Lecture 2: Linear Algebra

Mathematics for Machine Learning

Roadmap

Why Should I Care About Linear Algebra?

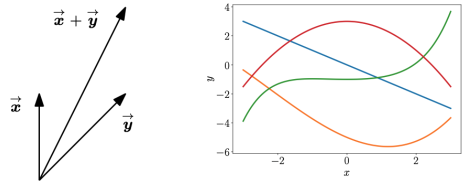
- Data = vectors: Every dataset is a matrix, every data point is a vector
- Neural networks: Every layer computes $\mathbf{x} \mapsto \mathbf{A}\mathbf{x} + \mathbf{b}$ — that's linear algebra!
- Computer vision: Image transformations (rotation, scaling) are matrix multiplications
- Dimensionality reduction: PCA, SVD — the backbone of data preprocessing

In one sentence: Linear algebra is the computational engine behind all of AI.

Roadmap

(1) Systems of Linear Equations
(2) Matrices
(3) Solving Systems of Linear Equations
(4) Vector Spaces
(5) Linear Independence
(6) Basis and Rank
(7) Linear Mappings
(8) Affine Spaces

- **Definition.** Algebra: a set of objects and rules to manipulate them
- Linear algebra:
  - Object: vectors $\mathbf{v}$
  - Operations: addition $(\mathbf{v} + \mathbf{w})$ and scalar multiplication $(k\mathbf{v})$
- **Example.** Examples of vectors:
  - Geometric vectors
  - Polynomials
  - Audio signals
  - Elements of $\mathbb{R}^n$



(a) Geometric vectors.   (b) Polynomials.

(a) Geometric vectors   (b) Polynomials

System of Linear Equations

- For unknown variables $(x_1, \cdots, x_n) \in \mathbb{R}^n$,

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

- Three cases of solutions:
  - No solution: $x_1 + x_2 + x_3 = 3, \quad x_1 - x_2 + 2x_3 = 2, \quad 2x_1 + 3x_3 = 1$
  - Unique solution: $x_1 + x_2 + x_3 = 3, \quad x_1 - x_2 + 2x_3 = 2, \quad x_2 + 3x_3 = 1$
  - Infinitely many: $x_1 + x_2 + x_3 = 3, \quad x_1 - x_2 + 2x_3 = 2, \quad 2x_1 + 3x_3 = 5$
- **Question.** Under what conditions does each case occur?

Matrix Representation

- A collection of linear equations:

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

- Matrix representation:

$$\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \Leftrightarrow \mathbf{Ax} = \mathbf{b}$$

  where $\mathbf{A}$ is the coefficient matrix, $\mathbf{x}$ is the variable vector, and $\mathbf{b}$ is the constant vector.

- Understanding $\mathbf{A}$ is the key to answering various questions about $\mathbf{Ax} = \mathbf{b}$.

Roadmap

Matrix: Addition and Multiplication

- For two matrices $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{m\times n}$,

$$\mathbf{A} + \mathbf{B} := \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n}$$

- For $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times k}$, the elements c_{ij} of $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{m\times k}$:

$$c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, k$$

- **Example.** $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}$, compute $\mathbf{AB}$ and $\mathbf{BA}$.

Identity Matrix and Matrix Properties

- **Definition.** A square matrix I_n with $I_{ii} = 1$ and $I_{ij} = 0$ for $i \neq j$:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Associativity: For $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times p}$, $\mathbf{C} \in \mathbb{R}^{p \times q}$: $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
- Distributivity: For $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$, $\mathbf{C}, D \in \mathbb{R}^{n \times p}$:
  (i) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
  (ii) $\mathbf{A}(\mathbf{C} + D) = \mathbf{AC} + \mathbf{A}D$
- Identity: $I_m\mathbf{A} = \mathbf{A}I_n = \mathbf{A}$

Inverse and Transpose

- **Definition.** For a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B}$ is the inverse of $\mathbf{A}$, denoted $\mathbf{A}^{-1}$, if

$$\mathbf{AB} = I_n = \mathbf{BA}$$

- Called regular/invertible/nonsingular, if it exists.
- If it exists, it is unique.
- For 2×2 matrix:

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

- **Definition.** For $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times m}$ with $b_{ij} = a_{ji}$ is the transpose of $\mathbf{A}$, denoted $\mathbf{A}^\top$.
- **Example.** For $\mathbf{A} = \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}$,

$$\mathbf{A}^\top = \begin{bmatrix} 0 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix}$$

- If $\mathbf{A} = \mathbf{A}^\top$, $\mathbf{A}$ is called symmetric.

- $\mathbf{A}\mathbf{A}^{-1} = I = \mathbf{A}^{-1}\mathbf{A}$
- $(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
- $(\mathbf{A} + \mathbf{B})^{-1} \neq \mathbf{A}^{-1} + \mathbf{B}^{-1}$ (common mistake!)
- $(\mathbf{A}^{\top})^{\top} = \mathbf{A}$
- $(\mathbf{A} + \mathbf{B})^{\top} = \mathbf{A}^{\top} + \mathbf{B}^{\top}$
- $(\mathbf{A}\mathbf{B})^{\top} = \mathbf{B}^{\top}\mathbf{A}^{\top}$
- If $\mathbf{A}$ is invertible, so is $\mathbf{A}^{\top}$.

Scalar Multiplication

- Multiplication by a scalar $\lambda \in \mathbb{R}$ to $\mathbf{A} \in \mathbb{R}^{m \times n}$
- **Example.** For $\mathbf{A} = \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}$, $3\mathbf{A} = \begin{bmatrix} 0 & 6 \\ 3 & -3 \\ 0 & 3 \end{bmatrix}$
- Associativity:
  - $(\lambda\psi)\mathbf{C} = \lambda(\psi\mathbf{C})$
  - $\lambda(\mathbf{BC}) = (\lambda\mathbf{B})\mathbf{C} = \mathbf{B}(\lambda\mathbf{C}) = (\mathbf{BC})\lambda$
  - $(\lambda\mathbf{C})^\top = \lambda\mathbf{C}^\top$
- Distributivity:
  - $(\lambda + \psi)\mathbf{C} = \lambda\mathbf{C} + \psi\mathbf{C}$
  - $\lambda(\mathbf{B} + \mathbf{C}) = \lambda\mathbf{B} + \lambda\mathbf{C}$

Quick Recap: Matrices

We just covered the building blocks of linear algebra — matrices.

