Worked Example: common linear maps (from Week 5 lectures).

- **Derivative operator:** $D : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$, $D(P) = P'$, is linear because $(P+Q)' = P' + Q'$ and $(\lambda P)' = \lambda P'$.
- **Multiplication operator:** Fix $Q \in \mathcal{P}(\mathbb{F})$; then $M_Q : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$, $M_Q(P) = QP$, is linear because polynomial multiplication distributes over addition and commutes with scalar multiplication.
- **Matrix operator:** Fix $A \in \mathbb{F}^{m \times n}$; then $T_A : \mathbb{F}^{n \times 1} \to \mathbb{F}^{m \times 1}$, $T_A(x) = Ax$, is linear by properties of matrix multiplication.
- **Backward-shift:** On $\mathbb{F}^\infty$, $\sigma_\leftarrow(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots)$ is linear (check componentwise).

These examples illustrate how linearity appears across many settings: polynomials, matrices, and sequence spaces.

3.2 Kernel (null space) and range (image) (Week 5)

Definition 3.3: Null space and range

Let V, W be vector spaces over $\mathbb{F}$ and let $T \in \mathcal{L}(V, W)$.

$$\begin{aligned} \text{Null}(T) &\coloneqq \{v \in V \mid T(v) = 0\} \quad \text{(also written } \ker(T)\text{)}, \\ \text{Range}(T) &\coloneqq \{T(v) \in W \mid v \in V\} \quad \text{(also written } \text{im}(T)\text{)}. \end{aligned}$$

The null space measures "how far T is from being injective": elements of $\text{Null}(T)$ are the vectors that T cannot distinguish from 0. The range measures "how much of W does T reach": it is the set of all possible outputs. Together with the Rank–Nullity Theorem (next subsection), these two subspaces completely determine the "information loss" and "coverage" of T.

Theorem 3.2: Null space and range are subspaces

Let V, W be vector spaces over $\mathbb{F}$ and $T \in \mathcal{L}(V, W)$. Then $\text{Null}(T)$ is a subspace of V, and $\text{Range}(T)$ is a subspace of W.

Proof. **Null space.** Since $T(0) = 0$, we have $0 \in \text{Null}(T)$. If $v, \tilde{v} \in \text{Null}(T)$ and $\lambda \in \mathbb{F}$, then

$$T(v + \tilde{v}) = T(v) + T(\tilde{v}) = 0 + 0 = 0, \qquad T(\lambda v) = \lambda T(v) = \lambda \cdot 0 = 0,$$

so $v + \tilde{v}, \lambda v \in \text{Null}(T)$.

Range. Since $0 = T(0) \in \text{Range}(T)$, it is nonempty. If $w = T(v)$ and $w' = T(v')$ are in the range, then for $\lambda \in \mathbb{F}$,

$$w + w' = T(v) + T(v') = T(v + v') \in \text{Range}(T), \qquad \lambda w = \lambda T(v) = T(\lambda v) \in \text{Range}(T).$$

Thus $\text{Range}(T)$ is a subspace of W. □

Extension (Beyond Syllabus: Preimages preserve subspaces)

Syllabus link: MH1201 Weeks 5–6 — Kernel and range.
Lemma. Let $T : V \to W$ be linear and let $U \leq W$ be a subspace. Then

$$T^{-1}(U) \coloneqq \{v \in V \mid T(v) \in U\}$$

is a subspace of V.
Proof. Since $0 \in U$ and $T(0) = 0$, we have $0 \in T^{-1}(U)$. Let $v, \tilde{v} \in T^{-1}(U)$ and $\lambda \in \mathbb{F}$. Then $T(v), T(\tilde{v}) \in U$. Because U is a subspace, $T(v) + T(\tilde{v}) \in U$ and $\lambda T(v) \in U$. By linearity,

$$T(v + \tilde{v}) = T(v) + T(\tilde{v}) \in U, \qquad T(\lambda v) = \lambda T(v) \in U,$$

so $v + \tilde{v}, \lambda v \in T^{-1}(U)$. Hence $T^{-1}(U)$ is a subspace. □

Worked Example

Worked Example: kernel and range of a polynomial map. Let $T : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}^2$ be defined by

$$\forall a, b, c \in \mathbb{R} : \qquad T(a + bX + cX^2) = (a + c,\ a - b).$$

Kernel. Solve $a + c = 0$ and $a - b = 0$, giving $(a, b, c) = (a, a, -a)$. Thus

$$\mathrm{Null}(T) = \{a(1 + X - X^2) \mid a \in \mathbb{R}\} = \mathrm{span}\{1 + X - X^2\}.$$

Range. For any $(u, v) \in \mathbb{R}^2$, take $a = 0$, $b = -v$, $c = u$, giving $T(uX^2 - vX) = (u, v)$. Hence $\mathrm{Range}(T) = \mathbb{R}^2$.

Worked Example

Worked Example (Tutorial 5, Question 3 style): kernel, range, and non-direct-sum. Define $T : \mathbb{R}^{2,2} \to \mathbb{R}^{2,2}$ by $T(A) = AJ - \mathrm{tr}(A)\, J$ where $J = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Explicit computation on $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ gives

$$T(A) = \begin{bmatrix} b - d & b - d \\ c - a & c - a \end{bmatrix}.$$

Range: Every output has equal columns, and any such matrix is attainable. Thus $\mathrm{Range}(T) = \left\{ \begin{bmatrix} u & u \\ v & v \end{bmatrix} \right\}$, with $\mathrm{Rank}(T) = 2$. **Kernel:** $T(A) = 0$ iff $b = d$ and $c = a$, so $\mathrm{Null}(T) = \left\{ \begin{bmatrix} a & b \\ a & b \end{bmatrix} \right\}$, with $\mathrm{Nullity}(T) = 2$. **Sanity check:** $\mathrm{Rank} + \mathrm{Nullity} = 2 + 2 = 4 = \dim(\mathbb{R}^{2,2})$. ✓
Non-direct-sum: $J = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ lies in both $\mathrm{Null}(T)$ and $\mathrm{Range}(T)$, so $\mathrm{Null}(T) \cap \mathrm{Range}(T) \neq \{0\}$ and the sum is *not* direct. This shows that $\mathbb{R}^{2,2} = \mathrm{Null}(T) \oplus \mathrm{Range}(T)$ does *not* hold in general.

3.3 Rank–Nullity theorem and injective/surjective criteria (Weeks 5–6)

Definition 3.4: Rank and nullity

Let V, W be vector spaces over $\mathbb{F}$ with V finite-dimensional, and let $T \in \mathcal{L}(V, W)$. Define

$$\mathrm{Nullity}(T) \coloneqq \dim(\mathrm{Null}(T)), \qquad \mathrm{Rank}(T) \coloneqq \dim(\mathrm{Range}(T)).$$

Theorem 3.3: Rank–Nullity Theorem

Let V, W be vector spaces over $\mathbb{F}$ with V finite-dimensional, and let $T \in \mathcal{L}(V, W)$. Then

$$\dim(V) = \mathrm{Rank}(T) + \mathrm{Nullity}(T).$$

Proof. Let $m \coloneqq \dim(\mathrm{Null}(T))$ and choose a basis $(b_1, \ldots, b_m)$ of $\mathrm{Null}(T)$. Extend it to a basis of V:

$$(b_1, \ldots, b_m, v_1, \ldots, v_{n-m}), \qquad n \coloneqq \dim(V).$$

We claim $(T(v_1), \ldots, T(v_{n-m}))$ is a basis of $\mathrm{Range}(T)$.

Independence. If $\sum_{i=1}^{n-m} \lambda_i T(v_i) = 0$, then $T(\sum_{i=1}^{n-m} \lambda_i v_i) = 0$, so $\sum_{i=1}^{n-m} \lambda_i v_i \in \text{Null}(T)$. But it also lies in $\text{span}\{v_1, \ldots, v_{n-m}\}$; by uniqueness of coordinates in the displayed basis, this forces $\sum_{i=1}^{n-m} \lambda_i v_i = 0$, hence all $\lambda_i = 0$.

Spanning. If $w \in \text{Range}(T)$, then $w = T(v)$ for some $v \in V$. Write $v = \sum_{j=1}^{m} \alpha_j b_j + \sum_{i=1}^{n-m} \beta_i v_i$. Since $T(b_j) = 0$,

$$w = T(v) = \sum_{i=1}^{n-m} \beta_i T(v_i) \in \text{span}\{T(v_1), \ldots, T(v_{n-m})\}.$$

Thus $\text{Rank}(T) = n - m$ and $\text{Nullity}(T) = m$, giving $n = (n - m) + m$.

□

The Rank–Nullity Theorem is one of the most frequently used results in linear algebra. It says that the "dimensions consumed by the kernel" plus the "dimensions captured in the range" always add up to the dimension of the domain.

Exam Focus

Using Rank–Nullity on exams.

- If you know any two of $\{\dim(V), \text{Rank}(T), \text{Nullity}(T)\}$, you can deduce the third.
- To show a map is injective, it suffices to show $\text{Nullity}(T) = 0$. To show surjectivity onto W, show $\text{Rank}(T) = \dim(W)$.
- Always state the hypothesis "V is finite-dimensional" before applying the theorem.
- Sanity check: after computing kernel and range independently, verify $\text{Rank} + \text{Nullity} = \dim(V)$.

Definition 3.5: Injective, surjective, bijective

Let $T \in \mathcal{L}(V, W)$.

- T is *injective* if $T(v) = T(\tilde{v}) \Rightarrow v = \tilde{v}$.
- T is *surjective onto* W if $\forall w \in W,\ \exists v \in V :\ T(v) = w$.
- T is *bijective* if it is both injective and surjective.

Injectivity means distinct inputs yield distinct outputs (no information loss); surjectivity means every element of the codomain is reached (full coverage). For linear maps, the null-space test (Theorem 3.4) provides a more efficient check for injectivity: instead of comparing all pairs, you only need to check that the kernel is trivial.

Key Formula

RREF criteria for injectivity/surjectivity (Tutorial 6, Question 2). Let V, W be finite-dimensional with ordered bases β, γ. Let $T \in \mathcal{L}(V, W)$ and $A = [T]_\beta^\gamma \in \mathbb{F}^{n \times m}$.

- T injective $\iff$ RREF of A has a pivot in every column ($\Leftrightarrow \text{Rank}(A) = m$, no free variables).
- T surjective onto W $\iff$ RREF of A has a pivot in every row ($\Leftrightarrow \text{Rank}(A) = n$).

These translate injectivity/surjectivity into concrete row-reduction checks that are fast to perform on exams.

Key Formula

Equal-dimension criterion (Tutorial 6, Question 3). If $\dim(V) = \dim(W) = n$ and $T \in \mathcal{L}(V, W)$, then

$$T \text{ injective} \iff T \text{ surjective onto } W \iff T \text{ bijective} \iff [T]_\beta^\gamma \text{ invertible.}$$

This is because Rank–Nullity forces $\text{Rank}(T) = n - \text{Nullity}(T)$, so nullity 0 is equivalent to rank $n = \dim(W)$.

Theorem 3.4: Null-space test for injectivity

Let $T \in \mathcal{L}(V, W)$. Then T is injective if and only if $\text{Null}(T) = \{0\}$.

Proof. If T is injective, then $T(v) = 0 = T(0)$ implies $v = 0$, so $\text{Null}(T) = \{0\}$. Conversely, if $\text{Null}(T) = \{0\}$ and $T(v) = T(\tilde{v})$, then $T(v - \tilde{v}) = 0$, hence $v - \tilde{v} = 0$ and $v = \tilde{v}$. □

Key Formula

Finite-dimensional shortcuts (state hypotheses!). Assume V, W are finite-dimensional.

- T injective $\iff \text{Nullity}(T) = 0 \iff \text{Rank}(T) = \dim(V)$.
- T surjective onto $W \iff \text{Rank}(T) = \dim(W)$.
- If $\dim(V) = \dim(W)$, then injective $\iff$ surjective $\iff$ bijective.

Definition 3.6: Invertible linear transformation

Let $T \in \mathcal{L}(V, W)$. We say T is *invertible* if there exists a function $S : W \to V$ such that

$$S \circ T = I_V \quad \text{and} \quad T \circ S = I_W.$$

In that case, S is unique and is denoted by T^{-1}.

Theorem 3.5: Invertibility $\iff$ bijectivity; inverse is linear

Let $T \in \mathcal{L}(V, W)$. Then T is invertible if and only if T is bijective. Moreover, if T is invertible, then $T^{-1} \in \mathcal{L}(W, V)$.

Proof. If T has inverse T^{-1}, then injectivity and surjectivity are immediate from $T^{-1} \circ T = I_V$ and $T \circ T^{-1} = I_W$. Conversely, if T is bijective, define $T^{-1}(w)$ to be the unique v with $T(v) = w$; then $T^{-1} \circ T = I_V$ and $T \circ T^{-1} = I_W$.

To show T^{-1} is linear: let $w_1, w_2 \in W$ and set $v_i = T^{-1}(w_i)$. Then $T(v_1 + v_2) = w_1 + w_2$, so $T^{-1}(w_1 + w_2) = v_1 + v_2$, and similarly $T^{-1}(\lambda w_1) = \lambda v_1$. Hence T^{-1} is linear.

□

Key Formula

Composition, injectivity, and surjectivity (Tutorial 5, Questions 6–7; Tutorial 6, Question 4). Let U, V, W be vector spaces and $S \in \mathcal{L}(U, V)$, $T \in \mathcal{L}(V, W)$.

- If $T \circ S$ is injective, then S is injective (but T need not be).
- If $T \circ S$ is surjective onto W, then T is surjective onto W (but S need not be).
- If S and T are both invertible, then $T \circ S$ is invertible and $(T \circ S)^{-1} = S^{-1} \circ T^{-1}$.
- The converse fails: $T \circ S$ invertible does *not* imply both S, T invertible individually.

Worked Example

Worked Example (Tutorial 5, Question 7): counterexamples for converses. Take $U = \mathbb{R}$, $V = \mathbb{R}^2$, $W = \mathbb{R}$. Define $S(t) = (t, 0)$ and $T(x, y) = x$. Then $(T \circ S)(t) = t$ is bijective.

- $T \circ S$ is injective, but T is *not* injective (since $T(0, 1) = 0 = T(0, 0)$).
- $T \circ S$ is surjective, but S is *not* surjective onto $\mathbb{R}^2$ (since $(0, 1) \notin S(\mathbb{R})$).

The key insight: $T \circ S$ injective only forces $\text{Range}(S) \cap \ker(T) = \{0\}$, not $\ker(T) = \{0\}$.

3.4 Matrix representations and composition (Week 6)

Definition 3.7: Matrix of a linear map

Let V, W be finite-dimensional over $\mathbb{F}$ with ordered bases

$$\beta = (b_1, \ldots, b_n) \text{ for } V, \qquad \gamma = (c_1, \ldots, c_m) \text{ for } W,$$

and let $T \in \mathcal{L}(V, W)$. Define $[T]_\beta^\gamma \in \mathbb{F}^{m \times n}$ by taking the i-th column to be $[T(b_i)]_\gamma$:

$$[T]_\beta^\gamma := \begin{bmatrix} | & & | \\ [T(b_1)]_\gamma & \cdots & [T(b_n)]_\gamma \\ | & & | \end{bmatrix}.$$

The matrix $[T]_\beta^\gamma$ encodes T completely (relative to the chosen bases): the j-th column tells you what T does to the j-th basis vector of the domain, expressed in the codomain basis. Changing bases will change the matrix, but the underlying linear map stays the same. This is why matrix representations depend on *two* ordered bases: one for the domain and one for the codomain.

Common Mistake

- Confusing $[T]_\beta^\gamma$ (a matrix in $\mathbb{F}^{m\times n}$) with T (an abstract linear map). The matrix is a *representation* that depends on the basis choice.
- Forgetting to use the correct basis for each side: the input vector must be in β-coordinates, and the output is in γ-coordinates.
- Dimension mismatch: $[T]_\beta^\gamma$ is $\dim(W) \times \dim(V)$, not $\dim(V) \times \dim(W)$.

Theorem 3.6: Fundamental coordinate identity

With the notation of Definition 3.7, for every $v \in V$,

$$[T(v)]_\gamma = [T]_\beta^\gamma\,[v]_\beta.$$

Proof. Write $v = \sum_{i=1}^n \lambda_i b_i$, so $[v]_\beta = (\lambda_1, \ldots, \lambda_n)^\mathsf{T}$. By linearity, $T(v) = \sum_{i=1}^n \lambda_i T(b_i)$, hence

$$[T(v)]_\gamma = \sum_{i=1}^n \lambda_i [T(b_i)]_\gamma = \begin{bmatrix} | & & | \\ [T(b_1)]_\gamma & \cdots & [T(b_n)]_\gamma \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = [T]_\beta^\gamma\,[v]_\beta.$$

□

Theorem 3.7: Composition corresponds to matrix multiplication

Let U, V, W be finite-dimensional over $\mathbb{F}$ with ordered bases α for U, β for V, and γ for W. Let $S \in \mathcal{L}(U, V)$ and $T \in \mathcal{L}(V, W)$. Then

$$[T \circ S]_\alpha^\gamma = [T]_\beta^\gamma\,[S]_\alpha^\beta.$$

Proof. For all $u \in U$,

$$[(T \circ S)(u)]_\gamma = [T(S(u))]_\gamma = [T]_\beta^\gamma [S(u)]_\beta = [T]_\beta^\gamma [S]_\alpha^\beta [u]_\alpha.$$

Also $[(T \circ S)(u)]_\gamma = [T \circ S]_\alpha^\gamma [u]_\alpha$. Since this holds for all u, the matrices must be equal.

□

The composition formula $[T \circ S]_\alpha^\gamma = [T]_\beta^\gamma [S]_\alpha^\beta$ requires the "inner" basis β to match: S outputs β-coordinates, and T reads β-coordinates. If the intermediate bases do not match, the formula fails.

Worked Example

Worked Example (Tutorial 5, Question 5 style): operator on $\mathcal{P}_2(\mathbb{R})$. Define $T \in \mathcal{L}(\mathcal{P}_2(\mathbb{R}))$ by $T(P) = P(X+1) - (P(X^2))''$. Write $P = a + bX + cX^2$.

- $P(X+1) = (a+b+c) + (b+2c)X + cX^2$.
- $(P(X^2))'' = (a + bX^2 + cX^4)'' = 2b + 12cX^2$.

Thus $T(P) = (a-b+c) + (b+2c)X - 11cX^2$, giving the standard matrix

$$[T]_{(1,X,X^2)} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & -11 \end{bmatrix}.$$

This is upper triangular with diagonal entries $1, 1, -11$ (all nonzero), so $\det = -11 \neq 0$ and T is bijective.

3.5 Change of coordinates and change of matrix representations (Week 7)

Definition 3.8: Change-of-coordinate matrix

Let V be a finite-dimensional vector space over $\mathbb{F}$. Let β and γ be ordered bases for V. The **change-of-coordinate matrix from β to γ** is

$$[I_V]_\beta^\gamma \;:=\; [I_V]_\beta^\gamma,$$

i.e. the matrix representation of the identity map $I_V : V \to V$ with domain basis β and codomain basis γ.

Why is this natural? Suppose you have a vector $v \in V$ whose β-coordinates you know, and you want its γ-coordinates. The identity map I_V does not change the vector itself, but different bases "see" the same vector differently. The matrix $[I_V]_\beta^\gamma$ is the translator: it converts β-coordinates to γ-coordinates. Its inverse $[I_V]_\gamma^\beta = ([I_V]_\beta^\gamma)^{-1}$ translates in the opposite direction.

Exam Focus

Practical computation of $[I_V]_\beta^\gamma$ (Week 7). If V has a convenient "standard" basis σ, form the matrices $P_\beta = [I_V]_\beta^\sigma$ and $P_\gamma = [I_V]_\gamma^\sigma$ by stacking basis vectors as columns, then

$$[I_V]_\beta^\gamma = P_\gamma^{-1} P_\beta.$$

An efficient row-reduction approach: form $[\,P_\gamma \mid P_\beta\,]$ and row-reduce to $[\,I_n \mid P_\gamma^{-1} P_\beta\,]$.

Theorem 3.8: Changing coordinate vectors

Let V be finite-dimensional with ordered bases β and γ. Then for every $v \in V$,

$$[v]_\gamma = [I_V]_\beta^\gamma \, [v]_\beta.$$

Proof. Apply Theorem 3.6 to $T = I_V$:

$$[I_V(v)]_\gamma = [I_V]_\beta^\gamma \, [v]_\beta.$$

Since $I_V(v) = v$, the result follows. □

Theorem 3.9: Inverse map and inverse matrix (Week 7)

Let V, W be finite-dimensional over $\mathbb{F}$ with ordered bases β for V and γ for W. Let $T \in \mathcal{L}(V, W)$. Then T is invertible from V to W if and only if $[T]_\beta^\gamma$ is an invertible square matrix, in which case

$$[T^{-1}]_\gamma^\beta = \left([T]_\beta^\gamma\right)^{-1}.$$

Proof. ($\Rightarrow$) If T is invertible, then $\dim(V) = \dim(W) = n$. Using composition–multiplication (Theorem 3.7),

$$[I_V]_\beta^\beta = [T^{-1} \circ T]_\beta^\beta = [T^{-1}]_\gamma^\beta [T]_\beta^\gamma, \quad [I_W]_\gamma^\gamma = [T \circ T^{-1}]_\gamma^\gamma = [T]_\beta^\gamma [T^{-1}]_\gamma^\beta.$$

Thus $[T]_\beta^\gamma$ is invertible with inverse $[T^{-1}]_\gamma^\beta$.

($\Leftarrow$) Suppose $[T]_\beta^\gamma \in \mathbb{F}^{n \times n}$ is invertible. Let $A := ([T]_\beta^\gamma)^{-1}$. By the standard existence-and-uniqueness result for a linear map specified on a basis (Weeks 5–6), there exists $S \in \mathcal{L}(W, V)$ with $[S]_\gamma^\beta = A$. Then

$$[S \circ T]_\beta^\beta = [S]_\gamma^\beta [T]_\beta^\gamma = A[T]_\beta^\gamma = I_n = [I_V]_\beta^\beta,$$

and similarly $[T \circ S]_\gamma^\gamma = I_n = [I_W]_\gamma^\gamma$. Hence $S \circ T = I_V$ and $T \circ S = I_W$, so T is invertible and $S = T^{-1}$. □

Theorem 3.10: Change of matrix representations under a change of coordinates (Week 7)

Let V be finite-dimensional and let β, γ be ordered bases for V. Let $T \in \mathcal{L}(V)$. Then

$$[T]_\gamma^\gamma = [I_V]_\beta^\gamma \, [T]_\beta^\beta \, [I_V]_\gamma^\beta.$$

Proof. For any $v \in V$,

$$[T(v)]_\gamma = [I_V]_\beta^\gamma [T(v)]_\beta \quad \text{and} \quad [T(v)]_\beta = [T]_\beta^\beta [v]_\beta \quad \text{and} \quad [v]_\beta = [I_V]_\gamma^\beta [v]_\gamma.$$

Substitute the last two into the first:

$$[T(v)]_\gamma = [I_V]_\beta^\gamma [T]_\beta^\beta [I_V]_\gamma^\beta [v]_\gamma.$$

But also $[T(v)]_\gamma = [T]_\gamma^\gamma [v]_\gamma$. Since this holds for all v, the matrices are equal.

□

This formula says: to change the basis from β to γ, you "sandwich" the β-representation between the change-of-coordinate matrices. Read right-to-left: first convert γ-coordinates to β-coordinates (multiply by

$[I_V]_\gamma^\beta$), then apply T in β-coordinates, then convert back to γ-coordinates (multiply by $[I_V]_\beta^\gamma$). This is exactly the similarity relation $[T]_\gamma^\gamma = Q^{-1}[T]_\beta^\beta Q$ with $Q = [I_V]_\gamma^\beta$.

Key Formula

Computing $[I_V]_\beta^\gamma$ via a standard basis σ (Week 7). Let σ be a standard ordered basis for V (e.g. the standard basis in $\mathbb{F}^n$). Then

$$[I_V]_\beta^\gamma = \left([I_V]_\gamma^\sigma\right)^{-1}[I_V]_\beta^\sigma, \qquad [I_V]_\gamma^\beta = \left([I_V]_\beta^\sigma\right)^{-1}[I_V]_\gamma^\sigma,$$

where $[I_V]_\beta^\sigma = \begin{bmatrix}[b_1]_\sigma & \cdots & [b_n]_\sigma\end{bmatrix}$ is formed by stacking basis vectors (in σ-coordinates) as columns.

Worked Example

Worked Example (Lecture Week 7 problem): compute $[I_V]_\beta^\alpha$ and coordinates.
Let $\alpha = (a_1, a_2, a_3)$ and $\beta = (b_1, b_2, b_3)$ be ordered bases of V such that

$$b_1 = 2a_1 - a_2 + a_3, \qquad b_2 = 3a_2 + a_3, \qquad b_3 = -3a_1 + 2a_3.$$

(1) Compute $[I_V]_\beta^\alpha$. By definition, the i-th column of $[I_V]_\beta^\alpha$ is $[b_i]_\alpha$. Thus

$$[b_1]_\alpha = \begin{bmatrix}2\\-1\\1\end{bmatrix}, \quad [b_2]_\alpha = \begin{bmatrix}0\\3\\1\end{bmatrix}, \quad [b_3]_\alpha = \begin{bmatrix}-3\\0\\2\end{bmatrix},$$

so

$$[I_V]_\beta^\alpha = \begin{bmatrix}2 & 0 & -3\\-1 & 3 & 0\\1 & 1 & 2\end{bmatrix}.$$

(2) Compute $[b_1 - 2b_2 + 2b_3]_\alpha$. Using linearity of coordinate expansion in the basis α,

$$[b_1 - 2b_2 + 2b_3]_\alpha = [b_1]_\alpha - 2[b_2]_\alpha + 2[b_3]_\alpha = \begin{bmatrix}2\\-1\\1\end{bmatrix} - 2\begin{bmatrix}0\\3\\1\end{bmatrix} + 2\begin{bmatrix}-3\\0\\2\end{bmatrix} = \begin{bmatrix}-4\\-7\\3\end{bmatrix}.$$

Worked Example

Worked Example (Lecture Week 7 problem): compute change-of-coordinate matrices in $\mathbb{R}^2$.
Let $\beta = ((-9,1),(-5,-1))$ and $\gamma = ((1,-4),(3,-5))$ be ordered bases for $\mathbb{R}^2$. Let σ be the standard basis. Then

$$[I_{\mathbb{R}^2}]_\beta^\sigma = \begin{bmatrix}-9 & -5\\1 & -1\end{bmatrix}, \qquad [I_{\mathbb{R}^2}]_\gamma^\sigma = \begin{bmatrix}1 & 3\\-4 & -5\end{bmatrix}.$$

Hence

$$[I_{\mathbb{R}^2}]_\beta^\gamma = \left([I_{\mathbb{R}^2}]_\gamma^\sigma\right)^{-1}[I_{\mathbb{R}^2}]_\beta^\sigma = \begin{bmatrix}6 & 4\\-5 & -3\end{bmatrix}, \qquad [I_{\mathbb{R}^2}]_\gamma^\beta = \left([I_{\mathbb{R}^2}]_\beta^\sigma\right)^{-1}[I_{\mathbb{R}^2}]_\gamma^\sigma = \begin{bmatrix}-\frac{3}{2} & -2\\\frac{5}{2} & 3\end{bmatrix}.$$

(Indeed $[I]_\beta^\gamma[I]_\gamma^\beta = I_2$.)

Worked Example

Worked Example (Tutorial 6, Question 7): compute $[T]_\gamma^\gamma$ by similarity. Define $T \in \mathcal{L}(\mathbb{R}^3)$ by $T(v) = Av$ where $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 3 \\ 0 & -1 & 0 \end{bmatrix}$. Let $\gamma = (g_1, g_2, g_3)$ with $g_1 = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$, $g_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$, $g_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Form $P = [g_1\ g_2\ g_3] = \begin{bmatrix} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. Then

$$[T]_\gamma^\gamma = P^{-1}AP = \begin{bmatrix} 0 & 2 & 8 \\ -1 & 4 & 6 \\ 0 & -1 & -1 \end{bmatrix}.$$

Sanity check: $\operatorname{tr}([T]_\gamma^\gamma) = 0 + 4 + (-1) = 3 = \operatorname{tr}(A) = 2 + 1 + 0$ ✓, and $\det([T]_\gamma^\gamma) = \det(A)$ (both can be verified numerically).

Worked Example

Worked Example (Lecture Week 7/8 problem): diagonalizing a map via a tailored basis. Define $T \in \mathcal{L}(\mathbb{R}^3)$ by

$$T(x, y, z) = (2x + y,\ 5y + 3z,\ 8z).$$

Let $\beta = (b_1, b_2, b_3)$ with $b_1 = (1, 0, 0)$, $b_2 = (1, 3, 0)$, $b_3 = (1, 6, 6)$.

