第一章　线性代数概要与提高

引　言　线性代数是什么?

除非特别说明, 本书的讨论均假定是在复数域$\mathbb{C}$ 的某子域$\mathbb{F}$ 上进行的. 但读者容易看出, 大部分内容实际上并不需要“除法”运算甚至不需要限定在复数域内讨论, 因此这些内容可以不加任何修饰地移植到没有“除法”的 $\mathbb{C}$ 的适当子集上, 比如整数集合 $\mathbb{Z}$ 或者其它的系统, 比如有限域$\mathbb{F}_q$, 其中 $q=p^m$, p 是素数 (在编码, 密码, 通信等领域有限域甚至比复数域 $\mathbb{C}$ 更加重要, 实际上, 许多计算机软件系统并不处理任意实数或复数, 而是代之以某个较大的有限域 $\mathbb{F}_q$).

本章概括了后续章节要用到的线性代数的基本知识, 也引入了一些简单而自然的概念, 如满秩分解, 矩阵的谱 以及线性空间的构造等, 除了利用分块初等矩阵证明了 **Sylvester**[1] **不等式**以及矩阵对角化的主定理外, 其余结论均不加证明或仅略加提示, 请参看 ([12]).

本科阶段的线性代数课程讨论两个相关问题, 一个是引入矩阵来解线性方程组, 另一个是利用线性方程组来研究矩阵.

解线性方程组是线性代数课程的初等部分. 首先, 矩阵的引入给了线性方程组两种简洁的表达, 即 $Ax=b$(称为矩阵形式) (矩阵也给了二**次型**

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

一个简洁的表达, 即 $f(x)=x^TAx$). 因此线性方程组解的存在性取决于系数矩阵A 与增广矩阵(A,b) 的秩相等与否. 其次, 利用矩阵的运算, 还可以将线性方程组 $Ax=b$ 再次改写为向量形式

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

可以看出, 线性方程组的解实际上是向量 b 关于 A 的列 $\alpha_1,\cdots,\alpha_n$ 的组合系数. 最后, 通过研究齐次线性方程组 $Ax=0$ 的解的结构, 可以知道任意解均可表为任一基础解系 的线性组合, 而 $Ax=b$ 如果有解, 则可将其化为相应的齐次线性方程组.

矩阵运算的难点在于其“乘法”. 理解 n 阶方阵 A 的高次幂 A^m 是理解矩阵乘法的关键 (也是理解和化简二次型的关键). 利用特征值与特征向量, 一些矩阵可以化为对角形, 即存在可逆矩阵P 使得 $P^{-1}AP=D=\mathrm{diag}\,(\lambda_1,\lambda_2\cdots,\lambda_n)$, 因此

$$A^m=PD^mP^{-1}=P\mathrm{diag}\,(\lambda_1^m,\lambda_2^m\cdots,\lambda_n^m)P^{-1}.$$

特别地, 实对称矩阵可以正交对角化, 即存在正交矩阵 Q 使得

$$Q^{-1}AQ=Q^TAQ=\mathrm{diag}\,(\lambda_1,\lambda_2\cdots,\lambda_n).$$

于是利用坐标变换 $x=Qy$ 即可将实二次型 $f(x)=x^TAx$ 化为标准型

$$f=\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2.$$

本章即是上述两条主线的总结与提高.

[1] James Joseph Sylvester(1814-1897), 英国数学家, 著名护士 F.Nightingale (南丁格尔) 的老师, 是 19 世纪后半叶美国最著名的数学家, 美国历史最悠久的著名数学期刊 American Journal of Mathematics 的创始人, 去世前为英国牛津大学教授.

第一节　矩阵乘法与分块矩阵

设 m, n 是正整数. 数域 $\mathbb{F}$ 上 $m\times n$ 矩阵 (也称为 m 行 n 列矩阵, $m\times n$ 阶矩阵, $m\times n$ 型矩阵等) 全体记为 $\mathbb{F}^{m\times n}$, 其中零矩阵记为 0. 全体 n 阶方阵构成的集合记为 $M_n(\mathbb{F})$ 或 $\mathbb{F}^{n\times n}$, 其中单位矩阵记为 I_n 或 I(线性代数课程中多用 E). 设矩阵 $A=(a_{ij})\in\mathbb{F}^{m\times n}$, 将 A 的每个元素改为其共轭元素所得的矩阵称为 A 的共轭矩阵, 记为 $\bar{A}$, 即 $\bar{A}=(\bar{a}_{ij})_{m\times n}$. 显然, A 的**共轭转置**矩阵等于它的转置共轭矩阵, $(\bar{A})^T=\overline{(A^T)}$, 记为 $A^*=\bar{A}^T$(也常用 A^H, 其中的 H 是 Hermite [2]的首字母).

第 i 行第 j 列元素为 1, 其余元素均为 0 的 $m\times n$ 矩阵称为**基本矩阵**, 记为 E_{ij} , $1\le i\le m$, $1\le j\le n$. 借助于这些矩阵, 任意 $m\times n$ 矩阵 $A=(a_{ij})$ 均能唯一地表示成

$$\begin{aligned} A &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \\ &= a_{11}E_{11}+\cdots+a_{1n}E_{1n}+\cdots+a_{m1}E_{m1}+\cdots+a_{mn}E_{mn} \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}E_{ij}. \end{aligned}$$

因为

$$E_{ij}E_{kl}=\delta_{jk}E_{il},\quad 1\le i\le m,\, 1\le j,k\le p,\, 1\le l\le n,$$

其中

$$\delta_{jk}=\begin{cases} 1, & \text{如果 } j=k; \\ 0, & \text{其他.} \end{cases}$$

是 Kronecker[3] 符号. 于是矩阵的乘法可表为:

$$\begin{aligned} AB &= (\sum_{i=1}^{m}\sum_{j=1}^{p} a_{ij}E_{ij})(\sum_{i=1}^{p}\sum_{j=1}^{n} b_{ij}E_{ij}) \\ &= \sum_{i,j,k,l}(a_{ij}E_{ij})(b_{kl}E_{kl}) = \sum_{i,j,k,l}(a_{ij}b_{kl})(E_{ij}E_{kl}) \\ &= \sum_{i,j,k,l}(a_{ij}b_{kl})(\delta_{jk}E_{il}) = \sum_{i,j,k,l}(\delta_{jk}a_{ij}b_{kl})E_{il} \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n}(\sum_{k=1}^{p} a_{ik}b_{kj})E_{ij}, \end{aligned}$$

即乘积 AB 的第 i 行第 j 列的元素等于 $\displaystyle\sum_{k=1}^{p} a_{ik}b_{kj}$, 这正是矩阵乘法的“左行右列”规则, 是按每个“元素”来做“乘法”运算, 从中可以看出该“乘法”实际上不是通常的 (数字) 乘法, 而是数字乘法与加法的混合体.

[2]Charles Hermite(1822-1901), 法国著名数学家与数学教育家, 其最著名的学生是法国历史上最伟大的数学家 Henri Poincare (也是著名的物理学家和哲学家). 他第一个证明了自然对数的底 e 是超越数, 德国数学家 Carl Louis Ferdinand von Lindemann 利用同样的方法证明了圆周率 π 是超越数.

[3]Leopold Kronecker(1823 – 1891) 是著名德国数学家, 逻辑学家, 其名言: God made the integers; all else is the work of man. 此处的符号是最著名的数学符号之一, 被称为 Kronecker delta.

矩阵乘法还可以按照“行向量”与“列向量”来进行. 设 A 是 $m\times n$ 矩阵, 分别以 A_j, A^i 表示 A 的第 j 列和第 i 行, 则有

$$
\begin{aligned}
A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\
&= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&= x_1\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \\
&= x_1A_1 + x_2A_2 + \cdots + x_nA_n.
\end{aligned}
$$

因此, 矩阵乘一个列向量等于该矩阵所有列的线性组合, 组合系数即是该列向量的对应元素. 同理, 一个行向量左乘一个矩阵等于该矩阵所有行的线性组合, 组合系数即是该行向量的对应元素, 即有

$$
(y_1\ y_2\ \cdots\ y_m)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = y_1A^1 + y_2A^2 + \cdots + y_mA^m.
$$

于是两个矩阵的乘积 $C = AB$ 的行向量与列向量的结构为 (这实际上是优化的计算机算法):

$$
C_j = AB_j, \quad C^i = A^iB.
$$

即矩阵 AB 的第 j 列是 A 的列向量的线性组合, 组合系数恰为矩阵 B 的第 j 列的相应元素; AB 的第 i 行是 B 的行向量的线性组合, 组合系数恰为矩阵 A 的第 i 行的相应元素. 特别, 用 e_j 表示第 j 个**基本列向量**(第 j 个元素为 1 其余元素均为 0), 则 $Ae_j = A_j$, $e_i^TA = A^i$ 以及 $E_{ij} = e_ie_j^T$(此处默认 e_i 与 e_j^T 有合适的行数与列数).

例 1.1.1 (矩阵乘法的行列结构)

(1) $$\begin{pmatrix} 0 & 0 & x \\ 0 & y & 0 \\ z & 0 & 0 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} xa_{31} & xa_{32} & xa_{33} \\ ya_{21} & ya_{22} & ya_{23} \\ za_{11} & za_{12} & za_{13} \end{pmatrix}$$ (行结构);

(2) $$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ x & y & 0 \\ 0 & 1 & z \end{pmatrix} = \begin{pmatrix} xa_{12} & ya_{12} + a_{13} & za_{13} \\ xa_{22} & ya_{22} + a_{23} & za_{23} \\ xa_{32} & ya_{32} + a_{33} & za_{33} \end{pmatrix}$$ (列结构).

例 1.1.2 ($AB=0$ 的意义) 若 $AB=0$, 则 $AB_j=0$, 故 B 的每个列向量都是齐次线性方程组 $Ax=0$ 的解向量 (特别, 若 A 是方阵, 则 B 的每个非零列向量都是 A 的属于特征值 0 的特征向量). 同理, 由于 $A^iB=0$, 故 A 的每个行向量都是齐次线性方程组 $y^TB=0$ 的解向量或 $B^Ty=0$ 的解向量的转置.

例 1.1.3 (线性方程组的分类) 如果一个线性方程组有解, 则称它是相容方程组; 否则就称其为矛盾方程组. 方程组 $Ax=b$ 有解 $\Longleftrightarrow$ b 是系数矩阵 A 的列的线性组合 $\Longleftrightarrow$ $r(A)=r(A,b)$, 即系数矩阵的秩等于增广矩阵的秩. 特别地, 齐次线性方程组 $Ax=0$ 有非零解 $\Longleftrightarrow$ A 的列向量线性相关; 有唯一解 (即零解) $\Longleftrightarrow$ A 的列向量线性无关.