- Matrices represent linear systems as $\mathbf{A}\mathbf{x} = \mathbf{b}$
- Key operations: addition, multiplication, inverse, transpose
- Inverse exists only for square, non-singular matrices
- Next up: Solving systems using Gaussian elimination

Roadmap

Example: Gaussian Elimination

- $-3x + 2z = -1$
- $x - 2y + 2z = -5/3$
- $-x - 4y + 6z = -13/3$

Express as augmented matrix and apply Gaussian elimination:

$$\left[\begin{array}{ccc|c} -3 & 0 & 2 & -1 \\ 1 & -2 & 2 & -5/3 \\ -1 & -4 & 6 & -13/3 \end{array}\right] \xrightarrow{\text{row ops}} \left[\begin{array}{ccc|c} -3 & 0 & 2 & -1 \\ 0 & -2 & 8/3 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The two nonzero rows give $-3x + 2z = -1$ and $-2y + (8/3)z = -2$.

Parametrizing Solutions

- From $-3x + 2z = -1$ and $-2y + (8/3)z = -2$:

$$\begin{aligned} x &= (1/3) + (2/3)z \\ y &= 1 + (4/3)z \\ z &= z \end{aligned}$$

- Solution set:

$$\left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 2/3 \\ 4/3 \\ 1 \end{bmatrix} \;\middle|\; z \in \mathbb{R} \right\}$$

- Each value of z gives a different solution.

Form of Solution Sets

- General form: $\mathbf{x} = \mathbf{x}_p + t_1\mathbf{v}_1 + \cdots + t_k\mathbf{v}_k$
- $\mathbf{x}_p$: particular solution
- $\mathbf{v}_1, \ldots, \mathbf{v}_k$: basis of the solution space of $\mathbf{Ax} = \mathbf{0}$
- $t_1, \ldots, t_k$: free parameters

The solution set is a **particular solution** plus the **null space** of **A**.

Gaussian Elimination Algorithm

(1) Write the augmented matrix $[\mathbf{A}|\mathbf{b}]$

(2) Use row operations to transform to row echelon form

(3) Back-substitute to find solutions

Row operations: swap rows, multiply a row by a nonzero scalar, add a multiple of one row to another.

Roadmap

Groups

- Think of it as: a set where you can combine any two elements, undo operations, and do nothing (identity).
- **Definition.** A set G with an operation $\circ : G \times G \to G$ is a group $(G, \circ)$ if:
  (1) Closure: Combining two elements stays in G $\quad (a \circ b \in G)$
  (2) Associativity: Order of grouping doesn't matter $\quad ((a \circ b) \circ c = a \circ (b \circ c))$
  (3) Identity: There is a "do nothing" element e $\quad (a \circ e = a)$
  (4) Inverse: Every element can be undone $\quad (a \circ a^{-1} = e)$
- **Example.** $(\mathbb{R}, +)$ is a group: closure ✓, associativity ✓, identity $= 0$, inverse of a is $-a$.

- Think of it as: a collection of objects you can add together and scale, and everything behaves "nicely."
- **Definition.** A set V with addition $+$ and scalar multiplication $\cdot$ is a vector space $(V, +, \cdot)$ over $\mathbb{R}$ if:
  (1) $(V, +)$ is an abelian group (you can add, subtract, and order doesn't matter)
  (2) $\lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v}$ (scaling distributes over addition)
  (3) $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$ (adding scalars then scaling = scaling separately)
  (4) $\lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}$ (scaling twice = scaling by product)
  (5) $1 \cdot \mathbf{v} = \mathbf{v}$ (scaling by 1 does nothing)

Examples of Vector Spaces

- $\mathbb{R}^n$ with standard addition and scalar multiplication (columns of numbers)
- Polynomials of degree at most n (e.g., $3x^2 + 2x + 1$)
- Continuous functions $C([a, b])$ (smooth curves)
- Matrices $\mathbb{R}^{m \times n}$ (tables of numbers)

Key insight: Vectors are not just arrows! Anything you can add and scale that obeys the axioms is a vector space.

Vector Subspaces

- Think of it as: a smaller vector space living inside a bigger one.
- **Definition.** $U \subset V$ is a vector subspace of V if U is itself a vector space (same operations).
- Quick test — check three things:
  (1) Does U contain the zero vector? $(\mathbf{0} \in U?)$
  (2) Is U closed under addition? $(\mathbf{u} + \mathbf{v} \in U?)$
  (3) Is U closed under scalar multiplication? $(\lambda\mathbf{u} \in U?)$
- **Example.** Examples:
  - The solution set of $\mathbf{Ax} = \mathbf{0}$ is a subspace of $\mathbb{R}^n$ ✓
  - The solution of $\mathbf{Ax} = \mathbf{b}$ $(\mathbf{b} \neq \mathbf{0})$ is not a subspace (fails zero vector test!)

Quick Recap: Vector Spaces

Vector spaces are the abstract foundation — any set with addition and scalar multiplication that satisfies the axioms.

- Groups provide the algebraic structure underlying vector spaces
- A vector space needs closure under $+$ and $\cdot$ plus five axioms
- Subspaces are vector spaces living inside a larger space
- Next up: Linear independence — when are vectors "redundant"?