1. Compute $[T]_\beta^\beta$.
2. Compute $T^4(0, -3, 6)$.

Solution. (1) Compute $T(b_i)$ and express in the basis β:

$$T(b_1) = (2, 0, 0) = 2b_1, \qquad T(b_2) = (5, 15, 0) = 5b_2, \qquad T(b_3) = (8, 48, 48) = 8b_3.$$

Thus

$$[T]_\beta^\beta = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 8 \end{bmatrix}.$$

(2) First compute $[(0, -3, 6)]_\beta$. Solve

$$(0, -3, 6) = \alpha b_1 + \beta b_2 + \gamma b_3 = (\alpha + \beta + \gamma,\ 3\beta + 6\gamma,\ 6\gamma),$$

giving $\gamma = 1$, $\beta = -3$, $\alpha = 2$. Hence $[(0, -3, 6)]_\beta = (2, -3, 1)^\mathsf{T}$. Then

$$[T^4(0, -3, 6)]_\beta = \left([T]_\beta^\beta\right)^4 [(0, -3, 6)]_\beta = \operatorname{diag}(2^4, 5^4, 8^4) \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 32 \\ -1875 \\ 4096 \end{bmatrix}.$$

Convert back:

$$T^4(0, -3, 6) = 32b_1 - 1875b_2 + 4096b_3 = (2253,\ 18951,\ 24576).$$

Extension (Beyond Syllabus: the coordinate isomorphism $[\cdot]_\beta$ packages everything)

Syllabus link: MH1201 Weeks 7–8 — Change of coordinates.
Theorem. Let U, V be finite-dimensional over $\mathbb{F}$ with ordered bases α (for U) and β (for V). Define the map

$$\Phi_{\beta,\alpha} : \mathcal{L}(U,V) \to \mathbb{F}^{\dim(V)\times\dim(U)}, \qquad \Phi_{\beta,\alpha}(S) = [S]_\alpha^\beta.$$

Then $\Phi_{\beta,\alpha}$ is an isomorphism of vector spaces. Moreover, for $S \in \mathcal{L}(U,V)$ and $T \in \mathcal{L}(V,W)$ (with ordered bases γ for W),

$$[T \circ S]_\alpha^\gamma = [T]_\beta^\gamma [S]_\alpha^\beta.$$

Proof. *Linearity:* For any $S_1, S_2 \in \mathcal{L}(U,V)$ and $\lambda \in \mathbb{F}$ and any $u \in U$,

$$[(S_1 + S_2)(u)]_\beta = [S_1(u) + S_2(u)]_\beta = [S_1(u)]_\beta + [S_2(u)]_\beta,$$

so by the fundamental identity (Theorem 3.6) applied to each map,

$$[S_1 + S_2]_\alpha^\beta [u]_\alpha = [S_1]_\alpha^\beta [u]_\alpha + [S_2]_\alpha^\beta [u]_\alpha.$$

Since this holds for all u, $[S_1 + S_2]_\alpha^\beta = [S_1]_\alpha^\beta + [S_2]_\alpha^\beta$. Similarly $[\lambda S_1]_\alpha^\beta = \lambda [S_1]_\alpha^\beta$.
Injective: If $[S]_\alpha^\beta = 0$, then for all u, $[S(u)]_\beta = 0$, hence $S(u) = 0$. So $S = 0$.
Surjective: Given any matrix $A \in \mathbb{F}^{\dim(V)\times\dim(U)}$, define S on the basis $\alpha = (a_1, \ldots, a_n)$ by requiring $[S(a_j)]_\beta$ to be the j-th column of A, then extend linearly. This produces a linear map with matrix $[S]_\alpha^\beta = A$ (uniqueness/existence from basis specification).
Finally, the composition identity is exactly Theorem 3.7. □

Common Mistake

- Swapping the direction of $[I_V]_\beta^\gamma$: it converts *β-coordinates to γ-coordinates* via $[v]_\gamma = [I_V]_\beta^\gamma [v]_\beta$.
- Forgetting that $[I_V]_\gamma^\beta = \left([I_V]_\beta^\gamma\right)^{-1}$ (invertibility holds because I_V is invertible).

3.6 Similarity of square matrices (Week 8)

Definition 3.9: Similarity

Let $\mathbb{F}$ be a field and $n \in \mathbb{N}$. For $A, B \in \mathbb{F}^{n\times n}$, we say A **is similar to** B, written $A \sim B$, if there exists an invertible $Q \in \mathbb{F}^{n\times n}$ such that

$$B = Q^{-1}AQ.$$

Theorem 3.11: Similarity is an equivalence relation

Let $\mathbb{F}$ be a field and $n \in \mathbb{N}$. The relation $\sim$ on $\mathbb{F}^{n\times n}$ is an equivalence relation.

Proof. **Reflexive.** For any A, take $Q = I_n$, then $A = I_n^{-1} A I_n$, so $A \sim A$.
Symmetric. If $A \sim B$, then $B = Q^{-1}AQ$ for some invertible Q. Multiply by Q on the left and Q^{-1} on the right to get $A = QBQ^{-1}$, so $B \sim A$.
Transitive. If $A \sim B$ and $B \sim C$, then $B = Q^{-1}AQ$ and $C = P^{-1}BP$ for invertible P, Q. Substitute:

$$C = P^{-1}(Q^{-1}AQ)P = (QP)^{-1}A(QP),$$

so $A \sim C$.

□

Similarity partitions $\mathbb{F}^{n\times n}$ into equivalence classes. Each class consists of all matrices that represent the *same* linear operator under various basis choices. This is why similarity invariants (determinant, trace, eigenvalues, characteristic polynomial) are really properties of the underlying operator T, not of any particular matrix.

Theorem 3.12: Determinant and trace are similarity invariants

Let $\mathbb{F}$ be a field and $A, B \in \mathbb{F}^{n\times n}$ with $A \sim B$. Then

$$\det(A) = \det(B) \qquad \text{and} \qquad \operatorname{tr}(A) = \operatorname{tr}(B).$$

Proof. Let $B = Q^{-1}AQ$ with Q invertible. Using multiplicativity of determinant,

$$\det(B) = \det(Q^{-1}AQ) = \det(Q^{-1})\det(A)\det(Q) = \det(Q)^{-1}\det(A)\det(Q) = \det(A).$$

For trace, use $\operatorname{tr}(XY) = \operatorname{tr}(YX)$ (valid for square matrices of the same size):

$$\operatorname{tr}(B) = \operatorname{tr}(Q^{-1}AQ) = \operatorname{tr}(AQQ^{-1}) = \operatorname{tr}(A).$$

□

Theorem 3.13: Matrix representations under different ordered bases are similar

Let V be finite-dimensional over $\mathbb{F}$, and let β, γ be ordered bases for V. Let $T \in \mathcal{L}(V)$. Then

$$[T]_\beta^\beta \sim [T]_\gamma^\gamma.$$

Proof. By Theorem 3.10,

$$[T]_\gamma^\gamma = [I_V]_\beta^\gamma [T]_\beta^\beta [I_V]_\gamma^\beta.$$

Since $[I_V]_\gamma^\beta = \left([I_V]_\beta^\gamma\right)^{-1}$ (because I_V is invertible), this is exactly the form $Q^{-1}AQ$ with $Q = [I_V]_\gamma^\beta$ (equivalently $Q^{-1} = [I_V]_\beta^\gamma$). Thus the two matrices are similar. □

Common Mistake

- Confusing similarity with equality: $A \sim B$ means they represent the *same linear operator* in different coordinates (bases), not that $A = B$.
- When proving invariants, do not assume $\operatorname{tr}(XY) = \operatorname{tr}(X)\operatorname{tr}(Y)$ (false). The correct identity is $\operatorname{tr}(XY) = \operatorname{tr}(YX)$.

3.7 Eigenvalues, eigenvectors, eigenspaces (Week 8)

Definition 3.10: Eigenvalue

Let V be a vector space over a field $\mathbb{F}$ and let $T \in \mathcal{L}(V)$. A scalar $\lambda \in \mathbb{F}$ is an **eigenvalue** of T if any (hence all) of the following equivalent statements hold:

1. $\text{Null}(T - \lambda I_V) \neq \{0\}$;
2. $\{0\} \subsetneq \text{Null}(T - \lambda I_V)$;
3. $T - \lambda I_V$ is not injective.

Eigenvalues are "field-specific": they must come from the base field $\mathbb{F}$ of the vector space. The same matrix can have eigenvalues over $\mathbb{C}$ but none over $\mathbb{R}$ (see the rotation example below). This is why the problem statement should always specify the field.

Intuitively, λ is an eigenvalue if $T - \lambda I_V$ "collapses" some nonzero vector to zero, i.e. T acts on some direction simply as scaling by λ.

Key Formula

Field-specific warning (Week 8). Eigenvalues must lie in the base field $\mathbb{F}$ of the vector space. The same linear map can have no eigenvalues over $\mathbb{R}$ but have eigenvalues over $\mathbb{C}$.

Theorem 3.14: Zero eigenvalue $\iff$ non-injectivity

Let V be a vector space over $\mathbb{F}$ and $T \in \mathcal{L}(V)$. Then $0 \in \mathbb{F}$ is an eigenvalue of T if and only if T is not injective.

Proof. 0 is an eigenvalue $\iff \text{Null}(T - 0 \cdot I_V) = \text{Null}(T) \neq \{0\} \iff T$ is not injective by Theorem 3.4. □

Definition 3.11: Eigenvector and eigenspace

Let V be a vector space over $\mathbb{F}$ and let $T \in \mathcal{L}(V)$. A nonzero vector $v \in V$ is an **eigenvector** of T if there exists $\lambda \in \mathbb{F}$ such that

$$T(v) = \lambda v.$$

For an eigenvalue λ of T, the **eigenspace** of T for λ is

$$E_{T,\lambda} \coloneqq \text{Null}(T - \lambda I_V).$$

Then the eigenvectors for λ are exactly $E_{T,\lambda} \setminus \{0\}$.

The eigenspace $E_{T,\lambda}$ is a subspace of V (it is the null space of $T - \lambda I_V$). Its nonzero elements are the eigenvectors for λ. The zero vector is in $E_{T,\lambda}$ but is *never* called an eigenvector (by convention, eigenvectors must be nonzero). Note that the eigenvalue associated to a given eigenvector v is unique: if $T(v) = \lambda v$ and $T(v) = \mu v$ with $v \neq 0$, then $(\lambda - \mu)v = 0$ implies $\lambda = \mu$.

Theorem 3.15: Eigenvector–kernel equivalence

Let $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. For $v \in V$,

$$v \in E_{T,\lambda} \iff (T - \lambda I_V)(v) = 0 \iff T(v) = \lambda v.$$

Proof. Immediate from expanding $(T - \lambda I_V)(v) = T(v) - \lambda I_V(v) = T(v) - \lambda v$. □

Theorem 3.16: Computing eigenvalues via a matrix representation (finite-dimensional)

Let V be finite-dimensional over $\mathbb{F}$ with $\dim(V) = n$, and let $T \in \mathcal{L}(V)$. Let β be any ordered basis for V. Then $\lambda \in \mathbb{F}$ is an eigenvalue of T if and only if

$$\det\big([T]_\beta^\beta - \lambda I_n\big) = 0.$$

Proof. λ is an eigenvalue $\iff T - \lambda I_V$ is not injective $\iff [T - \lambda I_V]_\beta^\beta$ is not invertible $\iff [T]_\beta^\beta - \lambda [I_V]_\beta^\beta$ is not invertible $\iff [T]_\beta^\beta - \lambda I_n$ is not invertible $\iff \det([T]_\beta^\beta - \lambda I_n) = 0$. □

Definition 3.12: Characteristic polynomial function

Let V be finite-dimensional over $\mathbb{F}$ with $\dim(V) = n$ and let $T \in \mathcal{L}(V)$. Fix any ordered basis β of V. Define the **characteristic polynomial function** of T by

$$p_T(\lambda) := \det\big([T]_\beta^\beta - \lambda I_n\big) \qquad (\lambda \in \mathbb{F}).$$

For an $n \times n$ matrix A, the characteristic polynomial $p_T(\lambda) = \det(A - \lambda I_n)$ is a polynomial of degree n in λ.

- For $n = 2$: $\det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} = \lambda^2 - (a + d)\lambda + (ad - bc) = \lambda^2 - \operatorname{tr}(A)\lambda + \det(A)$.
- For $n = 3$: expand using cofactor expansion along a row or column.

The roots of p_T in $\mathbb{F}$ are the eigenvalues. If $\mathbb{F} = \mathbb{R}$ and the polynomial has no real roots, then T has no eigenvalues over $\mathbb{R}$.

Theorem 3.17: Basis-independence of p_T

Let V be finite-dimensional and $T \in \mathcal{L}(V)$. If β and γ are two ordered bases of V, then for all $\lambda \in \mathbb{F}$,

$$\det\big([T]_\beta^\beta - \lambda I_n\big) = \det\big([T]_\gamma^\gamma - \lambda I_n\big).$$

Hence p_T is well-defined (independent of the chosen ordered basis).

Proof. By Theorem 3.13, there exists invertible Q such that $[T]_\gamma^\gamma = Q^{-1}[T]_\beta^\beta Q$. Then for every λ,

$$[T]_\gamma^\gamma - \lambda I_n = Q^{-1}\big([T]_\beta^\beta - \lambda I_n\big)Q,$$

so the two matrices are similar. By Theorem 3.12, similar matrices have equal determinants. □

Key Formula

Eigenvalues = roots of p_T (finite-dimensional). If $\dim(V) = n$, then eigenvalues of T are exactly the $\mathbb{F}$-roots of p_T. In particular, T has at most n distinct eigenvalues in $\mathbb{F}$.

Worked Example

Worked Example (Lecture Week 8): eigenvalues of $0_{V\to V}$ and I_V. Let V be nonzero over $\mathbb{F}$.

- For the zero map $0_{V\to V}$, we have $0_{V\to V}(v) = 0$ for all v, so every nonzero v satisfies $0(v) = 0 \cdot v$. Thus the only eigenvalue is $\lambda = 0$.
- For the identity I_V, we have $I_V(v) = 1 \cdot v$ for all $v \neq 0$, so the only eigenvalue is $\lambda = 1$.

Worked Example

Worked Example (Lecture Week 8): rotation matrix over $\mathbb{R}$ vs $\mathbb{C}$. Let $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and define $T(v) = Av$. Then

$$\det(A - \lambda I_2) = \det \begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix} = \lambda^2 + 1.$$

Over $\mathbb{R}$, $\lambda^2 + 1 = 0$ has no roots, so T has no real eigenvalues. Over $\mathbb{C}$, the roots are $\lambda = \pm i$, so T has eigenvalues $\pm i$ in $\mathbb{C}$.

Worked Example

Worked Example (Lecture Week 8): verify an eigenvector quickly. Let $T(v) = Av$ on $\mathbb{R}^3$ where

$$A = \begin{bmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}.$$

Check $v = \begin{bmatrix} 16 \\ 4 \\ 1 \end{bmatrix}$:

$$Av = \begin{bmatrix} 0 \cdot 16 + 6 \cdot 4 + 8 \cdot 1 \\ \frac{1}{2} \cdot 16 \\ \frac{1}{2} \cdot 4 \end{bmatrix} = \begin{bmatrix} 32 \\ 8 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 16 \\ 4 \\ 1 \end{bmatrix}.$$

Hence v is an eigenvector with eigenvalue $\lambda = 2$.

Worked Example

Worked Example (Lecture Week 8): two eigenvectors without heavy computation. Let $T(v) = Av$ on $\mathbb{R}^3$ where

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}.$$

Let $u = (1,1,1)^{\mathsf{T}}$. Then

$$Au = \begin{bmatrix} 6 \\ 6 \\ 6 \end{bmatrix} = 6u,$$

so u is an eigenvector for $\lambda = 6$. Now take any nonzero w with $(1,2,3) \cdot w = 0$, e.g. $w = (2,-1,0)^{\mathsf{T}}$. Then $Aw = u\,(1,2,3) \cdot w = 0$, so w is an eigenvector for $\lambda = 0$. Thus we found two eigenvectors for distinct eigenvalues 6 and 0.

Theorem 3.18: Eigenvectors for distinct eigenvalues are linearly independent

Let V be a vector space over $\mathbb{F}$ and let $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_n$ be distinct eigenvalues of T and let $v_i \neq 0$ be an eigenvector for λ_i. Then $\{v_1, \ldots, v_n\}$ is linearly independent.

Proof. We use induction on n.
For $n = 1$, $\{v_1\}$ is independent since $v_1 \neq 0$.
Assume the statement true for $n-1$ and consider a relation

$$\mu_1 v_1 + \cdots + \mu_n v_n = 0.$$

Apply T:

$$0 = T(0) = \mu_1 T(v_1) + \cdots + \mu_n T(v_n) = \mu_1 \lambda_1 v_1 + \cdots + \mu_n \lambda_n v_n.$$

Multiply the original relation by λ_n and subtract:

$$0 = \mu_1(\lambda_1 - \lambda_n)v_1 + \cdots + \mu_{n-1}(\lambda_{n-1} - \lambda_n)v_{n-1}.$$

By distinctness, each $(\lambda_i - \lambda_n) \neq 0$, so by the induction hypothesis $\{v_1, \ldots, v_{n-1}\}$ is independent and hence

$$\mu_1(\lambda_1 - \lambda_n) = \cdots = \mu_{n-1}(\lambda_{n-1} - \lambda_n) = 0 \implies \mu_1 = \cdots = \mu_{n-1} = 0.$$

Then the original relation becomes $\mu_n v_n = 0$, so $\mu_n = 0$ since $v_n \neq 0$. Thus all coefficients are zero, proving independence. □

Theorem 3.19: Sum of distinct eigenspaces is a direct sum

Let V be a vector space over $\mathbb{F}$ and let $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_n$ be distinct eigenvalues. Then

$$E_{T,\lambda_1} + \cdots + E_{T,\lambda_n} = E_{T,\lambda_1} \oplus \cdots \oplus E_{T,\lambda_n}.$$

Proof. We show uniqueness of zero decomposition. Let $v_i \in E_{T,\lambda_i}$ and suppose

$$v_1 + \cdots + v_n = 0.$$

If all $v_i = 0$ we are done. Otherwise remove the zero terms and keep a nonempty sub-list of nonzero vectors, still satisfying a linear combination equal to 0:

$$v_{i_1} + \cdots + v_{i_k} = 0, \qquad v_{i_j} \neq 0.$$

But each v_{i_j} is an eigenvector for the distinct eigenvalue λ_{i_j}, so by Theorem 3.18 the set $\{v_{i_1}, \ldots, v_{i_k}\}$ is linearly independent, contradicting a nontrivial dependence relation. Hence all $v_i = 0$, so the sum is direct. □

This theorem is the foundation of diagonalization: if the eigenspaces span the whole space, then we can choose a basis of eigenvectors. The direct-sum property ensures that eigenvectors from different eigenspaces never "interfere" with each other, so combining bases of individual eigenspaces always yields a linearly independent set.

Worked Example

Worked Example (Lecture Week 8): eigenvalues/eigenspaces of a "swap" operator on $\mathbb{R}^{2,2}$. Define $T \in \mathcal{L}(\mathbb{R}^{2,2})$ by

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} c & d \\ a & b \end{bmatrix}.$$

Then $T^2 = I$, so any eigenvalue λ must satisfy $\lambda^2 = 1$, hence $\lambda \in \{1, -1\}$.
Eigenspace for $\lambda = 1$. Solve $T(A) = A$:

$$\begin{bmatrix} c & d \\ a & b \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \iff a = c,\ b = d,$$

so

$$E_{T,1} = \left\{ \begin{bmatrix} a & b \\ a & b \end{bmatrix} : a, b \in \mathbb{R} \right\}.$$

Eigenspace for $\lambda = -1$. Solve $T(A) = -A$:

$$\begin{bmatrix} c & d \\ a & b \end{bmatrix} = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix} \iff c = -a,\ d = -b,$$

so

$$E_{T,-1} = \left\{ \begin{bmatrix} a & b \\ -a & -b \end{bmatrix} : a, b \in \mathbb{R} \right\}.$$

Moreover, $\mathbb{R}^{2,2} = E_{T,1} \oplus E_{T,-1}$ by Theorem 3.19.

Worked Example

Worked Example (Tutorial 7, Question 5): eigenvalues/eigenspaces of $F(x,y)=(y,x+y)$. Define $F \in \mathcal{L}(\mathbb{R}^{2,1})$ by

$$F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} y \\ x+y \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$. Then

$$\det(A - \lambda I_2) = \det \begin{bmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} = \lambda^2 - \lambda - 1,$$

so eigenvalues are

$$\lambda_{\pm} = \frac{1 \pm \sqrt{5}}{2}.$$

For $\lambda = \lambda_{\pm}$, solve $(A - \lambda I_2)v = 0$:

$$\begin{bmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0 \iff y = \lambda x.$$

Hence

$$E_{F,\lambda} = \operatorname{span}\left\{\begin{bmatrix} 1 \\ \lambda \end{bmatrix}\right\} \quad \text{for } \lambda \in \left\{\frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2}\right\}.$$

Extension (Tutorial 7, Question 4, Method 2)

Syllabus link: MH1201 Weeks 7–8 — Eigenvalues and eigenvectors.
Lemma (rotation eigenvalues over $\mathbb{R}$). Fix $\theta \in \mathbb{R}$ and define $R_\theta \in \mathcal{L}(\mathbb{R}^2)$ by

$$R_\theta(x,y) = (x\cos\theta - y\sin\theta,\ x\sin\theta + y\cos\theta).$$

If $\theta \notin \mathbb{Z}\pi$, then R_θ has no real eigenvalues.
Proof. Suppose for contradiction that $\lambda \in \mathbb{R}$ is an eigenvalue with eigenvector $u \neq 0$. Then $R_\theta(u) = \lambda u$ implies $R_\theta(u)$ is a scalar multiple of u, hence $R_\theta(u)$ is parallel to u. But geometrically R_θ rotates every nonzero vector by angle θ. A rotation maps a nonzero vector to a parallel vector if and only if the rotation angle is 0 or π modulo 2π, i.e. $\theta \in \mathbb{Z}\pi$. This contradicts $\theta \notin \mathbb{Z}\pi$. Hence no real eigenvalue exists. □

Extension (Tutorial 7, Question 6, Method 2)

Syllabus link: MH1201 Weeks 7–8 — Eigenvalues and eigenspaces as invariant subspaces.
Lemma (invariant decomposition yields eigenstructure). Let $T \in \mathcal{L}(\mathbb{R}^{2,2})$ be defined by

$$T(A) = A^{\mathsf{T}} + 2\,\mathrm{tr}(A)\,I_2.$$

Define

$$\mathrm{Sym}_2 = \{A : A^{\mathsf{T}} = A\}, \qquad \mathrm{Skew}_2 = \{A : A^{\mathsf{T}} = -A\}.$$

Then $\mathbb{R}^{2,2} = \mathrm{Sym}_2 \oplus \mathrm{Skew}_2$, and:

$$E_{T,-1} = \mathrm{Skew}_2, \qquad E_{T,5} = \mathbb{R}I_2, \qquad E_{T,1} = \{A \in \mathrm{Sym}_2 : \mathrm{tr}(A) = 0\}.$$

In particular, $\mathbb{R}^{2,2}$ has a basis of eigenvectors of T.
Proof. *Step 1: direct sum decomposition.* For any A, write

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

where $\frac{1}{2}(A + A^{\mathsf{T}}) \in \mathrm{Sym}_2$ and $\frac{1}{2}(A - A^{\mathsf{T}}) \in \mathrm{Skew}_2$. Also $\mathrm{Sym}_2 \cap \mathrm{Skew}_2 = \{0\}$, so $\mathbb{R}^{2,2} = \mathrm{Sym}_2 \oplus \mathrm{Skew}_2$.
Step 2: eigenvalue on Skew_2. If $A \in \mathrm{Skew}_2$, then $A^{\mathsf{T}} = -A$ and $\mathrm{tr}(A) = 0$, hence

$$T(A) = A^{\mathsf{T}} + 2\,\mathrm{tr}(A)I_2 = -A.$$

Thus every nonzero $A \in \mathrm{Skew}_2$ is an eigenvector with eigenvalue -1, and $E_{T,-1} = \mathrm{Skew}_2$.
Step 3: reduce on Sym_2 *and split by trace.* If $A \in \mathrm{Sym}_2$, then $A^{\mathsf{T}} = A$, so $T(A) = A + 2\,\mathrm{tr}(A)I_2$. Now decompose

$$\mathrm{Sym}_2 = \mathbb{R}I_2 \oplus \mathrm{Sym}_2^0, \qquad \mathrm{Sym}_2^0 \coloneqq \{A \in \mathrm{Sym}_2 : \mathrm{tr}(A) = 0\}.$$

This is a direct sum since $\mathrm{tr}(I_2) = 2 \neq 0$ and $\mathrm{tr}(\cdot) = 0$ on Sym_2^0. If $A \in \mathrm{Sym}_2^0$, then $T(A) = A$, so $A \in E_{T,1}$ and $E_{T,1} = \mathrm{Sym}_2^0$. If $A = tI_2$, then $\mathrm{tr}(A) = 2t$ and

$$T(A) = tI_2 + 2(2t)I_2 = 5tI_2,$$

so $\mathbb{R}I_2 \subseteq E_{T,5}$ and equality holds by the decomposition (no other symmetric trace-nonzero directions remain).
Step 4: eigenbasis. Choose bases of Skew_2, $\mathbb{R}I_2$, and Sym_2^0; their union is an eigenbasis because the sum is direct and spans $\mathbb{R}^{2,2}$. For example,

$$\mathrm{Skew}_2 = \mathrm{span}\left\{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right\}, \quad \mathbb{R}I_2 = \mathrm{span}\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\}, \quad \mathrm{Sym}_2^0 = \mathrm{span}\left\{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right\}.$$

□

Exam Focus

Weeks 7–8 exam patterns.

- **Change-of-coordinates:** compute $[I_V]_\beta^\gamma$ by stacking basis vectors in a standard basis and inverting once. Then convert coordinates via $[v]_\gamma = [I_V]_\beta^\gamma [v]_\beta$.
- **Change of representation:** use $[T]_\gamma^\gamma = [I_V]_\beta^\gamma [T]_\beta^\beta [I_V]_\gamma^\beta$.
- **Similarity proofs:** show $B = Q^{-1}AQ$; for invariants use $\det(Q^{-1}AQ) = \det(A)$ and $\operatorname{tr}(Q^{-1}AQ) = \operatorname{tr}(A)$.
- **Eigenvalues:** compute $p_T(\lambda) = \det([T] - \lambda I)$ and solve $p_T(\lambda) = 0$ over the *base field*.
- **Eigenspaces:** for each eigenvalue λ, solve $(T - \lambda I)v = 0$ (a kernel computation).

Extension (Tutorial 6, Question 5): existence of polynomial solutions via triangularity

Syllabus link: MH1201 Weeks 5–6 — Rank–Nullity and invertibility.
Theorem. Let $n \in \mathbb{N}_0$ and $Q \in \mathcal{P}_n(\mathbb{R})$. Then there exists $P \in \mathcal{P}_n(\mathbb{R})$ such that

$$Q = (X^2 + X)P'' + 2XP' + P(3).$$

Proof strategy. Define $L : \mathcal{P}_n(\mathbb{R}) \to \mathcal{P}_n(\mathbb{R})$ by $L(P) = (X^2 + X)P'' + 2XP' + P(3)$. Compute $L(X^k)$ for each k: the leading term of $L(X^k)$ has degree k with coefficient $k(k+1)$ (for $k \geq 1$) and $L(1) = 1$. Hence the matrix of L w.r.t. the standard basis $(1, X, \ldots, X^n)$ is *lower triangular* with diagonal entries $1, 2, 6, 12, \ldots, n(n+1)$, all nonzero. A triangular matrix with nonzero diagonal is invertible, so L is an isomorphism and in particular surjective. Therefore, for every Q, the equation $L(P) = Q$ has a solution.
Key technique: When an operator on a polynomial space has a triangular matrix representation with nonzero diagonal, it is automatically invertible. This is a common pattern on MH1201 exams.

Extension (Tutorial 6, Question 6): isomorphism between upper-triangular and symmetric matrices

Syllabus link: MH1201 Weeks 5–6 — Invertibility and isomorphisms.
Theorem. Let $V = \{M \in \mathbb{R}^{n\times n} \mid M \text{ upper triangular}\}$ and $W = \{M \in \mathbb{R}^{n\times n} \mid M \text{ symmetric}\}$. Then $V \cong W$.
Proof. Both V and W have dimension $\frac{n(n+1)}{2}$ (each is determined by the upper-triangular entries). An explicit isomorphism $\Phi : V \to W$ is given by "mirror the upper triangle":

$$\Phi(U)_{ij} := \begin{cases} u_{ij}, & i \leq j, \\ u_{ji}, & i > j. \end{cases}$$

This is linear (entrywise), and its inverse $\Psi : W \to V$ extracts the upper-triangular part: $\Psi(S)_{ij} = s_{ij}$ for $i \leq j$ and 0 otherwise. Both compositions are the identity on their respective spaces.

