第5章　实 二 次 型

二次型是线性代数中的一个重要内容，它起源于二次曲线和二次曲面的主轴问题，现在在物理的势能与动能、微分几何的曲面法曲率、经济的效用函数、统计的置信椭圆上都有应用.

5.1　二次型的化简

5.1.1　二次型的定义

平面上的二次曲线，可用方程

$$a_{11}x^2+2a_{12}xy+a_{22}y^2+2a_{13}x+2a_{23}y+a_{33}=0$$

表示. 三维空间中的二次曲面，可用方程

$$a_{11}x^2+2a_{12}xy+2a_{13}xz+a_{22}y^2+2a_{23}yz+a_{33}z^2+2a_{14}x+2a_{24}y+2a_{34}z+a_{44}=0$$

表示. 它们都涉及式子 $\sum\limits_{i,j=1}^{n}a_{ij}x_ix_j$，即所谓的二次型，而此类问题的处理则涉及二次型的化简、二次型的性质. 看一个具体例子：

例 5.1.1　二次曲线方程

$$a_{11}x^2+2a_{12}xy+a_{22}y^2+2a_{13}x+2a_{23}y+a_{33}=0 \tag{5.1}$$

表示或椭圆、或双曲线、或抛物线、或退化的点和直线.

解　题中的二次曲线方程可表示成

$$f(x,y)=(x,y,1)\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{12}&a_{22}&a_{23}\\a_{13}&a_{23}&a_{33}\end{pmatrix}\begin{pmatrix}x\\y\\1\end{pmatrix}=0.$$

记

$$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{12}&a_{22}&a_{23}\\a_{13}&a_{23}&a_{33}\end{pmatrix},$$

不妨设

$$\begin{pmatrix}a_{11}&a_{12}\\a_{12}&a_{22}\end{pmatrix}\neq\boldsymbol{O}.$$

若 $a_{12}=0$, 可令 $\theta=0$，否则令

$$\begin{cases} \sin 2\theta = \dfrac{2a_{12}}{\sqrt{(a_{11}-a_{22})^2+4a_{12}^2}}, \\ \cos 2\theta = \dfrac{a_{11}-a_{22}}{\sqrt{(a_{11}-a_{22})^2+4a_{12}^2}}. \end{cases}$$

计算可得

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

先做旋转变换

$$\begin{cases} x = x'\cos\theta - y'\sin\theta, \\ y = x'\sin\theta + y'\cos\theta, \end{cases}$$

此时方程为

$$g(x',y') = (x',y',1)\begin{pmatrix} \lambda_1 & 0 & b_{13} \\ 0 & \lambda_2 & b_{23} \\ b_{13} & b_{23} & b_{33} \end{pmatrix}\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = 0.$$

令

$$d_1 = \begin{cases} -\dfrac{b_{13}}{\lambda_1}, & 若\lambda_1 \neq 0, \\ 0, & 若\lambda_1 = 0, \end{cases} \qquad d_2 = \begin{cases} -\dfrac{b_{23}}{\lambda_2}, & 若\lambda_2 \neq 0, \\ 0, & 若\lambda_2 = 0. \end{cases}$$

再作平移变换

$$\begin{cases} x' = x'' + d_1, \\ y' = y'' + d_2, \end{cases}$$

方程可化为

$$h(x'',y'') = \begin{cases} \lambda_1 x''^2 + \lambda_2 y''^2 + \lambda_3 = 0, & \lambda_1,\lambda_2 \neq 0, \\ 2b_{13}x'' + \lambda_2 y''^2 + \lambda_3 = 0, & \lambda_1 = 0, \lambda_2 \neq 0, \\ \lambda_1 x''^2 + 2b_{23}y'' + \lambda_3 = 0, & \lambda_1 \neq 0, \lambda_2 = 0. \end{cases}$$

若 λ_1,λ_2 同号，但与 $-\lambda_3$ 异号，则方程 (5.1) 为矛盾方程；
若 λ_1,λ_2 同号，但 $\lambda_3=0$，则方程 (5.1) 退化为一个点；
若 $\lambda_1,\lambda_2,-\lambda_3$ 同号，则方程 (5.1) 表示椭圆；
若 λ_1,λ_2 异号，$\lambda_3=0$，则方程 (5.1) 退化为两条直线；
若 λ_1,λ_2 异号，$\lambda_3\neq 0$，则方程 (5.1) 表示双曲线；
若 $\lambda_1=0$ 或 $\lambda_2=0$，则方程 (5.1) 表示抛物线. □

定义 1. 数域 F 上含有 n 个实变量 $x_1,x_2,\ldots,x_n$ 的二次齐次多项式 齐次多项

$$\begin{aligned} f(x_1,x_2,\ldots,x_n) = a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{1n}x_1x_n \\ + a_{22}x_2^2 + 2a_{23}x_2x_3 + \cdots + 2a_{2n}x_2x_n \\ + \cdots\cdots\cdots\cdots\cdots\cdots \\ + \quad a_{nn}x_n^2 \end{aligned} \tag{5.2}$$

称为数域 F 上的二次型，其中 $a_{ij}\in F$.

若 $F=R/C$, 则称 $f(x_1,x_2,\ldots,x_n)$ 为实二次型/复二次型。

称为**二次型**：若全部 $a_{ij} \in \mathbf{R}$，则称式 (5.2) 中的 f 为**实二次型**；若全部 $a_{ij} \in \mathbf{C}$，则称式 (5.2) 中的 f 为**复二次型**. 本书主要考虑**实**二次型.

如果在式 (5.2) 中令 $a_{ji} = a_{ij}$, 则二次型 f 可改写为

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) = & a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + \cdots + a_{1n}x_1x_n \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + a_{23}x_2x_3 + \cdots + a_{2n}x_2x_n \\ & + \cdots \\ & + a_{n1}x_nx_1 + a_{n2}x_nx_2 + a_{n3}x_nx_3 + \cdots + a_{nn}x_n^2. \end{aligned}$$

利用矩阵的乘法, 容易验证, 上式的右端等于

$$(x_1\ x_2\ \cdots\ x_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

于是, 若记

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

则二次型 (5.2) 可以简洁地记为

$$f(x_1, x_2, \cdots, x_n) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}, \tag{5.3}$$

称之为**二次型 f 的矩阵表示**, 其中 $\boldsymbol{A}$ 是对称矩阵.　$A^T = A$

显然, 给定二次型 (5.2), 则对称矩阵 $\boldsymbol{A}$ 是唯一确定的; 反之, 给定一个对称矩阵 $\boldsymbol{A}$, (5.2) 就唯一确定了一个二次型. 这表明, 对称矩阵 $\boldsymbol{A}$ 和二次型 f 是一一对应的. 显然，实对称矩阵和实二次型也是一一对应的.

定义2. **定义 5.1.2**(二次型的矩阵)　称 (5.3) 中的对称矩阵 $\boldsymbol{A}$ 为**二次型 f 的矩阵**，$\boldsymbol{A}$ 的秩称为**二次型 f 的秩**.

例 5.1.2　求实二次型

$$f(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 - 2x_1x_3 + 2x_2^2 + 2x_2x_3 + 2x_3^2$$

的矩阵表示.

解　容易写出 $f(x_1, x_2, x_3) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$, 其中

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

□

$f(x_1, x_2, \ldots, x_n)$ 的另一种写法：

$$\begin{aligned} f(x_1, x_2, \ldots, x_n) = & \boxed{a_{11}x_1^2} + \boxed{a_{12}x_1x_2} + \cdots + \boxed{a_{1n}x_1x_n} && \sum_{j=1}^n a_{1j}x_1x_j \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n && \sum_{j=1}^n a_{2j}x_2x_j \\ & + \cdots\cdots\cdots\cdots && \\ & + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 && \sum_{j=1}^n a_{nj}x_nx_j \end{aligned}$$

$$\sum_{j=1}^n \sum_{i=1}^n a_{ij}x_ix_j \qquad \sum_{i=1}^n a_{i1}x_ix_1 \quad \sum_{i=1}^n a_{i2}x_ix_2 \quad \cdots \quad \sum_{i=1}^n a_{in}x_ix_n \qquad \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j$$

同时证明了双重求和可换序!

任给数域 F 上的一个二次型

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j,$$

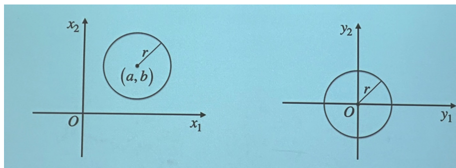
令 $A = (a_{ij})_{n\times n}$，则 $A^T = A$，且 $f(x_1, x_2, \ldots, x_n) = x^T A x$;

反之，任给数域 F 上的一个 n 阶对称阵 A, 则

$$f(x_1, x_2, \ldots, x_n) = x^T A x = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j$$

是数域 F 上的一个二次型 。

这表明数域 F 上的二次型与数域 F 上的对称阵一一对应。



$$f(x_1,x_2)=x_1^2-2ax_1+a^2+x_2^2-2bx_2+b^2$$

$$\begin{cases}x_1=y_1+a\\ x_2=y_2+b\end{cases}$$

$$f(x_1,x_2)=y_1^2+y_2^2$$

5.1.2　二次型的标准形

如同相似关系可以简化矩阵一样，也可以通过线性变换将二次型化成很简单的只有平方项的所谓标准形，从二次型矩阵来说就是对角矩阵.

定义 3. 替换

$$\begin{cases}x_1=p_{11}y_1+p_{12}y_2+\cdots+p_{1n}y_n\\ x_2=p_{21}y_1+p_{22}y_2+\cdots+p_{2n}y_n\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ x_n=p_{n1}y_1+p_{n2}y_2+\cdots+p_{nn}y_n\end{cases}$$

$f(x_1,x_2,\cdots,x_n)=x^{T}Ax \xlongequal{x=Py} f(y_1,y_2,\cdots,y_n)=y^{T}(P^{T}AP)y$

或

$$\begin{pmatrix}x_1\\ x_2\\ \cdots\\ x_n\end{pmatrix}=\underbrace{\begin{pmatrix}p_{11}&p_{12}&\cdots&p_{1n}\\ p_{21}&p_{22}&\cdots&p_{2n}\\ \cdots&\cdots&\cdots&\cdots\\ p_{n1}&p_{n2}&\cdots&p_{nn}\end{pmatrix}}_{P}\begin{pmatrix}y_1\\ y_2\\ \cdots\\ y_n\end{pmatrix}$$

称为由 $x_1,x_2,\ldots,x_n$ 到 $y_1,y_2,\ldots,y_n$ 的线性替换。

若 $|P|\neq 0$，则称线性替换 $x=Py$ 是非退化的；否则$x=Py$ 是退化的。

$$|\boldsymbol{P}|=\begin{vmatrix}c_{21}&c_{22}&\cdots&c_{2n}\\ \vdots&\vdots&&\vdots\\ c_{n1}&c_{n2}&\cdots&c_{nn}\end{vmatrix}\neq 0,$$

则称该线性变换为**非异线性变换**或**非退化线性变换**.

若 $|\boldsymbol{P}|=0$, 则称该线性变换为**奇异线性变换**或**退化线性变换**.

若 $\boldsymbol{P}$ 为正交矩阵, 则称该线性变换为**正交变换**.

定义4、**定义 5.1.4** (合同、合同变换)　设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是两个同阶方阵, 若存在一个可逆矩阵 $\boldsymbol{P}$, 使得有 $\boldsymbol{B}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$，则称 $\boldsymbol{A}$**合同于** $\boldsymbol{B}$. 称 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的**合同矩阵**, 而称 $\boldsymbol{P}$ 为 $\boldsymbol{A}$ 到 $\boldsymbol{B}$ 的**合同变换矩阵**.

容易证明, 矩阵的合同关系是一个等价关系，即满足下列三个性质:

(1) 自反性: $\boldsymbol{A}$ 与 $\boldsymbol{A}$ 合同;

(2) 对称性: 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同, 则 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 合同;

(3) 传递性: 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同, $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 合同, 则 $\boldsymbol{A}$ 与 $\boldsymbol{C}$ 合同.　(P可逆)

定义5、**定义 5.1.5** (二次型的标准形)　二次型 $f(x_1,x_2,\cdots,x_n)$ 经过非退化线性变换后得到一个只包含变量平方项的二次型 $d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2$，称为原二次型的**标准形**.

定理 5.1.1　*存在非退化的线性变换将实二次型化为标准形，且平方项系数可以任意次序排列；存在可逆矩阵将实对称矩阵合同变换为实对角矩阵，且对角元素可以任意次序排列.*

证明　由于实二次型的非退化线性变换与实对称矩阵的合同变换等价，故我们只证明第二个结论：存在可逆矩阵将实对称矩阵合同变换为实对角矩阵.

设对称矩阵 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$，则由定理 4.5.3 可知存在正交矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n),$$

即 $\boldsymbol{P}$ 将 $\boldsymbol{A}$ 合同变换为实对角矩阵.

再令 $\boldsymbol{D}=(\boldsymbol{e}_{i_1}\ \ \boldsymbol{e}_{i_2}\ \ \cdots\ \ \boldsymbol{e}_{i_n})$, 其中 $(i_1,i_2,\cdots,i_n)$ 为 $(1,2,\cdots,n)$ 的一个排列，则有

$$(\boldsymbol{PD})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{PD})=\boldsymbol{D}^{\mathrm{T}}\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)\boldsymbol{D}=\mathrm{diag}(\lambda_{i_1},\lambda_{i_2},\cdots,\lambda_{i_n}),$$

定理 1. 数域 F 上的任意一个二次型均可以经过非退化的线性替换化为标准型。

证明. 设

$$f(x_1,x_2,\dots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j,$$
$$a_{ij}=a_{ji}.$$

我们对变量个数 n 归纳。

当 $n=1$ 时，$f(x_1)=a_{11}x_1^2$，结论显然成立。

假设结论对 $n-1$个变量的二次型成立。

<u>情形 1</u>. a_{ii} $(1\le i\le n)$ 中至少有一个不等于 0，不妨设 $a_{11}\neq 0$.

$$\begin{aligned}
f(x_1,x_2,\dots,x_n)= {}& a_{11}x_1^2+a_{12}x_1x_2+\cdots+a_{1n}x_1x_n\\
&+a_{21}x_2x_1+a_{22}x_2^2+\cdots+a_{2n}x_2x_n\\
&+\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&+a_{n1}x_nx_1+a_{n2}x_nx_2+\cdots+a_{nn}x_n^2
\end{aligned}$$

$$=a_{11}x_1^2+\sum_{j=2}^{n}a_{1j}x_1x_j+\sum_{i=2}^{n}a_{i1}x_ix_1+\sum_{i=2}^{n}\sum_{j=2}^{n}a_{ij}x_ix_j$$

$$=a_{11}x_1^2+2\sum_{j=2}^{n}a_{1j}x_1x_j+\sum_{i=2}^{n}\sum_{j=2}^{n}a_{ij}x_ix_j$$

配方

$$=a_{11}\left(x_1+\sum_{j=2}^{n}\frac{a_{1j}}{a_{11}}x_j\right)^2-\frac{1}{a_{11}}\left(\sum_{j=2}^{n}a_{1j}x_j\right)^2+\sum_{i=2}^{n}\sum_{j=2}^{n}a_{ij}x_ix_j$$

都是 $x_2,\cdots,x_n$ 的二次型.

设

线性替换

$$\begin{cases}y_1=x_1+\frac{a_{12}}{a_{11}}x_2+\cdots+\frac{a_{1n}}{a_{11}}x_n\\ y_2=\qquad x_2\\ \cdots\cdots\\ y_n=\qquad\qquad x_n\end{cases}\iff\begin{cases}x_1=y_1-\frac{a_{12}}{a_{11}}y_2-\cdots-\frac{a_{1n}}{a_{11}}y_n\\ x_2=\qquad y_2\\ \cdots\cdots\\ x_n=\qquad\qquad y_n\end{cases}$$

亦即

$$\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}=\underbrace{\begin{pmatrix}1&-\frac{a_{12}}{a_{11}}&\cdots&-\frac{a_{1n}}{a_{11}}\\0&1&\cdots&0\\\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&1\end{pmatrix}}_{P_1}\begin{pmatrix}y_1\\y_2\\\cdots\\y_n\end{pmatrix}$$

显然，$x=P_1y$ 是非退化的线性替换，且

$$f(x_1,x_2,\dots,x_n)=a_{11}\left(x_1+\sum_{j=2}^{n}\frac{a_{1j}}{a_{11}}x_j\right)^2+\sum_{i=2}^{n}\sum_{j=2}^{n}b_{ij}x_ix_j$$
$$x=P_1y\downarrow$$
$$f(x_1,x_2,\dots,x_n)=a_{11}y_1^2+\sum_{i=2}^{n}\sum_{j=2}^{n}b_{ij}y_iy_j.$$

显然，

$$\sum_{i=2}^{n}\sum_{j=2}^{n}b_{ij}y_iy_j$$

是 $n-1$个变量的二次型。由归纳假定，存在非退化的线性替换

$$\begin{pmatrix}y_2\\y_3\\\cdots\\y_n\end{pmatrix}=\begin{pmatrix}c_{22}&c_{23}&\cdots&c_{2n}\\c_{32}&c_{33}&\cdots&c_{3n}\\\cdots&\cdots&\cdots&\cdots\\c_{n2}&c_{n3}&\cdots&c_{nn}\end{pmatrix}\begin{pmatrix}z_2\\z_3\\\cdots\\z_n\end{pmatrix}$$

使得

$$\sum_{i=2}^{n}\sum_{j=2}^{n}b_{ij}y_iy_j=d_2z_2^2+\cdots+d_nz_n^2.$$

令

$$\begin{pmatrix}y_1\\y_2\\\cdots\\y_n\end{pmatrix}=\underbrace{\begin{pmatrix}1&0&\cdots&0\\0&c_{22}&\cdots&c_{2n}\\\cdots&\cdots&\cdots&\cdots\\0&c_{n2}&\cdots&c_{nn}\end{pmatrix}}_{P_2}\begin{pmatrix}z_1\\z_2\\\cdots\\z_n\end{pmatrix}.$$

则 $y=P_2z$ 是非退化的线性替换，且

$$f(x_1,x_2,\dots,x_n)=a_{11}y_1^2+\sum_{i=2}^{n}\sum_{j=2}^{n}b_{ij}y_iy_j.$$
$$y=P_2z\downarrow$$
$$f(x_1,x_2,\dots,x_n)=a_{11}z_1^2+d_2z_2^2+\cdots+d_nz_n^2.$$

若令 $P=P_1P_2$，则 $x=Pz$，亦即

$$\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}=\begin{pmatrix}1&-\frac{a_{12}}{a_{11}}&\cdots&-\frac{a_{1n}}{a_{11}}\\0&1&\cdots&0\\\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&1\end{pmatrix}\begin{pmatrix}1&0&\cdots&0\\0&c_{22}&\cdots&c_{2n}\\\cdots&\cdots&\cdots&\cdots\\0&c_{n2}&\cdots&c_{nn}\end{pmatrix}\begin{pmatrix}z_1\\z_2\\\cdots\\z_n\end{pmatrix},$$

是非退化的线性替换，且

$$f(x_1,x_2,\dots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$$
$$x=Pz\downarrow$$
$$f(x_1,x_2,\dots,x_n)=a_{11}z_1^2+d_2z_2^2+\cdots+d_nz_n^2.$$

<u>情形 2</u>. $a_{ii}=0,\ 1\le i\le n$，但存在 $a_{1j}\neq 0$，$2\le j\le n$. 不妨设 $a_{12}\neq 0$.

令

$$\begin{cases}x_1=y_1+y_2\\x_2=y_1-y_2\\x_3=\qquad\qquad y_3\\\cdots\cdots\cdots\cdots\\x_n=\qquad\qquad\qquad y_n\end{cases}$$

亦即

$$\begin{pmatrix}x_1\\x_2\\x_3\\\cdots\\x_n\end{pmatrix}=\begin{pmatrix}1&1&0&\cdots&0\\1&-1&0&\cdots&0\\0&0&1&\cdots&0\\\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&1\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\\\cdots\\y_n\end{pmatrix}.$$

则

$$f(x_1,x_2,\dots,x_n)=2a_{12}(y_1+y_2)(y_1-y_2)+\cdots=2a_{12}y_1^2-2a_{12}y_2^2+\cdots$$

这样，**情形 2** 就转化为 **情形 1**.

<u>情形 3</u>. $a_{ii}=0$, $1\le i\le n$，且 $a_{1j}=0$，$2\le j\le n$.

此时 $f(x_1,x_2,\dots,x_n)$ 是只含有变量 $x_2,x_3,\dots,x_n$ 的 $n-1$元二次型，由归纳假定，存在非退化的线性替换

$$\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}=\begin{pmatrix}1&0&\cdots&0\\0&p_{22}&\cdots&p_{2n}\\\cdots&\cdots&\cdots&\cdots\\0&p_{n2}&\cdots&p_{nn}\end{pmatrix}\begin{pmatrix}y_1\\y_2\\\cdots\\y_n\end{pmatrix}$$

使 $f(x_1,x_2,\cdots,x_n)$ 为标准二次型、

故对角元素可任意排列. □

由定理 5.1.1 的证明可得化实二次型为标准形的最自然的方法，即用正交矩阵将实对称矩阵对角化的方法. 现将这种方法的计算步骤描述如下.

设 $\boldsymbol{A}$ 为实二次型 $f(\boldsymbol{x})$ 的矩阵.

(1) 求解矩阵 $\boldsymbol{A}$ 的特征方程

$$|\lambda \boldsymbol{E}-\boldsymbol{A}|=0,$$

解得特征值 $\lambda=\lambda_1,\lambda_2,\cdots,\lambda_r$;

(2) 对每一个特征值 $\lambda=\lambda_i(s_i$ 重), 求出齐次线性方程组

$$(\lambda_i\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta}$$

的基础解系 (即特征向量的极大无关组)$\boldsymbol{\xi}_{i_1},\cdots,\boldsymbol{\xi}_{i_{s_i}}$. 并标准正交化为

$$\boldsymbol{\eta}_{i_1},\cdots,\boldsymbol{\eta}_{i_{s_i}};$$

(3) 将标准正交化的特征向量作为列构成正交矩阵

$$\boldsymbol{P}=(\boldsymbol{\eta}_1\ \ \cdots\ \ \boldsymbol{\eta}_n),$$

则非退化线性变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 将实二次型 $f(\boldsymbol{x})$ 化为标准形

$$f(\boldsymbol{x})=g(\boldsymbol{y})=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y},$$

其中 $\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\cdots,\lambda_1,\lambda_2,\cdots,\lambda_2,\cdots,\lambda_r,\cdots,\lambda_r)$.

下面我们来看几个化标准形的例子.

例 5.1.3　用正交变换将实二次型

$$f(x_1,x_2,x_3)=x_1^2-2x_1x_2+8x_1x_3+x_2^2+8x_2x_3-4x_3^2$$

化为标准形, 并给出相应的线性变换.

解　该实二次型的矩阵形式为

$$f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x},\boldsymbol{A}=\begin{pmatrix}1&-1&4\\-1&1&4\\4&4&-4\end{pmatrix},\ \boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}.$$

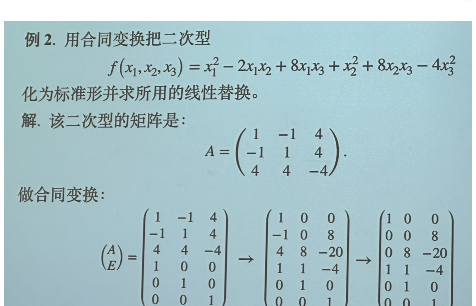
求 $\boldsymbol{A}$ 的特征值和标准正交的特征向量. 由

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1&1&-4\\1&\lambda-1&-4\\-4&-4&\lambda+4\end{vmatrix}=(\lambda-2)(\lambda-4)(\lambda+8)=0$$

解得特征值为 $\lambda=2,4,-8$.

对 $\lambda=2$，可求得单位特征向量 $\boldsymbol{\eta}_1=\dfrac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\\0\end{pmatrix}$.

对 $\lambda=4$，可求得单位特征向量 $\boldsymbol{\eta}_2=\dfrac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}$.



例 2. 用合同变换把二次型

$$f(x_1,x_2,x_3)=x_1^2-2x_1x_2+8x_1x_3+x_2^2+8x_2x_3-4x_3^2$$

化为标准形并求所用的线性替换。

解. 该二次型的矩阵是:

$$A=\begin{pmatrix}1&-1&4\\-1&1&4\\4&4&-4\end{pmatrix}.$$

做合同变换:

$$\begin{pmatrix}A\\E\end{pmatrix}=\begin{pmatrix}1&-1&4\\-1&1&4\\4&4&-4\\1&0&0\\0&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}1&0&0\\-1&0&8\\4&8&-20\\1&1&-4\\0&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}1&0&0\\0&0&8\\0&8&-20\\1&1&-4\\0&1&0\\0&0&1\end{pmatrix}$$

对 $\lambda=-8$，可求得单位特征向量 $\boldsymbol{\eta}_3=\dfrac{1}{\sqrt{6}}\begin{pmatrix}-1\\-1\\2\end{pmatrix}$.

令

$$\boldsymbol{P}=(\boldsymbol{\eta}_1\ \boldsymbol{\eta}_2\ \boldsymbol{\eta}_3)=\begin{pmatrix}-\dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{3}} & -\dfrac{1}{\sqrt{6}}\\ \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{3}} & -\dfrac{1}{\sqrt{6}}\\ 0 & \dfrac{1}{\sqrt{3}} & \dfrac{2}{\sqrt{6}}\end{pmatrix},$$

则有 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}=\boldsymbol{E}$, $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}=\operatorname{diag}(2,4,-8)$. 故在线性变换

$$\begin{cases} x_1=-\dfrac{1}{\sqrt{2}}y_1+\dfrac{1}{\sqrt{3}}y_2-\dfrac{1}{\sqrt{6}}y_3,\\ x_2=\dfrac{1}{\sqrt{2}}y_1+\dfrac{1}{\sqrt{3}}y_2-\dfrac{1}{\sqrt{6}}y_3,\\ x_3=\dfrac{1}{\sqrt{3}}y_2+\dfrac{2}{\sqrt{6}}y_3 \end{cases}$$

下，原实二次型化成的标准形为

$$g(y_1,y_2,y_3)=2y_1^2+4y_2^2-8y_3^2.$$

□

例 5.1.4　用正交变换将实二次型

$$f(x_1,x_2,x_3)=2x_1x_2+2x_1x_3-2x_2x_3$$

化为标准形, 并给出相应的线性变换.

解　该实二次型的矩阵形式为

$$f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax},\ \boldsymbol{A}=\begin{pmatrix}0&1&1\\1&0&-1\\1&-1&0\end{pmatrix},\ \boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}.$$

求 $\boldsymbol{A}$ 的特征值和标准正交的特征向量. 由

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda&-1&-1\\-1&\lambda&1\\-1&1&\lambda\end{vmatrix}=\begin{vmatrix}\lambda-1&-1&-1\\\lambda-1&\lambda&1\\0&1&\lambda\end{vmatrix}=(\lambda-1)^2(\lambda+2)=0$$

解得特征值为 $\lambda=1$(二重)，-2.

对 $\lambda=1$，可求得相互正交的单位特征向量 $\boldsymbol{\eta}_1=\dfrac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}$, $\boldsymbol{\eta}_2=\dfrac{1}{\sqrt{6}}\begin{pmatrix}1\\-1\\2\end{pmatrix}$.

对 $\lambda=-2$，可求得单位特征向量 $\boldsymbol{\eta}_3=\frac{1}{\sqrt{3}}\begin{pmatrix}-1\\1\\1\end{pmatrix}$.

令

$$\boldsymbol{P}=(\boldsymbol{\eta}_1\ \boldsymbol{\eta}_2\ \boldsymbol{\eta}_3)=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}}\end{pmatrix},$$

则有 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}=\boldsymbol{E},\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\,(1,1,-2)$. 故在线性变换

$$\begin{cases}x_1=\frac{1}{\sqrt{2}}y_1+\frac{1}{\sqrt{6}}y_2-\frac{1}{\sqrt{3}}y_3,\\ x_2=\frac{1}{\sqrt{2}}y_1-\frac{1}{\sqrt{6}}y_2+\frac{1}{\sqrt{3}}y_3,\\ x_3=\frac{2}{\sqrt{6}}y_2+\frac{1}{\sqrt{3}}y_3\end{cases}$$

下，原实二次型化成的标准形为

$$g(y_1,y_2,y_3)=y_1^2+y_2^2-2y_3^2.$$

□

除了用正交变换将实二次型化标准形外, 对一般二次型也可以用配方法化为标准形, 描述如下.

设二次型为

$$f(x_1,x_2,\cdots,x_n)=a_{11}x_1^2+2a_{12}x_1x_2+\cdots+2a_{1n}x_1x_n+a_{22}x_2^2+2a_{23}x_2x_3+\cdots+a_{nn}x_n^2.$$

反复对可能出现的以下两种情况进行处理.

情况 1 式中有非零平方项. 例如若非零平方项为 $a_{11}x_1^2$, 则将式中所有含 x_1 的项配成一个平方项 $a_{11}\left(x_1+\frac{a_{12}}{a_{11}}x_2+\cdots+\frac{a_{1n}}{a_{11}}x_n\right)^2$, 并令非退化线性变换为

$$\begin{cases}y_1=x_1+\frac{a_{12}}{a_{11}}x_2+\cdots+\frac{a_{1n}}{a_{11}}x_n,\\ y_2=x_2,\\ \quad\cdots\cdots\\ y_n=x_n,\end{cases}$$

则可将原式化为不含 x_1 也不含 y_1 的交叉项的式子.

情况 2 式中无非零平方项. 这时我们可以用一个线性变换配出平方项. 例如，若有非

零交叉项为 $2a_{12}x_1x_2$, 则作如下非退化线性变换

$$\begin{cases} x_1 = y_1 + y_2, \\ x_2 = y_1 - y_2, \\ x_3 = y_3, \\ \cdots\cdots \\ x_n = y_n, \end{cases}$$

就可将原式化为含有 y_1, y_2 的平方项的式子. 再按情况 1 进行处理.

每配成一个平方项，就消去一个元素如与 x_1 相关的所有交叉项，直到一系列的变换将所有的交叉项均消去即成标准形. 配方法所得的非退化线性变换在实际操作中可在所有配方配成后一次性求出.

下面是用配方法化标准形的例子.

例 5.1.5　用配方法将例 5.1.3 中的二次型

$$f(x_1, x_2, x_3) = x_1^2 - 2x_1x_2 + 8x_1x_3 + x_2^2 + 8x_2x_3 - 4x_3^2$$

化为标准形并指出所用的线性变换.

解　逐次对一个平方项及与该平方项有关的交叉项进行配方，

$$\begin{aligned} f(x_1, x_2, x_3) &= x_1^2 - 2x_1x_2 + 8x_1x_3 + x_2^2 + 8x_2x_3 - 4x_3^2 \\ &= (x_1 - x_2 + 4x_3)^2 + 16x_2x_3 - 20x_3^2 \\ &= (x_1 - x_2 + 4x_3)^2 - 20\left(x_3 - \frac{2}{5}x_2\right)^2 + \frac{16}{5}x_2^2. \end{aligned}$$

从而立即得到原二次型的标准形为

$$g(y_1, y_2, y_3) = y_1^2 - 20y_2^2 + \frac{16}{5}y_3^2.$$

所用线性变换为

$$\begin{cases} y_1 = x_1 - x_2 + 4x_3, \\ y_2 = x_3 - \dfrac{2}{5}x_2, \\ y_3 = x_2. \end{cases}$$

□

例 5.1.6　用配方法将例 5.1.4 中的二次型

$$f(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 - 2x_2x_3$$

化为标准形并指出所用的线性变换.

解　该二次型不含平方项，先做一个非退化线性变换从交叉项中产生平方项. 令

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

则

$$\begin{aligned} f(x_1,x_2,x_3) &= 2x_1x_2+2x_1x_3-2x_2x_3 \\ &= 2(y_1+y_2)(y_1-y_2)+2(y_1+y_2)y_3-2(y_1-y_2)y_3 \\ &= 2y_1^2-2y_2^2+4y_2y_3 \\ &= 2y_1^2-2(y_2-y_3)^2+2y_3^2. \end{aligned}$$

从而求得原二次型的标准形为

$$g(z_1,z_2,z_3)=2z_1^2-2z_2^2+2z_3^2.$$

所用线性变换为

$$\begin{cases} z_1=y_1=0.5x_1+0.5x_2, \\ z_2=y_2-y_3=0.5x_1-0.5x_2-x_3, \\ z_3=y_3=x_3. \end{cases}$$

□

注 1　配方过程中所用的线性变换都必须是非退化的.

我们还可用初等变换为工具将二次型化为标准形. 描述如下.

设二次型矩阵为

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

对矩阵

$$\boldsymbol{B}=\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{E} \end{pmatrix}$$

做一次列初等变换后接着做一次相应的行初等变换，重复这种做法，直到得到如下形式的矩阵

$$\begin{pmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{P} \end{pmatrix},$$

其中 $\boldsymbol{\Lambda}$ 为对角矩阵. 因为

$$\begin{pmatrix} \boldsymbol{P}^{\mathrm{T}} & \\ & \boldsymbol{E} \end{pmatrix}\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{E} \end{pmatrix}\boldsymbol{P}=\begin{pmatrix} \boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} \\ \boldsymbol{P} \end{pmatrix}=\begin{pmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{P} \end{pmatrix},$$

所以我们可用非退化线性变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$, 将二次型化为标准形:

$$f(\boldsymbol{x})=g(\boldsymbol{y})=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y}.$$

上述变换称之为合同变换法. 实际使用时通常是先用若干次列初等变换将某行非对角元化为 0，再依次做同样次数的相应行初等变换，反复进行. 下面来看用合同变换法化二次型为标准形的例子.

定理 1 的矩阵表述：

当 $a_{11} \neq 0$ 时，

$$f(x_1, x_2, \ldots, x_n) = x^T A x \xrightarrow{x = P_1 y} f(x_1, x_2, \ldots, x_n) = a_{11} y_1^2 + \sum_{i=2}^{n} \sum_{j=2}^{n} b_{ij} y_i y_j .$$

其中

$$P_1 = \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} .$$

也就是说

$$f(x_1, x_2, \ldots, x_n) = x^T \underline{A} x \xrightarrow{x = P_1 y} f(x_1, x_2, \ldots, x_n) = y^T \underline{P_1^T A P_1} y$$

令

$$\alpha = (a_{12}\ a_{13}\ \cdots\ a_{1n}), \qquad A_1 = \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix} .$$

则

$$A = \begin{pmatrix} a_{11} & \alpha \\ \alpha^T & A_1 \end{pmatrix}, \qquad P_1 = \begin{pmatrix} 1 & -a_{11}^{-1}\alpha \\ 0 & E_{n-1} \end{pmatrix} .$$

于是

$$P_1^T A P_1 = \begin{pmatrix} 1 & 0 \\ -a_{11}^{-1}\alpha^T & E_{n-1} \end{pmatrix} \begin{pmatrix} a_{11} & \alpha \\ \alpha^T & A_1 \end{pmatrix} \begin{pmatrix} 1 & -a_{11}^{-1}\alpha \\ 0 & E_{n-1} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & \alpha \\ 0 & A_1 - a_{11}^{-1}\alpha^T\alpha \end{pmatrix} \begin{pmatrix} 1 & -a_{11}^{-1}\alpha \\ 0 & E_{n-1} \end{pmatrix} = \begin{pmatrix} a_{11} & 0 \\ 0 & \underline{A_1 - a_{11}^{-1}\alpha^T\alpha} \end{pmatrix}$$

$A_1 - a_{11}^{-1}\alpha^T\alpha$ 是 $n-1$阶对称矩阵。由归纳假定，
存在 $n-1$ 阶可逆阵 B 使得 $B^T(A_1 - a_{11}^{-1}\alpha^T\alpha)B = D$ 为对角形。令

$$P_2 = \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix},$$

则有

$$\underline{P_2^T P_1^T A P_1} P_2 = \begin{pmatrix} 1 & 0 \\ 0 & B^T \end{pmatrix} \begin{pmatrix} a_{11} & 0 \\ 0 & \underline{A_1 - a_{11}^{-1}\alpha^T\alpha} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} a_{11} & 0 \\ 0 & D \end{pmatrix}$$

$$P_1 = \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

理解：相当于做了$n-1$次初等列变换.

$$= \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -\frac{a_{13}}{a_{11}} & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & 0 & 0 & \cdots & -\frac{a_{1n}}{a_{11}} \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} .$$

定理 1′. 数域 F 上的任意一个对称矩阵 A，均存在可逆阵 P 使得 P^TAP 为对角形矩阵。

如何求非退化的线性替换 $x = Py$ 使得

$$f(x_1, x_2, \ldots, x_n) = d_1y_1^2 + d_2y_2^2 + \cdots + d_ny_n^2 \ ?$$

求非退化线性替换的合同变换法

由于可逆阵可以分解为初等矩阵的乘积，比如 $P = P_1P_2\cdots P_s$，根据定理 1′，P^TAP 为对角形矩阵，亦即

$$P_s^T \cdots P_2^T P_1^T A P_1 P_2 \cdots P_s$$

为对角形矩阵。这表明对称阵 A 可以经过对称的初等变换化为对角形。

$$A \to P_1^TAP_1 \to P_2^TP_1^TAP_1P_2 \to \cdots \to P_s^T \cdots P_2^TP_1^TAP_1P_2 \cdots P_s$$

$$\begin{array}{ccc}
\begin{pmatrix} A \\ E \end{pmatrix} & & \\
\begin{pmatrix} P_1^T & \\ & E \end{pmatrix} \Big\downarrow P_1 & & \Big\downarrow \\
\begin{pmatrix} P_1^TAP_1 \\ P_1 \end{pmatrix} & & \cdots\cdots\cdots\cdots\cdots \\
\begin{pmatrix} P_2^T & \\ & E \end{pmatrix} \Big\downarrow P_2 & & \begin{pmatrix} P_s^T & \\ & E \end{pmatrix} \Big\downarrow P_s \\
\begin{pmatrix} P_2^TP_1^TAP_1P_2 \\ P_1P_2 \end{pmatrix} & & \begin{pmatrix} P_s^T \cdots P_2^TP_1^TAP_1P_2 \cdots P_s \\ P_1P_2 \cdots P_s \end{pmatrix}
\end{array}$$

例 2. **例 5.1.7**　用合同变换法将例 5.1.3 中的二次型

$$f(x_1,x_2,x_3)=x_1^2-2x_1x_2+8x_1x_3+x_2^2+8x_2x_3-4x_3^2$$

化为标准形并指出所用的线性变换.