设 A 是 n 阶方阵, $f(x)=a_0+a_1x+\cdots+a_mx^m$ 是 $\mathbb{F}$ 上的一个多项式, 称 n 阶方阵

$$a_0I+a_1A+\cdots+a_mA^m$$

为 A 的多项式, 记为 $f(A)$. 易知, 同一方阵的两个多项式是可以交换的, 即若

$$f(x)=a_0+a_1x+\cdots+a_mx^m,\ g(x)=b_0+b_1x+\cdots+b_lx^l,$$

则 $f(A)g(A)=g(A)f(A)$.

n 阶方阵 $A=(a_{ij})$ 的行列式记为 $|A|$ (另一个通用记号是 $\det A$), 它具有性质 $|AB|=|A|\,|B|$. 方阵 A 的迹 $\operatorname{tr}A$ 是 A 的对角线元素之和 $\displaystyle\sum_{i=1}^n a_{ii}$. 矩阵的迹具有下列基本性质:

命题 1.1.1 设 $A=(a_{ij})$, B 均为 n 阶方阵, λ 是数, 则
(1) $\operatorname{tr}(A+B)=\operatorname{tr}A+\operatorname{tr}B$;
(2) $\operatorname{tr}(\lambda A)=\lambda(\operatorname{tr}A)$;
(3) $\operatorname{tr}(AB)=\operatorname{tr}(BA)$ (此仅需 A, B 分别为 $m\times n$ 和 $n\times m$ 矩阵即可);
(4) $\operatorname{tr}A^T=\operatorname{tr}A$;
(5) $\operatorname{tr}(AA^*)=\sum\limits_{i,j=1}^n|a_{ij}|^2$. 特别地, $\operatorname{tr}(AA^*)=0 \iff A=0$.

其中性质 (5) 是因为 AA^* 的第 j 个对角线元素为 $\displaystyle\sum_{k=1}^n|a_{jk}|^2$, 即 A 的第 j 行作为 n 维向量的长度的平方.

与矩阵密切相关的另一个数字是**矩阵的秩**. 矩阵 A 的所有不为零的子式的最高阶数称为矩阵 A 的秩, 记为 $r(A)$. 约定零矩阵的秩是 0.

对任意 n 阶方阵 $A=(a_{ij})$, 去掉第 i 行第 j 列后所剩余的 $n-1$ 阶方阵的行列式称为元素 a_{ij} 的**余子式**, 记为 M_{ij}. 而 $(-1)^{i+j}M_{ij}$ 称为元素 a_{ij} 的**代数余子式**, 记为 A_{ij}. n 阶方阵

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

称为方阵 A 的**伴随矩阵**, 记为 $\operatorname{adj} A$(如果仅在实数域范围内讨论, 则常用符号 A^*). 伴随矩阵的重要性由下式体现:

$$A(\operatorname{adj} A) = (\operatorname{adj} A)A = |A|I. \tag{1.1.1}$$

对 n 阶方阵而言, “秩为 n”(也称为“**满秩**”), “**非奇异**”与“**可逆**”是等价的三个概念. 可逆矩阵的逆矩阵是唯一的, 记为 A^{-1}. 逆矩阵具有下述性质:

命题 1.1.2 (1) $A^{-1} = \frac{1}{|A|}\operatorname{adj} A$;

(2) $(A^{-1})^{-1} = A$;

(3) $(A^T)^{-1} = (A^{-1})^T$;

(4) $(AB)^{-1} = B^{-1}A^{-1}$;

(5) 若数 $\lambda \neq 0$, 则 $(\lambda A)^{-1} = \lambda^{-1}A^{-1}$;

(6) $|A^{-1}| = |A|^{-1}$;

(7) 设 A 是 $m \times n$ 矩阵, P 是 m 阶可逆方阵, Q 是 n 阶可逆方阵, 则

$$r(A) = r(PA) = r(AQ) = r(PAQ),$$

即可逆矩阵与任何矩阵乘积的秩等于该矩阵的秩.

矩阵秩的另一个极端是1. 一个非零矩阵 A 的秩为1 $\Longleftrightarrow$ A 是一个非零列矩阵与一个非零行矩阵的乘积, 即存在列向量 α, β 使得 $A = \alpha\beta^T$. 因此秩为1的方阵的高次幂可以如下算出:

$$A^m = (\alpha\beta^T)^m = (\beta^T\alpha)^{m-1}\alpha\beta^T = (\beta^T\alpha)^{m-1}A.$$

(注意上式中 $\beta^T\alpha$ 是数.)

对矩阵的和与乘积的秩的估计由下述不等式给出.

定理 1.1.1 (1) $r(A+B) \leq r(A) + r(B)$;

(2)(Sylvester 不等式) 设 A, B 分别为 $m \times p$, $p \times n$ 矩阵, 则

$$r(A) + r(B) - p \leq r(AB) \leq \min\{r(A), r(B)\}. \tag{1.1.2}$$

定理1.1.1(1) 中的不等式和 (2) 中右边的不等式都容易用矩阵的和与乘积的行列结构来证明, 但 (2) 中左边的不等式则有更好的办法, 即利用分块矩阵.

分块初等矩阵是由分块单位矩阵经过一次行 (或列) 初等变换得到的分块矩阵. 它们的作用与普通初等矩阵一样, 即设 P 是一个初等分块矩阵, A 是一个适当分块的分块矩阵, 则 PA 等于对 A 的行作与 P 匹配的初等变换. 类似地, 用一个初等分块矩阵 Q 右乘 A 则是对 A 的列作与 Q 匹配的初等变换.

现考虑下面的矩阵等式

$$\begin{pmatrix} I_m & A \\ 0 & I_p \end{pmatrix} \begin{pmatrix} AB & 0 \\ 0 & I_p \end{pmatrix} \begin{pmatrix} I_n & 0 \\ -B & I_p \end{pmatrix} = \begin{pmatrix} 0 & A \\ -B & I_p \end{pmatrix}.$$

右端的矩阵的秩显然不小于

$$r\begin{pmatrix} 0 & A \\ -B & 0 \end{pmatrix} = r(A) + r(B),$$

而左端矩阵的秩恰好是 $r(AB)+p$, Sylvester 不等式得证.

分块对角矩阵是分块矩阵的最简形式, 它们有一种非常简洁的记法, 即

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix} = A_1 \oplus A_2 \oplus \cdots \oplus A_s = \sum_{i=1}^{s} \oplus A_i,$$

这种表示称为矩阵的**直和**, 每个子矩阵 A_i 称为一个直和项.

例 1.1.4 (分块对角矩阵乘分块矩阵) 设 $A = A_1 \oplus A_2\,, B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$. 则

$$AB = \begin{pmatrix} A_1B_{11} & A_1B_{12} \\ A_2B_{21} & A_2B_{22} \end{pmatrix}, \quad BA = \begin{pmatrix} B_{11}A_1 & B_{12}A_2 \\ B_{21}A_1 & B_{22}A_2 \end{pmatrix}.$$

例 1.1.4 显然可以推广到任意有限多直和项的情形, 它实际上是对角矩阵与一般矩阵乘积的推广 (请注意其中的“左行右列”规则):

$$\begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \begin{pmatrix} \alpha_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \lambda_1 a_{11} & \lambda_1 a_{12} \\ \lambda_2 a_{21} & \lambda_2 a_{22} \end{pmatrix}, \quad \begin{pmatrix} \alpha_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 a_{11} & \lambda_2 a_{12} \\ \lambda_1 a_{21} & \lambda_2 a_{22} \end{pmatrix}.$$

分块矩阵是研究矩阵的强有力工具, 值得多加学习.

思考题
1. 秩为 0 的 n 阶矩阵只有 1 个. 秩为 1 的矩阵与秩为 2 的矩阵是否可以比较多少?
2. 当 $n \geq 2$ 时, n 阶可逆矩阵与不可逆矩阵都是无限的. 是否存在某种方式可以比较它们的多少?
3. 试给出矩阵秩的一种直观意义.

第二节 线性方程组与 n 维线性空间 $\mathbb{F}^n$

如果线性方程组 $Ax=b$ 有解 x_0, 则 $b=Ax_0$, 从而原方程组可化为 $A(x-x_0)=0$, 因此解线性方程组的根本在于解齐次线性方程组.

定理 1.2.1 ($Ax=0$ 的解的结构) 设 α, β 是 $Ax=0$ 的两个解向量, $\lambda \in \mathbb{F}$, 则
(1) $\alpha+\beta$ 也是 $Ax=0$ 的解;
(2) $\lambda\alpha$ 也是 $Ax=0$ 的解.

由此定理知齐次线性方程组的任意有限个解 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的线性组合

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s$$

也是解, 因此最好可以找到能够表示所有解的一组向量, 这就是极大线性无关组. 为此, 需要**线性无关**的概念.

定义 1.2.1 (线性无关与线性相关) 设 $S = \{\alpha_1, \alpha_2, \cdots, \alpha_s\}$ 是 $\mathbb{F}^n$ 的一组向量, 如果线性方程组 $x_1\alpha_1 + x_2\alpha_2 + \cdots + x_s\alpha_s = 0$ 仅有零解, 则称向量组 S 是线性无关的. 否则就称 S 是**线性相关**的.

定义 1.2.2 (极大线性无关组) 设 $S=\{\alpha_1,\alpha_2,\cdots,\alpha_s,\cdots\}$ 是 $\mathbb{F}^n$ 的一组向量 (有限或无限), 设其中部分向量 $M=\{\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}\}$ 满足下列条件:

(1) M 线性无关;

(2) S 中任何向量均能由 M 线性表示;

则称 M 是 S 的一个**极大线性无关组**. 特别, 齐次线性方程组 $Ax=0$ 的解集的一个极大线性无关组称为该方程组的一个**基础解系**.

向量组的极大线性无关组可能不唯一 (何时唯一?), 但每个极大线性无关组包含的向量个数均相同, 称为该向量组的秩. 矩阵的列向量组与行向量组的秩分别称为该矩阵的**列秩**与**行秩**, 它们与该矩阵的秩相等.

定理 1.2.2 (齐次线性方程组的基本定理) 齐次线性方程组 $A_{m\times n}x=0$ 的任何一个基础解系恰含 $n-r(A)$ 个解向量, 它的全体解为:

$$x=c_1\alpha_1+c_2\alpha_2+\cdots+c_{n-r}\alpha_{n-r}$$

其中 $c_1,c_2,\cdots,c_{n-r}$ 是任意常数, $\alpha_1,\alpha_2,\cdots,\alpha_{n-r}$ 是该方程组的一个基础解系.

定理 1.2.3 (线性方程组的基本定理) 设线性方程组 $A_{m\times n}x=b$ 有解, 则它的全体解为:

$$x=x_0+c_1\alpha_1+c_2\alpha_2+\cdots+c_{n-r}\alpha_{n-r}$$

其中 x_0 是其任意一个解 (称为**特解**), $c_1,c_2,\cdots,c_{n-r}$ 是任意常数, $\alpha_1,\alpha_2,\cdots,\alpha_{n-r}$ 是 $Ax=0$ 的一个基础解系.

当系数矩阵 $A=0$ 时, 齐次线性方程组 $Ax=0$ 的解是整个 $\mathbb{F}^n$, 它是2维平面或3维空间的自然推广, 因此称为 n 维线性空间或向量空间. 此时的基础解系称为该线性空间的**基**, 一组基包含的向量个数称为该线性空间的**维数**, 比如标准单位向量 $e_1,\cdots,e_n$ 构成 $\mathbb{F}^n$ 的一组基, 因此 $\mathbb{F}^n$ 的维数是 n. 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 $\mathbb{F}^n$ 的一组基, 则 $\mathbb{F}^n$ 的任何向量 α 均可唯一地表示为

$$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n,$$

有序数组 $(x_1,x_2,\cdots,x_n)$ 称为向量 α 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的**坐标**.