Worked Example

Worked Example (Tutorial 5, Question 4): substitution operator on polynomials. Let $Q \in \mathcal{P}(\mathbb{F})$ and define $T : \mathcal{P}_n(\mathbb{F}) \to \mathcal{P}_{n\cdot\deg(Q)}(\mathbb{F})$ by $T(P) = P(Q)$. Then T is linear because

$$T(P_1 + P_2) = (P_1 + P_2)(Q) = P_1(Q) + P_2(Q) = T(P_1) + T(P_2), \quad T(cP) = (cP)(Q) = cP(Q) = cT(P).$$

To verify T maps into $\mathcal{P}_{n\cdot\deg(Q)}(\mathbb{F})$: if $\deg(P) \leq n$ and $P \neq 0$, then $\deg(P(Q)) = \deg(P) \cdot \deg(Q) \leq n \cdot \deg(Q)$ (using the identity $\deg(P(Q)) = \deg(P) \cdot \deg(Q)$ for nonzero polynomials).

Worked Example

Worked Example (Tutorial 5, Question 2): linearity of the matrix-of-a-map operator. The map $M : \mathcal{L}(V, W) \to \mathbb{F}^{n\times m}$, $M(T) = [T]_\alpha^\beta$, is linear. **Proof sketch:** Column i of $M(T)$ is $[T(v_i)]_\beta$. Since the coordinate map $[\,\cdot\,]_\beta$ is linear,

$$M(S+T)_{*,i} = [(S+T)(v_i)]_\beta = [S(v_i)]_\beta + [T(v_i)]_\beta = M(S)_{*,i} + M(T)_{*,i}.$$

Similarly $M(aT) = aM(T)$. Hence M is linear. This result underlies the isomorphism $\mathcal{L}(V, W) \cong \mathbb{F}^{n\times m}$.

QRS@NTU_crsntu

4 Matrix Representations and Change of Basis

4.1 Ordered bases and coordinate vectors

Let $\mathbb{F}$ be a field, and let V be a finite-dimensional $\mathbb{F}$-vector space with $\dim(V) = n$.

Definition 4.1: Ordered basis and coordinate vector

An *ordered basis* of V is an ordered n-tuple $\beta = (b_1, \ldots, b_n)$ such that $\{b_1, \ldots, b_n\}$ is a basis of V. Given an ordered basis $\beta = (b_1, \ldots, b_n)$ and a vector $v \in V$, there exists a *unique* column vector

$$[v]_\beta = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{F}^{n,1} \quad \text{such that} \quad v = \sum_{i=1}^{n} c_i\, b_i.$$

We call $[v]_\beta$ the *coordinate vector* of v relative to β.

Ordering matters: the same set of vectors $\{b_1, b_2\}$ produces *different* coordinate maps depending on whether the ordered basis is (b_1, b_2) or (b_2, b_1). If $v = 3b_1 + 5b_2$, then $[v]_{(b_1,b_2)} = \binom{3}{5}$ but $[v]_{(b_2,b_1)} = \binom{5}{3}$. This is why we always specify "ordered basis" rather than just "basis" when working with coordinates.

Theorem 4.1: The coordinate map is a linear isomorphism

Fix an ordered basis $\beta = (b_1, \ldots, b_n)$ for V and define

$$[\cdot]_\beta : V \to \mathbb{F}^{n,1}, \qquad v \mapsto [v]_\beta.$$

Then $[\cdot]_\beta$ is a linear isomorphism.

Proof. Linearity follows from uniqueness of coordinates: if $v = \sum c_i b_i$ and $w = \sum d_i b_i$, then $v + w = \sum (c_i + d_i) b_i$ and $\lambda v = \sum (\lambda c_i) b_i$. Hence $[v + w]_\beta = [v]_\beta + [w]_\beta$ and $[\lambda v]_\beta = \lambda [v]_\beta$.

Bijectivity: surjective since any column (c_i) maps to $\sum c_i b_i$, and injective since $[v]_\beta = 0$ implies $v = \sum 0 \cdot b_i = 0$. □

This theorem tells us that the abstract vector space V is "the same" as $\mathbb{F}^{n,1}$ once we fix a basis—every linear-algebraic question about V can be translated to a concrete column-vector question in $\mathbb{F}^{n,1}$. The coordinate map $[\cdot]_\beta$ is the translator: it preserves addition, scalar multiplication, and all subspace relationships. This is the engine behind representing linear maps as matrices (next subsection).

Common Mistake: vectors vs. coordinates

$v \in V$ is *not* the same object as $[v]_\beta \in \mathbb{F}^{n,1}$. Always write the basis when you mean coordinates, and check whether your computation lives in V or in $\mathbb{F}^{n,1}$.

4.2 Matrix representation of a linear map

Let V, W be finite-dimensional $\mathbb{F}$-vector spaces with ordered bases $\beta = (b_1, \ldots, b_n)$ for V and $\gamma = (c_1, \ldots, c_m)$ for W.

Definition 4.2: Representation matrix $[T]_{\gamma\beta}$

Let $T : V \to W$ be linear. The *matrix representation* of T with respect to (β, γ) is the matrix

$$[T]_{\gamma\beta} \in \mathbb{F}^{m\times n} \quad \text{defined by} \quad \text{the } j\text{th column of } [T]_{\gamma\beta} \text{ is } [T(b_j)]_\gamma.$$

Equivalently,

$$[T]_{\gamma\beta} = \begin{bmatrix} | & & | \\ [T(b_1)]_\gamma & \cdots & [T(b_n)]_\gamma \\ | & & | \end{bmatrix}.$$

Why this construction is natural: we want a matrix A such that $A[v]_\beta = [T(v)]_\gamma$ for all v. By linearity, it suffices to pin down A's action on each standard basis vector e_j of $\mathbb{F}^{n,1}$ (which corresponds to $b_j \in V$). The j-th column of A records $[T(b_j)]_\gamma$, i.e. what T does to the j-th domain basis vector, expressed in the codomain basis. Everything else follows by linearity.

Exam Focus: size check for $[T]_{\gamma\beta}$

The matrix $[T]_{\gamma\beta}$ has size $\dim(W) \times \dim(V)$, *not* $\dim(V) \times \dim(W)$. The number of rows equals the dimension of the codomain, and the number of columns equals the dimension of the domain. Quick mnemonic: it multiplies a $\dim(V)$-vector on the right and produces a $\dim(W)$-vector on the left.

Theorem 4.2: Fundamental coordinate identity

For every $v \in V$,

$$[T(v)]_\gamma \;=\; [T]_{\gamma\beta}\,[v]_\beta.$$

Proof. Write $v = \sum_{j=1}^{n} x_j b_j$, so $[v]_\beta = (x_j)_{j=1}^{n}$. By linearity,

$$T(v) = \sum_{j=1}^{n} x_j\, T(b_j).$$

Applying $[\cdot]_\gamma$ and using linearity of the coordinate map,

$$[T(v)]_\gamma = \sum_{j=1}^{n} x_j\, [T(b_j)]_\gamma.$$

But the right-hand side is exactly the matrix-vector product $[T]_{\gamma\beta}[v]_\beta$ by the column definition of $[T]_{\gamma\beta}$. □

This identity is the single most important equation in the chapter. It says: "to find the γ-coordinates of $T(v)$, multiply the representation matrix by the β-coordinates of v." Every subsequent result (composition, change of basis, similarity) is built by chaining this identity.

Worked Example (Tutorial 7, Question 1): one-dimensional domain ⇒ one column

Let V be finite-dimensional over $\mathbb{F}$, fix $v \in V$, and define $T_v : \mathbb{F} \to V$ by $T_v(\lambda) = \lambda v$. Let β be an ordered basis for V, and use the ordered basis $(1_{\mathbb{F}})$ for $\mathbb{F}$.
Since $\dim(\mathbb{F}) = 1$, the matrix $[T_v]_{\beta,(1_{\mathbb{F}})}$ is $n \times 1$ and its unique column is

$$[T_v(1_{\mathbb{F}})]_\beta = [v]_\beta.$$

Hence

$$[T_v]_{\beta,(1_{\mathbb{F}})} = [v]_\beta.$$

This example illustrates a useful principle: the representation matrix of a map *from* $\mathbb{F}$ is just a single column—the coordinate vector of the image of $1_{\mathbb{F}}$. Conversely, a map *to* $\mathbb{F}$ (i.e. a linear functional) is represented by a single *row*.

Worked Example (Tutorial 7, Question 2): nested span condition ⇒ upper-triangular

Let V, W be n-dimensional over $\mathbb{F}$, with ordered bases $\beta = (v_1, \ldots, v_n)$ and $\gamma = (w_1, \ldots, w_n)$. Suppose $T \in \mathcal{L}(V, W)$ satisfies

$$\forall k \in \{1, \ldots, n\} : \quad T(v_k) \in \operatorname{span}(w_1, \ldots, w_k).$$

Let $A = [T]_{\gamma\beta} = (a_{ik})$. By definition, the kth column is $[T(v_k)]_\gamma = (a_{1k}, \ldots, a_{nk})^{\mathsf{T}}$.
The condition $T(v_k) \in \operatorname{span}(w_1, \ldots, w_k)$ means the coefficients of $w_{k+1}, \ldots, w_n$ are 0, i.e.

$$a_{ik} = 0 \quad \text{for all } i > k.$$

Thus all entries below the diagonal are 0, so A is upper-triangular.

Exam Focus: when does a matrix representation come out upper/lower-triangular?

The key idea: the structure of $[T]_{\gamma\beta}$ depends entirely on how $T(b_j)$ relates to the *initial segments* of γ.

- **Upper-triangular:** $T(v_k) \in \operatorname{span}(w_1, \ldots, w_k)$ for each $k \implies [T]_{\gamma\beta}$ upper-triangular.
- **Lower-triangular:** $T(v_k) \in \operatorname{span}(w_k, \ldots, w_n)$ for each $k \implies [T]_{\gamma\beta}$ lower-triangular.
- **Diagonal:** If $T(v_k) = \lambda_k w_k$ for each $k \implies [T]_{\gamma\beta} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

This is why choosing a "good basis" (e.g. one that diagonalizes T) simplifies computations enormously.

4.3 Change of coordinates and transition matrices

Now let V be finite-dimensional with $\dim(V) = n$, and let $\beta = (b_1, \ldots, b_n)$ and $\gamma = (c_1, \ldots, c_n)$ be ordered bases for V.

Definition 4.3: Change-of-coordinate matrix

The *change-of-coordinate matrix from β to γ* is the matrix

$$[I_V]_{\gamma\beta} \in \mathbb{F}^{n\times n},$$

i.e. the matrix representation of the identity map $I_V : V \to V$ with domain basis β and codomain basis γ.

Why use the identity map? We want to translate coordinates without changing the underlying vector. The identity map I_V does exactly that: $I_V(v) = v$, but the input is read in β-coordinates and the output is written in γ-coordinates. Hence $[I_V]_{\gamma\beta}$ converts β-coordinates to γ-coordinates, which is why it is called a "change-of-coordinate" matrix. The j-th column of $[I_V]_{\gamma\beta}$ is $[I_V(b_j)]_\gamma = [b_j]_\gamma$, i.e. the γ-coordinates of the j-th β-basis vector.

Key Formula: coordinate conversion

For every $v \in V$,

$$[v]_\gamma = [I_V]_{\gamma\beta}\,[v]_\beta.$$

In particular, $[I_V]_{\beta\gamma} = \left([I_V]_{\gamma\beta}\right)^{-1}$ and

$$[I_V]_{\gamma\beta}\,[I_V]_{\beta\gamma} = I_n = [I_V]_{\beta\gamma}\,[I_V]_{\gamma\beta}.$$

The inverse relation $[I_V]_{\beta\gamma} = ([I_V]_{\gamma\beta})^{-1}$ is automatic: $[I_V]_{\gamma\beta}$ represents the identity from $(\beta \to \gamma)$, and composing the identity from $(\gamma \to \beta)$ and then $(\beta \to \gamma)$ gives the identity on γ-coordinates. Hence their product is I_n, and the two matrices are mutual inverses.

Worked Example (Lecture Week 7): computing $[I_V]_{\alpha\beta}$ from basis relations

Let $\alpha = (a_1, a_2, a_3)$ and $\beta = (b_1, b_2, b_3)$ be ordered bases for V with

$$b_1 = 2a_1 - a_2 + a_3, \qquad b_2 = 3a_2 + a_3, \qquad b_3 = -3a_1 + 2a_3.$$

Then the columns of $[I_V]_{\alpha\beta}$ are $[b_j]_\alpha$. Hence

$$[I_V]_{\alpha\beta} = [\,[b_1]_\alpha \quad [b_2]_\alpha \quad [b_3]_\alpha] = \begin{bmatrix} 2 & 0 & -3 \\ -1 & 3 & 0 \\ 1 & 1 & 2 \end{bmatrix}.$$

Also,

$$[b_1 - 2b_2 + 2b_3]_\alpha = [b_1]_\alpha - 2[b_2]_\alpha + 2[b_3]_\alpha = \begin{bmatrix} -4 \\ -7 \\ 3 \end{bmatrix}.$$

Note how the linearity of the coordinate map was used in the second part: we did *not* need to expand $b_1 - 2b_2 + 2b_3$ in terms of a_i from scratch. Instead, we simply took the linear combination of the already-computed columns. This is a time-saving technique on exams.

Key Formula: computing $[I_V]_{\gamma\beta}$ via a standard basis

Let σ be a "standard" ordered basis for V (when available, e.g. $\mathbb{F}^n$). Then

$$[I_V]_{\sigma\beta} = [[b_1]_\sigma \quad \cdots \quad [b_n]_\sigma], \qquad [I_V]_{\sigma\gamma} = [[c_1]_\sigma \quad \cdots \quad [c_n]_\sigma],$$

and

$$[I_V]_{\gamma\beta} = \left([I_V]_{\sigma\gamma}\right)^{-1}[I_V]_{\sigma\beta}.$$

The logic: we cannot directly convert from β to γ when neither is standard. Instead, we route through the standard basis σ: first express β-vectors in σ-coordinates (easy), then convert σ-coordinates to γ-coordinates by inverting $[I_V]_{\sigma\gamma}$. The composition gives $[I_V]_{\gamma\beta} = [I_V]_{\gamma\sigma}[I_V]_{\sigma\beta} = ([I_V]_{\sigma\gamma})^{-1}[I_V]_{\sigma\beta}$.

Worked Example (Lecture Week 7): change of basis in $\mathbb{R}^2$

Let $\beta = ((-9,1),(-5,-1))$ and $\gamma = ((1,-4),(3,-5))$ be ordered bases of $\mathbb{R}^2$, and let σ be the standard basis.
Compute

$$[I_{\mathbb{R}^2}]_{\sigma\beta} = \begin{bmatrix} -9 & -5 \\ 1 & -1 \end{bmatrix}, \qquad [I_{\mathbb{R}^2}]_{\sigma\gamma} = \begin{bmatrix} 1 & 3 \\ -4 & -5 \end{bmatrix}.$$

Then

$$[I_{\mathbb{R}^2}]_{\gamma\beta} = \left([I_{\mathbb{R}^2}]_{\sigma\gamma}\right)^{-1}[I_{\mathbb{R}^2}]_{\sigma\beta} = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}, \qquad [I_{\mathbb{R}^2}]_{\beta\gamma} = \left([I_{\mathbb{R}^2}]_{\gamma\beta}\right)^{-1} = \begin{bmatrix} -\frac{3}{2} & -2 \\ \frac{5}{2} & 3 \end{bmatrix}.$$

Exam Focus: fast row-reduction trick for $[I_V]_{\gamma\beta}$

In $\mathbb{F}^n$ with standard basis σ, you can compute $[I_V]_{\gamma\beta}$ without explicitly inverting: row-reduce the augmented matrix

$$[\,[I_V]_{\sigma\gamma} \mid [I_V]_{\sigma\beta}\,] \longrightarrow [\,I_n \mid [I_V]_{\gamma\beta}\,].$$

Similarly, swapping β, γ gives $[I_V]_{\beta\gamma}$.

This trick works because row-reducing $[P_\gamma \mid P_\beta]$ to $[I_n \mid \cdot\,]$ is equivalent to left-multiplying both blocks by P_γ^{-1}, and the right block becomes $P_\gamma^{-1}P_\beta = [I_V]_{\gamma\beta}$. On exams this avoids a separate matrix inversion step, saving time and reducing errors.

Worked micro-example: row-reduction approach in $\mathbb{R}^2$

Using the bases from the previous example, form

$$[\,[I]_{\sigma\gamma} \mid [I]_{\sigma\beta}\,] = \left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array}\right].$$

$R_2 \leftarrow R_2 + 4R_1$:

$$\left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 7 & -35 & -21 \end{array}\right].$$

$R_2 \leftarrow \frac{1}{7}R_2$:

$$\left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 1 & -5 & -3 \end{array}\right].$$

$R_1 \leftarrow R_1 - 3R_2$:

$$\left[\begin{array}{cc|cc} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array}\right].$$

Reading off the right block: $[I_V]_{\gamma\beta} = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}$, matching the earlier result.

Common Mistake: "from β to γ" direction error

By definition, $[I_V]_{\gamma\beta}$ is the matrix of I_V with *input coordinates in* β and *output coordinates in* γ. So it satisfies $[v]_\gamma = [I_V]_{\gamma\beta}[v]_\beta$. A common error is to use the inverse direction unintentionally (especially when reading off columns).

Exam Focus: "read off columns" rule for $[I_V]_{\gamma\beta}$

The j-th column of $[I_V]_{\gamma\beta}$ is $[b_j]_\gamma$, i.e. the γ-coordinates of the j-th vector in β. *Not* the other way around. If you are given the α-coordinates of β-vectors directly (as in the Lecture Week 7 example), you can write down the columns immediately. Otherwise, route through a standard basis: columns of $[I_V]_{\sigma\beta}$ are $[b_j]_\sigma$ (easy to write down), then left-multiply by $([I_V]_{\sigma\gamma})^{-1}$.

4.4 Representation matrices under basis change

Let $T \in \mathcal{L}(V)$ and let β, γ be ordered bases for V.

Theorem 4.3: Change-of-basis formula for matrix representations

$$[T]_{\gamma\gamma} = [I_V]_{\gamma\beta}\,[T]_{\beta\beta}\,[I_V]_{\beta\gamma}.$$

Equivalently, if we set $Q \coloneqq [I_V]_{\beta\gamma}$ (so $Q^{-1} = [I_V]_{\gamma\beta}$), then

$$[T]_{\gamma\gamma} = Q^{-1}\,[T]_{\beta\beta}\,Q.$$

Proof. For any $v \in V$,

$$[T(v)]_\gamma = [I_V]_{\gamma\beta}\,[T(v)]_\beta = [I_V]_{\gamma\beta}\,[T]_{\beta\beta}\,[v]_\beta = [I_V]_{\gamma\beta}\,[T]_{\beta\beta}\,[I_V]_{\beta\gamma}\,[v]_\gamma.$$

Since this holds for all $[v]_\gamma \in \mathbb{F}^{n,1}$, the representing matrices must be equal: $[T]_{\gamma\gamma} = [I_V]_{\gamma\beta}[T]_{\beta\beta}[I_V]_{\beta\gamma}$. □

Read the sandwich formula right-to-left: (1) convert γ-input to β-coordinates via $[I_V]_{\beta\gamma}$, (2) apply T in β-coordinates via $[T]_{\beta\beta}$, (3) convert the β-output back to γ-coordinates via $[I_V]_{\gamma\beta}$. This explains *why* the formula has the form $Q^{-1}AQ$: we "wrap" the β-representation with coordinate translators.

Exam Focus: which Q to use in $Q^{-1}[T]_{\beta\beta}Q$

In the formula $[T]_{\gamma\gamma} = Q^{-1}[T]_{\beta\beta}Q$, the matrix $Q = [I_V]_{\beta\gamma}$ has columns $[c_j]_\beta$ (the β-coordinates of the γ-basis vectors). If both bases are given in standard coordinates, then:

$$Q = [I_V]_{\beta\gamma} = ([I_V]_{\sigma\beta})^{-1}[I_V]_{\sigma\gamma} = P_\beta^{-1}P_\gamma,$$

where P_β and P_γ stack the respective basis vectors as columns in the standard basis. When β happens to be the standard basis σ, this simplifies to $Q = P_\gamma$ (columns are γ-vectors in standard coordinates), and $[T]_{\gamma\gamma} = P_\gamma^{-1}[T]_{\sigma\sigma}P_\gamma$.

Worked Example (Tutorial 7, Question 7 style): similarity via $P^{-1}AP$ in $\mathbb{R}^3$

Define $T \in \mathcal{L}(\mathbb{R}^3)$ by $T(v) = Av$ where $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 3 \\ 0 & -1 & 0 \end{bmatrix}$. Let $\gamma = (g_1, g_2, g_3)$ with $g_1 = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$, $g_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$, $g_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Here $\beta = \sigma$ (standard basis), so $Q = [I_V]_{\sigma\gamma} = P_\gamma = [\, g_1 \; g_2 \; g_3 \,] = \begin{bmatrix} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. Then

$$[T]_{\gamma\gamma} = P_\gamma^{-1}AP_\gamma.$$

Since P_γ is upper-triangular with diagonal $(-1, 1, 1)$, its inverse is easy to compute:

$$P_\gamma^{-1} = \begin{bmatrix} -1 & 2 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Computing AP_γ first, then left-multiplying by P_γ^{-1}:

$$AP_\gamma = \begin{bmatrix} -2 & 5 & 3 \\ -1 & 3 & 5 \\ 0 & -1 & -1 \end{bmatrix}, \qquad [T]_{\gamma\gamma} = P_\gamma^{-1}(AP_\gamma) = \begin{bmatrix} 0 & 2 & 8 \\ -1 & 4 & 6 \\ 0 & -1 & -1 \end{bmatrix}.$$

Sanity check: $\operatorname{tr}([T]_{\gamma\gamma}) = 0 + 4 + (-1) = 3 = \operatorname{tr}(A) = 2 + 1 + 0$ ✓ and det should also be preserved (both equal -11).

4.5 Composition, identity, and inverse maps

Let U, V, W be finite-dimensional with ordered bases α, β, γ respectively. Let $S \in \mathcal{L}(U, V)$ and $T \in \mathcal{L}(V, W)$.

Key Formula: composition corresponds to multiplication

$$[T \circ S]_{\gamma\alpha} \;=\; [T]_{\gamma\beta}\,[S]_{\beta\alpha}.$$

Also, $[I_V]_{\beta\beta} = I_n$ for any ordered basis β of V.

The formula requires the "inner" basis β to match: S outputs β-coordinates, and T reads β-coordinates. If the intermediate bases disagree, the formula fails. This matching condition is analogous to matrix multiplication requiring compatible inner dimensions.

Theorem 4.4: Inverse map $\leftrightarrow$ inverse matrix

Let $T \in \mathcal{L}(V, W)$ with $\dim(V) = \dim(W) = n$, and let β and γ be ordered bases of V and W. Then T is invertible iff $[T]_{\gamma\beta} \in \mathbb{F}^{n\times n}$ is invertible, and in that case

$$[T^{-1}]_{\beta\gamma} = \big([T]_{\gamma\beta}\big)^{-1}.$$

Proof. If T is invertible, then $T^{-1} \circ T = I_V$ and $T \circ T^{-1} = I_W$. Taking representation matrices,

$$[I_V]_{\beta\beta} = [T^{-1} \circ T]_{\beta\beta} = [T^{-1}]_{\beta\gamma}[T]_{\gamma\beta}, \qquad [I_W]_{\gamma\gamma} = [T \circ T^{-1}]_{\gamma\gamma} = [T]_{\gamma\beta}[T^{-1}]_{\beta\gamma}.$$

Thus $[T]_{\gamma\beta}$ has a two-sided inverse $[T^{-1}]_{\beta\gamma}$ and is invertible, with $[T^{-1}]_{\beta\gamma} = ([T]_{\gamma\beta})^{-1}$.

Conversely, if $[T]_{\gamma\beta}$ is invertible, let $A = ([T]_{\gamma\beta})^{-1}$ and define $S \in \mathcal{L}(W, V)$ by $[S]_{\beta\gamma} = A$ (existence/uniqueness of a linear map given its matrix). Then $[S \circ T]_{\beta\beta} = I_n$ and $[T \circ S]_{\gamma\gamma} = I_n$, hence $S \circ T = I_V$ and $T \circ S = I_W$, so T is invertible and $S = T^{-1}$. □

Note the subscript swap: $[T^{-1}]_{\beta\gamma}$ has domain basis γ (the codomain basis of T) and codomain basis β (the domain basis of T). This makes sense because T^{-1} maps $W \to V$, reversing the roles of domain and codomain.

Worked Example (Lecture Week 7): using a "good basis" to simplify powers

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$T(x, y, z) = (2x + y,\ 5y + 3z,\ 8z).$$

Let $\beta = \big((1, 0, 0), (1, 3, 0), (1, 6, 6)\big)$.
Step 1: compute $[T]_{\beta\beta}$. Compute T on basis vectors and express in β-coordinates (equivalently, compute $[T]_{\beta\beta} = [I]_{\beta\sigma}[T]_{\sigma\sigma}[I]_{\sigma\beta}$). One finds

$$[T]_{\beta\beta} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 8 \end{bmatrix}.$$

Step 2: compute $T^4(0, -3, 6)$. First solve for $[v]_\beta$ where $v = (0, -3, 6)$:

$$[v]_\beta = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}.$$

Then

$$[T^4(v)]_\beta = \big([T]_{\beta\beta}\big)^4 [v]_\beta = \begin{bmatrix} 2^4 & 0 & 0 \\ 0 & 5^4 & 0 \\ 0 & 0 & 8^4 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 32 \\ -1875 \\ 4096 \end{bmatrix}.$$

Converting back to standard coordinates gives

$$T^4(0, -3, 6) = (2253,\ 18951,\ 24576).$$

This example demonstrates the key advantage of a diagonalizing basis: powers of a diagonal matrix are trivial to compute ($D^k = \mathrm{diag}(\lambda_1^k, \ldots, \lambda_n^k)$), whereas computing A^4 for a general matrix requires three matrix multiplications. The strategy is always: (1) change to the good basis, (2) do the easy computation, (3) change back.

Extension (Lecture Week 7): the Fundamental Theorem of Finite-Dimensional Vector Spaces

Syllabus link: MH1201 Week 7 — Matrix representations as isomorphisms.
Theorem. Let U, V, W be finite-dimensional over $\mathbb{F}$ with ordered bases α, β, γ. Let $S \in \mathcal{L}(U, V)$ and $T \in \mathcal{L}(V, W)$. Then the "matrix-of-a-map" operators

$$[\cdot]_{\beta\alpha} : \mathcal{L}(U, V) \to \mathbb{F}^{\dim V \times \dim U}, \quad [\cdot]_{\gamma\beta} : \mathcal{L}(V, W) \to \mathbb{F}^{\dim W \times \dim V}$$

are all *linear isomorphisms*. In particular,

$$\dim(\mathcal{L}(V, W)) = \dim(V) \cdot \dim(W).$$

Moreover, the diagram

$$\mathcal{L}(U, V) \xrightarrow{T\circ(\cdot)} \mathcal{L}(U, W) \qquad \longleftrightarrow \qquad \mathbb{F}^{\dim V \times \dim U} \xrightarrow{[T]_{\gamma\beta}\cdot(\cdot)} \mathbb{F}^{\dim W \times \dim U}$$

commutes: taking matrices turns composition into multiplication.
Proof sketch. *Linearity:* For $S_1, S_2 \in \mathcal{L}(U, V)$ and $\lambda \in \mathbb{F}$, column j of $[S_1 + S_2]_{\beta\alpha}$ is $[(S_1 + S_2)(a_j)]_\beta = [S_1(a_j)]_\beta + [S_2(a_j)]_\beta$, so $[S_1 + S_2]_{\beta\alpha} = [S_1]_{\beta\alpha} + [S_2]_{\beta\alpha}$. Similarly for scalar multiplication. *Injective:* $[S]_{\beta\alpha} = 0$ implies $S(a_j) = 0$ for all j, hence $S = 0$. *Surjective:* Given any $A \in \mathbb{F}^{m\times n}$, define S on basis α by requiring $[S(a_j)]_\beta$ to be the j-th column of A, then extend by linearity. □

4.6 Similarity viewpoint and basis-independence statements

Definition 4.4: Similarity of square matrices

Let $n \in \mathbb{N}$ and let $A, B \in \mathbb{F}^{n\times n}$. We say that A is *similar* to B, and write $A \sim B$, if there exists an invertible $Q \in \mathbb{F}^{n\times n}$ such that

$$B = Q^{-1}AQ.$$

Similarity is the matrix-level translation of "same linear operator, different basis." If $A = [T]_{\beta\beta}$ and $B = [T]_{\gamma\gamma}$ for the same $T \in \mathcal{L}(V)$, then $B = Q^{-1}AQ$ with $Q = [I_V]_{\beta\gamma}$. Thus the similarity class of a matrix captures all the basis-independent properties of the underlying operator.

Theorem 4.5: Similarity is an equivalence relation

The relation $\sim$ on $\mathbb{F}^{n\times n}$ is reflexive, symmetric, and transitive.

Proof. Reflexive: $A = I^{-1}AI$ so $A \sim A$.
Symmetric: if $B = Q^{-1}AQ$ with Q invertible, then $A = QBQ^{-1} = (Q^{-1})^{-1}B(Q^{-1})$, so $B \sim A$.
Transitive: if $B = Q^{-1}AQ$ and $C = R^{-1}BR$ with Q, R invertible, then

$$C = R^{-1}(Q^{-1}AQ)R = (QR)^{-1}A(QR),$$

so $A \sim C$. □

Theorem 4.6: Similarity invariants: determinant and trace

If $A \sim B$, then $\det(A) = \det(B)$ and $\operatorname{tr}(A) = \operatorname{tr}(B)$.