解　该二次型的矩阵形式为

$$f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x},\ \boldsymbol{A}=\begin{pmatrix}1&-1&4\\-1&1&4\\4&4&-4\end{pmatrix},\ \boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}.$$

做如下合同变换：

$$\begin{pmatrix}\boldsymbol{A}\\\boldsymbol{E}\end{pmatrix}=\begin{pmatrix}1&-1&4\\-1&1&4\\4&4&-4\\1&0&0\\0&1&0\\0&0&1\end{pmatrix}\xrightarrow[c_3-4c_1]{c_2+c_1}\begin{pmatrix}1&0&0\\-1&0&8\\4&8&-20\\1&1&-4\\0&1&0\\0&0&1\end{pmatrix}\xrightarrow[r_3-4r_1]{r_2+r_1}\begin{pmatrix}1&0&0\\0&0&8\\0&8&-20\\1&1&-4\\0&1&0\\0&0&1\end{pmatrix}$$

$$\xrightarrow{c_2+c_3}\begin{pmatrix}1&0&0\\0&8&8\\0&-12&-20\\1&-3&-4\\0&1&0\\0&1&1\end{pmatrix}\xrightarrow{r_2+r_3}\begin{pmatrix}1&0&0\\0&-4&-12\\0&-12&-20\\1&-3&-4\\0&1&0\\0&1&1\end{pmatrix}\xrightarrow{c_3-3c_2}\begin{pmatrix}1&0&0\\0&-4&0\\0&-12&16\\1&-3&5\\0&1&-3\\0&1&-2\end{pmatrix}$$

$$\xrightarrow{r_3-3r_2}\begin{pmatrix}1&0&0\\0&-4&0\\0&0&16\\1&-3&5\\0&1&-3\\0&1&-2\end{pmatrix}.$$

则在线性变换

$$\begin{cases}x_1=y_1-3y_2+5y_3,\\x_2=y_2-3y_3,\\x_3=y_2-2y_3\end{cases}$$

下，原二次型化成的标准形为

$$g(y_1,y_2,y_3)=y_1^2-4y_2^2+16y_3^2.$$

□

注 2　用例 5.1.7 的合同变换法还可以将一个复对称矩阵化为复对角矩阵，从而复二次型可以经合同变换化为标准形.

5.1.3　二次型的规范形

二次型的标准形不唯一，我们希望确定一个最简标准形，并且是唯一的，这就是所谓的规范形.

定义6. **定义 5.1.6**(实 (复) 二次型的规范形)　实二次型 $f(x_1,x_2,\cdots,x_n)$ 经过非退化的实线性变换得到如下形式的二次型

$$z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2, r\leqslant n,$$

称为原二次型的**实规范形**，r 称为该**二次型的秩**；复二次型 $f(x_1,x_2,\cdots,x_n)$ 经过非退化的复线性变换得到如下形式的二次型

$$z_1^2+z_2^2+\cdots+z_r^2, r\leqslant n,$$

称为原二次型的**复规范形**，r 称为该二次型的秩.

定理 3. 对任意的复二次型 $f(x_1,x_2,\ldots,x_n)=x^TAx$，存在非退化线性替换将其化为规范形 $z_1^2+z_2^2+\cdots+z_r^2$.

证明. 根据定理 1′，存在可逆矩阵 P_1 使得

$$P_1^TAP_1=\begin{pmatrix} d_1 & & & & & \\ & \ddots & & & & \\ & & d_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

为二次型矩阵的秩.

知存在可逆矩阵 $\boldsymbol{P}$ 使

令 $P_2=\mathrm{diag}(\frac{1}{\sqrt{d_1}}\cdots\frac{1}{\sqrt{d_r}}\ 1\cdots1)$，有

$$P_2^TP_1^TAP_1P_2=\begin{pmatrix} \frac{1}{\sqrt{d_1}} & & & & & \\ & \ddots & & & & \\ & & \frac{1}{\sqrt{d_r}} & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}^T\begin{pmatrix} d_1 & & & & & \\ & \ddots & & & & \\ & & d_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{d_1}} & & & & & \\ & \ddots & & & & \\ & & \frac{1}{\sqrt{d_r}} & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$$

$$=\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

$\cdot, \dfrac{1}{\sqrt{b_{rr}}},1,\cdots,1\Big)$

$\dfrac{1}{\sqrt{b_{rr}}},1,\cdots,1\Big)$

□

(惯性定理)

定理 4. 对任意的实二次型 $f(x_1,x_2,\ldots,x_n)=x^TAx$，存在非退化线性替换将其化为规范形。

$$z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2,$$

阵的

且规范形是唯一的，其中 r 是二次型的秩。

证明. 前半部分容易证明，下证唯一性。

设 $f(x_1,x_2,\dots,x_n)$ 经过非退化的线性替换 $x=By$ 化成

$$f(x_1,x_2,\dots,x_n)=y_1^2+\cdots+y_q^2-y_{q+1}^2-\cdots-y_r^2,$$

经过非退化的线性替换 $x=Cz$ 化成

$$f(x_1,x_2,\dots,x_n)=z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2.$$

于是,

$$y_1^2+\cdots+y_q^2-y_{q+1}^2-\cdots-y_r^2=z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2, \tag{1}$$

$$y=B^{-1}Cz.$$

令

$$B^{-1}C=\begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{pmatrix}.$$

则 $y=B^{-1}Cz$ 就是

$$\begin{cases} y_1=d_{11}z_1+d_{12}z_2+\cdots+d_{1n}z_n \\ y_2=d_{21}z_1+d_{22}z_2+\cdots+d_{2n}z_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n=d_{n1}z_1+d_{n2}z_2+\cdots+d_{nn}z_n \end{cases} \tag{2}$$

若$q<p$，则如下齐次线性方程组

$$\begin{cases} d_{11}z_1+\cdots+d_{1p}z_p+d_{1,p+1}z_{p+1}+\cdots+d_{1n}z_n=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ d_{q1}z_1+\cdots+d_{qp}z_p+d_{q,p+1}z_{p+1}+\cdots+d_{qn}z_n=0 \\ z_{p+1}=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ z_n=0 \end{cases} \tag{3}$$

有非零解。设

$$(z_1 \ \cdots \ z_p \ z_{p+1} \ \cdots \ z_n)^T = (k_1 \ \cdots \ k_p \ \underline{k_{p+1} \ \cdots \ k_n})^T$$

是 (3) 的一个非零解，则必有 $k_{p+1} = \cdots = k_n = 0$.

将此非零解代入 (2) 可得

$$(\underline{y_1 \ \cdots \ y_q} \ y_{q+1} \ \cdots \ y_n)^T = (\underline{0 \ \cdots \ 0} \ s_{q+1} \ \cdots \ s_n)^T.$$

$$y_1 = \cdots = y_q = 0.$$

将 $(0 \ \cdots \ 0 \ s_{q+1} \ \cdots \ s_n)^T$ 及 $(k_1 \ \cdots \ k_p \ 0 \ \cdots \ 0)^T$ 代入 (1) 可得

(1) 式左端 $= -s_{q+1}^2 - \cdots - s_r^2 \leq 0.$

(1) 式右端 $= k_1^2 + \cdots + k_p^2 > 0,$

这是个矛盾，故有 $q \geq p$. 同理可得 $q \leq p$，从而 $q = p$. □

定义7. **定义 5.1.7** (正惯性指数、负惯性指数) 若实二次型的实规范形为

$$z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2, \qquad r \leqslant n,$$

则称 p 为原二次型的**正惯性指数**；称 $r-p$ 为原二次型的**负惯性指数**.

推论1. **推论 5.1.4** 若实二次型矩阵 $\boldsymbol{A}$ 合同于对角矩阵 $\boldsymbol{B} = \operatorname{diag}(b_{11}, \cdots, b_{nn})$，则正对角元个数为实二次型的正惯性指数，负对角元个数为实二次型的负惯性指数，非零对角元个数为二次型的秩.

证明 设有可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B} = \operatorname{diag}(b_{11}, \cdots, b_{nn})$，设 $b_{11}, \cdots, b_{nn}$ 中 $b_{i_1 i_1}, \cdots, b_{i_p i_p}$ 为 p 个正对角元，$b_{i_{p+1} i_{p+1}}, \cdots, b_{i_r i_r}$ 为 $r-p$ 个负对角元，$b_{i_{r+1} i_{r+1}} = \cdots = b_{i_n i_n} = 0$. 令 $\boldsymbol{D} = \boldsymbol{P}\boldsymbol{C}\boldsymbol{\Lambda}$，其中

$$\boldsymbol{C} = (\boldsymbol{e}_{i_1} \cdots \boldsymbol{e}_{i_n}), \quad \boldsymbol{\Lambda} = \operatorname{diag}(1/\sqrt{|b_{i_1 i_1}|}, \cdots, 1/\sqrt{|b_{i_r i_r}|}, 1, \cdots, 1),$$

则

$$\boldsymbol{D}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{D} = \boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} \operatorname{diag}(b_{11}, \cdots, b_{nn}) \boldsymbol{C}\boldsymbol{\Lambda} = \operatorname{diag}(\boldsymbol{E}_p, -\boldsymbol{E}_{r-p}, \boldsymbol{O}_{n-r}),$$

故正惯性指数为 p，负惯性指数为 $r-p$，秩为 r. □

推论2. **推论 5.1.5** 实二次型矩阵 $\boldsymbol{A}$ 的正特征值个数为正惯性指数，负特征值个数为负惯性指数，非零特征值个数为二次型的秩.

证明 因为实二次型矩阵 $\boldsymbol{A}$ 为实对称矩阵，故存在正交矩阵 $\boldsymbol{P}$ 使得

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \operatorname{diag}(\lambda_1, \cdots, \lambda_n),$$

其中 $\lambda_1, \cdots, \lambda_n$ 为 $\boldsymbol{A}$ 的特征值，由推论 5.1.4 可知，其中正数个数为正惯性指数，负数个数为负惯性指数，非零个数为秩. 即正特征值个数为正惯性指数，负特征值个数为负惯性指数，非零特征值个数为秩. □

例 5.1.8 求实二次型

$$\begin{aligned} & f(x_1, x_2, x_3, x_4) \\ & = 2x_1^2 - 4x_1x_2 - 12x_1x_3 + 8x_1x_4 + 9x_2^2 + 18x_2x_3 - 10x_2x_4 + 23x_3^2 - 20x_3x_4 + 4x_4^2 \end{aligned}$$

的惯性指数.

解 用合同变换法求该二次型的标准形. 易知二次型矩阵为

$$\boldsymbol{A} = \begin{pmatrix} 2 & -2 & -6 & 4 \\ -2 & 9 & 9 & -5 \\ -6 & 9 & 23 & -10 \\ 4 & -5 & -10 & 4 \end{pmatrix}.$$

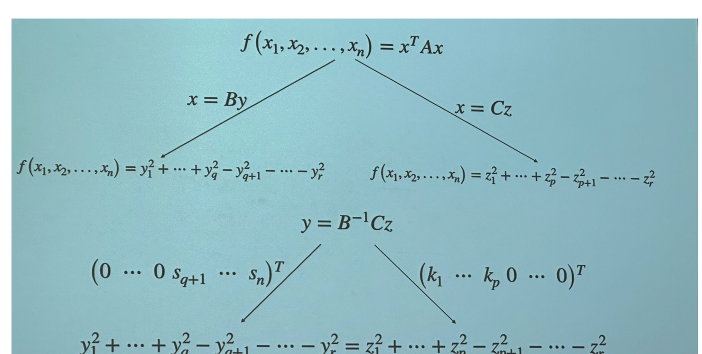
对任意的实二次型 $f(x_1, x_2, \dots, x_n) = x^T A x$,
如何按定义验证 $f(x_1, x_2, \dots, x_n) = x^T A x$ 是否正定?

$f(x_1, x_2, \dots, x_n) = x^T A x$ 的所有可能规范形

$$z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2,$$

$$z_1^2 + z_2^2 + \cdots + z_n^2$$

$$\begin{aligned} & z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2 \\ = & z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2 + 0z_{r+1}^2 + \cdots + 0z_n^2 \end{aligned}$$



对矩阵 $\boldsymbol{A}$ 对称地进行列和行的初等变换如下

$$\boldsymbol{A}\xrightarrow[c_4-2c_1]{\substack{c_2+c_1\\c_3+3c_1}}\begin{pmatrix}2&0&0&0\\-2&7&3&-1\\-6&3&5&2\\4&-1&2&-4\end{pmatrix}\xrightarrow[r_4-2r_1]{\substack{r_2+r_1\\r_3+3r_1}}\begin{pmatrix}2&0&0&0\\0&7&3&-1\\0&3&5&2\\0&-1&2&-4\end{pmatrix}$$

$$\xrightarrow[c_4+\frac{1}{7}c_2]{c_3-\frac{3}{7}c_2}\begin{pmatrix}2&0&0&0\\0&7&0&0\\0&3&\frac{26}{7}&\frac{17}{7}\\0&-1&\frac{17}{7}&-\frac{29}{7}\end{pmatrix}\xrightarrow[r_4+\frac{1}{7}r_2]{r_3-\frac{3}{7}r_2}\begin{pmatrix}2&0&0&0\\0&7&0&0\\0&0&\frac{26}{7}&\frac{17}{7}\\0&0&\frac{17}{7}&-\frac{29}{7}\end{pmatrix}$$

$$\xrightarrow{c_4-\frac{17}{26}c_3}\begin{pmatrix}2&0&0&0\\0&7&0&0\\0&0&\frac{26}{7}&0\\0&0&\frac{17}{7}&-\frac{149}{26}\end{pmatrix}\xrightarrow{r_4-\frac{17}{26}r_3}\begin{pmatrix}2&0&0&0\\0&7&0&0\\0&0&\frac{26}{7}&0\\0&0&0&-\frac{149}{26}\end{pmatrix}.$$

故正惯性指数为 3，负惯性指数为 1.　□

例 5.1.9　求实二次型

$$f(x_1,x_2,x_3)=x_1^2+4x_1x_2+6x_1x_3+4x_2^2+4x_2x_3+x_3^2$$

的惯性指数.

解　易知二次型矩阵为

$$\boldsymbol{A}=\begin{pmatrix}1&2&3\\2&4&2\\3&2&1\end{pmatrix}.$$

我们只要求 $\boldsymbol{A}$ 的特征值. 由

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1&-2&-3\\-2&\lambda-4&-2\\-3&-2&\lambda-1\end{vmatrix}=(\lambda+2)(\lambda^2-8\lambda+8)=0$$

解得特征值为 $\lambda=-2,4+2\sqrt{2},4-2\sqrt{2}$. 故该二次型的正惯性指数为 2，负惯性指数为 1. □

5.2　正定二次型

在所有实二次型中，有一类实二次型具有非负性，即只要 $\boldsymbol{x}\neq\boldsymbol{\theta}$，就有 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0$. 这类二次型及其矩阵在抽象空间的度量方面有着重要的应用.

定义8. **定义 5.2.1** (正定二次型、正定矩阵)　设 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为实二次型，若当实向量 $\boldsymbol{x}\neq\boldsymbol{\theta}$ 时都有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0$，则称 f 为**正定二次型**，称 $\boldsymbol{A}$ 为**正定矩阵**;

当 $\boldsymbol{x}\neq\boldsymbol{\theta}$ 时都有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}<0$，则称 f 为**负定二次型**，称 $\boldsymbol{A}$ 为**负定矩阵**;

当 $\boldsymbol{x} \neq \boldsymbol{\theta}$ 时都有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \geqslant 0$, 则称 f 为**半正定二次型**，称 $\boldsymbol{A}$ 为**半正定矩阵**;
当 $\boldsymbol{x} \neq \boldsymbol{\theta}$ 时都有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \leqslant 0$, 则称 f 为**半负定二次型**，称 $\boldsymbol{A}$ 为**半负定矩阵**.
注　有时为了强调正定矩阵的对称性, 也称对称正定矩阵.

例 5.2.1　说明 $\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ 为正定矩阵，$\boldsymbol{B} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ 为非正定矩阵.

定理 5. 设实二次型 $f(x) = x^T Ax$ 经过非退化的线性替换 $x = Py$ 化为 $g(y) = y^T P^T APy$. 则 $f(x) = x^T Ax$ 正定 $\Longleftrightarrow g(y) = y^T P^T APy$ 正定。

证明. $(\Rightarrow)$ $\forall y \neq 0$，由于 P 可逆，故 $Py \neq 0$. 于是

$$g(y) = y^T P^T APy = \left(\underline{Py}\right)^T A \left(\underline{Py}\right) = x^T Ax > 0.$$

即 $g(y) = y^T P^T APy$ 是正定的。

$(\Leftarrow)$ $\forall x \neq 0$，由于 P 可逆，故 $P^{-1}x \neq 0$. 于是

$$f(x) = x^T Ax = \left(\underline{P^{-1}x}\right)^T \underline{P^T AP} \left(\underline{P^{-1}x}\right) = y^T \underline{P^T AP} y > 0.$$

即 $f(x) = x^T Ax$ 是正定的。

显然，f 为负定二次型或半负定二次型的充要条件是 $-f$ 为正定二次型或半正定二次型. 因此，下面主要讨论正定和半正定二次型，也就是讨论正定和半正定矩阵.

定义9. **定义 5.2.2** (矩阵的顺序主子式和主子式)　矩阵 $\boldsymbol{A} = (a_{ij})_{n\times n}$ 的左上角 i 行 i 列 $(1 \leqslant i \leqslant n)$ 构成的行列式

$$\begin{vmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{ii} \end{vmatrix}$$

共有n个

称为矩阵 $\boldsymbol{A}$ 的i **阶顺序主子式**. 矩阵 $\boldsymbol{A}$ 的 $i_1, i_2, \cdots, i_k$ 行和 $i_1, i_2, \cdots, i_k$ 列 $(1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n)$ 的元素构成的行列式

$$\begin{vmatrix} a_{i_1i_1} & a_{i_1i_2} & \cdots & a_{i_1i_k} \\ a_{i_2i_1} & a_{i_2i_2} & \cdots & a_{i_2i_k} \\ \vdots & \vdots & & \vdots \\ a_{i_ki_1} & a_{i_ki_2} & \cdots & a_{i_ki_k} \end{vmatrix}$$

共有 $\sum_{i=1}^{n} C_n^i$ 个

称为矩阵 $\boldsymbol{A}$ 的k **阶主子式**.

定理6 **定理 5.2.1**　若 $\boldsymbol{A}$ 为 n 阶的实对称矩阵，则下列条件互为等价:

实对称阵若正定，其主对角线上元素全为正.

(1) $\boldsymbol{A}$ 为正定矩阵;

(2) $\boldsymbol{A}$ 的特征值均为正;

定理 7. 设 A 为 n 阶实对称矩阵。则

$$A \text{ 正定} \iff A \text{ 的各阶顺序主子式均为正。}$$

证明. ($\Rightarrow$) $\forall\ 1 \le k \le n$，令

$$A_k = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix}.$$

对任意的 $y = (y_1\ y_2\ \cdots\ y_k)^T \neq 0$,

易见 $x = (y\ 0)^T = (y_1\ \cdots\ y_k\ 0\ \cdots\ 0)^T \neq 0$ 且

$$y^T A_k y = x^T A x > 0.$$

这说明 A_k 是正定的。根据定理 6，A_k 的特征值均为正，从而 $|A_k| > 0$.

($\Leftarrow$) 对 n 归纳。$n = 1$ 时结论显然成立。

假设 $n = m$ 时结论成立。当 $n = m + 1$ 时，令

$$A = \begin{pmatrix} A_m & \alpha \\ \alpha^T & a_{nn} \end{pmatrix}.$$

由归纳假设，A_m 正定。故存在 m 阶可逆阵 B 使得 $B^T A_m B = E_m$. 令

$$P = \begin{pmatrix} B & 0 \\ 0 & 1 \end{pmatrix},$$

则有

$$P^T A P = \begin{pmatrix} B^T & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A_m & \alpha \\ \alpha^T & a_{nn} \end{pmatrix} \begin{pmatrix} B & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} E_m & B^T\alpha \\ \alpha^T B & a_{nn} \end{pmatrix}.$$

令

$$Q = \begin{pmatrix} E_m & -B^T\alpha \\ 0 & 1 \end{pmatrix},$$

则有

$$Q^T P^T A P Q = \begin{pmatrix} E_m & 0 \\ -\alpha^T B & 1 \end{pmatrix} \begin{pmatrix} E_m & B^T\alpha \\ \alpha^T B & a_{nn} \end{pmatrix} \begin{pmatrix} E_m & -B^T\alpha \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} E_m & 0 \\ 0 & a_{nn} - \alpha^T B B^T \alpha \end{pmatrix}$$

由于 $a_{nn} - \alpha^T B B^T \alpha = |Q^T||P^T||A||P||Q| = |P|^2|Q|^2|A| > 0$,

故 A 的正惯性指数为 n，从而 A 正定。

(3) $\boldsymbol{A}$ 的正惯性指数为 n;

定理7. (4) $\boldsymbol{A}$ 的各阶顺序主子式均为正.

证明 先证明 (1),(2),(3) 互为等价，设 $\boldsymbol{A} \in \mathbf{R}^{n\times n}$.

(1)$\Longrightarrow$(2)：因为实对称矩阵的特征值均为实数，并有实的特征向量，故设 λ 为 $\boldsymbol{A}$ 的一个实特征值，$\boldsymbol{\xi} \neq \boldsymbol{\theta}$ 为属于 λ 的实特征向量. 则 $\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\xi} = \boldsymbol{\xi}^{\mathrm{T}}(\lambda\boldsymbol{\xi}) = \lambda\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\xi}$，因为 $\boldsymbol{A}$ 为正定矩阵，故有 $\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\xi} > 0$，从而 $\lambda\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\xi} > 0$，而 $\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\xi} > 0$，所以可得 $\lambda > 0$，由 λ 的任意性，故有 $\boldsymbol{A}$ 的特征值均为正.

(2)$\Longrightarrow$(3)：由推论 5.1.5 可知 $\boldsymbol{A}$ 的正惯性指数为 n.

(3)$\Longrightarrow$(1)：因为 $\boldsymbol{A}$ 的正惯性指数为 n，故存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{E}$. 对任意 $\boldsymbol{x} \in \mathbf{R}^n, \boldsymbol{x} \neq \boldsymbol{\theta}$，令 $\boldsymbol{y} = \boldsymbol{P}^{-1}\boldsymbol{x}$，则 $\boldsymbol{y} \neq \boldsymbol{\theta}$，$\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}$，且

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{y} = \sum_{i=1}^{n} y_i^2 > 0,$$

由 $\boldsymbol{x}$ 的任意性，$\boldsymbol{A}$ 为正定矩阵.

(1)$\Longrightarrow$(4)：令

$$\boldsymbol{A}_i = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i} \\ a_{21} & a_{22} & \cdots & a_{2i} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ii} \end{pmatrix}, \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \end{pmatrix} \neq \boldsymbol{\theta}, \boldsymbol{x} = \begin{pmatrix} \boldsymbol{y} \\ 0 \end{pmatrix},\quad i = 1, 2, \cdots, n,$$

则 $\boldsymbol{x} \neq \boldsymbol{\theta}$ 且 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{A}_i\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0$，故 $\boldsymbol{A}_i$ 正定. 由 (1)、(2) 等价性知 $\boldsymbol{A}_i$ 的特征值均为正数，从而 $|\boldsymbol{A}_i| > 0$.

(4)$\Longrightarrow$(1)：用数学归纳法. 当 $n = 1$ 时，$\boldsymbol{A} = (a_{11})_{1\times 1}$，易知结论成立.

假设 $n = m$ 时结论成立. 当 $n = m + 1$ 时，令对称矩阵

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_m & \boldsymbol{u} \\ \boldsymbol{u}^{\mathrm{T}} & s \end{pmatrix} \in \mathbf{R}^{(m+1)\times(m+1)},$$

其中 $\boldsymbol{A}_m \in \mathbf{R}^{m\times m}, \boldsymbol{u} \in \mathbf{R}^m, s \in \mathbf{R}$，则 $\boldsymbol{A}_m$ 对称. 由 (4) 的条件可知 $\boldsymbol{A}_m$ 各顺序主子式均大于 0，根据归纳假设得 $\boldsymbol{A}_m$ 正定，进一步由 (1)、(2) 的等价性可得 $\boldsymbol{A}_m$ 的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 均大于 0。

因 $\boldsymbol{A}_m$ 正定，故存在正交矩阵 $\boldsymbol{P} \in \mathbf{R}^{m\times m}$ 使得

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}_m\boldsymbol{P} = \operatorname{diag}(\lambda_1, \cdots, \lambda_m) = \boldsymbol{\Lambda}.$$

于是我们有关系

$$\begin{pmatrix} \boldsymbol{P}^{\mathrm{T}} & \boldsymbol{0} \\ -\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}_m^{-1} & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{A}_m & \boldsymbol{u} \\ \boldsymbol{u}^{\mathrm{T}} & s \end{pmatrix} \begin{pmatrix} \boldsymbol{P} & -\boldsymbol{A}_m^{-1}\boldsymbol{u} \\ \boldsymbol{0} & 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Lambda} & \boldsymbol{0} \\ \boldsymbol{0} & s - \boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}_m^{-1}\boldsymbol{u} \end{pmatrix} = \boldsymbol{S}.$$

从上述式子可知 $\boldsymbol{A}$ 合同于对角矩阵 $\boldsymbol{S}$，并且 $|\boldsymbol{A}| = |\boldsymbol{\Lambda}|(s - \boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}_m^{-1}\boldsymbol{u})$.

由 (4) 的条件 $|\boldsymbol{A}|>0$ 及 $|\boldsymbol{\Lambda}|=\lambda_1\lambda_2\cdots\lambda_m>0$ 可得 $s-\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}_m^{-1}\boldsymbol{u}>0$，于是 $\boldsymbol{S}$ 的对角元均大于 0，由推论 5.1.4 知 $\boldsymbol{A}$ 的正惯性指数为 $m+1$，由 (1)、(3) 的等价性得 $\boldsymbol{A}$ 正定. □

用同样的方法可以证明关于半正定判别的定理:

定理8. **定理 5.2.2**　*若 $\boldsymbol{A}$ 为实对称矩阵，则下列条件互为等价:*

(1) *$\boldsymbol{A}$ 为半正定矩阵;*

(2) *$\boldsymbol{A}$ 的特征值均大于等于零;*

(3) *$\boldsymbol{A}$ 的正惯性指数为 $\mathrm{r}(\boldsymbol{A})$;*

(4) *$\boldsymbol{A}$ 的各阶主子式非负.*

证明作为习题五的题 24.

不能改为顺序主子式.

e.g. $\begin{pmatrix} 0 & & & \\ & -1 & & \\ & & \ddots & \\ & & & -1 \end{pmatrix}$

例 5.2.3　用顺序主子式判定

$$\boldsymbol{A}=\begin{pmatrix} 1 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}$$

是否正定.

解　显然，$\boldsymbol{A}$ 为实对称矩阵. 又

$$\det(1)=|(1)_{1\times1}|=1>0,\quad \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix}=2>0,$$

$$\begin{vmatrix} 1 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 3 \end{vmatrix}=\begin{vmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{vmatrix}=4>0,$$

各阶顺序主子式均 >0，故 $\boldsymbol{A}$ 为正定矩阵. □

例 5.2.4　用特征值判定

$$\boldsymbol{A}=\begin{pmatrix} 1 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}$$

是否正定.

解　显然，$\boldsymbol{A}$ 为实对称矩阵. 又

$$\begin{vmatrix} \lambda-1 & -1 & 1 \\ -1 & \lambda-3 & 1 \\ 1 & 1 & \lambda-3 \end{vmatrix}=\begin{vmatrix} \lambda-1 & -1 & 0 \\ -1 & \lambda-3 & \lambda-2 \\ 1 & 1 & \lambda-2 \end{vmatrix}=\begin{vmatrix} \lambda-1 & -1 & 0 \\ -2 & \lambda-4 & 0 \\ 1 & 1 & \lambda-2 \end{vmatrix}$$

$$=(\lambda-2)\begin{vmatrix} \lambda-1 & -1 \\ -2 & \lambda-4 \end{vmatrix}=(\lambda-2)(\lambda^2-5\lambda+2),$$

特征值为: $\lambda=2,\dfrac{5+\sqrt{17}}{2},\dfrac{5-\sqrt{17}}{2}$，均 >0，故 $\boldsymbol{A}$ 为正定矩阵. □

例 5.2.5　用标准形判定

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}$$

是否正定.

解　显然，A 为实对称矩阵. 用合同变换法化成标准形

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix} \xrightarrow[c_3+c_1]{c_2-c_1} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix} \xrightarrow[r_3+r_1]{r_2-r_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

对角线元素都 >0，得正惯性指数为 3，故 A 为正定. □

例8. **例 5.2.6**　t 取何值时，二次型

$$f(x_1,x_2,x_3) = 2x_1^2 + 2x_1x_2 + 4x_1x_3 + tx_2^2 - 2x_2x_3 + 3x_3^2$$

是正定二次型?

解　二次型矩阵为

$$A = \begin{pmatrix} 2 & 1 & 2 \\ 1 & t & -1 \\ 2 & -1 & 3 \end{pmatrix}.$$

用顺序主子式判别法:

$$\det(2) = 2 > 0, \quad \begin{vmatrix} 2 & 1 \\ 1 & t \end{vmatrix} = 2t - 1 > 0,$$

$$\begin{vmatrix} 2 & 1 & 2 \\ 1 & t & -1 \\ 2 & -1 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 0 \\ 1 & t-0.5 & -2 \\ 2 & -2 & 1 \end{vmatrix} = 2t - 9 > 0.$$

解两个关于 t 的不等式得到解集:$t > 4.5$, 故当 $t \in (4.5, +\infty)$ 时，该二次型为正定二次型. □

例9. **例 5.2.7**　证明: A 为正定矩阵当且仅当 A 有分解 $A = D^{\mathrm{T}}D$(D 可逆).　A正定 $\Leftrightarrow$ A合同于E

证明　设 A 为 n 阶矩阵，由定理 5.2.1 可知 A 为正定矩阵当且仅当 A 的正惯性指数为 n. 即存在可逆矩阵 P, 使得

$$P^{\mathrm{T}}AP = E.$$

令 $D = P^{-1}$, 则 D 可逆, 且有

$$D^{\mathrm{T}}D = (P^{-1})^{\mathrm{T}}P^{-1} = (P^{\mathrm{T}})^{-1}EP^{-1} = (P^{\mathrm{T}})^{-1}P^{\mathrm{T}}APP^{-1} = A.$$ □

例10. **例 5.2.8**　证明: 若实对称矩阵 A 满足关系式 $(A-E)(A-2E)=O$, 则 A 正定.

证明　展开关系式得　不能保证特征值为1或2.

$$A^2 - 3A + 2E = O.$$

设 λ 是 $\boldsymbol{A}$ 的特征值，$\boldsymbol{\xi}$ 是属于 λ 的特征向量，则有

$$(\boldsymbol{A}^2-3\boldsymbol{A}+2\boldsymbol{E})\boldsymbol{\xi}=(\lambda^2-3\lambda+2)\boldsymbol{\xi}=\boldsymbol{\theta}.$$

得

$$\lambda^2-3\lambda+2=(\lambda-1)(\lambda-2)=0$$

此即 $\boldsymbol{A}$ 的特征值或是 1 或是 2，均大于零，由定理5.2.1 可知 $\boldsymbol{A}$ 是正定矩阵. □

例 5.2.9 若 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 对称正定，证明：$(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta})^2\leqslant(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha})(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta})$，其中 $\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbf{R}^n$ 为任意列向量.

证明 当 $\boldsymbol{\alpha}=\boldsymbol{\theta}$ 时，结论显然.

当 $\boldsymbol{\alpha}\neq\boldsymbol{\theta}$ 时，令 $\boldsymbol{\xi}=t\boldsymbol{\alpha}+\boldsymbol{\beta}$，则有

$$\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\xi}=(t\boldsymbol{\alpha}+\boldsymbol{\beta})^{\mathrm{T}}\boldsymbol{A}(t\boldsymbol{\alpha}+\boldsymbol{\beta})=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}t^2+2\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta}t+\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta}.$$

由 $\boldsymbol{A}$ 的正定性，知 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}>0,\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\xi}\geqslant0$，所以有

$$(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha})t^2+(2\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta})t+(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta})\geqslant0.$$

再由二次方程根的判别准则得

$$\boldsymbol{\Delta}=(2\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta})^2-4(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha})(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta})\leqslant0,$$

即

$$(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta})^2\leqslant(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha})(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta}).$$ □

习 题 五

1. 用正交变换法求非退化线性变换，将下列二次型化为标准形:

(1) $f(x_1,x_2,x_3)=7x_1^2+4x_1x_2+4x_1x_3+3x_2^2+4x_2x_3+3x_3^2$;

(2) $f(x_1,x_2,x_3)=3x_1^2+2x_1x_2+4x_1x_3+3x_2^2+4x_2x_3+2x_3^2$;

(3) $f(x_1,x_2,x_3)=2x_1^2-4x_1x_2+4x_1x_3-x_2^2+2x_2x_3-x_3^2$;

(4) $f(x_1,x_2,x_3)=2x_1x_2+2x_1x_3+2x_2x_3$.

2. 用配方法求非退化线性变换，将下列二次型化为标准形:

(1) $f(x_1,x_2,x_3)=3x_1^2-2x_1x_2+4x_1x_3+2x_2^2+x_3^2$;

(2) $f(x_1,x_2,x_3)=3x_2^2+6x_2x_3+3x_3^2$;

(3) $f(x_1,x_2,x_3)=2x_1x_2+4x_1x_3-2x_2x_3$;

(4) $f(x_1,x_2,x_3,x_4)=(x_1+x_2)^2+(x_2+x_3)^2+(x_3+x_4)^2+(x_4+x_1)^2$.

