Contents

Vectors and Vector Spaces

Table of Contents

Interactive Visualizations

HTML Interactive Demos

Chapter 1 Interactive Hub - Navigate all visualizations for this chapter

Individual Lesson Visuals:

- **Vector Basics** - Magnitude & direction fundamentals
- **Vector Operations** - Addition, scaling, dot products with cosine relationships
- **Linear Independence** - Mathematical analysis of independence
- **Abstract Spaces** - Beyond arrows: polynomials, functions, matrices

Complete Experiences:

- **Interactive Demo Suite** - Complete Python visualization suite in HTML format

---

Introduction to Vectors

A **vector** is a mathematical object that has both magnitude and direction. In $\mathbb{R}^n$, a vector can be represented as an ordered list of n real numbers:

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

Geometric Interpretation

- In $\mathbb{R}^2$: vectors represent points in a plane or arrows from origin
- In $\mathbb{R}^3$: vectors represent points in 3D space
- In $\mathbb{R}^n$: vectors represent points in n-dimensional space

Examples

$$\mathbf{u} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 2 \\ 0 \\ -1 \\ 3 \end{pmatrix}$$

Interactive: Explore vector fundamentals with **Vector Basics** - see how magnitude and direction define everything about a vector.

Vector Operations

Vector Addition

For vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$:

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix}$$

Properties:

- **Commutative:** $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
- **Associative:** $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
- **Identity:** $\mathbf{v} + \mathbf{0} = \mathbf{v}$

Scalar Multiplication

For scalar $c \in \mathbb{R}$ and vector $\mathbf{v} \in \mathbb{R}^n$:

$$c\mathbf{v} = \begin{pmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{pmatrix}$$

Properties:

- **Distributive:** $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
- **Distributive:** $(c + d)\mathbf{v} = c\mathbf{v} + d\mathbf{v}$
- **Associative:** $(cd)\mathbf{v} = c(d\mathbf{v})$
- **Identity:** $1\mathbf{v} = \mathbf{v}$

Dot Product (Inner Product)

For vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$:

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Properties:

- **Commutative:** $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- **Distributive:** $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
- **Homogeneous:** $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v})$
- **Positive definite:** $\mathbf{v} \cdot \mathbf{v} \geq 0$, with equality iff $\mathbf{v} = \mathbf{0}$

Vector Magnitude (Norm)

The **magnitude** or **norm** of vector $\mathbf{v}$:

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

Unit Vectors

A **unit vector** has magnitude 1. Any non-zero vector $\mathbf{v}$ can be normalized:

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$$

Interactive: Practice vector operations with **Vector Operations** - see addition, scaling, dot products, and the cosine relationship in action.

Vector Spaces

A **vector space** V over field $\mathbb{F}$ (usually $\mathbb{R}$ or $\mathbb{C}$) is a set with two operations:

> **Field Definition:** A **field** $\mathbb{F}$ is a set equipped with two operations (addition and multiplication) that satisfy the usual arithmetic properties: commutativity, associativity, distributivity, existence of additive and multiplicative identities (0 and 1), and existence of additive inverses for all elements and multiplicative inverses for all non-zero elements. Common examples include the real numbers $\mathbb{R}$, complex numbers $\mathbb{C}$, and rational numbers $\mathbb{Q}$. - Vector addition: $+ : V \times V \to V$ - Scalar multiplication: $\cdot : \mathbb{F} \times V \to V$

Vector Space Axioms

For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and scalars $a, b \in \mathbb{F}$:

1. **Closure under addition**: $\mathbf{u} + \mathbf{v} \in V$
2. **Commutativity**: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. **Associativity**: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
4. **Zero vector**: $\exists \mathbf{0} \in V$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$
5. **Additive inverse**: $\forall \mathbf{v} \in V, \exists (-\mathbf{v}) \in V$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$
6. **Closure under scalar multiplication**: $a\mathbf{v} \in V$
7. **Scalar multiplication identity**: $1\mathbf{v} = \mathbf{v}$
8. **Scalar multiplication associativity**: $(ab)\mathbf{v} = a(b\mathbf{v})$
9. **Distributivity**: $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$
10. **Distributivity**: $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$

Examples of Vector Spaces

- $\mathbb{R}^n$: n-dimensional real coordinate space
- $\mathbb{C}^n$: n-dimensional complex coordinate space
- P_n: polynomials of degree at most n
- $C[a, b]$: continuous functions on interval $[a, b]$
- $M_{m\times n}$: $m \times n$ matrices

Concrete examples you can picture

- $\mathbb{R}^2$: all 2D arrows from the origin. Examples: $\begin{pmatrix}1\\2\end{pmatrix}$, $\begin{pmatrix}-3\\0.5\end{pmatrix}$. Under addition and scalar multiplication you stay in $\mathbb{R}^2$.
- $\mathbb{R}^3$: all 3D arrows. Examples: $\begin{pmatrix}0\\2\\-1\end{pmatrix}$, $\begin{pmatrix}4\\4\\4\end{pmatrix}$.
- $\mathbb{C}^2$: complex 2-vectors. Example: $\begin{pmatrix}1+i\\-2i\end{pmatrix}$.
- P_2: all polynomials of degree ≤ 2. General element: $a_0 + a_1x + a_2x^2$. Concrete examples: $1 - 3x + 2x^2$, $\frac{1}{2} + x$.
- $C[0, 1]$: continuous real-valued functions on $[0, 1]$. Examples: $f(x) = \sin x$, $g(x) = x^2$, $h(x) = e^x$ (restricted to $[0, 1]$).
- $M_{2\times 2}$: all 2×2 real matrices. Example: $\begin{pmatrix}1 & 2\\0 & -1\end{pmatrix}$.
- **Subspaces of** $\mathbb{R}^n$: any line or plane through the origin. Example in $\mathbb{R}^2$: all multiples of $\begin{pmatrix}2\\1\end{pmatrix}$; in $\mathbb{R}^3$: the xy-plane $\{(x, y, 0)\}$.

  This is essentially saying that a vector space is any collection of "vectors" (not necessarily geometric arrows) where adding two of them and scaling by numbers keeps you inside the same collection, and the usual algebraic rules hold.

Interactive: Discover abstract vector spaces with **Abstract Vector Spaces** - see how polynomials, functions, and matrices can all be "vectors".

Linear Independence

A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is **linearly independent** if the only solution to:

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

is $c_1 = c_2 = \cdots = c_k = 0$.

Linear Dependence

Vectors are **linearly dependent** if there exist scalars $c_1, c_2, \dots, c_k$ (not all zero) such that:

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

Key Properties

- **Zero vector dependency:** Any set containing the zero vector is linearly dependent
- **Two-vector dependency:** A set of two vectors is linearly dependent iff one is a scalar multiple of the other
- **Dimension limit:** In $\mathbb{R}^n$, any set of more than n vectors is linearly dependent

Visual intuition and concrete examples (in $\mathbb{R}^2$)

- **Linearly independent example.** Take $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. They point in two different directions and span the whole plane. The only way to make $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}$ is $c_1 = c_2 = 0$.
- **Linearly dependent example.** Take $\mathbf{w}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\mathbf{w}_2 = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = 2\mathbf{w}_1$. They lie on the same line through the origin (no new direction). Here $2\mathbf{w}_1 - \mathbf{w}_2 = \mathbf{0}$ with nonzero coefficients, so they are dependent.

  Intuitively: linear independence means "each vector brings a new direction you couldn't already get by mixing the others". Dependence means "at least one vector is redundant because it can be built from the rest".

Quick mental picture

- Two independent vectors in $\mathbb{R}^2$ form a grid (a parallelogram basis). Two dependent vectors line up on the same straight line.

Interactive: Explore independence concepts with **Linear Independence** - see when vectors add genuine new possibilities versus when they're just echoes.

Basis and Dimension

Span

The **span** of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is:

$$\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} = \{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k : c_i \in \mathbb{F}\}$$

Intuition: The span of a set of vectors = all possible linear combinations of them. Think of it as the space you can reach by scaling and adding those vectors.

