计量经济学

第 4 讲：线性代数复习

授课人：刘　岩

中山大学商学院

2025 年 3 月 5 日

本讲内容

1. 欧式空间
2. 线性变换与矩阵
3. 矩阵的逆与特征值
4. 对称矩阵

更多参考资料见：http://www.liuyanecon.com/book-linearalgebra/
及 3blue1brown 线性代数系列视频

本节内容

1. 欧式空间
2. 线性变换与矩阵
3. 矩阵的逆与特征值
4. 对称矩阵

欧式空间中的向量和运算

$\forall n \in \mathbb{N}$，n-维欧式空间 (Euclidean space) $\mathbb{R}^n$ 任意一点 $\boldsymbol{x}$ 可表示为 n-维（列）向量

$$\boldsymbol{x} = [x_1, \ldots, x_n]^\mathsf{T}, \quad x_i \in \mathbb{R}, i = 1, \ldots, n$$

$\mathbb{R}^n$ 具有线性空间结构，即存在零向量 $\mathbf{0} = [0, \ldots, 0]$，且可自然定义如下两个运算

1. 数乘：$\forall \boldsymbol{x} \in \mathbb{R}^n$ 以及 $\alpha \in \mathbb{R}$，$\alpha\boldsymbol{x} = [\alpha x_1, \ldots, \alpha x_n]^\mathsf{T} \in \mathbb{R}^n$
2. 加法：$\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$，$\boldsymbol{x} + \boldsymbol{y} = [x_1 + y_1, \ldots, x_n + y_n]^\mathsf{T} \in \mathbb{R}^n$

$\mathbb{R}^n$ 对数乘与加法运算封闭；并且 $\forall \boldsymbol{x} \in \mathbb{R}^n$，存在其加法的逆 $-\boldsymbol{x} = [-x_1, \ldots, -x_n]^\mathsf{T}$，使得 $\boldsymbol{x} + (-\boldsymbol{x}) = \mathbf{0}$；此外，数乘与加法满足自然的结合律、交换律与分配率

线性空间一般概念

- 对于任意集合 $\mathcal{X}$ 以及数域 $\mathbb{K}$，若 $\mathcal{X}$ 中向量对 $\mathbb{K}$ 中元素的数乘以及向量间加法运算封闭，即运算结果仍在该集合中，且存在零元、逆并满足运算结合律、交换律和分配率，则称为线性空间 (linear space)
  - 数域 $\mathbb{K}$ 为一个集合：其元素可以进行加、减、乘、除等运算，且运算结果还在该集合中，即对运算封闭；如 $\mathbb{Q}, \mathbb{R}, \mathbb{C}$
- 若 $\mathcal{Y} \subset \mathcal{X}$，且 $\mathcal{Y}$ 对数乘与加法封闭，同时含有零元、逆并满足运算定律，则称 $\mathcal{Y}$ 为 $\mathcal{X}$ 的线性子空间 (linear subspace)
  - 在 $\mathbb{R}^3$ 中，任何过元点的 2 维平面均为线性子空间

欧式空间上的内积与模长

- $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$，定义
$$\boldsymbol{x} \cdot \boldsymbol{y} = x_1 y_1 + \cdots + x_n y_n \in \mathbb{R}$$
为 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积 (inner product)，其中 $x_i y_i$ 表示实数乘积
- 内积可看做函数 $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$，且满足对称性，即 $f(\boldsymbol{x}, \boldsymbol{y}) = f(\boldsymbol{y}, \boldsymbol{x})$，以及双线性性 (bilinearity)：
  1. $\forall \alpha \in \mathbb{R}, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$，有 $f(\alpha \boldsymbol{x}, \boldsymbol{y}) = \alpha f(\boldsymbol{x}, \boldsymbol{y})$
  2. $\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$，有 $f(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}) = f(\boldsymbol{x}, \boldsymbol{z}) + f(\boldsymbol{y}, \boldsymbol{z})$
- $\forall \boldsymbol{x} \in \mathbb{R}^n$，定义
$$||\boldsymbol{x}|| = (\boldsymbol{x} \cdot \boldsymbol{x})^{\frac{1}{2}} = \sqrt{x_1^2 + \cdots x_n^2}$$
为 $\boldsymbol{x}$ 的模长 (norm)，满足 $||\boldsymbol{x}|| = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$ 及如下两条性质
  1. $\forall \alpha \in \mathbb{R}, \boldsymbol{x} \in \mathbb{R}^n$，有 $||\alpha \boldsymbol{x}|| = |\alpha| \times ||\boldsymbol{x}||$
  2. $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$，有三角不等式 $||\boldsymbol{x} + \boldsymbol{y}|| \le ||\boldsymbol{x}|| + ||\boldsymbol{y}||$

