第二章 向量和矩阵基础

第 5 讲 矩阵的基本特征

黄定江

DaSE @ ECNU

djhuang@dase.ecnu.edu.cn

1. 5.1 行列式
2. 5.2 迹和二次型
3. 5.3 特征值与特征向量

5.1.1 二阶行列式

考虑一个二阶矩阵 $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

我们将其作用在

$\boldsymbol{x}^{(1)} = (0,1)^{\mathrm{T}}, \boldsymbol{x}^{(2)} = (0,0)^{\mathrm{T}}$

$\boldsymbol{x}^{(3)} = (1,1)^{\mathrm{T}}, \boldsymbol{x}^{(4)} = (1,0)^{\mathrm{T}}$

$$\boldsymbol{y}^{(1)} = \boldsymbol{A}\boldsymbol{x}^{(1)} = (a_{12}, a_{22})^{\mathrm{T}}$$
$$\boldsymbol{y}^{(2)} = \boldsymbol{A}\boldsymbol{x}^{(2)} = (0,0)^{\mathrm{T}}$$
$$\boldsymbol{y}^{(3)} = \boldsymbol{A}\boldsymbol{x}^{(3)} = (a_{11}+a_{12}, a_{21}+a_{22})^{\mathrm{T}}$$
$$\boldsymbol{y}^{(4)} = \boldsymbol{A}\boldsymbol{x}^{(4)} = (a_{11}, a_{21})^{\mathrm{T}}$$

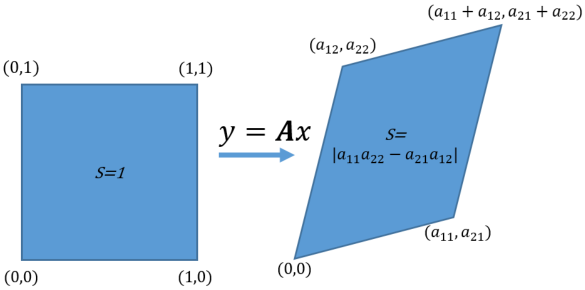

图 1: 矩阵 $\boldsymbol{A}$ 对正方形的作用效果。可以看到作用后的区域的面积为 $|a_{11}a_{22} - a_{21}a_{12}|$

5.1.1 二阶行列式

定义 1

对于二阶矩阵 $\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$，定义其**二阶行列式**为主对角元素乘积与副对角元素乘积的差，记为 $\det(\boldsymbol{A})=a_{11}a_{22}-a_{21}a_{12}$，或者记为 $|\boldsymbol{A}|=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{21}a_{12}$

例 1

给定一个方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2, \end{cases}$$

当 $a_{11}a_{22} - a_{12}a_{21} \neq 0$ 时，此方程组有唯一解，即

$$x_1 = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}, x_2 = \frac{a_{11}b_2 - a_{12}b_1}{a_{11}a_{22} - a_{12}a_{21}}.$$

利用行列式重新表述，当二阶行列式 $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$ 时，该方程组有唯一解，

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}}, x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{12} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}}.$$

排列

定义 2

n 个不同的元素排成一列，叫做这 n 个元素的**全排列**。一个排列中，如果一个大元素在小元素前，则称这两个数构成一个**逆序**。一个排列中存在的所有逆序的数目称为排列的**逆序数**。排列 $j_1, j_2, \cdots, j_n$ 的逆序数记为 $\tau(j_1, j_2, \cdots, j_n)$ 如果逆序数为奇数，称这个排列为**奇排列**；如果逆序数为偶数，称这个排列为**偶排列**。

例 2

- $\tau(3,2,1,4) = 3$，*3,2,1,4* 为奇排列
- $\tau(1,3,2,4) = 1$，*1,3,2,4* 为奇排列
- $\tau(3,1,2,4) = 2$，*3,1,2,4* 为偶排列

