Chapter 2 Vector Spaces and Matrices

1. Vector and Matrix
2. Rank of Matrix
3. Linear Equations
4. Inner Product and Norm

Notation

- set: If $\mathcal{X}$ is a set and $\boldsymbol{x}$ is an element of $\mathcal{X}$, we write $\boldsymbol{x} \in \mathcal{X}$.
- max/min: maximum/minimum point of $\mathcal{X}$.
- sup/inf: supremum/infimum of $\mathcal{X}$. e.g., $\inf\limits_{\boldsymbol{x}} \exp(\boldsymbol{x}) = 0$, $\sup\limits_{\boldsymbol{x}} \arctan(\boldsymbol{x}) = \frac{\pi}{2}$
- For $\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$, we use x_i as its ith coordinate. Note: Throughout the text we adopt the convention that the term vector (without the qualifier row or column) refers to a column vector.
- subspace: $\emptyset \neq \mathcal{V} \subset \mathbb{R}^n$ is called subspace if $a\boldsymbol{x} + b\boldsymbol{y} \in \mathcal{V}$, $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$.
- affine set: A set $\mathcal{X} = \bar{\boldsymbol{x}} + \mathcal{V}$, where $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ and $\mathcal{V} \subseteq \mathbb{R}^n$ is a subspace.

Definition (linear dependent)

A set of vectors $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k$ are said to be linearly independent if the equality

$$\alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 + \cdots \alpha_k \boldsymbol{a}_k = 0$$

implies that all coefficients α_i $(i = 1, \ldots, k)$, are equal to zero. A set of the vectors $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k$ are linearly dependent if it is not linearly independent.

★ The set composed of the single vector 0 is always linearly dependent.

Vector

Definition (linear combination)

A vector $\boldsymbol{a}$ is said to be a linear combination of vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ if there are scalars $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$ such that

$$\boldsymbol{a} = \alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 + \cdots \alpha_k \boldsymbol{a}_k.$$

Proposition (property of linear dependent)

A set of vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ are linearly dependent if and only if one of the vectors from the set is a linear combination of the remaining vectors.

Definition (span subspace)

Let $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ be vectors in $\mathbb{R}^n$. The set of all their linear combinations is called the span of $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ and is denoted

$$\text{span}[\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k] = \left\{ \sum_{i=1}^{k} \alpha_i \boldsymbol{a}_i \mid \alpha_1, \ldots, \alpha_k \in \mathbb{R} \right\}.$$

Definition (basis of the subspace)

Given a subspace $\mathcal{V}$, any set of linearly independent vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k\} \subset \mathcal{V}$ such that $\mathcal{V} = \text{span}[\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k]$ is referred to as a basis of the subspace $\mathcal{V}$.

★ All the bases of a subspace $\mathcal{V}$ contain the same number of vectors. This number is called the dimension of $\mathcal{V}$, denoted by dim$\mathcal{V}$.

Proposition (property of linear dependent)

If $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k\}$ is a basis of $\mathcal{V}$, then any vector $\boldsymbol{a} \in \mathcal{V}$ can be represented uniquely as

$$\boldsymbol{a} = \alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 + \cdots \alpha_k \boldsymbol{a}_k,$$

where the coefficients $\alpha_i \in \mathbb{R}$, $i = 1, 2, \ldots, k$. The α_i, $i = 1, 2, \ldots, k$, are called the coordinates of a with respect to the basis $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k\}$.

Definition (natural basis)

The natural (or canonical) basis of $\mathbb{R}^n$ is the column vectors of $n \times n$ identity.

Matrices are rectangular array of numbers. A matrix with m rows and n columns is called an $m \times n$ matrix, and we write

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \boldsymbol{A}^\top = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$ is its transpose.

Definition (range and null of $\boldsymbol{A} \in \mathbb{R}^{m\times n}$)

- Range of $\boldsymbol{A}$: $\mathcal{R}(\boldsymbol{A}) = \{\boldsymbol{y} \in \mathbb{R}^m \mid \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}, \ \forall \boldsymbol{x} \in \mathbb{R}^n\}$.
- Null of $\boldsymbol{A}$: $\mathcal{N}(\boldsymbol{A}) = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = 0\}$.