3. 用合同变换法求非退化线性变换，将下列二次型化为标准形:

(1) $f(x_1,x_2,x_3)=x_1^2+4x_1x_2-6x_1x_3+4x_2^2+9x_3^2$.

(2) $f(x_1,x_2,x_3,x_4)=x_1^2-2x_1x_2+4x_1x_3-4x_1x_4+2x_2^2+2x_2x_3+4x_2x_4$
$+11x_3^2+4x_3x_4-6x_4^2$;

证明. 设 $A=(a_{ij})_{n\times n}$，$a_{ij}=-a_{ji}$.

对 n 归纳。当 $n=1,2$ 时容易验证命题成立。

设 $n\geq 3$.

<u>情形 1</u>. 反对角线上有一个元素不为零，不妨设 $a_{n1}\neq 0$.

此时 $a_{1n}=-a_{n1}\neq 0$.

分析:

$$\begin{pmatrix} 0 & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & \boxed{\begin{matrix} 0 & \cdots & a_{2,n-1} \\ \cdots & \cdots & \cdots \\ a_{n-1,2} & \cdots & 0 \end{matrix}} & & & a_{2n} \\ \cdots & & & & \cdots \\ a_{n-1,1} & & & & a_{n-1,n} \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & 0 \end{pmatrix}$$

令

$$P=\begin{pmatrix} 1 & -\frac{a_{n2}}{a_{n1}} & \cdots & -\frac{a_{n,n-1}}{a_{n1}} & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix},\quad Q=\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & -\frac{a_{12}}{a_{1n}} & \cdots & -\frac{a_{1,n-1}}{a_{1n}} & 1 \end{pmatrix},$$

则

$$P^T=\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -\frac{a_{n2}}{a_{n1}} & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -\frac{a_{n,n-1}}{a_{n1}} & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix},\quad Q^T=\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & -\frac{a_{12}}{a_{1n}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & -\frac{a_{1,n-1}}{a_{1n}} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

于是

$$Q^TP^TAPQ=\begin{pmatrix} 0 & 0 & a_{1n} \\ 0 & A_1 & 0 \\ a_{n1} & 0 & 0 \end{pmatrix}.$$

由于 $(Q^TP^TAPQ)^T=-Q^TP^TAPQ$，故 A_1 是反对称阵。

根据归纳假定，存在 $n-2$ 阶可逆阵 R 使得

$$R^TA_1R=\begin{pmatrix} & & 1 \\ & 0 & \\ -1 & & \end{pmatrix}.$$

令

$$C=\begin{pmatrix} 1 & & \\ & R & \\ & & 1 \end{pmatrix},$$

4. 证明反对称矩阵均合同于形式为

$$\begin{pmatrix} & & & & & & & & 1 \\ & & & & & & & \iddots & \\ & & & & & & 1 & & \\ & & & & & 0 & & & \\ & & & & \iddots & & & & \\ & & & 0 & & & & & \\ & & -1 & & & & & & \\ & \iddots & & & & & & & \\ -1 & & & & & & & & \end{pmatrix}$$

的矩阵.

5. 求下列二次型的正惯性指数和负惯性指数:

(1) $f(x_1,x_2,x_3)=4x_1^2+4x_1x_2+4x_1x_3+4x_2^2+4x_2x_3-x_3^2$;

(2) $f(x_1,x_2,x_3,x_4)=x_1^2+2x_1x_2+2x_2^2+2x_2x_3+3x_3^2+2x_3x_4+4x_4^2$;

(3) $f(x_1,\cdots,x_n)=\sum\limits_{i=1}^n x_i^2+2\sum\limits_{1\leqslant i<j\leqslant n}x_ix_j$;

(4) $f(x_1,\cdots,x_n)=\sum\limits_{i=1}^n(x_i-\bar{x})^2$，其中 $\bar{x}=\dfrac{\sum\limits_{i=1}^n x_i}{n}, n\geqslant 2$.

5(4). 因为

$$\sum_{i=1}^n(x_i-\bar{x})^2$$
$$=\sum_{i=1}^n\left(-\frac{1}{n}x_1-\cdots-\frac{1}{n}x_{i-1}+\frac{n-1}{n}x_i-\frac{1}{n}x_{i+1}-\cdots\frac{1}{n}x_n\right)^2$$
$$=\frac{n-1}{n}\sum_{i=1}^n x_i^2-\frac{2}{n}\sum_{1\leq i<j\leq n}x_ix_j$$

故 $f(x_1,x_2,\ldots,x_n)$ 的矩阵是:

$$A=\begin{pmatrix} \frac{n-1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & \frac{n-1}{n} & \cdots & -\frac{1}{n} \\ \cdots & \cdots & \cdots & \cdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & \frac{n-1}{n} \end{pmatrix}.$$

6. 设 $\boldsymbol{A}$ 为实对称矩阵，λ_1,λ_n 为 $\boldsymbol{A}$ 的最小和最大特征值，证明：对任意的 $t\in[\lambda_1,\lambda_n]$，存在单位向量 $\boldsymbol{x}$ 使得 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=t$.

6. 设 ξ_i 是属于 λ_i 的单位特征向量，$i=1,n$. 令

$$\xi(\theta)=\sin\theta\xi_1+\cos\theta\xi_n.$$

则 $\left|\xi(\theta)\right|=1$, 且

$$\xi(\theta)^TA\xi(\theta)=\left(\sin\theta\xi_1+\cos\theta\xi_n\right)^T\left(\sin\theta\lambda_1\xi_1+\cos\theta\lambda_n\xi_n\right)$$
$$=\sin^2\theta\lambda_1\xi_1^T\xi_1+\cos^2\theta\lambda_n\xi_n^T\xi_n$$
$$=\sin^2\theta\lambda_1+\cos^2\theta\lambda_n=\lambda_1+\cos^2\theta\left(\lambda_n-\lambda_1\right).$$

$\forall t\in[\lambda_1,\lambda_n]$, 令

$$\theta=\arccos\sqrt{\frac{t-\lambda_1}{\lambda_n-\lambda_1}},$$

则有 $\xi(\theta)^TA\xi(\theta)=t$.

7. 设 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$ 为列满秩矩阵，证明 $\boldsymbol{B}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 为对称正定矩阵.

8. 设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{R}^{n\times n}$ 对称正定，证明 $k_1\boldsymbol{A}+k_2\boldsymbol{B},k_1>0,k_2>0$ 也对称正定.

9. 设 $\boldsymbol{A}\in\mathbf{R}^{m\times m},\boldsymbol{B}\in\mathbf{R}^{n\times n}$ 对称正定，证明存在 $\boldsymbol{\xi}\in\mathbf{R}^{m+n},\boldsymbol{\xi}\neq\boldsymbol{\theta}$，使得 $\boldsymbol{\xi}^{\mathrm{T}}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{0} & -\boldsymbol{B}\end{pmatrix}\boldsymbol{\xi}=0$.

9. 设 $A=(a_{ij})_{m\times m}$, $B=(b_{ij})_{n\times n}$. 由于 A,B 正定，故有

$$a_{11}>0,\ b_{11}>0.$$

令

$$\xi=\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{a_{11}}}\\0\\\cdots\\0\\\frac{1}{\sqrt{b_{11}}}\\0\\\cdots\\0\end{pmatrix},$$

则有

$$\xi^T\begin{pmatrix}A & \\ & -B\end{pmatrix}\xi=x^TAx-y^TBy=1-1=0.$$

10. 设 $\boldsymbol{A}\in\mathbf{R}^{m\times m},\boldsymbol{B}\in\mathbf{R}^{n\times n}$ 均为对称矩阵，$\begin{pmatrix}\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{B}\end{pmatrix}$ 对称正定，证明 $\boldsymbol{A},\boldsymbol{B}$ 均为对称正定矩阵.

11. 设 $\boldsymbol{A}$ 为实对称正定矩阵，$\boldsymbol{P}$ 为可逆矩阵，用定义证明 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 也是对称正定矩阵；反之 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 正定，则方阵 $\boldsymbol{P}$ 可逆.

12. 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为正定矩阵，证明 $\boldsymbol{B}=(a_{ij}b_ib_j)_{n\times n}$ 当 $b_1,\cdots,b_n$ 全为非零常数时为正定.

13. 判别下列矩阵的正定性:

(1) $\begin{pmatrix}1 & 1 & 0\\ 1 & 3 & 1\\ 0 & 1 & 2\end{pmatrix}$;　(2) $\begin{pmatrix}1 & & & & -1\\ & 1 & & -1 & \\ & & 1 & & \\ & -1 & & 2 & \\ -1 & & & & 3\end{pmatrix}$.

14. 判别下列二次型的正定性:

(1) $f(x_1,x_2,x_3)=x_1^2+2x_1x_2-2x_1x_3+2x_2^2+2x_2x_3+6x_3^2$;

(2) $f(x_1,x_2,x_3)=2x_1^2-4x_1x_2+8x_1x_3+6x_2^2-4x_2x_3+4x_3^2$;

(3) $f(x_1,\cdots,x_n)=\sum\limits_{i=1}^{n}x_i^2+\sum\limits_{1\leqslant i<j\leqslant n}x_ix_j$;

(4) $f(x_1,\cdots,x_n)=2\sum\limits_{i=1}^{n}x_i^2+2\sum\limits_{i=1}^{n-1}x_ix_{i+1}$.

15. 讨论 t 取何值时下列二次型是正定的:

(1) $f(x_1,x_2,x_3)=10x_1^2+2tx_1x_2+6x_1x_3+5x_2^2+4x_2x_3+x_3^2$;

(2) $f(x_1,x_2,x_3)=(t+1)x_1^2-2x_1x_2+(t+2)x_2^2-2x_2x_3+(t+1)x_3^2$;

(3) $f(x_1,x_2,x_3)=t(x_1^2+x_2^2+x_3^2)+2x_1x_2-2x_2x_3+2x_1x_3+x_3^2$.

16. 设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 对称正定，证明: 对任意 $r>0$ 和 $1\leqslant i_1<i_2<\cdots<i_r\leqslant n$ 有

$$\begin{vmatrix} a_{i_1i_1} & a_{i_1i_2} & \cdots & a_{i_1i_r}\\ a_{i_2i_1} & a_{i_2i_2} & \cdots & a_{i_2i_r}\\ \vdots & \vdots & & \vdots\\ a_{i_ri_1} & a_{i_ri_2} & \cdots & a_{i_ri_r}\end{vmatrix}>0.$$

17. 证明：若 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 对称正定，则 $\boldsymbol{A}^*$ 也对称正定.

18. 证明 $\boldsymbol{A}$ 为正定矩阵的充要条件是 $\boldsymbol{A}$ 有分解 $\boldsymbol{A}=\boldsymbol{B}^2$，其中 $\boldsymbol{B}$ 为对称正定矩阵.

19. 设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 对称正定，$\boldsymbol{B}\in\mathbf{R}^{n\times n}$ 为对称矩阵，证明存在合同变换矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$ 与 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{BP}$ 均为对角矩阵.

20. 若 $\boldsymbol{A}$ 正定，证明 $|\boldsymbol{E}+\boldsymbol{A}|>1$.

21. 已知实对称矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^3-6\boldsymbol{A}^2+11\boldsymbol{A}-6\boldsymbol{E}=\boldsymbol{O}$. 证明 $\boldsymbol{A}$ 是正定矩阵.

22. 设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 对称正定, n 维非零向量组 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r\}$ 满足：$\boldsymbol{\beta_i}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta}_j=0, i\neq j,\ i,j=1,2,\cdots,r$，证明该向量组线性无关.

23. 设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 对称正定，函数

$$f(x_1,\cdots,x_n)=\begin{vmatrix}-\boldsymbol{A} & \boldsymbol{x}\\ \boldsymbol{x}^{\mathrm{T}} & 0\end{vmatrix},$$

其中 $\boldsymbol{x}=(x_1,\cdots,x_n)^{\mathrm{T}}$，证明函数 $f(x_1,\cdots,x_n)$ 当 n 为偶数时是正定二次型，而当 n 为奇数时是负定二次型.

24*. 证明书中定理 5.2.2.

18. (⇒) 若 A 正定，则存在正交阵 P 使得

$$P^TAP=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}=\begin{pmatrix}\sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n}\end{pmatrix}\begin{pmatrix}\sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n}\end{pmatrix},$$

其中 $\lambda_i>0, (1\le i\le n)$ 是 A 的特征值。于是

$$A=P\begin{pmatrix}\sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n}\end{pmatrix}P^TP\begin{pmatrix}\sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n}\end{pmatrix}P^T=B^2,$$

B 是正定矩阵。

19. 因为 A 正定，故存在可逆阵 D 使得 $A=D^TD$. 令

$$C=\left(D^{-1}\right)^TBD^{-1},$$

则 C 为实对称矩阵。故存在正交阵 Q 使得

$$Q^TCQ=\begin{pmatrix}d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n\end{pmatrix}.$$

令 $P=D^{-1}Q$，则 P 可逆且

$$P^TAP=\left(D^{-1}Q\right)^TD^TD\left(D^{-1}Q\right)=E,$$

$$P^TBP=\left(D^{-1}Q\right)^TB\left(D^{-1}Q\right)=Q^TCQ=\begin{pmatrix}d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n\end{pmatrix}.$$

第 6 章　线性空间与线性变换

6.1　线性空间的定义

6.1.1　线性空间的概念

数学学习的一个重要的特征是抽象化. 从直线、平面等具体的几何空间到向量空间, 为了准确地描述线性空间的概念, 需要引入数域的概念.

定义1、**定义 6.1.1** (数环)　设 $\Re$ 是非空数集, 其中任何两个数之和、差与积仍属于 $\Re$(即 $\Re$ 关于加、减、乘法运算是封闭的), 则称 $\Re$ 是一个**数环**.

由数环的定义可以看出:

(1) 任何数环 $\Re$ 必含有 0, 因为若 $a \in \Re$, 则 $a - a = 0 \in \Re$;

(2) 若 $a \in \Re$, 因为 $0 - a = -a \in \Re$, 则 $-a \in \Re$. 只含一个 0 的数集 $\Re = \{0\}$ 显然是个数环, 而且由上述数环的性质可知, $\Re = \{0\}$ 是最小的数环.

定义2、**定义 6.1.2** (数域)　若 K 是至少含有两个互异数的数环, 且其中任何两数 a 与 b 之商 $(b \neq 0)$ 仍属于 K, 则称 K 是一个**数域**.

根据数域的定义有:

(1) 数域关于加、减、乘、除 (分母不为零) 四则运算是封闭的;

(2) 任何数域 K 中必含有 0 与 1, 因为 K 中至少有一个非零数 $a, a/a = 1 \in K$;

(3) 若 $a \neq 0$, 则有 $1/a = a^{-1} \in K$.

下面给出一些常见的数环和数域:

全体整数构成一个数环; 全体有理数构成一个数域, 而且是最小的数域; 全体实数构成一个数域, 叫实数域, 记为 $\mathbf{R}$; 全体复数构成一个数域, 叫复数域, 记为 $\mathbf{C}$.

读者可自行验证: 若集合 S 由形如 $a + b\sqrt{2}$ (其中 a, b 是有理数) 的数的全体构成, 则 S 是一个数域, 且包含了有理数域.

不一定是数集

定义3、**定义 6.1.3** (线性空间)　设 V 是一个非空集合, K 是一个数域, V 满足以下两个条件:

(1) 在 V 中定义一个封闭的加法运算, 即当 $x, y \in V$ 时, 有唯一的 $z = x + y \in V$, 且此加法运算满足下面 4 条性质:

1) $x + y = y + x$ (交换律);

2) $x + (y + z) = (x + y) + z$ (结合律);

3) 存在**零元素** $\mathbf{0} \in V$, 对 V 中任一元素 x 都有 $x + \mathbf{0} = x$;

(逆元素) 4) 存在**负元素**: 对任一元素 $x \in V$, 存在一个元素 $y \in V$, 使得 $x + y = \mathbf{0}$, 称 y 为 x 的负元素 (或相反元素), 记为 $-x$, 即 $x + (-x) = \mathbf{0}$.

(2) 在 V 中定义一个封闭的数乘运算, 即当 $x \in V, \lambda \in K$ 时, 有唯一的 $\lambda x \in V$, 且此数乘运算满足下面 4 条性质:

定义加法

1) $(\lambda + \mu)x = \lambda x + \mu x$ (分配律);

V

V中元素叫做向量

K

K中元素是数字

2) $\lambda(x \oplus y) = \lambda x \oplus \lambda y$ (数因子分配律);

3) $\lambda(\mu x) = (\overset{\text{正常乘法}}{\lambda\mu})x$ (结合律);

4) $1 \cdot x = x$.

其中 x, y, z 是 V 中的任意元素, λ, μ 是数域 K 中的任意数, 1 是数域 K 中的单位数.

我们称 V 是数域 K 上的**线性空间**, 记为 $V(K)$. 当 K 是实数域 $\mathbf{R}$ 时, 称 V 为**实线性空间**; 当 K 是复数域 $\mathbf{C}$ 时, 称 V 为**复线性空间**.

为掌握线性空间这个重要的概念, 我们给出一些例子.

例 1. 实数域 R 上全体 n 维向量，对通常意义下向量加法和数乘构成实数域 R 上的线性空间。

$V = R^n$
$\alpha = (a_1\ a_2\ \cdots\ a_n)^T$
$\beta = (b_1\ b_2\ \cdots\ b_n)^T$

$K = R$
k, ℓ

加法:
$\alpha + \beta = (a_1 + b_1\ \cdots\ a_n + b_n)^T$
$\alpha + \beta = \beta + \alpha$
$\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$
$\alpha + 0 = \alpha$
$\alpha + (-\alpha) = 0$

数乘:
$k \cdot \alpha = (k \cdot a_1\ \cdots\ k \cdot a_n)^T$
$(k + \ell) \cdot \alpha = k \cdot \alpha + \ell \cdot \alpha$
$k \cdot (\alpha + \beta) = k \cdot \alpha + k \cdot \beta$
$k \cdot (\ell \cdot \alpha) = (k \cdot \ell) \cdot \alpha$
$1 \cdot \alpha = \alpha$

例 2. 数域 K 上全体 $m \times n$ 矩阵，对通常意义下矩阵加法和数乘构成数域 K 上的线性空间。

V
$A = (a_{ij})_{m\times n}$
$B = (b_{ij})_{m\times n}$

K
k, ℓ

加法: $A + B = (a_{ij} + b_{ij})_{m\times n}$
$A + B = B + A$
$A + (B + C) = (A + B) + C$
$A + O = A$
$A + (-A) = O$

数乘: $k \cdot A = (k \cdot a_{ij})_{m\times n}$
$(k + \ell) \cdot A = k \cdot A + \ell \cdot A$
$k \cdot (A + B) = k \cdot A + k \cdot B$
$k \cdot (\ell \cdot A) = (k \cdot \ell) \cdot A$
$1 \cdot A = A$

例 3. 设 $V = R^+$为正实数集，$K = R$ 为实数域。
V 中的加法和数乘定义为:

$$a \oplus b = ab,\ \forall a, b \in V = R^+,$$
$$\lambda \circ a = a^\lambda,\ \forall a \in V = R^+,\ \lambda \in K = R.$$

证明 $V = R^+$ 对这两种运算构成数域 $K = R$ 上的线性空间。

证明. 首先验证两种运算的封闭性:

$$\forall a, b \in V,\quad a \oplus b = ab \in V,$$
$$\forall a \in V, \lambda \in K,\quad \lambda \circ a = a^\lambda \in V.$$

其次验证两种运算各自满足4条性质:

(1) $a \oplus b = a \cdot b = b \cdot a = b \oplus a$;

(2) $(a \oplus b) \oplus c = (a \cdot b) \oplus c = a \cdot b \cdot c = a \cdot (b \cdot c) = a \oplus (b \oplus c)$;

(3) $\forall a \in V,\ a \oplus 1 = a \cdot 1 = a$, 即 1 是 V 的零元素;

(4) $\forall a \in V,\ a \oplus \frac{1}{a} = a \cdot \frac{1}{a} = 1$, 即 a 的负元素是 $1/a$

(5) $\lambda \circ (a \oplus b) = \lambda \circ (a \cdot b) = (a \cdot b)^\lambda = a^\lambda \cdot b^\lambda = (\lambda \circ a) \oplus (\lambda \circ b)$;

(6) $(\lambda + \mu) \circ a = a^{\lambda+\mu} = a^\lambda \oplus a^\mu = (\lambda \circ a) \oplus (\mu \circ a)$;

(7) $\lambda \circ (\mu \circ a) = \lambda \circ (a^\mu) = (a^\mu)^\lambda = a^{\mu\cdot\lambda} = a^{\lambda\cdot\mu} = (\lambda \cdot \mu) \circ a$;

(8) $1 \circ a = a^1 = a$.

因此， V 构成数域 K 上的线性空间。

的集合, 在 P 上定义通常意义下的向量加法和

$$\lambda \circ \boldsymbol{a} = \mathbf{0}, \qquad \lambda \in K, \quad \boldsymbol{a} \in P.$$

虽然 P 对两种运算封闭, 但由于 $1 \circ \boldsymbol{a} = \mathbf{0} \neq \boldsymbol{a}$, 故 P 不是线性空间.

可以举出各种各样不同的线性空间的例子, 由于线性空间是向量空间的抽象和推广, 为几何直观, 也可以把线性空间叫做向量空间, 其中的每个元素也称为向量. 但要注意到, 这里

的向量不一定是前几章中用有序数组表示的向量, 它以考察的对象 (数学的、物理的等) 为向量.

另外, 对同一个集合, 若定义不同的线性运算 (即线性空间上的加法与数乘运算), 就构成不同的线性空间. 线性运算是线性空间的本质属性, 它反映了线性空间中元素之间的代数结构.

6.1.2　线性空间的性质

由线性空间的定义可以得到线性空间的下述基本性质:

性质 1　线性空间的零元素是唯一的.

证明　设 $\mathbf{0}_1, \mathbf{0}_2$ 是线性空间 V 的两个零元素, 则

$$\mathbf{0}_1 = \mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_2 + \mathbf{0}_1 = \mathbf{0}_2,$$

即零元素唯一. □

性质 2　线性空间中任一元素的负元是唯一的.

证明　设 x_1, x_2 均为 $x \in V$ 的负元素, 则

$$x_1 = x_1 + (x + x_2) = (x_1 + x) + x_2 = x_2.$$ □

性质 3. 设 $0, 1, k \in K$, $\alpha, -\alpha, 0 \in V$. 则

(1) $0\alpha = 0$;

(2) $(-1)\alpha = -\alpha$;

(3) $k0 = 0$;

(4) $k\alpha = 0 \Rightarrow k = 0$　或　$\alpha = 0$.

证明. (1) $\forall \beta \in V$,

$$\begin{aligned}\beta + 0\alpha &= \beta + 0 + 0\alpha = \beta + (-\alpha + \alpha) + 0\alpha \\ &= \beta + (-\alpha) + (\alpha + 0\alpha) = \beta + (-\alpha) + (1 + 0)\alpha \\ &= \beta + (-\alpha) + \alpha = \beta + 0 = \beta.\end{aligned}$$

(2) 由于

$$\alpha + (-1)\alpha = (1 + (-1))\alpha = 0\alpha = 0,$$

根据性质 2, $(-1)\alpha = -\alpha$.

(3) $k0 = k(00) = (k0)0 = 00 = 0$.

(4) 若 $k \neq 0$, 则

$$\alpha = 1\alpha = \left(\frac{1}{k}k\right)\alpha = \frac{1}{k}(k\alpha) = \frac{1}{k}0 = 0.$$

以 $\lambda\mathbf{0} = \mathbf{0}$.

$$x = 1x = \left(\frac{1}{\lambda}\lambda\right)x = \frac{1}{\lambda}(\lambda x) = \frac{1}{\lambda}\mathbf{0} = \mathbf{0},$$

矛盾, 所以 $\lambda \neq 0$ 与 $x \neq \mathbf{0}$ 不能同时成立. □

只含一个元素的线性空间称为**零空间**, 显见, 仅有的那个元素是零元素.

上述性质看上去似乎很显然, 为什么还要予以证明呢? 因为一般的线性空间内的加法和数乘运算常常是以非常抽象的方式给出的, 与我们熟知的通常的运算相差甚远, 所以, 严格意义而言, 必须从逻辑上加以证明.

6.2　线性空间的基、维数与坐标

6.2.1　基与坐标

线性空间形形色色, 为讨论及运算的方便, 提出两个问题:

定义 4. 设 $\alpha_1, \alpha_2, \ldots, \alpha_m, \beta \in V(K)$. 若存在 $k_1, k_2, \ldots, k_m \in K$ 使得

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m,$$

则称 β 是 $\alpha_1, \alpha_2, \ldots, \alpha_m$ 的一个线性组合或可由其线性表示。

定义 5. 设 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 及 $\beta_1, \beta_2, \ldots, \beta_t$ 是线性空间$V(K)$中的两组向量。
若每个 α_i 都可由 $\beta_1, \beta_2, \ldots, \beta_t$ 线性表示，
则称向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 可由 $\beta_1, \beta_2, \ldots, \beta_t$ 线性表示；
若两个向量组可以互相线性表示，则称这两个向量组等价。

定义 6. 设 $\alpha_1, \ldots, \alpha_m \in V(K)$. 若存在不全为零的数 $k_1, \ldots, k_m \in K$ 使得

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m = 0,$$

则称 $\alpha_1, \alpha_2, \ldots, \alpha_m$ 线性相关，否则称其线性无关。

线性相关性一些简单性质。

1. 一个向量 α 线性相关 $\Longleftrightarrow$ α 是 V 中的零元素；

$$k\alpha = 0 \overset{k \neq 0}{\Longrightarrow} \alpha = 0.$$

2. 包含零元素的向量组线性相关：

$$0 \cdot \alpha_1 + 0 \cdot \alpha_2 + \cdots + 0 \cdot \alpha_m + 1 \cdot 0 = 0.$$

3. 一个向量组中部分向量线性相关，则该向量组线性相关。

$$k_1 \cdot \alpha_1 + k_2 \cdot \alpha_2 + \cdots + k_r \cdot \alpha_r = 0.$$
$$\Downarrow$$
$$k_1 \cdot \alpha_1 + \cdots + k_r \cdot \alpha_r + \underline{0 \cdot \alpha_{r+1} + \cdots + 0 \cdot \alpha_m} = 0.$$

简单记为：部分相关，整体必相关

4. 若一个向量组线性无关，则其任一部分向量组线性无关。

简单记为：整体无关，部分必无关

5. 向量组 $\alpha_1, \alpha_2, \ldots, \alpha_m\ (m \geq 2)$ 线性相关的充要条件是：
其中至少有一个向量可由其余向量线性表示。

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m = 0.$$

(3) 当 $m \geqslant 2$ 时, 向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$
可以由该向量组中其余的向量线性表出;

(4) 若某向量组线性无关, 则它的任意一部
若某向量组中有一个子向量组线性相关, 则该

定义7. **定义 6.2.2** (维数) 设 V 是数域 K 上的线性空间,

(1) 如果在 V 中可以找到任意多个线性无关的向量, 则称 V 是**无限维线性空间**;

(2) 如果存在有限多个向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n (n \geqslant 1) \in V$, 满足:

1) $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 线性无关;

2) V 中任一向量都可由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 线性表出.

则称 V 是**有限维线性空间**, 称 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 是 V 的一组**基**(或**基底**), $\boldsymbol{\alpha}_i$ 叫第 i 个**基向量**, 基向量的个数 n 称为线性空间 V 的**维数**, 记为 $\dim(V) = n$, 并称 V 是 n 维线性空间.

例 6.2.1 设 $K^{m\times n}$ 是例 6.1.1 给出的线性空间, 由于 $K^{m\times n}$ 中任一矩阵 $\boldsymbol{A} = (a_{ij})_{m\times n}$ 都可表示为

$$\boldsymbol{A} = (a_{ij})_{m\times n} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\boldsymbol{E}_{ij},$$

其中 $\boldsymbol{E}_{ij}$ 表示在第 i 行, 第 j 列交叉处的元素为 1, 其余元素均为 0 的 $m \times n$ 的矩阵 (称矩阵单位), 容易证得 $\{\boldsymbol{E}_{ij} : i = 1, 2, \cdots, m; j = 1, 2, \cdots, n\}$ 是线性无关的, 所以 $K^{m\times n}$ 是 $m \cdot n$ 维的.

例 6.2.2 设有例 6.1.2 中的线性空间 $P_n[x]$, 注意到向量组 $1, x, x^2, \cdots, x^n$ 是线性无关的 (易证), 且 $P_n[x]$ 中任一向量都可以表示为

$$p(x) = a_0 \cdot 1 + a_1 x + a_2 \cdot x^2 + \cdots + a_n \cdot x^n.$$

因此, $P_n[x]$ 是 $n+1$ 维的.

例 6.2.3 易见: 齐次线性方程组 $\boldsymbol{Ax} = \boldsymbol{\theta}$ 的基础解系是其解空间的一组基.

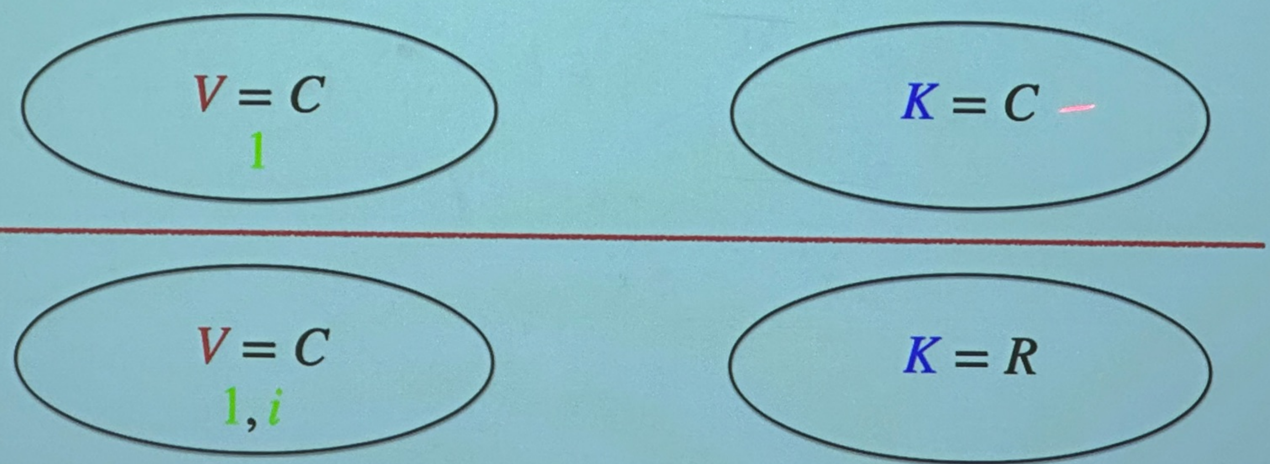
例 6.2.4 设 $\mathbf{C}$ 是复数域, 若将其看做复线性空间, 那么它是 1 维的; 若将其看做实线性空间, 则它是 2 维的, 因为 $\{1, i\}$ 是一组基.

可以证明, 线性空间里不同的基所含的向量的个数是相等的, 即线性空间的维数不因基的选择而改变. 另外, 由上述例子可见, 线性空间的基、维数跟相关的数域有关.

(无限维的例子)

例 4. 设 V 是实数域上关于文字 x 的一元多项式的全体，对多项式的加法以及数字与多项式的乘法构成实数域 R 的线性空间。该线性空间的一组基是 $1, x, x^2, \ldots, x^n, \ldots$

例 5. 设 $V=C$，$K=C$，C 是全体复数的集合。若定义 V 中的加法为复数的加法，数乘为复数的乘法， 则 $V=C$ 构成数域 $K=C$ 上的线性空间。此线性空间的维数为 1，$1\in V$ 构成 V 的一组基； 若把数域换成实数域 $K=R$，其余不变，则 $V=C$ 构成数域 $K=R$ 上的线性空间，且此线性空间的维数为 2, $1, i \in V$ 构成 V 的一组基。



定理 1. 设 $\alpha_1,\ldots,\alpha_r \in V(K)$ 线性无关, $\beta_1,\ldots,\beta_s \in V(K)$. 若 $\alpha_1,\ldots,\alpha_r$ 可由 $\beta_1, \ldots, \beta_s$ 线性表示，则有 $r \le s$.

证明： 设

$$\begin{cases} \alpha_1 = q_{11}\beta_1 + q_{21}\beta_2 + \cdots + q_{s1}\beta_s \\ \alpha_2 = q_{12}\beta_1 + q_{22}\beta_2 + \cdots + q_{s2}\beta_s \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_r = q_{1r}\beta_1 + q_{2r}\beta_2 + \cdots + q_{sr}\beta_s \end{cases}$$

也就是

$$(\alpha_1\ \alpha_2\ \cdots\ \alpha_r) = (\beta_1\ \beta_2\ \cdots\ \beta_s)\begin{array}{c} \begin{array}{cccc} \xi_1 & \xi_2 & & \xi_r \end{array} \\ \begin{pmatrix} \boxed{\begin{matrix} q_{11} \\ q_{21} \\ \cdots \\ q_{s1} \end{matrix}} & \boxed{\begin{matrix} q_{12} \\ q_{22} \\ \cdots \\ q_{s2} \end{matrix}} & \begin{matrix} \cdots \\ \cdots \\ \cdots \\ \cdots \end{matrix} & \boxed{\begin{matrix} q_{1r} \\ q_{2r} \\ \cdots \\ q_{sr} \end{matrix}} \end{pmatrix} \end{array} Q$$

若 $r>s$, 则 $r(Q)<r$, 从而 $\xi_1, \xi_2, \ldots, \xi_r \in K^s$ 线性相关！

即存在不全为零的数 $k_1, k_2, \ldots, k_r \in K$ 使得 $k_1\xi_1 + k_2\xi_2 + \cdots + k_r\xi_r = 0$. 由于

$$
(\alpha_1\ \alpha_2\ \cdots\ \alpha_r)\begin{pmatrix} k_1 \\ k_2 \\ \cdots \\ k_r \end{pmatrix} = (\beta_1\ \beta_2\ \cdots\ \beta_s)\begin{array}{c} \begin{array}{cccc} \xi_1 & \xi_2 & & \xi_r \end{array} \\ \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1r} \\ q_{21} & q_{22} & \cdots\ Q & q_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ q_{s1} & q_{s2} & \cdots & q_{sr} \end{pmatrix} \end{array} \begin{pmatrix} k_1 \\ k_2 \\ \cdots \\ k_r \end{pmatrix}
$$

故

$$
(\alpha_1\ \alpha_2\ \cdots\ \alpha_r)\begin{pmatrix} k_1 \\ k_2 \\ \cdots \\ k_r \end{pmatrix} = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_r\alpha_r = 0.
$$

这表明 $\alpha_1, \alpha_2, \ldots, \alpha_r$ 线性相关，矛盾。所以必有 $r \leq s$.

例 6.2.5　用 $C_{[a,b]}$ 表示闭区间上所有连续函数的集合, 因无法找到有限个连续函数作为它的基, 所以它是无限维的线性空间. 无限维线性空间的讨论超出了“线性代数”这门课程的范围, 有兴趣或想进一步研究的读者可参阅“泛函分析”课程的相关知识.

定理2. **定理 6.2.1**　设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 是 n 维线性空间 V 的一组基, 对任意的 $\boldsymbol{\alpha} \in V$, $\boldsymbol{\alpha}$ 可以唯一地由这一组基线性表出.

证明　设在基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 下, $\boldsymbol{\alpha}$ 有两种表示式:

证明: 设 α 在基 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 下有两种表示:

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n, \tag{1}$$

$$\alpha = y_1\alpha_1 + y_2\alpha_2 + \cdots + y_n\alpha_n, \tag{2}$$

则由 (1) + (2) × (−1) 可得

$$(x_1 - y_1)\,\alpha_1 + (x_2 - y_2)\,\alpha_2 + \cdots + (x_n - y_n)\,\alpha_n = 0.$$

于是 $x_i = y_i$, $1 \le i \le n$.