Cauchy-Schwartz 不等式

- Cauchy-Schwartz 不等式：$\forall x, y \in \mathbb{R}^n$，有 $|x \cdot y| \le ||x|| \times ||y||$

证明 $||x|| = 0$ 时不等式显然成立，故仅考虑 $||x|| > 0$ 的情形；此时，$\forall t \in \mathbb{R}$，$0 \le ||tx + y||^2 = (tx + y) \cdot (tx + y) = ||x||^2 t^2 + 2(x \cdot y)t + ||y||^2$，故由 $||x|| > 0$ 知前式最右作为 t 的 2 次函数恒大于等于 0，其判别式

$$\Delta = 4(x \cdot y)^2 - 4||x||^2||y||^2 \le 0 \Leftrightarrow |x \cdot y| \le ||x|| \times ||y||$$

- 由 Cauchy-Schwartz 即可证明 $||\cdot||$ 的三角不等式

$$||x + y||^2 = ||x||^2 + 2(x \cdot y) + ||y||^2 \le ||x||^2 + 2||x|| \times ||y|| + ||y||^2 = (||x|| + ||y||)^2$$

注 Cauchy-Schwartz 不等式是正定对称双线性函数的一般性质，即对 $f(x, x) \ge 0$ 的对称双线性函数 $f(\cdot, \cdot)$，有 $|f(x, y)|^2 \le f(x, x) \times f(y, y)$

欧式空间的基

- 定义 $\mathbb{R}^n$ 中单位向量

$$e_i = [0, \dots, 0, \underbrace{1}_{\text{第}i\text{位}}, 0, \dots, 0]^\top, \quad i = 1, \dots, n$$

则 $\forall \boldsymbol{x} = [x_1, \dots, x_n]^\top \in \mathbb{R}^n$，有如下线性表示 (linear representation)

$$\boldsymbol{x} = x_1 \boldsymbol{e}_1 + \cdots + x_n \boldsymbol{e}_n$$

- $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ 称为 $\mathbb{R}^n$ 的一组基向量 (basis)
- $\mathbb{R}^n$ 中的基向量并不唯一，如 $\mathbb{R}^2$ 中，$\tilde{\boldsymbol{e}}_1 = [1, 1]^\top, \tilde{\boldsymbol{e}}_2 = [-1, 1]^\top$ 也是一组基向量
- 基向量的判定标准：向量之间是否线性独立

向量间的线性独立性

- 给定 $m \geq 2$ 个向量 $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \in \mathbb{R}^n$，若存在不全为 0 的实数 $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$，使得

$$\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_m \boldsymbol{x}_m = \mathbf{0}$$

  则称 $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ 线性相关 (linearly dependent)
- 若否，即对任意非全为零的实数 $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$，

$$\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_m \boldsymbol{x}_m \neq \mathbf{0}$$

  则称 $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ 线性无关或线性独立 (linearly independent)
- $\mathbb{R}^n$ 中任意 n 个线性独立的向量组 $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$，都构成一组基向量
  - $\mathbb{R}^n$ 中线性无关的向量组最多包含 n 个向量，等价的，线性无关向量的最大个数等于线性空间的维数 (dimension)

本节内容

1 欧式空间

2 线性变换与矩阵

3 矩阵的逆与特征值

4 对称矩阵

线性变换

- 考虑两个线性空间 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 间的映射，$\mathcal{T}:\mathbb{R}^n \longrightarrow \mathbb{R}^m$，若 $\mathcal{T}$ 满足如下性质
  1. $\forall \alpha \in \mathbb{R}, \boldsymbol{x} \in \mathbb{R}^n$，都有 $\mathcal{T}(\alpha\boldsymbol{x}) = \alpha\mathcal{T}(\boldsymbol{x})$
  2. $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$，都有 $\mathcal{T}(\boldsymbol{x}+\boldsymbol{y}) = \mathcal{T}(\boldsymbol{x}) + \mathcal{T}(\boldsymbol{y})$