Geometric illustrations (use arrows for vectors and label the spaces): - In $\mathbb{R}^2$: - span of one nonzero vector $\vec{v} \to$ a line through the origin along $\vec{v}$ - span of two independent vectors $\vec{u}, \vec{v} \to$ the entire plane $\mathbb{R}^2$ - In $\mathbb{R}^3$: - span of one nonzero vector $\vec{v} \to$ a line through the origin - span of two independent vectors $\vec{u}, \vec{v} \to$ a plane through the origin - span of three independent vectors $\vec{u}, \vec{v}, \vec{w} \to$ the whole space $\mathbb{R}^3$

Real-life analogy: Span is like the territory covered when you combine directions. The pair {east, north} lets you reach any place on a flat floor; one single direction only gives a line of reachable points.

Basis

A **basis** for vector space V is a set of vectors that: 1. Spans V 2. Is linearly independent

Intuition: A basis is the simplest set of building blocks of the space. It is the smallest independent set of vectors that covers the whole space. If any vector in the set could be built from the others, it would be redundant — and the set would not be a basis.

Consistency note for diagrams: we draw vectors as arrows (e.g., $\vec{u}, \vec{v}$) and label the resulting space (line, plane, or whole space) accordingly.

> Key rule (highlight): If $\dim(V) = n$, then any linearly independent set of n vectors is a basis, and any spanning set of n vectors is also a basis.

Standard Basis for $\mathbb{R}^n$

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Dimension

The **dimension** of vector space V is the number of vectors in any basis for V.

$$\dim(V) = \text{number of vectors in basis}$$

Important Theorems

1. **Basis Theorem**: All bases of a finite-dimensional vector space have the same number of vectors
2. **Dimension Theorem**: If $\dim(V) = n$, then:
   - Any linearly independent set of n vectors is a basis

- Any spanning set of n vectors is a basis
- Any linearly independent set has at most n vectors
- Any spanning set has at least n vectors

Practice Problems

Problem 1

Given vectors $\mathbf{u} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix}$:

a) Calculate $\mathbf{u} + \mathbf{v}$
b) Calculate $3\mathbf{u} - 2\mathbf{v}$
c) Calculate $\mathbf{u} \cdot \mathbf{v}$
d) Find $\|\mathbf{u}\|$ and $\|\mathbf{v}\|$

Problem 2

Determine if the vectors $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ are linearly independent.

Problem 3

Find a basis for the subspace spanned by:

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \mathbf{w}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{w}_3 = \begin{pmatrix} 3 \\ 3 \\ 1 \\ 1 \end{pmatrix}$$

Next: Matrix Operations and Properties

Solutions

Solution to Problem 1

Given $\mathbf{u} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix}$:

- $\mathbf{u} + \mathbf{v} = \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix}$

- $3\mathbf{u} - 2\mathbf{v} = 3\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} - 2\begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix} = \begin{pmatrix} 4 \\ -11 \\ 13 \end{pmatrix}$
- $\mathbf{u} \cdot \mathbf{v} = 2 \cdot 1 + (-1) \cdot 4 + 3 \cdot (-2) = -8$
- $\|\mathbf{u}\| = \sqrt{2^2 + (-1)^2 + 3^2} = \sqrt{14}$, $\|\mathbf{v}\| = \sqrt{1^2 + 4^2 + (-2)^2} = \sqrt{21}$

Solution to Problem 2

Let $A = [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3]$. With $\mathbf{v}_1 = (1, 2, 1)^T$, $\mathbf{v}_2 = (2, 1, 3)^T$, $\mathbf{v}_3 = (1, -1, 2)^T$.

Compute $\det A = 0$, so the set is linearly dependent. In fact, $\mathbf{v}_3 = \mathbf{v}_2 - \mathbf{v}_1$, i.e., $\mathbf{v}_1 - \mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0}$ with nonzero coefficients.

Solution to Problem 3

Given $\mathbf{w}_1 = (1, 2, 1, 0)^T$, $\mathbf{w}_2 = (2, 1, 0, 1)^T$, $\mathbf{w}_3 = (3, 3, 1, 1)^T$.

Observe $\mathbf{w}_3 = \mathbf{w}_1 + \mathbf{w}_2$. A basis for $\operatorname{span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is therefore $\{\mathbf{w}_1, \mathbf{w}_2\}$ (dimension 2).

Matrices and Matrix Operations

Table of Contents

Introduction to Matrices

A **matrix** is a rectangular array of numbers arranged in rows and columns. An $m \times n$ matrix has m rows and n columns:

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

Notation

- A_{ij} or a_{ij}: element in row i, column j
- $A \in \mathbb{R}^{m \times n}$: A is an $m \times n$ real matrix
- $A \in \mathbb{C}^{m \times n}$: A is an $m \times n$ complex matrix

Examples

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & -1 \\ 0 & 3 \\ 1 & 4 \end{pmatrix}, \quad C = (5)$$

What does each entry in a matrix mean?

The meaning depends on context. A matrix is a rectangular array of numbers, but its entries can represent very different things:

1. **Data Table (rows = objects, columns = features).** Example:

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

   Row 1 could be student #1's scores on 3 exams; row 2 student #2's scores.

2. **Linear Transformation (mapping vectors from one space to another).** Example:

$$B = \begin{pmatrix} 2 & -1 \\ 0 & 3 \\ 1 & 4 \end{pmatrix}$$

   - The first column $(2, 0, 1)^T$ is where $(1, 0)^T$ (the x-axis) gets mapped.
   - The second column $(-1, 3, 4)^T$ is where $(0, 1)^T$ (the y-axis) gets mapped. Multiplying Bx shows how a vector in the plane is transformed into $\mathbb{R}^3$.

3. **Graph (Adjacency Matrix).** For a graph with 3 nodes:

$$G = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

   - Entry $g_{12} = 1$ means an edge between node 1 and 2.
   - Zeros on the diagonal mean no self-loops. Thus, matrices can encode relationships and networks.

Consistency note: Keep arrows for vectors and label rows/columns when helpful.

Interactive: Explore what matrices represent with **Matrix Introduction** - see how the same mathematical object can be data, transformations, or networks.

Matrix Operations

Matrix Addition

For matrices $A, B \in \mathbb{R}^{m \times n}$:

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

Example:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix}$$

Scalar Multiplication

For scalar $c \in \mathbb{R}$ and matrix $A \in \mathbb{R}^{m \times n}$:

$$(cA)_{ij} = c \cdot A_{ij}$$

Example:

$$3 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 9 & 12 \end{pmatrix}$$

Matrix Multiplication

For $A \in \mathbb{R}^{m \times p}$ and $B \in \mathbb{R}^{p \times n}$, the product $AB \in \mathbb{R}^{m \times n}$:

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

Key Points:

- **Dimension compatibility:** Number of columns in A must equal number of rows in B
- **Not commutative:** Matrix multiplication is **not commutative**: $AB \neq BA$ (in general)
- **Associative:** Matrix multiplication is **associative**: $(AB)C = A(BC)$

Example:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$$

Interactive: Master matrix operations with **Matrix Operations** - see how addition combines effects, multiplication composes transformations, and transpose changes perspective.

Matrix Transpose

The **transpose** of matrix $A \in \mathbb{R}^{m \times n}$ is $A^T \in \mathbb{R}^{n \times m}$:

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

Properties:

- **Double transpose:** $(A^T)^T = A$

- **Addition transpose:** $(A+B)^T = A^T + B^T$
- **Scalar transpose:** $(cA)^T = cA^T$
- **Product transpose:** $(AB)^T = B^T A^T$

Example:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

Special Types of Matrices

Square Matrices

A matrix $A \in \mathbb{R}^{n \times n}$ where number of rows equals number of columns.

Identity Matrix

The $n \times n$ **identity matrix** I_n has 1's on the diagonal and 0's elsewhere:

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Property: $AI = IA = A$ for any compatible matrix A.

Zero Matrix

The **zero matrix** O has all entries equal to 0:

$$O = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

Diagonal Matrix

A square matrix where all non-diagonal entries are zero:

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

Symmetric Matrix

A square matrix A is **symmetric** if $A = A^T$:

$$A_{ij} = A_{ji} \text{ for all } i, j$$

Skew-Symmetric Matrix

A square matrix A is **skew-symmetric** if $A = -A^T$:

$$A_{ij} = -A_{ji} \text{ for all } i, j$$

Upper/Lower Triangular Matrices

- **Upper triangular**: $A_{ij} = 0$ for $i > j$
- **Lower triangular**: $A_{ij} = 0$ for $i < j$

Examples:

$$U = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{pmatrix}$$

Interactive: Meet the matrix personalities with **Special Matrices** - see how identity preserves, diagonal stretches, symmetric reflects, and more.