Roadmap

(1) Systems of Linear Equations
(2) Matrices
(3) Solving Systems of Linear Equations
(4) Vector Spaces
(5) Linear Independence
(6) Basis and Rank
(7) Linear Mappings
(8) Affine Spaces

Linear Independence

- Think of it as: can you build one vector from the others? If yes $\rightarrow$ dependent. If no $\rightarrow$ independent.
- **Definition.** $\mathbf{v} = \lambda_1 \mathbf{x}_1 + \cdots + \lambda_k \mathbf{x}_k$ is a linear combination of $\mathbf{x}_1, \ldots, \mathbf{x}_k$.
- **Definition.** If $\mathbf{0} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i$ has a solution other than all $\lambda_i = 0$, the vectors are linearly dependent. If the only solution is $\lambda_1 = \cdots = \lambda_k = 0$, they are linearly independent.
- In plain English:
  - Independent $=$ no vector is "redundant" — you need all of them
  - Dependent $=$ at least one vector can be built from the others

Checking Linear Independence

- Use Gaussian elimination to get row echelon form
- All column vectors are linearly independent iff all columns are pivot columns
- **Example.** Check if $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} -1 \\ -2 \\ 1 \\ 1 \end{bmatrix}$ are linearly independent.
- Form matrix and reduce to row echelon form
- Every column is a pivot column $\Rightarrow \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ are linearly independent.

Linear Combinations of Linearly Independent Vectors

- Vector space V with k linearly independent vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k$
- m linear combinations $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$. **Question.** Are they linearly independent?
- If $\mathbf{x}_j = \mathbf{B}\boldsymbol{\lambda}_j$, then $\sum_{j=1}^{m} \psi_j \mathbf{x}_j = \mathbf{B} \sum_{j=1}^{m} \psi_j \boldsymbol{\lambda}_j$
- Key result: $\{\mathbf{x}\}$ linearly independent $\Leftrightarrow$ $\{\boldsymbol{\lambda}\}$ linearly independent

Example: Checking Linear Independence

- $\mathbf{x}_1 = \mathbf{b}_1 - 2\mathbf{b}_2 + \mathbf{b}_3 - \mathbf{b}_4$
- $\mathbf{x}_2 = -4\mathbf{b}_1 - 2\mathbf{b}_2 + 4\mathbf{b}_4$
- $\mathbf{x}_3 = 2\mathbf{b}_1 + 3\mathbf{b}_2 - \mathbf{b}_3 - 3\mathbf{b}_4$
- $\mathbf{x}_4 = 17\mathbf{b}_1 - 10\mathbf{b}_2 + 11\mathbf{b}_3 + \mathbf{b}_4$
- Matrix form and row reduction:

$$\mathbf{A} = \begin{bmatrix} 1 & -4 & 2 & 17 \\ -2 & -2 & 3 & -10 \\ 1 & 0 & -1 & 11 \\ -1 & -4 & -3 & 1 \end{bmatrix} \xrightarrow{\text{row reduce}} \begin{bmatrix} 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & -15 \\ 0 & 0 & 1 & -18 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- The last column is not a pivot column $\Rightarrow \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ are linearly dependent.

Roadmap

(1) Systems of Linear Equations
(2) Matrices
(3) Solving Systems of Linear Equations
(4) Vector Spaces
(5) Linear Independence
(6) Basis and Rank
(7) Linear Mappings
(8) Affine Spaces

Generating Set and Basis

- Think of it as: a basis is the smallest set of building blocks that can construct every vector in the space.
- **Definition.** If every $\mathbf{v} \in V$ can be written as a linear combination of $\mathbf{x}_1, \dots, \mathbf{x}_k$, they form a generating set (they span V): $V = \text{span}[\mathbf{x}_1, \dots, \mathbf{x}_k]$
- **Definition.** A basis B is a generating set with no redundant vectors ($=$ minimal generating set).
- The number of basis vectors is the dimension of V.
- Key properties:
  - B is linearly independent AND spans V
  - Every vector $\mathbf{x} \in V$ has a unique representation using basis vectors

Examples of Bases

- **Example.** Different bases for $\mathbb{R}^3$:

$$B_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}, \quad B_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

$$B_3 = \left\{ \begin{bmatrix} 0.5 \\ 0.8 \\ 0.4 \end{bmatrix}, \begin{bmatrix} 1.8 \\ 0.3 \\ 0.3 \end{bmatrix}, \begin{bmatrix} -2.2 \\ -1.3 \\ 3.5 \end{bmatrix} \right\}$$

- Not a basis (linearly independent, but not maximal in $\mathbb{R}^4$):

$$\mathcal{A} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ -4 \end{bmatrix} \right\}$$

Determining a Basis

- Goal: Find a basis of a subspace $U = \text{span}[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m]$
  (1) Construct a matrix $\mathbf{A} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_m]$
  (2) Find the row-echelon form of $\mathbf{A}$
  (3) Collect the pivot columns
- Logic: Pivot columns tell us which set of vectors is linearly independent.

Rank (1)

- Think of it as: rank $=$ how many "truly different" columns the matrix has.
- **Definition.** The rank of $\mathbf{A} \in \mathbb{R}^{m \times n}$, denoted $\text{rk}(\mathbf{A})$, is the number of linearly independent columns.
  - Same as the number of linearly independent rows
- **Example.** $\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 3 & 5 & 0 \end{bmatrix} \xrightarrow{\text{row reduce}} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$
- 2 pivot rows $\Rightarrow \text{rk}(\mathbf{A}) = 2$ (only 2 of 3 columns are truly independent)

Rank (2)

- The columns of $\mathbf{A}$ span a subspace U with $\dim(U) = \mathrm{rk}(\mathbf{A})$, and a basis of U can be found by Gaussian elimination.
- For all $\mathbf{A} \in \mathbb{R}^{n \times n}$: $\mathrm{rk}(\mathbf{A}) = n$ iff $\mathbf{A}$ is regular (invertible).
- The system $\mathbf{Ax} = \mathbf{b}$ is solvable iff $\mathrm{rk}(\mathbf{A}) = \mathrm{rk}(\mathbf{A}|\mathbf{b})$.
- For $\mathbf{A} \in \mathbb{R}^{m \times n}$, the solution space of $\mathbf{Ax} = \mathbf{0}$ has dimension $n - \mathrm{rk}(\mathbf{A})$.
- Full rank: $\mathrm{rk}(\mathbf{A}) = \min(m, n)$.