  则称 $\mathcal{T}$ 为 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性变换 (linear transformation)
  - $\mathcal{T}$ 为花体字母 T
- 给定 $\mathbb{R}^n$ 中的单位基 $\{\boldsymbol{e}_j\}_{j=1}^n$ 与 $\mathbb{R}^m$ 中的单位基 $\{\boldsymbol{f}_i\}_{i=1}^m$，则 $\forall j = 1, \ldots, n$，$\mathcal{T}(\boldsymbol{e}_j)$ 可由 $\{\boldsymbol{f}_i\}_{i=1}^m$ 线性表示

$$\mathcal{T}(\boldsymbol{e}_j) = t_{1j}\boldsymbol{f}_1 + \cdots + t_{mj}\boldsymbol{f}_m = \sum_{i=1}^m t_{ij}\boldsymbol{f}_i, \quad t_{ij} \in \mathbb{R}, i = 1, \ldots, m$$

  其中，$\mathcal{T}(\boldsymbol{e}_j)$ 的第 i 个分量为 t_{ij}，对应 $t_{ij}\boldsymbol{f}_i$

线性变换的矩阵表示

- $\forall \boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n \in \mathbb{R}^n$，由线性变换的性质可知

$$\mathscr{T}(\boldsymbol{x}) = x_1\mathscr{T}(\boldsymbol{e}_1) + \cdots + x_n\mathscr{T}(\boldsymbol{e}_n) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} t_{ij}\boldsymbol{f}_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} t_{ij}x_j\right)\boldsymbol{f}_i$$

故 $\mathscr{T}(\boldsymbol{x})$ 的第 i 个分量为 $\sum_{j=1}^{n} t_{ij}x_j$

- 将 $\{t_{ij}\}$ 排列为如下矩阵形式

$$\boldsymbol{T} = \begin{bmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{m1} & \cdots & t_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n}$$

线性变换的矩阵表示

- 定义矩阵与向量的乘法为

$$\boldsymbol{T}\boldsymbol{x} = \begin{bmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{m1} & \cdots & t_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} t_{1j} x_j \\ \vdots \\ \sum_{j=1}^{n} t_{mj} x_j \end{bmatrix}$$

  则 $\mathcal{T}(\boldsymbol{x}) = \boldsymbol{T}\boldsymbol{x}$，即线性变换 $\mathcal{T}$ 可以表示为矩阵与向量的乘积

- 此时 $\boldsymbol{T}$ 称作变换矩阵

结论 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性变换 $\mathcal{T}$ 与 $m \times n$ 矩阵 $\boldsymbol{T}$ 一一对应，$\boldsymbol{T}$ 的元素 t_{ij}，由 $\mathbb{R}^n$ 单位基向量 $\{\boldsymbol{e}_j\}$ 在线性变换 $\mathcal{T}$ 下的象，通过 $\mathbb{R}^m$ 单位基向量 $\{\boldsymbol{f}_i\}$ 的线性表示完全决定，t_{ij} 等于 $\mathcal{T}(\boldsymbol{e}_j)$ 的第 i 个分量

线性变换的复合与矩阵乘法

- 给定线性变换 $\mathscr{A}:\mathbb{R}^m\to\mathbb{R}^k$ 和 $\mathscr{B}:\mathbb{R}^n\to\mathbb{R}^m$，相应的变换矩阵为

$$A_{k\times m}=\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{km} \end{bmatrix},\quad B_{m\times n}=\begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$$

- 复合变换 $\mathscr{C}=\mathscr{A}\circ\mathscr{B}:\mathbb{R}^n\to\mathbb{R}^k, \boldsymbol{x}\mapsto\mathscr{A}(\mathscr{B}(x))$ 对应的变换矩阵 $C_{k\times n}$ 为

$$C=AB=\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{km} \end{bmatrix}\begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}=\begin{bmatrix} \sum_{j=1}^m a_{1j}b_{j1} & \cdots & \sum_{j=1}^m a_{1j}b_{jn} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^m a_{kj}b_{j1} & \cdots & \sum_{j=1}^m a_{kj}b_{jn} \end{bmatrix}$$