Matrix Properties

Trace

The **trace** of square matrix $A \in \mathbb{R}^{n\times n}$:

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

Properties:

- **Linearity:** $\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$
- **Scalar multiplication:** $\operatorname{tr}(cA) = c \cdot \operatorname{tr}(A)$
- **Cyclic property:** $\operatorname{tr}(AB) = \operatorname{tr}(BA)$
- **Transpose invariance:** $\operatorname{tr}(A^T) = \operatorname{tr}(A)$

Matrix Inverse

For square matrix $A \in \mathbb{R}^{n\times n}$, the **inverse** A^{-1} satisfies:

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

Properties:

- **Double inverse:** $(A^{-1})^{-1} = A$
- **Product inverse:** $(AB)^{-1} = B^{-1}A^{-1}$

- **Transpose inverse:** $(A^T)^{-1} = (A^{-1})^T$
- **Determinant inverse:** $\det(A^{-1}) = \frac{1}{\det(A)}$

Determinant

For square matrix $A \in \mathbb{R}^{n\times n}$, the **determinant** $\det(A)$ is a scalar.

For 2×2 matrices:

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

For 3×3 matrices (cofactor expansion):

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \det \begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

Properties:

- **Multiplicativity:** $\det(AB) = \det(A)\det(B)$
- **Transpose invariance:** $\det(A^T) = \det(A)$
- **Scalar scaling:** $\det(cA) = c^n \det(A)$ for $n \times n$ matrix
- **Invertibility criterion:** A is invertible iff $\det(A) \neq 0$

Interactive: Understand geometric invariants with **Determinant & Trace** - see how determinant measures area scaling and trace connects to eigenvalues.

Systems of Linear Equations

A system of m linear equations in n unknowns can be written as:

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

Matrix Form

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

where:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Solution Methods

1. Matrix Inverse (when A is square and invertible)

$$\mathbf{x} = A^{-1}\mathbf{b}$$

2. Gaussian Elimination Transform the augmented matrix $[A|\mathbf{b}]$ to row echelon form.

3. Cramer's Rule (for square systems with $\det(A) \neq 0$)

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

where A_i is A with column i replaced by $\mathbf{b}$.

Solution Types

1. **Unique solution**: $\det(A) \neq 0$ (for square systems)
2. **No solution**: System is inconsistent
3. **Infinitely many solutions**: System is underdetermined

Practice Problems

Problem 1

Given matrices:

$$A = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

Calculate: a) $A + B$ b) AB c) BA d) A^T e) $\det(A)$

Problem 2

Find the inverse of:

$$C = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$$

Problem 3

Solve the system using matrix methods:

$$2x + y = 5$$
$$x + 3y = 7$$

Problem 4

For what values of k does the following system have: a) A unique solution b) No solution c) Infinitely many solutions

$$\begin{aligned} x + 2y &= 3 \\ 2x + ky &= 6 \end{aligned}$$

Next: Linear Transformations

Solutions

Solution to Problem 1

Given $A = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$:

- a) $A + B = \begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix}$
- b) $AB = \begin{pmatrix} 2 \cdot 1 + 1 \cdot 0 & 2 \cdot 2 + 1 \cdot 1 \\ 3 \cdot 1 + 4 \cdot 0 & 3 \cdot 2 + 4 \cdot 1 \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ 3 & 10 \end{pmatrix}$
- c) $BA = \begin{pmatrix} 1 \cdot 2 + 2 \cdot 3 & 1 \cdot 1 + 2 \cdot 4 \\ 0 \cdot 2 + 1 \cdot 3 & 0 \cdot 1 + 1 \cdot 4 \end{pmatrix} = \begin{pmatrix} 8 & 9 \\ 3 & 4 \end{pmatrix}$
- d) $A^T = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$
- e) $\det(A) = 2 \cdot 4 - 1 \cdot 3 = 5$

Solution to Problem 2

$C = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$, $\det C = 3 \cdot 1 - 1 \cdot 2 = 1$. Hence

$$C^{-1} = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix}.$$

Solution to Problem 3

Solve $\begin{cases} 2x + y = 5 \\ x + 3y = 7 \end{cases}$. From the first, $y = 5 - 2x$. Substitute: $x + 3(5 - 2x) = 7 \Rightarrow x + 15 - 6x = 7 \Rightarrow -5x = -8 \Rightarrow x = \frac{8}{5}$. Then $y = 5 - \frac{16}{5} = \frac{9}{5}$.

Solution to Problem 4

System $\begin{cases} x + 2y = 3 \\ 2x + ky = 6 \end{cases}$ has coefficient matrix $\begin{pmatrix} 1 & 2 \\ 2 & k \end{pmatrix}$ with determinant $k - 4$.

- a) Unique solution iff $k \neq 4$.
- b) No solution when $k = 4$ and RHS inconsistent. Here second equation equals $2(x + 2y) = 6$, which is consistent with first, so "no solution" never occurs; it gives infinitely many.
- c) Infinitely many solutions when $k = 4$ (dependent equations).

Linear Transformations

Table of Contents

Definition of Linear Transformations

Philosophical Hook

"A linear transformation is a law of change: a way the world reshapes while preserving proportion and relation."

A **linear transformation** (or linear map) $T : V \to W$ between vector spaces V and W satisfies two essential principles:

1. **Additivity**: $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V$
2. **Homogeneity**: $T(c\mathbf{v}) = cT(\mathbf{v})$ for all $c \in \mathbb{F}$, $\mathbf{v} \in V$

Equivalently:

$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)$ for all scalars c_1, c_2 and vectors $\mathbf{v}_1, \mathbf{v}_2$.

Important Consequences

- The zero vector maps to the zero vector: $T(\mathbf{0}) = \mathbf{0}$
- Negatives are preserved: $T(-\mathbf{v}) = -T(\mathbf{v})$
- Linear combinations are preserved: $T\left(\sum_{i=1}^{n} c_i\mathbf{v}_i\right) = \sum_{i=1}^{n} c_iT(\mathbf{v}_i)$

Sparks of Imagination

- **Physics**: A linear transformation is like a law of nature that treats combinations fairly: if you double the cause, you double the effect; if you add two causes, you add their effects.

- **Sociology**: Imagine a policy that impacts individuals consistently: the combined effect on society is just the sum of the effects on each person. That's linearity.
- **Philosophy**: It encodes proportional justice — actions scale, combinations add, nothing gets distorted beyond recognition.
- **Everyday life**: Volume knobs, dimming lights, scaling images — all linear until limits are reached.

Visual Intuition

- **Grid Mapping**: Place a square grid in the plane. A linear transformation turns it into a slanted grid, but lines remain straight, and parallel lines remain parallel.
- **Basis Vectors**: Watch what happens to the simplest arrows ($\mathbf{e}_1$, $\mathbf{e}_2$). Where they go determines everything else.
- **Fabric Analogy**: Imagine a sheet of elastic cloth. A linear transformation stretches, rotates, or shears it — but never tears, curves, or wrinkles.

Interactive: Experience transformation as change with **Linear Transformation Basics -** see how linearity preserves structure while reshaping space.

Matrix Representation

Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ can be represented by an $m \times n$ matrix A such that:

$$T(\mathbf{x}) = A\mathbf{x}$$

Finding the Matrix Representation

If $T : \mathbb{R}^n \to \mathbb{R}^m$, then the matrix A has columns:

$$A = [T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \cdots \mid T(\mathbf{e}_n)]$$

where $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is the standard basis for $\mathbb{R}^n$.

Example

Find the matrix for $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by:

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + y \\ x - y \\ 3x \end{pmatrix}$$

Solution:

$$T \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \quad T \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

Therefore: $A = \begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 3 & 0 \end{pmatrix}$

Common Linear Transformations

1. Scaling (Dilation)

Scale by factor k in all directions:

$$T(\mathbf{x}) = k\mathbf{x}, \quad A = kI = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$$

2. Reflection

Reflection across x-axis:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Reflection across y-axis:

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

Reflection across line $y = x$:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

3. Rotation

Counterclockwise rotation by angle θ:

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Common angles: - 90°: $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ - 180°: $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ - 270°: $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

4. Shearing

Horizontal shear:

$$A = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

Vertical shear:

$$A = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$$

5. Projection

Orthogonal projection onto x-axis:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Orthogonal projection onto line through origin with direction u:

$$P = \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}}$$

> Note: This formula requires $\mathbf{u} \neq \mathbf{0}$. If you normalize $\mathbf{u}$ to a unit vector $\hat{\mathbf{u}} = \mathbf{u}/\|\mathbf{u}\|$, then the projection matrix simplifies to $P = \hat{\mathbf{u}}\,\hat{\mathbf{u}}^T$.