Proof. Let $B = Q^{-1}AQ$ with Q invertible. Then

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

For trace, use cyclicity $\operatorname{tr}(XY) = \operatorname{tr}(YX)$ (valid for square matrices of the same size):

$$\operatorname{tr}(B) = \operatorname{tr}(Q^{-1}AQ) = \operatorname{tr}(AQQ^{-1}) = \operatorname{tr}(A).$$

□

Beyond determinant and trace, similarity also preserves: eigenvalues (with multiplicities), the characteristic polynomial, rank, and nullity. On exams, trace and determinant are the fastest invariants to check. If $\operatorname{tr}(A) \neq \operatorname{tr}(B)$ or $\det(A) \neq \det(B)$, you can immediately conclude $A \not\sim B$ without further work.

Exam Focus: proving two matrices are NOT similar

To show $A \not\sim B$, it suffices to exhibit *one* similarity invariant that differs. The two cheapest invariants to compute are:

- tr: sum of diagonal entries (no computation beyond addition).
- det: for 2×2 matrices this is $ad - bc$.

If both match, try eigenvalues (roots of the characteristic polynomial). To show $A \sim B$, you typically must *find* an explicit invertible Q with $B = Q^{-1}AQ$, or show both represent the same operator.

Theorem 4.7: Matrix representations of one operator are similar

Let V be finite-dimensional and $T \in \mathcal{L}(V)$. For any ordered bases β, γ of V,

$$[T]_{\beta\beta} \sim [T]_{\gamma\gamma}.$$

Proof. By Theorem 3.10, $[T]_{\gamma\gamma} = Q^{-1}[T]_{\beta\beta}Q$ with $Q = [I_V]_{\beta\gamma}$ invertible. Hence $[T]_{\beta\beta} \sim [T]_{\gamma\gamma}$. □

Extension (Tutorial 7, Question 3): infinite-dimensional phenomena

Syllabus link: MH1201 Weeks 7–8 — Eigenvalues and dimension-dependent results.
Observation. The right-shift operator $\sigma_{\rightarrow} \in \mathcal{L}(\mathbb{R}^{\infty})$, defined by $\sigma_{\rightarrow}(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$, is injective but *not* surjective. This does not contradict the finite-dimensional result "injective $\Leftrightarrow$ surjective when $\dim(V) = \dim(W)$" because $\mathbb{R}^{\infty}$ is *infinite-dimensional*; rank–nullity (and hence the equal-dimension criterion) does not apply.
Moreover, $\sigma_{\rightarrow}$ has *no* real eigenvalues. If $\sigma_{\rightarrow}(x) = \lambda x$ for $x \neq 0$, comparing coordinates gives $0 = \lambda x_1$ and $x_k = \lambda x_{k+1}$ for all $k \geq 1$. Whether $\lambda = 0$ or $\lambda \neq 0$, an induction forces $x_k = 0$ for all k, contradicting $x \neq 0$.
Takeaway: Finite-dimensionality is an essential hypothesis for (i) the injective $\Leftrightarrow$ surjective equivalence, and (ii) the existence of eigenvalues guaranteed by the characteristic polynomial.

4.7 Common workflows (what to do on exams)

Exam Focus: three high-frequency workflows

1. **Compute $[T]_{\gamma\beta}$ from $T(b_i)$.** Compute $T(b_i) \in W$, then write each as a γ-linear combination; those coefficient columns form $[T]_{\gamma\beta}$.
2. **Compute a change-of-coordinate matrix.** To find $[I_V]_{\gamma\beta}$, write each b_i as a γ-linear combination; columns are $[b_i]_\gamma$. If both are given in standard coordinates, use
$$[I_V]_{\gamma\beta} = \left([I_V]_{\sigma\gamma}\right)^{-1}[I_V]_{\sigma\beta} \quad \text{or} \quad [\,[I_V]_{\sigma\gamma} \mid [I_V]_{\sigma\beta}\,] \to [\,I \mid [I_V]_{\gamma\beta}\,].$$
3. **Fast evaluation of $T(v)$ using matrices.** Convert v into coordinates in the domain basis ($[v]_\beta$), multiply by $[T]_{\gamma\beta}$ to get $[T(v)]_\gamma$, then (if needed) convert back to a concrete vector.

Exam Focus: computing $T^k(v)$ via a diagonalizing basis

If you can find an ordered basis β such that $[T]_{\beta\beta}$ is diagonal (or at least triangular), then:

1. Compute $[v]_\beta$ by solving $P_\beta[v]_\beta = v$ (where P_β stacks β-vectors as columns).
2. Compute $[T^k(v)]_\beta = ([T]_{\beta\beta})^k[v]_\beta$ — trivial for diagonal matrices.
3. Convert back: $T^k(v) = P_\beta[T^k(v)]_\beta$.

This is the strategy used in the "good basis" example ($T(x,y,z) = (2x+y, 5y+3z, 8z)$), and it generalizes to any diagonalizable operator.

Common Mistake: mixing up domain/codomain bases in $[T]_{\gamma\beta}$

$[T]_{\gamma\beta}$ means: *input coordinates in β (domain)* and *output coordinates in γ (codomain)*. A fast sanity check: the matrix size must be $(\dim W) \times (\dim V)$.

Common Mistake: forgetting the intermediate basis in composition

When computing $[T \circ S]_{\gamma\alpha} = [T]_{\gamma\beta}[S]_{\beta\alpha}$, the intermediate basis β must be the *same* in both matrices. If S is represented w.r.t. β' and T w.r.t. β, and $\beta' \neq \beta$, you cannot simply multiply the matrices. You must first insert a change-of-coordinate matrix: $[T]_{\gamma\beta}[I_V]_{\beta\beta'}[S]_{\beta'\alpha}$.

5 Eigenvalues, Diagonalization, and Jordan Canonical Form

Throughout, let $\mathbb{F}$ be a field (typically $\mathbb{R}$ or $\mathbb{C}$), and let V be an $\mathbb{F}$-vector space. Write $\mathcal{L}(V)$ for the space of linear operators $T : V \to V$ and I_V for the identity operator on V.

5.1 Eigenvalues, eigenvectors, and eigenspaces

Definition 5.1: Eigenvector and eigenvalue

Let $T \in \mathcal{L}(V)$. A vector $v \in V$ is an *eigenvector* of T if

$$v \neq 0 \quad \text{and} \quad \exists\, \lambda \in \mathbb{F} \text{ such that } T(v) = \lambda v.$$

Any such λ is called an *eigenvalue* of T (field-specific: λ must lie in the base field $\mathbb{F}$).

Geometrically, an eigenvector is a nonzero vector whose direction is preserved (or reversed) by T: the operator merely *scales* it by the factor λ. Every other vector gets "rotated" or "sheared" by T, but eigenvectors stay on their own line through the origin. This is why eigenvectors are so useful—they identify the directions along which T acts most simply.

Note that the eigenvalue λ associated to an eigenvector v is *unique*: if $T(v) = \lambda v = \mu v$ with $v \neq 0$, then $(\lambda - \mu)v = 0$ forces $\lambda = \mu$.

Definition 5.2: Eigenspace

Let $T \in \mathcal{L}(V)$ and let $\lambda \in \mathbb{F}$. The *eigenspace* of T for λ is

$$E_{T,\lambda} := \ker(T - \lambda I_V).$$

Then the set of eigenvectors for λ is $E_{T,\lambda} \setminus \{0\}$.

The eigenspace $E_{T,\lambda}$ is always a *subspace* of V (it is the kernel of the linear map $T - \lambda I_V$). It collects all vectors that T scales by λ, together with the zero vector. The scalar λ is an eigenvalue of T precisely when $E_{T,\lambda}$ is nontrivial (i.e. contains a nonzero vector).

Theorem 5.1: Equivalent characterizations of an eigenvalue

Let $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. The following are equivalent:

(a) λ is an eigenvalue of T.

(b) $E_{T,\lambda} \neq \{0\}$.

(c) $T - \lambda I_V$ is not injective.

Proof. (a)⇔(b): by definition, λ is an eigenvalue iff $\exists\, v \neq 0$ with $(T - \lambda I_V)(v) = 0$, i.e. $\ker(T - \lambda I_V) \neq \{0\}$.
(b)⇔(c): a linear map is injective iff its kernel is $\{0\}$. □

When V is finite-dimensional, we can add more equivalences to this list. In that case, $T - \lambda I_V$ is not injective ⇔ $T - \lambda I_V$ is not surjective ⇔ $T - \lambda I_V$ is not invertible ⇔ $\det([T]_\beta^\beta - \lambda I_n) = 0$. The last condition leads to the characteristic polynomial (next subsection).

Theorem 5.2: Zero eigenvalue ⇔ non-injectivity

Let $T \in \mathcal{L}(V)$. Then 0 is an eigenvalue of T if and only if T is not injective.

Proof. Apply Theorem 5.1 with $\lambda = 0$: 0 is an eigenvalue iff $\ker(T) \neq \{0\}$ iff T is not injective. □

This gives a quick test: if T has trivial kernel, then 0 is *not* an eigenvalue. Conversely, any non-injective operator automatically has 0 as an eigenvalue. In finite dimensions, 0 is an eigenvalue iff $\det([T]_\beta^\beta) = 0$.

Worked Example (Lecture Week 9): eigenvalues of $0_{V\to V}$ and I_V

Let $T = 0_{V\to V}$. Then $T(v) = 0$ for all v. The eigenvalue equation is $0 = \lambda v$ for some $v \neq 0$, which forces $\lambda = 0$. So 0 is the *only* eigenvalue of $0_{V\to V}$ and $E_{T,0} = V$.
Let $T = I_V$. Then $I_V(v) = \lambda v$ with $v \neq 0$ forces $\lambda = 1$. Hence 1 is the *only* eigenvalue of I_V and $E_{I_V,1} = V$.

Worked Example (Lecture Week 8): rotation matrix over $\mathbb{R}$ vs $\mathbb{C}$

Let $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and define $T(v) = Av$. Then $p_T(\lambda) = \det(A - \lambda I_2) = \lambda^2 + 1$.

- Over $\mathbb{R}$: $\lambda^2 + 1 = 0$ has no real roots, so T has *no* real eigenvalues.
- Over $\mathbb{C}$: $\lambda^2 + 1 = 0$ gives $\lambda = \pm i$, so T has eigenvalues i and $-i$.

This is the same matrix acting on the same vectors, but the available eigenvalues depend entirely on the base field.

Worked Example (Lecture Week 8): spotting eigenvectors without computation

Let $T(v) = Av$ on $\mathbb{R}^3$ where $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$.

- Let $u = (1, 1, 1)^\mathsf{T}$. Then $Au = (6, 6, 6)^\mathsf{T} = 6u$, so u is an eigenvector for $\lambda = 6$.
- For eigenvalue 0: pick any nonzero w with $(1, 2, 3) \cdot w = 0$, e.g. $w = (2, -1, 0)^\mathsf{T}$. Then $Aw = 0 = 0 \cdot w$, so w is an eigenvector for $\lambda = 0$.

Key insight: Since $\operatorname{rank}(A) = 1$, we have $\dim(\ker(A)) = 2$, so $\lambda = 0$ has geometric multiplicity 2. We found eigenvectors for two distinct eigenvalues without computing any determinant.

Common Mistake: eigenvalues depend on the base field

A linear operator on a real vector space may have *no* eigenvalues over $\mathbb{R}$ even though it has eigenvalues over $\mathbb{C}$. Always check whether the course is working over $\mathbb{R}$ or $\mathbb{C}$ before claiming eigenvalues exist.

Common Mistake: the zero vector is never an eigenvector

By definition, eigenvectors are *nonzero*. Although $T(0) = \lambda \cdot 0$ for every λ, the vector 0 is not an eigenvector. The eigenspace $E_{T,\lambda}$ *includes* 0 (it is a subspace), but the set of *eigenvectors* for λ is $E_{T,\lambda} \setminus \{0\}$.

5.2 Characteristic polynomial and computing eigenvalues

Now assume V is finite-dimensional over $\mathbb{F}$ with $\dim(V) = n$, and fix an ordered basis β of V.

Definition 5.3: Characteristic polynomial function

Let $T \in \mathcal{L}(V)$ with $\dim(V) = n$. The *characteristic polynomial function* of T is

$$p_T(\lambda) := \det([T]_\beta^\beta - \lambda I_n) \qquad (\lambda \in \mathbb{F}),$$

where $[T]_\beta^\beta \in \mathbb{F}^{n\times n}$ is the matrix representation of T in the basis β.

The characteristic polynomial is a degree-n polynomial in λ. It encodes all the eigenvalue information of T: its roots (in $\mathbb{F}$) are exactly the eigenvalues. The leading term is $(-\lambda)^n$ (from the diagonal product of $A - \lambda I_n$), and the constant term is $\det(A) = p_T(0)$.

Theorem 5.3: Basis-independence of p_T

The function p_T does not depend on the chosen ordered basis β.

Proof. Let γ be another ordered basis. Then $[T]_\gamma^\gamma = Q^{-1}[T]_\beta^\beta Q$ for some invertible Q. Hence, for any $\lambda \in \mathbb{F}$,

$$[T]_\gamma^\gamma - \lambda I_n = Q^{-1}([T]_\beta^\beta - \lambda I_n)Q,$$

so $\det([T]_\gamma^\gamma - \lambda I_n) = \det([T]_\beta^\beta - \lambda I_n)$. □

This is why we can speak of "the" characteristic polynomial of T without specifying a basis. The proof uses two facts: (1) matrix representations in different bases are similar, and (2) similar matrices have equal determinants.

Theorem 5.4: Eigenvalues are roots of p_T

Let $T \in \mathcal{L}(V)$ with $\dim(V) = n$. Then $\lambda \in \mathbb{F}$ is an eigenvalue of T if and only if

$$p_T(\lambda) = 0.$$

In particular, T has at most n eigenvalues in $\mathbb{F}$.

Proof. Fix a basis β and set $A = [T]_\beta^\beta$. Then

$$(T - \lambda I_V)(v) = 0 \text{ with } v \neq 0 \iff (A - \lambda I_n)[v]_\beta = 0 \text{ with } [v]_\beta \neq 0.$$

Thus λ is an eigenvalue iff $A - \lambda I_n$ is not invertible, iff $\det(A - \lambda I_n) = 0$, i.e. $p_T(\lambda) = 0$. □

The chain of equivalences used in the proof is worth memorizing for exams:

$$\lambda \text{ eigenvalue } \Leftrightarrow T-\lambda I_V \text{ not injective } \Leftrightarrow [T]_\beta^\beta - \lambda I_n \text{ not invertible } \Leftrightarrow \det([T]_\beta^\beta - \lambda I_n) = 0.$$

Each step uses a different foundational result: the definition of eigenvalue, the matrix representation preserving injectivity, and the determinant criterion for invertibility.

Worked Example (Tutorial 9, Question 1): eigenvalues of $L_A(v) = Av$

Let $A \in \mathbb{F}^{n\times n}$ and define $L_A \in \mathcal{L}(\mathbb{F}^{n,1})$ by $L_A(v) = Av$. Then λ is an eigenvalue of L_A iff $\det(A-\lambda I_n) = 0$, and

$$p_{L_A}(\lambda) = \det(A - \lambda I_n).$$

Moreover, L_A is diagonalizable iff A is similar to a diagonal matrix $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ with λ_i eigenvalues of A.

Exam Focus: fast eigenvalue computation

For an upper-/lower-triangular matrix A, the eigenvalues are the diagonal entries (with algebraic multiplicities). This is often the fastest way to read off eigenvalues before checking eigenspaces for diagonalizability.

Exam Focus: trace/determinant shortcuts for 2×2 matrices

For $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the characteristic polynomial is $\lambda^2 - \operatorname{tr}(A)\lambda + \det(A) = \lambda^2 - (a+d)\lambda + (ad-bc)$. The eigenvalues satisfy $\lambda_1 + \lambda_2 = \operatorname{tr}(A)$ and $\lambda_1\lambda_2 = \det(A)$. This avoids expanding the determinant from scratch every time.

Exam Focus: 3×3 characteristic polynomial via cofactor expansion

For upper-triangular A, read eigenvalues off the diagonal. Otherwise, expand $\det(A - \lambda I_3)$ along the row/column with the most zeros. A useful identity: for any $n \times n$ matrix,

$$p_T(\lambda) = (-1)^n\big(\lambda^n - \operatorname{tr}(A)\lambda^{n-1} + \cdots + (-1)^n \det(A)\big).$$

For $n = 3$: $p_T(\lambda) = -\lambda^3 + \operatorname{tr}(A)\lambda^2 - (\text{sum of } 2\times 2 \text{ cofactors})\lambda + \det(A)$.

5.3 Eigenspaces: direct sum structure and dimension bounds

Theorem 5.5: Sum of distinct eigenspaces is a direct sum

Let $T \in \mathcal{L}(V)$ and let $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$ be distinct eigenvalues of T. Then

$$E_{T,\lambda_1} + \cdots + E_{T,\lambda_m} = E_{T,\lambda_1} \oplus \cdots \oplus E_{T,\lambda_m}.$$

Proof. Let $v_i \in E_{T,\lambda_i}$ and suppose $v_1 + \cdots + v_m = 0$. We show $v_1 = \cdots = v_m = 0$.

If $m = 1$, the statement is immediate. Assume $m \geq 2$. Apply T to the relation to get

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

Subtract $\lambda_m(v_1 + \cdots + v_m) = 0$ to obtain

$$(\lambda_1 - \lambda_m)v_1 + \cdots + (\lambda_{m-1} - \lambda_m)v_{m-1} = 0.$$

All scalars $\lambda_i - \lambda_m$ for $i < m$ are nonzero. By repeating the same argument inductively (reducing the number of summands), we conclude $v_1 = \cdots = v_{m-1} = 0$, and then $v_m = 0$ from $v_1 + \cdots + v_m = 0$. Thus the sum is direct. □

Why is this result so important? It says that eigenvectors from different eigenspaces can *never* be redundant: combining bases of the individual eigenspaces always yields a linearly independent set. This is the structural backbone of diagonalizability—if these independent eigenvectors span all of V, then T is diagonalizable.

Theorem 5.6: Sum of eigenspace dimensions is bounded by $\dim(V)$

Let V be finite-dimensional and $T \in \mathcal{L}(V)$. Then

$$\sum_{\lambda \text{ eigenvalue of } T} \dim(E_{T,\lambda}) \leq \dim(V).$$

Proof. There are finitely many eigenvalues of T in $\mathbb{F}$ (Theorem 5.4). By Theorem 5.5, the sum of all eigenspaces is a direct sum, so

$$\dim\Big(\sum_{\lambda} E_{T,\lambda}\Big) = \sum_{\lambda} \dim(E_{T,\lambda}).$$

Since $\sum_\lambda E_{T,\lambda}$ is a subspace of V, its dimension is at most $\dim(V)$. □

Extension (Beyond Syllabus: Distinct eigenvalues ⇒ independent eigenvectors)

Syllabus link: MH1201 Weeks 9–10 — Eigenvalues and Eigenvectors.
Lemma. If $T \in \mathcal{L}(V)$ and $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues with eigenvectors $v_i \in E_{T,\lambda_i} \setminus \{0\}$, then $\{v_1, \ldots, v_m\}$ is linearly independent.
Proof. Suppose $a_1v_1 + \cdots + a_mv_m = 0$ with $a_i \in \mathbb{F}$. This is a vector in $E_{T,\lambda_1} + \cdots + E_{T,\lambda_m}$ equal to 0. By Theorem 5.5, the sum is direct, hence each summand must be 0: $a_iv_i = 0$ for all i. Since $v_i \neq 0$, we get $a_i = 0$ for all i. □

5.4 Diagonalization: definitions, criteria, and multiplicities

Definition 5.4: Diagonalizability

Let V be a vector space (not necessarily finite-dimensional) and let $T \in \mathcal{L}(V)$. We say that T is *diagonalizable* if V has a basis consisting of eigenvectors of T.

The definition says that every vector in V can be written as a linear combination of eigenvectors. Intuitively, a diagonalizable operator is one that acts "independently along each direction" in a suitable basis: T simply scales each basis vector, and the matrix becomes diagonal.

Theorem 5.7: Diagonal operator matrix representation

Let V be finite-dimensional with $\dim(V) = n$ and let $T \in \mathcal{L}(V)$. Then T is diagonalizable if and only if there exists an ordered basis β of V such that $[T]_\beta^\beta$ is diagonal.

Proof. ($\Rightarrow$) If T is diagonalizable, let $\beta = (v_1, \ldots, v_n)$ be an eigenbasis with $T(v_i) = \lambda_i v_i$. Then the ith column of $[T]_\beta^\beta$ is $[T(v_i)]_\beta = [\lambda_i v_i]_\beta = \lambda_i e_i$, so $[T]_\beta^\beta = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

($\Leftarrow$) If $[T]_\beta^\beta = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ for $\beta = (v_1, \ldots, v_n)$, then

$$[T(v_i)]_\beta = [T]_\beta^\beta [v_i]_\beta = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) e_i = \lambda_i e_i = [\lambda_i v_i]_\beta,$$

so $T(v_i) = \lambda_i v_i$. Hence β is an eigenbasis and T is diagonalizable. □

Exam Focus: reading diagonalization as $A = PDP^{-1}$

If the eigenbasis is $\beta = (v_1, \ldots, v_n)$ and the standard basis is σ, then:

$$P = [I_V]_{\sigma\beta} = [\, v_1 \ \cdots \ v_n \,], \quad D = [T]_\beta^\beta = \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$

and the change-of-basis formula gives

$$[T]_\sigma^\sigma = PDP^{-1}.$$

The columns of P are the eigenvectors, and the diagonal entries of D are the corresponding eigenvalues (in the *same order*).

Theorem 5.8: Fundamental Theorem of Diagonalizability (finite dimensions)

Let V be finite-dimensional and let $T \in \mathcal{L}(V)$. The following are equivalent:

(a) T is diagonalizable.

(b) $V = \sum_{\lambda \text{ eigenvalue of } T} E_{T,\lambda}$.

(c) $\dim(V) = \sum_{\lambda \text{ eigenvalue of } T} \dim(E_{T,\lambda})$.

Proof. (b)$\Leftrightarrow$(c): by Theorem 5.6, the sum $\sum_\lambda E_{T,\lambda}$ is a (finite) direct sum, so

$$\dim\Big(\sum_\lambda E_{T,\lambda}\Big) = \sum_\lambda \dim(E_{T,\lambda}).$$

Since $\sum_\lambda E_{T,\lambda} \subseteq V$, equality of subspaces holds iff dimensions are equal, hence (b)$\Leftrightarrow$(c).

(a)$\Rightarrow$(b): let B be a basis of V consisting of eigenvectors. For each eigenvalue λ, let $B_\lambda = B \cap E_{T,\lambda}$. Then $B = \bigcup_\lambda B_\lambda$, so

$$V = \operatorname{span}(B) = \sum_\lambda \operatorname{span}(B_\lambda) \subseteq \sum_\lambda E_{T,\lambda} \subseteq V,$$

hence equality holds.

(b)$\Rightarrow$(a): for each eigenvalue λ, choose a basis B_λ for $E_{T,\lambda}$ and set $B = \bigcup_\lambda B_\lambda$. Then $\operatorname{span}(B) = \sum_\lambda E_{T,\lambda} = V$, so B spans V and consists of eigenvectors; extract a basis of V from B. Thus T is diagonalizable. □

In the last step of (b)$\Rightarrow$(a), we claimed that $B = \bigcup_\lambda B_\lambda$ is already a basis (not just a spanning set). This is because the sum $\sum_\lambda E_{T,\lambda}$ is direct: combining bases of the individual eigenspaces produces a linearly independent set. So $|B| = \sum_\lambda |B_\lambda| = \sum_\lambda \dim(E_{T,\lambda}) = \dim(V)$, and B spans, hence B is a basis.

Theorem 5.9: Sufficient criterion: n distinct eigenvalues

Let V be n-dimensional and let $T \in \mathcal{L}(V)$. If T has n distinct eigenvalues in $\mathbb{F}$, then T is diagonalizable.

Proof. Let $\lambda_1, \ldots, \lambda_n$ be distinct eigenvalues. Then each $E_{T,\lambda_i} \neq \{0\}$, so $\dim(E_{T,\lambda_i}) \geq 1$. Hence $\sum_{i=1}^{n} \dim(E_{T,\lambda_i}) \geq n = \dim(V)$. By Theorem 5.6, the sum of these dimensions is also $\leq \dim(V)$, so equality holds. Now apply Theorem 5.8. □

The converse is *false*: an operator can be diagonalizable with repeated eigenvalues (e.g. I_V has only eigenvalue 1 but is trivially diagonalizable). The key condition is that *geometric multiplicity equals algebraic multiplicity* for every eigenvalue, not that all eigenvalues are distinct.

Definition 5.5: Algebraic vs. geometric multiplicity

Let V be finite-dimensional and let $T \in \mathcal{L}(V)$. For an eigenvalue λ of T:

- The *algebraic multiplicity* of λ is the multiplicity of λ as a root of p_T (the highest power of $(x - \lambda)$ dividing $p_T(x)$).
- The *geometric multiplicity* of λ is $\dim(E_{T,\lambda})$.

The algebraic multiplicity counts how many times λ appears as a root of the characteristic polynomial. The geometric multiplicity counts how many linearly independent eigenvectors exist for λ. The key inequality is:

$$1 \leq \text{geometric multiplicity of } \lambda \leq \text{algebraic multiplicity of } \lambda.$$

The lower bound holds because λ is an eigenvalue (so at least one eigenvector exists). The upper bound is a standard result. Diagonalizability fails precisely when the geometric multiplicity is *strictly less than* the algebraic multiplicity for some eigenvalue.

Theorem 5.10: Multiplicity characterization of diagonalizability

Let V be finite-dimensional and $T \in \mathcal{L}(V)$. Then T is diagonalizable if and only if:

(a) p_T factors completely into linear factors over $\mathbb{F}$, and

(b) for every eigenvalue λ, the geometric multiplicity of λ equals its algebraic multiplicity.

Condition (a) is automatic over $\mathbb{C}$ (every polynomial splits), but can fail over $\mathbb{R}$ (e.g. $\lambda^2 + 1$). Condition (b) is the one you must check by computing eigenspace dimensions.

Exam Focus: diagonalizability checklist

Given $T \in \mathcal{L}(V)$ with $\dim(V) = n$:

1. Compute eigenvalues: solve $p_T(\lambda) = 0$ (or use triangular/trace/det shortcuts).
2. For each eigenvalue λ, compute $E_{T,\lambda} = \ker(T - \lambda I_V)$ and $\dim(E_{T,\lambda})$.
3. If $\sum_\lambda \dim(E_{T,\lambda}) = n$, then T is diagonalizable; otherwise it is not.
4. If diagonalizable, form an eigenbasis by concatenating bases of the eigenspaces; then $[T]_\beta^\beta$ is diagonal.

Common Mistake: diagonalizability does not require distinct eigenvalues

Having n distinct eigenvalues is *sufficient* but not *necessary* for diagonalizability. The identity I_V is trivially diagonalizable with only one eigenvalue ($\lambda = 1$, geometric multiplicity n). The correct test is always: sum of geometric multiplicities $= n$.

5.5 Core applications: powers, sums of powers, and conjugacy

Key Formula: powers via diagonalization

If $A \in \mathbb{F}^{n\times n}$ is diagonalizable, i.e. $A = PDP^{-1}$ with $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, then

$$A^k = PD^kP^{-1}, \qquad D^k = \operatorname{diag}(\lambda_1^k, \dots, \lambda_n^k) \quad (k \in \mathbb{N}_0).$$

More generally, for $p(t) = \sum_{j=0}^{m} a_j t^j$,

$$p(A) = \sum_{j=0}^{m} a_j A^j = P\,p(D)\,P^{-1}, \qquad p(D) = \operatorname{diag}(p(\lambda_1), \dots, p(\lambda_n)).$$

The formula $A^k = PD^kP^{-1}$ follows from telescoping: $A^k = (PDP^{-1})^k = PD(P^{-1}P)D\cdots DP^{-1} = PD^kP^{-1}$. The general polynomial version uses the same idea together with linearity: $p(A) = \sum a_j A^j = \sum a_j PD^jP^{-1} = P(\sum a_j D^j)P^{-1} = Pp(D)P^{-1}$.