一般地, 齐次线性方程组 $Ax=0$ 的解集是 $\mathbb{F}^n$ 的关于向量加法与数乘 向量两种运算均封闭的特殊子集合, 称为 $\mathbb{F}^n$ 的**子空间**, 其维数为 $n-r(A)$. 值得指出的是, n 维线性空间 $\mathbb{F}^n$ 是一般抽象线性空间的原型, 对 $\mathbb{F}^n$ 的几乎所有研究方法和结果均可以不加改动地移植到任何一个线性空间中去, 这正是本章第四节的内容.

计算线性方程组的解和矩阵的秩等需要合适的办法, 这就是Gauss[4]消元法或初等变换. 初等变换可以通过下面的初等矩阵来实现.

1. (重排变换) 交换第 i 行 (列) 与第 j 行 (列) 的初等矩阵为 $I-E_{ii}-E_{jj}+E_{ij}+E_{ji}$;
2. (倍乘变换) 给第 i 行 (列) 乘以非零数 a 的初等矩阵为 $I+(a-1)E_{ii}$;
3. (倍加变换) 将第 j 行 (列) 的 a 倍加到第 i 行 (列) 的初等矩阵为 $I+aE_{ij}$.

注. 初学者务必仔细写出以上三种初等矩阵的 n 阶形式, 并确实搞清其结构.

[4]Johann Carl Friedrich Gauss(1777-1855), 德国著名数学家, 史称数学王子, 对数学的众多分支以及统计学, 物理学, 天文学, 大地测量学, 地理学以及电磁学等有重要贡献.

矩阵的一次行初等变换相当于左乘一个初等矩阵, 而矩阵的一次列初等变换相当于右乘一个初等矩阵. 比如, 设有按列分块矩阵 $A=(A_1,A_2,\cdots,A_n)$, 则

$$A(I+(a-1)E_{ii})=A+(a-1)AE_{ii}=(A_1,\cdots,A_{i-1},aA_i,A_{i+1},\cdots,A_n),$$

此即是将 A 的第 i 列乘 a 而其余列均不变的矩阵.

Gauss 消元法或初等变换的基本目的是将任意一个矩阵通过适当的初等变换化成结构比较简单的另一个矩阵, 其中由行初等变换能够得到的最简形式称为该矩阵的**Hermite 标准形**或**简化行阶梯形**.

定义 1.2.3 (Hermite 标准形) 设 $m\times n$ 矩阵 H 的秩为 r 且满足以下条件:

(1) 它的非零行恰为前 r 行, 且这 r 行的第一个非零元 (称为该行的先导元素) 为 1;

(2) 非零行的先导元素的列标随行标严格递增; 即若设第 k 行的先导元素出现在第 j_k 列, 则 $j_1<j_2<\cdots<j_r$;

(3) 非零行的先导元素所在的列的其它位置的元素为零, 即第 j_k 列为标准单位列向量 e_k;

则称 H 是 Hermite 标准形或简化行阶梯形.

定理 1.2.4 任一矩阵 A 都可经过一系列行初等变换化为 Hermite 标准形 H_A, 且相应的线性方程组同解, 即若 $H_A=PA$, 则 $Ax=b$ 与 $H_Ax=Pb$ 同解.

利用 Hermite 标准形可以得到矩阵的满秩分解, 其定义如下.

定义 1.2.4 设 $A\neq 0$. 如果列满秩矩阵 L 与行满秩矩阵 R 使得 $A=LR$, 则称 $A=LR$ 是 A 的一个满秩分解.

矩阵的满秩分解是不唯一的, 但矩阵 A 的满秩分解中的矩阵 L 与 R 的秩显然等于矩阵 A 的秩.

例 1.2.1 设 A 的秩为 r, 则存在可逆矩阵 P,Q 使得

$$A=P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q=\left(P\begin{pmatrix} I_r \\ 0 \end{pmatrix}\right)\left(\begin{pmatrix} I_r & 0 \end{pmatrix}Q\right) \tag{1.2.1}$$

注意 $P\begin{pmatrix} I_r \\ 0 \end{pmatrix}$ 是矩阵 P 的前 r 列构成的矩阵, 而 $\begin{pmatrix} I_r & 0 \end{pmatrix}Q$ 是矩阵 Q 的前 r 行构成的矩阵, 故它们分别是列满秩与行满秩的矩阵, 因此公式 (1.2.1) 给出矩阵 A 的一个满秩分解.

不过公式 (1.2.1) 给出的满秩分解需要记录变换矩阵 P 与 Q, 因此需要较大的计算量和较多的存储空间. 下面介绍一种不需要记录变换矩阵而只需要最后的 Hermite 标准形 H_A 的计算满秩分解的方法.

设矩阵 $A=(a_{ij})_{m\times n}$ 经过行初等变换化为 Hermite 标准形 H_A. 由 H_A 的定义, 可设

$$H_A=\begin{pmatrix} H_r \\ 0 \end{pmatrix},$$

其中 $r=r(A)$, H_r 为 $r\times n$ 的行满秩的 Hermite 标准形矩阵. 设 H_r 的第 k 行的先导元素所在的列标为 j_k, 则 $j_1<j_2<\cdots<j_r$, 且 H_r 的 $j_1, j_2, \cdots, j_r$ 列组成一个 r 阶单位矩阵, 故 A 的 $j_1, j_2, \cdots, j_r$ 列线性无关, 且

$$A=(A_{j_1},A_{j_2},\cdots,A_{j_r})H_r. \tag{1.2.2}$$

则公式 (1.2.2) 是 A 的一个满秩分解.

例 1.2.2 求下列矩阵的满秩分解:

$$A=\begin{pmatrix} 3 & -1 & 3 & 2 & 5 \\ 5 & -3 & 2 & 3 & 4 \\ 1 & -3 & -5 & 0 & -7 \\ 2 & -2 & -1 & 1 & -1 \end{pmatrix}$$

.

解 将 A 施行行初等变换化为 Hermite 标准形:

$$A=\begin{pmatrix} 3 & -1 & 3 & 2 & 5 \\ 5 & -3 & 2 & 3 & 4 \\ 1 & -3 & -5 & 0 & -7 \\ 2 & -2 & -1 & 1 & -1 \end{pmatrix} \to \begin{pmatrix} 0 & 8 & 18 & 2 & 26 \\ 0 & 12 & 27 & 3 & 39 \\ 1 & -3 & -5 & 0 & -7 \\ 0 & 4 & 9 & 1 & 13 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & -3 & -5 & 0 & -7 \\ 0 & 4 & 9 & 1 & 13 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & -3 & -5 & 0 & -7 \\ 0 & 1 & \frac{9}{4} & \frac{1}{4} & \frac{13}{4} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & \frac{7}{4} & \frac{3}{4} & \frac{11}{4} \\ 0 & 1 & \frac{9}{4} & \frac{1}{4} & \frac{13}{4} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = H_A.$$

所以

$$A=\begin{pmatrix} 3 & -1 \\ 5 & -3 \\ 1 & -3 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 1 & 0 & \frac{7}{4} & \frac{3}{4} & \frac{11}{4} \\ 0 & 1 & \frac{9}{4} & \frac{1}{4} & \frac{13}{4} \end{pmatrix}.$$

满秩分解可以用来计算低秩方阵的高次幂 (另一应用是求矩阵的广义逆, 见第六章).

例 1.2.3 求矩阵 A 的满秩分解, 并求 $A^n (n \geq 1)$, 其中

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

解 易知 A 的第一、第二列分别是第四、第三列的 -1 倍, 故 A 的满秩分解为

$$A=\begin{pmatrix} -3 & 2 \\ -2 & 1 \\ -2 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix} = LR.$$

于是, 当 $n \geq 1$ 时

$$A^n = (LR)^n = L(RL)^{n-1}R = \begin{pmatrix} -3 & 2 \\ -2 & 1 \\ -2 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 0 & -1 \end{pmatrix}^{n-1} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -2 \\ 2 & -3 \\ 2 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} (-2)^{n-1} & (-1)^{n-1} - (-2)^{n-1} \\ 0 & (-1)^{n-1} \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{3}{2}(-2)^n & \frac{1}{2}(-3(-2)^n + 2(-1)^n) & \frac{1}{2}(3(-2)^n - 2(-1)^n) & -\frac{3}{2}(-2)^n \\ (-2)^n & (-2)^n - (-1)^n & -(-2)^n + (-1)^n & -(-2)^n \\ (-2)^n & -(-2)^n & (-2)^n & -(-2)^n \\ \frac{1}{2}(-2)^n & -\frac{1}{2}(-2)^n & \frac{1}{2}(-2)^n & -\frac{1}{2}(-2)^n \end{pmatrix}.$$

例 1.2.4 设 A 是秩为 1 的 n 阶方阵, 试求 $A^n (n \geq 1)$.

解 设 $A = LR$ 为 A 的满秩分解, 其中 L, R 分别为列, 行向量. 则 RL 为一阶方阵, 且 $RL = \operatorname{tr}(RL) = \operatorname{tr}(LR) = \operatorname{tr} A$. 从而当 $n \geq 1$ 时, 有

$$A^n = (LR)^n = L(RL)^{n-1}R = (\operatorname{tr} A)^{n-1}LR = (\operatorname{tr} A)^{n-1}A.$$

思考题

1. 齐次线性方程组的解的几何意义是什么? 非齐次线性方程组的解与其对应的齐次线性方程组的解的几何意义是什么?
2. 设 A 是 $m \times n$ 矩阵, b 是非零向量. 问线性方程组 $AX = b$ 有解还是无解的可能性更大? 为什么?
3. 初等变换的几何意义是什么?
4. 试给出满秩分解的一种直观意义.

第三节 特征值与矩阵的相似对角化

例 1.2.4 表明利用满秩分解可以计算一些小秩矩阵的高次幂, 但该方法对于大秩特别是满秩矩阵无效. 因此需要更为普遍的方法, 这就是对角化方法. 确切地说, 设 A 是 n 阶复矩阵, 若存在对角矩阵 $D = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 与可逆矩阵 $P = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ 使得 $A = PDP^{-1}$, 则称 A 与 D 相似, 同时称 A 可以相似对角化或简称为可以对角化. (一般称矩阵 $P^{-1}AP$ 与矩阵 A 是相似的矩阵.) 如此则有 $A^m = PD^mP^{-1}$, 因此 A 的任意高次幂可以较方便地求出. 由于 $AP = PD$, 故 $A\alpha_i = \lambda_i\alpha_i$, $1 \leq i \leq n$. 所以 α_i 是齐次线性方程组 $Ax = \lambda_i x$ 的非零解, 即数 λ_i 与非零向量 α_i 均满足齐次线性方程组

$$Ax = \lambda x \tag{1.3.1}$$

因此方程组 (1.3.1) 的系数矩阵 $\lambda I - A$ 不可逆, 故诸 λ_i 均是多项式方程

$$|\lambda I - A| = 0$$

的根, 称为 A 的特征值或特征根 (物理等学科多称为**本征值**), 复系数多项式 $|\lambda I - A|$ 称为 A 的**特征多项式**. 对 A 的每个特征值 λ, 诸 α_i 均是线性方程组 $(\lambda I - A)x = 0$ 的非零解, 称为 A

的属于特征值 λ 的特征向量, 该齐次线性方程组的解集称为矩阵 A 的特征值 λ 的**特征子空间**, 记为 V_λ. 其维数 (等于 $n-r(\lambda I-A)$) 称为特征值 λ 的**几何重数**.