Interactive: Explore transformation types with **Common Transformations** - see scaling, rotation, reflection, shearing, and projection in action.

Properties of Linear Transformations

Rank and Nullity

For linear transformation $T : V \to W$: - **Rank**: $\text{rank}(T) = \dim(\text{Im}(T))$ - **Nullity**: $\text{nullity}(T) = \dim(\text{Ker}(T))$

Rank-Nullity Theorem

$$\dim(V) = \text{rank}(T) + \text{nullity}(T)$$

Types of Linear Transformations

Injective (One-to-One) T is **injective** if $T(\mathbf{u}) = T(\mathbf{v}) \Rightarrow \mathbf{u} = \mathbf{v}$

Equivalent conditions: - $\text{Ker}(T) = \{\mathbf{0}\}$ - $\text{nullity}(T) = 0$ - Columns of matrix A are linearly independent

Surjective (Onto) T is **surjective** if for every $\mathbf{w} \in W$, there exists $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{w}$

Equivalent conditions: - $\text{Im}(T) = W$ - $\text{rank}(T) = \dim(W)$ - Columns of matrix A span $\mathbb{R}^m$

Bijective (Isomorphism) T is **bijective** if it's both injective and surjective.

Equivalent conditions: - T has an inverse T^{-1} - Matrix A is invertible (square and $\det(A) \neq 0$)

Kernel and Image

Kernel (Null Space)

The **kernel** of $T : V \to W$ is:

$$\text{Ker}(T) = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\}$$

For matrix A, this is the **null space**:

$$\text{Null}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$$

Image (Range/Column Space)

The **image** of $T : V \to W$ is:

$$\text{Im}(T) = \{T(\mathbf{v}) : \mathbf{v} \in V\}$$

For matrix A, this is the **column space**:

$$\text{Col}(A) = \text{span}\{\text{columns of } A\}$$

Finding Kernel and Image

Example: Find kernel and image of $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \\ 1 & 2 & 2 \end{pmatrix}$

Kernel: Solve $A\mathbf{x} = \mathbf{0}$

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Row reduction gives: $\text{Ker}(A) = \text{span}\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \right\}$

Image: Find basis for column space

$$\text{Im}(A) = \text{span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right\}$$

Interactive: Visualize kernel and image with **Kernel and Image** - see directions that vanish versus directions preserved, and understand the rank-nullity theorem geometrically.

Composition and Inverse

Composition of Linear Transformations

If $T_1 : U \to V$ and $T_2 : V \to W$ are linear, then $(T_2 \circ T_1) : U \to W$ is linear:

$$(T_2 \circ T_1)(\mathbf{u}) = T_2(T_1(\mathbf{u}))$$

Matrix representation: If T_1 has matrix A_1 and T_2 has matrix A_2, then $T_2 \circ T_1$ has matrix $A_2 A_1$.

Inverse Linear Transformation

If $T : V \to W$ is bijective, then $T^{-1} : W \to V$ exists and is linear.

Matrix representation: If T has matrix A, then T^{-1} has matrix A^{-1}.

Change of Basis

If $T : V \to V$ and we change from basis $\mathcal{B}$ to basis $\mathcal{C}$, the matrix changes:

$$[T]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}P$$

where P is the change of basis matrix from $\mathcal{B}$ to $\mathcal{C}$.

Interactive: Master composition and change of basis with **Composition and Change of Basis** - see how transformations combine and how the same transformation looks different in different coordinate systems.

Practice Problems

Problem 1

Determine which of the following are linear transformations:

a) $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ 2x-y \end{pmatrix}$

b) $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 \\ y \end{pmatrix}$

c) $T : \mathbb{R}^2 \to \mathbb{R}$, $T\begin{pmatrix} x \\ y \end{pmatrix} = 3x - 2y + 1$

Problem 2

Find the matrix representation of the linear transformation that: - Rotates by 45° counterclockwise - Then reflects across the x-axis

Problem 3

For $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 3 & 1 \end{pmatrix}$:

a) Find $\mathrm{Ker}(A)$
b) Find $\mathrm{Im}(A)$
c) Verify the rank-nullity theorem

Problem 4

Given $T : \mathbb{R}^3 \to \mathbb{R}^2$ with matrix $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix}$:

a) Is T injective?
b) Is T surjective?
c) Find a vector in $\mathrm{Ker}(T)$ if it exists

Next: Eigenvalues and Eigenvectors

Solutions

Solution to Problem 1

a) Linear (additivity and homogeneity both hold).
b) Not linear (contains x^2).
c) Not linear (constant term $+1$ violates $T(0) = 0$ and additivity).

Solution to Problem 2

Rotation by 45° then reflect across x-axis: $A = R_{45}\, R_{\text{x-ref}}$ or, applying reflection after rotation, $A = R_{\text{x-ref}}\, R_{45}$ depending on stated order. With reflection after rotation:

$$R_{45} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}, \quad R_{\text{x-ref}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A = R_{\text{x-ref}} R_{45} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}.$$

Solution to Problem 3

For $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 1 & 2 & 2 \end{pmatrix}$ (from the example): Row reduction shows $\mathrm{rank}(A) = 2$.

- Kernel: Solve $A\mathbf{x} = 0$ gives $\mathbf{x} = t \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$. So $\mathrm{Ker}(A) = \mathrm{span}\{(2, -1, 1)^T\}$.

- Image/column space: Columns satisfy $c_1 = c_2 - c_3$, so a basis is $\{c_2, c_3\} = \{(2,1,2)^T, (1,3,2)^T\}$.

Solution to Problem 4

Composition corresponds to matrix multiplication. If T_1 has A_1 and T_2 has A_2, then $(T_2 \circ T_1)$ has matrix $A_2 A_1$ (order matters).

Eigenvalues and Eigenvectors

Table of Contents

Definition and Basic Concepts

Eigenvalues and Eigenvectors

For square matrix $A \in \mathbb{R}^{n \times n}$: - **Eigenvector**: A non-zero vector $\mathbf{v} \in \mathbb{R}^n$ such that $A\mathbf{v} = \lambda \mathbf{v}$ for some scalar λ - **Eigenvalue**: The scalar λ corresponding to eigenvector $\mathbf{v}$

Geometric Interpretation

An eigenvector is a direction that remains unchanged (up to scaling) under the linear transformation represented by A.

> Intuition: A stretches or squashes space differently in different directions. Along certain special directions (the eigenvectors), the action is a pure stretch by a factor λ (possibly a flip if $\lambda < 0$). This is why eigenvalues show up in applications like PCA (principal directions of variance) and stability of dynamical systems (growth/decay rates along invariant directions).

Eigenspace

The **eigenspace** E_λ corresponding to eigenvalue λ is:

$$E_\lambda = \{\mathbf{v} \in \mathbb{R}^n : A\mathbf{v} = \lambda \mathbf{v}\} = \mathrm{Ker}(A - \lambda I)$$

Finding Eigenvalues and Eigenvectors

Step 1: Find Eigenvalues

Solve the **characteristic equation**:

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

The **characteristic polynomial** is:

$$p(\lambda) = \det(A - \lambda I)$$

Step 2: Find Eigenvectors

For each eigenvalue λ_i, solve:

$$(A - \lambda_i I)\mathbf{v} = \mathbf{0}$$

Example: 2×2 Matrix

Find eigenvalues and eigenvectors of $A = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$

Step 1: Characteristic polynomial

$$\det(A - \lambda I) = \det \begin{pmatrix} 3-\lambda & 1 \\ 0 & 2-\lambda \end{pmatrix} = (3-\lambda)(2-\lambda) = 0$$

Eigenvalues: $\lambda_1 = 3$, $\lambda_2 = 2$

Step 2: Eigenvectors

For $\lambda_1 = 3$:

$$(A - 3I)\mathbf{v} = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Eigenvector: $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

For $\lambda_2 = 2$:

$$(A - 2I)\mathbf{v} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Eigenvector: $\mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

Animation: Watch "Eigenvalues & Eigenvectors" demo to see how eigenvectors remain on the same line while being scaled by their eigenvalues.