注 矩阵乘积须维数吻合：左乘矩阵的列数等于右乘矩阵的行数

分块矩阵乘积：列-行形式

- 将 $\boldsymbol{A}_{k\times m}$ 写为列向量的形式，$\boldsymbol{B}_{m\times n}$ 写为行向量的形式，则 $\boldsymbol{AB}$ 可写为

$$\begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_m \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_m \end{bmatrix} = \boldsymbol{a}_1\boldsymbol{b}_1 + \cdots + \boldsymbol{a}_m\boldsymbol{b}_m = \sum_{i=1}^{m} \underbrace{\boldsymbol{a}_i\boldsymbol{b}_i}_{k\times n} = \sum_{j=1}^{m} \begin{bmatrix} a_{1j}b_{j1} & \cdots & a_{1j}b_{jn} \\ \vdots & \ddots & \vdots \\ a_{kj}b_{j1} & \cdots & a_{kj}b_{jn} \end{bmatrix}$$

  - 此处 $\boldsymbol{b}_i$ 为行向量形式，与默认列向量的惯例不同
- 注意，每一个 $\boldsymbol{a}_j\boldsymbol{b}_j$ 都是 $k\times n$ 矩阵

注 同样维数矩阵的加法，就是对应元素相加

分块矩阵乘积：一般形式

- 更一般的情况，假设 $\boldsymbol{A}_{k\times m}$ 与 $\boldsymbol{B}_{m\times n}$ 可写为 $p\times q$ 与 $q\times r$ 分块形式

$$\boldsymbol{A}_{k\times m}=\begin{bmatrix}\boldsymbol{A}_{11} & \cdots & \boldsymbol{A}_{1q}\\ \vdots & \ddots & \vdots\\ \boldsymbol{A}_{p1} & \cdots & \boldsymbol{A}_{pq}\end{bmatrix},\quad \boldsymbol{B}_{m\times n}=\begin{bmatrix}\boldsymbol{B}_{11} & \cdots & \boldsymbol{B}_{1r}\\ \vdots & \ddots & \vdots\\ \boldsymbol{B}_{q1} & \cdots & \boldsymbol{B}_{qr}\end{bmatrix}$$

$\boldsymbol{A}_{ij},\boldsymbol{B}_{j\ell}$ 维数分别为 $k_i\times m_j, m_j\times n_\ell$，$i=1,\ldots,p, j=1,\ldots,q, \ell=1,\ldots,r$，且各分块矩阵维数满足 $\sum_{i=1}^{p}k_i=k,\sum_{j=1}^{q}m_j=m,\sum_{\ell=1}^{r}n_\ell=n$，则有

$$\boldsymbol{A}\boldsymbol{B}=\begin{bmatrix}\sum_{j=1}^{q}\boldsymbol{A}_{1j}\boldsymbol{B}_{j1} & \cdots & \sum_{j=1}^{q}\boldsymbol{A}_{1j}\boldsymbol{B}_{jr}\\ \vdots & \ddots & \vdots\\ \sum_{j=1}^{q}\boldsymbol{A}_{pj}\boldsymbol{B}_{j1} & \cdots & \sum_{j=1}^{q}\boldsymbol{A}_{pj}\boldsymbol{B}_{jr}\end{bmatrix}$$

矩阵的转置与乘积

- 给定 $m \times n$ 矩阵 $\boldsymbol{A} = [a_{ij}]_{1\le i\le m, 1\le j\le n}$，定义一个 $n \times m$ 矩阵 $\boldsymbol{B}$，其 j 行 i 列元素为 a_{ij}，即

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix}$$

  称 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的转置，记作 $\boldsymbol{B} = \boldsymbol{A}^{\mathsf{T}}$

- 给定 $\boldsymbol{A}_{k\times m}$ 与 $\boldsymbol{B}_{m\times n}$，则矩阵乘积的转置等于交换顺序后矩阵转置再乘积：

$$(\boldsymbol{A}\boldsymbol{B})^{\mathsf{T}} = \boldsymbol{B}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}}$$

注 即便维数吻合，矩阵乘积一般也不满足交换律：除对特定矩阵，否则 $\boldsymbol{A}\boldsymbol{B} \neq \boldsymbol{B}\boldsymbol{A}$

