Math Review: Linear Algebra

Jianan Yang

Peking University

February 27, 2025

Overview

Linear (Vector) Spaces

In mathematics, "linear" means that a property is preserved by addition and multiplication by a constant.

A *linear space* is a set X for which $x + y$ (addition) and αx (multiplication by α) are defined, where $x, y \in X$ and $\alpha \in \mathbb{R}$. N-vector on the *Euclidean space* $\mathbb{R}^N$:

$$x = (x_1, \dots, x_N).$$

Here addition and multiplication by a constant are defined component-wise:

$$(x_1, \dots, x_N) + (y_1, \dots, y_N) := (x_1 + y_1, \dots, x_N + y_N)$$

$$\alpha(x_1, \dots, x_N) := (\alpha x_1, \dots, \alpha x_N).$$

(The symbol ":=" means that we define the left-hand side by the right-hand side.)

Linear Functions

A *linear function* is a function that preserves linearity. Thus $f\colon \mathbb{R}^N \to \mathbb{R}$ is linear if

$$f(x+y) = f(x) + f(y),$$

$$f(\alpha x) = \alpha f(x).$$

An obvious example of a linear function $f\colon \mathbb{R}^N \to \mathbb{R}$ is

Linear Functions

A *linear function* is a function that preserves linearity. Thus $f \colon \mathbb{R}^N \to \mathbb{R}$ is linear if

$$f(x+y) = f(x) + f(y),$$

$$f(\alpha x) = \alpha f(x).$$

An obvious example of a linear function $f \colon \mathbb{R}^N \to \mathbb{R}$ is

$$f(x) = a_1 x_1 + \cdots + a_N x_N = \sum_{n=1}^{N} a_n x_n,$$

where $a_1, \ldots, a_N$ are numbers. In fact we can show that all linear functions are of this form.

Proof:...

Inner Product and Euclidean Norm

Let $x = (x_1, \ldots, x_N)$ and $y = (y_1, \ldots, y_N)$ be two vectors. Then

$$\langle x, y \rangle := x_1 y_1 + \cdots + x_N y_N = \sum_{n=1}^{N} x_n y_n$$

is called the *inner product* (also *vector product*) of x and y. The (*Euclidean*) *norm* of x is defined by

$$\|x\| := \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + \cdots + x_N^2}.$$

The Euclidean norm is also called the L^2 norm.

Linearity of Inner Product

Fixing x, the inner product $\langle x, y \rangle$ is linear in y, so we have

Linearity of Inner Product

Fixing x, the inner product $\langle x, y \rangle$ is linear in y, so we have

$$\langle x, y_1 + y_2 \rangle = \langle x, y_1 \rangle + \langle x, y_2 \rangle,$$

$$\langle x, \alpha y \rangle = \alpha \langle x, y \rangle.$$

The same holds for x as well, fixing y. So the inner product is a *bilinear* function of x and y.

Geometric Interpretation of Inner Product

In $\mathbb{R}^2$, for vectors $x = (x_1, x_2)$ and $y = (y_1, y_2)$:

Geometric Interpretation of Inner Product

In $\mathbb{R}^2$, for vectors $x = (x_1, x_2)$ and $y = (y_1, y_2)$:

$$\langle x, y\rangle = x_1 y_1 + x_2 y_2 = \|x\| \|y\| \cos\theta$$

where θ is the angle between vectors. The value of $\cos\theta$ determines:

$$\cos\theta \begin{cases} > 0, & \text{(acute angle)} \\ = 0, & \text{(right angle)} \\ < 0, & \text{(obtuse angle)} \end{cases}$$

Geometric Interpretation of Inner Product

In $\mathbb{R}^2$, for vectors $x = (x_1, x_2)$ and $y = (y_1, y_2)$:

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 = \|x\| \|y\| \cos \theta$$

where θ is the angle between vectors. The value of $\cos \theta$ determines:

$$\cos \theta \begin{cases} > 0, & \text{(acute angle)} \\ = 0, & \text{(right angle)} \\ < 0, & \text{(obtuse angle)} \end{cases}$$

Two vectors are:

- *orthogonal* if $\langle x, y \rangle = 0$
- form an acute angle if $\langle x, y \rangle > 0$
- form an obtuse angle if $\langle x, y \rangle < 0$

This geometric intuition extends to $\mathbb{R}^N$: vectors $x, y \in \mathbb{R}^N$ are orthogonal if $\langle x, y \rangle = 0$.

