# Lecture 1 Some Results in Linear Algebra

YANG Nian

yangnian@nju.edu.cn
Department of Finance and Insurance
Nanjing University

November 12, 2024

## Contents



## 1 Vectors, Matrices and Related Concepts

### Vectors and Matrices

- $\mathbb{R}$: the set of **real numbers** (also referred to as **scalars**).
- $\mathbb{R}^{m\times n}$: the set of **matrices** with $m$ rows and $n$ columns (of dimension $m \times n$); or the set of $m \times n$ rectangular arrays whose components are from $\mathbb{R}$.
- An element (a matrix) $A \in \mathbb{R}^{m\times n}$ can be written as

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

  where $a_{ij} \in \mathbb{R}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$.
- **Row vector**: a matrix with $m = 1$.

- **Column vector**: a matrix with $n = 1$. The word *vector* will always mean a column vector unless otherwise stated.
- $\mathbb{R}^n$: the set of all $n$-dimensional ($n$-dim) **real vectors**.

**Matrix Algebra**

- **Addition, Subtraction, Multiplication**
- The **transpose** of an $m$-by-$n$ matrix $A$ is the $n$-by-$m$ matrix $A^T$ formed by defining $(A^T)_{ij} = A_{ji}$
- **Matrix Multiplication**: $AB$ is defined only for an $m$-by-$n$ matrix $A$ and an $n$-by-$l$ matrix $B$, i.e., $(AB)_{ij} = \sum_{r=1}^{n} a_{ir}b_{rj}$
- Law of Matrix Algebra
  - **Associative laws**
    $(A + B) + C = A + (B + C)$, $(AB)C = A(BC)$
  - **Commutative law for addition**
    $A + B = B + A$
  - **Distributive laws**
    $A(B + C) = AB + AC$, $(A + B)C = AC + BC$
  - In general, commutative law for multiplication does not hold, i.e., $AB \neq BA$
- *Examples of matrix multiplications*: (a) system of equations, (b) covariance matrix of a random vector, (c) OLS.

**Elementary Matrices**

- Elementary matrices
  (a) $T_{ij}$: obtained by swapping row $i$ and row $j$.
  (b) $D_i(k)$: a diagonal matrix with diagonal entries 1 everywhere except in the $i$-th position, which is $k$.
  (c) $L_{ij}(k)$: the identity matrix but with a $k$ in the $(i, j)$ position.
- **Left (resp. right) multiplication** by an elementary matrix represents **elementary row (resp. column) operations**.
  (a) Row switching, that is interchanging two rows of a matrix.
  (b) Row multiplication, the is multiplying all entries of a row by a non-zero constant.
  (c) Row addition, that is adding a row to another.
- Elementary row operations used in **Gaussian elimination** to reduce a matrix to **row echelon form**.

**Rank and Trace**

- We say a finite collection $\mathcal{C} = \{x^1, x^2, \ldots, x^m\}$ of vectors in $\mathbb{R}^n$ is **linearly dependent** if there exists scalars $\alpha_1$,...,$\alpha_m \in \mathbb{R}$, not all of them zero, such that $\sum_{i=1}^{m} \alpha_i x^i = 0$. $\mathcal{C}$ is said to be **linearly independent** if it is not linearly dependent.
- The **rank** of a matrix $A$, $\text{rank}(A)$, is the **maximal number of linearly independent columns** of $A$.
  - $\text{rank}(A)$ is also the dimension of the vector space generated (or spanned) by its columns.
  - Column rank = Row rank.
- The **trace of a square matrix** $A$, denoted by $\text{tr}(A)$, is defined to be the sum of elements on the main diagonal of $A$.
  (a) The trace is a linear mapping.
  (b) A matrix and its transpose have the same trace.
  (c) The trace of a square matrix, which is the production of two matrices, can be rewritten as the sum of entry-wise products of their elements.
  (d) The trace is invariant under cyclic permutations.

**Inverse and Determinant**

- An $n$-by-$n$ square matrix $A$ is called **invertible** (also **nonsingular** or **nondegenerate**), if there exists an $n$-by-$n$ square matrix $B$ such that

$$AB = BA = I_n$$

  where $I_n$ denotes the $n$-by-$n$ identity matrix.
- The **determinant** is a **scalar value** that is a function of **the entries of a square matrix** $A$.
  - The determinant of $A$ is denoted by $\det(A)$, $\det A$, or $|A|$.
  - $\det(AB) = \det(A)\det(B)$
  - $\det(I_n) = 1$
  - $\det(A^{-1}) = \big[\det(A)\big]^{-1}$
- A square matrix $A$ is nonsingular if and only $\det(A) \neq 0$.