矩阵的秩

- 矩阵 $\boldsymbol{A} = [a_{ij}]_{m\times n}$ 可以看做 m 个行向量 $\{\boldsymbol{a}_{i\cdot}\}_{i=1}^{m}$ 或 n 个列向量 $\{\boldsymbol{a}_{\cdot j}\}_{j=1}^{n}$ 构成

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_{1\cdot} \\ \vdots \\ \boldsymbol{a}_{m\cdot} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_{\cdot 1} & \cdots & \boldsymbol{a}_{\cdot n} \end{bmatrix}$$

- 矩阵线性无关的行向量的个数等于线性无关的列向量的个数，这一个数称为矩阵的秩 (rank)，记作 $\mathrm{rank}(\boldsymbol{A})$
  - $\mathrm{rank}(\boldsymbol{A}) \le \min\{m, n\}$；若 $\mathrm{rank}(\boldsymbol{A}) = \min\{m, n\}$，则称其为满秩 (full rank) 矩阵
- 矩阵 $\boldsymbol{A}_{m\times n}$ 的秩与以 $\boldsymbol{A}$ 为系数矩阵、$\boldsymbol{b}_{m\times 1}$ 为值向量的线性方程组 $\boldsymbol{A}\boldsymbol{x}_{n\times 1} = \boldsymbol{b}$ 的解有密切联系

本节内容

1 欧式空间

2 线性变换与矩阵

3 矩阵的逆与特征值

4 对称矩阵

方阵与单位阵

- $n \times n$ 矩阵称为方阵 (square matrix)
  - 方阵 $\boldsymbol{A}$ 对应的线性变换 $\mathscr{A}: \mathbb{R}^n \to \mathbb{R}^n$ 可看做 $\mathbb{R}^n$ 自身的线性变换
  - $\mathbb{R}^n$ 上的线性变换可分为拉伸、旋转、反射与剪切四种
    - 更多讨论，见：耿修瑞，《矩阵之美（基础篇）》，科学出版社，2023
    - 注意，平移不是线性变换；线性变换需保持元点不动
- 以下矩阵称为 n-阶单位阵 (identity matrix)：

$$\boldsymbol{I} = \mathrm{diag}(\underbrace{1, \ldots, 1}_{n\text{个}}) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{n \times n}$$

- 对任意 $n \times n$ 方阵 $\boldsymbol{A}$，有 $\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}$

矩阵的逆

- 给定 $n \times n$ 矩阵 $\boldsymbol{A}$，若存在 $n \times n$ 矩阵 $\boldsymbol{B}$，满足 $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{I}_n$ 与 $\boldsymbol{B}\boldsymbol{A} = \boldsymbol{I}_n$，则称 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的逆矩阵 (inverse matrix)，称 $\boldsymbol{A}$ 可逆 (invertible)，逆矩阵记作 $\boldsymbol{A}^{-1}$
  - 朱慧坚、丁玖在返朴公众号的文章《从反函数的观点看逆矩阵》（2025 年 1 月 23 日）[https://mp.weixin.qq.com/s/c8kpLRuK5S04zntk_ddB8A]，从几何角度重新解释了逆矩阵的含义，并由此说明若 $\boldsymbol{B}$ 满足 $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{I}$，必有 $\boldsymbol{B}\boldsymbol{A} = \boldsymbol{I}$
- 矩阵的逆若存在，必唯一：若 $\boldsymbol{B}, \boldsymbol{C}$ 均为 $\boldsymbol{A}$ 的逆，则 $\boldsymbol{B} = \boldsymbol{C}$
- 乘积矩阵的逆矩阵为逆矩阵交换顺序后的乘积：若 $n \times n$ 方阵 $\boldsymbol{A}, \boldsymbol{B}$ 可逆，则

$$(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

  即 $\boldsymbol{A}\boldsymbol{B}$ 也可逆
  - 直接计算 $\boldsymbol{A}\boldsymbol{B}\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$ 与 $\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{B}$ 即可验证

矩阵可逆的等价条件：给定 n 阶方阵 A，以下条件相互等价

1. A 可逆
2. A 满秩，即 $\mathrm{rank}(A)=n$
3. A 行列式不为 0，即 $\det(A)\neq 0$

注 若定义在 n 阶方阵集合上的函数 $f:\mathbb{R}^n\times\cdots\times\mathbb{R}^n\,(\equiv\mathbb{R}^{n^2})\to\mathbb{R}$，即变元为方阵 A 的 n 个 n-维列向量（或行向量）$\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n\in\mathbb{R}^n$，满足如下 3 条性质：