5.1.2 行列式

定义 3

n 阶行列式

$$\det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \qquad (1)$$

等于所有取自不同行、不同列的 n 个元素的乘积

$$a_{1j_1} \cdots a_{nj_n} \qquad (2)$$

的代数和，其中 $j_1, \cdots, j_n$ 是 $1, \cdots, n$ 的一个排列，当 $j_1, \cdots, j_n$ 是偶排列时，项(2)前面带正号；当 $j_1, \cdots, j_n$ 是奇排列时，项(2)前面带负号，

即

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \sum_{j_1,\cdots,j_n} (-1)^{\tau(j_1,\cdots,j_n)} a_{1j_1} \cdots a_{nj_n} \qquad (3)$$

其中 $\sum_{j_1,j_2,\cdots,j_n}$ 表示对 $1, 2, \cdots, n$ 的所有排列求和。n 阶行列式由 $n!$ 项组成。常把行列式(1)简记为 $D = |\boldsymbol{A}| = |a_{ij}|$ 或 $\det(a_{ij})$。

行列式的函数和几何意义

- 一个 $n \times n$ 的方阵 A 的行列式是一个函数，将矩阵空间 $\mathbb{R}^{n\times n}$ 中的矩阵 A 映射为 $\mathbb{R}$ 中的一个实数，记作 $\det(A)$ 或 $|A|$
- 3 维空间中，3 阶方阵的行列式对应于一个平行六面体的体积
- n 阶行列式实际上对应 n 维空间中图形的有向体积

行列式的性质

定义 4

n 阶行列式中，划去元素 a_{ij} 所在第 i 行和第 j 列元素，剩余的元素按原来的次序组成的 $n-1$ 阶行列式称为元素 a_{ij} 的**余子式**，记成 M_{ij}。令 $A_{ij}=(-1)^{i+j}M_{ij}$，称 A_{ij} 是元素 a_{ij} 的**代数余子式**。

行列式的按行 (列) 展开

$$
\begin{aligned}
\det(\boldsymbol{A}) &= \sum_{j=1}^{n} a_{ij}(-1)^{i+j}\det(M_{ij}) = \sum_{j=1}^{n} a_{ij}A_{ij} \\
&= \sum_{i=1}^{n} a_{ij}(-1)^{i+j}\det(M_{ij}) = \sum_{i=1}^{n} a_{ij}A_{ij}
\end{aligned}
$$

n 阶行列式等于其任一行 (列) 元素与对应的代数余子式两两乘积之和。

例 3

计算 $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 \\ 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{vmatrix}$

我们将这个行列式按列展开可得

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 \\ 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 1 & 1 \\ 2 & 0 & 2 \\ 0 & 1 & 1 \end{vmatrix}$$

再分别按行展开

$$\begin{vmatrix} 2 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 1 & 1 \\ 2 & 0 & 2 \\ 0 & 1 & 1 \end{vmatrix} = 2\begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} + 2\begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} + \begin{vmatrix} 0 & 2 \\ 1 & 1 \end{vmatrix} - \begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix} + \begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} = 2+4-2-2+2=4$$

克莱姆法则

定理 1

设线性方程组为

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},$$

我们记 $\boldsymbol{b}$ 为常数列，$|\boldsymbol{A}|_j$ 为用常数列 $\boldsymbol{b}$ 代替 $\boldsymbol{A}$ 中的第 j 列，其余列不变所得矩阵的行列式。则若 $|\boldsymbol{A}| \neq 0$，则线性方程组有唯一解，且

$$\boldsymbol{x}_1 = \frac{|\boldsymbol{A}|_1}{|\boldsymbol{A}|}, \boldsymbol{x}_2 = \frac{|\boldsymbol{A}|_2}{|\boldsymbol{A}|}, \ldots, \boldsymbol{x}_n = \frac{|\boldsymbol{A}|_n}{|\boldsymbol{A}|}$$