**Partitioned/Block Matrices**

- A **partitioned matrix or block matrix** is a matrix which has been partitioned into submatrices by horizontal and/or vertical lines which extend along entire rows or columns.
- Block matrix multiplication
- Block matrix inversion
- Block matrix determinant
- Block diagonal matrices

# 2 Vector Spaces, Inner Product and Norms

**Vector Space (Linear Space)**

A **vector space V** (i.e. **linear space**) over $\mathbb{R}$ (or $\mathbb{C}$, etc) is a non-empty set of (column) **vectors** with rules for **vector addition** and **scalar multiplication**, which has the following properties:

- The linear combination of vectors stay in the space.
  - If $u, v \in \mathbf{V}$ and $\alpha, \beta \in \mathbb{R}$, then $\alpha u + \beta v \in \mathbf{V}$.
- Eight axioms: $u, v, w \in \mathbf{V}$ and $\alpha, \beta \in \mathbb{R}$
  - Vector addition:
    (a) Associativity: $u + (v + w) = (u + v) + w$
    (b) Commutativity: $u + v = v + u$
    (c) Existence of **zero vector** s.t. $\mathbf{0} + v = v$ with $\mathbf{0} \in \mathbf{V}$
    (d) Additive inverse: for $v \in V$, $\exists -v \in \mathbf{V}$, s.t. $v + (-v) = 0$.
  - Scalar multiplication
    (e) Compatibility: $\alpha(\beta v) = (\alpha\beta)v$
    (f) Multiplicative identity: $1v = v$ with $1 \in \mathbb{R}$
  - Addition and Multiplication: Distributivity
    (g) $\alpha(u + v) = \alpha u + \alpha v$
    (h) $(\alpha + \beta)v = \alpha v + \beta v$

*Examples of vector spaces*: Euclidean space $\mathbb{R}^n$, $C[0, 1]$, etc.

### Normed Vector Spaces and Inner Product Spaces

- A **normed vector space** is a vector space $\mathbf{V}$ over $\mathbb{R}$ with a map $\|\cdot\| : \mathbf{V} \to \mathbb{R}$, called **norm** that satisfies four conditions below for all $u, v \in \mathbf{V}$ and $\alpha \in \mathbb{R}$
  (a) **(Non-Negativity)** $\|v\| \geq 0$
  (b) **(Positivity)** $\|v\| = 0$ if and only if $v = 0$
  (c) **(Absolute homogeneity)** $\|\alpha v\| = |\alpha|\|v\|$
  (d) **(Subadditivity/Triangle inequality)** $\|v + u\| \leq \|v\| + \|u\|$
- An **inner product space** is a vector space $\mathbf{V}$ over $\mathbb{R}$ with a map $\langle \cdot, \cdot \rangle : \mathbf{V} \times \mathbf{V} \to \mathbb{R}$, called **inner product** that satisfies four conditions below for all $u, v, w \in \mathbf{V}$ and $\alpha \in \mathbb{R}$
  (a) $\langle v, v \rangle \geq 0$, and $\langle v, v \rangle = 0$ if and only if $v = 0$
  (b) $\langle v, u + w \rangle = \langle v, u \rangle + \langle v, w \rangle$
  (c) $\langle \alpha v, u \rangle = \alpha \langle v, u \rangle$
  (d) $\langle v, u \rangle = \langle u, v \rangle$
  - An inner product is a generalization of the **dot product** on finite dimensional vector spaces.
  - An inner product space could be a normed linear space, but the reverse is not true.
  - *Examples*: Euclidean norm, $\ell_p$-norm, matrix norms, etc.

### Euclidean Norm

- $\mathbb{R}^n$: a typical Euclidean vector space equipped with **dot product**, as the inner product on $\mathbb{R}^n$, defined below

$$\langle x, y \rangle := x^T y = \sum\nolimits_{i=1}^{n} x_i y_i, \text{ for all } x, y \in \mathbb{R}^n.$$

- Two vectors $x, y \in \mathbb{R}^n$ are **orthogonal** if $\langle x, y \rangle = 0$.
- The dot (inner) product induces a norm $\|\cdot\|_2$ on $\mathbb{R}^n$, called the **Euclidean norm**, for all $x \in \mathbb{R}^n$

$$\|x\|_2 := \sqrt{\langle x, x \rangle} = \sqrt{x^T x} = \left( \sum\nolimits_{i=1}^{n} x_i^2 \right)^{1/2}.$$