Worked Example (Tutorial 8, Question 1): Fibonacci via diagonalization

Let $F \in \mathcal{L}(\mathbb{R}^{2,1})$ be $F\left(\binom{x}{y}\right) = \binom{y}{x+y}$, i.e. $F(v) = Av$ with

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

Then $p_A(\lambda) = \lambda^2 - \lambda - 1$, with eigenvalues

$$\varphi = \frac{1+\sqrt{5}}{2}, \qquad \psi = \frac{1-\sqrt{5}}{2}.$$

An eigenbasis is $\left(\binom{1}{\varphi}, \binom{1}{\psi}\right)$, so $A = P\operatorname{diag}(\varphi, \psi)P^{-1}$ and

$$F^n\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} F_n \\ F_{n+1} \end{bmatrix}, \qquad F_n = \frac{\varphi^n - \psi^n}{\sqrt{5}} \quad (n \in \mathbb{N}_0).$$

The Fibonacci example illustrates the full diagonalization workflow: (1) find eigenvalues via p_A, (2) find eigenvectors, (3) form P and D, (4) decompose the input vector in the eigenbasis, (5) apply D^n (trivial for a diagonal matrix), (6) convert back. The induction proof $F^n\binom{0}{1} = \binom{F_n}{F_{n+1}}$ uses the recursion $F_{n+2} = F_{n+1} + F_n$ (which is encoded in the matrix product $A\binom{F_n}{F_{n+1}} = \binom{F_{n+1}}{F_n + F_{n+1}}$).

Worked Example (Tutorial 8, Question 2): diagonalization on $\mathcal{P}_2(\mathbb{R})$

Define $T \in \mathcal{L}(\mathcal{P}_2(\mathbb{R}))$ by

$$T(a + bX + cX^2) = (a + b + c) + bX + (b + 2c)X^2.$$

In the standard basis $(1, X, X^2)$, $[T] = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$. The characteristic polynomial is $\chi(\lambda) = (1-\lambda)^2(2-\lambda)$, giving eigenvalues 1 (alg. mult. 2) and 2.
Eigenspaces: $E_{T,1} = \text{span}\{1, X^2 - X\}$ (dim 2), $E_{T,2} = \text{span}\{1 + X^2\}$ (dim 1). Since $2 + 1 = 3 = \dim(\mathcal{P}_2(\mathbb{R}))$, T is diagonalizable with eigenbasis $\beta = (1, X^2 - X, 1 + X^2)$.
Iterates: Decompose $P = a + bX + cX^2 = (a - b - c) \cdot 1 + (-b)(X^2 - X) + (b + c)(1 + X^2)$. Then

$$T^n(P) = (a - b - c) + (-b)(X^2 - X) + (b + c) \cdot 2^n(1 + X^2).$$

Worked Example (Lecture Week 9): diagonalizable vs. defective (upper triangular 3×3)

Define $T \in \mathcal{L}(\mathbb{R}^3)$ by

$$T(x, y, z) = (6x + 3y + 4z,\ 6y + 2z,\ 7z).$$

In the standard basis, T has upper-triangular matrix

$$A = \begin{bmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{bmatrix},$$

so eigenvalues are 6 (algebraic multiplicity 2) and 7.
Compute $E_{T,6} = \ker(A - 6I_3)$:

$$A - 6I_3 = \begin{bmatrix} 0 & 3 & 4 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow\ z = 0,\ y = 0,\ x \text{ free}.$$

Hence $\dim(E_{T,6}) = 1 < 2$, so $\sum_\lambda \dim(E_{T,\lambda}) < 3$ and T is *not* diagonalizable.

This example is the prototypical "defective" case: the eigenvalue 6 has algebraic multiplicity 2 but geometric multiplicity 1. The eigenspace $E_{T,6}$ is only one-dimensional, so there is no way to find a second linearly independent eigenvector for $\lambda = 6$. The deficiency is $2 - 1 = 1$, and T cannot be diagonalized.

Worked Example (Lecture Week 9): diagonalization on $\mathbb{R}^{2,2}$

Define $T \in \mathcal{L}(\mathbb{R}^{2,2})$ by

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} c & d \\ a & b \end{bmatrix}.$$

Then $T^2 = I$, so the only possible eigenvalues are ± 1 (since $T(v) = \lambda v \Rightarrow v = T^2(v) = \lambda^2 v$). Indeed,

$$E_{T,1} = \left\{\begin{bmatrix} a & b \\ a & b \end{bmatrix}\right\} = \text{span}\left\{\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\right\}, \quad E_{T,-1} = \left\{\begin{bmatrix} a & b \\ -a & -b \end{bmatrix}\right\} = \text{span}\left\{\begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}\right\}.$$

Thus $\dim(E_{T,1}) + \dim(E_{T,-1}) = 2 + 2 = 4 = \dim(\mathbb{R}^{2,2})$, so T is diagonalizable.

Exam Focus: involutions ($T^2 = I$) are always diagonalizable (over fields where char $\neq 2$)

If $T^2 = I_V$, then every eigenvalue satisfies $\lambda^2 = 1$, so $\lambda \in \{1, -1\}$. The polynomial $t^2 - 1 = (t-1)(t+1)$ annihilates T, and since $1 \neq -1$ (assuming $\mathrm{char}(\mathbb{F}) \neq 2$), the eigenspaces for 1 and -1 together span V, so T is diagonalizable. More generally, if $p(T) = 0$ for a polynomial p that splits into *distinct* linear factors over $\mathbb{F}$, then T is diagonalizable.

Worked Example (Lecture Week 9): explicit formula for $\begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix}^n$

Let $M = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix}$. Then

$$\chi_M(\lambda) = \det(M - \lambda I_2) = \lambda^2 - 3\lambda + 2 = (\lambda - 1)(\lambda - 2),$$

so eigenvalues are $1, 2$ (distinct), hence M is diagonalizable.

Eigenvectors: for $\lambda = 1$, take $v_1 = \binom{-2}{1}$; for $\lambda = 2$, take $v_2 = \binom{-1}{1}$. Let $P = [v_1\ v_2] = \begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix}$, so $P^{-1} = \begin{bmatrix} -1 & -1 \\ 1 & 2 \end{bmatrix}$. Then

$$M^n = P \begin{bmatrix} 1 & 0 \\ 0 & 2^n \end{bmatrix} P^{-1} = \begin{bmatrix} 2 - 2^n & 2 - 2^{n+1} \\ -1 + 2^n & -1 + 2^{n+1} \end{bmatrix} \qquad (n \in \mathbb{N}_0).$$

Exam Focus: sanity-checking $M^n = PD^nP^{-1}$

After computing $M^n = PD^nP^{-1}$, always verify:

- $n = 0$: does $M^0 = I$ match? (Substitute $n = 0$ in your formula.)
- $n = 1$: does $M^1 = M$ match? (Substitute $n = 1$.)
- Trace: $\mathrm{tr}(M^n) = \lambda_1^n + \lambda_2^n + \cdots$ (since trace is a similarity invariant).

These quick checks catch most computational errors.

Extension (Tutorial 8, Question 3, Method 2)

Syllabus link: MH1201 Weeks 9–10 — Diagonalization (computational uses).
Lemma (Polynomial functional calculus on eigenvectors). Let $A \in \mathbb{F}^{n\times n}$ and let $p(t) = \sum_{k=0}^{m} a_k t^k \in \mathbb{F}[t]$. If $Av = \lambda v$ for some $v \neq 0$, then $p(A)v = p(\lambda)v$.
Proof. Using $A^k v = \lambda^k v$ (proved by induction on k),

$$p(A)v = \Big(\sum_{k=0}^{m} a_k A^k\Big)v = \sum_{k=0}^{m} a_k A^k v = \sum_{k=0}^{m} a_k \lambda^k v = p(\lambda)v.$$

□

Consequence: If A is diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_n$, then $p(A)$ is also diagonalizable (in the *same* eigenbasis) with eigenvalues $p(\lambda_1), \ldots, p(\lambda_n)$. Hence $\det(p(A)) = p(\lambda_1) \cdots p(\lambda_n)$.

Worked Example (Tutorial 8, Question 3): determinant of $\sum_{k=1}^{n} A^k$

Let $A \in \mathbb{C}^{3\times 3}$ satisfy that $A - I_3$, $A - eI_3$, and $A - iI_3$ are not invertible. Then $1, e, i$ are eigenvalues of A; they are distinct, so A is diagonalizable over $\mathbb{C}$.
Let $S_n = \sum_{k=1}^{n} A^k = p_n(A)$ where $p_n(t) = \sum_{k=1}^{n} t^k$. By the extension above, the eigenvalues of S_n are $p_n(1) = n$, $p_n(e) = \frac{e^{n+1}-e}{e-1}$, and $p_n(i) = \frac{i^{n+1}-i}{i-1}$, so

$$\det\left(\sum_{k=1}^{n} A^k\right) = n \cdot \frac{e^{n+1}-e}{e-1} \cdot \frac{i^{n+1}-i}{i-1} \qquad (n \in \mathbb{N}).$$

The key reasoning is: $A - \lambda I$ not invertible $\Rightarrow$ λ is an eigenvalue. Three distinct eigenvalues in a 3×3 matrix force diagonalizability. Then the polynomial functional calculus lets us evaluate $\det(p(A))$ as a product of $p(\lambda_i)$ without ever computing P or P^{-1}.

Extension (Tutorial 8, Question 4, Method 1)

Syllabus link: MH1201 Weeks 9–10 — Diagonalization (structure under distinct eigenvalues).
Lemma (Conjugacy from matched distinct eigenvalues in dimension 3). Let V be a 3-dimensional real vector space, and let $S, T \in \mathcal{L}(V)$. Assume S and T each have *three distinct* real eigenvalues $\lambda_1, \lambda_2, \lambda_3$. Then there exists an invertible $R \in \mathcal{L}(V)$ such that $S = R^{-1} \circ T \circ R$.
Proof. For each $j \in \{1, 2, 3\}$, pick nonzero eigenvectors $s_j \in E_{S,\lambda_j}$ and $t_j \in E_{T,\lambda_j}$. Distinct eigenvalues imply $\{s_1, s_2, s_3\}$ and $\{t_1, t_2, t_3\}$ are linearly independent, hence bases of V. Define $R \in \mathcal{L}(V)$ by $R(t_j) = s_j$ for $j = 1, 2, 3$ and extend linearly; then R is invertible. For each j,

$$(S \circ R)(t_j) = S(s_j) = \lambda_j s_j = \lambda_j R(t_j) = (R \circ T)(t_j),$$

so $S \circ R = R \circ T$ on a basis, hence on all of V. Right-compose with R^{-1} to get $S = R \circ T \circ R^{-1}$, equivalently $S = R^{-1} \circ T \circ R$ after renaming $R \leftarrow R^{-1}$. □

Extension (Tutorial 8, Question 4, Method 2): conjugacy via shared diagonal form

Syllabus link: MH1201 Weeks 9–10 — Diagonalization and similarity.
Observation. If S and T in $\mathcal{L}(V)$ are both diagonalizable with the *same* set of eigenvalues (counted with multiplicities), then $S \sim T$.
Proof. Both $[S]_{\beta_S}^{\beta_S}$ and $[T]_{\beta_T}^{\beta_T}$ equal the same diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ (with eigenvalues listed in matching order). By transitivity of similarity, $[S]_{\sigma}^{\sigma} \sim \Lambda \sim [T]_{\sigma}^{\sigma}$, hence $S \sim T$.
Concretely, the invertible operator that conjugates T to S is the change-of-basis map from β_T to β_S. □

5.6 Edge cases and diagnostic traps

Common Mistake: "repeated eigenvalue ⇒ diagonalizable" (false)

A repeated eigenvalue does *not* guarantee enough eigenvectors. Always compute $\dim(E_{T,\lambda})$ for repeated eigenvalues and compare with algebraic multiplicity.

Worked Example (Lecture Week 9): $D + D^2$ on $\mathcal{P}_3(\mathbb{R})$ is not diagonalizable

Let $D : \mathcal{P}_3(\mathbb{R}) \to \mathcal{P}_3(\mathbb{R})$ be the derivative operator and consider $T = D + D^2$. Since $D^4 = 0$ on $\mathcal{P}_3(\mathbb{R})$, D is nilpotent; hence $T = D(I + D)$ is also nilpotent. Therefore the only eigenvalue of T is 0.
If T were diagonalizable with only eigenvalue 0, then it would be similar to the zero diagonal matrix, hence $T = 0$, contradiction (e.g. $T(X) = 1 \neq 0$). Thus T is not diagonalizable.

This example demonstrates a powerful shortcut: if T is nilpotent and nonzero, it is *automatically not diagonalizable*. The argument: nilpotent $\Rightarrow$ only eigenvalue is $0 \Rightarrow$ if diagonalizable then $T \sim 0 \Rightarrow T = 0$, contradiction. This avoids any eigenspace dimension computation.

Exam Focus: nilpotent operators and diagonalizability

A linear operator T is *nilpotent* if $T^k = 0$ for some $k \in \mathbb{N}$. Key facts about nilpotent operators:

- The only eigenvalue of a nilpotent operator is 0 (if $T(v) = \lambda v$ and $T^k = 0$, then $\lambda^k v = 0$, so $\lambda = 0$).
- A nilpotent operator is diagonalizable if and only if it is the zero operator.
- Consequently, any nonzero nilpotent operator is *defective* (not diagonalizable).

Worked micro-example: the simplest non-diagonalizable matrix

The matrix $N = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is nilpotent ($N^2 = 0$) and nonzero, hence not diagonalizable. Explicitly: eigenvalue 0 has algebraic multiplicity 2 but $\ker(N) = \text{span}\{e_1\}$ has dimension 1. So geometric multiplicity $<$ algebraic multiplicity, confirming non-diagonalizability.

Exam Focus: solvability of $(T - \lambda I)v = b$

For $T \in \mathcal{L}(V)$ on finite-dimensional V:

$$\lambda \text{ is not an eigenvalue of } T \iff T - \lambda I_V \text{ is invertible.}$$

So if λ is not an eigenvalue, the equation $(T - \lambda I_V)v = b$ has a *unique* solution for every $b \in V$ (e.g. Tutorial 8, Question 5).

This is a frequently tested technique: to show an equation $T(v) = \lambda v + b$ has a solution, rewrite it as $(T - \lambda I)v = b$ and argue that $T - \lambda I$ is invertible (because λ is not an eigenvalue). The argument requires knowing all eigenvalues of T—if λ is not among them, the system is guaranteed to have a unique solution.

Worked Example (Tutorial 8, Question 5): using the spectrum to solve $T(v) = 8v + b$

Let $T \in \mathcal{L}(\mathbb{R}^3)$ with eigenvalues 6 and 7 only. To show there exist x, y, z with $T(x, y, z) = (6 + 8x,\ 7 + 8y,\ 13 + 8z)$, rewrite as

$$(T - 8I_3)(x, y, z) = (6, 7, 13).$$

Since 8 is not an eigenvalue of T, the operator $T - 8I_3$ is invertible. Hence the equation $(T - 8I_3)v = (6, 7, 13)$ has a (unique) solution $v = (x, y, z) \in \mathbb{R}^3$. □

Common Mistake: confusing algebraic and geometric multiplicity

When checking diagonalizability, you need *both* multiplicities. A common error is to compute only eigenvalues (algebraic multiplicities) and declare the operator diagonalizable. You must also check that each eigenspace has dimension equal to the algebraic multiplicity. For instance, $A = \begin{bmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{bmatrix}$ has eigenvalue 6 with algebraic multiplicity 2, but $\dim(E_{T,6}) = 1$, so A is *not* diagonalizable.

Exam Focus: complete diagonalizability workflow summary

1. **Find eigenvalues:** Solve $\det(A - \lambda I) = 0$. Record algebraic multiplicities.
2. **Find eigenspaces:** For each λ, row-reduce $A - \lambda I$ and find $\ker(A - \lambda I)$. Record geometric multiplicities (= number of free variables = $n - \text{rank}(A - \lambda I)$).
3. **Test:** If $\sum$(geometric mult.) $= n$, diagonalizable. Otherwise, not.
4. **Build P and D:** Columns of P are eigenvectors (one basis per eigenspace, concatenated). Diagonal of D has eigenvalues in the *same column order as* P.
5. **Apply:** $A^k = PD^kP^{-1}$, $p(A) = Pp(D)P^{-1}$.

QRS@NTU

6 Inner Product Spaces

6.1 Inner products and norms

Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. An *inner product space* is a pair $(V, \langle \cdot, \cdot \rangle)$ where V is an $\mathbb{F}$-vector space and $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ satisfies the axioms below.

Definition 6.1: Inner product (real vs. complex conventions)

Let V be a vector space over $\mathbb{F}$. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ is an *inner product* if:

(a) **(Sesquilinearity)** For all $u, u_1, u_2, v, v_1, v_2 \in V$ and $a, b \in \mathbb{F}$,

$$\langle au_1 + bu_2,\, v \rangle = \overline{a}\, \langle u_1, v \rangle + \overline{b}\, \langle u_2, v \rangle, \qquad \langle u,\, av_1 + bv_2 \rangle = a\, \langle u, v_1 \rangle + b\, \langle u, v_2 \rangle.$$

(Over $\mathbb{R}$, $\overline{a} = a$, so the inner product is bilinear.)

(b) **(Conjugate symmetry)** $\langle u, v \rangle = \overline{\langle v, u \rangle}$ for all $u, v \in V$. (Over $\mathbb{R}$ this reduces to symmetry: $\langle u, v \rangle = \langle v, u \rangle$.)

(c) **(Positive definiteness)** $\langle u, u \rangle \geq 0$ for all $u \in V$, and $\langle u, u \rangle = 0$ iff $u = 0$.

Definition 6.2: Norm induced by an inner product; distance

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. Define the induced norm $\| \cdot \| : V \to \mathbb{R}_{\geq 0}$ by

$$\|u\| \coloneqq \sqrt{\langle u, u \rangle}.$$

Define the induced distance by $d(u, v) \coloneqq \|u - v\|$.

Theorem 6.1: Basic norm properties

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space with induced norm $\| \cdot \|$. Then for all $u, v \in V$ and $\lambda \in \mathbb{F}$:

$$\|u\| \geq 0, \quad \|u\| = 0 \iff u = 0, \qquad \|\lambda u\| = |\lambda|\, \|u\|, \qquad \|u + v\| \leq \|u\| + \|v\|.$$

Examples (Lecture Week 11): standard/weighted/integral/Frobenius inner products

- $\mathbb{R}^n$: $\langle x, y \rangle = \sum_{k=1}^{n} x_k y_k$.
- $\mathbb{C}^n$: $\langle w, z \rangle = \sum_{k=1}^{n} \overline{w_k}\, z_k$ (standard convention).
- (Weighted) $\mathbb{C}^n$: if $c_1, \ldots, c_n > 0$, then $\langle w, z \rangle = \sum_{k=1}^{n} c_k\, \overline{w_k}\, z_k$.
- $\mathcal{P}_m(\mathbb{R})$: $\langle P, Q \rangle = \int_a^b P(x) Q(x)\, dx$ (e.g. $[-1, 1]$), which is positive definite on polynomials.
- $\mathbb{R}^{n \times n}$ (Frobenius): $\langle A, B \rangle_F = \operatorname{tr}(A^{\mathsf{T}} B) = \sum_{i,j} a_{ij} b_{ij}$.

Worked Example (Tutorial 10, Question 4(a)): inner products from Hermitian positive definite matrices

Let $H \in \mathbb{C}^{n\times n}$ be *Hermitian* ($H^\dagger = H$) and *positive definite* ($u^\dagger Hu > 0$ for all $u \neq 0$). Define $\langle u, v\rangle := u^\dagger Hv$. Then $\langle \cdot, \cdot\rangle$ is an inner product on $\mathbb{C}^{n,1}$:

- sesquilinearity comes from $(au_1 + bu_2)^\dagger = \overline{a}\, u_1^\dagger + \overline{b}\, u_2^\dagger$ and matrix distributivity;
- conjugate symmetry is $\langle u, v\rangle = u^\dagger Hv = \overline{v^\dagger H^\dagger u} = \overline{\langle v, u\rangle}$;
- positive definiteness is exactly $u^\dagger Hu > 0$ for $u \neq 0$.

Common Mistake: complex linearity in the wrong slot

In complex spaces, *exactly one slot* is linear; the other is conjugate-linear. Before expanding $\langle au+bv,\, w\rangle$ or $\langle w,\, au + bv\rangle$, decide which slot is conjugated.

6.2 Cauchy–Schwarz and the triangle inequality

Theorem 6.2: Cauchy–Schwarz inequality

Let $(V, \langle \cdot, \cdot\rangle)$ be an inner product space. Then for all $u, v \in V$,

$$|\langle u, v\rangle| \leq \|u\|\, \|v\|.$$

Theorem 6.3: Equality condition in Cauchy–Schwarz

For $u, v \in V$, equality $|\langle u, v\rangle| = \|u\|\, \|v\|$ holds if and only if u and v are linearly dependent.

Extension (Lecture Week 11: Cauchy–Schwarz via quadratic discriminant)

Syllabus link: MH1201 Weeks 11–12 — Inner product spaces; Cauchy–Schwarz.
Theorem. In any real/complex inner product space $(V, \langle\cdot,\cdot\rangle)$,

$$\forall u, v \in V: \quad |\langle u, v\rangle| \le \|u\|\,\|v\|.$$

Proof. If $v = 0$, then $\langle u, v\rangle = 0$ and the inequality holds. Assume $v \neq 0$. Choose $\alpha \in \mathbb{F}$ with $|\alpha| = 1$ and $\alpha\langle u, v\rangle = |\langle u, v\rangle|$. (If $\langle u, v\rangle = 0$ take $\alpha = 1$; otherwise take $\alpha = \frac{|\langle u,v\rangle|}{\langle u,v\rangle}$.)
For $t \in \mathbb{R}$, define

$$f(t) \coloneqq \langle u + t\alpha v,\, u + t\alpha v\rangle.$$

By positive definiteness, $f(t) \ge 0$ for all $t \in \mathbb{R}$. Expand using sesquilinearity and conjugate symmetry:

$$\begin{aligned} f(t) &= \langle u, u\rangle + \langle u, t\alpha v\rangle + \langle t\alpha v, u\rangle + \langle t\alpha v, t\alpha v\rangle \\ &= \|u\|^2 + t\alpha\,\langle u, v\rangle + t\overline{\alpha}\,\langle v, u\rangle + t^2|\alpha|^2\,\langle v, v\rangle \\ &= \|u\|^2 + t\alpha\,\langle u, v\rangle + t\overline{\alpha}\,\overline{\langle u, v\rangle} + t^2\|v\|^2. \end{aligned}$$

Since $\alpha\langle u, v\rangle = |\langle u, v\rangle|$, we also have $\overline{\alpha}\,\overline{\langle u, v\rangle} = |\langle u, v\rangle|$. Hence

$$f(t) = \|v\|^2 t^2 + 2|\langle u, v\rangle|\, t + \|u\|^2.$$

This is a real quadratic polynomial in t that is nonnegative for all $t \in \mathbb{R}$, so its discriminant is ≤ 0:

$$(2|\langle u, v\rangle|)^2 - 4\|v\|^2\|u\|^2 \le 0.$$

Therefore $|\langle u, v\rangle|^2 \le \|u\|^2\|v\|^2$, and taking square roots gives $|\langle u, v\rangle| \le \|u\|\,\|v\|$. □

Theorem 6.4: Triangle inequality from Cauchy–Schwarz

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space. Then for all $u, v \in V$,

$$\|u + v\| \le \|u\| + \|v\|.$$

Proof. Compute

$$\|u + v\|^2 = \langle u + v, u + v\rangle = \|u\|^2 + \langle u, v\rangle + \langle v, u\rangle + \|v\|^2 \le \|u\|^2 + 2|\langle u, v\rangle| + \|v\|^2.$$

By Cauchy–Schwarz, $|\langle u, v\rangle| \le \|u\|\,\|v\|$, hence

$$\|u + v\|^2 \le (\|u\| + \|v\|)^2.$$

Taking square roots yields $\|u + v\| \le \|u\| + \|v\|$. □

6.3 Orthogonality, orthonormal bases, and Fourier coefficients

Definition 6.3: Orthogonality; orthogonal and orthonormal sets

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space.

- $u \perp v$ means $\langle u, v\rangle = 0$.
- A subset $S \subseteq V$ is *orthogonal* if distinct vectors are pairwise orthogonal.
- A subset $S \subseteq V$ is *orthonormal* if it is orthogonal and $\|u\| = 1$ for all $u \in S$.

Theorem 6.5: Orthogonal nonzero sets are linearly independent

Let $S \subseteq V$ be an orthogonal set consisting of nonzero vectors. Then S is linearly independent.

Proof. Take distinct $u_1, \ldots, u_n \in S$ and suppose $\lambda_1 u_1 + \cdots + \lambda_n u_n = 0$. Fix i and take inner product with u_i:

$$0 = \langle u_i, 0\rangle = \left\langle u_i, \sum_{j=1}^{n} \lambda_j u_j \right\rangle = \sum_{j=1}^{n} \lambda_j \langle u_i, u_j\rangle = \lambda_i \langle u_i, u_i\rangle.$$

Since $\langle u_i, u_i\rangle = \|u_i\|^2 \neq 0$, we get $\lambda_i = 0$. Thus all $\lambda_i = 0$, so the set is linearly independent. □

Definition 6.4: Orthonormal basis

An *orthonormal basis* (ONB) of $(V, \langle\cdot,\cdot\rangle)$ is a basis β of V that is orthonormal.

Theorem 6.6: Fourier coefficient formula in an ONB

Let β be an orthonormal basis of an inner product space $(V, \langle\cdot,\cdot\rangle)$. Then for every $v \in V$,

$$v = \sum_{u \in \beta} \langle u, v\rangle\, u,$$

where only finitely many terms are nonzero if β is infinite (in this course, typically β is finite).

Proof. Write $v = \sum_{u\in\beta} \lambda_u u$ in the basis β. Fix $\tilde{u} \in \beta$ and take inner product with $\tilde{u}$:

$$\langle \tilde{u}, v\rangle = \left\langle \tilde{u}, \sum_{u\in\beta} \lambda_u u \right\rangle = \sum_{u\in\beta} \lambda_u \langle \tilde{u}, u\rangle = \lambda_{\tilde{u}} \langle \tilde{u}, \tilde{u}\rangle = \lambda_{\tilde{u}},$$

since β is orthonormal. Hence $\lambda_u = \langle u, v\rangle$ for all $u \in \beta$. □

Worked Example (Lecture Week 11): checking orthogonality in $\mathbb{R}^3$

In $(\mathbb{R}^3, \cdot)$, let

$$u_1 = (1, 1, 0), \quad u_2 = (1, -1, 1), \quad u_3 = (-1, 1, 2).$$

Compute dot products:

$$u_1 \cdot u_2 = 1 - 1 + 0 = 0, \qquad u_1 \cdot u_3 = -1 + 1 + 0 = 0, \qquad u_2 \cdot u_3 = -1 - 1 + 2 = 0.$$

Hence $\{u_1, u_2, u_3\}$ is an orthogonal subset of $\mathbb{R}^3$.

Worked Example (Lecture Week 11): Fourier coefficients in $P_2(\mathbb{R})$

On $P_2(\mathbb{R})$, define $\langle P, Q\rangle = \int_{-1}^{1} P(x)Q(x)\,dx$. Let

$$\beta = \left(\frac{1}{\sqrt{2}},\ \sqrt{\frac{3}{2}}\,X,\ \sqrt{\frac{5}{8}}\,(3X^2 - 1)\right),$$

which is an orthonormal basis (as given in the lecture problem). For $f = 1 + 4X^2$, the β-Fourier coefficients are $c_k = \langle \beta_k, f\rangle$:

$$c_0 = \frac{1}{\sqrt{2}} \int_{-1}^{1} (1 + 4x^2)\,dx = \frac{14}{3\sqrt{2}}, \qquad c_1 = \sqrt{\frac{3}{2}} \int_{-1}^{1} x(1 + 4x^2)\,dx = 0,$$

$$c_2 = \sqrt{\frac{5}{8}} \int_{-1}^{1} (3x^2 - 1)(1 + 4x^2)\,dx = \sqrt{\frac{5}{8}} \int_{-1}^{1} (12x^4 - x^2 - 1)\,dx = \sqrt{\frac{5}{8}} \cdot \frac{32}{15} = \frac{16\sqrt{5}}{15\sqrt{2}}.$$

Thus

$$1 + 4X^2 = c_0\beta_1 + c_1\beta_2 + c_2\beta_3.$$

6.4 Orthogonal projection onto a vector

Definition 6.5: Orthogonal projection onto a nonzero vector

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space and let $u \in V$ with $u \neq 0$. Define the orthogonal projection of $v \in V$ onto u by

$$\mathrm{proj}_u(v) := \frac{\langle u, v\rangle}{\langle u, u\rangle}\, u.$$

Theorem 6.7: Projection error is orthogonal

If $u \neq 0$, then $u \perp \big(v - \mathrm{proj}_u(v)\big)$.

Proof. Compute

$$\Big\langle u,\ v - \mathrm{proj}_u(v)\Big\rangle = \langle u, v\rangle - \Big\langle u, \frac{\langle u, v\rangle}{\langle u, u\rangle} u\Big\rangle = \langle u, v\rangle - \frac{\langle u, v\rangle}{\langle u, u\rangle}\langle u, u\rangle = 0.$$

□

Common Mistake: using projection formula without checking $u \neq 0$

$\text{proj}_u(v)$ is only defined for $u \neq 0$ (division by $\langle u, u\rangle$). In Gram–Schmidt, the algorithm guarantees each new vector is nonzero; do not assume this unless you have justified it.