这一结论，我们称为**克莱姆法则**。

行列式的性质

性质 1

交换矩阵相邻两行 (或两列) 改变矩阵行列式的符号。

性质 2

交换矩阵两行 (或两列) 改变矩阵行列式的符号。

性质 3

行列式关于矩阵的每行 (或每列) 是线性的。

对于列的情况是

$$\begin{vmatrix} a_{11} & \dots & k_1 a_{1k} + k_2 b_{1k} & \dots & a_{1n} \\ a_{21} & \dots & k_1 a_{2k} + k_2 b_{2k} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & k_1 a_{nk} + k_2 b_{1k} & \dots & a_{nn} \end{vmatrix}$$

$$= k_1 \begin{vmatrix} a_{11} & \dots & a_{1k} & \dots & a_{1n} \\ a_{21} & \dots & a_{2k} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nk} & \dots & a_{nn} \end{vmatrix} + k_2 \begin{vmatrix} a_{11} & \dots & b_{1k} & \dots & a_{1n} \\ a_{21} & \dots & b_{2k} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & b_{1k} & \dots & a_{nn} \end{vmatrix}$$

性质 4

如果某一行 (列) 元素全为零，行列式为零.

推论 1

- 如果行列式有两行 (列) 相等，则行列式为零；
- 如果行列式中有一行 (列) 是另一行 (列)k 倍，则行列式为零；
- 如果将行列式的某一行 (列)k 倍加到另一行 (列)，则行列式的值不变。

定理 2

n 阶行列式任意一行 (列) 元素与另一行 (列) 相应元素的代数余子式的两两乘积之和等于零，即当 $i \neq k$ 时，有

$$\sum_{k=1}^{n} a_{ik}A_{jk} = \begin{cases} |\boldsymbol{A}|, & i = j, \\ 0, & i \neq j. \end{cases}$$

$$\sum_{k=1}^{n} a_{ki}A_{kj} = \begin{cases} |\boldsymbol{A}|, & i = j, \\ 0, & i \neq j. \end{cases}$$

行列式的性质

性质 5

设 $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$

- $\det(\boldsymbol{A}) = \det(\boldsymbol{A}^{\mathrm{T}})$
- $\det(a\boldsymbol{A}) = a^n \det(\boldsymbol{A})$
- $\det(\boldsymbol{A}\boldsymbol{B}) = \det(\boldsymbol{B}\boldsymbol{A}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$
- $\boldsymbol{A}$ 是可逆矩阵当且仅当 $\det(\boldsymbol{A}) \neq 0$
- 若 $\boldsymbol{A}$ 是可逆矩阵，则 $\det(\boldsymbol{A}^{-1}) = \frac{1}{\det(\boldsymbol{A})}$

伴随矩阵

定义 5

矩阵

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

称为 $\boldsymbol{A}$ 的**伴随矩阵**，其中伴随矩阵 $\boldsymbol{A}^*$ 的第 i 行第 j 列元素是矩阵的第 j 行第 i 列元素的代数余子式。

伴随矩阵

由定理2和伴随矩阵的定义：

定理 3

$$\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{A}^*\boldsymbol{A} = \begin{pmatrix} |\boldsymbol{A}| & 0 & \cdots & 0 \\ 0 & |\boldsymbol{A}| & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & |\boldsymbol{A}| \end{pmatrix} = |\boldsymbol{A}|\boldsymbol{I}$$

如果 $\boldsymbol{A}$ 可逆，则

$$\boldsymbol{A}^* = |\boldsymbol{A}|\boldsymbol{A}^{-1}, \quad \boldsymbol{A}^{-1} = \frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*$$

5.2.1 迹

定义 6

方阵 $A_{n\times n}$ 对角元素的和称为**迹**，记作 $\mathit{Tr}(A)=\sum_{i=0}^{n} a_{ii}$.