Quick Recap: Basis and Rank

A basis is the minimal set of vectors that spans the space. Rank counts the linearly independent columns.

- A basis is a maximally linearly independent spanning set
- Rank determines solvability of $\mathbf{Ax} = \mathbf{b}$
- Null space dimension $= n - \text{rk}(\mathbf{A})$
- Next up: Linear mappings — functions that preserve vector space structure

Roadmap

Linear Mapping (1)

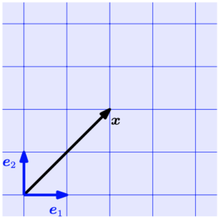
- Think of it as: a function that "plays nicely" with addition and scaling.
- **Definition.** $\Phi : V \to W$ is a linear mapping if:
  - $\Phi(\mathbf{x} + \mathbf{y}) = \Phi(\mathbf{x}) + \Phi(\mathbf{y})$ (adding then mapping = mapping then adding)
  - $\Phi(\lambda \mathbf{x}) = \lambda \Phi(\mathbf{x})$ (scaling then mapping = mapping then scaling)
- **Definition.** A mapping $\Phi : V \to W$ is called:
  - Injective (one-to-one): different inputs $\to$ different outputs (no two $\mathbf{x}$ map to same $\mathbf{w}$)
  - Surjective (onto): every output is hit (every $\mathbf{w} \in W$ is reachable)
  - Bijective: both injective and surjective (perfect one-to-one match)

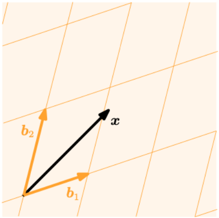
- Properties:
  - $\Phi(\mathbf{0}_V) = \mathbf{0}_W$
  - Φ is injective iff $\ker(\Phi) = \{\mathbf{0}_V\}$
  - Φ is surjective iff $\mathrm{Im}(\Phi) = W$
- **Definition.** Isomorphism: If Φ is bijective, then V and W are isomorphic, denoted $V \cong W$.
- **Theorem.** Two finite-dimensional vector spaces are isomorphic iff they have the same dimension.

- Let $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ be a basis of V.
- Every $\mathbf{x} \in V$ can be uniquely written as:

$$\mathbf{x} = \alpha_1 \mathbf{b}_1 + \cdots + \alpha_n \mathbf{b}_n$$

- $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$ is the coordinate of $\mathbf{x}$ w.r.t. B.
- Basis change $\Rightarrow$ Coordinate change





- Let $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ be a basis of V and $C = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ be a basis of W.
- For a linear mapping $\Phi : V \to W$, the transformation matrix $\mathbf{A}_\Phi$ is defined such that:

$$\Phi(\mathbf{x})_C = \mathbf{A}_\Phi \mathbf{x}_B$$

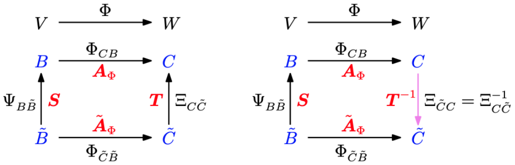
- The columns of $\mathbf{A}_\Phi$ are the coordinates of $\Phi(\mathbf{b}_1), \ldots, \Phi(\mathbf{b}_n)$ with respect to C.

- Two bases $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ and $\tilde{B} = (\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_n)$ of V.
- Change of basis matrix $\mathbf{S}$:

$$\tilde{\mathbf{b}}_j = \sum_{i=1}^{n} s_{ij} \mathbf{b}_i$$

- For a vector: $\mathbf{x}_B = \mathbf{S}\mathbf{x}_{\tilde{B}}$
- For a mapping $\Phi : V \to V$:

$$\mathbf{A}'_\Phi = \mathbf{S}^{-1} \mathbf{A}_\Phi \mathbf{S}$$



Basis Change: General Case

- For $\Phi : V \to W$ with bases $B, \tilde{B}$ of V and $C, \tilde{C}$ of W:
- (inter) transformation matrices $\mathbf{A}_\Phi$ from B to C and $\mathbf{A}'_\Phi$ from $\tilde{B}$ to $\tilde{C}$
- (intra) transformation matrices $\mathbf{S}$ from $\tilde{B}$ to B and T from $\tilde{C}$ to C
- **Theorem.** $\mathbf{A}'_\Phi = T^{-1}\mathbf{A}_\Phi\mathbf{S}$

- Kernel (null space): everything that maps to zero

  $$\ker(\Phi) = \{\mathbf{v} \in V \mid \Phi(\mathbf{v}) = \mathbf{0}_W\}$$

  "What gets destroyed?"
- Image (range): everything that can be reached "What outputs are possible?"
- V: domain, W: codomain

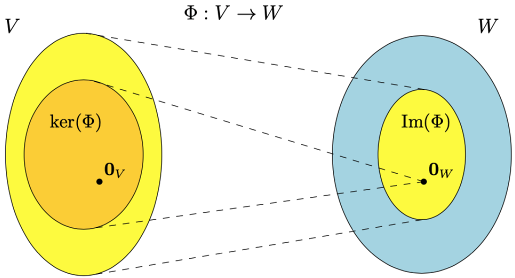


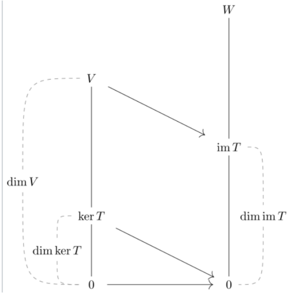
Image and Kernel: Properties

- $\ker(\Phi)$ is a subspace of V
- $\mathrm{Im}(\Phi)$ is a subspace of W
- Φ is injective iff $\ker(\Phi) = \{\mathbf{0}_V\}$
- Φ is surjective iff $\mathrm{Im}(\Phi) = W$
- Rank-Nullity Theorem:

$$\dim(\ker(\Phi)) + \dim(\mathrm{Im}(\Phi)) = \dim(V)$$

- Simplified for $\mathbf{A} \in \mathbb{R}^{m\times n}$:

$$\mathrm{rk}(\mathbf{A}) + \mathrm{nullity}(\mathbf{A}) = n$$

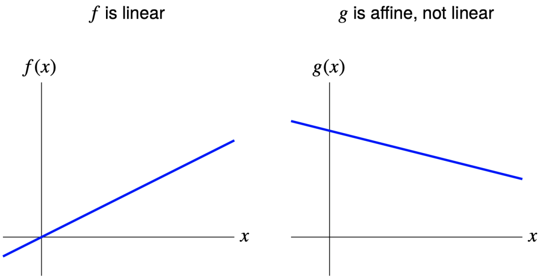


Roadmap

(1) Systems of Linear Equations
(2) Matrices
(3) Solving Systems of Linear Equations
(4) Vector Spaces
(5) Linear Independence
(6) Basis and Rank
(7) Linear Mappings
(8) Affine Spaces