6.5 Gram–Schmidt orthogonalization

Theorem 6.8: Gram–Schmidt orthogonalization process

Let $(V, \langle\cdot,\cdot\rangle)$ be a finite-dimensional inner product space with $\dim(V) = n \geq 1$, and let $\beta = (b_1, \ldots, b_n)$ be an ordered basis of V.
Define recursively:

$$v_1 := b_1, \qquad v_k := b_k - \sum_{j=1}^{k-1} \text{proj}_{v_j}(b_k) = b_k - \sum_{j=1}^{k-1} \frac{\langle v_j, b_k\rangle}{\langle v_j, v_j\rangle} v_j, \quad (2 \leq k \leq n).$$

Then $(v_1, \ldots, v_n)$ is an orthogonal ordered basis of V, and

$$\left(\frac{v_1}{\|v_1\|}, \ldots, \frac{v_n}{\|v_n\|}\right)$$

is an orthonormal ordered basis of V. Moreover, for each k,

$$\text{span}(v_1, \ldots, v_k) = \text{span}(b_1, \ldots, b_k).$$

Worked Example (Tutorial 10, Question 1): Gram–Schmidt on an orthogonal basis does nothing

If $(b_1, \ldots, b_n)$ is already orthogonal, then for $j < k$ we have $\langle b_j, b_k\rangle = 0$, so every projection term $\text{proj}_{b_j}(b_k)$ vanishes. Hence the Gram–Schmidt recursion gives $v_k = b_k$ for all k, so the output basis is exactly $(b_1, \ldots, b_n)$ (up to the optional normalization step).

Exam Focus: running Gram–Schmidt efficiently

- Keep the *unnormalized* orthogonal vectors v_k until the end; normalize once.
- Exploit symmetry/parity in integral inner products: odd integrands integrate to 0 on $[-a, a]$.
- In $\mathbb{R}^n$ with the dot product, compute projections using dot products; in matrix-defined inner products $\langle u, v\rangle = u^{\mathsf{T}} G v$ or $u^{\dagger} H v$, precompute Gv or Hv once.

Extension (Tutorial 10, Question 2, Method 1)

Syllabus link: MH1201 Weeks 11–12 — Gram–Schmidt; orthonormal bases.
Lemma (Gram–Schmidt preserves a nested-span flag). Let $\beta = (b_1, \dots, b_n)$ be a basis and let $\gamma = (g_1, \dots, g_n)$ be the orthonormal basis obtained by Gram–Schmidt. Then for each k,

$$g_k \in \operatorname{span}(b_1, \dots, b_k), \qquad \operatorname{span}(g_1, \dots, g_k) = \operatorname{span}(b_1, \dots, b_k).$$

Proof. Let $(v_1, \dots, v_n)$ be the unnormalized Gram–Schmidt vectors and $g_k = v_k/\|v_k\|$. We prove by induction on k that $v_k \in \operatorname{span}(b_1, \dots, b_k)$ and $\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(b_1, \dots, b_k)$.
For $k = 1$, $v_1 = b_1 \in \operatorname{span}(b_1)$ and spans agree.
Assume the claim holds for $k-1$. Then each v_j for $j < k$ lies in $\operatorname{span}(b_1, \dots, b_{k-1})$, so each projection $\operatorname{proj}_{v_j}(b_k)$ is a scalar multiple of v_j and thus lies in $\operatorname{span}(b_1, \dots, b_{k-1})$. Therefore

$$v_k = b_k - \sum_{j=1}^{k-1} \operatorname{proj}_{v_j}(b_k) \in \operatorname{span}(b_1, \dots, b_k).$$

Also, clearly $v_k \in \operatorname{span}(b_1, \dots, b_k)$ implies $\operatorname{span}(v_1, \dots, v_k) \subseteq \operatorname{span}(b_1, \dots, b_k)$. Conversely, $b_k = v_k + \sum_{j=1}^{k-1} \operatorname{proj}_{v_j}(b_k) \in \operatorname{span}(v_1, \dots, v_k)$, and by induction $b_1, \dots, b_{k-1} \in \operatorname{span}(v_1, \dots, v_{k-1}) \subseteq \operatorname{span}(v_1, \dots, v_k)$. Hence $\operatorname{span}(b_1, \dots, b_k) \subseteq \operatorname{span}(v_1, \dots, v_k)$, so equality holds.
Finally, g_k is a nonzero scalar multiple of v_k, so it lies in the same span, and span equalities persist under nonzero rescaling. □
Consequence (upper-triangularity preserved). If a linear operator T satisfies $T(\operatorname{span}(b_1, \dots, b_k)) \subseteq \operatorname{span}(b_1, \dots, b_k)$ for all k, then the same holds with b_i replaced by g_i, so the matrix of T in γ is upper triangular.

6.6 Orthogonal complements

Definition 6.6: Orthogonal complement of a nonempty set

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and let $A \subseteq V$ be nonempty. Define

$$A^{\perp} := \{v \in V : \ \forall u \in A, \ \langle u, v \rangle = 0\}.$$

Theorem 6.9: Orthogonal complements are subspaces

If $A \subseteq V$ is nonempty, then $A^{\perp}$ is a vector subspace of V.

Proof. For all $u \in A$, $\langle u, 0 \rangle = 0$, so $0 \in A^{\perp}$. If $v, w \in A^{\perp}$ and $\lambda \in \mathbb{F}$, then for all $u \in A$,

$$\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle = 0 + 0 = 0, \qquad \langle u, \lambda v \rangle = \lambda \langle u, v \rangle = \lambda \cdot 0 = 0.$$

Hence $v + w, \lambda v \in A^{\perp}$, so $A^{\perp}$ is a subspace. □

Theorem 6.10: Basic properties of orthogonal complements

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space and let $A, B \subseteq V$ be nonempty.

(a) $\{0\}^\perp = V$ and $V^\perp = \{0\}$.

(b) $A \cap A^\perp \subseteq \{0\}$.

(c) If $A \subseteq B$ then $B^\perp \subseteq A^\perp$.

(d) $A^\perp = \operatorname{span}(A)^\perp$.

Theorem 6.11: Computing $W^\perp$ from a basis

Let $W \leq V$ be a subspace and let B be a basis of W. Then

$$W^\perp = \{v \in V : \ \forall b \in B, \ \langle b, v\rangle = 0\}.$$

Proof. If $v \in W^\perp$, then $\langle b, v\rangle = 0$ for all $b \in B \subseteq W$. Conversely, assume $\langle b, v\rangle = 0$ for all $b \in B$. Any $w \in W$ can be written $w = \sum_{b\in B} \lambda_b b$, hence

$$\langle w, v\rangle = \Big\langle \sum_{b\in B} \lambda_b b, \ v \Big\rangle = \sum_{b\in B} \overline{\lambda_b}\, \langle b, v\rangle = 0,$$

so $v \in W^\perp$. □

Key Procedure (Lecture Week 12): finding $A^\perp$ for a finite set A

To compute $A^\perp$ for a nonempty finite set $A = \{a_1, \ldots, a_m\}$:

1. Replace A by a basis B of $\operatorname{span}(A)$ (take a maximal linearly independent subset).
2. Solve the linear system in the unknown $v \in V$:

$$\forall b \in B : \quad \langle b, v\rangle = 0.$$

The solution set is $A^\perp$.

Theorem 6.12: Orthogonal direct sum decomposition (finite-dimensional)

Let $(V, \langle\cdot,\cdot\rangle)$ be a finite-dimensional inner product space and let $W \leq V$. Then

$$V = W \oplus W^\perp, \qquad \text{hence} \qquad \dim(V) = \dim(W) + \dim(W^\perp).$$

Proof. Choose an orthonormal basis $(b_1, \ldots, b_m)$ of W (existence via Gram–Schmidt). For $v \in V$, define

$$p := \sum_{k=1}^{m} \langle b_k, v\rangle\, b_k \in W, \qquad q := v - p.$$

We show $q \in W^\perp$. Let $w \in W$ and write $w = \sum_{j=1}^m \lambda_j b_j$. Then

$$\begin{aligned}\langle w, q\rangle &= \Big\langle \sum_{j=1}^m \lambda_j b_j,\ v - \sum_{k=1}^m \langle b_k, v\rangle b_k \Big\rangle \\ &= \sum_{j=1}^m \overline{\lambda_j}\,\langle b_j, v\rangle - \sum_{j=1}^m \sum_{k=1}^m \overline{\lambda_j}\,\langle b_k, v\rangle\,\langle b_j, b_k\rangle \\ &= \sum_{j=1}^m \overline{\lambda_j}\,\langle b_j, v\rangle - \sum_{j=1}^m \overline{\lambda_j}\,\langle b_j, v\rangle\,\langle b_j, b_j\rangle = 0,\end{aligned}$$

since $\langle b_j, b_k\rangle = 0$ for $j \neq k$ and $\langle b_j, b_j\rangle = 1$. Thus $q \in W^\perp$ and $v = p + q \in W + W^\perp$, so $V = W + W^\perp$.

Finally, if $x \in W \cap W^\perp$, then $\langle x, x\rangle = 0$, hence $x = 0$. So the sum is direct: $V = W \oplus W^\perp$. □

Theorem 6.13: Double orthogonal complements

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space and let $W \leq V$.

(a) Always: $W \subseteq W^{\perp\perp}$.

(b) If $\dim(V) < \infty$, then $W = W^{\perp\perp}$.

Proof. (a) Let $w \in W$. Take any $u \in W^\perp$. Then $\langle u, w\rangle = 0$ by definition of $W^\perp$. By conjugate symmetry, $\langle w, u\rangle = \overline{\langle u, w\rangle} = 0$, hence $w \in W^{\perp\perp}$.

(b) Assume $\dim(V) < \infty$. Apply Theorem 6.12 to W and to $W^\perp$:

$$\dim(V) = \dim(W) + \dim(W^\perp), \qquad \dim(V) = \dim(W^\perp) + \dim(W^{\perp\perp}).$$

Subtracting gives $\dim(W) = \dim(W^{\perp\perp})$. Together with $W \subseteq W^{\perp\perp}$, finite-dimensionality implies $W = W^{\perp\perp}$. □

Worked Example (Tutorial 10, Question 5(c)): strict inclusion in infinite dimensions

In $V = \ell^2$ with $\langle x, y\rangle = \sum_{k=1}^\infty x_k y_k$, let

$$W = c_{00} \coloneqq \{x \in \ell^2 :\ x \text{ has finite support}\}.$$

If $z \in W^\perp$, then for every standard basis vector $e_k \in c_{00}$,

$$0 = \langle z, e_k\rangle = z_k,$$

so $z = 0$ and $W^\perp = \{0\}$. Hence $W^{\perp\perp} = V$, but $W \neq V$, so $W \subsetneq W^{\perp\perp}$.

6.7 Orthogonal projection onto a subspace

Definition 6.7: Orthogonal projection operator onto a subspace

Let $(V, \langle\cdot,\cdot\rangle)$ be a finite-dimensional inner product space and let $W \leq V$. For $v \in V$, write the unique decomposition $v = p + q$ with $p \in W$ and $q \in W^\perp$ (Theorem 6.12). Define $P_W : V \to W$ by $P_W(v) \coloneqq p$.

Key Formula: projection using an orthonormal basis of W

If $(b_1, \ldots, b_m)$ is an orthonormal basis of W, then for all $v \in V$,

$$P_W(v) = \sum_{k=1}^{m} \langle b_k, v \rangle \, b_k, \qquad v - P_W(v) \in W^{\perp}.$$

Extension (Beyond Syllabus: projection as the unique minimizer)

Syllabus link: MH1201 Weeks 11–12 — Orthogonal complements; orthogonal projection.
Theorem. Let $(V, \langle \cdot, \cdot \rangle)$ be finite-dimensional, let $W \leq V$, and let P_W be the orthogonal projection. Then $P_W(v)$ is the unique point in W minimizing $\|v - w\|$ over $w \in W$.
Proof. Fix $v \in V$ and write $v = P_W(v) + r$ with $r \in W^{\perp}$. For any $w \in W$, we have $v - w = (P_W(v) - w) + r$ where $P_W(v) - w \in W$ and $r \in W^{\perp}$, hence $(P_W(v) - w) \perp r$. Therefore

$$\|v - w\|^2 = \|P_W(v) - w\|^2 + \|r\|^2 \geq \|r\|^2 = \|v - P_W(v)\|^2,$$

with equality iff $\|P_W(v) - w\| = 0$, i.e. $w = P_W(v)$. Thus $P_W(v)$ is the unique minimizer. □

Worked Example (Tutorial 10, Question 6): $V^{\perp}$ for a hyperplane and orthogonal decomposition in $\mathbb{R}^4$

Let $V = \{(w, x, y, z) \in \mathbb{R}^4 : \; w - y + z = 0\}$ in $(\mathbb{R}^4, \cdot)$. Write the constraint as $n \cdot (w, x, y, z) = 0$ with $n = (1, 0, -1, 1)$. Hence

$$V^{\perp} = \operatorname{span}\{n\} = \operatorname{span}\{(1, 0, -1, 1)\}.$$

For $p = (1, 2, 3, 4)$, the projection onto $V^{\perp}$ is

$$b = \operatorname{proj}_n(p) = \frac{p \cdot n}{n \cdot n} \, n = \frac{2}{3}(1, 0, -1, 1) = \left(\frac{2}{3}, 0, -\frac{2}{3}, \frac{2}{3}\right),$$

and $a = p - b = \left(\frac{1}{3}, 2, \frac{11}{3}, \frac{10}{3}\right) \in V$ gives $p = a + b$.

6.8 Frobenius inner product and orthogonal complements in matrix spaces

Worked Example (Tutorial 10, Question 7): diagonal matrices and $D^{\perp}$

Let $D = \{A \in \mathbb{R}^{n \times n} : \; A \text{ is diagonal}\}$ in $(\mathbb{R}^{n \times n}, \langle \cdot, \cdot \rangle_F)$. If $X = (x_{ij}) \in D^{\perp}$, then for every diagonal $\Delta = \operatorname{diag}(d_1, \ldots, d_n)$,

$$0 = \langle X, \Delta \rangle_F = \sum_{i=1}^{n} x_{ii} d_i,$$

so $x_{ii} = 0$ for all i. Thus $D^{\perp}$ is the subspace of zero-diagonal matrices, with basis $\{E_{ij} : i \neq j\}$.

Worked Example (Lecture Week 12): upper-triangular matrices and $U^\perp$

Let $U = \{A \in \mathbb{R}^{n\times n} : A \text{ is upper triangular}\}$ in $(\mathbb{R}^{n\times n}, \langle\cdot,\cdot\rangle_F)$. Then $X \in U^\perp$ iff $\langle X, A\rangle_F = \sum_{i,j} x_{ij}a_{ij} = 0$ for all upper triangular A. Choosing $A = E_{ij}$ with $i \le j$ forces $x_{ij} = 0$ for all $i \le j$. Hence $U^\perp$ is the subspace of strictly lower-triangular matrices, with basis $\{E_{ij} : i > j\}$.

6.9 Least squares viewpoint (if assessed in the module)

Lecture Week 12: linear least squares and normal equations

Given data $(x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R}^2$, approximate by a line $f(x) = mx + c$ by minimizing

$$E(m, c) := \sum_{k=1}^{n} |y_k - (mx_k + c)|^2.$$

Let

$$y := \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \qquad A := \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}, \qquad b := \begin{bmatrix} m \\ c \end{bmatrix},$$

so $E(m, c) = \|y - Ab\|^2$ in $(\mathbb{R}^{n,1}, \cdot)$ and $W := \text{range}(A) = \{Ab : b \in \mathbb{R}^{2,1}\}$ is a subspace. Then the following are equivalent:

1. b minimizes $\|y - Ab\|^2$.
2. Ab is the orthogonal projection of y onto W, i.e. $y - Ab \in W^\perp$.
3. (**Normal equations**) $A^\mathsf{T}A\,b = A^\mathsf{T}y$.

Indeed, $y - Ab \in W^\perp$ means $\langle Ab', y - Ab\rangle = 0$ for all b', equivalently $A^\mathsf{T}(y - Ab) = 0$.

Exam Focus: orthogonal complement computations

- Replace A by a basis of $\text{span}(A)$ before imposing orthogonality equations.
- For $W = \{x : Mx = 0\}$ in $(\mathbb{R}^n, \cdot)$, a fast route is $W^\perp = \text{span}\{\text{rows of } M\}$ (view constraints as dot products with normals).
- In Frobenius spaces, test against matrix units E_{ij} to force individual entries to vanish.

Common Mistake: "orthogonal" vs. "orthonormal"

Orthogonal means pairwise perpendicular; orthonormal additionally requires each vector has norm 1. Fourier coefficient formula $v = \sum\langle u, v\rangle u$ requires an *orthonormal* basis. For an orthogonal (not normalized) basis, coefficients involve divisions by $\langle u, u\rangle$.

7 Comprehensive Course Summary and Final Review

This section is a course-wide synthesis of MH1201. It is intended to sit *after* the six main chapters, not to replace them. The goal is to compress the module into a logically connected, exam-ready narrative that preserves the course's main abstractions, theorem hypotheses, standard workflows, and recurrent traps.

Exam Focus: what the recent finals reward most strongly

Across the uploaded finals, the heaviest recurrent demands are:

- moving fluently between an *abstract linear operator* and its *matrix representation* relative to a specified ordered basis;
- computing Null(T), Range(T), eigenvalues, eigenspaces, and diagonalizability from that representation;
- using a diagonalizing basis to compute $T^k(v)$ or A^k;
- handling *non-standard inner products*, then applying Gram–Schmidt, Fourier coefficients, orthogonal complements, projections, and QR-style factorization;
- proving dimension statements cleanly, especially by basis extension/extraction, rank–nullity, and dimension formulas for sums or orthogonal complements.

A strong exam answer in MH1201 is usually not long because it is computationally heavy; it is strong because it uses the *right structural theorem at the right moment*.

7.1 The logical architecture of MH1201

The course is best viewed as a three-stage progression.

Stage 1: Algebraic structure of spaces. One first learns what vector spaces and subspaces are, then how spaces are built from spans, how redundancy is detected by linear dependence, and how bases and dimension quantify the true number of independent directions.

Stage 2: Linear maps as the central objects. The course then shifts from sets of vectors to maps $T : V \to W$. The central questions become: what does T kill, what can T reach, when is T invertible, and how do coordinates convert T into a matrix?

Stage 3: Structure theory and geometry. After that, one asks whether an operator can be simplified by a good basis: first via eigenvalues/eigenvectors and diagonalization, then via inner products, orthogonality, orthonormal bases, projections, and best approximation.

Policy Application

Unifying principle. Almost every major theorem in MH1201 can be understood as answering one of the following:

- What is the right basis?
- What is the right invariant?
- What is the right decomposition?

Span, basis, direct sum, rank–nullity, similarity, diagonalization, and orthogonal decomposition are all manifestations of this single organising theme.

7.2 Vector spaces, subspaces, span, and direct sums

1. Vector spaces and the subspace viewpoint

A vector space over $\mathbb{F}$ is a set V together with addition and scalar multiplication satisfying the usual axioms. In this course, the most important ambient examples are

$$\mathbb{F}^n, \qquad \mathbb{F}^{m,n}, \qquad P_n(\mathbb{F}), \qquad \mathcal{F}(S,\mathbb{F}).$$

In computations, however, the more important object is usually a *subspace* of a known ambient vector space.

Definition 7.1: Subspace

Let V be a vector space over $\mathbb{F}$ and let $W \subseteq V$. Then W is a subspace of V if it is a vector space under the restricted operations inherited from V.

Theorem 7.1: Subspace Test

For $W \subseteq V$,

$$W \leq V \iff \big(W \neq \varnothing,\ u, v \in W \Rightarrow u+v \in W,\ c \in \mathbb{F},\ u \in W \Rightarrow cu \in W\big).$$

Equivalently, one may use the one-line version

$$W \leq V \iff \big(W \neq \varnothing \text{ and } \forall u, v \in W, \forall c \in \mathbb{F} :\ cu+v \in W\big).$$

Exam Focus: fast subspace proofs

There are three standard proof patterns.

1. Apply the Subspace Test directly.
2. Recognise the set as a kernel of a linear map.
3. Show the set equals $\mathrm{Span}(S)$ for some S.

The second method is especially high-yield: whenever a set is defined by *homogeneous linear conditions*, it is usually a kernel.

Common Mistake: checking closure but forgetting the origin

A subset can be closed under certain operations and still fail to be a subspace if $0 \notin W$. Affine sets such as

$$\{x \in \mathbb{R}^n : a^{\mathsf{T}}x = b\}, \qquad b \neq 0,$$

typically fail precisely because they do not pass through the origin.

2. Span and generated structure

Definition 7.2: Span

For $S \subseteq V$,

$$\mathrm{Span}(S) = \{\text{all finite linear combinations of vectors from } S\}.$$

The word *finite* matters. This is not cosmetic; it is part of the definition.

Theorem 7.2: Span as smallest containing subspace

For every $S \subseteq V$,

$$\mathrm{Span}(S) \leq V, \qquad S \subseteq \mathrm{Span}(S),$$

and if $S \subseteq W \leq V$, then

$$\mathrm{Span}(S) \subseteq W.$$

Thus $\mathrm{Span}(S)$ is the smallest subspace of V containing S.

Key Formula: basic span identities

For $A, B \subseteq V$,

$$A \subseteq B \Rightarrow \mathrm{Span}(A) \subseteq \mathrm{Span}(B), \qquad \mathrm{Span}(A \cup B) = \mathrm{Span}(A) + \mathrm{Span}(B),$$

$$\mathrm{Span}(\mathrm{Span}(A)) = \mathrm{Span}(A), \qquad \mathrm{Span}(\varnothing) = \{0\}.$$

3. Sums, intersections, and direct sums

If $U, W \leq V$, then both $U \cap W$ and $U + W$ are subspaces. The crucial distinction is between an ordinary sum and a *direct* sum.

Definition 7.3: Direct sum

For subspaces $V_1, \ldots, V_n \leq V$,

$$W = V_1 \oplus \cdots \oplus V_n$$

means that $W = V_1 + \cdots + V_n$ and each $w \in W$ has a unique decomposition

$$w = v_1 + \cdots + v_n, \qquad v_i \in V_i.$$

Theorem 7.3: Direct-sum criteria

For two subspaces $U, W \leq V$,

$$V = U \oplus W \iff V = U + W \text{ and } U \cap W = \{0\}.$$

More generally,

$$V_1 + \cdots + V_n \text{ is direct} \iff \Big(\sum_{i=1}^{n} v_i = 0,\ v_i \in V_i \Rightarrow v_1 = \cdots = v_n = 0\Big).$$

The general zero-sum criterion is the safer one for $n \geq 3$.

Common Mistake: pairwise trivial intersections are not enough for $n \geq 3$

For more than two subspaces, it is *false* that

$$V_i \cap V_j = \{0\}\ \forall i \neq j$$

automatically implies the sum is direct. One must check the full direct-sum condition.

7.3 Linear independence, bases, coordinates, and dimension

1. Linear independence and dependence

Definition 7.4: Linear independence

A subset $S \subseteq V$ is linearly independent if every finite relation

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$$

with distinct $v_i \in S$ forces

$$\lambda_1 = \cdots = \lambda_n = 0.$$

Otherwise S is linearly dependent.

Theorem 7.4: Dependence iff one vector lies in the span of the others

A set S is linearly dependent if and only if one vector of S lies in the span of finitely many others from S.

This theorem explains why dependence is exactly the same as redundancy.

Key Facts: quick tests for independence

$$\varnothing \text{ is linearly independent}, \qquad \{v\} \text{ is linearly independent} \iff v \neq 0,$$
$$S \text{ linearly independent} \Rightarrow \text{every subset of } S \text{ is linearly independent},$$
$$S \text{ dependent} \Rightarrow \text{every superset of } S \text{ is dependent}.$$

For vectors $v_1, \ldots, v_n \in \mathbb{F}^{m,1}$, linear independence is equivalent to the homogeneous system

$$[\, v_1 \ \cdots \ v_n \,]\lambda = 0$$

having only the trivial solution.

2. Bases and coordinates

Definition 7.5: Basis and coordinates

A basis of V is a set $B \subseteq V$ that is both linearly independent and spanning. If $\beta = (b_1, \ldots, b_n)$ is an ordered basis and

$$v = \lambda_1 b_1 + \cdots + \lambda_n b_n,$$

then the coordinate vector of v relative to β is

$$[v]_\beta = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Theorem 7.5: Uniqueness in a basis

If B is a basis of V, then every $v \in V$ admits a unique finite expansion in vectors of B.

This uniqueness is exactly what makes coordinate vectors meaningful.

Common Mistake: forgetting that order matters

A basis is a set, but coordinates require an *ordered* basis. Reordering the same basis changes the coordinate vector and changes the matrix representation of a linear map.

3. Dimension and finite-dimensional structure

Definition 7.6: Dimension

If V has a finite basis, then $\dim(V)$ is the number of vectors in any basis of V.

Theorem 7.6: Finite-dimensional size logic

If $\dim(V) = n$, then:

- every linearly independent subset of V has size at most n;
- every spanning subset of V has size at least n;
- any linearly independent set of n vectors is a basis;
- any spanning set of n vectors is a basis.

These are among the highest-yield finite-dimensional results in the course.

Theorem 7.7: Basis extension and extraction

Assume V is finite-dimensional.

- Every linearly independent set extends to a basis of V.
- Every spanning set of V contains a basis of V.

Exam Focus: the two standard basis constructions

To *build* a basis, start from an independent set and add vectors outside the current span. To *prune* to a basis, start from a spanning set and remove redundant vectors one by one.

4. Dimension formulae

Theorem 7.8: Dimension of a sum

If $U, W \leq V$ and V is finite-dimensional, then

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

Key Formula: direct sums and dimensions

If $V = U \oplus W$, then

$$\dim(V) = \dim(U) + \dim(W).$$

More generally, the direct-sum condition is equivalent to saying that dimensions add *without overlap*.

This formula is one of the main dimension bridges in the course. It is routinely used to show:

- an intersection is trivial or nontrivial;
- a sum is the whole space;
- a proposed decomposition is direct;
- a missing dimension can be computed from known pieces.

7.4 Linear transformations, kernel, range, rank, and invertibility

1. Linearity and the space $\mathcal{L}(V, W)$

Definition 7.7: Linear transformation

A map $T : V \to W$ is linear if

$$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$$

for all $u, v \in V$ and all $\alpha, \beta \in \mathbb{F}$.

The set $\mathcal{L}(V, W)$ of all linear maps from V to W is itself a vector space under pointwise addition and scalar multiplication. This matters conceptually because one can study linear maps with the same structural tools used to study vectors.

2. Null space and range

Definition 7.8: Kernel and range

For $T \in \mathcal{L}(V, W)$,

$$\text{Null}(T) = \ker(T) = \{v \in V : T(v) = 0\}, \qquad \text{Range}(T) = \{T(v) : v \in V\}.$$

Theorem 7.9: Kernel and range are subspaces

If $T \in \mathcal{L}(V, W)$, then

$$\text{Null}(T) \leq V, \qquad \text{Range}(T) \leq W.$$

These two subspaces measure the two most important structural effects of T:

- $\text{Null}(T)$: what T collapses to 0;
- $\text{Range}(T)$: what outputs T can actually produce.

3. Rank–Nullity and injective/surjective logic

Definition 7.9: Rank and nullity

If $T : V \to W$ is linear and V is finite-dimensional, define

$$\text{Rank}(T) = \dim(\text{Range}(T)), \qquad \text{Nullity}(T) = \dim(\text{Null}(T)).$$

Theorem 7.10: Rank–Nullity Theorem

If $T : V \to W$ is linear and V is finite-dimensional, then

$$\dim(V) = \text{Rank}(T) + \text{Nullity}(T).$$

Key Formula: injective, surjective, invertible

For linear $T : V \to W$,

$$T \text{ injective} \iff \text{Null}(T) = \{0\}.$$

If V and W are finite-dimensional, then

$$T \text{ surjective} \iff \text{Rank}(T) = \dim(W).$$

If moreover $\dim(V) = \dim(W)$, then the following are equivalent:

$$T \text{ injective} \iff T \text{ surjective} \iff T \text{ bijective} \iff T \text{ invertible}.$$

Exam Focus: the equal-dimension shortcut

Whenever $T : V \to W$ with $\dim(V) = \dim(W)$, you should immediately think:

$$\ker(T) = \{0\} \iff \text{Range}(T) = W \iff T \text{ invertible}.$$

This shortcut appears repeatedly in both theory questions and computational questions.

4. Invertible linear maps and isomorphism

A linear map $T : V \to W$ is invertible if there exists a map $T^{-1} : W \to V$ such that

$$T^{-1}T = I_V, \qquad TT^{-1} = I_W.$$

If T is invertible, then T^{-1} is also linear.

This is the correct notion of two vector spaces being "the same" from the standpoint of linear algebra: they are *isomorphic*.

7.5 Matrices, coordinates, change of basis, and similarity

1. Matrix representations

Fix ordered bases $\beta = (v_1, \ldots, v_n)$ of V and $\gamma = (w_1, \ldots, w_m)$ of W. For $T \in \mathcal{L}(V, W)$, the matrix of T relative to (β, γ) is the unique matrix $[T]_\beta^\gamma \in \mathbb{F}^{m,n}$ satisfying

$$[T(v)]_\gamma = [T]_\beta^\gamma [v]_\beta \qquad \forall v \in V.$$

Theorem 7.11: Fundamental coordinate identity

If $[T]_\beta^\gamma$ is the matrix representation of T, then for each basis vector $v_j \in \beta$, the j-th column of $[T]_\beta^\gamma$ is

$$[T(v_j)]_\gamma.$$

Equivalently,

$$[T(v)]_\gamma = [T]_\beta^\gamma [v]_\beta$$

for every $v \in V$.