- The **Cauchy-Schwartz** inequality states $\|x^T y\|_2 \leq \|x\|_2 \|y\|_2$.
- Other norms on $\mathbb{R}^n$
  - For $p \geq 1$, define the $\ell_p$**-norm** $\|\cdot\|_p$ via $\|x\|_p = (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}$ for all $x \in \mathbb{R}^n$.
  - The $\ell_\infty$-norm: $\|x\|_\infty := \max_{1 \leq i \leq n}\{|x_i|\} = \lim_{p \to \infty} \|x\|_p$.
- **Geometric interpretation of the inner (dot) product.**

**Linear Subspaces, Dimension and Bases**

- A non-empty set $S$ of $\mathbb{R}^n$ is called a **subspace** of $\mathbb{R}^n$ if $\alpha x + \beta y \in S$ whenever $x, y \in S$ and $\alpha, \beta \in \mathbb{R}$.
- The **span** of a finite collection $\mathcal{C} = \{x^1, \ldots, x^m\}$ of vectors in $\mathbb{R}^n$ is defined as
$$\operatorname{span}(\mathcal{C}) := \left\{\sum_{i=1}^{m} \alpha_i x^i : \alpha_1, \ldots, \alpha_m \in \mathbb{R}\right\}. \tag{1}$$
$\operatorname{span}(\mathcal{C})$ is a subspace of $\mathbb{R}^n$.
- Given a subspace $S$ of $\mathbb{R}^n$ with $S \neq \{\mathbf{0}\}$, a **basis** of $S$ is a linearly independent collection of vectors spanning $S$.
  - The **number of the basis** of $S$, $\dim(S)$, is called the **dimension** of the subspace $S$.
  - Every basis of a given subspace $S$ has the same number.
  - By definition, the dimension of the subspace $\{\mathbf{0}\}$ is zero.
- The **orthogonal complement** $S^\perp$ of $S$ is defined as $S^\perp := \{y \in \mathbb{R}^n : x^T y = 0\}$.
  - $S^\perp$ is a subspace of $\mathbb{R}^n$, $\dim(S) + \dim(S^\perp) = n$, $(S^\perp)^\perp = S$.
  - $\forall x \in \mathbb{R}^n$, there is a unique decomposition $x = x^1 + x^2$, where $x^1 \in S$ and $x^2 \in S^\perp$.

**Range, Nullspace, Orthogonal Decomposition**

- Let $A \in \mathbb{R}^{m\times n}$. The **column space** of $A$ is a subspace of $\mathbb{R}^m$ spanned by the columns of $A$.
  - The column space is also known as the **range** of $A$, denoted by
  $$\mathcal{R}(A) := \{Ax : x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m.$$
  - The system $Ax = b$ is solvable. $\Leftrightarrow b \in \mathcal{R}(A)$.
  - The **row space** of $A$ is the subspace of $\mathbb{R}^n$ spanned by the rows of $A$, i.e., $\mathcal{R}(A^T)$.
  - $\operatorname{rank}(A) = \dim\big(\mathcal{R}(A)\big) = \dim\big(\mathcal{R}(A^T)\big) \le \min\{m, n\}$.
- The **nullspace** (or **kernel**) of $A$ is the set
$$\mathcal{K}(A) := \{x \in \mathbb{R}^n : Ax = 0\} \subseteq \mathbb{R}^n.$$
- **Orthogonal decomposition** induced by $A$
$$\mathcal{R}(A) = \big(\mathcal{K}(A^T)\big)^\perp, \quad \mathcal{R}(A^T) = \big(\mathcal{K}(A)\big)^\perp.$$
- **Orthogonalization**: linearly independent vectors to orthogonal vectors
- **Projection** of a vector onto a subspace

# 3 Eigenvalues and Eigenvectors

### Eigenvalues and Eigenvectors

- Consider $A \in \mathbb{R}^{n\times n}$. We say that $\lambda \in \mathbb{C}$ is an **eigenvalue** of $A$ with corresponding **eigenvector** $u \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ if $Au = \lambda u$.
  - The zero vector $\mathbf{0} \in \mathbb{R}^n$ cannot be an eigenvector.
  - Given $A \in \mathbb{R}^{n\times n}$, there are exactly $n$ eigenvalues (counting multiplicities).
- **Eigendecomposition (the eigenvectors diagonalize a matrix)**. Suppose $A \in \mathbb{R}^{n\times n}$ has $n$ linearly independent eigenvectors $u_1,\ldots,u_n \in \mathbb{R}^n$ associated with eigenvalues $\lambda_1,\ldots,\lambda_n$, respectively. Let $U := [u_1, \ldots, u_n]$ and $\Lambda = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$. Then

$$U^{-1}AU = \Lambda.$$

- Consider the set of eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ of $A \in \mathbb{R}^{n\times n}$.
  (a) $\mathrm{tr}(A) = \sum_{i=1}^n \lambda_i$, and $\det(A) = \Pi_{i=1}^n \lambda_i$.
  (b) The eigenvalues of $A^T$ are the same as those of $A$.
  (c) For any $c \in \mathbb{R}$, the eigenvalues of $cI + A$ are $c + \lambda_1, \ldots, c + \lambda_n$.
  (d) For any integer $k \geq 1$, the eigenvalues of $A^k$ are $\lambda_1^k, \ldots, \lambda_n^k$.
  (e) If $|A| \neq 0$, then the eigenvalues of $A^{-1}$ are $\lambda_1^{-1}, \ldots, \lambda_n^{-1}$.