性质 6

$\boldsymbol{A},\boldsymbol{B}\in\mathbb{R}^{n\times n}$ 的迹满足以下性质：

- $\mathit{Tr}(\boldsymbol{A}+\boldsymbol{B})=\mathit{Tr}(\boldsymbol{A})+\mathit{Tr}(\boldsymbol{B})$
- $\mathit{Tr}(\alpha\boldsymbol{A})=\alpha\,\mathit{Tr}(\boldsymbol{A}),\alpha\in\mathbb{R}$
- $\mathit{Tr}(\boldsymbol{I}_n)=n$
- $\mathit{Tr}(\boldsymbol{A}\boldsymbol{B})=\mathit{Tr}(\boldsymbol{B}\boldsymbol{A})$

性质 7

- 迹的循环置换不变性，即

$$Tr(\boldsymbol{ABC}) = Tr(\boldsymbol{CAB}) = Tr(\boldsymbol{BCA})$$

  其中 $\boldsymbol{A} \in \mathbb{R}^{k\times l}, \boldsymbol{B} \in \mathbb{R}^{l\times m}, \boldsymbol{C} \in \mathbb{R}^{m\times n}$。

- 这个性质可以扩展到任意多矩阵的乘积即

$$Tr(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_n) = Tr(\boldsymbol{A}_n\boldsymbol{A}_1\cdots\boldsymbol{A}_{n-1}) = \cdots = Tr(\boldsymbol{A}_2\boldsymbol{A}_3\cdots\boldsymbol{A}_1) \tag{4}$$

- 作为(4)的特例，对于两个非方阵 $\boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{B} \in \mathbb{R}^{n\times m}$：

$$Tr(\boldsymbol{AB}) = Tr(\boldsymbol{BA})$$

推论 2

相似矩阵的迹是相等的，因为

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

5.2.2 二次型

定义 7

一个系数在数域 $\mathbb{K}$ 上的 $x_1, x_2, \cdots, x_n$ 的二次齐次多项式

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) = & a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n \\ & + a_{22}x_2^2 + 2a_{23}x_2x_3 + \cdots + 2a_{2n}x_2x_n \\ & + \cdots\cdots + a_{nn}x_n^2 \end{aligned} \tag{5}$$

称为数域 $\mathbb{K}$ 上的n **元二次型**，简称**二次型**，当 $\mathbb{K}$ 为 $\mathbb{R}$ 或 $\mathbb{C}$ 时，分别称为**实二次型**或**复二次型**。

5.2.2 二次型

定义 8

对称矩阵是转置和自己相等的矩阵，即

$$A = A^{\mathrm{T}} \tag{6}$$

二次型（5）的系数排成的对称矩阵

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

称为所给**二次型的矩阵**，其中 $a_{ij} = a_{ji}, i, j = 1, 2, \cdots, n$，若令 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$，则所给的二次型可表示为：

$$f(x_1, x_2, \cdots, x_n) = \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} = Tr(\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x})$$

二次型之间的变换

如果对 $\boldsymbol{x}$ 做可逆的线性变换，即 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}, \boldsymbol{y}=\boldsymbol{C}^{-1}\boldsymbol{x}$，那么 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}\boldsymbol{y}$。通过线性变换，我们把二次型 $f(x_1,x_2,\cdots,x_n)=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 化为新的二次型 $f=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}\boldsymbol{y}$，显然，新的二次型矩阵为 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$，仍然是对称矩阵。

定义 9

设 $\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶矩阵，若存在可逆矩阵 $\boldsymbol{C}$ 使得

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{B}$$

则称 $\boldsymbol{A}$ 合同于 $\boldsymbol{B}$（$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是合同矩阵），记作 $\boldsymbol{A}\simeq\boldsymbol{B}$。

合同矩阵的性质

性质 8

合同矩阵的性质：

- 反身性：$\boldsymbol{A} \simeq \boldsymbol{A}$
- 对称性：$\boldsymbol{A} \simeq \boldsymbol{B}$，则 $\boldsymbol{B} \simeq \boldsymbol{A}$
- 传递性：$\boldsymbol{A} \simeq \boldsymbol{B}$，$\boldsymbol{B} \simeq \boldsymbol{C}$，则 $\boldsymbol{A} \simeq \boldsymbol{C}$