1. 多重线性性：$\forall\boldsymbol{b}\in\mathbb{R}^n, i=1,\ldots,n$，有
   $f(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_i+\boldsymbol{b},\ldots,\boldsymbol{a}_n)=f(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_i,\ldots,\boldsymbol{a}_n)+f(\boldsymbol{a}_1,\ldots,\boldsymbol{b},\ldots,\boldsymbol{a}_n)$，且
   $\forall\alpha\in\mathbb{R}, i=1,\ldots,n$，有 $f(\boldsymbol{a}_1,\ldots,\alpha\boldsymbol{a}_i,\ldots,\boldsymbol{a}_n)=\alpha f(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_i,\ldots,\boldsymbol{a}_n)$
2. 反对称性：$\forall i,j=1,\ldots,n, i\neq j$，有
   $f(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_i,\ldots,\boldsymbol{a}_j,\ldots,\boldsymbol{a}_n)=-f(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_j,\ldots,\boldsymbol{a}_i,\ldots,\boldsymbol{a}_n)$
3. 归一化：对单位基向量组 $\{\boldsymbol{e}_i\}_{1\leq i\leq n}$ 有 $f(\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n)=1$

则称 $f(\cdot,\ldots,\cdot)$ 为行列式 (determinant) 函数，A 的行列式记作 $\det(A)$

矩阵的特征多项式与特征值

- 给定 n 阶方阵 $\boldsymbol{A} = [a_{ij}]_{1\le i,j\le n}$，对变元 λ 定义

$$\mathcal{P}(\lambda) = \det(\lambda\boldsymbol{I} - \boldsymbol{A}) = \det\begin{bmatrix} \lambda - a_{11} & \cdots & -a_{1n} \\ \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & \lambda - a_{nn} \end{bmatrix}$$

  即 $\mathcal{P}(\lambda)$ 为 λ 的 n 阶多项式，称为 $\boldsymbol{A}$ 的特征多项式 (characteristic polynomial)
  - 行列式的计算只涉及矩阵元素的加、减、乘，故 $\mathcal{P}(\lambda)$ 的所有系数，均为 $\boldsymbol{A}$ 的元素的进行计算所得
- 特征多项式的零点，即 $\mathcal{P}(\lambda) = 0$ 的根，称为 $\boldsymbol{A}$ 的特征值 (characteristic value)

代数基本定理与特征值

定理 1

$\forall n \in \mathbb{N}$，复数域 $\mathbb{C}$ 上的任意 n 阶多项式

$$\mathcal{P}_n(z) = a_n z^n + \cdots + a_1 z + a_0, \quad z \in \mathbb{C}$$

其中系数 $a_n, \ldots, a_0 \in \mathbb{C}$ 且 $a_n \neq 0$，在 $\mathbb{C}$ 中一定有 n 个零点（含重根）：$z_1, \ldots, z_n \in \mathbb{C}$ 使得 $\mathcal{P}_n(z_i) = 0, i = 1, \ldots, n$

- 作为代数基本定理的应用，任意 n 阶实数方阵 $\boldsymbol{A} \in \mathbb{R}^{n^2}$ 在 $\mathbb{C}$ 中一定有 n 个特征值：$\lambda_1, \ldots, \lambda_n \in \mathbb{C}$
- 实矩阵 $\boldsymbol{A}$ 的特征多项式 $\mathcal{P}(\lambda)$ 系数全为实数，此时若有复特征值 $\lambda = a + bi \in \mathbb{C}$，则其复共轭 $\bar{\lambda} = a - bi$ 亦为特征值，故复特征值总是成对出现