Properties of Eigenvalues and Eigenvectors

Basic Properties

1. **Trace**: $\text{tr}(A) = \sum_{i=1}^{n} \lambda_i$ (sum of eigenvalues)
2. **Determinant**: $\det(A) = \prod_{i=1}^{n} \lambda_i$ (product of eigenvalues)
3. **Zero eigenvalue**: $\lambda = 0$ is an eigenvalue iff A is singular
4. **Eigenvectors**: Eigenvectors corresponding to different eigenvalues are linearly independent

Algebraic and Geometric Multiplicity

- **Algebraic multiplicity**: Multiplicity of λ as a root of characteristic polynomial
- **Geometric multiplicity**: $\dim(E_\lambda) = \text{nullity}(A - \lambda I)$

Important: Geometric multiplicity $\leq$ Algebraic multiplicity

Similar Matrices

Matrices A and B are **similar** if $B = P^{-1}AP$ for some invertible P.

Properties of similar matrices: - Same eigenvalues (with same multiplicities) - Same trace and determinant - Same characteristic polynomial

---

Diagonalization

Definition

Matrix A is **diagonalizable** if it's similar to a diagonal matrix:

$$A = PDP^{-1}$$

where D is diagonal and P is invertible.

Diagonalization Theorem

An $n \times n$ matrix A is diagonalizable iff: 1. A has n linearly independent eigenvectors, OR 2. For each eigenvalue, geometric multiplicity = algebraic multiplicity

Diagonalization Process

1. Find all eigenvalues of A
2. For each eigenvalue, find a basis for its eigenspace
3. If total number of basis vectors equals n, then A is diagonalizable
4. Form P using eigenvectors as columns
5. Form D with corresponding eigenvalues on diagonal

Example

Diagonalize $A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}$

Step 1: Find eigenvalues

$$\det(A - \lambda I) = (4-\lambda)(1-\lambda) + 2 = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3) = 0$$

Eigenvalues: $\lambda_1 = 2$, $\lambda_2 = 3$

Step 2: Find eigenvectors

For $\lambda_1 = 2$: $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

For $\lambda_2 = 3$: $\mathbf{v}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

Step 3: Form matrices

$$P = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

Verification: $A = PDP^{-1}$

Applications

Powers of Matrices

If $A = PDP^{-1}$, then:

$$A^k = PD^kP^{-1} = P \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix} P^{-1}$$

Systems of Differential Equations

The system $\mathbf{x}' = A\mathbf{x}$ has solution:

$$\mathbf{x}(t) = e^{At}\mathbf{x}_0$$

If $A = PDP^{-1}$, then:

$$e^{At} = Pe^{Dt}P^{-1} = P \begin{pmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n t} \end{pmatrix} P^{-1}$$

Principal Component Analysis (PCA)

In data analysis, eigenvalues and eigenvectors of covariance matrices identify: - **Eigenvectors**: Principal directions of data variation - **Eigenvalues**: Amount of variance in each direction

Stability Analysis

For dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$: - **Stable**: All $|\lambda_i| < 1$ - **Unstable**: Any $|\lambda_i| > 1$ - **Marginally stable**: All $|\lambda_i| \leq 1$ with some $|\lambda_i| = 1$

Markov Chains

For stochastic matrix P: - Eigenvalue $\lambda = 1$ always exists - Corresponding eigenvector gives steady-state distribution - Other eigenvalues determine convergence rate

---

Special Types of Matrices

Symmetric Matrices

For symmetric matrix $A = A^T$: - All eigenvalues are real - Eigenvectors corresponding to different eigenvalues are orthogonal - Always diagonalizable: $A = QDQ^T$ where Q is orthogonal

Orthogonal Matrices

For orthogonal matrix Q (where $Q^TQ = I$): - All eigenvalues have absolute value 1 - Over the complex numbers, eigenvalues are of the form $e^{i\theta}$ (points on the unit circle). - Over the reals, eigenvalues are either $+1$ or -1 when a real eigenvalue exists; 2×2 rotation blocks correspond to complex-conjugate pairs $e^{\pm i\theta}$.

Positive Definite Matrices

Matrix A is **positive definite** if: - A is symmetric - All eigenvalues are positive - $\mathbf{x}^TA\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$

Practice Problems

Problem 1

Find eigenvalues and eigenvectors of:

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

Problem 2

Determine if the following matrix is diagonalizable:

$$B = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

Problem 3

If A has eigenvalues $\lambda_1 = 2$, $\lambda_2 = -1$, $\lambda_3 = 3$, find: a) $\text{tr}(A)$ b) $\det(A)$ c) Eigenvalues of A^2 d) Eigenvalues of A^{-1} (if it exists)

Problem 4

For the matrix $C = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}$: a) Find eigenvalues and eigenvectors b) Diagonalize C c) Compute C^{10}

Problem 5

Show that if A is invertible, then λ is an eigenvalue of A iff $\frac{1}{\lambda}$ is an eigenvalue of A^{-1}.

Next: Inner Product Spaces and Orthogonality

Solutions

Solution to Problem 1

For $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$, the characteristic polynomial is $\lambda^2 - 3\lambda - 4 = (\lambda - 4)(\lambda + 1)$, so $\lambda_1 = 4$, $\lambda_2 = -1$. Eigenvectors: - For $\lambda = 4$: $(A - 4I)v = 0 \Rightarrow \begin{pmatrix} -3 & 2 \\ 3 & -2 \end{pmatrix} v = 0$, take $v_1 = (2, 3)^T$. - For $\lambda = -1$: $(A + I)v = 0 \Rightarrow \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} v = 0$, take $v_2 = (1, -1)^T$.

Solution to Problem 2

Diagonalizable iff there are n linearly independent eigenvectors. Equivalently, for each eigenvalue, geometric multiplicity equals algebraic multiplicity; in total the sum of geometric multiplicities equals n.

Solution to Problem 3

If A is invertible and $Av = \lambda v$ with $v \neq 0$, then $v = A^{-1}(\lambda v) = \lambda A^{-1}v$, so $A^{-1}v = \frac{1}{\lambda}v$. Conversely, if $A^{-1}v = \mu v$, multiply by A to get $v = \mu A v$, hence $Av = \frac{1}{\mu}v$.

Solution to Problem 4

For $C = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}$, characteristic polynomial: $\lambda^2 - 7\lambda + 6 = (\lambda - 1)(\lambda - 6)$. Eigenvectors: - $\lambda = 6$: $(C - 6I)v = 0 \Rightarrow \begin{pmatrix} -1 & -2 \\ -2 & -4 \end{pmatrix} v = 0$, take $v_1 = (2, -1)^T$. - $\lambda = 1$: $(C - I)v = 0 \Rightarrow \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} v = 0$, take $v_2 = (1, 2)^T$. Diagonalization: $P = [v_1 \ v_2]$, $D = \mathrm{diag}(6, 1)$, and $C = PDP^{-1}$.

Solution to Problem 5

If A is invertible, then $Av = \lambda v \Leftrightarrow A^{-1}v = \frac{1}{\lambda}v$. Thus eigenvalues of A^{-1} are reciprocals of nonzero eigenvalues of A, with the same eigenvectors.

Inner Product Spaces and Orthogonality

Table of Contents

Inner Product Spaces

An **inner product** on a real vector space V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ satisfying:

Inner Product Axioms

For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and scalars $c \in \mathbb{R}$:

1. **Symmetry**: $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
2. **Linearity in first argument**:
    - $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$
    - $\langle \mathbf{u} + \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle$
3. **Positive definiteness**: $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, with equality iff $\mathbf{v} = \mathbf{0}$

Standard Inner Product in $\mathbb{R}^n$

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i$$

Induced Norm

The inner product induces a **norm**:

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

Cauchy-Schwarz Inequality

For any vectors $\mathbf{u}, \mathbf{v}$ in an inner product space:

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|$$

Equality holds iff $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent.