Properties of Inner Product and Norm

Cauchy-Schwarz Inequality:

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

Properties of the Norm:

1. Non-negativity: $\|x\| \geq 0$, with equality if and only if $x = 0$
2. Homogeneity: $\|\alpha x\| = |\alpha|\|x\|$
3. Triangle Inequality: $\|x + y\| \leq \|x\| + \|y\|$

Matrix: Linear Maps $\mathbb{R}^N \to \mathbb{R}^M$

A *linear map* $f : \mathbb{R}^N \to \mathbb{R}^M$ associates to each $x \in \mathbb{R}^N$ a vector $f(x) \in \mathbb{R}^M$ and preserves:

- Addition: $f(x + y) = f(x) + f(y)$
- Scalar multiplication: $f(\alpha x) = \alpha f(x)$

For $f(x) = (f_1(x), \ldots, f_M(x))$, each component $f_m(x)$ is a linear function:

Matrix: Linear Maps $\mathbb{R}^N \to \mathbb{R}^M$

A *linear map* $f\colon \mathbb{R}^N \to \mathbb{R}^M$ associates to each $x \in \mathbb{R}^N$ a vector $f(x) \in \mathbb{R}^M$ and preserves:

- Addition: $f(x+y) = f(x) + f(y)$
- Scalar multiplication: $f(\alpha x) = \alpha f(x)$

For $f(x) = (f_1(x), \ldots, f_M(x))$, each component $f_m(x)$ is a linear function:

$$f_m(x) = a_{m1}x_1 + \cdots + a_{mN}x_N$$

where $a_{m1}, \ldots, a_{mN}$ are constants. This represents the m-th element of the output vector.

Matrix Representation of Linear Maps

A linear map $f\colon \mathbb{R}^N \to \mathbb{R}^M$ corresponds to an array of numbers $\{a_{mn}\}$ where $1 \leq m \leq M$ and $1 \leq n \leq N$, written as a matrix:

$$A = (a_{mn}) = \begin{bmatrix} a_{11} & \cdots & a_{1n} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} & \cdots & a_{mN} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{Mn} & \cdots & a_{MN} \end{bmatrix}$$

For an $M \times N$ matrix A and an N-vector x, the m-th element of Ax is:

$$(Ax)_m = a_{m1}x_1 + \cdots + a_{mN}x_N$$

Thus, $f(x) = Ax$ represents a linear map $\mathbb{R}^N \to \mathbb{R}^M$, and the set of all $M \times N$ matrices can be identified as $\mathbb{R}^{MN}$.

Composition of Linear Maps

Consider two linear maps with their matrix representations:

- $f : \mathbb{R}^N \to \mathbb{R}^M$ with matrix $A = (a_{mn})$, so $f(x) = Ax$
- $g : \mathbb{R}^M \to \mathbb{R}^L$ with matrix $B = (b_{lm})$, so $g(y) = By$

Their composition $h = g \circ f$ where $h(x) := g(f(x))$ is a linear map:

$$\mathbb{R}^N \xrightarrow{f} \mathbb{R}^M \xrightarrow{g} \mathbb{R}^L$$

Since $h(x) = g(f(x)) = B(Ax)$, the matrix $C = (c_{ln})$ representing h has elements:

$$c_{ln} = \sum_{m=1}^{M} b_{lm} a_{mn}$$

Therefore, $C = BA$ represents the composition of the maps.