同时满足反身性、对称性和传递性的关系叫做等价关系。等价矩阵、相似矩阵、合同矩阵都是矩阵的一种等价关系。

引理 1

对称矩阵 $\boldsymbol{A}$ 的下列变换都是合同变换:

- 交换 $\boldsymbol{A}$ 的第 i 行和第 j 行，再交换其第 i 列和第 j 列;
- 将 $\boldsymbol{A}$ 的第 i 行乘以非零常数 k，再将其第 i 列乘以 k;
- 将 $\boldsymbol{A}$ 的第 i 行乘以 k 加到第 j 行，再将其第 i 列乘以 k 加到第 j 列。

引理 2

设 $\boldsymbol{A}$ 是属于 $\mathbb{K}$ 上的非零对称矩阵，则必存在非奇异矩阵 $\boldsymbol{C}$，使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 的第 1 行第 1 个元素不为零。

二次型合同于对角矩阵

定理 4

设 $\boldsymbol{A}$ 是数域 $\mathbb{K}$ 上的 n 阶对称阵，则必然存在 $\mathbb{K}$ 上的 n 阶非奇异矩阵 $\boldsymbol{C}$，使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 为对角矩阵。即对称矩阵 $\boldsymbol{A}$ 必然合同于对角矩阵。

证明.

当 $\boldsymbol{A}$ 为 1 阶矩阵时，结论显然成立。假设对 k 阶矩阵，结论成立。对 $k+1$ 阶矩阵 $\boldsymbol{A}$，

$$\boldsymbol{A} \overset{\text{由引理 }2}{\simeq} \underbrace{\begin{pmatrix} a'_{11} & \cdots & a'_{1,k+1} \\ \vdots & \ddots & \vdots \\ a'_{k+1,1} & \cdots & a'_{k+1,k+1} \end{pmatrix}}_{\boldsymbol{A}\text{ 合同于左上角元 } a'_{11}\neq 0 \text{ 的矩阵}} \overset{\text{由引理 }1}{\simeq} \underbrace{\begin{pmatrix} a'_{11} & 0 & \cdots & 0 \\ 0 & b_{11} & \cdots & b_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{k1} & \cdots & b_{kk} \end{pmatrix}}_{\substack{\text{第 2 行减去第 1 行的 } a'_{12}/a'_{11} \text{ 倍，} \\ \text{第 2 列减去第 1 列的 } a'_{12}/a'_{11} \text{ 倍 } \cdots}} \overset{\text{由假设}}{\simeq} \underbrace{\begin{pmatrix} a'_{11} & \boldsymbol{0}_{1\times k} \\ \boldsymbol{0}_{k\times 1} & \boldsymbol{\Lambda}_{k\times k} \end{pmatrix}}_{k\text{ 阶子矩阵合同于对角矩阵}}$$

标准型

定理 5

数域 $\mathbb{K}$ 上任意一个二次型都可经过非退化的线性替换化为平方和

$$d_1 y_1^2 + d_2 y_2^2 + \cdots + d_n y_n^2$$

的形式，它称为所给二次型的**标准形**。

用初等变换法可以将二次型化为标准形。设二次型 $f = f(x_1, x_2, \cdots, x_n)$ 的矩阵为 $\boldsymbol{A}$，作初等变换

$$\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{I} \end{pmatrix} \xrightarrow[\text{对 } \boldsymbol{I} \text{ 只作其中的初等列变换}]{\text{对 } \boldsymbol{A} \text{ 作成对的初等行、列变换}} \begin{pmatrix} \boldsymbol{D} \\ \boldsymbol{C} \end{pmatrix}$$

其中 $\boldsymbol{D}$ 是对角矩阵 $\boldsymbol{D} = [d_1, d_2, \cdots, d_n]$，$\boldsymbol{C}$ 是非退化的线性替换矩阵，此时，$f = d_1 y_1^2 + d_2 y_2^2 + \cdots + d_n y_n^2$。