Triangle Inequality

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

Orthogonality and Orthonormal Sets

Orthogonal Vectors

Vectors $\mathbf{u}$ and $\mathbf{v}$ are **orthogonal** if:

$$\langle \mathbf{u}, \mathbf{v} \rangle = 0$$

Notation: $\mathbf{u} \perp \mathbf{v}$

Orthogonal Sets

A set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is **orthogonal** if:

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0 \text{ for all } i \neq j$$

Orthonormal Sets

A set is **orthonormal** if it's orthogonal and each vector has unit length:

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Properties of Orthogonal Sets

1. **Linear Independence**: Any orthogonal set of non-zero vectors is linearly independent
2. **Pythagorean Theorem**: If $\mathbf{u} \perp \mathbf{v}$, then $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$
3. **Orthogonal Basis**: An orthogonal basis simplifies many computations

Coordinates in Orthonormal Basis

If $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$ is an orthonormal basis for V, then:

$$\mathbf{v} = \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{q}_i \rangle \mathbf{q}_i$$

Gram-Schmidt Process

The **Gram-Schmidt process** converts a linearly independent set into an orthogonal set.

Algorithm

Given linearly independent vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$:

Step 1: $\mathbf{u}_1 = \mathbf{v}_1$

Step 2: $\mathbf{u}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \mathbf{u}_1$

Step 3: $\mathbf{u}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \mathbf{u}_1 - \frac{\langle \mathbf{v}_3, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle} \mathbf{u}_2$

General Step:

$$\mathbf{u}_k = \mathbf{v}_k - \sum_{j=1}^{k-1} \frac{\langle \mathbf{v}_k, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle} \mathbf{u}_j$$

Normalization

To get orthonormal vectors:

$$\mathbf{q}_k = \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}$$

Example

Orthogonalize $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

Step 1: $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

Step 2:

$$\mathbf{u}_2 = \begin{pmatrix}1\\1\\0\end{pmatrix} - \frac{2}{3}\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}1/3\\1/3\\-2/3\end{pmatrix}$$

Step 3:

$$\mathbf{u}_3 = \begin{pmatrix}1\\0\\0\end{pmatrix} - \frac{1}{3}\begin{pmatrix}1\\1\\1\end{pmatrix} - \frac{1/3}{2/3}\begin{pmatrix}1/3\\1/3\\-2/3\end{pmatrix} = \begin{pmatrix}1/2\\-1/2\\0\end{pmatrix}$$

Orthogonal Projections

Projection onto a Line

The **orthogonal projection** of $\mathbf{v}$ onto line spanned by $\mathbf{u}$ is:

$$\text{proj}_{\mathbf{u}}\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{u}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u}$$

Projection onto a Subspace

For subspace W with orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}$:

$$\text{proj}_W\mathbf{v} = \sum_{i=1}^{k} \langle \mathbf{v}, \mathbf{q}_i\rangle \mathbf{q}_i$$

Projection Matrix

If $Q = [\mathbf{q}_1 \mid \mathbf{q}_2 \mid \cdots \mid \mathbf{q}_k]$ has orthonormal columns:

$$P = QQ^T$$

Then $P\mathbf{v} = \text{proj}_W\mathbf{v}$ for any vector $\mathbf{v}$.

Properties of Projection Matrices

1. **Idempotent**: $P^2 = P$
2. **Symmetric**: $P^T = P$
3. **Eigenvalues**: Only 0 and 1

Orthogonal Complement

The **orthogonal complement** of subspace W is:

$$W^\perp = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{w}\rangle = 0 \text{ for all } \mathbf{w} \in W\}$$

Fundamental theorem: $V = W \oplus W^\perp$ (direct sum)

QR Decomposition

Every matrix $A \in \mathbb{R}^{m \times n}$ with linearly independent columns can be factored as:

$$A = QR$$

where: - $Q \in \mathbb{R}^{m \times n}$ has orthonormal columns - $R \in \mathbb{R}^{n \times n}$ is upper triangular with positive diagonal entries

Construction via Gram-Schmidt

If $A = [\mathbf{a}_1 \mid \mathbf{a}_2 \mid \cdots \mid \mathbf{a}_n]$, apply Gram-Schmidt to get $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$.

Then:

$$Q = [\mathbf{q}_1 \mid \mathbf{q}_2 \mid \cdots \mid \mathbf{q}_n]$$

$$R = \begin{pmatrix} \langle \mathbf{a}_1, \mathbf{q}_1 \rangle & \langle \mathbf{a}_2, \mathbf{q}_1 \rangle & \cdots & \langle \mathbf{a}_n, \mathbf{q}_1 \rangle \\ 0 & \langle \mathbf{a}_2, \mathbf{q}_2 \rangle & \cdots & \langle \mathbf{a}_n, \mathbf{q}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \langle \mathbf{a}_n, \mathbf{q}_n \rangle \end{pmatrix}$$

Applications of QR Decomposition

1. **Solving linear systems**: $A\mathbf{x} = \mathbf{b} \Rightarrow R\mathbf{x} = Q^T\mathbf{b}$
2. **Least squares problems**
3. **Computing eigenvalues** (QR algorithm)
4. **Numerical stability** in computations

Least Squares Solutions

For overdetermined system $A\mathbf{x} = \mathbf{b}$ (more equations than unknowns), find $\mathbf{x}$ that minimizes $\|A\mathbf{x} - \mathbf{b}\|^2$.

Normal Equations

The least squares solution satisfies:

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

Solution: $\mathbf{x} = (A^T A)^{-1} A^T \mathbf{b}$ (if $A^T A$ is invertible)

Geometric Interpretation

The least squares solution makes $A\mathbf{x}$ the orthogonal projection of $\mathbf{b}$ onto $\text{Col}(A)$:

$$A\mathbf{x} = \text{proj}_{\text{Col}(A)}\mathbf{b}$$

QR Method for Least Squares

If $A = QR$, then:

$$\mathbf{x} = R^{-1}Q^T\mathbf{b}$$

This is more numerically stable than using normal equations.

Applications

1. **Linear regression**: Fitting lines/curves to data
2. **Data fitting**: Polynomial approximation
3. **Signal processing**: Noise reduction
4. **Computer graphics**: Surface fitting

Orthogonal Matrices

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

Properties

1. **Preserves lengths**: $\|Q\mathbf{x}\| = \|\mathbf{x}\|$
2. **Preserves angles**: $\langle Q\mathbf{u}, Q\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$
3. **Determinant**: $\det(Q) = \pm 1$
4. **Eigenvalues**: All have absolute value 1

Examples

- **Rotation matrices**: $\det(Q) = 1$
- **Reflection matrices**: $\det(Q) = -1$
- **Householder reflectors**: $H = I - 2\mathbf{u}\mathbf{u}^T$ where $\|\mathbf{u}\| = 1$

Practice Problems

Problem 1

Apply the Gram-Schmidt process to orthogonalize:

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

Problem 2

Find the QR decomposition of:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Problem 3

Find the orthogonal projection of $\mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ onto the subspace spanned by:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Problem 4

Solve the least squares problem for the overdetermined system:

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$$

Problem 5

Show that if Q is orthogonal, then $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ for any vector $\mathbf{x}$.

This completes the core linear algebra sequence. Next topics might include: Singular Value Decomposition, Spectral Theory, or applications to specific fields.

Solutions

Solution to Problem 1 (Gram–Schmidt)

Compute

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{u}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Then

$$\mathbf{u}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 - \frac{\langle \mathbf{v}_3, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\mathbf{u}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \tfrac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - 0 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 0 \\ -1/2 \end{pmatrix}.$$

Orthonormal set: $\mathbf{q}_1 = \frac{1}{\sqrt{2}}(1, 0, 1)^T$, $\mathbf{q}_2 = (0, 1, 0)^T$, $\mathbf{q}_3 = \frac{1}{\sqrt{2}}(1, 0, -1)^T$.

Solution to Problem 2 (QR of A)

Apply Gram–Schmidt to columns of A to form Q with orthonormal columns and set $R = Q^T A$. Concretely, the first column normalizes $(1, 1, 0)^T$; the second subtracts its projection and normalizes; then compute R entries via inner products $r_{ij} = \langle a_j, q_i \rangle$.

Solution to Problem 3 (Projection)

With $\mathbf{b} = (1, 2, 3)^T$, $\mathbf{v}_1 = (1, 1, 0)^T$, $\mathbf{v}_2 = (1, 0, 1)^T$, orthonormalize to $Q = [\mathbf{q}_1\ \mathbf{q}_2]$ as in Solution 1, then $\text{proj}_{\text{span}\{v_1, v_2\}}(\mathbf{b}) = QQ^T\mathbf{b}$.