矩阵的特征向量

- 若 λ_i 为 n 阶方阵 $\boldsymbol{A}$ 的特征值，则存在非零 n 阶向量 $\boldsymbol{c}_i \neq \boldsymbol{0}$，使得 $\boldsymbol{A}\boldsymbol{c}_i = \lambda_i \boldsymbol{c}_i$，此时称 $\boldsymbol{c}_i$ 为矩阵 $\boldsymbol{A}$（与特征 λ_i 对应的）特征向量 (characteristic vector)
  - 若 λ_i 为 $\boldsymbol{A}$ 的特征值，则 $\det(\lambda_i \boldsymbol{I} - \boldsymbol{A}) = 0$，故 $\mathrm{rank}(\lambda_i \boldsymbol{I} - \boldsymbol{A}) < n$，此时线性方程组 $(\lambda_i \boldsymbol{I} - \boldsymbol{A})\boldsymbol{c} = \boldsymbol{0}$ 一定有非零解
  - 若 λ_i 为复特征值，则对应的特征向量 $\boldsymbol{c}_i \in \mathbb{C}^n$ 是 n 维复向量
- 对应于不同特征值的特征向量，线性无关
- 特征值为重根时，对应的特征向量可能线性相关，也可能线性无关
  - 单位阵 $\boldsymbol{I}_2$ 的特征值显然有重根 1，同时任意两个线性无关的向量都是特征向量；而矩阵 $\begin{bmatrix} 3 & 2 \\ -2 & -1 \end{bmatrix}$ 特征多项式同为 $\lambda^2 - 2\lambda + 1$，但线性无关特征向量只有 $[1, -1]^{\mathrm{T}}$

特征值分解：定义

- 若 n 阶方阵 $\boldsymbol{A}$ 的特征值 $\lambda_1, \ldots, \lambda_n$ 对应了 n 个线性无关的特征向量 $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$，满足 $\boldsymbol{A}\boldsymbol{c}_i = \lambda_i \boldsymbol{c}_i, i = 1, \ldots, n$，定义 n 阶方阵

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_n \end{bmatrix}$$

则 $\boldsymbol{C}$ 可逆，且有

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{C} &= \boldsymbol{A}\begin{bmatrix} \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}\boldsymbol{c}_1 & \cdots & \boldsymbol{A}\boldsymbol{c}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\boldsymbol{c}_1 & \cdots & \lambda_1\boldsymbol{c}_n \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} = \boldsymbol{C} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}
\end{aligned}$$

$$\Longrightarrow \boldsymbol{A} = \boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^{-1}, \quad \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

特征值分解：条件

- 上式称为矩阵 $\boldsymbol{A}$ 的特征值分解，亦称 $\boldsymbol{A}$ 可对角化
- 可对角化的一个常见充分条件：若 $\boldsymbol{A}$ 有 n 个互不相同的特征值 $\lambda_1,\ldots,\lambda_n$，满足 $\lambda_i \neq \lambda_j, \forall i \neq j$，则各特征值对应的特征向量一定线性无关

证明 反设 $\boldsymbol{c}_1,\ldots,\boldsymbol{c}_m$ 为具有线性相关性的最少个数的特征向量组，即存在 $\alpha_1,\ldots,\alpha_m \in \mathbb{C}$ 全不为 0（否则可以将系数为零的特征向量去掉），使得

$$\alpha_1\boldsymbol{c}_1 + \cdots + \alpha_n\boldsymbol{c}_n = \mathbf{0}$$

上式两端左乘 $\boldsymbol{A}$，得到

$$\alpha_1\lambda_1\boldsymbol{c}_1 + \cdots + \alpha_m\lambda_m\boldsymbol{c}_m = \mathbf{0}$$

将前一式乘 λ_m 再减去后一式，得

$$\alpha_1(\lambda_m - \lambda_1)\boldsymbol{c}_1 + \cdots + \alpha_m(\lambda_m - \lambda_m)\boldsymbol{c}_m = \mathbf{0}$$

此时得到了 $m-1$ 个线性相关的特征向量组，与 m 的最小性矛盾

本节内容

1 欧式空间

2 线性变换与矩阵

3 矩阵的逆与特征值

4 对称矩阵

实对称矩阵：定义与性质

- 给定 n 阶实方阵 $\boldsymbol{A}$，若 $\boldsymbol{A} = \boldsymbol{A}^{\mathsf{T}}$，则称 $\boldsymbol{A}$ 为实对称矩阵，简称对称矩阵
  - n 维随机向量 $[X_1, \ldots, X_n]^{\mathsf{T}}$ 的协方差矩阵 $\Sigma = [\mathrm{cov}(X_i, X_j)]_{1\le i,j\le n}$ 为实对称矩阵
- n 阶实对称矩阵 $\boldsymbol{A}$ 一定有 n 个实特征值 $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$，且不论特征值是否有重根，$\boldsymbol{A}$ 一定有 n 个线性无关的特征向量 $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n \in \mathbb{R}^n$，故 $\boldsymbol{A}$ 一定可对角化

$$\boldsymbol{A} = \boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^{-1}, \quad \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