Definition (trace of square matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$: $\mathrm{tr}(\boldsymbol{A}) = \sum_{i=1}^n a_{ii}$)

properties of trace:

- $\mathrm{tr}(\alpha\boldsymbol{A} + \beta\boldsymbol{B}) = \alpha\mathrm{tr}(\boldsymbol{A}) + \beta\mathrm{tr}(\boldsymbol{B})$; $\mathrm{tr}(\boldsymbol{A}^T) = \mathrm{tr}(\boldsymbol{A})$; $\mathrm{tr}(\boldsymbol{AB}) = \mathrm{tr}(\boldsymbol{BA})$;
- $\mathrm{tr}(\boldsymbol{A}) = \sum_{i=1}^n \lambda_i$ the eigenvalues of $\boldsymbol{A}$ are denoted by $\lambda_1, \cdots, \lambda_n$.

Definition (rank of matrix)

The maximal number of linearly independent columns of $\boldsymbol{A}$ is called the rank of the matrix $\boldsymbol{A}$, denoted rank$\boldsymbol{A}$.

★ rank$\boldsymbol{A}$= dimension of span$[\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n]$=span$[\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m]$.

preserving rank, i.e., rank$\boldsymbol{A}$ is invariant under some operations

1. Multiplication of the columns of $\boldsymbol{A}$ by nonzero scalars.
2. Interchange of the columns.
3. Addition to a given column a linear combination of other columns.
4. Multiplied by invertible matrix.
5. Rigid transform (e.g., transform, rotation, flip, reflective, $\cdots$)
6. ....

★ rank$\boldsymbol{A}$ is invariant when operation is applied to rows.

Matrix

Definition (determinant)

For any square matrix, e.g., $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, a scalar called the determinant of $\boldsymbol{A}$, denoted $\det\boldsymbol{A}$ or $|A|$.

★ Properties and calculations of determinant (cf. linear algebra textbook).

Proposition

If matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ $(m > n)$ has a nonzero nth-order minor, then the columns of $\boldsymbol{A}$ are linearly independent; i.e., $\operatorname{rank}\boldsymbol{A} = n$.

Definition

A nonsingular matrix is a square matrix satisfying $\det\boldsymbol{A} \neq 0$. Suppose that $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. Then, $\boldsymbol{A}$ is nonsingular if and only if there is a $\boldsymbol{B} \in \mathbb{R}^{n\times n}$ such that $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}_n$. The matrix $\boldsymbol{B}$ is called inverse of $\boldsymbol{A}$.

Linear Equations

Definition (linear equations)

$\boldsymbol{Ax} = \boldsymbol{b}$ or $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b}$.

Theorem (existence, uniqueness of solution)

The linear equations $\boldsymbol{Ax} = \boldsymbol{b}$ with $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, it has

- *solution* $\iff$ *rank*$\boldsymbol{A}$ = *rank*$[\boldsymbol{A}, \boldsymbol{b}]$.
- *unique solution* $\iff$ *rank*$\boldsymbol{A}$ = *rank*$[\boldsymbol{A}, \boldsymbol{b}] = n$.
- *infinite solutions* $\iff$ *rank*$\boldsymbol{A}$ = *rank*$[\boldsymbol{A}, \boldsymbol{b}] < n$.

Theorem

The linear equations $\boldsymbol{Ax} = \boldsymbol{b}$ with $\boldsymbol{A} \in \mathbb{R}^{m\times n}$. The solution set of $\boldsymbol{Ax} = \boldsymbol{b}$ can be formulated by

- $\boldsymbol{x} = \boldsymbol{x}^0 + k_1\boldsymbol{\xi}_1 + k_2\boldsymbol{\xi}_2 + \cdots + k_{n-r}\boldsymbol{\xi}_{n-r}$;
- $\boldsymbol{x} = \boldsymbol{x}^0 + \mathcal{N}(\boldsymbol{A})$;
- $\boldsymbol{x} = \boldsymbol{A}^\dagger\boldsymbol{b}$; $\cdots$

Inner Products and Norms

Definition (Euclidean inner product)

For $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{R}^n$, the Euclidean inner product is $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i y_i = \boldsymbol{x}^\top \boldsymbol{y}$.

Definition (inner product)

The inner product is a real-valued function $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ with the following properties:

1. Nonnegativity: $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$, and $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0 \iff \boldsymbol{x} = 0$.
2. Symmetry: $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$.
3. Additivity: $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$.
4. Homogeneity: $\langle r\boldsymbol{x}, \boldsymbol{y} \rangle = r\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ for every $r \in \mathbb{R}$.