Common Mistake: mixing up rows, columns, and basis roles

If $T : V \to W$, then

$$[T]_\beta^\gamma \in \mathbb{F}^{(\dim W)\times(\dim V)}.$$

The *columns* come from the images of *domain-basis* vectors expressed in the *codomain basis*. Many sign and dimension errors come from reversing this.

2. Composition and matrix multiplication

If $T : U \to V$ and $S : V \to W$, then

$$[S \circ T]_\alpha^\delta = [S]_\beta^\delta [T]_\alpha^\beta$$

provided the middle basis β is the same on both sides. Matrix multiplication is therefore not an arbitrary formal rule; it is the coordinate shadow of composition of linear maps.

3. Change of coordinates

Definition 7.10: Change-of-coordinate matrix

If β and γ are ordered bases of the same vector space V, then

$$[I_V]_\beta^\gamma$$

is the change-of-coordinate matrix from β-coordinates to γ-coordinates. Thus

$$[v]_\gamma = [I_V]_\beta^\gamma [v]_\beta.$$

Key Formula: coordinate-change inverse relation

$$[I_V]_\gamma^\beta = \left([I_V]_\beta^\gamma\right)^{-1}.$$

This is one of the most frequently mishandled pieces of notation in the course. Always read the subscripts and superscripts as *from* the lower basis *to* the upper basis.

4. Change of matrix representation and similarity

If $T \in \mathcal{L}(V)$ and β, γ are ordered bases of V, then

$$[T]_\gamma^\gamma = [I_V]_\beta^\gamma \, [T]_\beta^\beta \, [I_V]_\gamma^\beta.$$

Thus two matrices representing the same operator in different ordered bases are similar.

Definition 7.11: Similarity

Matrices $A, B \in \mathbb{F}^{n,n}$ are similar if there exists an invertible matrix Q such that

$$B = Q^{-1}AQ.$$

Theorem 7.12: Similarity invariants used in MH1201

If $A \sim B$, then A and B have the same:

- determinant,
- trace,
- characteristic polynomial,
- eigenvalues (with algebraic multiplicities),
- rank,
- nullity.

Exam Focus: how to show two matrices are not similar

The quickest route is often to exhibit a similarity invariant that differs:

$$\det(A) \neq \det(B), \quad \operatorname{tr}(A) \neq \operatorname{tr}(B), \quad p_A(\lambda) \neq p_B(\lambda),$$

or a different rank/nullity/eigenvalue profile.

7.6 Eigenvalues, eigenspaces, diagonalization, and operator structure

1. Eigenvalues and eigenspaces

Definition 7.12: Eigenvalue and eigenspace

Let $T \in \mathcal{L}(V)$. A scalar $\lambda \in \mathbb{F}$ is an eigenvalue of T if there exists $v \neq 0$ such that

$$T(v) = \lambda v.$$

The corresponding eigenspace is

$$E_{T,\lambda} = \operatorname{Null}(T - \lambda I_V).$$

Theorem 7.13: Equivalent eigenvalue tests

For $T \in \mathcal{L}(V)$,

$$\lambda \text{ is an eigenvalue} \iff E_{T,\lambda} \neq \{0\} \iff T - \lambda I_V \text{ is not injective.}$$

If V is finite-dimensional and β is an ordered basis, then

$$\lambda \text{ is an eigenvalue} \iff \det\big([T]_\beta^\beta - \lambda I\big) = 0.$$

Common Mistake: the field matters

A real matrix may have no real eigenvalues but still have complex eigenvalues. Eigenvalues are always computed *in the specified field*. The standard 2×2 rotation matrix is the canonical example over $\mathbb{R}$ versus $\mathbb{C}$.

2. Characteristic polynomial

Definition 7.13: Characteristic polynomial

If V is finite-dimensional and β is an ordered basis of V, define

$$p_T(\lambda) = \det([T]_\beta^\beta - \lambda I).$$

This is independent of the choice of basis.

The basis-independence matters because the characteristic polynomial is an operator invariant, not a coordinate accident.

Key Formula: characteristic roots and easy checks

$$\lambda \text{ eigenvalue} \iff p_T(\lambda) = 0.$$

For $A \in \mathbb{F}^{2,2}$,

$$p_A(\lambda) = \lambda^2 - (\operatorname{tr} A)\lambda + \det(A),$$

so trace and determinant are often enough to speed up the eigenvalue computation.

3. Direct-sum structure of eigenspaces

Theorem 7.14: Distinct eigenspaces add directly

If $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of T, then

$$E_{T,\lambda_1} + \cdots + E_{T,\lambda_k} = E_{T,\lambda_1} \oplus \cdots \oplus E_{T,\lambda_k}.$$

In particular, eigenvectors corresponding to distinct eigenvalues are linearly independent.

This is one of the most important bridges in the course: the direct-sum machinery from the first half of the course reappears as the structural backbone of eigenvalue theory.

4. Diagonalization

Definition 7.14: Diagonalizability

An operator $T \in \mathcal{L}(V)$ is diagonalizable if there exists an ordered basis of V consisting of eigenvectors of T. Equivalently, there exists an invertible P and diagonal D such that

$$[T]_\sigma^\sigma = PDP^{-1}$$

for some ordered basis σ.

Theorem 7.15: Fundamental diagonalization criterion

For finite-dimensional V,

$$T \text{ is diagonalizable} \iff V = \bigoplus_\lambda E_{T,\lambda}.$$

Equivalently,

$$T \text{ is diagonalizable} \iff \sum_\lambda \dim(E_{T,\lambda}) = \dim(V),$$

where the sum runs over distinct eigenvalues in the field.

Theorem 7.16: Multiplicity criterion

Assume the characteristic polynomial splits over $\mathbb{F}$. Then T is diagonalizable if and only if, for each eigenvalue λ,

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

Key Formula: distinct-eigenvalue criterion

If $\dim(V) = n$ and T has n distinct eigenvalues in $\mathbb{F}$, then T is diagonalizable.

Common Mistake: repeated eigenvalue does not force failure

The statement

$$\text{“repeated eigenvalue”} \Rightarrow \text{“not diagonalizable”}$$

is false. Repeated eigenvalues only mean that extra checking is needed. What matters is whether enough linearly independent eigenvectors exist.

Exam Focus: complete diagonalizability workflow

Given $A \in \mathbb{F}^{n,n}$, the standard workflow is:

1. compute $p_A(\lambda)$;
2. find the eigenvalues in the correct field;
3. compute each eigenspace $\mathrm{Null}(A - \lambda I)$;
4. compare the total eigenspace dimension with n, or compare algebraic and geometric multiplicities;
5. if diagonalizable, form P from an eigenbasis and D from the matching eigenvalues in the same column order.

5. Powers and polynomial functions of operators

If

$$A = PDP^{-1},$$

then

$$A^k = PD^kP^{-1}$$

for every $k \in \mathbb{N}$, and more generally

$$p(A) = P\,p(D)\,P^{-1}$$

for polynomials p. This is why diagonalization is computationally useful and not just theoretically elegant.

Worked Illustration: why diagonalization simplifies iteration

If $\beta = (v_1, \ldots, v_n)$ is an eigenbasis with $T(v_i) = \lambda_i v_i$, and

$$v = c_1v_1 + \cdots + c_nv_n,$$

then

$$T^k(v) = c_1\lambda_1^k v_1 + \cdots + c_n\lambda_n^k v_n.$$

Thus repeated application of T becomes scalar exponentiation once one has moved into an eigenbasis.

6. Edge cases and Jordan-type failure

The uploaded chapter is titled "Eigenvalues, Diagonalization, and Jordan Canonical Form", but the course emphasis in the notes is much more heavily on:

- eigenspaces and direct-sum structure,
- diagonalizability criteria,
- defective examples,
- nilpotent diagnostics,
- solving operator equations using spectral information.

So the correct revision priority is diagonalization and its failure modes, not a full Jordan algorithm.

Exam Focus: nilpotent operators

If T is nilpotent, then the only possible eigenvalue is 0. Hence a nonzero nilpotent operator cannot be diagonalizable, since a diagonalizable operator with only eigenvalue 0 would have to be the zero operator.

7.7 Inner products, orthogonality, orthonormal bases, and projections

1. Inner products and induced norm

Definition 7.15: Inner product

An inner product on a vector space V over $\mathbb{F}$ is a map

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$$

satisfying linearity/sesquilinearity, conjugate symmetry, and positive definiteness.

The course uses both standard Euclidean inner products and weighted or matrix-defined inner products such as

$$\langle x, y \rangle = x^{\mathsf{T}} A y$$

for a symmetric positive definite matrix A, as well as the Frobenius inner product on matrix spaces:

$$\langle A, B \rangle_F = \operatorname{tr}(A^{\mathsf{T}} B).$$

Definition 7.16: Induced norm

The norm induced by an inner product is

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

Theorem 7.17: Cauchy–Schwarz and triangle inequality

For all $u, v \in V$,

$$|\langle u, v \rangle| \le \|u\| \, \|v\|, \qquad \|u + v\| \le \|u\| + \|v\|.$$

Equality in Cauchy–Schwarz holds if and only if u and v are linearly dependent.

Common Mistake: complex linearity in the wrong slot

In complex vector spaces, one slot is conjugate-linear. Before expanding an inner product, fix the course convention and stick to it consistently.

2. Orthogonality and orthonormal bases

Definition 7.17: Orthogonality and ONB

Vectors $u, v \in V$ are orthogonal if $\langle u, v\rangle = 0$. A set is orthogonal if distinct vectors are pairwise orthogonal. A set is orthonormal if it is orthogonal and every vector has norm 1. An orthonormal basis is an orthonormal set that is also a basis.

Theorem 7.18: Orthogonal nonzero sets are linearly independent

Any orthogonal set of nonzero vectors is linearly independent.

Theorem 7.19: Fourier coefficient formula

If β is an orthonormal basis of V, then every $v \in V$ has the expansion

$$v = \sum_{u\in\beta} \langle u, v\rangle\, u.$$

Key Formula: orthogonal versus orthonormal expansions

If $(u_1, \ldots, u_n)$ is *orthogonal* but not normalized, then

$$v = \sum_{k=1}^{n} \frac{\langle u_k, v\rangle}{\langle u_k, u_k\rangle}\, u_k.$$

If $(u_1, \ldots, u_n)$ is *orthonormal*, this simplifies to

$$v = \sum_{k=1}^{n} \langle u_k, v\rangle\, u_k.$$

Common Mistake: orthogonal $\neq$ orthonormal

Many exam errors come from using the orthonormal formula when the basis has not been normalized.

3. Projection onto a vector and Gram–Schmidt

Definition 7.18: Projection onto a nonzero vector

For $u \neq 0$, define

$$\mathrm{proj}_u(v) = \frac{\langle u, v\rangle}{\langle u, u\rangle} u.$$

The point is that $v - \mathrm{proj}_u(v)$ is orthogonal to u, so projection separates the component of v along u from the error perpendicular to u.

Theorem 7.20: Gram–Schmidt

Given a basis $(b_1, \dots, b_n)$ of a finite-dimensional inner product space, Gram–Schmidt constructs an orthogonal basis $(v_1, \dots, v_n)$ via

$$v_1 = b_1, \qquad v_k = b_k - \sum_{j=1}^{k-1} \operatorname{proj}_{v_j}(b_k) \quad (k \geq 2),$$

and then produces an orthonormal basis by normalizing the v_k.

Exam Focus: efficient Gram–Schmidt

A good exam solution keeps the v_k orthogonal but unnormalized until the end. Normalizing too early usually creates unnecessary radicals and increases the chance of algebra mistakes.

4. Orthogonal complements

Definition 7.19: Orthogonal complement

For $A \subseteq V$,

$$A^{\perp} = \{v \in V : \langle v, a \rangle = 0 \text{ for all } a \in A\}.$$

Theorem 7.21: Core orthogonal-complement facts

For $A \subseteq V$ and finite-dimensional $W \leq V$:

- $A^{\perp}$ is a subspace;
- $A^{\perp} = \operatorname{Span}(A)^{\perp}$;
- $V = W \oplus W^{\perp}$;
- $\dim(W) + \dim(W^{\perp}) = \dim(V)$;
- $W^{\perp\perp} = W$.

The relation $V = W \oplus W^{\perp}$ is the geometric strengthening of direct-sum theory from the first half of the course.

5. Orthogonal projection onto a subspace

Definition 7.20: Orthogonal projection onto a subspace

Let $W \leq V$ be finite-dimensional. For each $v \in V$, write uniquely

$$v = w + z, \qquad w \in W,\ z \in W^{\perp}.$$

Then the orthogonal projection of v onto W is

$$\mathrm{proj}_W(v) = w.$$

Key Formula: projection using an orthonormal basis of W

If $(u_1, \ldots, u_m)$ is an orthonormal basis of W, then

$$\mathrm{proj}_W(v) = \sum_{k=1}^{m} \langle u_k, v \rangle u_k.$$

Policy Application

Best-approximation interpretation. The vector $\mathrm{proj}_W(v)$ is the unique point in W closest to v. This is the conceptual reason that least-squares problems are really projection problems in disguise.

6. Least squares, QR, and Frobenius geometry

The course notes explicitly connect projection to least squares and, through Gram–Schmidt, to QR decomposition.

If the columns of A are linearly independent and $A = QR$ with Q having orthonormal columns, then solving

$$\min_x \|Ax - b\|^2$$

amounts to projecting b onto $\mathrm{Range}(A)$. In coordinate form this leads to the normal equations

$$A^{\mathsf{T}}Ax = A^{\mathsf{T}}b.$$

Similarly, under the Frobenius inner product on $\mathbb{R}^{n,n}$, one can compute orthogonal complements of natural matrix subspaces such as diagonal, upper-triangular, symmetric, or skew-symmetric matrices. This is exactly the same orthogonal-complement theory as before, just carried out in a matrix space.

QRS@NTU — qrsntu.org

7.8 Standard problem-solving workflows

Exam Focus: high-frequency workflows and what each answer should look like

Workflow 1: Testing a subset for subspace status. Either apply the Subspace Test directly, or identify the set as a kernel/span. *Expected answer form:* a short closure proof, or a one-line kernel argument.

Workflow 2: Testing linear independence. Set a linear combination equal to 0, or row-reduce a matrix with the vectors as columns. *Expected answer form:* the trivial-solution conclusion.

Workflow 3: Proving a set is a basis. Show independence plus spanning, or use the finite-dimensional shortcut "n vectors in an n-dimensional space". *Expected answer form:* one theorem invocation plus a clean verification.

Workflow 4: Computing coordinates. Solve

$$v = \lambda_1 b_1 + \cdots + \lambda_n b_n$$

and package the coefficients into $[v]_\beta$. *Expected answer form:* a column vector.

Workflow 5: Computing $[T]_\beta^\gamma$. Find $T(b_j)$ for each domain basis vector b_j, express each image in the codomain basis γ, and place those coordinate vectors as columns. *Expected answer form:* a matrix whose columns are image coordinates.

Workflow 6: Using rank–nullity. Compute either kernel or range explicitly, then use

$$\dim(V) = \text{Rank}(T) + \text{Nullity}(T)$$

to infer the missing quantity and deduce injective/surjective behavior. *Expected answer form:* dimension arithmetic plus a property conclusion.

Workflow 7: Changing basis. Use $[I_V]_\beta^\gamma$ to convert coordinates, then use

$$[T]_\gamma^\gamma = [I_V]_\beta^\gamma [T]_\beta^\beta [I_V]_\gamma^\beta.$$

Expected answer form: a similarity transform with correct basis direction.

Workflow 8: Computing eigenvalues and eigenspaces. Find roots of the characteristic polynomial, then compute each null space $\text{Null}(A - \lambda I)$. *Expected answer form:* eigenvalues followed by spanning descriptions of eigenspaces.

Workflow 9: Checking diagonalizability. Either show n distinct eigenvalues, or show the total eigenspace dimension is n, or compare algebraic versus geometric multiplicity. *Expected answer form:* one decisive criterion, not several pages of unnecessary work.

Workflow 10: Running Gram–Schmidt / finding QR. Produce orthogonal vectors first, normalize them, form Q from the orthonormal columns, and recover R from coordinate relations or $R = Q^\top A$ in the Euclidean case. *Expected answer form:* explicit orthonormal basis and upper-triangular R.

Workflow 11: Computing $W^\perp$ and $\text{proj}_W(v)$. Turn orthogonality conditions into linear equations; if needed, first replace the spanning set of W by an orthonormal basis. *Expected answer form:* a spanning description of $W^\perp$ and a projection formula/value.

Workflow 12: Solving linear operator equations on polynomial spaces. Represent the operator in the standard ordered basis, work in coordinates, and convert back to polynomial form. *Expected answer form:* either an explicit polynomial, or an inverse operator formula written in polynomial language.

7.9 Common mistakes and hidden assumptions

Common Mistake Cluster I: foundational structure

Do not confuse:

- **subset** with **subspace**;
- **span** with **basis**;
- **sum** with **direct sum**;
- **finite-dimensional theorem** with **general theorem without hypotheses**.

Always track the field $\mathbb{F}$, the ambient space, and whether the statement is being made for sets, ordered bases, matrices, or operators.

Common Mistake Cluster II: maps and matrices

Do not confuse:

- T with $[T]_\beta^\gamma$;
- the *standard matrix* with the matrix in a nonstandard basis;
- domain basis versus codomain basis;
- similarity of matrices with equality of matrices.

The matrix is a representation of the map, not the map itself.

Common Mistake Cluster III: spectral and inner-product issues

Do not confuse:

- eigenvalue with eigenvector;
- algebraic multiplicity with geometric multiplicity;
- repeated eigenvalue with non-diagonalizability;
- orthogonal with orthonormal;
- arbitrary basis with orthonormal basis;
- projection onto a vector with projection onto a subspace.

In complex spaces, also track the conjugation convention in the inner product.

7.10 Representative worked illustrations

Worked Illustration 1: proving a direct sum by dimensions

Suppose V is finite-dimensional, $V = U + W$, and

$$\dim(U) + \dim(W) = \dim(V).$$

Then the dimension formula gives

$$\dim(U \cap W) = \dim(U) + \dim(W) - \dim(U + W) = 0.$$

Hence $U \cap W = \{0\}$, so

$$V = U \oplus W.$$

Why this matters. This is the cleanest standard dimension-based direct-sum proof.

Worked Illustration 2: injective versus surjective via Rank–Nullity

Let $T : V \to W$ be linear with $\dim(V) = \dim(W) = n$. If $\text{Null}(T) = \{0\}$, then $\text{Nullity}(T) = 0$. By Rank–Nullity,

$$n = \text{Rank}(T) + 0,$$

so $\text{Rank}(T) = n = \dim(W)$. Hence $\text{Range}(T) = W$, so T is surjective. Therefore T is invertible. **Why this matters.** This is the canonical equal-dimension shortcut.

Worked Illustration 3: diagonalizability with a repeated eigenvalue

Suppose $A \in \mathbb{R}^{3,3}$ has characteristic polynomial

$$p_A(\lambda) = (\lambda - 2)^2(\lambda - 5).$$

Then 2 has algebraic multiplicity 2, while 5 has algebraic multiplicity 1. If

$$\dim E_{A,2} = 1, \qquad \dim E_{A,5} = 1,$$

then the total eigenspace dimension is

$$1 + 1 = 2 < 3.$$

Hence A is not diagonalizable. **Why this matters.** A repeated eigenvalue does not settle the question by itself; one must compute the eigenspace.

Worked Illustration 4: projection onto a subspace

Let $W \leq V$ have orthonormal basis (u_1, u_2). Then for every $v \in V$,

$$\mathrm{proj}_W(v) = \langle u_1, v\rangle u_1 + \langle u_2, v\rangle u_2.$$

The error

$$v - \mathrm{proj}_W(v)$$

lies in $W^\perp$. Hence $\mathrm{proj}_W(v)$ is the unique vector in W closest to v. **Why this matters.** This single formula is the source of Fourier expansion, orthogonal decomposition, and least-squares best approximation.

QRS@NTU — qrsntu.org

QRS@NTU — qrsntu.org

7.11 Final comprehensive revision checklist

Final Revision Checklist: "Given ⋯, be able to ⋯"

1. Given a subset of a known vector space, be able to determine whether it is a subspace and justify the answer with the correct closure and origin logic.
2. Given a family of vectors or functions, be able to compute or describe its span and explain why that span is the smallest containing subspace.
3. Given subspaces U, W, be able to distinguish $U + W$ from $U \oplus W$, and decide directness correctly.
4. Given vectors $v_1, \ldots, v_n$, be able to test linear independence by a homogeneous system or by the span criterion for dependence.
5. Given a linearly independent set in a finite-dimensional space, be able to extend it to a basis; given a spanning set, be able to extract a basis from it.
6. Given a finite-dimensional space and a family of vectors of the correct size, be able to use the "independent $\Leftrightarrow$ spanning" shortcut appropriately.
7. Given $U, W \leq V$, be able to apply
$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$$
to compute unknown dimensions or prove direct-sum statements.
8. Given a map $T : V \to W$, be able to test linearity directly.
9. Given a linear map, be able to compute $\mathrm{Null}(T)$, $\mathrm{Range}(T)$, rank, and nullity.
10. Given $\dim(V)$, rank, or nullity data, be able to deploy Rank–Nullity and deduce injective/surjective/invertible conclusions.
11. Given an ordered basis β, be able to compute coordinate vectors $[v]_\beta$.
12. Given T and ordered bases β, γ, be able to compute $[T]_\beta^\gamma$ from images of basis vectors.
13. Given two ordered bases of the same vector space, be able to compute $[I_V]_\beta^\gamma$ and use it to convert coordinates.
14. Given two matrix representations of the same operator in different bases, be able to relate them by similarity.
15. Given $A \in \mathbb{F}^{n,n}$, be able to compute the characteristic polynomial and identify eigenvalues in the correct field.
16. Given an eigenvalue λ, be able to compute $E_{A,\lambda} = \mathrm{Null}(A - \lambda I)$.
17. Given algebraic and geometric multiplicities, be able to decide diagonalizability and explain *why*.
18. Given a diagonalizable matrix $A = PDP^{-1}$, be able to compute A^k, $p(A)$, or $T^k(v)$ efficiently.
19. Given a nilpotent or otherwise special operator, be able to use structural shortcuts to reason about diagonalizability.
20. Given a proposed bilinear or sesquilinear form, be able to verify whether it defines an inner product, including symmetry/Hermitian symmetry and positive definiteness.
21. Given vectors in an inner product space, be able to compute norms, test orthogonality, and apply Cauchy–Schwarz when needed.
22. Given an orthogonal or orthonormal set, be able to expand vectors correctly and not confuse the two coefficient formulas.

Policy Application

Last-week revision advice. For the final stretch, revise in this order:

1. basis / dimension / direct sums;
2. kernel / range / rank–nullity / invertibility;
3. matrix representations and change of basis;
4. eigenvalues / eigenspaces / diagonalization;
5. inner products / Gram–Schmidt / orthogonal complements / projections.

That order follows the actual logical dependency structure of the course.

A Formula & Key Results Reference

Chapter 1: Vector Spaces and Subspaces

Definitions

- **Vector space over $\mathbb{F}$.** A nonempty set V with addition and scalar multiplication satisfying the usual axioms.
- **Subspace.** For $W \subseteq V$,
$$W \leq V \iff W \text{ is a vector space under the restricted operations from } V.$$
- **Linear combination / span.** For $S \subseteq V$,
$$\operatorname{span}(S) = \{\text{all finite linear combinations of vectors from } S\}.$$
Convention: $\operatorname{span}(\varnothing) = \{0\}$.
- **Sum and intersection of subspaces.** For $V_1, \dots, V_n \leq V$,
$$V_1 + \cdots + V_n = \{v_1 + \cdots + v_n : \ v_i \in V_i\}, \qquad \bigcap_{i=1}^{n} V_i = \{v \in V : \ v \in V_i \ \forall i\}.$$
- **Direct sum.**
$$W = V_1 \oplus \cdots \oplus V_n$$
means $W = V_1 + \cdots + V_n$ and each $w \in W$ has a unique decomposition $w = v_1 + \cdots + v_n$ with $v_i \in V_i$.

Theorems

- **Subspace Test.**
$$W \leq V \iff \Big(W \neq \varnothing \ \wedge \ u, v \in W \Rightarrow u + v \in W \ \wedge \ c \in \mathbb{F}, \ u \in W \Rightarrow cu \in W\Big).$$
- **One-line subspace test.**
$$W \leq V \iff \Big(W \neq \varnothing \ \wedge \ \forall u, v \in W, \forall c \in \mathbb{F} : \ cu + v \in W\Big).$$

- **Span is a subspace and smallest containing subspace.**

$$S \subseteq \operatorname{span}(S), \qquad \operatorname{span}(S) \leq V, \qquad S \subseteq W \leq V \Rightarrow \operatorname{span}(S) \subseteq W.$$

- **Basic span identities.**

$$A \subseteq B \Rightarrow \operatorname{span}(A) \subseteq \operatorname{span}(B), \qquad \operatorname{span}(\operatorname{span}(A)) = \operatorname{span}(A), \qquad \operatorname{span}(A \cup B) = \operatorname{span}(A) + \operatorname{span}(B).$$

- **Intersection and sum of subspaces are subspaces.** If $V_i \leq V$, then $\bigcap_i V_i \leq V$ and $V_1 + \cdots + V_n \leq V$.
- **Union criterion.** For $W_1, W_2 \leq V$,

$$W_1 \cup W_2 \leq V \iff (W_1 \subseteq W_2) \text{ or } (W_2 \subseteq W_1).$$

- **Direct-sum zero-sum criterion.**

$$V_1 + \cdots + V_n = V_1 \oplus \cdots \oplus V_n \iff \Big(\sum_{i=1}^{n} v_i = 0,\ v_i \in V_i \Rightarrow v_1 = \cdots = v_n = 0\Big).$$

- **Two-subspace criterion.**

$$W = U \oplus Z \iff \big(W = U + Z\big) \ \wedge \ \big(U \cap Z = \{0\}\big).$$

Equivalences / Identities

- **Derived vector-space identities.**

$$-(-u) = u, \qquad 0_{\mathbb{F}} u = 0_V, \qquad (-1)u = -u, \qquad c0_V = 0_V.$$

- **Immediate subspace consequences.** If $W \leq V$, then

$$0 \in W, \qquad u \in W \Rightarrow -u \in W, \qquad u, v \in W \Rightarrow u - v \in W.$$

- **Fast disproof.**

$$0 \notin W \Rightarrow W \not\leq V.$$

- **Kernel viewpoint.** If $T : V \to U$ is linear, then

$$\ker(T) = \{v \in V : T(v) = 0\} \leq V.$$

  Homogeneous linear conditions usually define subspaces.

- **Affine-subspace criterion.** For fixed $b \in V$,

$$\{\lambda_1 a_1 + \cdots + \lambda_n a_n + b : \lambda_i \in \mathbb{F}\} \leq V \iff b \in \operatorname{span}\{a_1, \ldots, a_n\}.$$

Algorithms / Templates

- **Subspace proof routes.** Use: (i) Subspace Test; (ii) identify as a kernel/solution space of homogeneous linear equations; (iii) show $W = \operatorname{span}(S)$.
- **Direct-sum workflow.** Show $V = U + W$ and $U \cap W = \{0\}$; or use the zero-sum criterion.
- **Sum-map reformulation.** For $U, W \leq V$, define $\phi : U \times W \to V$ by $\phi(u, w) = u + w$. Then

$$\operatorname{Im}(\phi) = U + W, \qquad \ker(\phi) \cong U \cap W.$$

  Hence $U \oplus W$ iff ϕ is injective.

Chapter 2: Finite-Dimensional Vector Spaces

Definitions

- **Linear independence.** $S \subseteq V$ is linearly independent (LI) iff every finite relation

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$$

  with distinct $v_i \in S$ forces $\lambda_1 = \cdots = \lambda_n = 0$.

- **Basis.** $B \subseteq V$ is a basis iff B is LI and $\operatorname{span}(B) = V$.

- **Ordered basis and coordinates.** If $\mathcal{B} = (b_1, \ldots, b_n)$ is an ordered basis and

$$v = \sum_{i=1}^{n} c_i b_i,$$

  then

$$[v]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{F}^{n,1}.$$

- **Finite-dimensional / dimension.** V is finite-dimensional iff it has a finite basis; then

$$\dim(V) = |B| \quad \text{for any basis } B.$$

Theorems

- **Dependence criterion.** A set S is linearly dependent iff one vector in S lies in the span of the others.

- **Unique expansion in a basis.** If B is a basis of V, every $v \in V$ has a unique expression as a finite linear combination of vectors in B.

- **Basis existence.** Every vector space has a basis.

- **Size-Limitation Lemma.** If V has a basis of size n, then every LI subset of V has size at most n.

- **Well-defined dimension.** In a finite-dimensional space, all bases have the same size.