Linear vs. Affine Function

- Key distinction:
  - Linear: $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$
  - Affine: $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$
- In informal usage, people sometimes call affine functions "linear"
- Affine = linear + translation



Affine Subspace

- Think of it as: take a subspace (line/plane through the origin) and slide it somewhere else.
- **Definition.** Let U be a subspace of V and $\mathbf{x}_0 \in V$. The set

$$L = \mathbf{x}_0 + U = \{\mathbf{x}_0 + \mathbf{u} \mid \mathbf{u} \in U\}$$

  is an affine subspace with support point $\mathbf{x}_0$ and direction space U.
- Parametric form:

$$\mathbf{x} = \underbrace{\mathbf{x}_0}_{\text{start here}} + \underbrace{\lambda_1 \mathbf{b}_1 + \cdots + \lambda_k \mathbf{b}_k}_{\text{move in these directions}}$$

An affine subspace is a **shifted** vector subspace — like a line or plane that doesn't pass through the origin.

Examples of Affine Subspaces

- Lines in $\mathbb{R}^2$ or $\mathbb{R}^3$
- Planes in $\mathbb{R}^3$
- Solution sets of linear equations $\mathbf{A}\mathbf{x} = \mathbf{b}$ (for $\mathbf{b} \neq \mathbf{0}$)
- Hyperplanes: $\mathbf{a}^\top \mathbf{x} = c$

Quick Recap: Affine Spaces

Affine subspaces generalize vector subspaces by allowing a shift away from the origin.

- Linear: $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$ vs. Affine: $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$
- Solution sets of $\mathbf{A}\mathbf{x} = \mathbf{b}$ are affine subspaces
- Parametric form: particular solution + direction space
- That completes our tour of linear algebra!

Concepts at a Glance

Concept	What It Means	Why It Matters
Vector Space	Set closed under $+$ and $\cdot$	Foundation of linear algebra
Basis	Minimal spanning set	Defines coordinates
Rank	$\#$ linearly indep. columns	Determines solvability
Kernel	Null space of $\mathbf{A}$	Solution space of $\mathbf{A}\mathbf{x} = \mathbf{0}$
Image	Range of mapping	What outputs are reachable
Linear Mapping	Structure-preserving map	Transforms between spaces
Affine Subspace	Shifted subspace	Solution set of $\mathbf{A}\mathbf{x} = \mathbf{b}$

Pattern: Every concept connects back to solving $\mathbf{A}\mathbf{x} = \mathbf{b}$ — the central problem of linear algebra.

Common Mistakes to Avoid

(1) Matrix multiplication is NOT commutative ($\mathbf{AB} \neq \mathbf{BA}$ in general)
Always check dimensions and order carefully

(2) $(\mathbf{A} + \mathbf{B})^{-1} \neq \mathbf{A}^{-1} + \mathbf{B}^{-1}$
Inverse does not distribute over addition

(3) Linear independence $\neq$ orthogonality
Vectors can be linearly independent without being perpendicular

(4) Rank tells you about solvability, not the solution itself
Use rank to determine IF solutions exist, then solve

(5) Affine $\neq$ Linear
$f(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$ is affine (has translation), not linear

Key Takeaways

(1) Linear algebra studies vectors and linear mappings between vector spaces

(2) Matrices are the computational tool for representing linear systems

(3) Gaussian elimination $\rightarrow$ row echelon form $\rightarrow$ rank, basis, solutions

(4) Every vector space has a basis — a minimal spanning set

(5) Linear mappings preserve structure; change of basis transforms coordinates

(6) Rank-Nullity Theorem: $\dim(\ker) + \dim(\mathrm{Im}) = \dim(V)$

Vectors $\rightarrow$ Linear Independence $\rightarrow$ Basis $\rightarrow$ Rank $\rightarrow$ Linear Mappings $\rightarrow$ Transformations

Review Question 1: Systems of Linear Equations

Question. A system $\mathbf{Ax} = \mathbf{b}$ can have three types of solutions.

(a) What are the three types?

(b) Under what conditions on $\mathbf{A}$ and $\mathbf{b}$ does each type occur?

(c) If $\mathbf{A}$ is a 3×3 matrix with $\text{rk}(\mathbf{A}) = 2$, how many free variables are there?

Hint: Compare $\text{rk}(\mathbf{A})$ with $\text{rk}(\mathbf{A}|\mathbf{b})$ and with n (number of unknowns). Free variables $= n - \text{rk}(\mathbf{A})$.

Review Question 2: Matrix Operations

Question. Consider matrices $\mathbf{A} \in \mathbb{R}^{2\times 3}$ and $\mathbf{B} \in \mathbb{R}^{3\times 4}$.

(a) What are the dimensions of $\mathbf{AB}$?

(b) Can you compute $\mathbf{BA}$? Why or why not?

(c) Is matrix multiplication commutative? Give a counterexample.

Hint: For $\mathbf{AB}$ to work, # columns of $\mathbf{A}$ must equal # rows of $\mathbf{B}$. Result has rows of $\mathbf{A}$, columns of $\mathbf{B}$.

Question. Let $\mathbf{A}$ be an invertible $n \times n$ matrix.