### Real Symmetric Matrices and Positive Semidefinite Matrices

- **Real symmetric matrices**: $\mathbb{S}^n = \{A \in \mathbb{R}^{n\times n} : A = A^T\}$
- $A \in \mathbb{S}^n \Leftrightarrow \langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in \mathbb{R}^n$.
- $A \in \mathbb{S}^n$ has **real** eigenvalues and **orthogonal** eigenvectors associate with different eigenvalues.
- **The Spectral Theorem.** For $A \in \mathbb{S}^n$, there exists an **orthogonal** matrix $U \in \mathbb{R}^{n\times n}$ (i.e. $UU^T = U^TU = I$), and a diagonal matrix $\Lambda \in \mathbb{R}^{n\times n}$ such that

$$A = U\Lambda U^T = \sum_{i=1}^n \lambda_i u_i u_i^T.$$

- The following statements are equivalent for $A \in \mathbb{S}^n$.
  (a) $A$ is positive semidefinite.
  (b) All eigenvalues of $A$ are non-negative.
  (c) There exists a unique $A^{1/2} \in \mathbb{S}_+^n$ such that $A = A^{1/2}A^{1/2}$.
  (d) There exists an $k \times n$ matrix $B$, where $k = \mathrm{rank}(A)$ such that $A = B^TB$.

**Positive Definite Matrices $\mathbb{S}^n_{++}$**

- **Positive Semidefinite/Definite Matrix**: $\mathbb{S}^n_{+}/\mathbb{S}^n_{++}$

$$\mathbb{S}^n_{+} = \left\{A \in \mathbb{S}^n : x^T A x \geq 0 \text{ if } x \in \mathbb{R}^n \backslash \{\mathbf{0}\}\right\}$$
$$\mathbb{S}^n_{++} = \left\{A \in \mathbb{S}^n : x^T A x > 0 \text{ if } x \in \mathbb{R}^n \backslash \{\mathbf{0}\}\right\}$$

- The following statements are equivalent for $A \in \mathbb{S}^n$.
  (a) $A$ is positive definite.
  (b) $A^{-1}$ exists and is positive definite.
  (c) All eigenvalues of $A$ are positive.
  (d) There exists a unique $n \times n$ $A^{1/2} \in \mathbb{S}^n_{++}$ such that $A = A^{1/2}A^{1/2}$.

- **Schur Complement**. Consider $X \in \mathbb{S}^n$ as

$$X = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}, \tag{2}$$

where $A \in \mathbb{S}^k$ and $|A| \neq 0$, the matrix

$$D = C - B^T A^{-1} B$$

is called the Schur complement of $A$ in $X$.

# 4 Singular Values and Singular Vectors*

**Singular Values and Singular Vectors***

- Interpret $A \in \mathbb{R}^{m\times n}$ as a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$
- Given $A \in \mathbb{R}^{m\times n}$ with $\text{rank}(A) = r \geq 1$. Then, there exist orthogonal matrices $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ such that

$$A = U\Lambda V^T,$$

which is the **Singular Value Decomposition (SVD)** of $A$.

  - Let $q = \min\{m, n\}$. $\Lambda \in \mathbb{R}^{m\times n}$ has $\Lambda_{ij} = 0$ for $i \neq j$, and $\Lambda_{11} \geq \Lambda_{11} \geq \cdots \geq \Lambda_{rr} > \Lambda_{r+1,r+1} = \cdots = \Lambda_{qq} = 0$.
  - $\Lambda_{11}, \ldots, \Lambda_{qq}$ are called the **the singular values** of $A$.
  - The columns of $U$ (resp. $V$) are called the **left** (resp. **right**) **singular vectors** of $A$

- The SVD can also be written as

$$A = \sum\nolimits_{i=1}^{r} \Lambda_{ii} u_i v_i^T,$$

where $u_i$ (resp. $v_i$) is the $i$-th column of the matrix $U$ (resp. $V$), for $i = 1, \ldots, r$.