的系

□

若任给一组有序数 $x_1, x_2, \cdots, x_n$, 则对应有向量

$$\boldsymbol{\alpha} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n.$$

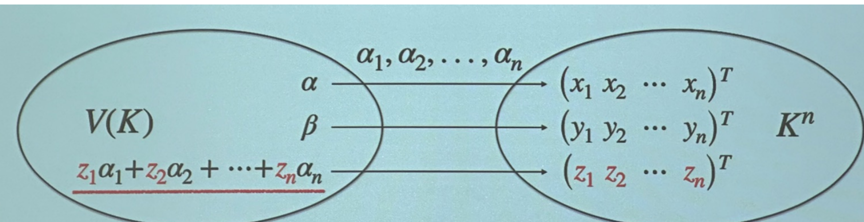
由此可见, V 中的向量 $\boldsymbol{\alpha}$ 和有序数组 $(x_1, x_2, \cdots, x_n)$ 之间存在着一一对应关系, 于是可以用有序数组来表示向量, 从而使得向量之间的运算转换成我们熟悉的常规运算, 带来很大的方便. 因此有下面的定义.

定义8. **定义 6.2.3** (坐标)　设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 是 n 维线性空间 V 的一组基, 对任意的 $\boldsymbol{\alpha} \in V$, 若有一组有序数 $x_1, x_2, \cdots, x_n$ 使得 $\boldsymbol{\alpha}$ 可表示为

$$\boldsymbol{\alpha} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n,$$

这组有序数就称为向量 $\boldsymbol{\alpha}$ 在基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 下的**坐标**, 记为

$$\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \quad 或 \quad \boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}.$$



可表示为

最后一条性

任意取定 $V(K)$ 的一组基，$V(K)$ 中的向量与 K^n 中元素按照:

$$向量 \mapsto 坐标 \tag{6.1}$$

构成一一对应。

注意 K^n 中的元素是“数组”,

如果在 K^n 中定义: 加法为数组的加法，数乘为数字与数组的乘法，

则 K^n 构成数域 K 上的线性空间。

例 6. 设 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 是线性空间 $V(K)$ 的一组基。则向量

$$\begin{aligned} \beta_1 &= p_{11}\alpha_1 + p_{21}\alpha_2 + \cdots + p_{n1}\alpha_n \\ \beta_2 &= p_{12}\alpha_1 + p_{22}\alpha_2 + \cdots + p_{n2}\alpha_n \\ &\cdots\cdots\cdots\cdots \\ \beta_n &= p_{1n}\alpha_1 + p_{2n}\alpha_2 + \cdots + p_{nn}\alpha_n \end{aligned}$$

构成 $V(K)$ 一组基的充要条件是

$$|P| = \begin{vmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{vmatrix} \neq 0.$$

证明. 只需证明 $\beta_1, \beta_2, \ldots, \beta_n$ 线性无关。设

$$k_1\beta_1 + k_2\beta_2 + \cdots + k_n\beta_n = 0.$$

亦即

$$(k_1p_{11} + k_2p_{12} + \cdots + k_np_{1n})\,\alpha_1 + \cdots + (k_1p_{n1} + k_2p_{n2} + \cdots + k_np_{nn})\,\alpha_n = 0.$$

由于 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性无关，故有

$$\begin{cases} k_1p_{11} + k_2p_{12} + \cdots + k_np_{1n} = 0 \\ k_1p_{21} + k_2p_{22} + \cdots + k_np_{2n} = 0 \\ \cdots\cdots\cdots\cdots \\ k_1p_{n1} + k_2p_{n2} + \cdots + k_np_{nn} = 0 \end{cases}$$

即 $(k_1\ k_2\ \cdots\ k_n)^T$ 是齐次线性方程组 $Px = 0$ 的解。故

$$\beta_1, \beta_2, \ldots, \beta_n \text{ 线性无关} \iff Px = 0 \text{ 只有零解} \iff |P| \neq 0.$$

$$\begin{cases} \cdots_{12} + \cdots + k_na_{1n} = 0, \\ k_1a_{21} + k_2a_{22} + \cdots + k_na_{2n} = 0, \\ \quad \cdots\cdots \\ k_1a_{n1} + k_2a_{n2} + \cdots + k_na_{nn} = 0. \end{cases}$$

而该方程组有唯一零解的充要条件是系数行列式 $D = |a_{ij}|_{n\times n} \neq 0$. □

上面的例子实际上给出了从一个已知基底构造另外基底的方法.

正如前面所述, 有了坐标概念后, 抽象线性空间中的元素就具体化、数量化了, 而且把元素与元素之间的运算变成了对坐标的运算, 其意义十分重要.

例 6.2.8　在线性空间 $P_n[x]$ 中, 多项式

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

在基 $1, x, x^2, \cdots, x^n$ 下的坐标是 $(a_0, a_1, \cdots, a_n)$, 若取另一组基

$$1, (x-a), (x-a)^2, \cdots, (x-a)^n,$$

则多项式 $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ 按泰勒公式展开为

$$p(x) = p(a) + p'(a)(x-a) + \frac{p''(a)}{2!}(x-a)^2 + \cdots + \frac{p^{(n)}(a)}{n!}(x-a)^n.$$

因此, $p(x)$ 在新的一组基 $1, (x-a), (x-a)^2, \cdots, (x-a)^n$ 下的坐标是

$$\left(p(a), p'(a), \frac{p''(a)}{2!}, \cdots, \frac{p^{(n)}(a)}{n!}\right).$$

6.2.2　基变换与坐标变换

由上面可以发现, 一个向量的坐标与所取的基底有关, 那么同一个向量在两组不同的基底下的两组坐标之间有什么样的关系呢?

定理 3. 设 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 与 $\beta_1, \beta_2, \ldots, \beta_n$ 是线性空间 $V(K)$ 的两组基。

$$(\beta_1\ \beta_2\ \cdots\ \beta_n) = (\alpha_1\ \alpha_2\ \cdots\ \alpha_n)\, P,$$

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

若 $\alpha \in V(K)$ 在两组基 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 与 $\beta_1, \beta_2, \ldots, \beta_n$ 下的坐标分别是 $(x_1\ x_2\ \cdots\ x_n)^T$ 与 $(y_1\ y_2\ \cdots\ y_n)^T$，则

$$\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix} = P^{-1} \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}.$$

矩阵 P 称为由基 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 到 $\beta_1, \beta_2, \ldots, \beta_n$ 的**过渡矩阵**。

证明. 由于

$$\alpha = (\alpha_1\ \alpha_2\ \cdots\ \alpha_n) \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}, \qquad \alpha = \underline{(\beta_1\ \beta_2\ \cdots\ \beta_n)} \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix},$$

故有

$$\alpha = \underline{(\alpha_1\ \alpha_2\ \cdots\ \alpha_n) P} \boxed{\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}} = (\alpha_1\ \alpha_2\ \cdots\ \alpha_n) \boxed{\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}}, \tag{6.6}$$

因此，

$$\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix} = P^{-1} \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}.$$

$$\begin{pmatrix} \vdots \\ y_n \end{pmatrix} = P \begin{pmatrix} \vdots \\ x_n \end{pmatrix},$$

即式 (6.5). □

定理 6.2.2 中, $\boldsymbol{P}$ 称为从基底 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 到基底 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n$ 的**过渡矩阵**, 称式 (6.3) 为**基底变换公式**, 称式 (6.6) 为**坐标变换公式**.

由上述可见, 过渡矩阵 $\boldsymbol{P}$ 中的第 j 个列向量 $\boldsymbol{p}_j = (a_{1j}, \cdots, a_{nj})^{\mathrm{T}}$ 是 $\boldsymbol{\beta}_j$ 在基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 下的坐标向量.

由例6.2.7 可知，过渡矩阵 $\boldsymbol{P}$ 是可逆矩阵，从而由式 (6.3) 可得

$$(\boldsymbol{\alpha}_1\boldsymbol{\alpha}_2\cdots\boldsymbol{\alpha}_n)=(\boldsymbol{\beta}_1\boldsymbol{\beta}_2\cdots\boldsymbol{\beta}_n)\boldsymbol{P}^{-1},$$

即 $\boldsymbol{P}^{-1}$ 是从基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 到基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的过渡矩阵.

设 $\boldsymbol{e}_1=(1,0,\cdots,0)^{\mathrm{T}},\boldsymbol{e}_2=(0,1,\cdots,0)^{\mathrm{T}},\cdots,\boldsymbol{e}_n=(0,\cdots,0,1)^{\mathrm{T}}$ 是 $\mathbf{R}^n$ 的一组向量, 易见它们是线性无关的, 故是一组基. 因为对任意一个向量 $\boldsymbol{\alpha}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in\mathbf{R}^n$, 有

$$\boldsymbol{\alpha}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_n\boldsymbol{e}_n.$$

故称 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 是 $\mathbf{R}^n$ 的**自然基**.

例7、**例 6.2.9** 设 $\mathbf{R}^4$ 中的向量 $\boldsymbol{\alpha}$ 在基底 $\boldsymbol{\alpha}_1=(1,2,-1,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,-1,1,1)^{\mathrm{T}},\boldsymbol{\alpha}_3=(-1,2,1,1)^{\mathrm{T}},\boldsymbol{\alpha}_4=(-1,-1,0,1)^{\mathrm{T}}$ 下的坐标是 $(x_1,x_2,x_3,x_4)^{\mathrm{T}}$, 求 $\boldsymbol{\alpha}$ 在另一组基 $\boldsymbol{\beta}_1=(2,1,0,1)^{\mathrm{T}},\boldsymbol{\beta}_2=(0,1,2,2)^{\mathrm{T}},\boldsymbol{\beta}_3=(-2,1,1,2)^{\mathrm{T}},\boldsymbol{\beta}_4=(1,3,1,2)^{\mathrm{T}}$ 下的坐标.

解

$$(\boldsymbol{\alpha}_1\boldsymbol{\alpha}_2\boldsymbol{\alpha}_3\boldsymbol{\alpha}_4)=(\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3\boldsymbol{e}_4)\begin{pmatrix}1&1&-1&-1\\2&-1&2&-1\\-1&1&1&0\\0&1&1&1\end{pmatrix}=(\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3\boldsymbol{e}_4)\boldsymbol{A},$$

$$(\boldsymbol{\beta}_1\boldsymbol{\beta}_2\boldsymbol{\beta}_3\boldsymbol{\beta}_4)=(\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3\boldsymbol{e}_4)\begin{pmatrix}2&0&-2&1\\1&1&1&3\\0&2&1&1\\1&2&2&2\end{pmatrix}=(\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3\boldsymbol{e}_4)\boldsymbol{B}.$$

于是

$$(\boldsymbol{\beta}_1\boldsymbol{\beta}_2\boldsymbol{\beta}_3\boldsymbol{\beta}_4)=(\boldsymbol{\alpha}_1\boldsymbol{\alpha}_2\boldsymbol{\alpha}_3\boldsymbol{\alpha}_4)\boldsymbol{A}^{-1}\boldsymbol{B},$$

其中 $\boldsymbol{P}=\boldsymbol{A}^{-1}\boldsymbol{B}$ 为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 到 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$ 的过渡矩阵.

由坐标变换公式 (6.5), $\boldsymbol{\alpha}$ 在 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$ 下的新坐标是

$$(x_1',x_2',x_3',x_4')=(x_1,x_2,x_3,x_4)(\boldsymbol{P}^{-1})^{\mathrm{T}}.$$

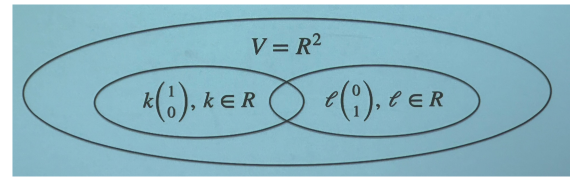
计算可得

$$(\boldsymbol{P}^{-1})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{B}^{\mathrm{T}})^{-1}=\begin{pmatrix}0&-1&0&1\\1&1&0&-1\\-1&0&0&1\\1&0&1&-1\end{pmatrix}.$$

故得

$$\begin{aligned}x_1'&=x_2-x_3+x_4, & x_2'&=-x_1+x_2,\\ x_3'&=x_4, & x_4'&=x_1-x_2+x_3-x_4.\end{aligned}$$

□



6.3　线性空间的子空间

6.3.1　子空间

在三维空间直角坐标系 O-xyz 中, 所有以坐标原点 O 为起点的向量关于通常的加法和数乘构成了一个实数域上的三维线性空间. 可以看出, 任一过 O 点的一条直线或一个平面都是三维几何空间的子集, 它们关于向量加法和数乘的运算分别构成一个一维和二维的线性空间. 这种现象也可以推广到一般的线性空间, 即线性空间的子集在原线性空间规定的加法和数乘运算下, 仍可构成线性空间, 这种性质, 对深入研究线性空间的本质、分解与构造非常有用. 于是, 我们引入如下子空间的概念.

定义9. **定义 6.3.1** (子空间)　设 W 是数域 K 上的线性空间 V 的一个非空子集, 若 W 关于 V 上的加法和数乘也构成数域 K 上的一个线性空间, 则称 W 是 V 的一个**线性子空间**, 简称**子空间**, 记为 $W \subseteq V$, 若 $W \neq V$, 记为 $W \subset V$.

容易证明如下定理

定理 6.3.1　线性空间 V 的一个非空子集 W 是 V 的子空间的充要条件是

(1) 对任意的 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in W$, 有 $\boldsymbol{\alpha} + \boldsymbol{\beta} \in W$, 即对加法封闭;

(2) 对任意的 $\boldsymbol{\alpha} \in W, \lambda \in K$, 有 $\lambda\boldsymbol{\alpha} \in W$, 即对数乘封闭.

亦可把上面两个条件合并得到.

定理4、**定理 6.3.2**　线性空间 V 的一个非空子集 W 是 V 的子空间的充要条件是对任意的 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in W$,　$\lambda, \mu \in K$, 有 $\lambda\boldsymbol{\alpha} + \mu\boldsymbol{\beta} \in W$.

由此可见, 验证一个非空子集是否是子空间, 只需验证这个子集对原线性空间的加法和数乘运算保持封闭就可以了, 其余 8 条规则自然满足.

每个线性空间至少有两个子空间, 一个是仅由零向量构成的, 称为**零子空间**; 一个是它自身, 我们称它们为**平凡子空间**, 而其他的子空间称为非平凡子空间 (或**真子空间**).

显然, n 维线性空间 V 的任一子空间的维数都小于等于 n.

设 $\alpha_1, \alpha_2, \ldots, \alpha_r$ 是线性空间 $V(K)$ 中任意 r 个向量。令

$$W = \{\alpha \mid \alpha = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_r\alpha_r,\ k_1, k_2, \ldots, k_r \in K\}.$$

容易验证 W 构成 V 的子空间。

这个子空间称为由向量 $\alpha_1, \alpha_2, \ldots, \alpha_r$ 生成的子空间，记为

$$span\{\alpha_1, \alpha_2, \ldots, \alpha_r\} \quad 或 \quad L\{\alpha_1, \alpha_2, \ldots, \alpha_r\}.$$

特别地, 零子空间是由零向量**生成的子空间** $\mathrm{span}\{\mathbf{0}\}$.

设

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \quad \cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0, \end{cases} \qquad (a_{ij} \in \mathbf{R})$$

是实数域 $\mathbf{R}$ 上的齐次线性方程组, 它的全体解向量是 $\mathbf{R}^n$ 中的一个非空子集 W, 显然, W 中的解向量满足定理6.3.1 的条件, 因此 W 是 $\mathbf{R}^n$ 的一个子空间, 这个子空间称为该齐次方程组的**解空间**. 该齐次方程组的任一组基础解系是 W 的一组基, 若方程组的系数矩阵的秩为 r, 则

$$\dim(W) = n - r.$$

6.3.2 子空间的交与和

线性空间的子空间有两种重要的运算 —— 交与和, 这可看成是由子空间生成的子空间.

定义10、**定义 6.3.3** (子空间的交与和) 设 W_1, W_2 是数域 K 上线性空间 V 的两个子空间, 定义 W_1 与 W_2 的交为

$$W_1 \cap W_2 = \{\boldsymbol{\alpha} \mid \boldsymbol{\alpha} \in W_1, \boldsymbol{\alpha} \in W_2\},$$ 两子空间的交一定存在

W_1 与 W_2 的和为

一般情况下 $W_1+W_2 \neq W_1 \cup W_2$, 事实上 $W_1 \cup W_2 \subseteq W_1+W_2$

$$W_1 + W_2 = \{\boldsymbol{\gamma} \mid \boldsymbol{\gamma} = \boldsymbol{\alpha} + \boldsymbol{\beta}, \text{对所有的} \boldsymbol{\alpha} \in W_1, \boldsymbol{\beta} \in W_2\}.$$

定理5、**定理 6.3.3** 数域 K 上线性空间 V 的两个子空间 W_1 与 W_2 的交与和仍是 V 的子空间.

证明 由于每一个子空间都含有零向量, 所以 W_1 与 W_2 的交也含有零向量, 即 $W_1 \cap W_2$ 是非空的集合. 设 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in W_1 \cap W_2$, 则 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in W_1$ 且 $\boldsymbol{\alpha}, \beta \in W_2$, 因为 W_1, W_2 都是 V 的子空间, 所以, 对任意的 $k_1, k_2 \in K$, 有

$$k_1\boldsymbol{\alpha} + k_2\boldsymbol{\beta} \in W_1, \qquad k_1\boldsymbol{\alpha} + k_2\boldsymbol{\beta} \in W_2.$$

因此 $k_1\boldsymbol{\alpha} + k_2\boldsymbol{\beta} \in W_1 \cap W_2$ 故 $W_1 \cap W_2$ 是 V 的子空间.

显然 $W_1 + W_2$ 非空, 对任意的 $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2 \in W_1 + W_2$, 存在 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in W_1, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \in W_2$ 使得

$$\boldsymbol{\gamma}_1 = \boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1, \qquad \boldsymbol{\gamma}_2 = \boldsymbol{\alpha}_2 + \boldsymbol{\beta}_2.$$

因 W_1, W_2 都是子空间, 则对任意的 $k_1, k_2 \in K$, 有

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 \in W_1, \qquad k_1\boldsymbol{\beta}_1 + k_2\boldsymbol{\beta}_2 \in W_2.$$

于是

$$k_1\boldsymbol{\gamma}_1 + k_2\boldsymbol{\gamma}_2 = (k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2) + (k_1\boldsymbol{\beta}_1 + k_2\boldsymbol{\beta}_2) \in W_1 + W_2.$$

因此, $W_1 + W_2$ 是 V 的子空间. □

例 6.3.1 在三维几何空间 $\mathbf{R}^3$ 中, 以 W_1 表示过坐标原点 O 的一条直线, W_2 表示一个过 O 点且与 W_1 垂直的平面, 则 $W_1 \cap W_2 = \{\mathbf{0}\}, W_1 + W_2 = \mathbf{R}^3$.

需要注意的是, 子空间的交与集合的交的概念是一致的, 但子空间的和与集合的并的概念并不一致. 如上例所示, $W_1 \cap W_2 = \{O\}$, 但 $W_1 \cup W_2 = \{$ 经过原点的直线与垂直平面上的点 $\}$.

子空间的交与和满足交换律和结合律, 即若 W_1, W_2, W_3 是线性空间 V 的子空间, 则

$$\begin{aligned}
&W_1 + W_2 = W_2 + W_1, \\
&(W_1 + W_2) + W_3 = W_1 + (W_2 + W_3), \\
&W_1 \cap W_2 = W_2 \cap W_1, \\
&(W_1 \cap W_2) \cap W_3 = W_1 \cap (W_2 \cap W_3).
\end{aligned}$$

于是, 可以定义多个子空间的交与和

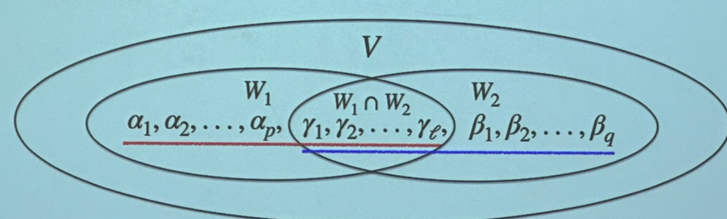
$$\begin{aligned}
&W_1 \cap W_2 \cap \cdots \cap W_m = (W_1 \cap W_2 \cap \cdots \cap W_{m-1}) \cap W_m, \\
&W_1 + W_2 + \cdots + W_m = (W_1 + W_2 + \cdots + W_{m-1}) + W_m.
\end{aligned}$$

下面给出子空间的交与和的维数公式.

定理6、**定理 6.3.4**　若 W_1, W_2 是线性空间 V 的两个有限维子空间, 则

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2).$$

证明. 设 $\gamma_1, \gamma_2, \ldots, \gamma_\ell$ 是 $W_1 \cap W_2$ 的一组基(可以没有)。 则其可以扩充为 W_1 的一组基 $\alpha_1, \ldots, \alpha_p, \gamma_1, \ldots, \gamma_\ell$, 以及 W_2 的一组基 $\gamma_1, \ldots, \gamma_\ell, \beta_1, \ldots, \beta_q$.



$$\alpha_1, \alpha_2, \ldots, \alpha_p, \gamma_1, \gamma_2, \ldots, \gamma_\ell, \beta_1, \beta_2, \ldots, \beta_q$$

只需证明 $\alpha_1, \alpha_2, \ldots, \alpha_p, \gamma_1, \gamma_2, \ldots, \gamma_\ell, \beta_1, \beta_2, \ldots, \beta_q$ 是 $W_1 + W_2$ 的一组基。易见, $W_1 + W_2$ 中的向量均可由该向量组线性表示。设

$$s_1\alpha_1 + \cdots + s_p\alpha_p + k_1\gamma_1 + \cdots + k_\ell\gamma_\ell + t_1\beta_1 + \cdots + t_q\beta_q = 0. \qquad (1)$$

令 $\xi = s_1\alpha_1 + \cdots + s_p\alpha_p + k_1\gamma_1 + \cdots + k_\ell\gamma_\ell$, 则 $\xi \in W_1$ 且

$$\xi = -t_1\beta_1 - \cdots - t_q\beta_q \in W_2.$$

这表明 $\xi \in W_1 \cap W_2$, 从而可由 $\gamma_1, \gamma_2, \ldots, \gamma_\ell$ 线性表示, 设为

$$\xi = c_1\gamma_1 + c_2\gamma_2 + \cdots + c_\ell\gamma_\ell.$$

由 (1) 得 $c_1\gamma_1 + c_2\gamma_2 + \cdots + c_\ell\gamma_\ell + t_1\beta_1 + \cdots + t_q\beta_q = 0$, 从而有

$$t_1 = \cdots = t_q = 0.$$

代入 (1) 可得 $s_1 = \cdots = s_p = k_1 = \cdots = k_\ell = t_1 = \cdots = t_q = 0$.

令

$$\alpha_1=\begin{pmatrix}1\\0\\0\end{pmatrix},\quad \alpha_2=\begin{pmatrix}0\\1\\0\end{pmatrix},\quad \alpha_3=\begin{pmatrix}0\\0\\1\end{pmatrix},\quad \alpha=\begin{pmatrix}1\\1\\1\end{pmatrix}.$$

易见，$W_1=L\{\alpha_1,\alpha_2\}, W_2=L\{\alpha_2,\alpha_3\}, W_3=L\{\alpha_3\}$ 均是 R^3 的子空间，且

$$W_1+W_2=R^3,\qquad W_1+W_3=R^3.$$

$$\alpha=\underbrace{(\alpha_1+\alpha_2)}_{W_1}+\underbrace{\alpha_3}_{W_2}=\underbrace{\alpha_1}_{W_1}+\underbrace{(\alpha_2+\alpha_3)}_{W_2}\qquad \alpha=\underbrace{(\alpha_1+\alpha_2)}_{W_1}+\underbrace{\alpha_3}_{W_3}$$

$_l=0(l=1,2,\cdots,n_2-$
性无关性得 $k_i=0(i=$
系数必须全部为零，这就
从而证明了定理成立. □
(作为习题, 读者自证).
表示成 $\gamma=\alpha+\beta,\alpha\in$
$\in W_2$, 使 $\gamma=\alpha'+\beta'$.

定义11. **定义 6.3.4** (直和)　若 W_1+W_2 中任一向量都只能唯一地表示为子空间 W_1 的一个向量与子空间 W_2 的一个向量的和, 则称 W_1+W_2 是**直和**(或**直接和**), 记为 $W_1\oplus W_2$ 或 $W_1\dotplus W_2$.

若 $W=W_1\oplus W_2$, 则称在 W 内 W_1 是 W_2 的**补空间**, 或 W_2 是 W_1 的**补空间**.

定理7. **定理 6.3.5**　*W_1+W_2 是直和的充要条件是 $W_1\cap W_2=\{0\}$.*

证明. ($\Leftarrow$) 若 $\exists\,\gamma\in W_1+W_2$ 有两种不同的表示方式，即存在$\alpha_1,\alpha_2\in W_1$ 及 $\beta_1,\beta_2\in W_2$，且 $\alpha_1\neq\alpha_2$，$\beta_1\neq\beta_2$ 使得

$$\gamma=\alpha_1+\beta_1,\qquad \gamma=\alpha_2+\beta_2,$$

则有

$$(\alpha_1-\alpha_2)+(\beta_1-\beta_2)=0,\implies \alpha_1-\alpha_2=-(\beta_1-\beta_2)\neq 0.$$

这说明 $\alpha_1-\alpha_2\neq 0$ 且 $\alpha_1-\alpha_2\in W_1\cap W_2$，矛盾。

($\Rightarrow$) 假设 $W_1\cap W_2\neq\{0\}$，则存在 $\alpha\in W_1\cap W_2$ 且 $\alpha\neq 0$.
由于 $W_1\cap W_2$ 是线性空间，故 $-\alpha\in W_1\cap W_2$. 于是

$$0+0=\alpha+(-\alpha)=0\in W_1+W_2,$$

即 0 有两种不同表示法，这与 W_1+W_2 是直和矛盾。

表示成 W_1 的
$,\boldsymbol{\beta}_1\neq\boldsymbol{\beta}_2$, 使
少存在一向量
$\boldsymbol{\alpha}\neq\mathbf{0}$. 因 $W_1\cap W_2$ 是线性空间, 故有 $-\boldsymbol{\alpha}\in W_1\cap W_2$, 则有 $\mathbf{0}=\boldsymbol{\alpha}+(-\boldsymbol{\alpha})$, 而 $\mathbf{0}$ 又可表示成 W_1 中的零向量 $\mathbf{0}$ 与 W_2 中的零向量 $\mathbf{0}$ 之和 $\mathbf{0}=\mathbf{0}+\mathbf{0}$, 即 W_1+W_2 中零向量表示法不唯一, 矛盾. 定理得证. □

推论1. **推论 6.3.6**　*W_1+W_2 是直和的充要条件是* $\dim(W_1+W_2)=\dim W_1+\dim W_2$.

证明　由定理 6.3.4 维数公式立刻推得. □

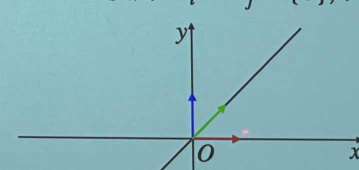
子空间直和的概念可推广到有限多个子空间.

定义 12. 设 $W_1,W_2,\ldots,W_m$ 是线性空间 V 的子空间。若

$$W_1\cap W_2=\{0\},$$
$$(W_1+W_2)\cap W_3=\{0\},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$(W_1+\cdots+W_{m-1})\cap W_m=\{0\},$$

则称 $W_1+W_2+\cdots+W_m$ 为直和，记为 $W_1\oplus W_2\oplus\cdots\oplus W_m$.

注意. 定义中的条件不可以替换为 $W_i\cap W_j=\{0\}, i\neq j$.



$V=W_1\oplus W_2\oplus\cdots\oplus W_m$.

例8. **例 6.3.2**　设 F^n 是数域 F 上的 n 维列向量空间，$\boldsymbol{A}$ 是 F 上的 n 阶方阵，令

$$V_1=\{\boldsymbol{A}\boldsymbol{x}:\forall\boldsymbol{x}\in F^n\},\quad V_2=\{\boldsymbol{x}:\boldsymbol{A}\boldsymbol{x}=\mathbf{0},\boldsymbol{x}\in F^n\},$$

试证: (1) V_1, V_2 是 F^n 的子空间; (2) 若 $\boldsymbol{A}$ 是幂等矩阵，即 $\boldsymbol{A}^2 = \boldsymbol{A}$, 则 $F^n = V_1 \oplus V_2$.

证明　(1) 易见 V_1, V_2 都是 F^n 的非空子集，$\forall k, l \in F$, 和 $\forall \boldsymbol{x}, \boldsymbol{y} \in F^n$，因

$$k\boldsymbol{A}\boldsymbol{x} + l\boldsymbol{A}\boldsymbol{y} = \boldsymbol{A}(k\boldsymbol{x} + l\boldsymbol{y}) \in V_1,$$

故 V_1 是 F^n 的子空间;

$\forall \boldsymbol{\xi}, \boldsymbol{\eta} \in V_2$, 有 $\boldsymbol{A}\boldsymbol{\xi} = \boldsymbol{0}, \boldsymbol{A}\boldsymbol{\eta} = \boldsymbol{0}$，且 $\forall k, l \in F$, 有

$$\boldsymbol{A}(k\boldsymbol{\xi} + l\boldsymbol{\eta}) = k\boldsymbol{A}\boldsymbol{\xi} + l\boldsymbol{A}\boldsymbol{\eta} = \boldsymbol{0},$$

故 $k\boldsymbol{\xi} + l\boldsymbol{\eta} \in V_2$, 即 V_2 是 F^n 的子空间.

(2) 当 $\boldsymbol{A}$ 是幂等矩阵时, 将 F^n 中任一向量表示成: $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x} + (\boldsymbol{x} - \boldsymbol{A}\boldsymbol{x})$, 注意到 $\boldsymbol{A}\boldsymbol{x} \in V_1$, 以及因 $\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{A}\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{0}$, 得 $\boldsymbol{x} - \boldsymbol{A}\boldsymbol{x} \in V_2$, 所以 $F^n \subseteq V_1 + V_2$, 从而 $F^n = V_1 + V_2$.

设 $\boldsymbol{\xi}$ 是 $V_1 \cap V_2$ 中的任一向量，因 $\boldsymbol{\xi} \in V_1$, 所以存在 $\boldsymbol{\eta} \in F^n$ 使得 $\boldsymbol{\xi} = \boldsymbol{A}\boldsymbol{\eta}$. 又因 $\boldsymbol{\xi} \in V_2$, 所以 $\boldsymbol{A}\boldsymbol{\xi} = \boldsymbol{0}$, 于是 $\boldsymbol{\xi} = \boldsymbol{A}\boldsymbol{\eta} = \boldsymbol{A}^2\boldsymbol{\eta} = \boldsymbol{A}(\boldsymbol{A}\boldsymbol{\eta}) = \boldsymbol{A}\boldsymbol{\xi} = \boldsymbol{0}$, 即 $V_1 \cap V_2 = \{\boldsymbol{0}\}$, 所以 $F^n = V_1 \oplus V_2$.

□

6.4　线 性 变 换

6.4.1　线性变换的概念

为了进一步研究线性空间, 本节考虑有限维线性空间之间的一类特殊的映射–线性映射和线性变换.

定义15、**定义 6.4.1** (线性变换)　设 V_1, V_2 都是数域 K 上线性空间, 根据某一规则 T, 对 V_1 中的任一元素 $\boldsymbol{\alpha}$, 有 V_2 中的唯一的元素 $\boldsymbol{\alpha}'$ 与之对应, 即 $T\boldsymbol{\alpha} = \boldsymbol{\alpha}'$, 则称 T 为 V_1 到 V_2 的**映射**.

如果 V_1 到 V_2 的映射 T 还满足

和的像=像的和

$$T(\boldsymbol{\alpha} \underset{V_1\text{中加法}}{+} \boldsymbol{\beta}) = T\boldsymbol{\alpha} \underset{V_2\text{中加法}}{+} T\boldsymbol{\beta}, \quad T(\underset{V_1\text{中数乘}}{\lambda\boldsymbol{\alpha}}) = \underset{V_2\text{中数乘}}{\lambda T\boldsymbol{\alpha}},$$

其中 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V_1, \lambda \in K$, 则称 T 为 V_1 到 V_2 的**线性映射**. 在 $V_1 = V_2 = V$ 时，称这个 T 为 V 上的**线性变换**.

例 6.4.1　对 $P_n[x]$ 中的多项式求导, 记为 $D[p(x)] = \dfrac{\mathrm{d}}{\mathrm{d}x}p(x)$, 不难验证 D 是 $P_n[x]$ 上的线性变换.

例 9. 设 $A = (a_{ij})_{m\times n} \in R^{m\times n}$. $\forall x \in R^n$，令 $Tx = Ax$, 则有

$$T(x+y) = A(x+y) = Ax + Ay = Tx + Ty, \qquad \forall x, y \in R^n,$$

$$T(kx) = A(kx) = kAx = kTx, \qquad \forall x \in R^n, k \in R,$$

故 T 是 R^n 到 R^m 上的线性映射。

则容易验证, $\boldsymbol{A}$ 是 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 的线性映射.

由定义 6.4.1 可以看出, 线性映射 T 具有下述性质:

(1) $T\boldsymbol{0} = \boldsymbol{0}$;

证: $T0 = T(00) = 0T(0) = 0$

(2) 对任意的 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m \in V_1$, 有

$$T(k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m) = k_1T\boldsymbol{\alpha}_1 + k_2T\boldsymbol{\alpha}_2 + \cdots + k_mT\boldsymbol{\alpha}_m;$$

(3) 若 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性相关, 则 $T\boldsymbol{\alpha}_1, T\boldsymbol{\alpha}_2, \cdots, T\boldsymbol{\alpha}_m$ 也线性相关.

证明 (1) 在 $T(\lambda\boldsymbol{\alpha}) = \lambda T\boldsymbol{\alpha}$ 中取 $\lambda = 0$ 即可得到 $T\mathbf{0} = \mathbf{0}$;

(2) 连续应用 $T(k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2) = k_1T\boldsymbol{\alpha}_1 + k_2T\boldsymbol{\alpha}_2$ 即可;

(3) 由刚证得的性质 (2), 根据线性相关的定义即得. □

若 $T\boldsymbol{\alpha}_1, T\boldsymbol{\alpha}_2, \cdots, T\boldsymbol{\alpha}_m$ 线性相关, 不一定能推出 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性相关, 即性质 (3) 的逆命题一般不成立.

下面给出几个特殊的线性变换, 在以后的讨论中是非常有用的.

(1) **数乘变换**: 设 k 是数域 K 内的一个常数, 对任意的 $\boldsymbol{\alpha} \in V$, 令 $T_k\boldsymbol{\alpha} = k\boldsymbol{\alpha}$, 容易验证 T_k 是 V 上的线性变换.

(2) **恒等变换**: 设 $T\boldsymbol{\alpha} = \boldsymbol{\alpha}$, 即 T 把 V 中的任意元素 $\boldsymbol{\alpha}$ 变成自身, 则称 T 为**恒等变换**或**单位变换**, 可记为 $\boldsymbol{I}$ 或 $\boldsymbol{E}$, 即 $\boldsymbol{I}\boldsymbol{\alpha} = \boldsymbol{\alpha}$.

(3) **零变换**: 设 $T\boldsymbol{\alpha} = \mathbf{0}$, 即 T 把 V 中的任意元素 $\boldsymbol{\alpha}$ 变成零元素, 则称 T 为零变换, 记为 T_0, 即 $T_0\boldsymbol{\alpha} = \mathbf{0}$.

定义 14. 设 V_1, V_2 都是数域 K 上的线性空间,

$$T: \ V_1 \mapsto V_2$$

是线性映射。则

$$Im(T) = \{T\alpha \mid \alpha \in V_1\}$$

称为 T 的像空间,

$$Ker(T) = \{\alpha \mid \alpha \in V_1, \ T\alpha = 0\}$$

称为 T 的核空间。

V_2 都是数域 K 上线性空间, 设线性映射 T :

空间, 记作 $\mathcal{R}(T)$ 或 $\mathrm{Im}(T)$, 简记为 $\mathcal{R}$.

$T\boldsymbol{\alpha} = \mathbf{0}', \boldsymbol{\alpha} \in V_1\}$

的零元.