Solution to Problem 4 (Least Squares via QR)

For $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}$ and $\mathbf{b} = (2, 3, 5)^T$, QR gives $A = QR$. Solve $R\mathbf{x} = Q^T\mathbf{b}$ to obtain $\mathbf{x} = (1, 1)^T$.

Solution to Problem 5 (Orthogonal preserves norm)

Since $Q^T Q = I$, we have $\|Q\mathbf{x}\|^2 = (Q\mathbf{x})^T(Q\mathbf{x}) = \mathbf{x}^T(Q^T Q)\mathbf{x} = \mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2$.

Spectral Theorem & Applications

Table of Contents

Statement of the Spectral Theorem

For finite-dimensional vector spaces:

- Over the reals: A real symmetric matrix $A = A^T$ is diagonalizable by an orthogonal matrix Q:
$$A = QDQ^T$$
where D is diagonal with real eigenvalues, and Q has orthonormal eigenvectors as columns.

- Over the complex numbers: A complex Hermitian matrix $A = A^*$ is diagonalizable by a unitary matrix U:

$$A = UDU^*$$

with D real diagonal; eigenvalues are real and eigenvectors can be chosen orthonormal.

Consequences

- Eigenvectors corresponding to distinct eigenvalues are orthogonal.
- All eigenvalues of Hermitian/symmetric matrices are real.
- Quadratic form $x^T Ax$ (or x^*Ax) can be analyzed via eigenvalues.

Hermitian/Symmetric Matrices

- Real symmetric: $A = A^T$.
- Complex Hermitian: $A = A^*$ (conjugate transpose).
- Positive semidefinite (PSD): $x^*Ax \geq 0$ for all x; equivalent to all eigenvalues ≥ 0.
- Positive definite (PD): $x^*Ax > 0$ for all $x \neq 0$; eigenvalues > 0.

QC connection: Observables (e.g., spin, energy) are represented by Hermitian operators; measurement outcomes are the real eigenvalues.

ML connection: Sample covariance matrices are symmetric PSD; their eigendecomposition underpins PCA.

Orthogonal/Unitary Diagonalization

If A is symmetric/Hermitian, there exists orthonormal eigenbasis $\{q_i\}_{i=1}^n$ and real eigenvalues $\{\lambda_i\}$ such that:

$$A = \sum_{i=1}^{n} \lambda_i q_i q_i^T \quad \text{(real)}, \qquad A = \sum_{i=1}^{n} \lambda_i q_i q_i^* \quad \text{(complex)}$$

This expansion is called the spectral decomposition.

Rayleigh Quotient and Variational Characterization

For unit vector x:

$$R_A(x) = \frac{x^*Ax}{x^*x}$$

- $\lambda_{\max} = \max_{\|x\|=1} R_A(x)$ - $\lambda_{\min} = \min_{\|x\|=1} R_A(x)$

Applications in QC/ML/Applied Math

- QC: Any state can be expanded in an observable's orthonormal eigenbasis; time evolution under a Hamiltonian uses eigen-decomposition.
- ML: PCA = eigen-decomposition of covariance $\Sigma = \frac{1}{m}X^TX$; principal directions are eigenvectors with largest eigenvalues.
- Stability: For linear dynamical system $x_{k+1} = Ax_k$, eigenvalues determine growth/decay/oscillation.
- Quadratic optimization: Optimize x^TAx under constraints via eigenvalues and eigenvectors.

Animation: Use "Eigenvalues & Eigenvectors" demo in `interactive-visualizations.py` to see invariant directions and scaling.

Worked Examples

Example 1: Diagonalization of a Symmetric Matrix

Let

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

Characteristic polynomial: $\det(A-\lambda I) = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3$. Eigenvalues: $\lambda_1 = 1, \lambda_2 = 3$. Eigenvectors: for $\lambda_1 = 1$, $v_1 = (1,-1)^T$; for $\lambda_2 = 3$, $v_2 = (1,1)^T$. Orthonormal basis: $q_1 = \frac{1}{\sqrt{2}}(1,-1)^T$, $q_2 = \frac{1}{\sqrt{2}}(1,1)^T$. Then $Q = [q_1\, q_2]$, $D = \operatorname{diag}(1,3)$, and $A = QDQ^T$.

Example 2: PCA via Spectral Decomposition

Given centered data matrix $X \in \mathbb{R}^{m\times n}$, covariance $\Sigma = \frac{1}{m}X^TX$. Eigen-decompose $\Sigma = Q\Lambda Q^T$. The top-k principal components are the first k columns of Q. Projected data: $Y = XQ_k$. Explained variance ratio: $\lambda_i / \sum_j \lambda_j$.

Practice Problems

1. Prove that a real symmetric matrix has orthogonal eigenvectors corresponding to distinct eigenvalues.
2. Let A be PSD. Show $\lambda_{\min} \geq 0$ using the Rayleigh quotient.
3. Diagonalize $\begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$ and classify as PD/PSD/indefinite.
4. For a covariance matrix Σ, explain why all eigenvalues are nonnegative and how PCA selects directions.
5. QC: For a Hermitian operator H, explain why measurement outcomes are real and relate eigenbasis to measurement probabilities.

Solutions

Solution to Problem 1

If A is real symmetric, there exists an orthonormal basis of eigenvectors; let x, y be eigenvectors with $Ax = \lambda x$, $Ay = \mu y$, $\lambda \neq \mu$. Then

$$\lambda\langle x, y\rangle = \langle Ax, y\rangle = \langle x, Ay\rangle = \mu\langle x, y\rangle,$$

so $(\lambda - \mu)\langle x, y\rangle = 0$ and hence $\langle x, y\rangle = 0$.

Solution to Problem 2

For unit x, $R_A(x) = x^T Ax$. Since A is PSD, $x^T Ax \geq 0$ for all x, so the extremal values of R_A (which equal eigenvalues) satisfy $\lambda_{\min} \geq 0$.

Solution to Problem 3

Compute the characteristic polynomial, find eigenvalues/eigenvectors, and orthonormalize. One finds eigenvalues $\{4, 3, 1\}$ with orthonormal eigenvectors; the matrix is PD because all eigenvalues are > 0.

Solution to Problem 4

Covariance matrices are symmetric and for any z, $z^T \Sigma z = \mathrm{Var}(z^T X) \geq 0$, so eigenvalues are ≥ 0. PCA selects eigenvectors with largest eigenvalues (directions of maximal variance).

Solution to Problem 5

Hermitian H has real eigenvalues and an orthonormal eigenbasis. Measurement outcomes are eigenvalues (real). If $|\psi\rangle = \sum_i c_i |h_i\rangle$ in the eigenbasis, probability of outcome λ_i is $|c_i|^2$.

Matrix Decompositions (SVD, QR, LU)

Table of Contents

Why Decompose?

Decompositions express a matrix as a product of simpler, structured matrices. This reveals geometry, improves numerical stability, and powers algorithms in QC/ML.

- Compression and denoising (SVD)
- Least-squares and orthogonalization (QR)
- Fast linear solves and reuse across right-hand sides (LU)

Singular Value Decomposition (SVD)

For any matrix $A \in \mathbb{R}^{m\times n}$, there exist orthogonal matrices $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$, and diagonal $\Sigma \in \mathbb{R}^{m\times n}$ with nonnegative entries $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$ such that:

$$A = U\,\Sigma\,V^T.$$

- Columns of U are left singular vectors; columns of V are right singular vectors.
- Singular values σ_i are square-roots of eigenvalues of A^TA (or AA^T).
- Best rank-k approximation (Eckart–Young):

  $$A_k = U_{:,1:k}\,\Sigma_{1:k,1:k}\,V_{:,1:k}^T$$

  minimizes $\|A - A_k\|_F$.

Geometric view: V rotates coordinates, Σ scales along orthogonal axes, U rotates to final position.

Animation: In `interactive-visualizations.py`, the "Matrix Transformations" demo visualizes orthogonal axes and anisotropic scaling.

QR Decomposition

For $A \in \mathbb{R}^{m\times n}$ with full column rank, QR decomposes as:

$$A = QR,$$

where $Q \in \mathbb{R}^{m\times n}$ has orthonormal columns and $R \in \mathbb{R}^{n\times n}$ is upper triangular with positive diagonal.

- Gram–Schmidt yields Q and R.
- Solving least squares: $\min_x \|Ax - b\|_2$ becomes $Rx = Q^Tb$.