标准型

例 4

用初等变换法化二次型 $f(x_1,x_2,x_3)=x_1x_2+x_1x_3-3x_2x_3$ 为标准形。

解

$f(x_1,x_2,x_3)$ 的二次型矩阵为

$$\boldsymbol{A}=\begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & -3/2 \\ 1/2 & -3/2 & 0 \end{pmatrix},$$

$$\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{I} \end{pmatrix}=\begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & -3/2 \\ 1/2 & -3/2 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/4 & 0 \\ 0 & 0 & 3 \\ 1 & -1/2 & 3 \\ 1 & 1/2 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

标准型

$$\boldsymbol{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/4 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \boldsymbol{C} = \begin{pmatrix} 1 & -1/2 & 3 \\ 1 & 1/2 & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

线性替换为

$$\begin{cases} x_1 = y_1 - \frac{1}{2}y_2 + 3y_3, \\ x_2 = y_1 + \frac{1}{2}y_2 - y_3, \\ x_3 = y_3 \end{cases}$$

由此得 $f(x_1, x_2, x_3) = y_1^2 - \frac{1}{4}y_2^2 + 3y_3^2$

二次型分类

定义 10

设 $f(x_1,x_2,\cdots,x_n)=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为 n 元实二次型，若对任一组不全为零的实数 $c_1,c_2,\cdots,c_n$ 都有

- $f(c_1,c_2,\cdots,c_n)>0(<0)$，则称 $f(x_1,x_2,\cdots,x_n)$ 为**正定二次型（负定二次型）**，此时称 $\boldsymbol{A}$ 为**正定矩阵（负定矩阵）**。
- $f(c_1,c_2,\cdots,c_n)\geq 0(\leq 0)$，则称 $f(x_1,x_2,\cdots,x_n)$ 为**正半定二次型（负半定二次型）**，此时称 $\boldsymbol{A}$ 为**正半定矩阵（负半定矩阵）**。
- 若 $f(x_1,x_2,\cdots,x_n)$ 既不是正半定的，又不是负半定的，则称 $f(x_1,x_2,\cdots,x_n)$ 为**不定二次型**。

正定二次型（负定二次型）必是正半定二次型（负半定二次型）。

例 5

$$f(x_1, x_2, \cdots, x_n) = x_1^2 + x_2^2 + \cdots + x_n^2$$

是正定型，

$$f(x_1, x_2, \cdots, x_n) = -x_1^2 - x_2^2 - \cdots - x_n^2$$

是负定型。正定型最小值为 0，负定型最大值为 0。

利用合同矩阵判断正定性

由于一个二次型矩阵合同于对角矩阵，存在线性变换使得：

$$f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

如果 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 全部为正，那么对任意不为零向量的 $\boldsymbol{x}$，$f(\boldsymbol{x}) > 0$ 恒成立；
如果 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 全部为负，那么对任意不为零向量的 $\boldsymbol{x}$，$f(\boldsymbol{x}) < 0$ 恒成立。

5.3.1 矩阵的特征系

- 任何一个矩阵 $\boldsymbol{A}$ 代表一个线性映射。如果矩阵 $\boldsymbol{A}$ 是方阵，则它是一个线性变换
- 线性变换包含了旋转和拉伸变换等，对于拉伸变换：假设有向量 $\boldsymbol{\mu}$ 作为线性变换 $\boldsymbol{A}$ 的输入时，所产生的输出与输入只相差一个比例因子 λ，即

$$\boldsymbol{A}\boldsymbol{\mu} = \lambda\boldsymbol{\mu}, \boldsymbol{\mu} \neq \boldsymbol{0} \tag{7}$$

- 意味着输入向量在线性变换下能够保持方向不变，所以 $\boldsymbol{\mu}$ 刻画了在线性变换中固定方向的向量特征，而 λ 则可视为线性变换对向量 $\boldsymbol{\mu}$ 方向上的固定增益