矩阵 A 的所有特征值的集合记为 $\sigma(A)$, 称为 A 的谱. 特征值的最大模称为 A 的**谱半径**, 记为 $\rho(A)$, 即

$$\rho(A)=\max\{|\lambda|:\lambda\in\sigma(A)\}.$$

从几何上看, 矩阵 A 的特征值全部位于以原点为圆心, 谱半径 $\rho(A)$ 为半径的圆盘内.

设 $\sigma(A)=\{\lambda_1,\lambda_2,\cdots,\lambda_s\}$, 且

$$|\lambda I-A|=\prod_{i=1}^{s}(\lambda-\lambda_i)^{n_i} \tag{1.3.2}$$

则正整数 n_i 称为特征值 λ_i 的**代数重数**.

定理 1.3.1 (特征值的性质) (1) 矩阵的行列式等于其所有特征值的积, 即 $|A|=\prod\limits_{i=1}^{s}(\lambda_i)^{n_i}$;

(2) 矩阵的迹等于其所有特征值的和, 即 $\operatorname{tr}(A)=\sum\limits_{i=1}^{s}n_i\lambda_i$;

(3) A 可逆 $\iff$ 0 不是 A 的特征值;

(4) 设 $f(x)$ 是任意多项式, λ 是 A 的一个特征值, α 是属于 λ 的特征向量, 则 $f(\lambda)$ 是 $f(A)$ 的一个特征值, α 是属于 $f(\lambda)$ 的特征向量;

(5) 设 A 可逆且其特征多项式为 (1.3.2), 则其逆矩阵的特征多项式为

$$|\lambda I-A^{-1}|=\prod_{i=1}^{s}(\lambda-\lambda_i^{-1})^{n_i},$$

且若 α 是 A 的属于特征值 λ 的特征向量, 则 α 也是 A^{-1} 的属于特征值 λ^{-1} 的特征向量;

(6) 任何特征值的几何重数不超过其代数重数;

(7) 相似矩阵具有相同的特征多项式 (因此具有相同的特征值).

定理 1.3.2 (特征向量的性质) (1) 属于不同特征值的特征向量线性无关;

(2) n 阶矩阵 A 可以对角化 $\iff$ A 有 n 个线性无关的特征向量 $\iff$ $\mathbb{F}^n$ 有一组由 A 的特征向量组成的基.

定理 1.3.3 (对角化主定理) 一个 n 阶矩阵 A 可以对角化 $\iff$ A 的每个特征值的代数重数与几何重数相等. 特别, 若 A 有 n 个不同的特征值, 则 A 可以对角化.

证 由于代数重数之和为 n, 故条件的充分性由 定理1.3.2 的 (2) 可得. 现证必要性. 设 A 可以对角化, 于是存在可逆矩阵 $P=(\alpha_1,\alpha_2,\cdots,\alpha_2)$ 与对角矩阵

$$D=\operatorname{diag}(\lambda_1,\cdots,\lambda_1,\lambda_2,\cdots,\lambda_2,\cdots,\lambda_s,\cdots,\lambda_s)$$

使得 $P^{-1}AP=D$, 即

$$(A\alpha_1,A\alpha_2,\cdots,A\alpha_n)=(\lambda_1\alpha_1,\cdots,\lambda_1\alpha_{n_1},\lambda_2\alpha_{n_1+1},\cdots,\lambda_2\alpha_{n_1+n_2},\cdots,\lambda_n\alpha_n),$$

故知 $\alpha_1, \cdots, \alpha_{n_1}$ 是属于 λ_1 的特征向量, $\alpha_{n_1+1}, \cdots, \alpha_{n_1+n_2}$ 是属于 λ_2 的特征向量, 等等. 因此对每个特征值 λ_i, $1 \le i \le n$, 其几何重数至少是其代数重数 n_i. 但由 定理1.3.1(6) 知任何特征值的几何重数最多是其代数重数, 故必相等. $\square$

思考题

1. 矩阵的特征向量和特征值有何直观意义?

2. 交换矩阵 A 的两行对其特征值与特征向量有何影响? 交换两列呢? 试总结之.

3. 如果同时交换矩阵 A 与 B 的相同两行 (比如同时交换第 1、2 行), 所得的矩阵相似, 那么 A 与 B 是否相似? 如果既交换 1、2 两行, 又交换 1、2 两列, 则又如何?

4. 能否有某种办法衡量有相同特征值的矩阵与无相同特征值的矩阵的多少? 你认为哪种多一些?

5. 能否有某种办法衡量可对角化的矩阵与不可对角化的矩阵的多少? 你认为哪种多一些?

第四节 线性空间

考察 n 维线性空间 $\mathbb{F}^n$ 可知, 其运算加法 "+" 满足下面五个条件:

(C) **封闭性**: 对任意 $\alpha, \beta \in \mathbb{F}^n$, $\alpha + \beta \in \mathbb{F}^n$;

(A1) **结合律**: 对任意 $\alpha, \beta, \gamma \in \mathbb{F}^n$, $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$;

(A2) **交换律**: 对任意 $\alpha, \beta \in \mathbb{F}^n$, $\alpha + \beta = \beta + \alpha$;

(A3) 存在**零向量**: 即存在元素 0, 使得对任意 $\alpha \in \mathbb{F}^n$, $\alpha + 0 = \alpha$;

(A4) 存在**负向量**: 对任意 $\alpha \in \mathbb{F}^n$, 存在一个向量, 记为 $-\alpha$, 使得 $\alpha + (-\alpha) = 0$.

带有满足上述条件的加法 "+" 的集合 V 广泛存在, 称为**加群**或**交换群** (顾名思义, 去掉交换律 (A2) 的系统就称为**群**, 不过此种情况下, 运算一般称为 "乘法" 而非 "加法", 运算符号也用 "$\bullet$" 或省略而不用 "+"), 记为 $(V, +)$ (此时 V 中的元素一般就称为 "元素" 而不是 "向量"), 在不至于混淆的情况下, 经常省略 $(V, +)$ 的加法符号 "+" 而简称 V 是加群. 比如我们在中小学就已经熟悉的整数集合 $\mathbb{Z}$, 有理数域 $\mathbb{Q}$, 实数域 $\mathbb{R}$ 与复数域 $\mathbb{C}$ 等均是加群 (其中的加法 "+" 均指普通加法), 但自然数集合 $\mathbb{N}$ 不是 (为什么?). 在线性代数课程中, 我们又知道数域 $\mathbb{F}$ 上的 $m \times n$ 阶矩阵的加法也满足上述五条性质, 因此 $\mathbb{F}^{m \times n}$ 是一个加群. 在高等数学或数学分析中, 闭区间上的连续函数全体 $C[a, b]$ 在函数的加法下 (即 $(f + g)(x) = f(x) + g(x)$) 也构成一个加群. 一般而言, 封闭性条件 (C) 是默认的, 即对任何非空集合 V, 称某种规则 R 是 V 上的 (二元)**运算**, 则该规则必须满足封闭性, 即对任意 $x, y \in V$, $xRy \in V$.

线性空间 $\mathbb{F}^n$ 的另一个特征是 $\mathbb{F}$ 中的数 (或称数字, 数量) a 与 $\mathbb{F}^n$ 中的元素 (称为向量) α 可以作 "乘法", 称为 "数乘", 记为 $a \bullet \alpha$(经常简记为 $a\alpha$), 注意这仍是 $\mathbb{F}^n$ 中的元素. $\mathbb{F}^n$ 中的数乘满足下列四个条件 (以下将 $\mathbb{F}^n$ 记为 V):

(B1) 数乘的结合律: 设 $a, b \in \mathbb{F}$, $\alpha \in V$, 有 $a(b\alpha) = (ab)\alpha$;

(B2) 数乘关于向量加法的分配律: 设 $a \in \mathbb{F}$, $\alpha, \beta \in V$, 有 $a(\alpha + \beta) = a\alpha + a\beta$;

(B3) 数乘关于数的加法的分配律: 设 $a, b \in \mathbb{F}$, $\alpha \in V$, 有 $(a + b)\alpha = a\alpha + b\alpha$;

(B4) 数乘的初始条件: $1 \bullet \alpha = \alpha$, 其中 $1 \in \mathbb{F}$.

由于数乘是两个不同的集合之间的联系, 所以需要与各自的 "运算" (即 V 中的加法 "+" 和 $\mathbb{F}$ 中的加法 "+" 和乘法) 保持和谐 (警示: 两个加法, 同一符号 "+", 不同含义! 参考下面的 例1.4.2), 这就是条件 (B1)-(B3) 的含义. 条件 (B4) 一是为了使数乘有意义 (否则, 令 $1 \bullet \alpha = 0$, 则将有 $a\alpha = 0$, 任意 $a \in \mathbb{F}$), 二是给数乘一个自然的初始规则 (有兴趣的读者可以研究如下问题: 若将条件 (B4) 改为 $1 \bullet \alpha = 2\alpha$ 将如何?). 同加法一样, 数乘是普通乘法的概括与抽象, 并不仅限于我们熟悉的系统, 比如所有数域的乘法, $\mathbb{F}^{m \times n}$ 中的数乘矩阵等等. 所以, 所谓非空集合 V 的 "加法" 泛指满足 (A1)-(A4) 的所有运算, 而 "数乘" 则是数域对于 V 的

“加法”满足 (B1)-(B4) 的所有运算, 因此“加法”与“数乘”均不限于我们所知道的普通加法与普通乘法, 请参考下面的 例 1.4.2.

定义 1.4.1 设 $(V,+)$ 是一个加群, 如果定义了数域 $\mathbb{F}$ 中的数与 V 中元素 (称为向量) 的数乘 (记为 $\bullet$, 一般省略不写), 则称 $(V,+,\bullet)$ 是数域 $\mathbb{F}$ 上的**线性空间**(或**向量空间**), 简称 V 是 $(\mathbb{F})$ 线性空间[5]. 特别, 当 $\mathbb{F}=\mathbb{R}$ 或 $\mathbb{C}$ 时, V 称为实空间或复空间. 一般地, 数域 $\mathbb{F}$ 称为线性空间 V 的**基域**. 本书除特别指明, 所有线性空间的基域均是指数域 $\mathbb{F}$(所以我们经常略去定语“数域 $\mathbb{F}$ 上的”), 而将 $\mathbb{F}$ 中的元素统称为“数”.

以下列出最常见的几种线性空间.