- **Basis extension / extraction.** If $\dim(V) < \infty$:

  any LI subset extends to a basis,
  any spanning set contains a basis.

- **Subspaces of finite-dimensional spaces are finite-dimensional.** If $W \leq V$ and $\dim(V) < \infty$, then $\dim(W) \leq \dim(V)$.

- **Dimension formula.** If $U, W \leq V$ and $\dim(V) < \infty$, then

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

- **Direct-sum dimension formula.** If $V = U \oplus W$ and $\dim(V) < \infty$, then

$$\dim(V) = \dim(U) + \dim(W).$$

Equivalences / Identities

- **Quick LI facts.**

$$\varnothing \text{ is LI}, \qquad \{v\} \text{ is LI} \iff v \neq 0, \qquad S \text{ dependent} \Rightarrow \text{ every superset dependent}.$$

- **Matrix LI test in $\mathbb{F}^m$.** For $A = [v_1 \ \cdots \ v_n]$,

$$\{v_1, \ldots, v_n\} \text{ LI} \iff A\lambda = 0 \text{ has only the trivial solution} \iff \text{every column of } \operatorname{rref}(A) \text{ is a pivot column}.$$

- **Dimension bounds in an n-dimensional space.**

$$\text{LI set} \Rightarrow |S| \leq n, \qquad \text{spanning set} \Rightarrow |S| \geq n.$$

- **Exact-n criterion.** If $\dim(V) = n$ and $|S| = n$, then

$$S \text{ basis} \iff S \text{ spans } V \iff S \text{ is LI}.$$

- **Standard dimensions.**

$$\dim(\mathbb{F}^n) = n, \qquad \dim(\mathcal{P}_n(\mathbb{F})) = n + 1, \qquad \dim(\mathbb{F}^{m,n}) = mn.$$

- **Common matrix-subspace dimensions.**

$$\dim\{A \in \mathbb{R}^{n,n} : A^{\mathsf{T}} = A\} = \frac{n(n+1)}{2}, \qquad \dim\{A \in \mathbb{R}^{n,n} : A^{\mathsf{T}} = -A\} = \frac{n(n-1)}{2},$$

$$\dim\{A \in \mathbb{R}^{n,n} : A \text{ upper triangular}\} = \frac{n(n+1)}{2}, \qquad \dim\{A \in \mathbb{F}^{n,n} : \operatorname{tr}(A) = 0\} = n^2 - 1.$$

- **Direct sum via dimensions.** For $U, W \leq V$ with $\dim(V) < \infty$,

$$V = U \oplus W \iff \big(V = U + W\big) \ \wedge \ \big(\dim(U) + \dim(W) = \dim(V)\big).$$

Algorithms / Templates

- **Basis test.** Prove spanning + LI; or, if the number of vectors equals $\dim(V)$, prove only one of the two.
- **Coordinates in $\mathbb{R}^n$.** If $P = [b_1 \ \cdots \ b_n]$, then

$$[v]_{\mathcal{B}} = P^{-1}v.$$

Equivalently row-reduce $[\, b_1 \ \cdots \ b_n \mid v \,]$.

- **Extend LI set to a basis.** Append vectors from a known spanning set or standard basis and discard redundant additions.
- **Extract a basis from a spanning set.** Remove vectors that lie in the span of earlier/remaining vectors until LI remains.
- **Dimension problems.** Compute $\dim(U \cap W)$ via

$$\dim(U \cap W) = \dim(U) + \dim(W) - \dim(U + W).$$

Chapter 3: Linear Transformations

Definitions

- **Linear map.** $T : V \to W$ is linear iff

$$T(u+v) = T(u) + T(v), \qquad T(cu) = cT(u) \quad (\forall u, v \in V,\ c \in \mathbb{F}),$$

equivalently,

$$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v).$$

- **Space of linear maps.**

$$\mathcal{L}(V, W) = \{T : V \to W :\ T \text{ linear}\}, \qquad \mathcal{L}(V) = \mathcal{L}(V, V).$$

- **Kernel / range.**

$$\text{Null}(T) = \ker(T) = \{v \in V : T(v) = 0\}, \qquad \text{Range}(T) = \text{im}(T) = \{T(v) : v \in V\}.$$

- **Rank / nullity.**

$$\text{Rank}(T) = \dim(\text{Range}(T)), \qquad \text{Nullity}(T) = \dim(\text{Null}(T)).$$

- **Injective / surjective / bijective.** Standard meanings for a map $T : V \to W$.
- **Invertible linear map.** $T \in \mathcal{L}(V, W)$ is invertible iff there exists $S : W \to V$ such that

$$S \circ T = I_V, \qquad T \circ S = I_W.$$

Theorems

- $\mathcal{L}(V, W)$ **is a vector space.** Under pointwise addition and scalar multiplication.
- **Kernel and range are subspaces.**

$$\text{Null}(T) \leq V, \qquad \text{Range}(T) \leq W.$$

- **Preimage of a subspace is a subspace.** If $U \leq W$, then

$$T^{-1}(U) = \{v \in V : T(v) \in U\} \leq V.$$

- **Rank–Nullity.** If $\dim(V) < \infty$ and $T \in \mathcal{L}(V, W)$, then

$$\dim(V) = \text{Rank}(T) + \text{Nullity}(T).$$

- **Injectivity via kernel.**

$$T \text{ injective} \iff \text{Null}(T) = \{0\}.$$

- **Invertibility $\iff$ bijectivity.** A linear map is invertible iff it is bijective; if invertible, then T^{-1} is linear.

Equivalences / Identities

- **Linearity quick checks.**
$$T(0) = 0, \qquad T(-v) = -T(v).$$
- **Finite-dimensional criteria.**
$$T \text{ injective} \iff \text{Nullity}(T) = 0, \qquad T \text{ surjective} \iff \text{Rank}(T) = \dim(W).$$
- **Equal-dimension criterion.** If $\dim(V) = \dim(W)$, then
$$T \text{ injective} \iff T \text{ surjective} \iff T \text{ bijective} \iff T \text{ invertible}.$$
- **Composition facts.**
$$S, T \text{ invertible} \Rightarrow (T \circ S)^{-1} = S^{-1} \circ T^{-1}.$$
Also,
$$T \circ S \text{ injective} \Rightarrow S \text{ injective}, \qquad T \circ S \text{ surjective} \Rightarrow T \text{ surjective}.$$

Algorithms / Templates

- **Linearity test.** Check $T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$ directly.
- **Kernel / range workflow.** Compute $\ker(T)$ and $\text{im}(T)$; then apply Rank–Nullity and infer injectivity/surjectivity.
- **Polynomial / matrix operator workflow.** On $\mathcal{P}_n$ or matrix spaces, evaluate T on a chosen basis; kernel and range then become linear-system questions in coordinates.

Chapter 4: Matrix Representations and Change of Basis

Definitions

- **Ordered basis and coordinate map.** For $\beta = (b_1, \ldots, b_n)$,
$$[v]_\beta = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \iff v = \sum_{i=1}^{n} c_i b_i.$$
- **Matrix representation of a linear map.** For $T : V \to W$ with ordered bases $\beta = (b_1, \ldots, b_n)$ for V and $\gamma = (c_1, \ldots, c_m)$ for W,
$$[T]_\beta^\gamma = [[T(b_1)]_\gamma \quad \cdots \quad [T(b_n)]_\gamma] \in \mathbb{F}^{m \times n}.$$
- **Change-of-coordinate matrix.** For ordered bases β, γ of V,
$$[I_V]_\beta^\gamma$$
converts β-coordinates to γ-coordinates.
- **Similarity.** For $A, B \in \mathbb{F}^{n \times n}$,
$$A \sim B \iff \exists Q \text{ invertible such that } B = Q^{-1} A Q.$$

Theorems

- **Coordinate map is a linear isomorphism.**
$$[\cdot]_\beta : V \to \mathbb{F}^{n,1}$$
is a linear isomorphism.
- **Fundamental coordinate identity.**
$$[T(v)]_\gamma = [T]_\beta^\gamma [v]_\beta \qquad (\forall v \in V).$$
- **Composition corresponds to multiplication.** If $S : U \to V$ and $T : V \to W$ are linear, then
$$[T \circ S]_\alpha^\gamma = [T]_\beta^\gamma [S]_\alpha^\beta.$$
- **Coordinate conversion.**
$$[v]_\gamma = [I_V]_\beta^\gamma [v]_\beta, \qquad [I_V]_\gamma^\beta = \left([I_V]_\beta^\gamma\right)^{-1}.$$
- **Inverse map / inverse matrix.** If $T \in \mathcal{L}(V)$ is invertible, then
$$[T^{-1}]_\beta^\beta = \left([T]_\beta^\beta\right)^{-1}.$$
- **Change-of-basis formula for one operator.** For $T \in \mathcal{L}(V)$ and ordered bases β, γ of V,
$$[T]_\gamma^\gamma = [I_V]_\beta^\gamma [T]_\beta^\beta [I_V]_\gamma^\beta.$$
- **Similarity invariants.** If $A \sim B$, then
$$\det(A) = \det(B), \qquad \operatorname{tr}(A) = \operatorname{tr}(B).$$
- **Matrix representations in different bases are similar.**
$$[T]_\beta^\beta \sim [T]_\gamma^\gamma.$$

Equivalences / Identities

- **Size check for $[T]_\beta^\gamma$.**
$$[T]_\beta^\gamma \in \mathbb{F}^{(\dim W)\times(\dim V)}.$$
Rows = codomain dimension; columns = domain dimension.
- **Read-off-columns rule.** The jth column of $[T]_\beta^\gamma$ is $[T(b_j)]_\gamma$.
- **Change of basis via a standard basis σ.**
$$[I_V]_\beta^\gamma = \left([I_V]_\gamma^\sigma\right)^{-1} [I_V]_\beta^\sigma.$$
If $V = \mathbb{F}^n$, then
$$[I_V]_\beta^\sigma = \begin{bmatrix} [b_1]_\sigma & \cdots & [b_n]_\sigma \end{bmatrix}.$$
- **Triangular-shape criterion.** If
$$T(v_k) \in \operatorname{span}(w_1, \ldots, w_k) \quad (\forall k),$$
then $[T]_\beta^\gamma$ is upper triangular for $\beta = (v_1, \ldots, v_n)$ and $\gamma = (w_1, \ldots, w_n)$.
- 2×2 **characteristic identity.** For $A \in \mathbb{F}^{2\times 2}$,
$$\det(A - \lambda I_2) = \lambda^2 - \operatorname{tr}(A)\lambda + \det(A).$$

Algorithms / Templates

- **Build $[T]_\beta^\gamma$.** Compute $T(b_j)$ for each domain basis vector, express each in the codomain basis, use these coordinate columns.
- **Build $[I_V]_\beta^\gamma$.** Express each b_j in the basis γ; those coordinate vectors are the columns.
- **Fast row-reduction trick for basis change.** If $P_\beta = [I_V]_\beta^\sigma$ and $P_\gamma = [I_V]_\gamma^\sigma$, then

$$[I_V]_\beta^\gamma = P_\gamma^{-1} P_\beta.$$

- **Exam workflow for iterates.** If $[T]_\beta^\beta$ is diagonal or triangular in a well-chosen basis, compute powers there first, then convert back.

Chapter 5: Eigenvalues, Diagonalization, and Structure Theory

Definitions

- **Eigenvalue / eigenvector.** For $T \in \mathcal{L}(V)$,

$$T(v) = \lambda v \quad \text{with } v \neq 0.$$

- **Eigenspace.**

$$E_{T,\lambda} = \ker(T - \lambda I_V).$$

- **Characteristic polynomial.** If $\dim(V) = n$,

$$p_T(\lambda) = \det([T]_\beta^\beta - \lambda I_n),$$

  independent of basis β.
- **Diagonalizable.** $T \in \mathcal{L}(V)$ is diagonalizable iff V has a basis consisting of eigenvectors of T.
- **Algebraic / geometric multiplicity.** For an eigenvalue λ:

$$\text{algmult}(\lambda) = \text{multiplicity of } \lambda \text{ as a root of } p_T, \qquad \text{geomult}(\lambda) = \dim(E_{T,\lambda}).$$

Theorems

- **Eigenvalue equivalences.**

$$\lambda \text{ eigenvalue} \iff E_{T,\lambda} \neq \{0\} \iff T - \lambda I_V \text{ not injective.}$$

  In finite dimensions,

$$\lambda \text{ eigenvalue} \iff T - \lambda I_V \text{ not invertible} \iff \det([T]_\beta^\beta - \lambda I) = 0.$$

- **Zero eigenvalue criterion.**

$$0 \text{ eigenvalue of } T \iff T \text{ not injective.}$$

- **Characteristic polynomial is basis-independent.**

- **Eigenvalues are roots of p_T.** In particular, a linear operator on an n-dimensional space has at most n eigenvalues in $\mathbb{F}$.
- **Distinct eigenvalues give LI eigenvectors.**
- **Distinct eigenspaces sum directly.**

$$E_{T,\lambda_1} + \cdots + E_{T,\lambda_k} = E_{T,\lambda_1} \oplus \cdots \oplus E_{T,\lambda_k} \qquad (\lambda_i \text{ distinct}).$$

- **Dimension bound for eigenspaces.**

$$\sum_{\lambda \text{ eigenvalue}} \dim(E_{T,\lambda}) \leq \dim(V).$$

- **Fundamental Theorem of Diagonalizability.** For finite-dimensional V,

$$T \text{ diagonalizable} \iff V = \bigoplus_{\lambda} E_{T,\lambda} \iff \dim(V) = \sum_{\lambda} \dim(E_{T,\lambda}).$$

- **Distinct-eigenvalue criterion.** If $\dim(V) = n$ and T has n distinct eigenvalues in $\mathbb{F}$, then T is diagonalizable.
- **Multiplicity criterion.** T is diagonalizable iff:
  1. p_T splits completely over $\mathbb{F}$;
  2. for every eigenvalue λ,
  $$\text{geomult}(\lambda) = \text{algmult}(\lambda).$$

Equivalences / Identities

- **Diagonalization as $A = PDP^{-1}$.** If $\beta = (v_1, \ldots, v_n)$ is an eigenbasis with eigenvalues $\lambda_1, \ldots, \lambda_n$, then

$$P = [I_V]_\beta^\sigma = [\, v_1 \ \cdots \ v_n \,], \qquad D = [T]_\beta^\beta = \text{diag}(\lambda_1, \ldots, \lambda_n),$$

  so
  $$[T]_\sigma^\sigma = PDP^{-1}.$$
  Columns of P are eigenvectors, in the same order as diagonal entries of D.
- **Powers / polynomials via diagonalization.** If $A = PDP^{-1}$, then

$$A^k = PD^kP^{-1}, \qquad D^k = \text{diag}(\lambda_1^k, \ldots, \lambda_n^k).$$

  More generally, for $p(t) = \sum a_j t^j$,
  $$p(A) = P\,p(D)\,P^{-1}, \qquad p(D) = \text{diag}(p(\lambda_1), \ldots, p(\lambda_n)).$$
- **Field dependence.** Eigenvalues are roots in the base field only.
- **Multiplicity inequalities.**
$$1 \leq \text{geomult}(\lambda) \leq \text{algmult}(\lambda).$$
- **Nilpotent diagnostic.** If $T^m = 0$ for some $m \geq 1$, then the only possible eigenvalue is 0. Hence a nonzero nilpotent operator is not diagonalizable.
- **Solvability of $(T - \lambda I)v = b$.** In finite dimensions,

$$\lambda \text{ not an eigenvalue} \iff T - \lambda I \text{ invertible} \iff (T - \lambda I)v = b \text{ has a unique solution for every } b.$$

QRS@NTU · qrsntu.com

Algorithms / Templates

- **Eigenvalue workflow.** Compute $p_T(\lambda)$, solve $p_T(\lambda) = 0$ in the correct field, then compute each eigenspace $\ker(T - \lambda I)$.
- **Diagonalizability checklist.**
  1. Find eigenvalues.
  2. Compute $\dim(E_{T,\lambda})$ for each eigenvalue.
  3. Check whether the eigenspace dimensions sum to $\dim(V)$.
  4. If yes, concatenate eigenspace bases to form an eigenbasis.
- **Iterates / recurrence workflow.** After diagonalizing $A = PDP^{-1}$, move the initial vector into the eigenbasis, apply D^n, then convert back.
- **Fast exam checks.**
  - Triangular matrix $\Rightarrow$ eigenvalues are diagonal entries.
  - Repeated eigenvalue $\not\Rightarrow$ diagonalizable.
  - Distinct eigenvalues $\Rightarrow$ diagonalizable.

Chapter 6: Inner Product Spaces

Definitions

- **Inner product.** A map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ satisfying conjugate symmetry, linearity in the chosen slot, and positive-definiteness.
- **Norm / distance.**
$$\|v\| = \sqrt{\langle v, v \rangle}, \qquad d(u, v) = \|u - v\|.$$
- **Orthogonality.**
$$u \perp v \iff \langle u, v \rangle = 0.$$
Orthogonal set: pairwise orthogonal. Orthonormal set: orthogonal and each vector has norm 1.
- **Orthonormal basis (ONB).** A basis that is orthonormal.
- **Projection onto a nonzero vector.**
$$\mathrm{proj}_u(v) = \frac{\langle u, v \rangle}{\langle u, u \rangle} u \qquad (u \neq 0).$$
- **Orthogonal complement.** For nonempty $A \subseteq V$,
$$A^{\perp} = \{v \in V : \ \langle u, v \rangle = 0 \ \forall u \in A\}.$$
- **Orthogonal projection onto a subspace.** If $V = W \oplus W^{\perp}$, then for $v = p + q$ with $p \in W$, $q \in W^{\perp}$,
$$P_W(v) = p.$$

Theorems

- **Basic norm properties.**

$$\|v\| \geq 0, \qquad \|v\| = 0 \iff v = 0, \qquad \|cv\| = |c|\,\|v\|.$$

- **Cauchy–Schwarz.**

$$|\langle u, v\rangle| \leq \|u\|\,\|v\|.$$

  Equality iff u, v are linearly dependent.

- **Triangle inequality.**

$$\|u + v\| \leq \|u\| + \|v\|.$$

- **Orthogonal nonzero sets are LI.**

- **Fourier coefficient formula in an ONB.** If β is an ONB, then

$$v = \sum_{u \in \beta} \langle u, v\rangle u.$$

  For a finite ONB $\beta = (u_1, \ldots, u_n)$,

$$v = \sum_{i=1}^{n} \langle u_i, v\rangle u_i.$$

- **Projection error is orthogonal.**

$$u \perp \big(v - \mathrm{proj}_u(v)\big) \qquad (u \neq 0).$$

- **Gram–Schmidt.** For a basis $(b_1, \ldots, b_n)$,

$$v_1 = b_1, \qquad v_k = b_k - \sum_{j=1}^{k-1} \mathrm{proj}_{v_j}(b_k) = b_k - \sum_{j=1}^{k-1} \frac{\langle v_j, b_k\rangle}{\langle v_j, v_j\rangle} v_j \quad (2 \leq k \leq n).$$

  Then $(v_1, \ldots, v_n)$ is orthogonal and

$$\mathrm{span}(v_1, \ldots, v_k) = \mathrm{span}(b_1, \ldots, b_k) \quad (\forall k).$$

  Normalization gives an ONB.

- **Orthogonal complements are subspaces.**

- **Basic orthogonal-complement properties.**

$$\{0\}^\perp = V, \qquad V^\perp = \{0\}, \qquad A \subseteq B \Rightarrow B^\perp \subseteq A^\perp, \qquad A^\perp = \mathrm{span}(A)^\perp.$$

- **Orthogonal direct sum decomposition.** If $\dim(V) < \infty$ and $W \leq V$, then

$$V = W \oplus W^\perp, \qquad \dim(V) = \dim(W) + \dim(W^\perp).$$

- **Double orthogonal complement.** If $\dim(V) < \infty$ and $W \leq V$, then

$$W^{\perp\perp} = W.$$

- **Projection formula for a subspace.** If $(b_1, \ldots, b_m)$ is an ONB of W, then

$$P_W(v) = \sum_{k=1}^{m} \langle b_k, v\rangle b_k, \qquad v - P_W(v) \in W^\perp.$$

Equivalences / Identities

- **Orthogonal vs orthonormal.** Orthogonal means pairwise perpendicular; orthonormal additionally requires norm 1.
- **Fourier formula for an orthogonal basis $\{u_i\}$.**

$$v = \sum_i \frac{\langle u_i, v\rangle}{\langle u_i, u_i\rangle} u_i.$$

  The simpler coefficient formula requires an ONB.

- **Computing $W^\perp$ from a basis.** If B is a basis of W, then

$$W^\perp = \{v \in V : \ \langle b, v\rangle = 0 \ \forall b \in B\}.$$

- **Frobenius inner product.** On $\mathbb{R}^{n,n}$,

$$\langle A, B\rangle_F = \operatorname{tr}(A^\mathsf{T} B) = \sum_{i,j} a_{ij} b_{ij}.$$

  Useful orthogonal-complement facts:

$$D = \{\text{diagonal matrices}\} \Rightarrow D^\perp = \{A : a_{ii} = 0 \ \forall i\},$$
$$U = \{\text{upper triangular matrices}\} \Rightarrow U^\perp = \{\text{strictly lower triangular matrices}\}.$$

- **Least squares / normal equations.** For $y \in \mathbb{R}^{n,1}$ and $A \in \mathbb{R}^{n\times m}$,

$$\min_b \|y - Ab\|^2 \iff y - Ab \in \operatorname{Range}(A)^\perp \iff A^\mathsf{T} A\, b = A^\mathsf{T} y.$$

Algorithms / Templates

- **Gram–Schmidt workflow.** Keep the orthogonal vectors unnormalized until the end; normalize only once.
- **Find $A^\perp$ for a finite set A.**
  1. Replace A by a basis of $\operatorname{span}(A)$.
  2. Solve the equations $\langle b, v\rangle = 0$ for all basis vectors b.
- **Projection onto a subspace.**
  1. Find an ONB of W.
  2. Use $P_W(v) = \sum \langle b_k, v\rangle b_k$.
  3. The error vector is $v - P_W(v) \in W^\perp$.
- **Hyperplane shortcut in Euclidean spaces.** If

$$W = \{x \in \mathbb{R}^n : Mx = 0\},$$

  then

$$W^\perp = \operatorname{span}\{\text{rows of } M\}.$$

- **Frobenius-complement workflow.** Test against matrix units E_{ij} to force specific entries to vanish.

B Notation Reference

This appendix provides a comprehensive notation table for the course. Use this section to clarify symbols and conventions.

Symbol	Meaning	Context
Fields and Vector Spaces		
$\mathbb{F}$	Field (usually $\mathbb{R}$ or $\mathbb{C}$)	General field notation
V, W, U	Vector spaces over $\mathbb{F}$	Abstract spaces
$\mathbb{R}^n, \mathbb{C}^n$	n-dimensional real/complex vector space	Standard Euclidean space
$P_n(\mathbb{F})$	Vector space of polynomials of degree $\leq n$	Polynomial spaces
$M_{m\times n}(\mathbb{F})$	Vector space of $m \times n$ matrices	Matrix spaces
0 or $\mathbf{0}$	Zero vector	General vector spaces
Subspaces and Bases		
$W \leq V$	W is a subspace of V	Subspace relation
$\text{Span}(S)$	Span of set S	Linear combinations of vectors in S
$\dim(V)$	Dimension of vector space V	Number of vectors in a basis
$\mathcal{B}, \beta, \gamma$	Ordered bases for vector spaces	Used for matrix representations
$[v]_{\mathcal{B}}$	Coordinate vector of v relative to basis $\mathcal{B}$	Column vector in $\mathbb{F}^{n,1}$
$V = U \oplus W$	Internal direct sum	$V = U + W$ and $U \cap W = \{0\}$

QRS@NTU

Symbol	Meaning	Context
Linear Maps and Matrices		
$T : V \to W$	Linear transformation from V to W	Function notation
$\mathcal{L}(V, W)$	$\mathcal{L}(V, W) \coloneqq \{T : V \to W \mid T \text{ is linear}\}$	Space of all linear maps
$\text{Null}(T)$ or $\ker(T)$	Null space (kernel) of T	$\{v \in V \mid T(v) = 0\}$
$\text{range}(T)$ or $\text{im}(T)$	Range (image) of T	$\{T(v) \mid v \in V\}$
$\text{Nullity}(T)$	Nullity of $T \coloneqq \dim(\text{Null}(T))$	Rank-Nullity Theorem
$\text{Rank}(T)$	Rank of $T \coloneqq \dim(\text{range}(T))$	Rank-Nullity Theorem
I or I_V	Identity transformation or matrix	$I(v) = v$ for all v
$S \circ T$	Composition of S and T	$(S \circ T)(v) = S(T(v))$
T^{-1}	Inverse transformation	Exists if T is bijective
$[T]_\beta^\gamma$	Matrix of T relative to bases β and γ	Matrix representation
$[I_V]_\beta^\gamma$	Change-of-coordinate matrix from β to γ	Transforms $[v]_\beta$ to $[v]_\gamma$
$A \sim B$	A is similar to B	$\exists Q$ invertible: $B = Q^{-1}AQ$
A^T	Transpose of matrix A	$(A^\mathsf{T})_{ij} = A_{ji}$
A^* or $A^\dagger$	Conjugate transpose (Hermitian adjoint)	$(A^*)_{ij} = \overline{A_{ji}}$
$\text{tr}(A)$	Trace of matrix A	Sum of diagonal entries
$\det(A)$	Determinant of matrix A	Scalar value
Eigenvalues and Inner Products		
λ, v	Eigenvalue and Eigenvector	$T(v) = \lambda v, v \neq 0$
$E_{T,\lambda}$	Eigenspace for eigenvalue λ	$\text{Null}(T - \lambda I_V)$
$p_T(\lambda)$	Characteristic polynomial of T	$\det([T]_\beta^\beta - \lambda I_n)$
$m_A(\lambda)$	Minimal polynomial of A	Least degree monic polynomial with $m_A(A) = 0$
$J_k(\lambda)$	Jordan block of size k with eigenvalue λ	Jordan Canonical Form
$\langle u, v \rangle$	Inner product of u and v	Inner-product spaces
$\|v\|$	Norm of vector v	$\sqrt{\langle v, v \rangle}$
$u \perp v$	u is orthogonal to v	$\langle u, v \rangle = 0$
$W^\perp$	Orthogonal complement of W	$\{v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in W\}$
$\text{proj}_W(v)$	Orthogonal projection of v onto W	Projection operator
$\{e_1, \ldots, e_n\}$	Orthonormal basis	$\langle e_i, e_j \rangle = \delta_{ij}$
δ_{ij}	Kronecker delta	1 if $i = j$; 0 otherwise
Logic and Set Theory		
$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$	Standard number sets	Mathematical domains
$\iff$	If and only if	Logical equivalence
$\implies$	Implies	Logical implication
$\forall, \exists$	For all, There exists	Quantifiers
$\in, \subseteq$	Element of, Subset of	Set relations
$\cap, \cup, \emptyset$	Intersection, Union, Empty set	Set operations

Note: Throughout this document, we use standard mathematical notation. When in doubt, refer to this table or consult your textbook.

C Examination and Administrative Information

This appendix provides key information about assessments, examination policies, and revision strategies for MH1201.

C.1 Assessment Breakdown

Component	ILOs Assessed	Weight	Type
In-class quizzes	1–9	15%	Team, point-based
Midterm Exam	1–5	25%	Individual, point-based
Final Exam	1–9	60%	Individual, point-based

Notes:

- In-class quizzes are conducted nearly every tutorial in teams of three or four students.
- Your lowest two in-class quiz scores are dropped at the end of the semester.
- Online homework on WebAssign is for practice only and does not count toward your grade.

C.2 Important Dates

Event	Study Week	Date	Time / Venue
Midterm Exam	8	March 11, 2026 (Wed)	1330–1520, Hall C
Final Exam	—	April 27, 2026 (Mon)	1700–1900, TBA

C.3 Help-Sheet Policy for Exams

- Every student is allowed to consult **ONE double-sided A4-sized help-sheet** with **typed notes**.
- No attachments to the help-sheet either as additional sheets or as sticky notes are allowed.
- Use the **Formula & Key Results Reference** section (Appendix A) and the **Notation Reference** (Appendix B) of these notes as a starting point for your help sheet.

C.4 Study Resources

Recommended textbooks:

- *Linear Algebra: A Modern Introduction* (4th edition) by David Poole, Cengage, 2015. (Available on WebAssign)
- *Linear Algebra Done Right* (4th edition) by Sheldon Axler, Springer, 2024. (Free e-book: https://linear.axler.net/LADR4e.pdf)
- *Introduction to Linear Algebra* (6th edition) by Gilbert Strang, Wellesley-Cambridge Press, 2023.

Online resources:

- NTULearn course page: tutorial problem sets, online homework, announcements.
- WebAssign: online homework for practice (already paid for by NTU).

- Office hours: Dr. Leonard Huang (Tuesdays, 1400–1500, SPMS-05-08).
- Tutorial sessions: consult your tutor for clarification.

QRS@NTU resources:

- Visit https://qrsntu.org for additional resources, past-year papers, and revision materials.

QRS@NTU — qrsntu.org