5.3.1 矩阵的特征系

定义 11

给定 n 阶方阵 $\boldsymbol{A}$, 如果存在数 λ 和非零向量 $\boldsymbol{x}$，满足 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$，那么 λ 称为矩阵的**特征值**，$\boldsymbol{x}$ 称为 λ 对于矩阵 $\boldsymbol{A}$ 的**特征向量**。

若 $\boldsymbol{x}$ 是特征向量，$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$，则对 $k\neq 0$，$\boldsymbol{A}(k\boldsymbol{x})=\lambda(k\boldsymbol{x})$，$k\boldsymbol{x}$ 也是特征向量。因此方程:

$$(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\mathbf{0}$$

有无穷多解。需满足 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0$。

定义 12

$\lambda\boldsymbol{I}-\boldsymbol{A}$ 称为**特征矩阵**；行列式 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})$ 为 λ 的 n 次多项式，称为**特征多项式**；称方程 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0$ 为**特征方程**；特征方程的解为**特征根**，特征根也就是所求的特征值。

5.3.1 矩阵的特征系

定义 13

线性变换 $\boldsymbol{A}$ 的同一个特征值 λ_i 对应的特征向量构成的线性子空间，称为特征子空间 *(eigenspace)*，记为 $V_\lambda(\boldsymbol{A})$。

定义 14

上述定义的概念：特征值、特征向量、特征多项式、特征方程、特征根、特征子空间统称为**矩阵的特征系**。

定义 15

特征值的集合称为**矩阵的谱**，记为 $\lambda(\boldsymbol{A})$；矩阵的谱的绝对值的最大值 $r = \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|\}$ 叫做**矩阵谱半径**。

5.3.1 矩阵的特征系

例 6

根据定义求矩阵

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$$

的特征值和特征向量。

5.3.1 矩阵的特征系

矩阵 $\boldsymbol{A}$ 的特征方程为

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \begin{vmatrix} \lambda - 1 & -2 & -2 \\ -2 & \lambda - 1 & -2 \\ -2 & -2 & \lambda - 1 \end{vmatrix} = (\lambda + 1)^2(\lambda - 5) = 0,$$

故 $\boldsymbol{A}$ 的特征值为 $\lambda_1 = \lambda_2 = -1$（二重特征值），$\lambda_3 = 5$。

5.3.1 矩阵的特征系

对 $\lambda_1 = \lambda_2 = -1$，由 $(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \mathbf{0}$，得到方程

$$\begin{pmatrix} -2 & -2 & -2 \\ -2 & -2 & -2 \\ -2 & -2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

它有无穷多个解。

- 设 $x_2 = 1, x_3 = 0$，求出解为 $\boldsymbol{x} = [-1, 1, 0]^{\mathrm{T}}$，记为 $\boldsymbol{x}_1$，
- 设 $x_2 = 0, x_3 = 1$，求出解为 $\boldsymbol{x} = [-1, 0, 1]^{\mathrm{T}}$，记为 $\boldsymbol{x}_2$，
  - 则 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 是属于特征值 -1 的两个线性无关的特征向量，属于 -1 的全部特征向量为 $k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2, k_1, k_2 \in \mathbb{R}$。

同理，$\lambda_3 = 5$ 的一个特征向量为 $\boldsymbol{x}_3 = [1, 1, 1]^{\mathrm{T}}$，属于 5 的全部特征向量为 $k\boldsymbol{x}_3, k \in \mathbb{R}$。

5.3.2 矩阵特征值与特征向量的性质

性质 9

设 λ 是 $\boldsymbol{A}$ 的特征值，$\boldsymbol{x}_1, \boldsymbol{x}_2$ 都是 $\boldsymbol{A}$ 的属于 λ 的特征向量，则 $c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2$ 也是 $\boldsymbol{A}$ 的属于 λ 的特征向量。

性质 10

设 λ 是 $\boldsymbol{A}$ 的特征值，$\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的属于 λ 的特征向量，则

- $k\lambda$ 是 $k\boldsymbol{A}$ 的特征值 (k 为任意常数);
- λ^m 是 $\boldsymbol{A}^m$ 的特征值 (m 为正整数);
- 若 $\boldsymbol{A}$ 可逆，则 λ^{-1} 是 $\boldsymbol{A}^{-1}$ 的特征值。

5.3.2 矩阵特征值与特征向量的性质

性质 11

设 λ 是 $\boldsymbol{A}$ 的特征值，$\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的属于 λ 的特征向量，则 $f(\lambda)$ 是 $f(\boldsymbol{A})$ 的特征值，其中 f 是一多项式。

证明.