例 1.4.1 (1) 任何单元集 $\{\alpha\}$ 构成任何域上的线性空间, 只需将 α 定义成零向量即可;

(2) 任何域 $\mathbb{F}$ 按照其加法和乘法构成本身上的 1 维线性空间, 任何非零元素均构成 $\mathbb{F}$ 的一组基;

(3) 域 $\mathbb{F}$ 上的 $m\times n$ 矩阵全体按矩阵加法和数乘矩阵作成 $\mathbb{F}$ 上的 mn 维线性空间, 其一组基为全体基本矩阵 $E_{ij}, 1\le i\le m,\ 1\le j\le n$; 特别, 全体 n 阶方阵作成一个 n^2 维线性空间; 而全体 $n\times 1$ 阶矩阵, 即全体 n 维 (列) 向量构成 $\mathbb{F}$ 上的一个 n 维线性空间, 其一组基由所有标准向量构成, 即 $e_1, e_2, \cdots, e_n$;

(4) 系数取自 $\mathbb{F}$ 的全体一元多项式按多项式加法和数乘多项式作成 $\mathbb{F}$ 上的一个无限维线性空间 $\mathbb{F}[x]$, 其一组基为 $1, x, x^2, \cdots, x^n, \cdots$; 线性空间 $\mathbb{F}[x]$ 中次数小于 n 的多项式也构成 $\mathbb{F}$ 上的一个 n 维线性空间, 即 $\mathbb{F}[x]_n=\{f(x)\in\mathbb{F}[x]\,|\,\partial f(x)<n\}$($\partial f(x)$ 表示 $f(x)$ 的次数, 零多项式的次数可以约定为 $-\infty$), 其一组基为 $1, x, x^2, \cdots, x^{n-1}$;

(5) 我们已经知道 $C[a,b]$ 按照普通意义下的函数加法构成一个加群, 如果再按照普通意义定义数乘, 即对任何 $f\in C[a,b], \lambda\in\mathbb{R}$, 定义

$$(\lambda\bullet f)(x)=\lambda f(x), \forall x\in[a,b] \tag{1.4.1}$$

则 $C[a,b]$ 就是一个无限维实线性空间;

(6) 设 A 是任意非空集合, A 到数域 $\mathbb{F}$ 的所有函数 (映射) 的全体记为 $\mathbb{F}^A$. $\mathbb{F}^A$ 按照普通意义下的函数加法和数乘函数构成一个 $(\mathbb{F})$ 线性空间, 称为集合 A 上的**函数空间**;

(7) 设 A 是一个 $m\times n$ 矩阵, 则齐次线性方程组 $Ax=0$ 的所有解构成一个线性空间.

常常我们需要构造新的不熟悉的例子 (惟其如此, 才有可能“创新”), 试分析下面的例子:

例 1.4.2 (奇怪的加法与数乘) 设 $V=\{$所有正实数$\}, \mathbb{F}=\mathbb{R}$ 是实数域. 定义 V 中的加法运算为 $x\oplus y=xy$ (即通常的实数乘法); 定义 V 中元素与 $\mathbb{F}$ 中数的数乘运算为 $k\bullet x=x^k$(通常的幂运算). 则 $(V,\oplus,\bullet)$ 是实线性空间 (证明见习题 21).

我们在习题 22-23 中也设计了几个不同寻常的“加法”与“数乘”, 请读者仔细体会.

由线性空间的两种运算“加法”和“数乘”(合称为线性运算), 可以与 n 维线性空间 $\mathbb{F}^n$ 类似导出三个最为基本的概念: 线性组合, 线性相关与线性无关. 实际上, 我们稍后将看到, 本书讨论的线性空间与 $\mathbb{F}^n$ 的唯一差别是元素的名称 (此差别显然是无关紧要的).

---

[5]线性空间的一般定义由意大利数学家 Giuseppe Peano 于 1888 年给出.

定义 1.4.2 如果线性空间 V 中存在 n 个线性无关的向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得 V 中任意向量和它们线性相关, 则称 V 是 n 维线性空间, 向量组 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 称为 V 的一组**基**, 向量 α_j 称为**基向量**, 非负整数 n 称为 V 的维数, 记作 $\dim V$ (或更为准确地, $\dim_{\mathbb{F}} V$). 如果不存在这样的有限整数, 则线性空间 V 称为是**无限维线性空间**. $n = 0$ 时, 线性空间 V 没有基, 称为平凡线性空间或零线性空间, 简记为 0 (注意: 本书使用符号 0 表示数字 0, 向量 0 和矩阵 0 等不同对象).

例 1.4.1 中的 (1)-(4) 的基较为明显, 然而寻找该例 (5) 中的线性空间 $C[a, b]$ 的一组基就非常不容易了 (基的存在性便是很难的问题). 当 A 是有限集时, (6) 中的函数空间 $\mathbb{F}^A$ 的基较易构造, 见习题 24. 矩阵理论课程的讨论几乎只限于有限维线性空间, 无限维线性空间是泛函分析等课程的研究对象.

定理 1.4.1 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 n 维线性空间 V 的一组基, 则 V 中任意向量 α 均可唯一地表为线性组合

$$\alpha = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n. \tag{1.4.2}$$

由组合系数 $k_1, k_2, \cdots, k_n$ 确定的 n 元有序数组称为向量 α 关于基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的坐标, k_j 称为第 j 个坐标或分量. 本书将坐标写成列向量的形式, 即 $(k_1, k_2, \cdots, k_n)^T$.

一般来说, 线性空间的基不是唯一的 (对照极大线性无关组), 但线性空间的维数是唯一确定的 (见习题 26), 因为可以证明, 不同的基包含的向量个数相同 (对照向量组的秩), 这可由下面的定理推出 (对照线性代数课程的替换定理).

定理 1.4.2 n 维线性空间中任意 $n+1$ 个向量必线性相关.

证明轮廓 设这 $n+1$ 个向量的一个线性组合等于 0, 由于这 $n+1$ 个向量的每一个都可由一组基的基向量线性表示, 于是可以得到一个 $n+1$ 个未知数 n 个方程的齐次线性方程组, 因此该方程组必有非零解, 即这 $n+1$ 个向量线性相关.

推论 1.4.1 n 维线性空间中任意 n 个线性无关的向量均构成一组基, 且任何一组基恰含 n 个向量.

定理 1.4.3 n 维线性空间中任意 r 个线性无关的向量均能扩充成一组基.

设 V 是 n 维线性空间, $\alpha_1, \alpha_2, \cdots, \alpha_n$ 与 $\beta_1, \beta_2, \cdots, \beta_n$ 是 V 的两组基, 分别简称为 α– 基与 β– 基. 它们的关系可以用下面的方程组表示:

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

或用矩阵形式表达为

$$(\beta_1, \beta_2, \cdots, \beta_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)P, \tag{1.4.3}$$

其中 n 阶矩阵 $P = (p_{ij})$ 称为由 α– 基到 β– 基的**过渡矩阵**. 显然, 此时由 β– 基到 α– 基的过渡矩阵为 P^{-1}(过渡矩阵必是可逆矩阵, 见习题 27).

现在假设 $\gamma \in V$ 在 $\alpha-$ 基下的坐标为 $x = (x_1, x_2, \cdots, x_n)^T$, 关于 $\beta-$ 基的坐标为 $y = (y_1, y_2, \cdots, y_n)^T$. 问 x 与 y 有何联系? 为此, 将 $\gamma = y_1\beta_1 + y_2\beta_2 + \cdots + y_n\beta_n$ 代入公式 (1.4.3) 可得

$$\begin{aligned}\gamma = \ & y_1(p_{11}\alpha_1 + p_{21}\alpha_2 + \cdots + p_{n1}\alpha_n) \\ & + y_2(p_{12}\alpha_1 + p_{22}\alpha_2 + \cdots + p_{n2}\alpha_n) + \cdots \\ & + y_n(p_{n1}\alpha_1 + p_{n2}\alpha_2 + \cdots + p_{nn}\alpha_n) \\ = \ & (y_1p_{11} + y_2p_{12} + \cdots + y_np_{1n})\alpha_1 \\ & + (y_1p_{21} + y_2p_{22} + \cdots + y_np_{2n})\alpha_2 + \cdots \\ & + (y_1p_{n1} + y_2p_{n2} + \cdots + y_np_{nn})\alpha_n,\end{aligned}$$

即

$$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)(P(y_1, y_2, \cdots, y_n)^T),$$

由坐标的唯一性可知,

$$\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = P\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}, \quad \text{或} \quad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = P^{-1}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \tag{1.4.4}$$

公式 (1.4.4) 称为**坐标变换公式**.

例 1.4.3 设 $V = \mathbb{R}^2$, 取 $\alpha_1 = (1,0)^T$, $\alpha_2 = (0,1)^T$ 与 $\beta_1 = (1,1)^T$, $\beta_2 = (-1,1)^T$, 求由 $\alpha-$ 基到 $\beta-$ 基的过渡矩阵以及 $\gamma = (2,3)^T$ 在 $\beta-$ 基下的坐标.

解 由于 $\alpha-$ 基是标准基, 易知由 $\alpha-$ 基到 $\beta-$ 基的过渡矩阵为

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

写成分块矩阵有 $P=(\beta_1,\beta_2)$. 因此, $\gamma=(2,3)^T$ 在 $\beta-$ 基下的坐标为 $P^{-1}(2,3)^T = (5/2, 1/2)^T$.

例 1.4.4 求任意 $f(x) \in \mathbb{F}[x]_n$ 在标准基 $1, x, x^2, \cdots, x^{n-1}$ 下的坐标. 向量组 $1, x-a, (x-a)^2, \cdots, (x-a)^{n-1}$ 也构成 $\mathbb{F}[x]_n$ 的一组基, 求 $f(x)$ 在该组基下的坐标. 求从标准基到第二组基的过渡矩阵.

解 由高等数学或数学分析中的 Taylor[6] (泰勒) 公式可知:

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!}x^k,$$

即 $f(x)$ 在标准基下的坐标为

$$(f(0), f'(0), \cdots, \frac{f^{(k)}(0)}{k!}, \cdots, \frac{f^{(n-1)}(0)}{(n-1)!})^T.$$

[6]Brook Taylor(1685-1739), 英国数学家.

对任意 $a\in\mathbb{F}$, 仍由 Taylor 公式可知, $f(x)$ 在第二组基下的坐标为

$$(f(a), f'(a), \cdots, \frac{f^{(k)}(a)}{k!}, \cdots, \frac{f^{(n-1)}(a)}{(n-1)!})^T.$$

由于

$$(x-a)^k=\sum_{r=0}^{k} C_k^r x^r(-a)^{k-r},$$

故从标准基到第二组基的过渡矩阵 $P=(p_{ij})$ 由下式给出:

$$p_{ij}=\begin{cases} C_{j-1}^{i-1}(-a)^{j-i} & i\le j,\\ 0 & i>j.\end{cases}$$

思考题

1. 将线性空间的条件 (B4) 即 $1\bullet\alpha=\alpha$ 改为 $1\bullet\alpha=2\alpha$ 将如何?