定理 8. 设 V_1, V_2 都是数域 K 上的线性空间,

$$T: \ V_1 \mapsto V_2$$

是线性映射。则 $Im(T)$是 V_2的线性子空间, $Ker(T)$ 是V_1的线性子空间。

$\longrightarrow V_2$,

$T\boldsymbol{\alpha} = \boldsymbol{\alpha}', T\boldsymbol{\beta} = \boldsymbol{\beta}'$,

$$T\alpha = \alpha', \qquad T\beta = \beta'.$$

于是

$$k\alpha' + \ell\beta' = kT\alpha + \ell T\beta = T(k \cdot \alpha + \ell \cdot \beta) \in Im(T),$$

故 $Im(T)$ 是 V_2 的线性子空间。

(2) $\forall \alpha, \beta \in Ker(T)$, $k, \ell \in K$, 由于

$$T(k \cdot \alpha + \ell \cdot \beta) = kT\alpha + \ell T\beta = k0 + \ell 0 = 0,$$

故 $k \cdot \alpha + \ell \cdot \beta \in Ker(T)$, 从而 $Ker(T)$ 是 V_1 的线性子空间。

特别的, 若

$$T: \ V \mapsto V$$

是线性空间V 上的线性变换, 则 $Im(T)$ 与 $Ker(T)$都是V 的线性子空间。

$(\boldsymbol{\alpha} + \boldsymbol{\beta})$,

$\in \mathcal{R}$. 因此, $\mathcal{R}$ 是一个线性空间.

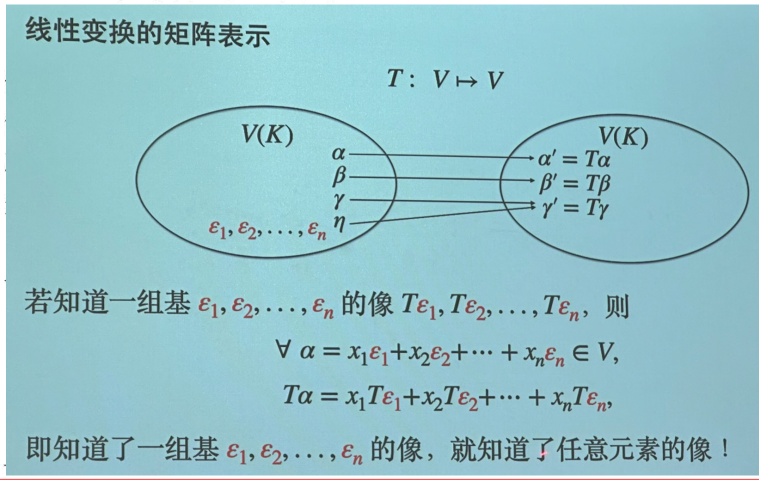
$+ \mathbf{0}' = \mathbf{0}'$,

$= \mathbf{0}'$,

□

例 6.4.3 设 $\boldsymbol{A} = (a_{ij}) \in M_{m\times n}(\mathbf{R})$ 是一个实数域上的 $m \times n$ 矩阵, 则 $T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ 定义了线性空间 $\mathbf{R}^n$ 到线性空间 $\mathbf{R}^m$ 的一个线性映射 T. 而齐次线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\theta}$ 的解空间就是线性映射 T 的核空间 $\ker(T)$, 它是 $\mathbf{R}^n$ 的一个子空间. 而 $\boldsymbol{A} = (\boldsymbol{\alpha}_1\boldsymbol{\alpha}_2\cdots\boldsymbol{\alpha}_n)$ 的列向量组生成的线性空间就是 T 的像空间 $\mathrm{Im}(T)$, 即 $\mathrm{Im}(T) = \mathrm{span}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$, 它是 $\mathbf{R}^m$ 的一个子空间. 如果 $\mathrm{r}(\boldsymbol{A}) = r$, 则还有 $\dim(\ker(T)) = n - r, \dim(\mathrm{Im}(T)) = r$.

6.4.2　线性变换的矩阵表示



…那给出以表示 T, 显然太不方
…一地线性表出, 线性变换又
…, 则 V 中的每个元素的像也
…变换和矩阵之间的一一对应

…变换, $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 V 的

…$_n$.

…下的像也随之确定了.

定理9. **定理 6.4.2**　设对 V 的两个线性变换 T_1 和 T_2, 有

$$T_1\varepsilon_i = T_2\varepsilon_i, \quad i = 1, 2, \cdots, n,$$

则 $T_1 = T_2$(这里两个线性变换相等是指它们对 V 的任一向量的像相等).

证明　对任一 $x \in V$, 因

$$\begin{aligned} T_1 x &= x_1 T_1\varepsilon_1 + x_2 T_1\varepsilon_2 + \cdots + x_n T_1\varepsilon_n \\ &= x_1 T_2\varepsilon_1 + x_2 T_2\varepsilon_2 + \cdots + x_n T_2\varepsilon_n \\ &= T_2 x \end{aligned}$$

所以 $T_1 = T_2$. □

定理 6.4.2 表明, 线性空间 V 上的任一线性变换 T 由它的基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 在 T 下的像 $T\varepsilon_1, T\varepsilon_2, \cdots, T\varepsilon_n$ 唯一决定.

设

$$T\varepsilon_i = a_{1i}\varepsilon_1 + a_{2i}\varepsilon_2 + \cdots + a_{ni}\varepsilon_n, \quad i = 1, 2, \cdots, n. \tag{6.10}$$

将上式写成矩阵的形式

$$(T\varepsilon_1 T\varepsilon_2 \cdots T\varepsilon_n) = (\varepsilon_1 \varepsilon_2 \cdots \varepsilon_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

令

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

$\boldsymbol{A}$ 的第 i 个列向量是基的像 $T\varepsilon_i$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标. 因此, 取定一组基后, 线性空间 V 上的任一线性变换 T 都有式 (6.10) 所确定的方阵 $\boldsymbol{A}$ 与之对应, $\boldsymbol{A}$ 称为线性变换 T 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵.

反之, 若给定一个 n 阶的矩阵 $\boldsymbol{A}$, 是否存在线性空间 V 上的某一线性变换 T, 使得它在给定的一组基下的矩阵就是 $\boldsymbol{A}$? 为回答这个问题, 先证明下面的引理.

引理1. **引理 6.4.1** 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是线性空间 V 的一组基, 任给 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\in V$, 则一定存在唯一的线性变换 T, 使得

$$T\varepsilon_i=\boldsymbol{\alpha}_i,\qquad i=1,2,\cdots,n.$$

证明. $\forall\ \xi=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n\in V$, 定义

$$T\xi=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n.$$

易见, $T\varepsilon_i=\alpha_i$, $i=1,2,\ldots,n$. $\forall\ \xi_1,\xi_2\in V$,

$$\xi_1=y_1\varepsilon_1+y_2\varepsilon_2+\cdots+y_n\varepsilon_n,$$
$$\xi_2=z_1\varepsilon_1+z_2\varepsilon_2+\cdots+z_n\varepsilon_n,$$

以及 $\forall\ k_1,k_2\in K$,

$$T(k_1\xi_1+k_2\xi_2)=T\big((k_1y_1+k_2z_1)\,\varepsilon_1+(k_1y_2+k_2z_2)\,\varepsilon_2+\cdots+(k_1y_n+k_2z_n)\,\varepsilon_n\big)$$
$$=(k_1y_1+k_2z_1)\,\alpha_1+(k_1y_2+k_2z_2)\,\alpha_2+\cdots+(k_1y_n+k_2z_n)\,\alpha_n$$

有

$$=(k_1y_1+k_2z_1)\,\alpha_1+(k_1y_2+k_2z_2)\,\alpha_2+\cdots+(k_1y_n+k_2z_n)\,\alpha_n$$
$$=k_1\,\underline{(y_1\alpha_1+y_2\alpha_2+\cdots+y_n\alpha_n)}+k_2\,\underline{(z_1\alpha_1+z_2\alpha_2+\cdots+z_n\alpha_n)}=k_1\underline{T\xi_1}+k_2\underline{T\xi_2},$$

即如此定义的 T 是线性变换。

根据定理 9，这个 T 是满足要求的唯一线性变换。

及

$$\begin{aligned}T(\boldsymbol{\alpha}+\boldsymbol{\beta})&=(a_1+b_1)\boldsymbol{\alpha}_1+(a_2+b_2)\boldsymbol{\alpha}_2+\cdots+(a_n+b_n)\boldsymbol{\alpha}_n\\&=(a_1\boldsymbol{\alpha}_1+a_2\boldsymbol{\alpha}_2+\cdots+a_n\boldsymbol{\alpha}_n)+(b_1\boldsymbol{\alpha}_1+b_2\boldsymbol{\alpha}_2+\cdots+b_n\boldsymbol{\alpha}_n)\\&=T\boldsymbol{\alpha}+T\boldsymbol{\beta},\end{aligned}$$
$$T(k\boldsymbol{\alpha})=ka_1\boldsymbol{\alpha}_1+ka_2\boldsymbol{\alpha}_2+\cdots+ka_n\boldsymbol{\alpha}_n=kT\boldsymbol{\alpha}.$$

即 T 是线性变换, 而且满足 $T\varepsilon_i=\boldsymbol{\alpha}_i$, $i=1,2,\cdots,n$, 再由定理 6.4.2, T 是唯一的. □

由此引理可得

定理 10. 设 $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ 是线性空间 $V(K)$ 的一组基。则V上的线性变换 T 按照规则 $(T\varepsilon_1\ T\varepsilon_2\ \cdots\ T\varepsilon_n)=(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)A$ 与数域 K 上 n 阶方阵 A 构成一一对应。

…变换 T 与 n 阶方阵 …$\boldsymbol{A}$ 的第 i 个列向量

证明. 只需证明，$\forall$ 数域 K上的 n 阶方阵 A，$\exists\ V$上的线性变换 T 使得

$$(T\varepsilon_1\ T\varepsilon_2\ \cdots\ T\varepsilon_n)=(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)A.$$

设

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

令

$$\alpha_i=a_{1i}\varepsilon_1+a_{2i}\varepsilon_2+\cdots+a_{ni}\varepsilon_n,\ i=1,2,\ldots,n.$$

根据引理 1，存在唯一线性变换 T 使得

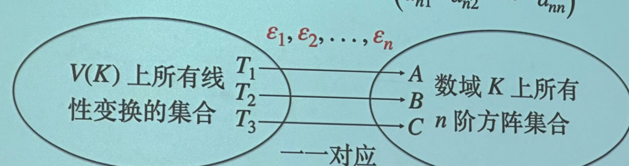
$$T\varepsilon_i=\alpha_i,\ i=1,2,\ldots,n. \quad (6.11)$$

易见，

$$(T\varepsilon_1\ T\varepsilon_2\ \cdots\ T\varepsilon_n)=(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}.$$

按照规则

$$(T\varepsilon_1\ T\varepsilon_2\ \cdots\ T\varepsilon_n)=(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&A\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}, \qquad \alpha_2\cdots\alpha_n)$$



得该定理. □

建立了一一对应:

$\forall\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n\in V,$

$$T\alpha=(T\varepsilon_1\ T\varepsilon_2\ \cdots\ T\varepsilon_n)\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}=(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)\boxed{A\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}}.$$

转化为对矩阵 $\boldsymbol{A}$ 的研究. 例如,

例 6.4.4　已知 $D: D[p(x)]=\dfrac{\mathrm{d}}{\mathrm{d}x}p(x)$ 是线性空间 $P_n[x]$ 上的线性变换, 求 D 在基 $1,x,x^2,\cdots,x^n$ 下的矩阵.

解

$$\begin{aligned}(D(1),D(x),\cdots,D(x^n))&=(0,1,\cdots,nx^{n-1})\\&=(1,x,x^2,\cdots,x^n)\begin{pmatrix}0&1&0&\cdots&0\\0&0&2&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&n\\0&0&0&\cdots&0\end{pmatrix},\end{aligned}$$

即 D 在基 $1,x,x^2,\cdots,x^n$ 下的矩阵是 $n+1$ 阶方阵

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

□

例 6.4.5 V 是一个 n 维的线性空间, 则 V 上的恒等变换 E 在任一组基下的矩阵都是单位矩阵 $\boldsymbol{E}_n$; V 上的数乘变换 T_k 在任一组基下的矩阵都是数量矩阵 $k\boldsymbol{E}_n$; 而零变换 T_0 在任一组基下的矩阵都是零矩阵.

例10~ **例 6.4.6** 设 T 是 $\mathbf{R}^3$ 到 $\mathbf{R}^3$ 的线性变换:

$$T: T(x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}) = x\boldsymbol{i} + y\boldsymbol{j}.$$

这其实是把空间向量投影到 xOy 面上的线性变换.

(1) 求 T 在基底 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 下的矩阵;

(2) 求 T 在基底 $\boldsymbol{\alpha} = \boldsymbol{i}, \boldsymbol{\beta} = \boldsymbol{j}, \boldsymbol{\gamma} = \boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k}$ 下的矩阵.

解 (1) 由 T 的定义可知

$$T(\boldsymbol{i}) = \boldsymbol{i}, \quad T(\boldsymbol{j}) = \boldsymbol{j}, \quad T(\boldsymbol{k}) = \boldsymbol{0},$$

即

$$T(\boldsymbol{ijk}) = (\boldsymbol{ijk}) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

故 T 在基底 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 下的矩阵是 $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

(2) 由题设

$$T(\boldsymbol{\alpha}) = \boldsymbol{i} = \boldsymbol{\alpha}, \quad T(\boldsymbol{\beta}) = \boldsymbol{j} = \boldsymbol{\beta}, \quad T(\boldsymbol{\gamma}) = \boldsymbol{i} + \boldsymbol{j} = \boldsymbol{\alpha} + \boldsymbol{\beta},$$

即

$$T(\boldsymbol{\alpha\beta\gamma}) = (\boldsymbol{\alpha\beta\gamma}) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

故 T 在基底 $\boldsymbol{\alpha} = \boldsymbol{i}, \boldsymbol{\beta} = \boldsymbol{j}, \boldsymbol{\gamma} = \boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k}$ 下的矩阵为 $\boldsymbol{B} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. □

上例说明, 同一个线性变换在不同的基底下, 所对应的矩阵是不同的, 我们自然要问: 不同的基底是否对应不同的矩阵, 它们之间有什么样的关系?

定理11 **定理 6.4.4**　在 n 维线性空间 V 中取定两组基底 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 与 $\omega_1, \omega_2, \cdots, \omega_n$, 设由基底 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\omega_1, \omega_2, \cdots, \omega_n$ 的过渡矩阵为

$$\boldsymbol{P} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \quad (\boldsymbol{P}\text{可逆}),$$

即 $(\omega_1 \omega_2 \cdots \omega_n) = (\varepsilon_1 \varepsilon_2 \cdots \varepsilon_n)\boldsymbol{P}$. 并设 V 上的线性变换 T 在这两组基底下的矩阵分别是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 则

$$\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}.$$

即 $\boldsymbol{A}$ 相似于 $\boldsymbol{B}$ $(\boldsymbol{A} \sim \boldsymbol{B})$.

证明　由

$$T(\varepsilon_1 \varepsilon_2 \cdots \varepsilon_n) = (\varepsilon_1 \varepsilon_2 \cdots \varepsilon_n)\boldsymbol{A},$$
$$T(\omega_1 \omega_2 \cdots \omega_n) = (\omega_1 \omega_2 \cdots \omega_n)\boldsymbol{B}$$

得

$$T(\omega_1 \omega_2 \cdots \omega_n) = (\varepsilon_1 \varepsilon_2 \cdots \varepsilon_n)\boldsymbol{P}\boldsymbol{B}. \tag{6.12}$$

故有

$$\underline{(T\omega_1\ T\omega_2\ \cdots\ T\omega_n) = (\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)PB}.$$

另一方面，由

$$(\omega_1, \omega_2, \ldots, \omega_n) = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix},$$

可得

$$(T\omega_1\ T\omega_2\ \cdots\ T\omega_n) = (T\varepsilon_1\ T\varepsilon_2\ \cdots\ T\varepsilon_n)\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix},$$

$$\underline{(T\omega_1\ T\omega_2\ \cdots\ T\omega_n) = (\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)AP}. \tag{6.13}$$

注意到在同一基底下, 坐标是唯一的, 由式 (6.12) 及式 (6.13) 得

$$\boldsymbol{P}\boldsymbol{B} = \boldsymbol{A}\boldsymbol{P}.$$

因 $\boldsymbol{P}$ 可逆, 于是有

$$\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}.$$

□

定理12. **定理 6.4.5**　设 $\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶方阵，则 $\boldsymbol{A}\sim\boldsymbol{B}$ 的充要条件是：它们是 n 维的线性空间 V 上的某个线性变换 T 在不同基底下的矩阵.

证明. ($\Leftarrow$) 根据定理 11.

($\Rightarrow$) 设 $B=P^{-1}AP$，$\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n$ 是线性空间 $V(K)$ 的一组基。

根据定理 10，存在$V(K)$ 上某个线性变换 T 使得：

$$(T\varepsilon_1\ T\varepsilon_2\ \cdots\ T\varepsilon_n)=(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)A.$$

令

$$(\omega_1,\omega_2,\dots,\omega_n)=(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n)P,$$

则 $\omega_1,\omega_2,\dots,\omega_n$ 也是$V(K)$ 的一组基，且

$$\begin{aligned}(T\omega_1\ T\omega_2\ \cdots\ T\omega_n)&=(T\varepsilon_1\ T\varepsilon_2\ \cdots\ T\varepsilon_n)P=(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)AP\\&=(\omega_1\ \omega_2\ \cdots\ \omega_n)P^{-1}AP=(\omega_1\ \omega_2\ \cdots\ \omega_n)B.\end{aligned}$$

必要性.

线性变换 T 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵,

$_2\cdots\varepsilon_n)\boldsymbol{P}$.

.4, T 在这组基底下的矩阵就是 $\boldsymbol{P}^{-1}\boldsymbol{AP}$, □

例11、**例 6.4.7**　设线性变换 T 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵是

$$\boldsymbol{A}=\begin{pmatrix}1&3&2&8\\5&7&0&1\\3&0&1&3\\1&1&2&2\end{pmatrix}.$$

试求 T 在基 $\boldsymbol{\omega}_1=\varepsilon_1,\boldsymbol{\omega}_2=\varepsilon_1+\varepsilon_2,\boldsymbol{\omega}_3=\varepsilon_1+\varepsilon_2+\varepsilon_3,\boldsymbol{\omega}_4=\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4$ 下的矩阵.

解　法一：由 $\boldsymbol{A}$ 的定义知

$$\begin{aligned}T\varepsilon_1&=\varepsilon_1+5\varepsilon_2+3\varepsilon_3+\varepsilon_4,\\T\varepsilon_2&=3\varepsilon_1+7\varepsilon_2+\varepsilon_4,\\T\varepsilon_3&=2\varepsilon_1+\varepsilon_3+2\varepsilon_4,\\T\varepsilon_4&=8\varepsilon_1+\varepsilon_2+3\varepsilon_3+2\varepsilon_4.\end{aligned}$$

于是

$$\begin{aligned}T\boldsymbol{\omega}_1&=T\varepsilon_1=\varepsilon_1+5\varepsilon_2+3\varepsilon_3+\varepsilon_4\\&=-4\boldsymbol{\omega}_1+2\boldsymbol{\omega}_2+2\boldsymbol{\omega}_3+\boldsymbol{\omega}_4,\\T\boldsymbol{\omega}_2&=T(\varepsilon_1+\varepsilon_2)=4\varepsilon_1+12\varepsilon_2+3\varepsilon_3+2\varepsilon_4\\&=-8\boldsymbol{\omega}_1+9\boldsymbol{\omega}_2+\boldsymbol{\omega}_3+2\boldsymbol{\omega}_4,\\T\boldsymbol{\omega}_3&=T(\varepsilon_1+\varepsilon_2+\varepsilon_3)=6\varepsilon_1+12\varepsilon_2+4\varepsilon_3+4\varepsilon_4\\&=-6\boldsymbol{\omega}_1+8\boldsymbol{\omega}_2+4\boldsymbol{\omega}_4,\\T\boldsymbol{\omega}_4&=T(\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4)=14\varepsilon_1+13\varepsilon_2+7\varepsilon_3+6\varepsilon_4\\&=\boldsymbol{\omega}_1+6\boldsymbol{\omega}_2+\boldsymbol{\omega}_3+6\boldsymbol{\omega}_4.\end{aligned}$$

因此 T 在基 $\boldsymbol{\omega}_1,\boldsymbol{\omega}_2,\boldsymbol{\omega}_3,\boldsymbol{\omega}_4$ 下的矩阵为

$$\boldsymbol{B}=\begin{pmatrix}-4&-8&-6&1\\2&9&8&6\\2&1&0&1\\1&2&4&6\end{pmatrix}.$$

解. 易见

$$\omega_1=(\varepsilon_1\ \varepsilon_2\ \varepsilon_3\ \varepsilon_4)\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\ \omega_2=(\varepsilon_1\ \varepsilon_2\ \varepsilon_3\ \varepsilon_4)\begin{pmatrix}1\\1\\0\\0\end{pmatrix},\ \omega_3=(\varepsilon_1\ \varepsilon_2\ \varepsilon_3\ \varepsilon_4)\begin{pmatrix}1\\1\\1\\0\end{pmatrix},\ \omega_4=(\varepsilon_1\ \varepsilon_2\ \varepsilon_3\ \varepsilon_4)\begin{pmatrix}1\\1\\1\\1\end{pmatrix}.$$

令 $(\omega_1,\omega_2,\omega_3,\omega_4)=(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)P$，则有

法二：根据定理 6.4.4, 利用相似原理, 由于从基底 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 到 $\omega_1,\omega_2,\omega_3,\omega_4$ 的过渡矩阵为

$$\boldsymbol{P}=\begin{pmatrix}1&1&1&1\\0&1&1&1\\0&0&1&1\\0&0&0&1\end{pmatrix}.$$

所以

$$\begin{aligned}\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}&=\begin{pmatrix}1&-1&0&0\\0&1&-1&0\\0&0&1&-1\\0&0&0&1\end{pmatrix}\begin{pmatrix}1&3&2&8\\5&7&0&1\\3&0&1&3\\1&1&2&2\end{pmatrix}\begin{pmatrix}1&1&1&1\\0&1&1&1\\0&0&1&1\\0&0&0&1\end{pmatrix}\\&=\begin{pmatrix}-4&-8&-6&1\\2&9&8&6\\2&1&0&1\\1&2&4&6\end{pmatrix}.\end{aligned}$$

□

6.5　线性变换的特征值和特征向量

6.5.1　线性变换的特征值和特征向量

我们知道, n 维的线性空间 V 上的某个线性变换 T 在不同基底下的矩阵是不同的, 能否找到一组基, 使得 T 在这组基下的矩阵具有尽可能简单的形式? 具体而言, 在本节我们要设法找到这样一组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n,$$

使得 T 在这组基下的矩阵具有对角形, 即

$$(T\varepsilon_1 T\varepsilon_2\cdots T\varepsilon_n)=(\varepsilon_1\varepsilon_2\cdots\varepsilon_n)\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix}.$$

亦即

$$\begin{aligned}T\varepsilon_1&=\lambda_1\varepsilon_1,\\T\varepsilon_2&=\lambda_2\varepsilon_2,\\&\cdots\cdots\\T\varepsilon_n&=\lambda_n\varepsilon_n.\end{aligned}$$

虽然上述过程并不一定能做到, 但也提供了一个方法: 找出满足

$$T\boldsymbol{\xi}=\lambda\boldsymbol{\xi}$$

的数 λ 和非零向量 $\boldsymbol{\xi}$, 这样就有可能使得 T 在这组基下的矩阵具有对角形.

定义15、**定义 6.5.1** (线性变换的特征值与特征向量)　设 V 是数域 K 上线性空间, T 是 V 上的一个线性变换, 若对 K 中的一个数 λ, 存在 V 的一个非零向量 $\boldsymbol{\xi}$, 使得

$$T\boldsymbol{\xi}=\lambda\boldsymbol{\xi},$$

则称 λ 是线性变换 T 的一个**特征值**, $\boldsymbol{\xi}$ 是 T 的属于 λ 的**特征向量**.

特征值和特征向量的概念不仅对线性变换的研究具有重要的意义, 而且在物理、力学和工程技术中也有实际的意义.

定理 6.5.1　设 V 是数域 K 上线性空间, T 是 V 上的一个线性变换, 则 T 的属于特征值 λ 的所有特征向量和零向量一起构成了 V 的一个子空间, 该子空间记为

$$V_\lambda=\{\boldsymbol{\xi}\in V: T\boldsymbol{\xi}=\lambda\boldsymbol{\xi}\}.$$

证明　任取 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2\in V_\lambda$, 当 $\boldsymbol{\xi}_1+\boldsymbol{\xi}_2\neq\mathbf{0}$ 时, 因

$$T(\boldsymbol{\xi}_1+\boldsymbol{\xi}_2)=T\boldsymbol{\xi}_1+T\boldsymbol{\xi}_2=\lambda\boldsymbol{\xi}_1+\lambda\boldsymbol{\xi}_2=\lambda(\boldsymbol{\xi}_1+\boldsymbol{\xi}_2),$$

即 $\boldsymbol{\xi}_1+\boldsymbol{\xi}_2$ 也是 T 的属于特征值 λ 的特征向量; 又若 $\boldsymbol{\xi}$ 是 T 的属于 λ 的特征向量, 对任意的 $k\in K, k\neq 0$, 因

$$T(k\boldsymbol{\xi})=kT\boldsymbol{\xi}=k\lambda\boldsymbol{\xi}=\lambda(k\boldsymbol{\xi}),$$

即 $k\boldsymbol{\xi}$ 也是 T 的属于特征值 λ 的特征向量. 这说明 V_λ 的元素对加法和数乘运算是封闭的, 于是 V_λ 是子空间. □

定义16·**定义 6.5.2** (特征子空间)　称 V_λ 为线性变换 T 对应于特征值 λ 的**特征子空间**.

由于线性变换 T 的属于 λ 的特征向量有无穷多个, 所以对应于特征值 λ 的特征子空间 V_λ 是一个非零子空间. 因此为了表示出 V_λ 的所有向量, 需要求出 V_λ 的一组基. 实际操作中, 我们可以先求出 T 的所有特征值, 然后再求出 T 的每个特征值对应的线性无关的特征向量 (V_λ 的一组基), 讨论如下.

线性空间 V 上一个线性变换 T 在选定的一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下对应于一个矩阵 $\boldsymbol{A}$, 即

$$(T\varepsilon_1 T\varepsilon_2\cdots T\varepsilon_n)=(\varepsilon_1\varepsilon_2\cdots\varepsilon_n)\boldsymbol{A},$$

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

设 $\lambda\in K$ 是 T 的特征值, ξ 是 T 的属于 λ 的特征向量, 即

$$T\xi=\lambda\xi,\qquad \xi\neq 0.$$

也就是

$$T\xi=\lambda\left(x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n\right)=\left(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n\right)\lambda\boxed{\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}}.$$

另一方面，由于 T 是线性变换，

$$\begin{aligned}T\xi&=T\left(x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n\right)\\&=\left(T\varepsilon_1\ T\varepsilon_2\ \cdots\ T\varepsilon_n\right)\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}=\left(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n\right)A\boxed{\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}}.\end{aligned}$$

由于向量 $T\xi$ 在基 $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ 下的坐标是唯一的，故有

$$A\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}=\lambda\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}.$$

(6.14)

反之，若 $\lambda\in K$ 是矩阵 A 的特征值，$x=\left(x_1\ x_2\ \cdots\ x_n\right)^T$ 是 A 的属于 λ 特征向量，则 $\xi=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n\in V$，$\xi\neq 0$，且

解的充要条件是

$$T\xi=T\left(x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n\right)=\left(T\varepsilon_1\ T\varepsilon_2\ \cdots\ T\varepsilon_n\right)\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}.$$

(6.15)

$$=\left(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n\right)A\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}=\left(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n\right)\lambda\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}=\lambda\xi.$$

$$\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$$

$$\left(T\varepsilon_1\ T\varepsilon_2\ \cdots\ T\varepsilon_n\right)=\left(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n\right)A$$

$$\begin{array}{ccc}T & \qquad & A\\ \lambda & & \lambda\end{array}$$

$$\xi=\left(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n\right)\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}\qquad x=\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix}$$

λ 是 T 的特征值 $\Longleftrightarrow$ $\lambda\in K$ 是 A 的特征值

ξ 是 T 的属于 λ 的特征向量 $\Longleftrightarrow$ x 是 A 的属于 λ 的特征向量

矩阵 A 称为 T 的**特征矩阵**，$\left|\lambda E-A\right|$ 称为 T 的**特征多项式**。

定理 13. 线性变换 T 的特征值与特征向量与基的选取无关。

值. 从而可知:

征向量, 则 λ 是矩阵 $\boldsymbol{A}$ 的特征

是矩阵 $\boldsymbol{A}$ 的特征值, 则齐次线

λ 的一个特征向量. 由此可见,

征值、特征向量的性质与计算而

得到.

计算线性变换 T 的特征值和特征向量的步骤如下:

(1) 给定 V 的一组基, 求出 T 在这组基下的矩阵 $\boldsymbol{A}$;

(2) 计算特征多项式 $p(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|$;

(3) 求 $p(\lambda)=0$ 的含于 K 的根 (T 的特征值)

$$\lambda_1,\lambda_2,\cdots,\lambda_n;$$

(4) 对每个 $\lambda_i(i=1,2\cdots,n)$, 求齐次线性方程组

$$(\lambda_i \boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta}$$

的一个基础解系;

(5) 以求出的基础解系为坐标写出 V 的一个向量组, 它就是 $V_{\lambda_i}(i=1,2,\cdots,n)$ 的一组基.

若计算 $\boldsymbol{A}$ 的特征值和特征向量, 只需直接进行第 (2)、(3)、(4) 步, 且第 (3) 步须求出 $p(\lambda)=0$ 的全部复根 (即 $\boldsymbol{A}$ 的全部特征值).

例 6.5.1　设数域 $K=\mathbf{R}$, 线性变换 T 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵是

$$\boldsymbol{A}=\begin{pmatrix}1&2&0\\2&1&0\\0&0&-1\end{pmatrix}.$$

求 T 的特征值和特征向量.

解　矩阵 $\boldsymbol{A}$ 的特征多项式是

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1&-2&0\\-2&\lambda-1&0\\0&0&\lambda+1\end{vmatrix}=(\lambda+1)^2(\lambda-3),$$

所以矩阵 $\boldsymbol{A}$ 的全部特征值是 $\lambda_1=-1,\lambda_2=-1,\lambda_3=3$, 它们都是实数，所以也是线性变换 T 的全部特征值.

对应于特征值 $\lambda_1=-1,\lambda_2=-1$, 求得齐次线性方程组

$$(\lambda_i\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta},\quad i=1,2$$

的一个基础解系为

$$\begin{pmatrix}-1\\1\\0\end{pmatrix},\quad\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

因此, 线性变换 T 属于特征值 -1 的两个线性无关的特征向量为

$$\boldsymbol{\xi}_1=\varepsilon_2-\varepsilon_1,\quad \boldsymbol{\xi}_2=\varepsilon_3.$$

对应于特征值 $\lambda_3=3$, 求得齐次线性方程组

$$(\lambda_3\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta}$$

的一个基础解系为

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

因此, 线性变换 T 属于特征值 3 的一个线性无关的特征向量为 $\boldsymbol{\xi}_3 = \varepsilon_1 + \varepsilon_2$. □

例 6.5.2　平面上的全体向量构成了实数域上的一个二维线性空间 $\mathbf{R}^2$, 在直角坐标系下取单位坐标向量 $\boldsymbol{e}_1, \boldsymbol{e}_2$ 作为一组基. 设转角为 θ 的旋转变换 $T_\theta : \mathbf{R}^2 \to \mathbf{R}^2$ 为

$$T_\theta(x, y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta),$$

则 T_θ 是 $\mathbf{R}^2$ 上的线性变换. T_θ 在基 $\boldsymbol{e}_1, \boldsymbol{e}_2$ 下的矩阵为

$$\boldsymbol{A} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

$\boldsymbol{A}$ 的特征多项式

$$p(\lambda) = |\lambda\boldsymbol{E} - \boldsymbol{A}| = \begin{vmatrix} \lambda - \cos\theta & \sin\theta \\ -\sin\theta & \lambda - \cos\theta \end{vmatrix} = \lambda^2 - 2\lambda\cos\theta + 1.$$

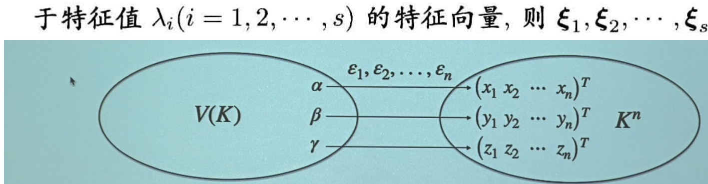
如果 $\theta \neq n\pi$, 该二次方程仅有复数根. 因 $\mathbf{R}^2$ 是实数域上的线性空间, 故 T_θ 没有特征值, 也没有特征向量. □

上述结果与解析几何中的结果是一致的, 即当 $\theta \neq n\pi$ 时, 平面上不存在非零向量 $\boldsymbol{\xi}$, 满足 $T\boldsymbol{\xi} = \lambda\boldsymbol{\xi}$, 亦即不存在非零向量 $\boldsymbol{\xi}$, 旋转 θ 角后仍然落在原向量所在的直线上.

在 n 维线性空间 V 上取定一组基后, 线性变换的特征值一定是它在这组基下的矩阵的特征值. 选择不同的基, 线性变换对应的矩阵就不同, 但据定理 6.4.5, 这些不同基下的矩阵是相似的. 而相似矩阵具有相同的特征多项式和相同的特征值, 从而知线性变换的特征值一定是它在任一基下的矩阵的特征值. 我们也把线性变换 T 在任一组基下的矩阵的特征多项式称为 T 的**特征多项式**, 而将 T 在任一组基下的矩阵的特征矩阵称为 T 的**特征矩阵**. 于是有下面的结论:

定理13 **定理 6.5.2**　有限维线性空间上的线性变换的特征值和特征多项式与所选基底无关.

定理15 **定理 6.5.3**　设 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是线性变换 T(或矩阵 $\boldsymbol{A}$) 的 s 个互异的特征值, $\boldsymbol{\xi}_i$ 是属于特征值 $\lambda_i (i = 1, 2, \cdots, s)$ 的特征向量, 则 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_s$ 线性无关.



定理 14. 设 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 是 n 维线性空间 $V(K)$ 的一组基，则 $\alpha_1, \alpha_2, \ldots, \alpha_s \in V(K)$ 线性无关当且仅当

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{n1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{n2} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1s} \\ a_{2s} \\ \cdots \\ a_{ns} \end{pmatrix},$$

在 K^n 中线性无关，其中

$$\alpha_i = a_{1i}\varepsilon_1 + a_{2i}\varepsilon_2 + \cdots + a_{ni}\varepsilon_n,\ i = 1,2,\ldots,s.$$

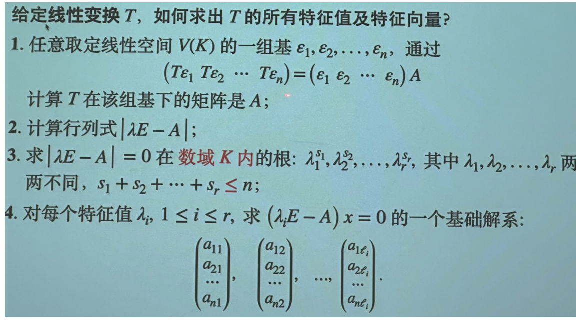
性无关.

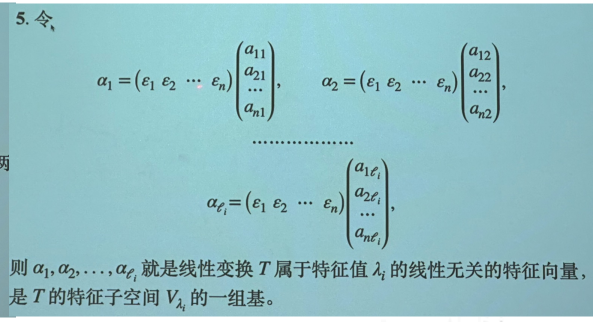
下面证明对应于 $s+1$ 个互异特征

性无关.

$\xi_{s+1} = \mathbf{0}$.　　(6.16)

$k_{s+1}\lambda_{s+1}\boldsymbol{\xi}_{s+1} = \mathbf{0}$.