(a) What is the inverse of $\mathbf{AB}$?

(b) Is $(\mathbf{A}+\mathbf{B})^{-1} = \mathbf{A}^{-1} + \mathbf{B}^{-1}$? Why or why not?

(c) What is the transpose of $\mathbf{AB}$?

Hint: For products, the inverse reverses order: $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$. Same idea for transpose: $(\mathbf{AB})^{\top} = \mathbf{B}^{\top}\mathbf{A}^{\top}$.

Review Question 4: Gaussian Elimination

Question. Given the augmented matrix $[\mathbf{A}|\mathbf{b}]$:

(a) What are the three allowed row operations?

(b) What does the row echelon form tell you about the system?

(c) How do you determine the number of solutions from the RREF?

Hint: The three row operations are: swap rows, multiply a row by a nonzero scalar, add a multiple of one row to another. Count pivots in the RREF — each pivot = one determined variable.

Review Question 5: Vector Spaces and Subspaces

Question. Determine which of the following are subspaces of $\mathbb{R}^3$:

(a) $U_1 = \{(x, y, z) : x + y + z = 0\}$

(b) $U_2 = \{(x, y, z) : x + y + z = 1\}$

(c) $U_3 = \{(x, y, z) : x^2 + y^2 = 0\}$

(d) For each, explain why or why not (check the three conditions).

Hint: Always check the zero vector first — it's the fastest way to rule out a set. Does $(0, 0, 0)$ satisfy the condition?

Review Question 6: Linear Independence

Question. Consider vectors in $\mathbb{R}^3$.

(a) How do you check if vectors are linearly independent?

(b) What is the geometric meaning of linear dependence in $\mathbb{R}^3$?

(c) Can 4 vectors in $\mathbb{R}^3$ be linearly independent? Why or why not?

Hint: Put vectors as columns, row reduce. If every column has a pivot $\rightarrow$ independent. In $\mathbb{R}^3$, at most 3 vectors can be independent (why?).

Review Question 7: Basis and Dimension

Question. Consider a subspace $U = \text{span}[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]$.

(a) How do you find a basis for U?

(b) Is the basis unique? If not, what is unique about it?

(c) If $\dim(U) = 2$, what does that mean geometrically in $\mathbb{R}^3$?

Hint: To find a basis, row reduce the matrix of spanning vectors. Pivot columns of the *original* matrix form a basis. The basis is not unique, but the dimension (number of basis vectors) always is.

Review Question 8: Rank

Question. Let $\mathbf{A} \in \mathbb{R}^{3\times 4}$.

(a) What is the maximum possible rank of $\mathbf{A}$?

(b) If $\mathrm{rk}(\mathbf{A}) = 2$, what is the dimension of the null space?

(c) How does the rank relate to the solvability of $\mathbf{Ax} = \mathbf{b}$?

Hint: Max rank $= \min(\text{rows}, \text{cols})$. Null space dimension $= n - \mathrm{rk}(\mathbf{A})$ where $n =$ number of columns.

Review Question 9: Linear Mappings

Question. Let $\Phi : V \to W$ be a linear mapping.

(a) What two properties must Φ satisfy to be linear?

(b) What is the kernel of Φ? What does it tell you about injectivity?

(c) State the Rank-Nullity Theorem and explain what it means.

Hint: The kernel is everything that maps to $\mathbf{0}$. If only $\mathbf{0}$ maps to $\mathbf{0}$, the mapping is injective (one-to-one).

Review Question 10: Affine Spaces

Question. Consider the equation $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{b} \neq \mathbf{0}$.

(a) Is the solution set a subspace? Why or why not?

(b) What kind of geometric object is the solution set?

(c) Write the solution set in parametric form using a particular solution and the null space.

Hint: The solution set is NOT a subspace (it doesn't contain $\mathbf{0}$). It's an affine subspace: $\{\mathbf{x}_p + \mathbf{v} : \mathbf{v} \in \ker(\mathbf{A})\}$ where $\mathbf{x}_p$ is any particular solution.

Theory Connection: Linear Algebra in Machine Learning

- Feature matrices: Each row is a data point, each column is a feature $\rightarrow \mathbf{X} \in \mathbb{R}^{N \times D}$
- Weight matrices: Model parameters transform inputs: $\hat{\mathbf{y}} = \mathbf{X}$
- Least squares: The optimal solves $\mathbf{X}^\top \mathbf{X} = \mathbf{X}^\top \mathbf{y}$ (normal equation)
- Regularization: Ridge regression adds λI to make $\mathbf{X}^\top \mathbf{X} + \lambda I$ invertible

ML models learn by finding the right matrix transformations from data to predictions.

Theory Connection: Linear Algebra in Computer Vision

- Images as matrices: A grayscale image is a matrix $\mathbf{A} \in \mathbb{R}^{H \times W}$
- Geometric transformations: Rotation, scaling, shearing are all matrix multiplications
- Convolution filters: Each filter is a small matrix applied to image patches
- Homogeneous coordinates: Affine transformations become linear in higher dimensions

Every image filter, transformation, and detection algorithm relies on matrix operations.

Theory Connection: Linear Algebra in Deep Learning

- Each layer: Computes $\mathbf{x} \mapsto \sigma(\mathbf{x} + \mathbf{b})$ — an affine map + nonlinearity
- Forward pass: Chain of matrix multiplications through layers
- Backpropagation: Gradient computation uses the chain rule on matrix products
- Attention mechanism: $\text{Attention}(Q, K, V) = \text{softmax}(\frac{QK^T}{\sqrt{d}})V$ — all matrices!

Deep learning is essentially repeated linear algebra with nonlinear activations.

Theory Connection: Linear Algebra in Optimization

- Gradient descent: Update rule $\mathbf{x}_{t+1} = \mathbf{x}_t - \eta \nabla f(\mathbf{x}_t)$ — vector operations
- Hessian matrix: Second-order information $H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$ guides convergence
- Quadratic forms: $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x} + c$ — eigenvalues determine curvature
- Positive definiteness: $\mathbf{A} \succ 0$ ensures the function has a unique minimum

Optimization algorithms are guided by the linear algebra of gradients and curvature.