2. 线性空间的定义实际上没有用到每个非零元素均有逆元这个条件. 如何改造线性空间的定义, 使其包括更多的系统, 比如包括通常加法和乘法下的整数集合 (去掉数域 $\mathbb{F}$ 中每个非零元素均有逆元的条件将得到**数环**的概念)?

3. 设 $u=u(x,y,z,t)$ 是未知函数, c 是常数, $\nabla^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}$ 是 **Laplace**[7] **算符**. **波动方程** $\frac{\partial^2 u}{\partial t^2}=c^2\nabla^2 u$ 的全体解是否构成线性空间? 若 u 与时间 t 无关, 则波动方程变为 **Laplace 方程** $\nabla^2 u=0$. 该方程的全体解是否构成线性空间? 总结之.

4. 试给出基与基向量一个直观的解释.

5. 试给出过渡矩阵的一种直观解释.

第五节 内积空间与正定二次型

由 定理1.3.3 知, 并非每个矩阵都可以对角化, 但任何实对称矩阵却可以以更精细的方式对角化. 为此需要正交矩阵的概念, 由于这个概念有很强的几何背景, 引出该概念的自然方式是先在一般实或复线性空间内引入长度, 而**内积**较长度更容易处理 (因为内积具有双线性性质), 因此有下面的定义.

定义 1.5.1 设 $\mathbb{F}$ 是实数域或复数域, V 是 $\mathbb{F}$ 上的线性空间. 若对 V 中任意两个向量 α,β, 都定义了 $\mathbb{F}$ 中一个数 (α,β)(称为向量 α 与 β 的内积), 使得

(1) (共轭对称性) $(\alpha,\beta)=\overline{(\beta,\alpha)}$, 其中 $\overline{(\beta,\alpha)}$ 是复数 (β,α) 的共轭复数;

(2) (正定性) $(\alpha,\alpha)\ge 0$, 且等号成立 $\iff \alpha=0$;

(3) (双线性) $(a\alpha+b\beta,\gamma)=a(\alpha,\gamma)+b(\beta,\gamma)$, 对任意 $\alpha,\beta,\gamma\in V,\ a,b\in\mathbb{F}$ 成立;

则称 V 为一个**内积空间**.

请注意, 由于内积的共轭对称性, 内积的第三个条件“双线性”仅对第一个变量成立, 而对第二个变量是“共轭双线性”的, 即

$$(\alpha, a\beta+b\gamma)=\bar{a}(\alpha,\beta)+\bar{b}(\alpha,\gamma).$$

若 $\mathbb{F}=\mathbb{R}$ 是实数域, 则内积确为对称的 (也称为可交换的), 即 $(\alpha,\beta)=(\beta,\alpha)$, 故此时的内积对两个变量均为“双线性”的; 有限维实内积空间又称为欧几里得空间或**欧氏空间**; 当 $\mathbb{F}=\mathbb{C}$ 是复数域, 则内积空间称为**酉空间**(或复内积空间). 除特别说明, 本书仅限于讨论有限维内积空间.

[7]Pierre-Simon Laplace(1749-1827), 著名法国数学家, 物理学家, 天文学家, 被称为法国的 Newton (牛顿).

例 1.5.1 设 $V=\mathbb{R}[x]_n$ 是次数小于 n 的实系数多项式构成的 n 维实线性空间, $a<b$ 是实数. 设 $f,g\in V$, 规定

$$(f,g)=\int_a^b f(x)g(x)\,\mathrm{d}\,x, \tag{1.5.1}$$

则 (f,g) 定义了 V 上的一个内积. (实际上, 公式 (1.5.1) 也是无限维空间 $C[a,b]$ 上的内积.)

非负实数 (α,α) 的算数平方根定义为向量 α 的**模**(或**范数**, 或长度), 记为

$$\|\alpha\|=\sqrt{(\alpha,\alpha)}.$$

特别, 模为 1 的向量称为**单位向量**或**标准向量** .

命题 1.5.1 (内积与范数的性质) (1) $(a\alpha,\,b\beta)=a\bar{b}(\alpha,\beta)$;
(2) $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$;
(3) $(0,\alpha)=(\alpha,0)=0$;
(4) $\|c\alpha\|=|c|\|\alpha\|$;
(5) (Cauchy-Schwarz [8]不等式)$|(\alpha\,,\beta)|\leq\|\alpha\|\|\beta\|$ 且等号成立 $\iff$ α 与 β 线性相关;
(6) (三角不等式) $\|\alpha+\beta\|\leq\|\alpha\|+\|\beta\|$.

由 Cauchy-Schwarz 不等式即可一般地定义“角度”, 即设 α 与 β 的夹角为 θ, 则

$$\theta=\arccos\frac{|(\alpha\,,\beta)|}{\|\alpha\|\|\beta\|} \tag{1.5.2}$$

此时 $0\leq\theta\leq\frac{\pi}{2}$ (参考本节思考题 2). 特别地, 若 $\mathbb{F}=\mathbb{R}$, 则此夹角可以更精确地定义为

$$\theta=\arccos\frac{(\alpha\,,\beta)}{\|\alpha\|\|\beta\|} \tag{1.5.3}$$

此时有 $0\leq\theta\leq\pi$.

内积具有比较明显的物理意义, 即“恒力作功”. 范数则揭示了内积的几何意义, 比如我们最熟悉的欧氏空间 $V=\mathbb{R}^n$ 中的通常内积一般定义为

$$(\alpha,\beta)=\beta^T\alpha=\alpha^T\beta=(\beta,\alpha);$$

此即是坐标乘积之和; 而 $U=\mathbb{C}^n$ 中的内积则定义为

$$(\alpha,\beta)=\beta^*\alpha=\overline{(\beta,\alpha)}=\overline{\alpha^*\beta}.$$

在 $\mathbb{R}^2$ 中有

$$(\alpha\,,\beta)=\|\alpha\|\|\beta\|\cos\theta, \tag{1.5.4}$$

[8]Augustin-Louis Cauchy(1789-1857), 法国著名数学家, 对数学的多个分支以及数学物理有重要贡献, 著名的 ϵ-δ- 语言即是其一大发明. Hermann Schwarz(1843-1921), 德国数学家. Cauchy-Schwarz 不等式首先由 Cauchy 于 1821 年发现, 再由俄国数学家 Victor Yakovlevich Bunyakovsky(1804-1889) 于 1859 年重新发现, 后由 Schwarz 于 1888 年再次重新发现.

其中 θ 是 α 与 β 的夹角. 由于

$$(\alpha-\beta\,,\alpha-\beta)=\|\alpha\|^2-2(\alpha,\,\beta)+\|\beta\|^2. \tag{1.5.5}$$

记 $a=\|\alpha\|$, $b=\|\beta\|$, $c=\|\alpha-\beta\|$. 则由等式 (1.5.4) 与 (1.5.5) 可得余弦定理:

$$c^2=a^2+b^2-2ab\cos\theta. \tag{1.5.6}$$

而当 $\cos\theta=0$ 即 $\theta=\frac{\pi}{2}$ 时, 即得勾股定理.

内积的引入使得线性空间的结构更加丰富多彩, 因为长度与角度实际上带来了“图形”与“位置”等概念 (由此线性空间能够更好地模拟真实的宇宙). 由角度的定义立即可知, 两个非零向量平行 (即线性相关) 当且仅当其夹角为 0 或 π (当 $\mathbb{F}=\mathbb{R}$ 时), 而两个非零向量垂直当且仅当其夹角为 $\frac{\pi}{2}$. 两个向量的相互位置中以“垂直”最为重要, 故有下面的定义.

定义 1.5.2 两个内积为 0 的向量 α 与 β 称为**正交**的, 记为 $\alpha\perp\beta$. 一组非零向量如果两两正交则称为**正交组**. 单位向量构成的正交组称为**标准正交组**.

命题 1.5.2 正交组必是线性无关的.

定理 1.5.1 内积空间必存在**标准正交基**.

证 对内积空间 V 的任意一个线性无关组 $\alpha_1,\,\alpha_2,\cdots,\,\alpha_s$, 利用 **Gram-Schmidt** [9] **正交化方法** 即可求得 V 的一个标准正交组, 即取

$$\beta_1=\alpha_1\,,\quad \gamma_1=\frac{\beta_1}{\|\beta_1\|}.$$

归纳地, 假设 $\gamma_1,\,\gamma_2,\cdots,\,\gamma_{k-1}$ 已被确定, 则取

$$\beta_k=\alpha_k-(\alpha_k,\gamma_{k-1})\gamma_{k-1}-(\alpha_k,\gamma_{k-2})\gamma_{k-2}-\cdots-(\alpha_k,\gamma_1)\gamma_1. \tag{1.5.7}$$

而令

$$\gamma_k=\frac{\beta_k}{\|\beta_k\|}.$$

则 $\gamma_1,\,\gamma_2,\cdots,\,\gamma_s$ 是 V 的一个标准正交组, 且

$$[\alpha_1,\alpha_2,\cdots,\alpha_{k-1}]=[\gamma_1,\gamma_2,\cdots,\gamma_{k-1}],\ 1\le k\le s.$$

因此, 对内积空间 V 的任意一组基 $\alpha_1,\,\alpha_2,\cdots,\,\alpha_n$, 利用 Gram-Schmidt 正交化方法可求得 V 的一个标准正交基. □

注意, 如果 β 是一个非零向量, α 是任意向量, 则向量 $\frac{(\alpha,\beta)}{\|\beta\|}\beta$ 恰好是向量 α 在向量 β 上的**投影向量**, 一般记为 $\mathrm{Proj}_\beta\alpha$. 特别, 若 β 是一个单位向量, 则 $\mathrm{Proj}_\beta\alpha=(\alpha,\beta)\beta$. 因此, Gram-Schmidt 正交化方法的几何意义是: 设 $\alpha_1,\alpha_2\in\mathbb{R}^3$ 线性无关, 则向量 $\alpha\in\mathbb{R}^3$ 在 α_1, α_2

[9] Jørgen Pedersen Gram(1850-1916), 丹麦著名数学家与精算学家, 曾任丹麦保险协会主席, 1883 年发现 Gram-Schmidt 正交化方法, 被自行车撞伤而亡. Erhard Schmidt(1876-1959), 德国数学家, 师从德国著名数学家 David Hilbert, 1907 年发现 Gram-Schmidt 正交化方法. 实际上该方法早在 1836 年即已被 Cauchy 所使用.

确定的平面上的投影等于它在该平面的一组正交坐标轴上的投影之和. 比如, 公式 (1.5.7) 中的被减各项恰好是 α_k 在诸 $\gamma_j\,(1 \le j \le k-1)$ 上的投影.

(注意: Gram-Schmidt 正交化方法是数值不稳定的, 在计算机上使用时需要改进.)

设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 $V = \mathbb{R}^n$ 或 $V = \mathbb{C}^n$ 的一组标准正交基, 则矩阵

$$Q = (\alpha_1, \alpha_2, \cdots, \alpha_n)$$

称为**酉矩阵**. 实的酉矩阵称为**正交矩阵**.

定理 1.5.2 (酉矩阵的刻画) 矩阵 Q 是酉矩阵 $\iff Q^*Q = I$. 特别, 实矩阵 A 是正交矩阵 $\iff Q^TQ = I \iff Q^{-1} = Q^T$.