- 更进一步，$\boldsymbol{A}$ 有 n 个彼此正交的单位特征向量，即 $\boldsymbol{c}_i \cdot \boldsymbol{c}_j = 0, \forall i \neq j$ 且 $||\boldsymbol{c}_i|| = 1, \forall 1 \le i \le n$，令 $\boldsymbol{C} = [\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n]$ 为 n 个列向量构成的方阵，计算可知

$$\boldsymbol{C}^{\mathsf{T}}\boldsymbol{C} = \begin{bmatrix} \boldsymbol{c}_1^{\mathsf{T}} \\ \vdots \\ \boldsymbol{c}_n^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{c}_1 \cdot \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_1 \cdot \boldsymbol{c}_n \\ \vdots & \ddots & \vdots \\ \boldsymbol{c}_n \cdot \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_n \cdot \boldsymbol{c}_n \end{bmatrix} = \boldsymbol{I}_n$$

实对称矩阵：定义与性质

- 由此可见，$\boldsymbol{A}$ 存在一组单位正交特征向量，且其构成的矩阵满足 $\boldsymbol{C}^{\mathsf{T}}\boldsymbol{C} = \boldsymbol{I}$，故 $\boldsymbol{C}^{-1} = \boldsymbol{C}^{\mathsf{T}}$
    - 满足上述性质的矩阵，称为正交矩阵
- 换言之，任意的实对称矩阵 $\boldsymbol{A}$，都可以通过正交矩阵 $\boldsymbol{C}$ 来对角化

$$\boldsymbol{A} = \boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^{\mathsf{T}}, \quad \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n), \boldsymbol{C}\boldsymbol{C}^{\mathsf{T}} = \boldsymbol{I}$$

正定二次型函数

- 给定 n 阶实对称矩阵 A，定义函数 $f:\mathbb{R}^n \to \mathbb{R}$ 如下
$$f(\boldsymbol{x}) = \boldsymbol{x}^\top A \boldsymbol{x}, \quad \boldsymbol{x}\text{为列向量}$$
  称 $f(\cdot)$ 为二次型 (quadratic form) 函数
- 若对任意 $\boldsymbol{x} \in \mathbb{R}^n$ 且 $\boldsymbol{x} \neq \mathbf{0}$，有 $f(\boldsymbol{x}) = \boldsymbol{x}^\top A \boldsymbol{x} > 0$，则称 $f(\cdot)$ 为正定 (positive definite) 二次型，称 A 为正定矩阵
- 若对任意 $\boldsymbol{x} \in \mathbb{R}^n$ 且 $\boldsymbol{x} \neq \mathbf{0}$，有 $f(\boldsymbol{x}) = \boldsymbol{x}^\top A \boldsymbol{x} \geq 0$，且存在 $\boldsymbol{y} \neq \mathbf{0}$ 使得 $f(\boldsymbol{y}) = 0$，则称 $f(\cdot)$ 为半正定 (semi-positive definite) 二次型，称 A 为半正定矩阵

正定矩阵的性质

给定 n 阶实对称矩阵 $\boldsymbol{A}$，下列性质相互等价

1. $\boldsymbol{A}$ 是正定矩阵
2. $\boldsymbol{x}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定二次型函数
3. $\boldsymbol{A}$ 的所有特征值 $\lambda_1, \dots, \lambda_n > 0$ 严格为正

对半正定矩阵，下列性质相互等价

1. $\boldsymbol{A}$ 是半正定矩阵
2. $\boldsymbol{x}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{x}$ 为半正定二次型函数
3. $\boldsymbol{A}$ 的部分特征值 $\lambda_1, \dots, \lambda_k > 0, k < n$ 严格为正，其余特征值 $\lambda_{k+1} = \cdots = \lambda_n = 0$

注 协方差矩阵至少是半正定矩阵；若 $X_1, \dots, X_n$ 互相线性独立，则协方差矩阵是正定矩阵