Theory Connection: Linear Algebra in Data Science

- PCA: Find directions of maximum variance $\to$ eigenvectors of $\mathbf{X}^\top\mathbf{X}$
- SVD: $\mathbf{A} = U\Sigma V^\top$ decomposes any matrix into rotations and scaling
- Dimensionality reduction: Project data onto top k principal components
- Recommender systems: Matrix factorization $\mathbf{A} \approx H^\top$ discovers latent features

Data science uses matrix decompositions to find structure hidden in high-dimensional data.

PhD View: Abstract Vector Spaces and Functional Analysis

- Infinite-dimensional spaces: Function spaces like $L^2(\mathbb{R})$ are vector spaces with inner product $\langle f, g \rangle = \int f(x)g(x)\,dx$
- Hilbert spaces: Complete inner product spaces — the natural setting for quantum mechanics and kernel methods
- Banach spaces: Complete normed spaces without necessarily having an inner product
- Reproducing Kernel Hilbert Spaces (RKHS): Foundation of kernel methods in ML
- Connection to ML: Kernel trick maps finite-dimensional data into infinite-dimensional feature spaces

PhD View: Matrix Decompositions

- SVD: $\mathbf{A} = U\Sigma V^\top$ — the most important decomposition; works for any matrix
- Eigendecomposition: $\mathbf{A} = Q\Lambda Q^{-1}$ — only for square diagonalizable matrices
- QR decomposition: $\mathbf{A} = QR$ — orthogonal $\times$ upper triangular; used in least squares
- LU decomposition: $\mathbf{A} = LU$ — lower $\times$ upper triangular; efficient for solving linear systems
- Cholesky: $\mathbf{A} = LL^\top$ — for symmetric positive definite matrices; fastest decomposition
- Applications: Recommender systems (SVD), PageRank (eigendecomposition), regression (QR)

PhD View: Tensor Algebra

- Tensors: Generalization of vectors (order 1) and matrices (order 2) to higher-order arrays
- Tensor products: $\mathbf{x} \otimes \mathbf{y}$ creates an order-2 tensor from two vectors
- Multilinear maps: Maps that are linear in each argument separately
- Tensor decompositions: CP decomposition, Tucker decomposition — generalizations of SVD
- In deep learning: Weight tensors in CNNs, attention tensors in transformers
- Einstein notation: $C_{ijk} = A_{il} B_{ljk}$ — compact notation for multilinear operations

PhD View: Numerical Linear Algebra

- Condition number: $\kappa(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$ — measures sensitivity to perturbations
- Numerical stability: Small changes in input $\to$ small changes in output (not always!)
- Iterative solvers: Conjugate gradient, GMRES — for large sparse systems where direct methods fail
- Randomized algorithms: Randomized SVD can approximate large matrices efficiently
- GPU computing: Matrix multiplications are embarrassingly parallel — key to deep learning speed
- Floating point: $\mathbf{A} + \epsilon \mathbf{I}$ regularization prevents numerical issues in practice

PhD View: Spectral Theory

- Spectral theorem: Symmetric matrices have real eigenvalues and orthogonal eigenvectors
- Spectral decomposition: $\mathbf{A} = \sum_i \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$ — matrix as sum of rank-1 components
- Graph Laplacian: $L = D - \mathbf{A}$ where D is degree matrix — eigenvalues reveal graph structure
- Spectral clustering: Use eigenvectors of graph Laplacian to cluster data
- PageRank: Finding the dominant eigenvector of the web's transition matrix
- Spectral gap: Difference between largest eigenvalues determines mixing time and connectivity

Math Tutorial: Introduction

- This tutorial section provides deeper explanations and worked examples
- Each topic from the main lecture is explored in greater detail
- Includes step-by-step solutions and practical applications
- Topics covered:
  1. Systems of Linear Equations — Detailed Examples
  2. Matrix Operations — Computational Methods
  3. Vector Spaces — Theoretical Foundations
  4. Linear Independence — Practical Techniques
  5. Basis and Rank — Advanced Concepts
  6. Linear Mappings — Transformation Matrices
  7. Affine Spaces — Geometric Interpretation

Problem: Solve the system:

$$\begin{cases} 2x + 3y - z = 8 \\ x - y + 2z = 1 \\ 3x + 2y + z = 7 \end{cases}$$

Solution: Write the augmented matrix:

$$\left[\begin{array}{ccc|c} 2 & 3 & -1 & 8 \\ 1 & -1 & 2 & 1 \\ 3 & 2 & 1 & 7 \end{array}\right] \xrightarrow{\text{row reduce}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

Answer: $x = 2,\ y = 1,\ z = -1$ (unique solution)

Tutorial: Systems of Linear Equations — Example 2

Problem: Determine the nature of solutions:

$$\begin{cases} x + 2y - z = 3 \\ 2x + 4y - 2z = 6 \\ 3x + 6y - 3z = 9 \end{cases}$$

Analysis: Equations 2 and 3 are multiples of equation 1.

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 2 & 4 & -2 & 6 \\ 3 & 6 & -3 & 9 \end{array}\right] \xrightarrow{\text{row reduce}} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Result: Infinitely many solutions: $\mathbf{x} = \begin{bmatrix} 3 - 2y + z \\ y \\ z \end{bmatrix}, \quad y, z \in \mathbb{R}$

Problem: Given $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}$, compute $\mathbf{AB}$, $\mathbf{BA}$, and $\mathbf{A}^{-1}$.