一般而言, 逆矩阵的计算是非常困难的问题, 但 定理 1.5.2 告诉我们, 酉矩阵的逆矩阵就是其共轭转置矩阵!

在第二章我们将看到, 当 $n = 2$ 时, 正交矩阵的几何意义几乎就是平面解析几何中的旋转变换. 由于可以利用旋转变换将中心在原点的有心二次曲线 (即椭圆和双曲线) 变成标准形式 (这是化简二次型的范例), 故利用正交矩阵就可以将实对称矩阵对角化. 需要注意的是, 与实对称矩阵相应的复数矩阵不是复对称矩阵, 而是复共轭对称矩阵, 习惯上称为 **Hermite 矩阵**, 即需要满足 $A^* = A$. Hermite 矩阵是最重要的一类矩阵, 比如量子力学中的物理量均是由 Hermite 矩阵来表达的, 而酉空间中的内积则由所谓“正定” Hermite 矩阵完全确定 (见下面的 定理 1.5.6).

定理 1.5.3 Hermite 矩阵的特征值均为实数, 且属于不同特征值的特征向量彼此正交 (证明见习题 20).

定理 1.5.4 (Hermite 矩阵的酉对角化) Hermite 矩阵 A 可以酉对角化, 即存在酉矩阵 U 使得 $U^*AU = D$ 是对角矩阵. 特别, 实对称矩阵可以正交对角化.

证明思路: 由 定理 1.5.3 可知, 可以选出 n 维线性空间 $\mathbb{F}^n(\mathbb{F} = \mathbb{R}$ 或 $\mathbb{C})$ 的一组标准正交基, 使得每个基向量都是该 Hermite 矩阵的特征向量, 以此组标准正交基为列作成的矩阵就是所需要的酉矩阵. 容易验证, 实对称矩阵具有实特征向量, 故定理成立.

定理 1.5.4 为化简一般 n 元 (Hermite) 二次型提供了理论依据.

设 $x = (x_1, \cdots, x_n)^T$ 是未定元 (即未知数) 向量, 称下面关于 $x_1, \cdots, x_n, \bar{x}_1, \cdots, \bar{x}_n$ 的复系数二次多项式

$$f(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}\bar{x}_i x_j$$

为 (复) Hermite 二次型, 简称二次型. 易知存在唯一的 n 阶 Hermite 矩阵 $A = (a_{ij})$ 使得 $f(x) = x^*Ax$, 该矩阵 A 称为二次型 $f(x)$ 的矩阵.

定义 1.5.3 设 $f(x) = x^*Ax$ 是复二次型, A 是 Hermite 矩阵, 若对任意非零向量 $\alpha \in \mathbb{C}^n$, 均有 $f(\alpha) = \alpha^*A\alpha > 0$, 则称 $f(x)$ 是正定二次型, A 是正定矩阵. 类似地可以定义半正定, 负定, 半负定二次型, 半正定, 负定, 半负定矩阵等. 对实二次型和实对称矩阵的定义类似.

定理 1.5.5 设 A 是 n 阶 Hermite 矩阵, 则下列条件等价:
(1) A 是正定的;
(2) $f(x) = x^*Ax$ 是正定二次型;
(3) A 的特征值均为正实数;
(4) 存在 $m \times n$ 阶列满秩矩阵 M 使得 $A = M^*M$;
(5) 存在 n 阶可逆矩阵 M 使得 $A = M^*M$;
(6) 存在 n 阶可逆矩阵 P 使得 $P^*AP = I$(即 A 与 I **合同**).

由于存在酉矩阵 U 与实对角矩阵 $D = \mathrm{diag}\,(\lambda_1\,, \cdots, \lambda_n)$ 使得 $U^*AU = D$ (故 $\lambda_1\,, \cdots, \lambda_n$ 是 A 的实特征值), 作坐标变换 $x = Uy\,(y = (y_1, y_2, \cdots, y_n)^T)$ 即可将 $f(x)$ 化简为下面的对角形式:

$$f(x) = (Uy)^*A(Uy) = \lambda_1|y_1|^2 + \lambda_2|y_2|^2 + \cdots + \lambda_n|y_n|^2 \tag{1.5.8}$$

上式具有重要意义, 因为由此不仅可以判断二次型 $f(x)$ 及其矩阵 A 的正定性等而且可以确定 (实) 二次型 $f(x)$ 所确定的几何体 $f(x) = 0$ 或 $f(x) = 1$ 的形状 (对照: 如果使用更广泛的坐标变换 $x = Py$, 其中 P 为可逆矩阵, 则还可以将 $f(x)$ 化简为比公式 (1.5.8) 更简单的形式, 但却失去了大部分几何意义).

正定矩阵与内积有密切的关系. 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 n 维内积空间 V 的一组基, 设 $a\,, \beta \in V$ 的坐标分别是 $x = (x_1,\, x_2,\, \cdots,\, x_n)^T$ 与 $y = (y_1,\, y_2,\, \cdots,\, y_n)^T$, 则有

$$(\alpha\,, \beta) = (\sum_{i=1}^{n} x_i\alpha_i, \sum_{j=1}^{n} y_j\alpha_j) = \sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha_i,\, \alpha_j)x_i\bar{y_j}. \tag{1.5.9}$$

记 $(\alpha_i,\, \alpha_j) = a_{ij}$, 则公式 (1.5.9) 变为

$$(\alpha\,, \beta) = y^*Ax \tag{1.5.10}$$

其中 $A = (a_{ij})$ 是 Hermite 矩阵. 由 定义 1.5.1(2) 的正定性可知, $A = (a_{ij})$ 还是正定矩阵, 称为基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的**度量矩阵**或 Gram[10]矩阵. 反过来, 如果 Hermite 矩阵 $A = (a_{ij})$ 是正定的, 则容易验证由公式 (1.5.10) 定义的二元向量函数确是 n 维线性空间 V 的一个内积. 因此有下述定理:

定理 1.5.6 (内积与正定矩阵) 设 V 是 n 维复线性空间, 则其上的内积与正定矩阵一一对应. 确切地说, 设 $(\alpha\,, \beta)$ 是 V 上的二元向量函数, 则 $(\alpha\,, \beta)$ 是内积 $\iff$ 存在正定 Hermite 矩阵 $A = (a_{ij})$ 使得 $(\alpha\,, \beta) = y^*Ax$, 其中 x 与 y 分别是 α 与 β 在某组基下的坐标.

由 定理 1.5.6 可知, 任何非平凡实或复线性空间上均可以定义无限多种不同的内积, 我们将在本书第五章看到这些“不同的”内积之间的关系 (结论是: 有限维实或复线性空间上的任意两个内积**本质上是完全相同的**!). 本章习题 29-42 展示了一些实或复线性空间上各种各样的内积.

标准正交基可以使内积的计算大为简化. 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是内积空间 V 的一个标准正交基, $\alpha, \beta \in V$, 则

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n,\ \beta = b_1\alpha_1 + b_2\alpha_2 + \cdots + b_n\alpha_n,$$

[10]即 Gram-Schmidt 正交化方法中的 J.P.Gram.

从而

$$\begin{aligned}(\alpha,\beta) &= (a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n, b_1\alpha_1 + b_2\alpha_2 + \cdots + b_n\alpha_n) \\ &= a_1\bar{b}_1 + a_2\bar{b}_2 + \cdots + a_n\bar{b}_n,\end{aligned}$$

即在标准正交基下, 向量的内积就是对应坐标的 (共轭) 乘积之和; 这与 $\mathbb{R}^n$ 或 $\mathbb{C}^n$ 的普通内积完全一致. 这里的本质是标准正交基的度量矩阵乃是单位矩阵, 因此其确定的二次型没有交叉项且平方项的系数均为 1.

对任意 n 阶复矩阵 $A = (a_{ij})$, 一般称

$$y^*Ax = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i\bar{y}_j \tag{1.5.11}$$

为矩阵 A(对应) 的 (n 维)**双线性型**或 $\mathbb{C}^n$ 上的双线性型 (注意, 它对第一个变元 x 是线性的, 对第二个变元 y 是共轭线性的). 二次型 x^*Ax 称为双线性函数 y^*Ax 的二次型. 按照二次型的正定, 半正定等可以定义相应的双线性型的正定性, 半正定性等概念. 于是内积不过是一种特殊的即正定的双线性型而已. 但注意由公式 (1.5.11) 定义的双线性型中的矩阵 A 不必是 Hermite 的, 因此 A 不必是该双线性型的二次型的矩阵.

例 1.5.2 实双线性型 $(x,y) = ax_1y_1 + bx_1y_2 + bx_2y_1 + cx_2y_2$ 的矩阵形式为:

$$(x,y) = (y_1, y_2)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = y^TGx.$$

因此该双线性型 定义 $\mathbb{R}^2$ 的一个内积 $\iff$ 其定义矩阵

$$G = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

是正定矩阵. (请思考: 如果 x_1y_2 与 x_2y_1 的系数不相等如何?)

例 1.5.3 (度量矩阵的行列式的几何意义) 设 $\alpha_1, \alpha_2 \in \mathbb{R}^2$, 考察行列式

$$D = \begin{vmatrix} (\alpha_1,\alpha_1) & (\alpha_2,\alpha_1) \\ (\alpha_1,\alpha_2) & (\alpha_2,\alpha_2) \end{vmatrix}$$

的几何意义.

解 设 α_1 与 α_2 的夹角为 θ. 直接计算可知 (注意 $(\alpha_2,\alpha_1) = (\alpha_1,\alpha_2)$)

$$D = \|\alpha_1\|^2\|\alpha_2\|^2 - (\alpha_1,\alpha_2)^2 = \|\alpha_1\|^2\|\alpha_2\|^2 - \|\alpha_1\|^2\|\alpha_2\|^2\cos^2\theta = \|\alpha_1\|^2\|\alpha_2\|^2\sin^2\theta$$

这恰好是以 α_1, α_2 为邻边的平行四边形的面积的平方. (对照: 在线性代数课程中, 三阶行列式的几何意义是平行六面体的有向体积.)

思考题

1. 将内积的正定性条件去掉将如何? 是否这是无稽之谈?
2. 比较公式 (1.5.2) 与 (1.5.3), 说明为什么不在公式 (1.5.2) 中将内积的绝对值符号去掉?
3. 正交性概念是通常垂直概念的推广. Gram-Schmidt 正交化方法在立体几何中有何解释?

4. 试给出标准正交基的一个直观解释.

5. 由标准正交基到另一组标准正交基的过渡矩阵有何特点?

6. 设 $\mathbb{F} = \mathbb{C}$ 或 $\mathbb{R}$. $\mathbb{F}$ 上的 n 元二次型全体是否构成 $\mathbb{F}$ 上的线性空间? n 维双线性型全体呢?

7. 试对 $\mathbb{F}$ 上的任意 m 维向量 x 与 n 维向量 y, 推广双线性型的概念. 这样的双线性型全体是否构成 $\mathbb{F}$ 上的线性空间?

8. 三阶度量矩阵的行列式有何几何解释?