对 (6.16) 两边同时作用 T 得

$$k_1\lambda_1\boldsymbol{\xi}_1 + k_2\lambda_2\boldsymbol{\xi}_2 + \cdots + k_s\lambda_s\boldsymbol{\xi}_s + k_{s+1}\lambda_{s+1}\boldsymbol{\xi}_{s+1} = \mathbf{0}.$$

从而有

$$k_1(\lambda_{s+1} - \lambda_1)\boldsymbol{\xi}_1 + \cdots + k_s(\lambda_{s+1} - \lambda_s)\boldsymbol{\xi}_s = \mathbf{0}.$$

由归纳假设, $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_s$ 线性无关, 则

$$k_i(\lambda_{s+1} - \lambda_i) = 0, \qquad i = 1, 2, \cdots, s.$$

因 $\lambda_{s+1} - \lambda_i \neq 0 (i \leqslant s)$, 故 $k_i = 0, (i = 1, 2, \cdots, s)$. 从而 (6.16) 变成

$$k_{s+1}\boldsymbol{\xi}_{s+1} = \mathbf{0}.$$

因 $\boldsymbol{\xi}_{s+1} \neq \mathbf{0}$, 所以 $k_{s+1} = 0$. 因此 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_s, \boldsymbol{\xi}_{s+1}$ 线性无关. 定理证毕. $\square$

6.5.2 线性变换的最简表示

对角矩阵是矩阵中最简单的一种, 本节将研究对于 n 维的线性空间 V 上的某个线性变换 T 是否存在一组基使得 T 在该组基下的矩阵为对角矩阵.

定义 6.5.3 设 T 是 n 维线性空间 V 上的一个线性变换, 若 T 在某组基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$ 下的矩阵为对角矩阵, 则称 T 有**最简表示**.

下面讨论什么样的线性变换有最简表示.

设 n 维线性空间 V 上的一个线性变换 T 在基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$ 下的矩阵为对角矩阵 $\boldsymbol{A}$, 即

$$(T\boldsymbol{\varepsilon}_1 T\boldsymbol{\varepsilon}_2 \cdots T\boldsymbol{\varepsilon}_n) = (\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_2 \cdots \boldsymbol{\varepsilon}_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

因此

$$T\boldsymbol{\varepsilon}_i = \lambda_i\boldsymbol{\varepsilon}_i, \qquad i = 1, 2, \cdots, n.$$

于是 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$ 是 T 的 n 个线性无关的特征向量.

另一方面, 如果 T 有 n 个线性无关的特征向量 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$,

$$T\boldsymbol{\varepsilon}_i = \lambda_i\boldsymbol{\varepsilon}_i, \qquad i = 1, 2, \cdots, n.$$

取 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$ 为 V 的一组基, 显见

$$(T\boldsymbol{\varepsilon}_1 T\boldsymbol{\varepsilon}_2 \cdots T\boldsymbol{\varepsilon}_n) = (\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_2 \cdots \boldsymbol{\varepsilon}_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

即 T 在基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$ 下的矩阵是对角矩阵.

上面的讨论证明了如下定理

定理 6.5.4　n 维线性空间 V 上的一个线性变换 T 有最简表示的充要条件是 T 有 n 个线性无关的特征向量.

现在我们来看在 n 维线性空间 V 上，什么样的线性变换 T 会有 n 个线性无关的特征向量, 定理6.5.3 给出了一个充分条件, 并由此得到:

推论 6.5.5　若线性变换 T 有 n 个互异的特征值, 则 T 有最简表示.

应用定理6.5.3 还可以证明

定理 6.5.6　若 λ_1, λ_2 是线性变换 T 的两个不同的特征值, $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_s$ 是 T 的属于 λ_1 的线性无关的特征向量, $\boldsymbol{\nu}_1, \boldsymbol{\nu}_2, \cdots, \boldsymbol{\nu}_t$ 是 T 的属于 λ_2 的线性无关的特征向量, 则 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_s; \boldsymbol{\nu}_1, \boldsymbol{\nu}_2, \cdots, \boldsymbol{\nu}_t$ 线性无关.

证明　参见定理 4.3.3. □

对于同一个特征值对应的特征向量, 有如下结论:

定理 6.5.7　若 λ_0 是线性变换 T 的 s 重特征值, 则 T 的属于 λ_0 的特征向量中, 线性无关的最大组包含的向量的个数不超过 s.

证明　略. □

定理 6.5.8　n 维线性空间上的线性变换 T 有最简表示 $\Longleftrightarrow$ T 有 n 个特征值 (包括重数)，且对 T 的每个 s_i 重特征值 λ_i, 其对应的特征矩阵 $\lambda_i \boldsymbol{E} - \boldsymbol{A}$ 的秩为 $n - s_i$.

证明　注意到 $(\lambda_i \boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{\theta}$ 的基础解系包含 $n - (n - s_i)$, 即 s_i 个线性无关的特征向量, 而 s_i 之和是 n, 由定理 4.3.3 可知, $\boldsymbol{A}$ 有 n 个线性无关的特征向量, 则 $\boldsymbol{A}$ 与对角矩阵相似，即 T 有最简表示. 反之亦然. □

由上面的定理, 判别线性变换 T 的矩阵是否有最简表示, 只需计算特征值与特征矩阵的秩即可.

例 6.5.3　设 V 上线性变换 T 在基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2$ 下的矩阵为

$$\boldsymbol{A} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

分别把 V 看作实线性空间和复线性空间, T 是否有最简表示? 若有，则求其最简表示.

解　若 V 是实线性空间,

$$\begin{aligned} p(\lambda) = |\lambda \boldsymbol{E} - \boldsymbol{A}| &= \begin{vmatrix} \lambda - 1 & 1 \\ -1 & \lambda - 1 \end{vmatrix} \\ &= \lambda^2 - 2\lambda + 2 = [\lambda - (1+i)][\lambda - (1-i)]. \end{aligned}$$

因此在实数域上 $\boldsymbol{A}$ 没有特征值, 故没有特征向量, 从而 T 也没有特征值与特征向量，所以没有最简表示.

若 V 是复线性空间, $\boldsymbol{A}$ 有特征值 $1+i$ 和 $1-i$，这也是 T 的特征值，且都是单特征值，故 T 有最简表示.

对应于特征值 $\lambda_1 = 1 + i$, 解齐次线性方程组

$$(\lambda_1 \boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{\theta},$$

得到相应的特征子空间的一个基础解系, $\boldsymbol{\xi}_1=(i,1)^{\mathrm{T}}$. 对应于特征值 $\lambda_2=1-i$, 解齐次线性方程组

$$(\lambda_2\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta},$$

得到相应的特征子空间的一个基础解系 $\boldsymbol{\xi}_2=(-i,1)^{\mathrm{T}}$. 二阶方阵 $\boldsymbol{A}$ 有两个线性无关的特征向量 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2$, 所以 $\boldsymbol{A}$ 可以对角化. 令

$$\boldsymbol{P}=\begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix},$$

则

$$\begin{aligned}\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}&=\frac{1}{2}\begin{pmatrix} -i & 1 \\ i & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}\\ &=\begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}.\end{aligned}$$

注意到 $\boldsymbol{A}$ 是 T 在基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2$ 下的矩阵，令 $(\boldsymbol{\eta}_1\boldsymbol{\eta}_2)=(\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_2)\boldsymbol{P}$, 则 T 的最简表示为

$$(T\boldsymbol{\eta}_1T\boldsymbol{\eta}_2)=(\boldsymbol{\eta}_1\boldsymbol{\eta}_2)\begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}.$$

□

6.5.3* 不变子空间

研究线性变换的一个重要的方法是把它所作用的空间进行分解, 这一节介绍线性变换的不变子空间的概念, 并由此说明线性变换的矩阵的化简与线性变换的内在联系.

定义 6.5.4　设 T 是线性空间 V 上的一个线性变换, W 是 V 的一个子空间, 如果对任意的 $\boldsymbol{\alpha}\in W$, 都有 $T\boldsymbol{\alpha}\in W$, 则称 W 是 T 的一个**不变子空间**.

例 6.5.4　线性空间 V 的任何一个子空间都是数乘变换的不变子空间, 因为子空间对数量乘法是封闭的.

例 6.5.5　零子空间 $\{\mathbf{0}\}$ 和整个线性空间 V, 对每个线性变换 T 而言显然都是 T 的不变子空间, 称 $\{\mathbf{0}\}$ 和 V 为 T 的**平凡不变子空间**. 除此之外的不变子空间称为 T 的**非平凡不变子空间**.

例 6.5.6　证明线性空间 V 上的线性变换 T 的像空间 $\mathcal{R}(T)$, 核空间 $\ker(T)$ 及特征子空间 V_λ 都是 T 的不变子空间.

证明　任取 $\boldsymbol{\alpha}\in\mathcal{R}(T)$, 因 $T\boldsymbol{\alpha}\in\mathcal{R}(T)$, 所以 $\mathcal{R}(T)$ 是 T 的不变子空间.

任取 $\boldsymbol{\alpha}\in\ker(T)$, 因 $T\boldsymbol{\alpha}\in\ker(T)$, 所以 $\ker(T)$ 是 T 的不变子空间.

设 V_λ 是 T 的任一特征子空间, 任取 $\boldsymbol{\alpha}\in V_\lambda$, 因 $T\boldsymbol{\alpha}=\lambda\boldsymbol{\alpha}\in V_\lambda$, 所以 V_λ 是 T 的不变子空间. □

例 6.5.7　证明线性变换 T 的不变子空间的和与交还是 T 的不变子空间.

证明　证明略, 留作习题. □

例 6.5.8　设线性空间 V 的子空间 W 由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 生成, 即

$$W=\operatorname{span}\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\},$$

则 W 是 T 的不变子空间的充要条件是 $T\boldsymbol{\alpha}_1,T\boldsymbol{\alpha}_2,\cdots,T\boldsymbol{\alpha}_m\in W$.

证明　必要性显然.

若 $T\boldsymbol{\alpha}_1,T\boldsymbol{\alpha}_2,\cdots,T\boldsymbol{\alpha}_m\in W$, 因 W 中的每个向量 $\boldsymbol{\alpha}$ 都可表示成

$$\boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_m\boldsymbol{\alpha}_m,$$

故有

$$T\boldsymbol{\alpha}=x_1T\boldsymbol{\alpha}_1+x_2T\boldsymbol{\alpha}_2+\cdots+x_mT\boldsymbol{\alpha}_m\in W.$$

□

下面讨论不变子空间和线性变换矩阵化简之间的关系.

设 T 是线性空间 V 上的一个线性变换, W 是 V 的关于 T 的一个不变子空间. 在 W 中取一组基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_m$, 并由此扩充为 V 的一组基

$$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_m,\quad \boldsymbol{\varepsilon}_{m+1},\cdots,\boldsymbol{\varepsilon}_n.$$

由于 $T\boldsymbol{\varepsilon}_1,T\boldsymbol{\varepsilon}_2,\cdots,T\boldsymbol{\varepsilon}_m\in W$, 所以有

$$\begin{aligned}
&T\boldsymbol{\varepsilon}_1=a_{11}\boldsymbol{\varepsilon}_1+a_{21}\boldsymbol{\varepsilon}_2+\cdots+a_{m1}\boldsymbol{\varepsilon}_m,\\
&T\boldsymbol{\varepsilon}_2=a_{12}\boldsymbol{\varepsilon}_1+a_{22}\boldsymbol{\varepsilon}_2+\cdots+a_{m2}\boldsymbol{\varepsilon}_m,\\
&\quad\cdots\cdots\\
&T\boldsymbol{\varepsilon}_m=a_{1m}\boldsymbol{\varepsilon}_1+a_{2m}\boldsymbol{\varepsilon}_2+\cdots+a_{mm}\boldsymbol{\varepsilon}_m,\\
&T\boldsymbol{\varepsilon}_{m+1}=a_{1m+1}\boldsymbol{\varepsilon}_1+a_{2m+1}\boldsymbol{\varepsilon}_2+\cdots+a_{mm+1}\boldsymbol{\varepsilon}_m\\
&\qquad\qquad+a_{m+1m+1}\boldsymbol{\varepsilon}_{m+1}+\cdots+a_{nm+1}\boldsymbol{\varepsilon}_n,\\
&\quad\cdots\cdots\\
&T\boldsymbol{\varepsilon}_n=a_{1n}\boldsymbol{\varepsilon}_1+a_{2n}\boldsymbol{\varepsilon}_2+\cdots+a_{mn}\boldsymbol{\varepsilon}_m\\
&\qquad\quad+a_{m+1,n}\boldsymbol{\varepsilon}_{m+1}+\cdots+a_{nn}\boldsymbol{\varepsilon}_n.
\end{aligned}$$

则 T 在基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_m,\boldsymbol{\varepsilon}_{m+1},\cdots,\boldsymbol{\varepsilon}_n$ 下的矩阵为

$$\begin{pmatrix}
a_{11} & \cdots & a_{1m} & a_{1,m+1} & \cdots & a_{1n}\\
\vdots & & \vdots & \vdots & & \vdots\\
a_{m1} & \cdots & a_{mm} & a_{m,m+1} & \cdots & a_{mn}\\
0 & \cdots & 0 & a_{m+1,m+1} & \cdots & a_{m+1,n}\\
\vdots & & \vdots & \vdots & & \vdots\\
0 & \cdots & 0 & a_{n,m+1} & \cdots & a_{nn}
\end{pmatrix}=\begin{pmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_3\\ \boldsymbol{O} & \boldsymbol{A}_2\end{pmatrix}. \tag{6.17}$$

其中，因为 W 是 V 的关于 T 的一个不变子空间，所以像 $T\varepsilon_1, T\varepsilon_2, \cdots, T\varepsilon_m$ 仍属于 W，故可由 W 的基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m$ 线性表出:

$$\begin{aligned}
T\varepsilon_1 &= a_{11}\varepsilon_1 + a_{21}\varepsilon_2 + \cdots + a_{m1}\varepsilon_m, \\
T\varepsilon_2 &= a_{12}\varepsilon_1 + a_{22}\varepsilon_2 + \cdots + a_{m2}\varepsilon_m, \\
&\cdots\cdots \\
T\varepsilon_m &= a_{1m}\varepsilon_1 + a_{2m}\varepsilon_2 + \cdots + a_{mm}\varepsilon_m.
\end{aligned}$$

故 T 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m, \varepsilon_{m+1}, \cdots, \varepsilon_n$ 下的矩阵具有 (6.17) 的形式，m 阶方阵 $\boldsymbol{A}_1$ 是 T 在 W 的基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m$ 下的矩阵.

另一方面，如果 T 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m, \varepsilon_{m+1}, \cdots, \varepsilon_n$ 下的矩阵是 (6.17)，容易证明由 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m$ 生成的子空间是 V 的关于 T 的一个不变子空间.

如果 V 分解成若干个 T 的不变子空间的直和:

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_m,$$

在 T 的每个不变子空间 W_i 中取基

$$\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{in_i}, \qquad i = 1, 2, \cdots, m,$$

把所有这些基合并在一起就成为 V 的一组基. 在这样的一组基下，T 的矩阵具有分块对角形

$$\begin{pmatrix} \boldsymbol{A}_1 & & & \\ & \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_m \end{pmatrix}, \tag{6.18}$$

其中，$\boldsymbol{A}_i (i = 1, 2, \cdots, m)$ 是 T 在 W_i 的基 $\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{in_i}$ 下的矩阵.

反之，若线性变换 T 在所有 $\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{in_i}$ 合并的 V 的基下的矩阵是分块对角矩阵 (6.18)，则由 $\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{in_i}$ 生成的子空间 W_i 是 T 的不变子空间，且 V 是 W_i 的直和.

习 题 六

1. 检验下列集合对于所指定的运算是否构成实数域上的线性空间:

(1) 次数等于 $n(n \geqslant 1)$ 的实系数多项式全体，对于多项式的加法和数乘;

(2) 同阶实对称 (反对称) 矩阵全体，对于矩阵的加法和数乘;

(3) 平面上不平行于某一向量的全体向量所组成的集合，对于向量的加法和数乘;

(4) 平面上全体向量，对于通常的加法和如下定义的数乘: $k\boldsymbol{\alpha} = \mathbf{0}$ (k 为实数，$\boldsymbol{\alpha}$ 为平面向量).

2. 设 $V = \{\boldsymbol{x} : \boldsymbol{x} = c_1 \sin t + c_2 \sin 2t + \cdots + c_n \sin nt, c_k \in \mathbf{R}, 0 \leqslant t \leqslant 2\pi\}$. V 中元素对于通常三角多项式的加法和数乘运算，是否构成实数域上的线性空间? 若 V 是 $\mathbf{R}$ 上的线性空间，证明 $\{\sin t, \sin 2t, \cdots, \sin nt\}$ 是 V 的一组基，试提出确定 c_k 的方法.

2. 若

$$x = c_1 \sin t + \cdots + c_k \sin kt + \cdots + c_n \sin nt,$$

则 $\forall k,\ 1 \le k \le n$,

$$x \sin kt = c_1 \sin t \sin kt + \cdots + c_k \sin kt \sin kt + \cdots + c_n \sin nt \sin kt.$$

两边取定积分可得

$$\begin{aligned}
&\int_0^{2\pi} x \sin kt dt \\
&= \int_0^{2\pi} (c_1 \sin t \sin kt + \cdots + c_k \sin kt \sin kt + \cdots + c_n \sin nt \sin kt)\, dt \\
&= \frac{c_1}{2}\int_0^{2\pi} (\cos(1-k)t - \cos(1+k)t)\, dt + \cdots + \\
&\quad \frac{c_k}{2}\int_0^{2\pi} (\cos 0 - \cos 2kt)\, dt + \cdots + \\
&\quad \frac{c_n}{2}\int_0^{2\pi} (\cos(n-k)t - \cos(n+k)t)\, dt = c_k\pi,
\end{aligned}$$

故

$$c_k = \frac{1}{\pi}\int_0^{2\pi} x \sin kt dt.$$

若

$$x = c_1 \sin t + \cdots + c_k \sin kt + \cdots + c_n \sin nt = 0,$$

则 $\forall k,\ 1 \le k \le n$，我们有

$$c_k = \frac{1}{\pi}\int_0^{2\pi} x \sin kt dt = 0.$$

这表明 $\sin t, \sin 2t, \ldots, \sin nt$ 线性无关。

3. 在 $\mathbf{R}^4$ 中有两组基

$$\boldsymbol{\alpha}_1 = (1,0,0,0)^{\mathrm{T}}, \qquad \boldsymbol{\alpha}_2 = (0,1,0,0)^{\mathrm{T}},$$
$$\boldsymbol{\alpha}_3 = (0,0,1,0)^{\mathrm{T}}, \qquad \boldsymbol{\alpha}_4 = (0,0,0,1)^{\mathrm{T}},$$

与

$$\boldsymbol{\beta}_1 = (2,1,-1,1)^{\mathrm{T}}, \qquad \boldsymbol{\beta}_2 = (0,3,1,0)^{\mathrm{T}},$$
$$\boldsymbol{\beta}_3 = (5,3,2,1)^{\mathrm{T}}, \qquad \boldsymbol{\beta}_4 = (6,6,1,3)^{\mathrm{T}}.$$

(1) 求由基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 到基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$ 的过渡矩阵;

(2) 求对两组基有相同坐标的非零向量.

4. 设四维线性空间 V 的两组基 $\Phi:\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 和 $\Psi:\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$ 满足

$$\begin{cases} \boldsymbol{\alpha}_1 + 2\boldsymbol{\alpha}_2 = \boldsymbol{\beta}_3, \\ \boldsymbol{\alpha}_2 + 2\boldsymbol{\alpha}_3 = \boldsymbol{\beta}_4, \\ \boldsymbol{\beta}_1 + 2\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_3, \\ \boldsymbol{\beta}_2 + 2\boldsymbol{\beta}_3 = \boldsymbol{\alpha}_4, \end{cases}$$

(1) 求由基 Φ 到基 Ψ 的过渡矩阵;

(2) 求向量 $\boldsymbol{\alpha} = 2\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2 + \boldsymbol{\beta}_3 + \boldsymbol{\beta}_4$ 在基 Φ 下的坐标;

(3) 判断是否存在非零元素 $\boldsymbol{\alpha} \in V$，使得 $\boldsymbol{\alpha}$ 在基 Φ 和基 Ψ 下的坐标相同.

5. V 表示在 $[0,1]$ 上有定义的所有实函数构成的线性空间, 下列哪些实函数集合构成 V 的子空间.

(1) 所有 $f(x)$, 使 $f(x)=0$;

(2) 所有 $f(x)$, 使 $f(x)=f(1-x)$;

(3) 所有 $f(x)$, 使 $2f(0)=f(1)$;

(4) 所有 $f(x)$, 使当 $0 \leqslant x \leqslant 1$ 时恒有 $f(x) \geqslant 0$.

6. 假定 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是 $\mathbf{R}^3$ 的一组基, 试求由

$$\boldsymbol{\alpha}_1' = \boldsymbol{\alpha}_1 - 2\boldsymbol{\alpha}_2 + 3\boldsymbol{\alpha}_3, \quad \boldsymbol{\alpha}_2' = 2\boldsymbol{\alpha}_1 + 3\boldsymbol{\alpha}_2 + 2\boldsymbol{\alpha}_3, \quad \boldsymbol{\alpha}_3' = 4\boldsymbol{\alpha}_1 + 13\boldsymbol{\alpha}_2$$

生成的子空间 $\mathrm{span}(\boldsymbol{\alpha}_1',\boldsymbol{\alpha}_2',\boldsymbol{\alpha}_3')$ 的基.

7. 设 V_1,V_2 都是线性空间 V 的有限维子空间, 且 $V_1 \subset V_2$. 证明: 如果 $\dim V_1 = \dim V_2$, 则 $V_1 = V_2$.

8. 证明: 所有 n 阶方阵空间是线性子空间 L_1 和 L_2 的直接和, 其中 L_1 是对称方阵子空间, L_2 是反对称方阵子空间.

9. 求由向量 $\boldsymbol{\xi}_i$ 生成的子空间与由向量 $\boldsymbol{\eta}_i$ 生成的子空间的交与和的基及维数, 其中

$$\boldsymbol{\xi}_1 = (1,2,-1,-2)^{\mathrm{T}}, \quad \boldsymbol{\xi}_2 = (3,1,1,1)^{\mathrm{T}}, \quad \boldsymbol{\xi}_3 = (-1,0,1,-1)^{\mathrm{T}};$$
$$\boldsymbol{\eta}_1 = (2,5,-6,-5)^{\mathrm{T}}, \quad \boldsymbol{\eta}_2 = (-1,2,-7,3)^{\mathrm{T}}.$$

10. 下列映射是不是线性的? 为什么? 哪些是单射? 哪些是满射?

(1) 在 $\mathbf{R}^3$ 中, $T(x,y,z) = (x+y, 2z, x)$;

(2) 在 n 阶方阵构成的线性空间 $M_{n\times n}$ 中, 设 $\boldsymbol{M}$ 为一固定 n 阶方阵, 定义 $T: T(\boldsymbol{A}) = \boldsymbol{M}\boldsymbol{A}$, $\boldsymbol{A}$ 为 $M_{n\times n}$ 中任一方阵;

7. 设 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 是 V_1 的一组基。
若 $dim(V_1) = dim(V_2)$, 则由于 $V_1 \subseteq V_2$,
$\alpha_1,\alpha_2,\ldots,\alpha_n$ 是 V_2 的一组基。于是,

$$V_1 = \{\alpha \mid k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n,\ k_1,k_2,\ldots,k_n \in K\} = V_2.$$

8. $\forall A = (a_{ij})_{n\times n} \in L$, 总有

$$A = \frac{A+A^T}{2} + \frac{A-A^T}{2}.$$

由于

$$\left(\frac{A+A^T}{2}\right)^T = \frac{A+A^T}{2}, \qquad \left(\frac{A-A^T}{2}\right)^T = -\frac{A-A^T}{2},$$

故

$$\frac{A+A^T}{2} \in L_1, \qquad \frac{A-A^T}{2} \in L_2.$$

这表明 $L = L_1 + L_2$.
若 $A \in L_1 \cap L_2$, 则有 $A = A^T = -A$, 即 $A = O$.
故 $L = L_1 \oplus L_2$.

9. 设 $\alpha \in L\{\xi_1,\xi_2,\xi_3\} \cap L\{\eta_1,\eta_2\}$, 则有

$$\alpha = k_1\xi_1 + k_2\xi_2 + k_3\xi_3 = \ell_1\eta_1 + \ell_2\eta_2,$$

亦即

$$\begin{pmatrix} 1 & 3 & -1 & -2 & 1 \\ 2 & 1 & 0 & -5 & -2 \\ -1 & 1 & 1 & 6 & 7 \\ -2 & 1 & -1 & 5 & -3 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ \ell_1 \\ \ell_2 \end{pmatrix} = 0.$$

$$\begin{pmatrix} 1 & 3 & -1 & -2 & 1 \\ 2 & 1 & 0 & -5 & -2 \\ -1 & 1 & 1 & 6 & 7 \\ -2 & 1 & -1 & 5 & -3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 3 & -1 & -2 & 1 \\ 0 & -5 & 2 & -1 & -4 \\ 0 & 4 & 0 & 4 & 8 \\ 0 & 7 & -3 & 1 & -1 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 3 & -1 & -2 & 1 \\ 0 & 1 & -2 & -3 & -4 \\ 0 & 4 & 0 & 4 & 8 \\ 0 & 7 & -3 & 1 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 5 & 7 & 13 \\ 0 & 1 & -2 & -3 & -4 \\ 0 & 0 & 8 & 16 & 24 \\ 0 & 0 & 11 & 22 & 27 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 0 & 5 & 7 & 13 \\ 0 & 1 & -2 & -3 & -4 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 11 & 22 & 27 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & -3 & -2 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & -6 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 0 & 0 & -3 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$\ell_1 = t,\ \ell_2 = 0.$

$k_1 = 3t,\ k_2 = -t,\ k_3 = -2t.$

$$L\{\xi_1,\xi_2,\xi_3\} \cap L\{\eta_1,\eta_2\} = \{k\eta_1 \mid k \in R\},$$
$$L\{\xi_1,\xi_2,\xi_3\} + L\{\eta_1,\eta_2\} = L\{\xi_1,\xi_2,\xi_3,\eta_2\}.$$

(3) 将复数域 $\mathbf{C}$ 视为实数域 $\mathbf{R}$ 上的线性空间, 定义 $T: T(z)=\bar{z}$,$\bar{z}$ 是复数 z 的共轭复数;

(4) 在 (3) 中, 若视 $\mathbf{C}$ 为复数域 $\mathbf{C}$ 上的线性空间, T 仍如 (3) 定义, 情况如何?

(5) 在 $\mathbf{R}^3$ 中, $T(x,y,z)=(x^2,x+y,z)$.

11. 设 T 是 V 的一个线性变换, 如果 $T^{k-1}\boldsymbol{\alpha}\neq\mathbf{0}$, 但 $T^k\boldsymbol{\alpha}=\mathbf{0}$.

(1) 证明: $\boldsymbol{\alpha},T\boldsymbol{\alpha},\cdots,T^{k-1}\boldsymbol{\alpha}(k\geqslant 1)$ 线性无关;

(2) 设 $W(\boldsymbol{\alpha})=\mathrm{span}\{\boldsymbol{\alpha},T\boldsymbol{\alpha},\cdots,T^{k-1}\boldsymbol{\alpha}\}$, 将 T 看成 $W(\boldsymbol{\alpha})$ 中线性变换, 试求 T 在基 $\boldsymbol{\alpha},T\boldsymbol{\alpha},\cdots,T^{k-1}\boldsymbol{\alpha}$ 下的矩阵.

11. (1) 设

$$\ell_0\alpha+\ell_1T\alpha+\cdots+\ell_{k-1}T^{k-1}\alpha=0. \qquad (1)$$

由于 $T^{k-1}\alpha\neq 0$ 且 $T^k\alpha=0$,
依次将 $T^{k-1},\ldots,T$ 作用于 (1) 式两边可得

$$\ell_0=0,\ \ell_1=0,\ldots,\ \ell_{k-2}=0.$$

再将 $\ell_0=0,\ \ell_1=0,\ldots,\ \ell_{k-2}=0$ 代入 (1) 式可得 $\ell_{k-1}=0$.
这表明 $\alpha,T\alpha,\ldots,T^{k-1}\alpha$ 线性无关。

(2) 令 $\beta_1=\alpha,\ \beta_2=T\alpha,\ldots,\ \beta_k=T^{k-1}\alpha$, 则有

$$T\beta_i=\beta_{i+1},\ 1\le i\le k-1,$$
$$T\beta_k=0.$$

于是

$$(T\beta_1\ \cdots\ T\beta_{k-1}\ T\beta_k)=(\beta_1\ \cdots\ \beta_{k-1}\ \beta_k)\begin{pmatrix}0&0&\cdots&0&0\\1&0&\cdots&0&0\\0&1&\cdots&0&0\\\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&1&0\end{pmatrix},$$

即 T 在基底 $\alpha,T\alpha,\ldots,T^{k-1}\alpha$ 下的矩阵是:

$$\begin{pmatrix}0&0&\cdots&0&0\\1&0&\cdots&0&0\\0&1&\cdots&0&0\\\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&1&0\end{pmatrix}.$$

12. 设 $\mathbf{R}$ 上的三维线性空间 V 的线性变换 T 在基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$ 下的矩阵为

$$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}.$$

(1) 求 T 在基 $\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_1$ 下的矩阵;

(2) 求 T 在基 $\boldsymbol{\varepsilon}_1,k\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$ 下的矩阵, 其中 $k\in\mathbf{R}$, 且 $k\neq 0$;

(3) 求 T 在基 $\boldsymbol{\varepsilon}_1+\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$ 下的矩阵.

13. 设三维线性空间 V 的两组基为 $\Phi:\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$; $\Psi:\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3$. 由基 Φ 变到基 Ψ 的过渡矩阵 $\boldsymbol{C}=\begin{pmatrix}1&0&1\\0&-1&0\\-1&0&1\end{pmatrix}$, 线性变换 T 满足

$$\begin{cases}T(\boldsymbol{\varepsilon}_1+2\boldsymbol{\varepsilon}_2+3\boldsymbol{\varepsilon}_3)=\boldsymbol{\eta}_1+\boldsymbol{\eta}_2,\\T(2\boldsymbol{\varepsilon}_1+\boldsymbol{\varepsilon}_2+2\boldsymbol{\varepsilon}_3)=\boldsymbol{\eta}_2+\boldsymbol{\eta}_3,\\T(\boldsymbol{\varepsilon}_1+3\boldsymbol{\varepsilon}_2+4\boldsymbol{\varepsilon}_3)=\boldsymbol{\eta}_1+\boldsymbol{\eta}_3.\end{cases}$$

(1) 求 T 在基 Ψ 下的矩阵 $\boldsymbol{A}$;

(2) 求 $T(\boldsymbol{\eta}_1)$ 在基 Φ 下的坐标.

14. 设 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 是线性空间 V 的一组基,

$$(T\alpha_1,T\alpha_2,\ldots,T\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)A.$$

则对任意的可逆阵 P,

$$(\beta_1,\beta_2,\ldots,\beta_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)P$$

也是 V 的一组基，且 T 在基 $\beta_1,\beta_2,\ldots,\beta_n$ 下的矩阵是

$$P^{-1}AP.$$

若 T 在任一组基下的矩阵都相同，即 $A=P^{-1}AP$, 也就是

$$PA=AP.$$

这说明 A 与所有的可逆矩阵可交换。设

$$A=(a_{ij})_{n\times n}.$$

第i行×2倍

令 $P=E(i(2))$, $1\le i\le n$, 则由 $E(i(2))A=AE(i(2))$ 可得

$$(2a_{i1},\ldots,2a_{ii},\ldots,2a_{in})=(a_{i1},\ldots,2a_{ii},\ldots,a_{in}),\ 1\le i\le n.$$

这表明

$$a_{ij}=0,\ i\neq j.$$

因此,

$$A=\begin{pmatrix}a_{11}&&&\\&a_{22}&&\\&&\ddots&\\&&&a_{nn}\end{pmatrix}.$$

令 $P=E(i,j)$, $1\le i,j\le n$, 则由 $E(i,j)A=AE(i,j)$ 可得

交换两行(列)

$$a_{ii}=a_{jj},\ i\neq j.$$

因此，A 是数量阵。

14. T 是数域 K 上的 n 维线性空间 V 的一个线性变换. 证明: T 在任意一组基下的矩阵都相同的充要条件是 T 是数乘变换.

15. 在线性空间 $\mathbf{R}^3$ 中, 定义线性变换 T 为

$$T(x_1,x_2,x_3)^{\mathrm{T}}=(-x_1-2x_2+2x_3,x_2,x_3)^{\mathrm{T}}.$$

求 T 的所有特征值和特征向量.

16. $T:\ R^4\mapsto R^3$, $\forall x\in R^4$, $Tx=Ax$,

$$A=\begin{pmatrix}1&-1&-1&1\\1&-1&1&-3\\1&-1&-2&3\end{pmatrix}.$$

取 R^4 的一组基 e_1,e_2,e_3,e_4, 则

$$Te_1=Ae_1=\begin{pmatrix}1\\1\\1\end{pmatrix},\qquad Te_2=Ae_2=\begin{pmatrix}-1\\-1\\-1\end{pmatrix},$$
$$Te_3=Ae_3=\begin{pmatrix}-1\\1\\-2\end{pmatrix},\qquad Te_4=Ae_4=\begin{pmatrix}1\\-3\\3\end{pmatrix}.$$

由于 $\forall Tx\in Im(T)$ 均可由 Te_1,Te_2,Te_3,Te_4 线性表示，故 Te_1,Te_2,Te_3,Te_4 的一个极大无关组就构成 $Im(T)$ 的一组基。计算易得 Te_1,Te_3 是 Te_1,Te_2,Te_3,Te_4 的一个极大无关组。

由

$$Ax=\begin{pmatrix}1&-1&-1&1\\1&-1&1&-3\\1&-1&-2&3\end{pmatrix}x=0$$

求得 $Ax=0$ 的一个基础解系:

$$\alpha_1=\begin{pmatrix}1\\1\\0\\0\end{pmatrix},\qquad \alpha_2=\begin{pmatrix}1\\0\\2\\1\end{pmatrix}.$$

因为 $Ker(T)$ 就是 $Ax=0$ 的通解，故

$$Ker(T)=L\{\alpha_1,\alpha_2\}.$$

也就是说 α_1,α_2 是 $Ker(T)$ 的一组基。

16. 给定矩阵 $\begin{pmatrix}1&-1&-1&1\\1&-1&1&-3\\1&-1&-2&3\end{pmatrix}$, 定义 $\mathbf{R}^4$ 到 $\mathbf{R}^3$ 的一个线性映射 T 为: $T\boldsymbol{x}=\boldsymbol{A}\boldsymbol{x},\forall\boldsymbol{x}\in\mathbf{R}^4$. 试分别求出像空间 $\mathrm{Im}(T)$ 与核空间 $\ker(T)$ 的一组基.