LU Decomposition

For many square matrices, we can factor:

$$A = LU,$$

with L unit lower triangular and U upper triangular (possibly with row permutations $PA = LU$).

- Efficient for solving $Ax = b$ with repeated right-hand sides.
- Foundation for many direct solvers in numerical linear algebra.

QC / ML Parallels

- QC: Density matrices ρ (PSD, trace 1) admit eigen/SVD-like decompositions revealing mixed-state probabilities and purifications. Partial traces relate to block-structured SVD.
- ML: SVD underpins compression, noise reduction, and recommendation systems (latent factor models). QR appears in stable least-squares; LU in fast batched solves.

Worked Examples

Example 1: SVD of a 2×2 Matrix

Let $A = \begin{pmatrix} 3 & 0 \\ 4 & 0 \end{pmatrix}$. Then $A^T A = \begin{pmatrix} 25 & 0 \\ 0 & 0 \end{pmatrix}$ with eigenvalues $25, 0$. Thus $\sigma_1 = 5, \sigma_2 = 0$. One valid SVD is

$$U = \begin{pmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{pmatrix}, \ \Sigma = \begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix}, \ V = I.$$

Example 2: QR via Gram–Schmidt

For $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$, apply Gram–Schmidt to obtain Q with orthonormal columns and compute $R = Q^T A$.

Example 3: LU with Partial Pivoting

Factor $A = \begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix}$. With permutation $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, we have $PA = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix} = LU$ with $L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$.

Practice Problems

1. Show that singular values are nonnegative and equal to the square-roots of eigenvalues of $A^T A$.
2. Compute the best rank-1 approximation of $\begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$.
3. Use QR to solve the least squares problem for $\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} x \approx \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$.
4. Perform LU factorization with partial pivoting for $\begin{pmatrix} 2 & 4 & 8 \\ 1 & 3 & 2 \\ 3 & 10 & 4 \end{pmatrix}$.
5. QC/ML: Explain how SVD underlies PCA and how eigen-decomposition of covariance relates to SVD of the data matrix.

Solutions

Solution to Problem 1

For any A, $A^T A$ is symmetric PSD. If $A^T A v = \lambda v$ with $\lambda \geq 0$, define $\sigma = \sqrt{\lambda} \geq 0$. Then singular values are $\sigma_i = \sqrt{\lambda_i(A^T A)} \geq 0$.

Solution to Problem 2

Let $B = \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$. The best rank-1 approximation is $\sigma_1 u_1 v_1^T$ using the top singular triplet. Here $\sigma_1 = 2$, $u_1 = (1, 0, 0)^T$, $v_1 = (1, 0)^T$, so $B_1 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Solution to Problem 3

QR via Gram–Schmidt yields Q with orthonormal columns and $R = Q^T A$. Solving $Rx = Q^T b$ gives the least-squares solution $x = (1, 1)^T$ for the given data.

Solution to Problem 4

Apply Gaussian elimination with partial pivoting to produce P, then read off L (unit lower) and U (upper). One valid factorization is obtainable with standard steps; determinants of leading principal minors are nonzero so factorization exists with a suitable permutation.

Solution to Problem 5

For centered data matrix X (rows as samples), $\Sigma = \frac{1}{m}X^TX$. If $X = U\Sigma_X V^T$ is the SVD, then $X^TX = V\Sigma_X^2 V^T$; columns of V (right singular vectors) are the eigenvectors of Σ (principal components) and squared singular values are the eigenvalues (variances along components).

Determinants & Advanced Properties

Table of Contents

Conceptual Role of the Determinant

While determinants are rarely used for direct computation in modern ML/QC (we prefer methods like Gaussian elimination or QR decomposition for stability and speed), they provide a profound **structural understanding** of matrices and linear transformations.

Think of the determinant not as a computational tool, but as a single number that summarizes deep geometric and algebraic properties of a matrix.

Geometric Interpretation: Scaling Factor

The determinant, `det(A)`, of a square matrix `A` is the **signed volume scaling factor** of the linear transformation represented by `A`.

- **In 2D**: `det(A)` is the signed area of the parallelogram formed by transforming the unit square.
  - If `det(A) = 2`, the area is doubled.
  - If `det(A) = 0.5`, the area is halved.
  - If `det(A) = 0`, the space collapses into a lower dimension (a line or a point), and information is lost.
- **In 3D**: `det(A)` is the signed volume of the parallelepiped formed by transforming the unit cube.

Orientation and Sign

- `det(A) > 0`: Preserves orientation (e.g., no reflection).

- `det(A) < 0`: Reverses orientation (a "flip" or reflection occurs).

Animation: The "Matrix Transformations" demo in `interactive-visualizations.py` visually shows this. The `det(A)` value is displayed, and you can see how the area of the grid changes and flips as the determinant changes sign.

Algebraic Properties & Invertibility

The most critical algebraic property of the determinant is its connection to matrix invertibility.

> A square matrix `A` is invertible **if and only if** `det(A) $\neq$ 0`.

This single fact connects many key concepts:

If `det(A) $\neq$ 0` (Invertible)	If `det(A) = 0` (Singular/Non-invertible)
`Ax = b` has a unique solution.	`Ax = b` has no solution or infinitely many.
The transformation is reversible.	The transformation is irreversible (information loss).
The columns of `A` are linearly independent.	The columns of `A` are linearly dependent.
The null space of `A` is just the zero vector.	The null space of `A` contains non-zero vectors.
All eigenvalues are non-zero.	At least one eigenvalue is zero.

Connection to Eigenvalues and Trace

For any square matrix `A`, the determinant and trace are directly related to its eigenvalues (λ_i):

- **Determinant**: The product of the eigenvalues.

$$\det(A) = \prod_{i=1}^{n} \lambda_i = \lambda_1 \lambda_2 \cdots \lambda_n$$

  This provides a beautiful intuition: since eigenvalues are the scaling factors in the principal directions, the total volume scaling factor (determinant) must be the product of these individual scaling factors.

- **Trace**: The sum of the eigenvalues.

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

These two properties are incredibly powerful theoretical tools for checking calculations and understanding matrix behavior without computing the eigenvalues directly.

Why It Matters: The Bridge to Advanced Topics

The determinant provides the theoretical glue that holds linear algebra together and serves as a bridge to more advanced fields:

- **Graph Theory & ML**: The determinant of graph-related matrices (like the Laplacian) can count spanning trees, revealing the graph's connectivity and robustness.
- **Probability & Statistics**: In change of variables for multivariable probability distributions, the determinant of the Jacobian matrix appears as the factor that preserves the total probability.
- **Quantum Computing**: While less prominent in introductory QC, determinants appear in advanced concepts like the calculation of fermionic path integrals and in the characterization of certain quantum gates and channels.
- **General Relativity**: The determinant of the metric tensor ($g_{\mu\nu}$) is used to define the volume element for integration in curved spacetime.

It closes the loop on our structural understanding, allowing us to confidently branch into these specialized areas.

Practice Problems

1. If a 3x3 matrix `A` has eigenvalues 2, 3, and -1, what is `det(A)` and `tr(A)`?
2. You apply a transformation `A` to a 1x1 square in 2D, and the resulting shape is a line. What can you say about `det(A)`?
3. If `det(A) = -1`, what does this imply about the transformation geometrically? Give an example of such a matrix.
4. Explain why a matrix with a zero eigenvalue cannot be invertible, using the determinant property.
5. Consider the transformation matrix for a 90-degree rotation in 2D. What is its determinant? Does this make sense geometrically? Why?

Solutions

Solution to Problem 1

Eigenvalues $2, 3, -1$ imply $\det(A) = 2 \cdot 3 \cdot (-1) = -6$ and $\operatorname{tr}(A) = 2+3+(-1) = 4$.

Solution to Problem 2

The image is 1D (a line), so area scales to zero; thus $\det(A) = 0$ (transformation is singular/collapses dimension).

Solution to Problem 3

$\det(A) = -1$ means orientation flips and area is preserved (magnitude 1). Example: a reflection matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ has determinant -1.

Solution to Problem 4

If $\lambda = 0$ is an eigenvalue, their product (the determinant) is zero, hence A cannot be invertible. Conversely, invertibility requires all eigenvalues nonzero, giving nonzero determinant.

Solution to Problem 5

For a 90° rotation, $R_{90} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ with $\det(R_{90}) = 1$. Geometrically this preserves area and orientation (a pure rotation), consistent with determinant 1.