Definition (orthogonal)

The vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are said to be orthogonal if $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$.

Definition (Euclidean norm)

The Euclidean norm of an $\boldsymbol{x} \in \mathbb{R}^n$ is defined as $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} = \sqrt{\boldsymbol{x}^\top \boldsymbol{x}}$.

Theorem (Cauchy-Schwarz inequality)

For any $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$, the Cauchy-Schwarz inequality $|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\| \|\boldsymbol{y}\|$ holds. Furthermore, "$=$" holds if and only if $\boldsymbol{x} = \alpha \boldsymbol{y}$ for some $\alpha \in \mathbb{R}$.

★ variation of Cauchy-Schwarz inequality (e.g., p-norm, Q-norm, $\cdots$)

Definition (norm)

The Euclidean norm of a vector $\|\boldsymbol{x}\|$ has the following properties:

1. Nonnegativity: $\|\boldsymbol{x}\| \geq 0$, and $\|\boldsymbol{x}\| = 0 \iff \boldsymbol{x} = 0$.
2. Homogeneity: $\|r\boldsymbol{x}\| = |r| \|\boldsymbol{x}\|$ for every $r \in \mathbb{R}$.
3. Triangle inequality: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

Inner Products and Norms

Theorem (Pythagorean theorem)

If $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal, i.e. $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$, then $\|\boldsymbol{x}+\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$.

General norms

$$\|\boldsymbol{x}\|_p = \begin{cases} (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}, & \text{if } 1 \le p < \infty, \\ \max\{|x_1|, |x_2|, \cdots, |x_n|\}, & \text{if } p = \infty, \end{cases}$$

Definition (equivalent of norm)

Let $\|\cdot\|_*$, $\|\cdot\|_\dagger$ be two norms in $\mathbb{R}^n$. For any $\boldsymbol{x} \in \mathbb{R}^n$, there exist $c_1 > 0$ and $c_2 > 0$ such that $c_1\|\boldsymbol{x}\|_* \le \|\boldsymbol{x}\|_\dagger \le c_2\|\boldsymbol{x}\|_*$.

Example

$\|\boldsymbol{x}\|_2 \le \|\boldsymbol{x}\|_1 \le \sqrt{n}\|\boldsymbol{x}\|_2$; $\|\boldsymbol{x}\|_\infty \le \|\boldsymbol{x}\|_2 \le \sqrt{n}\|\boldsymbol{x}\|_\infty$;
$\|\boldsymbol{x}\|_\infty \le \|\boldsymbol{x}\|_1 \le n\|\boldsymbol{x}\|_\infty$; $\|\boldsymbol{x}\|_\infty \le \|\boldsymbol{x}\|_2 \le n\|\boldsymbol{x}\|_1$;

Definition (continuity)

A function $\boldsymbol{f}:\mathbb{R}^n\to\mathbb{R}^m$ is continuous at $\boldsymbol{x}\in\mathbb{R}^n$ if: for all $\varepsilon>0$, there exists $\delta>0$ such that $\|\boldsymbol{y}-\boldsymbol{x}\|<\delta\Rightarrow\|\boldsymbol{f}(\boldsymbol{y})-\boldsymbol{f}(\boldsymbol{x})\|<\varepsilon$.

Definition (inner product in $\mathbb{C}^n$)

In the complex vector space $\mathbb{C}^n$, inner product $\langle\boldsymbol{x},\boldsymbol{y}\rangle=\sum_{i=1}^n x_i\bar{y}_i$, where the bar denotes complex conjugation.

Property

Inner product on $\mathbb{C}^n$ is a complex-valued function admitting properties:

1. $\langle\boldsymbol{x},\boldsymbol{x}\rangle\geq 0$, "$=$" holds if and only if $\boldsymbol{x}=0$.
2. $\langle\boldsymbol{x},\boldsymbol{y}\rangle=\overline{\langle\boldsymbol{y},\boldsymbol{x}\rangle}$.
3. $\langle\boldsymbol{x}+\boldsymbol{y},\boldsymbol{z}\rangle=\langle\boldsymbol{x},\boldsymbol{z}\rangle+\langle\boldsymbol{y},\boldsymbol{z}\rangle$.
4. $\langle r\boldsymbol{x},\boldsymbol{y}\rangle=r\langle\boldsymbol{x},\boldsymbol{y}\rangle$, where $r\in\mathbb{C}$.

Homework

Exercise in text book: 2.3, 2.5, 2.10