17. 已知 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4$ 是四维线性空间 V 的一组基, V 上的线性变换 T 在这组基下的矩

17. $T: V \mapsto V, \forall \alpha \in V, T\alpha = A\alpha$,

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ -1 & 2 & 1 & 3 \\ 1 & 2 & 5 & 5 \\ 2 & -2 & 1 & -2 \end{pmatrix}.$$

(1) 由于

$$(\varepsilon_1', \varepsilon_2', \varepsilon_3', \varepsilon_4') = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4) \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 3 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & -3 & 1 & 2 \end{pmatrix},$$

故 T 在基 $\varepsilon_1', \varepsilon_2', \varepsilon_3', \varepsilon_4'$ 下的矩阵是

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 3 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & -3 & 1 & 2 \end{pmatrix}^{-1} A \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 3 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & -3 & 1 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & -5 & 3 & 2 \\ \frac{2}{3} & -\frac{14}{3} & \frac{10}{3} & \frac{10}{3} \\ \frac{8}{3} & -\frac{56}{3} & \frac{40}{3} & \frac{40}{3} \\ \frac{2}{3} & \frac{13}{3} & -\frac{11}{3} & -\frac{14}{3} \end{pmatrix}.$$

(2) 由于

$$\begin{pmatrix} 1 & 0 & 2 & 1 \\ -1 & 2 & 1 & 3 \\ 1 & 2 & 5 & 5 \\ 2 & -2 & 1 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \\ 0 & -2 & -3 & -4 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & \frac{3}{2} & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

故 $T\varepsilon_1, T\varepsilon_2$ 是 $T\varepsilon_1, T\varepsilon_2, T\varepsilon_3, T\varepsilon_4$ 的极大无关组，
$(-2\ \ -3/2\ 1\ 0)^T$, $(-1\ \ -2\ 0\ 1)^T$ 是 $Ax = 0$ 的基础解系。
从而有：

$$Im(T) = L\{T\varepsilon_1, T\varepsilon_2\},$$

$$Ker(T) = L\left\{-2\varepsilon_1 - \frac{3}{2}\varepsilon_2 + \varepsilon_3, -\varepsilon_1 - 2\varepsilon_2 + \varepsilon_4\right\}.$$

阵为

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 2 & 1 \\ -1 & 2 & 1 & 3 \\ 1 & 2 & 5 & 5 \\ 2 & -2 & 1 & -2 \end{pmatrix}.$$

(1) 求 T 在基 $\varepsilon_1' = \varepsilon_1 - 2\varepsilon_2 + \varepsilon_4, \varepsilon_2' = 3\varepsilon_2 - \varepsilon_3 - 3\varepsilon_4, \varepsilon_3' = \varepsilon_3 + \varepsilon_4$, $\varepsilon_4' = 2\varepsilon_4$ 下的矩阵;

(2) 求 T 的像空间与核空间;

(3) 在 T 的核空间中选一组基, 把它扩充成 V 的一组基, 并求 T 在这组基下的矩阵;

(4) 在 T 的像空间中选一组基, 把它扩充成 V 的一组基, 并求 T 在这组基下的矩阵.

18*. 设 T_1, T_2 是数域 C 上的 n 维线性空间 V 上的线性变换, 且 $T_1T_2 = T_2T_1$, 证明: 如果 λ_0 是 T_1 的特征值, 那么 V_{λ_0} 是 T_2 的不变子空间.

19*. 证明: 线性空间 V 上的线性变换 T 的一维不变子空间必定是由 T 的某个特征向量所生成.

20*. 设 T 是数域 K 上线性空间 V 的一个线性变换, W_1, W_2 是 T 的不变子空间. 证明: $W_1 + W_2, W_1 \bigcap W_2, T(W_1) + T(W_2)$ 都是 T 的不变子空间.

第7章　内积空间

7.1　内积空间

回忆一下, 在空间解析几何中, 若 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 是 $\mathbf{R}^3$ 中的向量, 它们的数量积为

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}\cdot\boldsymbol{\beta}=|\boldsymbol{\alpha}||\boldsymbol{\beta}|\cos\theta=x_1y_1+x_2y_2+x_3y_3, \tag{7.1}$$

其中 θ 是 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 之间的夹角, (x_1,x_2,x_3) 是 $\boldsymbol{\alpha}$ 的坐标, (y_1,y_2,y_3) 是 $\boldsymbol{\beta}$ 的坐标. 易见数量积满足下列性质:

(1) $(\boldsymbol{\alpha},\boldsymbol{\beta})=(\boldsymbol{\beta},\boldsymbol{\alpha})$;

(2) $(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta})=(\boldsymbol{\alpha}_1,\boldsymbol{\beta})+(\boldsymbol{\alpha}_2,\boldsymbol{\beta})$;

(3) $(\lambda\boldsymbol{\alpha},\boldsymbol{\beta})=\lambda(\boldsymbol{\alpha},\boldsymbol{\beta}),\lambda$ 为任意实数;

(4) $(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0$, 当且仅当 $\boldsymbol{\alpha}=\mathbf{0}$ 时, $(\boldsymbol{\alpha},\boldsymbol{\alpha})=0$.

向量的长度和夹角都可以由数量积 (点积) 求出, 即

$$\|\boldsymbol{\alpha}\|=\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})},\qquad \cos\theta=\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{\|\boldsymbol{\alpha}\|\|\boldsymbol{\beta}\|}.$$

我们抽象出数量积所含有的代数性质, 在一般的线性空间中引入内积的概念.

定义 7.1.1 (实内积与欧氏空间)　设 V 是实数域 $\mathbf{R}$ 上的线性空间, 对 V 中的任意两个向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$, 由某种规则确定了一个实数, 记为 $(\boldsymbol{\alpha},\boldsymbol{\beta})$, 并满足下列条件:

(1) 对称性: $(\boldsymbol{\alpha},\boldsymbol{\beta})=(\boldsymbol{\beta},\boldsymbol{\alpha})$;

(2) 可加性: $(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta})=(\boldsymbol{\alpha}_1,\boldsymbol{\beta})+(\boldsymbol{\alpha}_2,\boldsymbol{\beta})$;

(3) 齐次性: $(\lambda\boldsymbol{\alpha},\boldsymbol{\beta})=\lambda(\boldsymbol{\alpha},\boldsymbol{\beta}),\forall\lambda\in\mathbf{R}$;

(4) 非负性: $(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0$, 当且仅当 $\boldsymbol{\alpha}=\mathbf{0}$ 时, $(\boldsymbol{\alpha},\boldsymbol{\alpha})=0$,

则实数 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 就称为向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 的**实内积**，有时简称**内积**.

定义了实内积的实数域 $\mathbf{R}$ 上的线性空间称为**实内积空间**, 并称有限维实内积空间为**欧几里得 (Euclid) 空间**，简称**欧氏空间**.

定义 7.1.1 中 (2), (3) 两条性质合起来称为线性性, 实内积满足的这几条性质叫**内积公理**. 由定义 7.1.1 容易得

(1) $(\boldsymbol{\alpha},\boldsymbol{\beta}+\boldsymbol{\gamma})=(\boldsymbol{\alpha},\boldsymbol{\beta})+(\boldsymbol{\alpha},\boldsymbol{\gamma})$;

(2) $(\boldsymbol{\alpha},\lambda\boldsymbol{\beta})=\lambda(\boldsymbol{\alpha},\boldsymbol{\beta}),\quad\forall\lambda\in\mathbf{R}$;

(3) $(\boldsymbol{\alpha},\mathbf{0})=0=(\mathbf{0},\boldsymbol{\beta})$;

(4)

$$\left(\sum_{i=1}^{m} a_i\boldsymbol{\alpha}_i, \sum_{j=1}^{n} b_j\boldsymbol{\beta}_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_ib_j(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_j)$$

$$= (a_1, a_2, \cdots, a_m)\begin{pmatrix} (\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1)(\boldsymbol{\alpha}_1,\boldsymbol{\beta}_2)\cdots(\boldsymbol{\alpha}_1,\boldsymbol{\beta}_n) \\ (\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2)\cdots(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_n) \\ \cdots\cdots \\ (\boldsymbol{\alpha}_m,\boldsymbol{\beta}_1)(\boldsymbol{\alpha}_m,\boldsymbol{\beta}_2)\cdots(\boldsymbol{\alpha}_m,\boldsymbol{\beta}_n) \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

其中 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j \in V, a_i, b_j \in \mathbf{R}, (i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$.

例 7.1.1　在 $\mathbf{R}^3$ 中定义实内积为 (7.1) 的数量积, 则 $\mathbf{R}^3$ 构成一个欧几里得空间.

例 7.1.2　闭区间 $[a, b](b > a)$ 上的实连续函数的全体按通常意义的加法和数乘运算构成无穷维线性空间 $C_{[a,b]}$, 对 $f(x), g(x) \in C_{[a,b]}$, 定义内积为

$$(f(x), g(x)) = \int_a^b f(x)g(x)\mathrm{d}x.$$

容易验证上述内积满足实内积定义的要求, 则 $C_{[a,b]}$ 就成为实内积空间.

7.1.1　长度、范数、夹角与正交性

类似于三维几何空间 $\mathbf{R}^3$, 实内积空间中抽象向量的长度、夹角等度量可用内积来表示.

定义 7.1.2 (长度、范数)　设 V 是实内积空间, $\boldsymbol{\alpha} \in V$, 则称 $\|\boldsymbol{\alpha}\| = \sqrt{(\boldsymbol{\alpha}, \boldsymbol{\alpha})}$ 为向量 $\boldsymbol{\alpha}$ 的**长度或范数**. 显然, 向量的长度有如下性质:

(1) $\|\boldsymbol{\alpha}\| = 0 \Leftrightarrow \boldsymbol{\alpha} = \mathbf{0}$;

(2) $\|\lambda\boldsymbol{\alpha}\| = \sqrt{\lambda^2(\boldsymbol{\alpha}, \boldsymbol{\alpha})} = |\lambda|\|\boldsymbol{\alpha}\|, \forall\lambda \in \mathbf{R}$;

(3) 记 $\boldsymbol{\alpha}^0 = \dfrac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|}$, 则 $\|\boldsymbol{\alpha}^0\| = 1$, 即 $\boldsymbol{\alpha}^0$ 是 $\boldsymbol{\alpha}$ 的单位向量.

定义 7.1.3 (夹角)　设 V 是实内积空间, 有非零向量 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$, 则它们的夹角 θ 定义为

$$\theta = \arccos\frac{(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\|\boldsymbol{\alpha}\|\|\boldsymbol{\beta}\|}, \qquad 0 \leqslant \theta \leqslant \pi.$$

注　若 $\boldsymbol{\beta} \neq 0$, 由 $(\boldsymbol{\alpha} + \lambda\boldsymbol{\beta}, \boldsymbol{\alpha} + \lambda\boldsymbol{\beta}) \geqslant 0$, 展开左端后, 取 $\lambda = -\dfrac{(\boldsymbol{\alpha}, \boldsymbol{\beta})}{(\boldsymbol{\beta}, \boldsymbol{\beta})}$, 容易得到 $\dfrac{(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\|\boldsymbol{\alpha}\|\|\boldsymbol{\beta}\|} \leqslant 1$. 这说明上述夹角的定义是有意义的.

定义 7.1.4 (正交)　设有向量 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$, 若 $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = 0$, 则称 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$**正交(垂直)**, 记为 $\boldsymbol{\alpha} \perp \boldsymbol{\beta}$.

显然, (1) 零向量与任何向量正交, 并且, 与自身正交的向量只能是零向量. (2) 若 $\boldsymbol{\alpha}, \boldsymbol{\beta}$ 是非零向量, 则 $\boldsymbol{\alpha} \perp \boldsymbol{\beta} \Longleftrightarrow \boldsymbol{\alpha}, \boldsymbol{\beta}$ 之间的夹角为 $\pi/2$.

定理 7.1.1　$\boldsymbol{\alpha} \perp \boldsymbol{\beta} \Longleftrightarrow \|\boldsymbol{\alpha} + \boldsymbol{\beta}\|^2 = \|\boldsymbol{\alpha}\|^2 + \|\boldsymbol{\beta}\|^2$.

证明

$$\begin{aligned}\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2&=(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\alpha}+\boldsymbol{\beta})=(\boldsymbol{\alpha},\boldsymbol{\alpha})+2(\boldsymbol{\alpha},\boldsymbol{\beta})+(\boldsymbol{\beta},\boldsymbol{\beta})\\&=\|\boldsymbol{\alpha}\|^2+2(\boldsymbol{\alpha},\boldsymbol{\beta})+\|\boldsymbol{\beta}\|^2,\end{aligned}$$

因此,

$$\boldsymbol{\alpha}\perp\boldsymbol{\beta}\Longleftrightarrow(\boldsymbol{\alpha},\boldsymbol{\beta})=0\Longleftrightarrow\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2=\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2.$$

□

上述定理在二维空间就是通常的商高定理：直角三角形的两直角边平方和等于斜边平方.

例 7.1.3 在二维欧氏空间 $\mathbf{R}^2$ 中, 定义两种不同的内积: ① $(x,y)_1=x^{\mathrm{T}}y$, ② $(x,y)_2=x^{\mathrm{T}}Cy$, 其中 $C=\begin{pmatrix}1&1\\1&2\end{pmatrix}$, $\forall x,y\in\mathbf{R}^2$.

可以验证上面给出的两种定义确实是内积, 因此在定义①下的 $\mathbf{R}^2$ 记为 $\mathbf{R}_1^2$, 在定义②下的 $\mathbf{R}^2$ 记为 $\mathbf{R}_2^2$. 试问向量 $\boldsymbol{\alpha}=(1,1)^{\mathrm{T}},\boldsymbol{\beta}=(-1,1)^{\mathrm{T}}$ 在这两个欧氏空间中是否正交?

解 由于

$$\begin{aligned}(\boldsymbol{\alpha},\boldsymbol{\beta})_1&=(1,1)(-1,1)^{\mathrm{T}}=0,\\(\boldsymbol{\alpha},\boldsymbol{\beta})_2&=(1,1)\begin{pmatrix}1&1\\1&2\end{pmatrix}(-1,1)^{\mathrm{T}}=1,\end{aligned}$$

故 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 在 $\mathbf{R}_1^2$ 中正交, 但在 $\mathbf{R}_2^2$ 中不正交. 这表明, 向量正交与否, 和该欧氏空间上定义的内积有关. □

正交性的概念可推广到多个向量的情形.

定义 7.1.5 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是实内积空间中的一组非零向量, 若它们两两正交, 则称 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是一个 **正交向量组**.

定理 7.1.2 *设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是一个正交向量组, 则它们线性无关.*

证明 设存在一组常数 $k_1,k_2,\cdots,k_n$, 使

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n=\mathbf{0}.$$

分别以 $\boldsymbol{\alpha}_i(i=1,2,\cdots,n)$ 与上式作内积

$$(\boldsymbol{\alpha}_i,k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n)=0,\qquad i=1,2,\cdots,n,$$

即

$$k_1(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_1)+k_2(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_2)+\cdots+k_n(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_n)=0.$$

由于当 $i\neq j$ 时, $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)=0$, 故得

$$k_i(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)=0,\qquad i=1,2,\cdots,n.$$

注意到 $\boldsymbol{\alpha}_i$ 非零, 即 $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)>0$, 则必有

$$k_i=0,\qquad i=1,2,\cdots,n.$$

所以 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关. □

类似于定理 7.1.1, 若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是一个正交向量组, 则

$$\|\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_n\|^2=\|\boldsymbol{\alpha}_1\|^2+\|\boldsymbol{\alpha}_2\|^2+\cdots+\|\boldsymbol{\alpha}_n\|^2.$$

7.1.2　酉空间

欧几里得空间是在实数域上定义内积的线性空间, 而酉空间则是对复数域上的线性空间进行讨论的. 由于酉空间的理论和欧氏空间非常相似, 有一套平行的理论, 故本节只给出酉空间中内积的定义及主要结论, 并不给出详细的说明, 读者可自证之.

定义 7.1.6 (复内积、酉空间)　设 V 是复数域 $\mathbf{C}$ 上的线性空间, 在 V 上定义一个二元复函数, 称为**复内积**, 记为 $(\boldsymbol{\alpha},\boldsymbol{\beta})$, 满足下列性质:

(1) $(\boldsymbol{\alpha},\boldsymbol{\beta})=\overline{(\boldsymbol{\beta},\boldsymbol{\alpha})}$, ($\overline{(\boldsymbol{\beta},\boldsymbol{\alpha})}$ 是 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 的共轭复数);

(2) $(\lambda\boldsymbol{\alpha},\boldsymbol{\beta})=\lambda(\boldsymbol{\alpha},\boldsymbol{\beta})$;

(3) $(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma})=(\boldsymbol{\alpha},\boldsymbol{\gamma})+(\boldsymbol{\beta},\boldsymbol{\gamma})$;

(4) $(\boldsymbol{\alpha},\boldsymbol{\alpha})$ 是非负实数, 当且仅当 $\boldsymbol{\alpha}=\mathbf{0}$ 时, $(\boldsymbol{\alpha},\boldsymbol{\alpha})=0$.

其中 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in V,\lambda$ 是复数. 我们称定义了复内积的线性空间为**复内积空间**, 有限维复内积空间称为**酉空间**.

对向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$, 若 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0$, 则称 $\boldsymbol{\alpha},\boldsymbol{\beta}$ **正交**(相互**垂直**).

欧氏空间与酉空间的区别不仅在于数域不同, 还在于内积定义中的条件 (1) 是不同的, 若不涉及这两方面, 它们相关的性质是类似的. 下面列出酉空间的主要性质:

(1) $(\boldsymbol{\alpha},\lambda\boldsymbol{\beta})=\overline{\lambda}(\boldsymbol{\alpha},\boldsymbol{\beta})$;

(2) $(\boldsymbol{\alpha},\boldsymbol{\beta}+\boldsymbol{\gamma})=(\boldsymbol{\alpha},\boldsymbol{\beta})+(\boldsymbol{\alpha},\boldsymbol{\gamma})$;

(3) $(\boldsymbol{\alpha},\mathbf{0})=(\mathbf{0},\boldsymbol{\alpha})=0$;

(4) $(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0$, 因此记 $\|\boldsymbol{\alpha}\|=\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})}$ 为向量 $\boldsymbol{\alpha}$ 的**长度**;

(5) $\left(\sum\limits_{i=1}^{m}\lambda_i\boldsymbol{\alpha}_i,\sum\limits_{j=1}^{n}\mu_j\boldsymbol{\beta}_j\right)=\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}\lambda_i\overline{\mu}_j(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j)$;

(6) 对任意的 $\boldsymbol{\alpha},\boldsymbol{\beta}$, **施瓦兹 (Schwartz) 不等式**成立, 即

$$|(\boldsymbol{\alpha},\boldsymbol{\beta})|\leqslant\|\boldsymbol{\alpha}\|\|\boldsymbol{\beta}\|.$$

上式当且仅当 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 线性相关时, 等号成立.

性质 (6) 的施瓦兹不等式证明如下:

证明　若 $\boldsymbol{\beta}=0$, 显然成立.

若 $\boldsymbol{\beta}\neq 0$, 由 $(\boldsymbol{\alpha}+\lambda\boldsymbol{\beta},\boldsymbol{\alpha}+\lambda\boldsymbol{\beta})\geqslant 0$, 展开左端, 取 $\lambda=-\dfrac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{(\boldsymbol{\beta},\boldsymbol{\beta})}$ 即可证得. □

例 7.1.4 在 $\mathbf{C}^n$ 上, 对于任意两个向量 $\boldsymbol{\alpha}=(x_1,\cdots,x_n)^{\mathrm{T}},\boldsymbol{\beta}=(y_1,\cdots,y_n)^{\mathrm{T}}$, 令

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}^{\mathrm{T}}\overline{\boldsymbol{\beta}}=x_1\overline{y_1}+\cdots+x_n\overline{y_n},$$

则 $\mathbf{C}^n$ 对于这个内积构成一个酉空间.

7.2 欧氏空间中的正交变换

7.2.1 欧氏空间的标准正交基

下面考虑在欧氏空间中向量的内积如何通过坐标进行计算, 从而引出度量矩阵的概念.

设 V 是欧氏空间, $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是它的一组基底, $\forall\boldsymbol{\alpha},\boldsymbol{\beta}\in V$:

$$\begin{aligned}\boldsymbol{\alpha}&=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n,\\ \boldsymbol{\beta}&=y_1\boldsymbol{\alpha}_1+y_2\boldsymbol{\alpha}_2+\cdots+y_n\boldsymbol{\alpha}_n,\end{aligned}$$

由内积的性质, 有

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\left(\sum_{i=1}^n x_i\boldsymbol{\alpha}_i,\sum_{j=1}^n y_j\boldsymbol{\alpha}_j\right)=\sum_{i=1}^n\sum_{j=1}^n(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)x_iy_j. \tag{7.2}$$

记

$$\begin{aligned}\boldsymbol{A}&=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\\&=\begin{pmatrix}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1)&(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)&\cdots&(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_n)\\(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1)&(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2)&\cdots&(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_n)\\\vdots&\vdots&&\vdots\\(\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_1)&(\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_2)&\cdots&(\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_n)\end{pmatrix}\\&=\begin{pmatrix}\boldsymbol{\alpha}_1\\\boldsymbol{\alpha}_2\\\vdots\\\boldsymbol{\alpha}_n\end{pmatrix}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n),\end{aligned}$$

$\boldsymbol{\alpha},\boldsymbol{\beta}$ 在基底 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 下的坐标为

$$X=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\qquad Y=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}.$$

那么 (7.2) 的内积就可表示成

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=X^{\mathrm{T}}\boldsymbol{A}Y. \tag{7.3}$$

我们称矩阵 $\boldsymbol{A}$ 为欧氏空间在基底 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 下的**度量矩阵**, 也称作**格拉姆 (Gram) 矩阵**.

(7.3) 说明, 向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 的内积可通过 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 在某基底下的坐标和度量矩阵来表示, 因此, 度量矩阵完全确定了内积, 而向量的长度、夹角这些度量又可以由内积进行表示, 这就是度量矩阵这个名字的含义.

定理 7.2.1　*度量矩阵是对称正定矩阵.*

证明　由于内积具有对称性, 所以度量矩阵是实对称的;

而对任意的非零向量 $\boldsymbol{\alpha}\in V$, 欧氏空间在基底 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 下的度量矩阵必满足

$$(\boldsymbol{\alpha},\boldsymbol{\alpha})=X^{\mathrm{T}}\boldsymbol{A}X>0.$$

因此, 度量矩阵是对称正定矩阵. □

定理 7.2.2　*欧氏空间中两组不同基底下的度量矩阵是合同的.*

证明　设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 是欧氏空间的两组不同的基底, 度量矩阵分别是 $\boldsymbol{A}$ 与 $\boldsymbol{B}$, 这两组不同基底之间的过渡矩阵是 $\boldsymbol{P}$, 即

$$(\boldsymbol{\beta}_1\boldsymbol{\beta}_2\cdots\boldsymbol{\beta}_n)=(\boldsymbol{\alpha}_1\boldsymbol{\alpha}_2\cdots\boldsymbol{\alpha}_n)\boldsymbol{P}. \tag{7.4}$$

由于

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix}(\boldsymbol{\alpha}_1\boldsymbol{\alpha}_2\cdots\boldsymbol{\alpha}_n),$$

$$\boldsymbol{B}=\begin{pmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\\ \vdots\\ \boldsymbol{\beta}_n\end{pmatrix}(\boldsymbol{\beta}_1\boldsymbol{\beta}_2\cdots\boldsymbol{\beta}_n).$$

将 (7.4) 代入上式得

$$\boldsymbol{B}=\boldsymbol{P}^{\mathrm{T}}\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix}(\boldsymbol{\alpha}_1\boldsymbol{\alpha}_2\cdots\boldsymbol{\alpha}_n)\boldsymbol{P}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}.$$

即不同基底下的度量矩阵是合同的. □

把空间解析几何中直角坐标系的概念推广到 n 维欧氏空间, 就得到标准正交基的概念.

定义 7.2.1 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是欧氏空间的一组基, 如果在这组基下的度量矩阵是单位矩阵, 则称 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是**标准正交基**.

欧氏空间的标准正交基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是一组长度为 1 的正交向量组.

我们知道, 正交向量组是线性无关的; 反之, 从一组线性无关的向量组出发, 必可构造出一组长度为 1 的正交向量组, 这就是线性无关向量组的正交规范化, 常用的正交规范化方法是下述定理 7.2.3 的证明中给出的施密特 (Schmidt) 正交化过程. 这里对一般向量的正交化过程与第 4 章定理 4.4.2 中给出正交化过程在形式上没有什么差别.

定理 7.2.3 对 n 维欧氏空间 V 的任意一组基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$, 都可以找到一组标准正交基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$, 使得

$$V=\operatorname{span}\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}=\operatorname{span}\{\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n\}.$$

证明 以下证明的过程也是**施密特 (Schmidt) 正交化过程**:

(1) 令 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1$, 显然 $\boldsymbol{\beta}_1\neq\boldsymbol{\theta}$.

(2) 令 $\boldsymbol{\beta}_2=k\boldsymbol{\beta}_1+\boldsymbol{\alpha}_2$, 由

$$0=(\boldsymbol{\beta}_2,\boldsymbol{\beta}_1)=(k\boldsymbol{\beta}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)=(k\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)+(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1),$$

得

$$k=-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}.$$

故

$$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1.$$

由 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1$ 的线性无关性立得 $\boldsymbol{\beta}_2\neq\boldsymbol{\theta}$.

(3) 令 $\boldsymbol{\beta}_3=k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+\boldsymbol{\alpha}_3$, 由 $\boldsymbol{\beta}_3\perp\boldsymbol{\beta}_2$, 且 $\boldsymbol{\beta}_3\perp\boldsymbol{\beta}_1$, 即

$$\begin{aligned}(\boldsymbol{\beta}_3,\boldsymbol{\beta}_2)&=(k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)\\&=(k_1\boldsymbol{\beta}_1,\boldsymbol{\beta}_2)+(k_2\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)+(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)=0,\\(\boldsymbol{\beta}_3,\boldsymbol{\beta}_1)&=(k_1\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)+(k_2\boldsymbol{\beta}_2,\boldsymbol{\beta}_1)+(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)\\&=k_1(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)+k_2(\boldsymbol{\beta}_2,\boldsymbol{\beta}_1)+(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)=0\end{aligned}$$

得

$$k_1=-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)},\qquad k_2=-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}.$$

故

$$\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2.$$

按上述方法, 已经构造出三个两两正交的向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$, 照此继续下去, 一般地, 假设已经作出了 r 个两两正交的非零向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$, 为求出第 $r+1$ 个非零正交向量, 令

$$\boldsymbol{\beta}_{r+1}=k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+\cdots+k_r\boldsymbol{\beta}_r+\boldsymbol{\alpha}_{r+1},$$

利用 r 个正交条件

$$(\boldsymbol{\beta}_{r+1}, \boldsymbol{\beta}_i) = 0, \qquad i = 1, 2, \cdots, r,$$

可解得

$$k_i = -\frac{(\boldsymbol{\alpha}_{r+1}, \boldsymbol{\beta}_i)}{(\boldsymbol{\beta}_i, \boldsymbol{\beta}_i)}, \qquad i = 1, 2, \cdots, r.$$

由于 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r$ 是 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 的线性组合和 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{r+1}$ 的线性无关性，容易得到 $\boldsymbol{\beta}_{r+1} \neq \boldsymbol{\theta}$.

由此可见, 当 $r = n - 1$ 时, 就求出了 n 个非零向量的正交向量组.

接着, 把所求得的正交向量组中的向量单位化, 令 $\varepsilon_i = \dfrac{\boldsymbol{\beta}_i}{\|\boldsymbol{\beta}_i\|}$, 即可得到一组标准正交基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$. □

由此有

定理 7.2.4　*任何 n 维欧氏空间都有正交基和标准正交基.*

标准正交基的存在性给我们讨论问题带来很大的方便. 就内积而言, 当 n 维欧氏空间 V 的基底取为一组标准正交基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 时, 由 (7.3) 可知, 向量 $\boldsymbol{\alpha}, \boldsymbol{\beta}$ 的内积有最简表示形式

$$(\boldsymbol{\alpha}, \boldsymbol{\beta}) = X^{\mathrm{T}} \boldsymbol{A} Y = X^{\mathrm{T}} Y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i,$$

且 $\boldsymbol{\alpha}, \boldsymbol{\beta}$ 的坐标可由内积表示出. 如 $\boldsymbol{\alpha}$ 的坐标 $x_1, x_2, \cdots, x_n$ 可表为

$$x_i = (\varepsilon_i, \boldsymbol{\alpha}), \qquad i = 1, 2, \cdots, n.$$

因此, 在讨论欧氏空间时, 我们一般总采用标准正交基. 例如大家熟知的三维几何空间中, 利用笛卡儿坐标系 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 作为标准正交基, 这时任意两个向量的内积的计算是最简单的.

例 7.2.1　试由四维欧氏空间 V 的一组基

$$\boldsymbol{\alpha}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\alpha}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\alpha}_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{\alpha}_4 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

出发构造一组标准正交基.

解　由施密特正交化方法, 令 $\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1$, 再令 $\boldsymbol{\beta}_2 = k\boldsymbol{\beta}_1 + \boldsymbol{\alpha}_2$, 由

$$k = -\frac{(\boldsymbol{\alpha}_2, \boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1)} = -\frac{1}{2}$$

得

$$\boldsymbol{\beta}_2 = -\frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}.$$

继续令 $\boldsymbol{\beta}_3 = k_1\boldsymbol{\beta}_1 + k_2\boldsymbol{\beta}_2 + \boldsymbol{\alpha}_3$, 由

$$k_1 = -\frac{(\boldsymbol{\alpha}_3, \boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1)} = \frac{1}{2}, \qquad k_2 = -\frac{(\boldsymbol{\alpha}_3, \boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2, \boldsymbol{\beta}_2)} = \frac{1}{3}$$

得

$$\boldsymbol{\beta}_3 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ 1 \end{pmatrix}.$$

又令 $\boldsymbol{\beta}_4 = k_1\boldsymbol{\beta}_1 + k_2\boldsymbol{\beta}_2 + k_3\boldsymbol{\beta}_3 + \boldsymbol{\alpha}_4$, 由

$$k_1 = -\frac{(\boldsymbol{\alpha}_4, \boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1)} = -1, \quad k_2 = -\frac{(\boldsymbol{\alpha}_4, \boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2, \boldsymbol{\beta}_2)} = \frac{2}{3}, \quad k_3 = -\frac{(\boldsymbol{\alpha}_4, \boldsymbol{\beta}_3)}{(\boldsymbol{\beta}_3, \boldsymbol{\beta}_3)} = -\frac{1}{2}$$

得

$$\boldsymbol{\beta}_4 = -\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \frac{2}{3}\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

最后将 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\beta}_4$ 单位化, 得

$$\boldsymbol{\varepsilon}_1 = \frac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\varepsilon}_2 = \frac{\boldsymbol{\beta}_2}{\|\boldsymbol{\beta}_2\|} = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ 0 \end{pmatrix},$$

$$\boldsymbol{\varepsilon}_3 = \frac{\boldsymbol{\beta}_3}{\|\boldsymbol{\beta}_3\|} = \begin{pmatrix} -\frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{12}} \\ \frac{3}{\sqrt{12}} \end{pmatrix}, \quad \boldsymbol{\varepsilon}_4 = \frac{\boldsymbol{\beta}_4}{\|\boldsymbol{\beta}_4\|} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

于是 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3, \boldsymbol{\varepsilon}_4$ 即为所求的一组标准正交基. □

例 7.2.2　设 V 是一个欧氏空间, $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 是 V 的一组基，已知基 $\boldsymbol{\alpha}_1=\varepsilon_1-\varepsilon_2$, $\boldsymbol{\alpha}_2=-\varepsilon_1+2\varepsilon_2$, $\boldsymbol{\alpha}_3=\varepsilon_2+2\varepsilon_3+\varepsilon_4$, $\boldsymbol{\alpha}_4=\varepsilon_1+\varepsilon_3+\varepsilon_4$ 的度量矩阵为

$$\boldsymbol{P}=\begin{pmatrix} 2 & -3 & 0 & 1 \\ -3 & 6 & 0 & -1 \\ 0 & 0 & 13 & 9 \\ 1 & -1 & 9 & 7 \end{pmatrix}.$$

(1) 求 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 的度量矩阵 $\boldsymbol{B}$;

(2) 求数 a, 使 $\boldsymbol{\alpha}=\varepsilon_1-\varepsilon_2+2\varepsilon_3$ 与 $\boldsymbol{\beta}=\varepsilon_1+a\varepsilon_2+2\varepsilon_3-\varepsilon_4$ 正交;

(3) 求一个向量 $\boldsymbol{\xi}_4$ 与 $\boldsymbol{\xi}_1=\varepsilon_1+\varepsilon_2-\varepsilon_3+\varepsilon_4$, $\boldsymbol{\xi}_2=\varepsilon_1-\varepsilon_2-\varepsilon_3+\varepsilon_4$, $\boldsymbol{\xi}_3=2\varepsilon_1+\varepsilon_2+\varepsilon_3+3\varepsilon_4$ 都正交;

(4) 求 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3,\boldsymbol{\xi}_4$ 的度量矩阵.

解　(1) 由 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 到 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的过渡矩阵为

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

所以由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 到 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 的过渡矩阵为 $\boldsymbol{C}^{-1}$, 于是基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 的度量矩阵为

$$\boldsymbol{B}=(\boldsymbol{C}^{-1})^{\mathrm{T}}\boldsymbol{P}\boldsymbol{C}^{-1}=\begin{pmatrix} 2 & 1 & 0 & -1 \\ 1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 1 \\ -1 & 0 & 1 & 3 \end{pmatrix}.$$

(2) 因为 $(\boldsymbol{\alpha},\boldsymbol{\beta})=(1,-1,2,0)\boldsymbol{B}(1,a,2,-1)^{\mathrm{T}}=-3a+10$, 所以 $a=\dfrac{10}{3}$.

(3) 设 $\boldsymbol{\gamma}=x_1\varepsilon_1+x_2\varepsilon_2+x_3\varepsilon_3+x_4\varepsilon_4$ 与 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$ 均正交，则有

$$\begin{cases} (1,1,-1,1)\boldsymbol{B}(x_1,x_2,x_3,x_4)^{\mathrm{T}}=0, \\ (1,-1,-1,1)\boldsymbol{B}(x_1,x_2,x_3,x_4)^{\mathrm{T}}=0, \\ (2,1,1,3)\boldsymbol{B}(x_1,x_2,x_3,x_4)^{\mathrm{T}}=0. \end{cases}$$

解之得一个非零解 $(11,-6,-1,0)^{\mathrm{T}}$. 令 $\boldsymbol{\xi}_4=11\varepsilon_1-6\varepsilon_2-\varepsilon_3$, 则此 $\boldsymbol{\xi}_4$ 与 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$ 均正交.

(4) 由 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 到 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3,\boldsymbol{\xi}_4$ 的过渡矩阵为

$$\boldsymbol{Q}=\begin{pmatrix} 1 & 1 & 2 & 11 \\ 1 & -1 & 1 & -6 \\ -1 & -1 & 1 & -1 \\ 1 & 1 & 3 & 0 \end{pmatrix}.$$

从而得 ξ_1,ξ_2,ξ_3,ξ_4 的度量矩阵为

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Q}=\begin{pmatrix}9&1&9&0\\1&1&3&0\\9&3&35&0\\0&0&0&172\end{pmatrix}.$$

□

7.2.2 欧氏空间中的正交变换

定义 7.2.2 设 V 是一个欧氏空间, T 是 V 上的线性变换, 如果对于任何向量 $\boldsymbol{x},\boldsymbol{y}\in V$, 变换 T 恒能使下式成立:

$$(T(\boldsymbol{x}),T(\boldsymbol{y}))=(\boldsymbol{x},\boldsymbol{y}), \tag{7.5}$$

则称 T 是 V 上的**正交变换**.

例如, 恒等变换是一个正交变换; 坐标平面上的旋转变换也是一个正交变换.

定理 7.2.5 *设 T 是欧氏空间 V 上的线性变换, 下面写出的任一条件都是使 T 成为正交变换的充要条件:*

(1) T 使向量长度保持不变, 即对任何 $\boldsymbol{x}\in V$, 有

$$(T(\boldsymbol{x}),T(\boldsymbol{x}))=(\boldsymbol{x},\boldsymbol{x}); \tag{7.6}$$

(2) 任一组标准正交基经 T 变换后的像仍是一组标准正交基;

(3) T 在任一组标准正交基下的矩阵 $\boldsymbol{A}$ 是正交矩阵.