Solution:

$$\mathbf{AB} = \begin{bmatrix} 4 & 6 \\ 10 & 12 \end{bmatrix}, \quad \mathbf{BA} = \begin{bmatrix} 2 & 4 \\ 10 & 14 \end{bmatrix}$$

Note: $\mathbf{AB} \neq \mathbf{BA}$ (matrix multiplication is not commutative)

For $\mathbf{A}^{-1}$: $\det(\mathbf{A}) = 1(4) - 2(3) = -2$

$$\mathbf{A}^{-1} = \frac{1}{-2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}$$

Tutorial: Vector Spaces — Subspace Verification

Problem: Verify that $U = \{(x, y, z) \in \mathbb{R}^3 : x + y - z = 0\}$ is a subspace of $\mathbb{R}^3$.
Solution: Check three conditions:

(1) Zero vector: $(0, 0, 0)$ satisfies $0 + 0 - 0 = 0$ ✓

(2) Closure under addition: If $\mathbf{u}, \mathbf{v}$ satisfy the equation:

$$(u_1 + v_1) + (u_2 + v_2) - (u_3 + v_3) = (u_1 + u_2 - u_3) + (v_1 + v_2 - v_3) = 0 \quad \checkmark$$

(3) Closure under scalar multiplication: If $\mathbf{u}$ satisfies and $\lambda \in \mathbb{R}$:

$$(\lambda u_1) + (\lambda u_2) - (\lambda u_3) = \lambda(u_1 + u_2 - u_3) = 0 \quad \checkmark$$

U is a subspace of $\mathbb{R}^3$

Tutorial: Linear Independence — Detailed Example

Problem: Determine if $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}$ are linearly independent.

Solution: Form matrix (vectors as columns) and row reduce:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 1 & 0 & 1 \end{bmatrix} \xrightarrow{\text{row reduce}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Analysis: Column 3 is not a pivot column $\Rightarrow$ vectors are linearly dependent.
From the RREF: $x_1 + x_3 = 0$ and $x_2 + x_3 = 0$, so $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2$.

Verification: $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix} = \mathbf{v}_3$ ✓

Tutorial: Basis and Dimension — Finding a Basis

Problem: Find a basis for $U = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \right\}$

Solution: Form matrix and reduce:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow{\text{row reduce}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Result: Pivot columns are 1 and 2, so a basis is:

$$B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$$\dim(U) = 2$$

Tutorial: Rank Calculation — Detailed Example

Problem: Find the rank of $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

Solution: Row reduce to echelon form:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 1 & 1 & 1 & 1 \end{bmatrix} \xrightarrow[R_3 - R_1]{R_2 - 2R_1} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & -2 & -3 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Number of pivot columns $= 2$, so $\text{rk}(\mathbf{A}) = 2$

Problem: Find the transformation matrix for $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $\Phi(x, y) = (2x + y, x - y)$ with respect to the standard basis.

Solution: Apply Φ to basis vectors:

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

The transformation matrix is:

$$\mathbf{A}_\Phi = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$$

Verification: $\begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + y \\ x - y \end{bmatrix}$ ✓

Tutorial: Kernel and Image — Detailed Example

Problem: For $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}$, find $\ker(\mathbf{A})$ and $\operatorname{Im}(\mathbf{A})$.

Kernel: Solve $\mathbf{A}\mathbf{x} = \mathbf{0}$. Row reduce: $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix}$

General solution: $x_1 = -2x_2 - 3x_3$

$$\ker(\mathbf{A}) = \operatorname{span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Image: Pivot columns of $\mathbf{A}$ form a basis: $\operatorname{Im}(\mathbf{A}) = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$

Rank-Nullity check: $\dim(\ker(\mathbf{A})) + \dim(\operatorname{Im}(\mathbf{A})) = 2 + 1 = 3 = n$ ✓

Tutorial: Basis Change — Detailed Example

Problem: Given bases $B = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\tilde{B} = \{\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2\}$ where:

$$\tilde{\mathbf{b}}_1 = \mathbf{b}_1 + \mathbf{b}_2, \quad \tilde{\mathbf{b}}_2 = \mathbf{b}_2$$

Find the change of basis matrix $\mathbf{S}$.

Solution: Express new basis in terms of old:

$$\mathbf{S} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

For a vector $\mathbf{x}$: $\mathbf{x}_B = \mathbf{S}\mathbf{x}_{\tilde{B}}$

Example. If $\mathbf{x}_{\tilde{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, then: $\mathbf{x}_B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

Tutorial: Affine Spaces — Geometric Example

Problem: Find the parametric equation of the line passing through $\mathbf{p} = (1, 2, 3)$ in direction $\mathbf{d} = (2, -1, 1)$.
Solution: The line is an affine subspace:

$$L = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + t \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}$$

Parametric form: $x = 1 + 2t, \quad y = 2 - t, \quad z = 3 + t$
Example. Points on the line:

- $t = 0$: $(1, 2, 3)$ (support point)
- $t = 1$: $(3, 1, 4)$
- $t = -1$: $(-1, 3, 2)$

Definition. For a square matrix $\mathbf{A}$, a non-zero vector $\mathbf{v}$ is an eigenvector with eigenvalue λ if:

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

Example. Find eigenvalues of $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$

Solution: Solve $\det(\mathbf{A} - \lambda I) = 0$: $(3-\lambda)(2-\lambda) = 0 \Rightarrow \lambda_1 = 3, \; \lambda_2 = 2$

For $\lambda_1 = 3$: $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ For $\lambda_2 = 2$: $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Tutorial: Summary and Key Takeaways

- Systems of Equations: Gaussian elimination is the fundamental tool
- Matrices: Represent linear transformations and systems
- Vector Spaces: Abstract structures with specific properties
- Linear Independence: Key concept for basis and dimension
- Rank: Determines solvability and solution space dimension
- Linear Mappings: Preserve vector space structure
- Kernel and Image: Fundamental subspaces of a transformation
- Basis Change: Allows computation in different coordinate systems
- Affine Spaces: Generalization of linear subspaces

Tutorial: Practice Problems

(1) Solve the system: $\begin{cases} x + y + z = 6 \\ 2x - y + z = 3 \\ x + y - z = 0 \end{cases}$

(2) Find the rank of $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

(3) Determine if $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ are linearly independent.

(4) Find the kernel and image of $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

(5) Find the transformation matrix for $\Phi(x, y) = (x + 2y, 3x - y)$

Thank You!

Questions?