Matrix Operations

Standard Matrix Algebra Rules:

- Matrix multiplication: $C = BA$ (defined by previous composition rule)
- Distributive property: $B(A_1 + A_2) = BA_1 + BA_2$
- Associative property: $A(BC) = (AB)C$

Square Matrices and the Identity Matrix

Square Matrix: An $M \times N$ matrix with $M = N$

Identity Matrix I:

- Corresponds to identity map $\mathrm{id} : \mathbb{R}^N \to \mathbb{R}^N$ where $\mathrm{id}(x) = x$
- Properties:
  - Square matrix with diagonal elements $= 1$
  - All non-diagonal elements $= 0$
  - For any square matrix A: $AI = IA = A$

Bijective Maps

For a map $f : \mathbb{R}^N \to \mathbb{R}^N$:

- *One-to-one* (injective):
- *Onto* (surjective):
- *Bijective*: both injective and surjective

Inverse Function:

- For bijective f, each y has unique x with $y = f(x)$
- Write $x = f^{-1}(y)$, called the inverse of f

Matrix Inverses

Matrix Inverse:

- For bijective linear map $f(x) = Ax$, inverse f^{-1} is also linear
- Matrix of f^{-1} is denoted A^{-1} (inverse of A)
- Key property: $AA^{-1} = A^{-1}A = I$

Uniqueness Proof: Suppose B, C are both inverses of A. Then:

Therefore, if an inverse exists, it must be unique.
A matrix that has an inverse is called *regular, nonsingular, invertible,* etc.

Matrix Determinant: Definition

Base Cases:

- For 1×1 matrix $A = (a)$:
- For 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$:

General Definition (Inductive):
For $N \times N$ matrix $A = (a_{mn})$, using any row or column:

$$\det A = \sum_{n=1}^{N} (-1)^{m+n} a_{mn} M_{mn} = \sum_{m=1}^{N} (-1)^{m+n} a_{mn} M_{mn}$$

where M_{mn} is the determinant of the $(N-1) \times (N-1)$ matrix obtained by removing row m and column n of A.

This definition is consistent (independent of choice of m, n).

Properties of Determinants

1. **Regularity and Inverse:**
   - A is regular (invertible) $\iff \det A \neq 0$
   - If invertible: $A^{-1} = \frac{1}{\det A}\tilde{A}$, where $\tilde{a}_{mn} = (-1)^{m+n} M_{mn}$
   - Example for 2×2:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \implies$$

Properties of Determinants (Cont'd)

2. **Multiplicativity:**

$$\det(AB) = (\det A)(\det B) \text{ for square matrices of same order}$$

3. **Block Triangular Matrices:**

$$\det \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix} = (\det A_{11})(\det A_{22})$$

where A_{11} and A_{22} are square matrices.

Row and Column Vector Notation

Row Vector ($1 \times N$ matrix):

$$x = (x_1, \dots, x_N)$$

Column Vector ($N \times 1$ matrix):

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

Matrix Dimensions:

- An N-row vector is a $1 \times N$ matrix
- An N-column vector is an $N \times 1$ matrix

Matrix Multiplication and Linear Maps

For linear map $f(x) = Ax$ where A is $M \times N$ and x is $N \times 1$:

$$Ax = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1N}x_N \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mN}x_N \\ \vdots \\ a_{M1}x_1 + \cdots + a_{MN}x_N \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} & \cdots & a_{mN} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{1n} & & a_{1N} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ \vdots \\ x_N \end{bmatrix}$$

The notation shows equivalence between:

- Linear map evaluation (Ax on left)
- Matrix multiplication (Ax on right)

Vector and Matrix Transposition

Vector Notation:

- Default: Vectors are column vectors unless specified otherwise
- Row vector: $(x_1, \ldots, x_N)$
- Column vector notation options: $(x_1, \ldots, x_N)'$ (using transpose)

Matrix Transpose:

- For $M \times N$ matrix $A = (a_{mn})$, its transpose A' is $N \times M$
- Elements: $(A')_{nm} = a_{mn}$
- Geometric interpretation: "Flip matrix diagonally"

Orthogonal Matrices

Definition: A square matrix P is *orthogonal* if:

$$P'P = PP' = I$$

Properties:

- Column vectors of P are orthogonal and have norm 1
- For orthogonal matrices: $P^{-1} = P'$

Verification: Writing out the elements of $P'P$ shows that:

Symmetric Matrices and Definiteness

Symmetric Matrix:

- $A = A'$ (elements symmetric about diagonal)

Types of Definiteness:

- *Positive Semidefinite*:
- *Positive Definite*:
  -
  -

Partial Order: For symmetric matrices A and B:

$$A \geq B \iff$$

Principal Minors and Positive Definiteness

Principal Minor:

- k-th principal minor $=$ determinant of submatrix formed by first k rows and columns
- For $A = (a_{mn})$ of size $N \times N$:
  - 1st principal minor: a_{11}
  - 2nd principal minor: $a_{11}a_{22} - a_{12}a_{21}$
  - N-th principal minor: $\det A$

Proposition: *(Test for Positive Definiteness)*

- For real symmetric matrix A:

$$A \text{ is positive definite } \iff$$

Proof:

Eigenvalues and Eigenvectors

Definition:
If A is a square matrix and there exist a number α and a nonzero vector v such that $Av = \alpha v$, then:

- v is an **eigenvector** of A
- α is the corresponding **eigenvalue**

Key Properties

$$Av = \alpha v \Leftrightarrow (\alpha I - A)v = 0$$

α is an eigenvalue of A if and only if $\det(\alpha I - A) = 0$

$\Phi_A(x) := \det(xI - A)$ is called the **characteristic polynomial** of A.

Complex Inner Products

- Even with real matrices, eigenvalues and eigenvectors can be complex
- **Complex inner product**: For vectors $x, y \in \mathbb{C}^N$:

$$\langle x, y \rangle =$$

where $\bar{x}$ is the complex conjugate and x^* is the adjoint (conjugate transpose)

Properties of Complex Inner Products

- $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (conjugate symmetry)
- For matrix A, its adjoint A^* is the conjugate transpose
- $\langle x, Ay \rangle = \langle A^* x, y \rangle$ because...

Hermitian Matrices

- **Hermitian matrices**: $A^* = A$
  - For real matrices, Hermitian $=$ symmetric
  - For Hermitian A, the quadratic form $\langle x, Ax\rangle$ is real:
- **Proposition:**
  - The eigenvalues of a Hermite matrix are real
- Since real symmetric matrices are Hermite, the eigenvalues of real symmetric matrices are all real (and so are eigenvectors).

Unitary Matrices and Basis Representation

Definition:
If U is a square matrix such that $U^* U = UU^* = I$, then U is called *unitary*. Real unitary matrices are orthogonal, by definition.

Alternative Basis Representation:
We usually take the standard basis $\{e_1, \ldots, e_N\}$ in $\mathbb{R}^N$. Suppose we take vectors $\{p_1, \ldots, p_N\}$, where the matrix $P = [p_1, \ldots, p_N]$ is regular. Let x be any vector and $y = P^{-1}x$. Then

$$x = PP^{-1}x = Py = y_1 p_1 + \cdots + y_N p_N,$$

The elements of y can be interpreted as the coordinates of x when we use the basis P

Matrix Representation

Matrix Representation in Alternative Basis:
For a basis P, a matrix A transforms to $B = P^{-1}AP$ when representing the linear map $x \mapsto Ax$.

- The linear map $x \mapsto Ax$ becomes $P^{-1}x \mapsto P^{-1}Ax = (P^{-1}AP)(P^{-1}x)$
- Finding P such that $P^{-1}AP$ has a simple form (e.g., diagonal) is often useful

Diagonalization

Theorem (Diagonalization of Symmetric Matrices):
Let A be real symmetric. Then there exists a real orthogonal matrix P such that

$$P^{-1}AP = P^{T}AP = \mathrm{diag}[\alpha_1, \ldots, \alpha_N],$$

where $\alpha_1, \ldots, \alpha_N$ are eigenvalues of A.
Similarly, Hermite matrices can be diagonalized by unitary matrices.

The End