设 $f(\boldsymbol{A}) = c_n\boldsymbol{A}^n + c_{n-1}\boldsymbol{A}^{n-1} + \cdots + c_1\boldsymbol{A} + c_0\boldsymbol{I}$。$\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的属于 λ 的特征向量，则 $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$，得

$$
\begin{aligned}
f(\boldsymbol{A})\boldsymbol{x} &= (c_n\boldsymbol{A}^n + c_{n-1}\boldsymbol{A}^{n-1} + \cdots + c_1\boldsymbol{A} + c_0\boldsymbol{I})\boldsymbol{x} \\
&= c_n\boldsymbol{A}^n\boldsymbol{x} + c_{n-1}\boldsymbol{A}^{n-1}\boldsymbol{x} + \cdots + c_1\boldsymbol{A}\boldsymbol{x} + c_0\boldsymbol{I}\boldsymbol{x} \\
&= c_n\lambda^n\boldsymbol{x} + c_{n-1}\lambda^{n-1}\boldsymbol{x} + \cdots + c_1\lambda\boldsymbol{x} + c_0\boldsymbol{x} \\
&= (c_n\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0)\boldsymbol{x} = f(\lambda)\boldsymbol{x}
\end{aligned}
$$

证毕 $\square$

5.3.2 矩阵特征值与特征向量的性质

性质 12

矩阵 $\boldsymbol{A}$ 的特征值的和为矩阵 $\boldsymbol{A}$ 的迹，矩阵 $\boldsymbol{A}$ 的特征值的积为矩阵 $\boldsymbol{A}$ 的行列式。

证明.

设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值，则有

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_0$$

根据行列式展开规则和根与系数关系：

$$c_{n-1} = -(\sum_{i=1}^{n} a_{ii}) = -(\sum_{i=1}^{n} \lambda_i)$$

$$c_0 = (-1)^n \det(\boldsymbol{A}) = (-1)^n \prod_{i=1}^{n} \lambda_i$$

证毕 $\square$

5.3.2 矩阵特征值与特征向量的性质

性质 13

若 $\boldsymbol{A}$ 相似 $\boldsymbol{B}$，$\boldsymbol{B}=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}, \boldsymbol{A}=\boldsymbol{Q}\boldsymbol{B}\boldsymbol{Q}^{-1}$

- 相似变换不改变矩阵的特征值，即若 λ 是 $\boldsymbol{A}$ 的特征值，λ 也是 $\boldsymbol{B}$ 的特征值，
- 相似变换改变矩阵的特征向量，若 $\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的特征向量，$\boldsymbol{Q}^{-1}\boldsymbol{x}$ 是 $\boldsymbol{B}$ 的特征向量，
- 相似矩阵有相同的特征方程，反之不成立。

对称矩阵的对角化

定理 6

设 $\boldsymbol{A}$ 为 n 阶对称矩阵，则有正交矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$。其中 $\boldsymbol{\Lambda}$ 是以 $\boldsymbol{A}$ 的 n 个特征值为对角元的对角阵。

此定理证明较为复杂，在此不予证明。

该定理说明，对称矩阵，等价、合同、相似于一个对角矩阵。

矩阵的基本特征小结

以矩阵作为变量的函数

- 秩
- 行列式
- 迹
- $\cdots$

矩阵的基本特征

- 特征值
- 特征向量
- 特征空间
- 谱半径
- $\cdots$

如何利用以矩阵作为变量的函数来建模？如何计算大规模矩阵的特征值等？