证明 (1) 只要让 $\boldsymbol{y}=\boldsymbol{x}$ 便知必要性成立. 为证明充分性, 分别对向量 $\boldsymbol{x},\boldsymbol{y},\boldsymbol{x}+\boldsymbol{y}$ 使用 (7.6) 即有

$$\begin{aligned}&(T(\boldsymbol{x}),T(\boldsymbol{x}))=(\boldsymbol{x},\boldsymbol{x}),\\&(T(\boldsymbol{y}),T(\boldsymbol{y}))=(\boldsymbol{y},\boldsymbol{y}),\\&(T(\boldsymbol{x}+\boldsymbol{y}),T(\boldsymbol{x}+\boldsymbol{y}))=(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{x}+\boldsymbol{y}),\end{aligned}$$

而

$$\begin{aligned}(T(\boldsymbol{x}+\boldsymbol{y}),T(\boldsymbol{x}+\boldsymbol{y}))&=(T(\boldsymbol{x})+T(\boldsymbol{y}),T(\boldsymbol{x})+T(\boldsymbol{y}))\\&=(T(\boldsymbol{x}),T(\boldsymbol{x}))+2(T(\boldsymbol{x}),T(\boldsymbol{y}))+(T(\boldsymbol{y}),T(\boldsymbol{y}))\\&=(\boldsymbol{x},\boldsymbol{x})+2(T(\boldsymbol{x}),T(\boldsymbol{y}))+(\boldsymbol{y},\boldsymbol{y}),\\(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{x}+\boldsymbol{y})&=(\boldsymbol{x},\boldsymbol{x})+2(\boldsymbol{x},\boldsymbol{y})+(\boldsymbol{y},\boldsymbol{y}),\end{aligned}$$

故有

$$(T(\boldsymbol{x}),T(\boldsymbol{y}))=(\boldsymbol{x},\boldsymbol{y}).$$

(2) 由正交变换的定义可知, 若 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 为 V 的标准正交基, 则有 $(T(\varepsilon_i),T(\varepsilon_j))=$

$(\varepsilon_i,\varepsilon_j)=\delta_{ij}$, 即基的像仍是一组标准正交基. 反之, 若 $\boldsymbol{x}=\sum_{i=1}^{n}x_i\varepsilon_i, \boldsymbol{y}=\sum_{j=1}^{n}y_j\varepsilon_j$, 则可推出

$$\begin{aligned}(T(\boldsymbol{x}),T(\boldsymbol{y}))&=\left(T\left(\sum_{i=1}^{n}x_i\varepsilon_i\right),T\left(\sum_{j=1}^{n}y_j\varepsilon_j\right)\right)\\&=\sum_{i,j=1}^{n}x_iy_j\delta_{ij}=\sum_{i=1}^{n}x_iy_i=\left(\sum_{i=1}^{n}x_i\varepsilon_i,\sum_{j=1}^{n}y_j\varepsilon_j\right)\\&=(\boldsymbol{x},\boldsymbol{y}),\end{aligned}$$

故 T 是正交变换.

(3) 假设 T 在标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为 $\boldsymbol{A}$, 则有

$$T(\varepsilon_i)=a_{1i}\varepsilon_1+a_{2i}\varepsilon_2+\cdots+a_{ni}\varepsilon_n,\qquad i=1,2,\cdots,n,$$

即

$$T(\varepsilon_1\varepsilon_2\cdots\varepsilon_n)=(\varepsilon_1\varepsilon_2\cdots\varepsilon_n)\boldsymbol{A},$$

其中

$$\boldsymbol{A}=(a_{ij})_{n\times n}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}.$$

另一方面, 有

$$\begin{aligned}(T(\varepsilon_i),T(\varepsilon_j))&=\left(\sum_{k=1}^{n}a_{ki}\varepsilon_k,\sum_{l=1}^{n}a_{lj}\varepsilon_l\right)\\&=\sum_{k,l=1}^{n}a_{ki}a_{lj}(\varepsilon_k,\varepsilon_l)=\sum_{k,l=1}^{n}a_{ki}a_{lj}\delta_{kl}\\&=\sum_{k=1}^{n}a_{ki}a_{kj},\quad i,j=1,2,\cdots,n.\end{aligned}$$

由上面 (2) 知

$$\begin{aligned}T\text{为正交变换}&\Longleftrightarrow(T(\varepsilon_i),T(\varepsilon_j))=\delta_{ij}\\&\Longleftrightarrow\sum_{k=1}^{n}a_{ki}a_{kj}=\delta_{ij},\qquad i,j=1,2,\cdots,n\\&\Longleftrightarrow\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{E},\quad\text{即 }\boldsymbol{A}\text{ 是正交矩阵.}\end{aligned}$$

□

例 7.2.3　设欧氏空间 $\mathbf{R}^4$ 中的一个标准正交基为 $\varepsilon_1=(1,0,0,0)$, $\varepsilon_2=(0,1,0,0)$,

$\varepsilon_3=(0,0,1,0)$, $\varepsilon_4=(0,0,0,1)$, 并设 $\boldsymbol{\alpha}_1=\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},0,0\right)$, $\boldsymbol{\alpha}_2=\left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2},0,0\right)$, $\boldsymbol{\alpha}_3=\left(0,0,\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)$, $\boldsymbol{\beta}_1=\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right)$, $\boldsymbol{\beta}_2=\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right)$, $\boldsymbol{\beta}_3=\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right)$. 试求 $\mathbf{R}^4$ 上的正交变换 T, 满足 $T\boldsymbol{\alpha}_i=\boldsymbol{\beta}_i, i=1,2,3$. 并给出 T 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 下的矩阵.

解 显然，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 及 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 都是标准正交向量组，我们首先将它们分别扩充为 $\mathbf{R}^4$ 的标准正交基. 为此，令 $\widetilde{\boldsymbol{\alpha}}_4=(u_1,u_2,u_3,u_4)$, $\widetilde{\boldsymbol{\beta}}_4=(v_1,v_2,v_3,v_4)$, 由 $(\widetilde{\boldsymbol{\alpha}}_4,\boldsymbol{\alpha}_i)=0,(\widetilde{\boldsymbol{\beta}}_4,\boldsymbol{\beta}_i)=0, i=1,2,3$ 解得 $\widetilde{\boldsymbol{\alpha}}_4=k(0,0,1,-1)$, $\widetilde{\boldsymbol{\beta}}_4=k(1,-1,1,-1)$, k 为任意非零常数. 将 $\widetilde{\boldsymbol{\alpha}}_4,\widetilde{\boldsymbol{\beta}}_4$ 分别单位化得 $\boldsymbol{\alpha}_4=\left(0,0,\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)$, $\boldsymbol{\beta}_4=\left(\frac{1}{2},-\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right)$, 则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 及 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$ 均为欧氏空间 $\mathbf{R}^4$ 的标准正交基. 再考虑标准正交基之间的过渡矩阵，设 $(\boldsymbol{\alpha}_1\boldsymbol{\alpha}_2\boldsymbol{\alpha}_3\boldsymbol{\alpha}_4)=(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4)\boldsymbol{P}$, 及 $(\boldsymbol{\beta}_1\boldsymbol{\beta}_2\boldsymbol{\beta}_3\boldsymbol{\beta}_4)=(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4)\boldsymbol{Q}$, 则

$$\boldsymbol{P}=\begin{pmatrix}\frac{\sqrt{2}}{2}&\frac{\sqrt{2}}{2}&0&0\\ \frac{\sqrt{2}}{2}&-\frac{\sqrt{2}}{2}&0&0\\ 0&0&\frac{\sqrt{2}}{2}&\frac{\sqrt{2}}{2}\\ 0&0&\frac{\sqrt{2}}{2}&-\frac{\sqrt{2}}{2}\end{pmatrix},\quad \boldsymbol{Q}=\begin{pmatrix}\frac{1}{2}&\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\\ \frac{1}{2}&-\frac{1}{2}&\frac{1}{2}&-\frac{1}{2}\\ \frac{1}{2}&-\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}\\ \frac{1}{2}&\frac{1}{2}&-\frac{1}{2}&-\frac{1}{2}\end{pmatrix},$$

且 $\boldsymbol{P},\boldsymbol{Q}$ 均为正交矩阵.

再设线性变换 T 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 下的矩阵为 $\boldsymbol{A}$, 即，$T(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4)=(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4)\boldsymbol{A}$. 下面依据题目要求 $T\boldsymbol{\alpha}_i=\boldsymbol{\beta}_i$，并利用上面求得的过渡矩阵去确定矩阵 $\boldsymbol{A}$. 由

$$\begin{aligned}T(\boldsymbol{\alpha}_1\boldsymbol{\alpha}_2\boldsymbol{\alpha}_3\boldsymbol{\alpha}_4)&=T(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4)\boldsymbol{P}\\&=(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4)\boldsymbol{AP}\\&=(\boldsymbol{\beta}_1\boldsymbol{\beta}_2\boldsymbol{\beta}_3\boldsymbol{\beta}_4)\\&=(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4)\boldsymbol{Q},\end{aligned}$$

得 $\boldsymbol{AP}=\boldsymbol{Q}$, 从而得 $\boldsymbol{A}=\boldsymbol{QP}^{-1}$. 由于 $\boldsymbol{P},\boldsymbol{Q}$ 都是正交矩阵，所以 $\boldsymbol{P}^{-1},\boldsymbol{A}=\boldsymbol{QP}^{-1}$ 也都是正交矩阵. 于是所确定的线性变换 $T(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4)=(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4)\boldsymbol{A}$ 即是满足题意的正交变换. 而 T 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 下的矩阵为

$$\boldsymbol{A}=\boldsymbol{QP}^{-1}=\begin{pmatrix}\frac{\sqrt{2}}{2}&0&\frac{\sqrt{2}}{2}&0\\ 0&\frac{\sqrt{2}}{2}&0&\frac{\sqrt{2}}{2}\\ 0&\frac{\sqrt{2}}{2}&0&-\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}&0&-\frac{\sqrt{2}}{2}&0\end{pmatrix}.$$

□

7.2.3* 酉空间中的酉变换

在酉空间中，与欧氏空间中的正交变换相平行的概念是酉变换, 而与正交矩阵相平行的是酉矩阵.

定义 7.2.3　设 σ 为酉空间 V 上的一个线性变换，若对于 $\boldsymbol{\alpha},\boldsymbol{\beta}\in V$ 都有 $(\sigma(\boldsymbol{\alpha}),\sigma(\boldsymbol{\beta}))=(\boldsymbol{\alpha},\boldsymbol{\beta})$, 则称 σ 为 V 上的一个**酉变换**.

定义 7.2.4　若一个 n 阶复方阵 $\boldsymbol{A}$ 满足 $\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{A}\overline{\boldsymbol{A}}^{\mathrm{T}}=\boldsymbol{E},\boldsymbol{E}$ 是 n 阶单位阵, 则称 $\boldsymbol{A}$ 是**酉矩阵**.

与定理 7.2.5 类似的，有

定理 7.2.6　*设 σ 是 n 维酉空间 V 上的一个线性变换，则下列陈述彼此等价:*

1) *σ 是酉变换;*

2) *若 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 是 V 的一个标准正交基，则 $\sigma(\boldsymbol{\alpha}_1),\cdots,\sigma(\boldsymbol{\alpha}_n)$ 也是 V 的一个标准正交基;*

3) *σ 在 V 的任一标准正交基下的矩阵是酉矩阵.*

证明　略. □

由第 4 章的讨论可知，对实数域上的矩阵，实对称矩阵有很多很好的性质，但它的实质是与欧氏空间上的对称变换有着密切的关系. 那么在复数域上怎样的矩阵具有与实对称矩阵类似的性质? 下面介绍酉空间上的对称变换及所谓 Hermite 矩阵的情况.

定义 7.2.5　若酉空间 V 上的一个线性变换 σ 满足: 对一切 $\boldsymbol{\alpha},\boldsymbol{\beta}\in V$ 都有 $(\sigma(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{\alpha},\sigma(\boldsymbol{\beta}))$, 则称 σ 为 V 上的一个**对称变换**.

定义 7.2.6　设 $\boldsymbol{A}$ 是复数域 $\mathbf{C}$ 上的 $n\times n$ 矩阵，若 $\overline{\boldsymbol{A}}^{\mathrm{T}}=\boldsymbol{A}$, 则称 $\boldsymbol{A}$ 是一个 **Hermite 矩阵**.

显然，实对称矩阵是 Hermite 矩阵的特殊情形. 由复内积的性质经计算可以证明:

定理 7.2.7　*设 σ 是 n 维酉空间 V 的一个线性变换，则 σ 是对称变换的充要条件是 σ 在 V 的任意标准正交基下的矩阵是 Hermite 矩阵.*

对称变换和 Hermite 矩阵还有下述性质.

定理 7.2.8　*设 σ 是 n 维酉空间 V 的一个对称变换，那么*

(1) σ 的特征值都是实数;

(2) σ 的属于不同特征值的特征向量彼此正交;

(3) 存在 V 的一个标准正交基，使得 σ 在这个基下的矩阵是实对角矩阵.

证明　我们只证明 (1)，其余的证明留作习题.

设 $\lambda\in\mathbf{C}$ 是 σ 的一个特征值, $\boldsymbol{\alpha}$ 是属于 σ 的一个特征向量. 则

$$\lambda(\boldsymbol{\alpha},\boldsymbol{\alpha})=(\lambda\boldsymbol{\alpha},\boldsymbol{\alpha})=(\sigma(\boldsymbol{\alpha}),\boldsymbol{\alpha})=(\boldsymbol{\alpha},\sigma(\boldsymbol{\alpha}))=(\boldsymbol{\alpha},\lambda\boldsymbol{\alpha})=\overline{\lambda}(\boldsymbol{\alpha},\boldsymbol{\alpha}).$$

因为 $(\boldsymbol{\alpha},\boldsymbol{\alpha})\neq 0$, 所以必须 $\lambda=\overline{\lambda}$, 即 λ 是实数. □

定理 7.2.9　*设 $\boldsymbol{A}$ 是一个 n 阶 Hermite 矩阵, 则存在一个 n 阶酉矩阵 $\boldsymbol{U}$, 使得 $\overline{\boldsymbol{U}}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$ 是一个实对角矩阵，即任意 Hermite 矩阵都“酉相似”于一个实对角矩阵.*

证明　证明与定理 4.5.4 的证明相似. □

类似 n 维欧氏空间的度量矩阵的定义，在酉空间 V 中取一组基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$, 构造矩阵

$$\boldsymbol{A}=\begin{pmatrix}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_n)\\ (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_n)\\ \vdots & \vdots & & \vdots\\ (\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_n)\end{pmatrix},$$

这个矩阵是由酉空间 V 中的内积以及基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 唯一决定的, 称其为在 V 的基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 下的**度量矩阵**.

注意到 $\boldsymbol{A}$ 的 (i,j) 元素是 $a_{ij}=(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)$, 而 $\overline{\boldsymbol{A}}^{\mathrm{T}}$ 的 (i,j) 元素是 $\overline{(\boldsymbol{\alpha}_j,\boldsymbol{\alpha}_i)}=(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)$, 即 $\overline{\boldsymbol{A}}^{\mathrm{T}}=\boldsymbol{A}$. 这说明，酉空间 V 的内积在 V 的任意一组基下的度量矩阵 $\boldsymbol{A}$ 是一个 Hermite 矩阵.

设 $\boldsymbol{A}=(a_{ij})_{n\times n}$, $\boldsymbol{\alpha}=\sum\limits_{i=1}^n x_i\boldsymbol{\alpha}_i$, $\boldsymbol{\beta}=\sum\limits_{i=1}^n y_i\boldsymbol{\alpha}_i$, 则

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\left(\sum_{i=1}^n x_i\boldsymbol{\alpha}_i,\sum_{j=1}^n y_j\boldsymbol{\alpha}_j\right)=\sum_{i=1}^n\sum_{j=1}^n x_i\overline{y_j}(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)=\sum_{i=1}^n\sum_{j=1}^n a_{ij}x_i\overline{y_j}.$$

特别地，当 $\boldsymbol{\beta}=\boldsymbol{\alpha}$ 时, 有

$$(\boldsymbol{\alpha},\boldsymbol{\alpha})=\sum_{i=1}^n\sum_{j=1}^n a_{ij}x_i\overline{x_j}.$$

由于 n 阶 Hermite 矩阵 $\boldsymbol{A}$ 酉相似于一个实对角矩阵 $\boldsymbol{D}=\mathrm{diag}(d_1,\cdots,d_n)$, 即存在一个酉矩阵 $\boldsymbol{U}$, 使得 $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{D}$. 令 $\boldsymbol{X}=\boldsymbol{U}\boldsymbol{Y}$, 其中 $\boldsymbol{Y}^{\mathrm{T}}=(y_1,\cdots,y_n)$, 则

$$\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=\overline{\boldsymbol{Y}}^{\mathrm{T}}\overline{\boldsymbol{U}}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}\boldsymbol{Y}=\overline{\boldsymbol{Y}}^{\mathrm{T}}\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}\boldsymbol{Y}=\overline{\boldsymbol{Y}}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{Y}=d_1y_1\overline{y_1}+\cdots+d_ny_n\overline{y_n}.$$

以上讨论证明了下面的定理:

定理 7.2.10 对于 Hermite 型 $f(x_1,\cdots,x_n)=\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$, 存在酉线性替换 $\boldsymbol{X}=\boldsymbol{U}\boldsymbol{Y}$($\boldsymbol{U}$ 是酉矩阵), 使得

$$f(x_1,\cdots,x_n)=d_1y_1\overline{y_1}+d_2y_2\overline{y_2}+\cdots+d_ny_n\overline{y_n},$$

其中 $d_1,d_2,\cdots,d_n$ 是 $\boldsymbol{A}$ 的全部特征值，它们都是实数.

□

定义 7.2.7 若对于一切 $\boldsymbol{\alpha}\in\mathbf{C}^n,\boldsymbol{\alpha}\neq\mathbf{0}$ 都有

$$\overline{\boldsymbol{\alpha}}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}>0,$$

则称 $\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 是一个**正定 Hermite 型**. 一个正定 Hermite 型 $\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 的矩阵 $\boldsymbol{A}$ 称为**正定 Hermite 矩阵**.

正定 Hermite 矩阵与第 5 章讨论的实正定矩阵有类似的结果.

定理 7.2.11　设 $\boldsymbol{A}$ 是一个 n 阶 Hermite 矩阵，则下列陈述彼此等价：

(1) $\boldsymbol{A}$ 是正定 Hermite 矩阵；

(2) 对于任意 n 阶复可逆矩阵 $\boldsymbol{P}$, $\overline{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{AP}$ 是正定 Hermite 矩阵；

(3) $\boldsymbol{A}$ 的特征值全大于零；

(4) 存在 n 阶可逆复矩阵 $\boldsymbol{P}$，使 $\overline{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{AP} = \boldsymbol{E}_n$；

(5) $\boldsymbol{A}$ 可以分解成 $\overline{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{Q}$, 其中 $\boldsymbol{Q}$ 是 n 阶可逆复矩阵；

(6) $\boldsymbol{A}$ 的所有顺序主子式全大于零.

证明　(1) $\Longrightarrow$ (2). 任取 $\boldsymbol{\alpha} \in \mathbf{C}^n, \boldsymbol{\alpha} \neq \mathbf{0}$, 有 $\boldsymbol{P\alpha} \neq \mathbf{0}$. 因为 $\boldsymbol{A}$ 是正定 Hermite 矩阵，所以

$$\overline{\boldsymbol{\alpha}}^{\mathrm{T}}(\overline{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{AP})\boldsymbol{\alpha} = (\overline{\boldsymbol{P\alpha}})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{P\alpha}) > 0.$$

因此 $\overline{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{AP}$ 是正定 Hermite 矩阵.

(2) $\Longrightarrow$ (3). 由假设, $\boldsymbol{A}$ 是 Hermite 矩阵. 于是存在酉矩阵 $\boldsymbol{U}$，使

$$\boldsymbol{U}^{-1}\boldsymbol{AU} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) = \boldsymbol{D},$$

其中 λ_i 是实数且都是 $\boldsymbol{A}$ 的特征值, $i = 1, 2, \cdots, n$. 由假设，$\boldsymbol{U}^{-1}\boldsymbol{AU}$ 是正定 Hermite 矩阵，由此推出 $\overline{\boldsymbol{e}_i}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{e}_i = \lambda_i > 0$, $i = 1, 2, \cdots, n$.

(3) $\Longrightarrow$ (4). 因为 $\boldsymbol{A}$ 是 Hermite 矩阵，所以存在酉矩阵 $\boldsymbol{U}$，使得

$$\boldsymbol{U}^{-1}\boldsymbol{AU} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n),$$

其中 $\lambda_i > 0$, $i = 1, 2, \cdots, n$. 令

$$\boldsymbol{Q} = \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n}),$$

则有 $\boldsymbol{U}^{-1}\boldsymbol{AU} = \boldsymbol{QQ}$，从而 $\boldsymbol{Q}^{-1}\boldsymbol{U}^{-1}\boldsymbol{AUQ}^{-1} = \boldsymbol{E}_n$. 令 $\boldsymbol{P} = \boldsymbol{UQ}^{-1}$，则 $\overline{\boldsymbol{P}}^{\mathrm{T}} = \overline{\boldsymbol{UQ}^{-1}}^{\mathrm{T}} = \boldsymbol{Q}^{-1}\overline{\boldsymbol{U}}^{\mathrm{T}} = \boldsymbol{Q}^{-1}\boldsymbol{U}^{-1}$. 于是 $\overline{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{AP} = \boldsymbol{E}_n$.

(4) $\Longrightarrow$ (5). 由假设 $\overline{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{AP} = \boldsymbol{E}_n$, 于是 $\boldsymbol{A} = (\overline{\boldsymbol{P}}^{\mathrm{T}})^{-1}\boldsymbol{P}^{-1}$. 令 $\boldsymbol{Q} = \boldsymbol{P}^{-1}$, 则

$$\overline{\boldsymbol{Q}}^{\mathrm{T}} = \overline{\boldsymbol{P}^{-1}}^{\mathrm{T}} = (\overline{\boldsymbol{P}}^{-1})^{\mathrm{T}} = (\overline{\boldsymbol{P}}^{\mathrm{T}})^{-1}.$$

所以 $\boldsymbol{A} = \overline{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{Q}$.

(5) $\Longrightarrow$ (1). 设 $\boldsymbol{A} = \overline{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{Q}$, 其中 $\boldsymbol{Q}$ 可逆. 任取 $\boldsymbol{\alpha} \in \mathbf{C}^n$ 且 $\boldsymbol{\alpha} \neq \mathbf{0}$, 有

$$\overline{\boldsymbol{\alpha}}^{\mathrm{T}}\boldsymbol{A\alpha} = \overline{\boldsymbol{\alpha}}^{\mathrm{T}}\overline{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{Q\alpha}.$$

设 $(\boldsymbol{Q\alpha})^{\mathrm{T}} = (c_1, \cdots, c_n)$, 则

$$\overline{\boldsymbol{\alpha}}^{\mathrm{T}}\boldsymbol{A\alpha} = c_1\overline{c_1} + \cdots + c_n\overline{c_n} = |c_1|^2 + \cdots + |c_n|^2 > 0.$$

所以 $\boldsymbol{A}$ 是正定 Hermite 矩阵.

(1) $\Longleftrightarrow$ (6). 类似于定理 5.2.1 的证明.

7.3 欧几里得空间的同构

定义 7.3.1(线性空间的同构) 数域 K 上的两个线性空间 V 和 W 称为**同构**, 如果由 V 到 W 有一个一一映射 f 满足以下两个条件:

(1) $f(\boldsymbol{\alpha}+\boldsymbol{\beta})=f(\boldsymbol{\alpha})+f(\boldsymbol{\beta})$;

(2) $f(k\boldsymbol{\alpha})=kf(\boldsymbol{\alpha})$.

这里 $\boldsymbol{\alpha},\boldsymbol{\beta}\in V, k\in K$, 这样的映射 f 称为**同构映射**.

定义 7.3.2(欧几里得空间的同构) 欧几里得空间 V 与 W 称为**同构**的, 如果由 V 到 W 有一个一一映射 f 满足以下三个条件:

(1) $f(\boldsymbol{\alpha}+\boldsymbol{\beta})=f(\boldsymbol{\alpha})+f(\boldsymbol{\beta})$;

(2) $f(k\boldsymbol{\alpha})=kf(\boldsymbol{\alpha})$;

(3) $(f(\boldsymbol{\alpha}),f(\boldsymbol{\beta}))=(\boldsymbol{\alpha},\boldsymbol{\beta})$,

这里 $\boldsymbol{\alpha},\boldsymbol{\beta}\in V, k\in K$, 这样的映射 f 称为 V 到 W 的**同构映射**.

由定义 7.3.1、7.3.2 可知, 以上同构关系都具有自反、对称、传递三性质 (读者作为练习自证).

定理 7.3.1 设 V 与 W 都是数域 K 上的有限维线性空间, 则它们同构的充要条件是它们的维数相同.

证明 必要性: 设 V 的维数为 n, $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 V 的一组基, 于是 $\boldsymbol{\alpha}_1\leftrightarrow\boldsymbol{\alpha}_1',\boldsymbol{\alpha}_2\leftrightarrow\boldsymbol{\alpha}_2',\cdots,\boldsymbol{\alpha}_n\leftrightarrow\boldsymbol{\alpha}_n'$, 这里 $\boldsymbol{\alpha}_1',\boldsymbol{\alpha}_2',\cdots,\boldsymbol{\alpha}_n'\in W$, 则 $\boldsymbol{\alpha}_1',\boldsymbol{\alpha}_2',\cdots,\boldsymbol{\alpha}_n'$ 必然线性无关. 若 $\boldsymbol{\beta}_1',\boldsymbol{\beta}_2',\cdots,\boldsymbol{\beta}_{n+1}'$ 为 W 中任意 $n+1$ 个向量, 而 $\boldsymbol{\beta}_j\leftrightarrow\boldsymbol{\beta}_j', j=1,2,\cdots,n+1$, 则因 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_{n+1}$ 线性相关, 所以, $\boldsymbol{\beta}_1',\boldsymbol{\beta}_2',\cdots,\boldsymbol{\beta}_{n+1}'$ 也必然线性相关, 所以 W 的维数是 n.

充分性: 如果 V 与 W 都是 K 上的 n 维线性空间, $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 与 $\boldsymbol{\alpha}_1',\cdots,\boldsymbol{\alpha}_n'$ 分别为它们的基底, 让

$$\sum_{i=1}^{n}k_i\boldsymbol{\alpha}_i\leftrightarrow\sum_{i=1}^{n}k_i\boldsymbol{\alpha}_i',$$

就建立了 V 与 W 的同构性. □

若以 K^n 表示分量为 K 中的数的全体 n 元向量的集合, 则有如下推论.

推论 7.3.2 数域 K 上的任何 n 维线性空间都同构于 K^n.

从定义可以看出, 欧几里得空间比一般的线性空间多了一个内积的内容, 因此, 两个欧几里得空间同构的条件是: 它们作为线性空间来说同构, 而且还要有对应的任何两对向量的内积相等, 但是我们有以下定理.

定理 7.3.3 两个有限维欧几里得空间同构的充要条件是它们有相同的维数.

证明 必要性: 因为 V 与 W 同构, 首先作为线性空间来说, 应该也同构, 所以其维数相等.

充分性: 假定它们的维数都是 n, 在 V 中取一个标准正交基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$, 在 W 中取

一个标准正交基 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$, 则当

$$\alpha = k_1\varepsilon_1 + k_2\varepsilon_2 + \cdots + k_n\varepsilon_k,$$
$$\alpha' = k_1\varepsilon_1' + k_2\varepsilon_2' + \cdots + k_n\varepsilon_n'$$

时, 令 $\alpha \leftrightarrow \alpha'$, 这种对应显然是一一对应的, 而且使 $\alpha \leftrightarrow \alpha', \beta \leftrightarrow \beta'$ 时, 有

$$k\alpha \leftrightarrow k\alpha', \qquad \alpha + \beta \leftrightarrow \alpha' + \beta'.$$

下面证明它们有相同的内积. 令

$$\beta = \sum_{i=1}^{n} b_i\varepsilon_i, \qquad \beta' = \sum_{i=1}^{n} b_i\varepsilon_i',$$

于是

$$(\alpha, \beta) = \sum_{i=1}^{n} k_i b_i, \qquad (\alpha', \beta') = \sum_{i=1}^{n} k_i b_i,$$

故有 $(\alpha, \beta) = (\alpha', \beta')$, 所以欧几里得空间 V 与 W 同构. □

我们知道, $\mathbf{R}$ 上的一切 n 元向量所构成的 n 维线性空间, 对

$$\alpha = (a_1, a_2, \cdots, a_n), \beta = (b_1, b_2, \cdots, b_n),$$

令

$$(\alpha, \beta) = a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$

则它成为一个欧几里得空间, 记为 $\mathbf{R}^n$.

由定理 7.3.3 可以看到, 任何 n 维欧几里得空间都与 $\mathbf{R}^n$ 同构. 如果我们把所有同构的空间看成一个空间, 那么 n 维欧几里得空间就是实数域上的全体 n 元向量所构成的空间. 当 $\alpha = (a_1, a_2, \cdots, a_n), \beta = (b_1, b_2, \cdots, b_n)$ 时, 内积定义为

$$(\alpha, \beta) = \sum_{i=1}^{n} a_i b_i.$$

习　题　七

1. 设 V 为实数域上的一个 n 维向量空间, $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 为其一组基, 两个向量

$$u = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n,$$
$$v = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n.$$

定义实函数 (u, v) 为

$$(u, v) = x_1y_1 + 2x_2y_2 + \cdots + nx_ny_n,$$

证明: V 在此规定下是一个欧氏空间.

2. 在 $\mathbf{R}^n$ 中, 设 $\alpha = (a_1, a_2, \cdots, a_n), \beta = (b_1, b_2, \cdots, b_n)$, 分别定义实数 (α, β) 如下:

(1) $(\alpha, \beta) = \left(\sum_{i=1}^{n} a_i^2 b_i^2\right)^{1/2}$;

(2) $(\boldsymbol{\alpha},\boldsymbol{\beta})=\left(\sum\limits_{i=1}^{n}a_i\right)\left(\sum\limits_{j=1}^{n}b_i\right)$;

(3) $(\boldsymbol{\alpha},\boldsymbol{\beta})=\left(\sum\limits_{i=1}^{n}ia_ib_i\right)$;

判断它们是否为 $\mathbf{R}^n$ 中向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的内积.

3. 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 是 m 个互相正交的非零向量, 试证 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关.

4. 证明: $P_2[x]=\{p(x):p(x)=a_0+a_1x+a_2x^2,a_i\in\mathbf{R},i=0,1,2\}$ 在合适的加法、数乘和内积下构成一个欧氏空间.

5. 设 $\boldsymbol{A}=(a_{ij})$ 是一个 n 阶正定矩阵, 而 $\boldsymbol{\alpha}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}},\boldsymbol{\beta}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$ 在 $\mathbf{R}^n$ 中定义实函数

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta}.$$

(1) 证明: 在这个意义下, $\mathbf{R}^n$ 构成一欧氏空间;

(2) 求向量 $\boldsymbol{e}_1=(1,0,\cdots,0),\boldsymbol{e}_2=(0,1,\cdots,0),\cdots,\boldsymbol{e}_n=(0,0,\cdots,1)$ 的度量矩阵.

6. 在复数域 $\mathbf{C}$ 上的 n 维向量空间 $\mathbf{C}^n$ 中, 任取

$$\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n),\qquad \boldsymbol{\beta}=(b_1,b_2,\cdots,b_n),$$

令内积

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=a_1\bar{b}_1+a_2\bar{b}_2+\cdots+a_n\bar{b}_n,$$

证明 $\mathbf{C}^n$ 在此内积下是一个酉空间.

7. 设欧氏空间 $P_2[x]$ 中的内积定义为

$$(f,g)=\int_{-1}^{1}f(x)g(x)\mathrm{d}x.$$

(1) 求基 $1,x,x^2$ 的度量矩阵;

(2) 用坐标与度量矩阵乘积形式计算 $f(x)=1-x+x^2$ 与 $g(x)=1-4x-5x^2$ 的内积.

8. 求齐次线性方程组

$$\begin{cases}2x_1+x_2-x_3+x_4-3x_5=0,\\ x_1+x_2-x_3+x_5=0\end{cases}$$

的解空间 (作为 $\mathbf{R}^5$ 的子空间) 的一组标准正交基.

9. 设 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 是欧氏空间 V 的标准正交基, $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 是 V 的任意 k 个向量, 证明: $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 两两正交的充要条件是

$$\sum_{s=1}^{n}(\boldsymbol{\alpha}_p,\boldsymbol{e}_s)(\boldsymbol{\alpha}_q,\boldsymbol{e}_s)=0,\quad (p,q=1,2,\cdots,k;p\neq q).$$

10. 在酉空间 $\mathbf{C}^3$ 中给定如下一组基

$$\boldsymbol{\alpha}_1=(i,-1,i)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=(1,0,i)^{\mathrm{T}},\quad \boldsymbol{\alpha}_3=(1,1,1)^{\mathrm{T}},$$

利用施密特正交方法把它正交化后再单位化, 求出 $\mathbf{C}^3$ 的一组标准正交基.

11. 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 为 n 维欧氏空间 V 的一组基, 证明任意两个向量

$$\boldsymbol{u}=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\quad \boldsymbol{v}=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$$

的内积可由等式

$$(\boldsymbol{u},\boldsymbol{v})=x_1y_1+\cdots+x_ny_n$$

表达的充要条件是: $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 为标准正交基.

12. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 是三维欧氏空间中一组标准正交基, 证明:

$$\boldsymbol{\alpha}_1=\frac{1}{3}(2\varepsilon_1+2\varepsilon_2-\varepsilon_3),\quad \boldsymbol{\alpha}_2=\frac{1}{3}(2\varepsilon_1-\varepsilon_2+2\varepsilon_3),\quad \boldsymbol{\alpha}_3=\frac{1}{3}(\varepsilon_1-2\varepsilon_2-2\varepsilon_3)$$

也是一组标准正交基.

13. 设 $\boldsymbol{\alpha}_1=(1,0,2,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(0,2,0,3)^{\mathrm{T}},\boldsymbol{\alpha}_3=(2,6,4,9)^{\mathrm{T}}$, 试把 $L\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$ 的基底扩充成欧氏空间 $\mathbf{R}^4$ 的一组基, 并将它标准正交化.

14. 设 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$ 是三维欧几里得空间 V 的一组基, 其度量矩阵

$$\boldsymbol{A}=\begin{pmatrix}1&-1&1\\-1&2&0\\1&0&4\end{pmatrix},$$

试求 V 的一组标准正交基, 并验证它的度量矩阵是单位矩阵.

15. 在几何空间 $\mathbf{R}^3$ 中取标准正交基 (三个坐标轴上的单位向量), 验证以下由六个三阶方阵表示的线性变换都是正交变换, 并给出其几何解释.

$$\boldsymbol{A}_1=\begin{pmatrix}1&&\\&1&\\&&1\end{pmatrix},\quad \boldsymbol{A}_2=\begin{pmatrix}-1&&\\&1&\\&&1\end{pmatrix},$$

$$\boldsymbol{A}_3=\begin{pmatrix}1&&\\&-1&\\&&-1\end{pmatrix},\quad \boldsymbol{A}_4=\begin{pmatrix}-1&&\\&-1&\\&&-1\end{pmatrix},$$

$$\boldsymbol{A}_5=\begin{pmatrix}1&0&0\\0&\cos\theta&-\sin\theta\\0&\sin\theta&\cos\theta\end{pmatrix},\quad \boldsymbol{A}_6=\begin{pmatrix}-1&0&0\\0&\cos\theta&-\sin\theta\\0&\sin\theta&\cos\theta\end{pmatrix}.$$

16. 设 $\boldsymbol{\alpha}_i=(a_{i1},a_{i2},\cdots,a_{in}),i=1,2,\cdots,n$ 是 n 维欧氏空间 $\mathbf{R}^n$ 的向量组, 证明:

$$\sum_{j=1}^{n}a_{ij}x_j=0\quad 与\quad \sum_{j=1}^{n}(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)x_j=0,\quad (i=1,2,\cdots,n)$$

的解空间同构.

参 考 文 献

(美) 利昂. 2004. 线性代数 (英文版). 6 版. 北京：机械工业出版社.

北京大学数学系几何与代数教研室前代数小组. 2003. 高等代数. 3 版. 北京：高等教育出版社.

陈仲，栗熙. 1998. 大学数学 (下册). 南京：南京大学出版社.

李小刚. 2006. 线性代数及其应用. 北京：科学出版社.

李正元，李永乐，袁荫棠. 2009. 数学复习全书 (理工类. 数学一). 11 版. 北京：国家行政学院出版社.

王巧玲等. 2000. 大学数学教程 (第二册). 南京：南京大学出版社.

张志涌，刘瑞桢，杨祖荫. 1997. 掌握和精通 MATLAB. 北京：北京航空航天大学出版社.

周伯壎. 1966. 高等代数. 北京：人民教育出版社.

Gilbert Strang. 2003. Introduction to Linear Algebra. 3rd ed. Cambridge: Wellesley-Cambridge Press.