9. 设 $(\bullet,\bullet)_i, i = 1,2$ 是 n 维实线性空间 V 上的两个不同的内积, $\alpha,\beta \in V$. 是否可能 $(\alpha,\beta)_1 = 0$ 但 $(\alpha,\beta)_2 \neq 0$? 是否可能 $(\alpha,\alpha)_1 < (\beta,\beta)_1$ 但 $(\alpha,\alpha)_2 > (\beta,\beta)_2$? 一般地, 这两个内积有何关系?

10. 试对 n 维实线性空间 V 上的双线性型讨论上题类似的问题?

第六节 应用: 网络流, 投入产出模型, 随机变量的独立性

线性方程组, 矩阵等概念和方法无论在工业生产, 信息技术还是经济社会科学领域均有非常广泛的应用, 本节我们仅作一简要介绍.

一. 网络流

网络流是最简单也最常见的实际问题, 比如经济学中的收支平衡, 车流量, 化学方程式配平等, 其数学模型是线性方程组, 著名的 Kirchhoff[11] 第一第二定律实际上就是平衡电路网络的线性方程组. 这类问题通常都涉及到较多的未知数, 常需要借助相应的数值计算软件 (如 Matlab, Maple, Mathematica 等) 予以解决.

例 1.6.1 下面是某一区域某天的 (平均) 交通流量图:

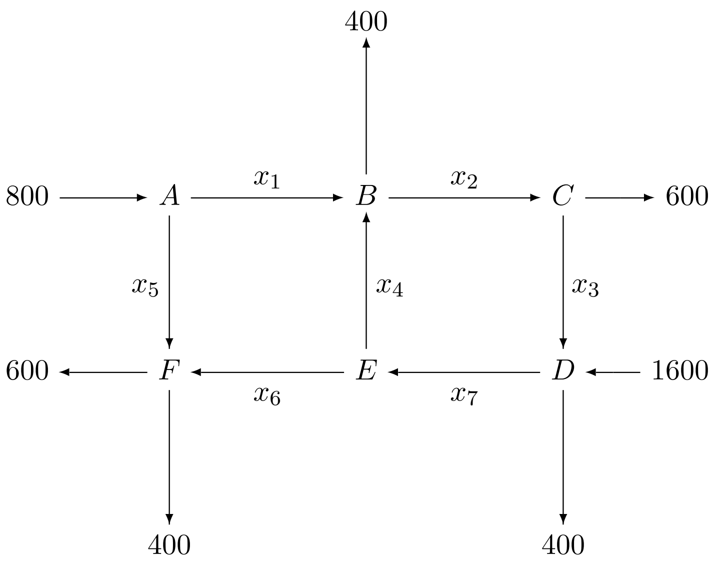

图 1.6.1

图中有 A, B, C, D, E, F 共 6 个路口, 已有 7 条街道记录了当天的平均车流量. 另有 7 处的平均车流量 $x_i\,(1 \leq i \leq 7)$ 未知, 试利用每个路口的进出车流量相等关系推算这 7 处的平均车流量.

分析 由于进入每个路口的流量和离开该路口的流量相等, 故可建立含有 7 个未知数 $x_i\,(1 \leq i \leq 7)$ 的 6 个方程, 因此这是求解线性方程组的问题, 见习题 51.

二. 投入产出模型

[11] Gustav Robert Kirchhoff(1824-1887), 德国著名物理学家.

以下我们简要介绍经济学中著名的Leontief[12]投入产出模型. 按W.Leontief的理论, 可以将一个国家的经济体系分为 n 个生产商品和服务的部门, 以及一些不生产产品或服务 (而仅仅消费商品与服务) 的部门. 用 $x \in \mathbb{R}^n$ 记录每一部门一年的产出, 称为**产出向量**; 用 $d \in \mathbb{R}^n$ 表示各种非生产部门所需求的商品与服务, 称为**最终需求向量**.

显然 x 与 d 的元素均非负, 称为非负向量. 类似地可以定义**非负矩阵**. Leontief的投入产出模型 (又称生产方程) 为

$$(\text{Leontief投入产出模型}) \qquad x = Cx + d \tag{1.6.1}$$

其中 $C \in M_n(\mathbb{R})$ 称为**消耗矩阵**, Cx 称为**中间需求**.

定理 1.6.1 (Leontief定理) 设Leontief投入产出模型中的消耗矩阵 C 和最终需求向量 d 均非负, C 的每一列的和均小于1, 则 $I-C$ 可逆, 产出向量 $x=(I-C)^{-1}d$ 非负并且是方程 (1.6.1) 的唯一解.

证明见习题52. 由于 C 的每一列的和均小于1, 故矩阵 $I-C$ 实际上是一个M**-矩阵** (即形如 $sI-B$ 的矩阵, 其中 B 为非负矩阵而 $s \geq \rho(B)$), 因此其逆矩阵 $(I-C)^{-1}$ 是一个非负矩阵. 实际上, 矩阵 $(I-C)^{-1}$ 的第 j 列表示当第 j 个部门的最终需求增加1单位时, 各部门需要增加的产出的数量.

Leontief生产方程的对偶形式是**价格方程**

$$(\text{Leontief价格方程}) \qquad p = C^T p + v \tag{1.6.2}$$

其中 p 称为**价格向量**, 它表示各部门产出的单位价格, v 称为**增值向量**, 它表示每单位产出的附加值. 经济中最重要的指标国内生产总值(GDP) 实际上可以表示为

$$GDP = p^T d = v^T x. \tag{1.6.3}$$

上式中第二个等号的证明见习题53.

三. 秩1矩阵与随机变量的独立性

概率统计的一个基本概念是随机变量的独立性. 设 X 与 Y 是两个随机变量, 如果

$$p(XY) = p(X)p(Y)$$

则称 X 与 Y 是独立的. 设二维离散随机变量 (X,Y) 的联合分布矩阵为 $P=(p_{ij})_{m\times n}$, 关于 X 和 Y 的边缘分布分别为 $P_X=(p_{i\bullet})_{m\times 1}$ 与 $P_Y=(p_{\bullet j})_{1\times n}$. 则有下述定理

定理 1.6.2 离散随机变量 X 与 Y 相互独立 $\iff$ 对任意 i,j 均有 $p_{ij}=p_{i\bullet}p_{\bullet j}$ $\iff$ 二维随机变量 (X,Y) 的联合分布矩阵 $P=(p_{ij})$ 的秩为1. (证明见习题55).

常见的概率统计教材或著作中仅指出了第一个等价, 但显然第二个等价更为简洁有效.

例 1.6.2 (1999年考研数学一试题) 设随机变量 X 与 Y 独立, 填充下表

[12] Wassily Wassilyovich Leontief(1905-1999), 著名俄裔美籍经济学家, 1973年获得诺贝尔经济学奖. 他的三个博士研究生Paul Samuelson, Robert Solow与Vernon Lomax Smith分别于1970, 1987和2002 年获得诺贝尔经济学奖.

(X,Y)	y_1	y_2	y_3	$p_{i\bullet}$
x_1		$1/8$		
x_2	$1/8$			
$p_{\bullet j}$	$1/6$			1

解 因为 $p_{11} = 1/6 - 1/8 = 1/24$ 以及 X 与 Y 独立, 故由 定理1.6.2 可知所有的列 (与行) 均成比例, 因此第2列是第1列的3倍, 最后一列是第1列的6倍, 于是所有数据均被确定.

四、图与矩阵

一个 (无向) 图 G 可以用集合 $G=(V,E)$ 来定义, 其中 V 是 G 的顶点集, E 是 G 的边集. 如果 V 与 E 都是有限集 (绝大多数有意义的情形如此), 则可以用矩阵来表示图 G. 为此, 设 $V=\{v_1,\cdots,v_n\}$, 设 v_i 与 v_j 之间的边的数目为 a_{ij}, $1\le i,j\le n$, 则称矩阵 $A(G)=(a_{ij})_{n\times n}$ 为图 G 的**邻接矩阵**.

例 1.6.3 设 G 是下面的图:

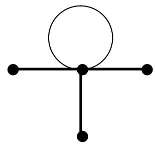

图 1.6.2

则其邻接矩阵为 (顶点从左至右, 从上到下编号)

$$A(G)=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

显然, 图的邻接矩阵是一个对称的非负整数矩阵, 而且每个图有唯一确定的邻接矩阵; 反过来, 每个对称的非负整数矩阵确定唯一一个图. 尚有其它多种图的矩阵表示法. 用矩阵表示图, 便于用代数方法与计算机研究处理较为复杂的图, 实际上研究图的现代理论即“代数图论”的相当部分的数学基础就是矩阵理论.

习 题 一

1. 计算:

$$(1)\ \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix}^n;\quad (2)\ \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}^n;\quad (3)\ \begin{pmatrix} a & 1 & & & \\ & a & 1 & & \\ & & a & 1 & \\ & & & a & 1 \\ & & & & a \end{pmatrix}^n.$$

2. 证明: 与任意 n 阶方阵可交换的矩阵必是纯量矩阵 λI.
3. 利用初等变换求 $A^{-1}B$ 及 CA^{-1}, 其中

$$A=\begin{pmatrix} 4 & 5 & 0 \\ 2 & 3 & 1 \\ 2 & 7 & -3 \end{pmatrix},\ B=\begin{pmatrix} 4 & 5 & 0 & 10 \\ 2 & 3 & 1 & -1 \\ 2 & 7 & 9 & -3 \end{pmatrix},\ C=\begin{pmatrix} 4 & 5 & 0 \\ 2 & 3 & 1 \\ 2 & 7 & 9 \\ -2 & 3 & 7 \end{pmatrix}.$$

4. 设 $A,B\in M_n$, 证明: $\operatorname{adj}(AB)=\operatorname{adj}(B)\operatorname{adj}(A)$.

5. 证明: 对任意矩阵 A, 有 $r(A^*A) = r(AA^*) = r(A)$.

6. 证明: 对任意 n 阶矩阵 A, 有 $r(A^n) = r(A^{n+1})$.

7. 设 ω 是 n 次本原单位根 (可设 $\omega = e^{2\pi i/n} = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$), 试求 Fourier [13]矩阵

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)(n-1)} \end{pmatrix}$$

的逆矩阵.

8. 设 $A = (a_{ij})_{n\times n}$ 是可逆的对称实矩阵. 证明: 二次型

$$f(x_1, x_2, \cdots, x_n) = \begin{vmatrix} 0 & x_1 & \cdots & x_n \\ -x_1 & a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ -x_n & a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

的矩阵是 A 的伴随矩阵.

9. 证明矩阵秩的 Frobenius[14] 不等式: $r(AB) + r(BC) \le r(B) + r(ABC)$.

10. 证明行初等变换不改变矩阵的列向量之间的线性关系.

11. 设 A 是 n 阶矩阵, 对任意 $0 \neq x \in \mathbb{F}^n$ 均有 $Ax \neq x$, 证明 $I - A$ 可逆并求其逆.

12. 设 n 阶矩阵 A 可逆, x 与 y 是 n 维列向量. 如果 $(A + xy^*)^{-1}$ 可逆, 证明 **Sherman-Morrison**[15] **公式**:

$$(A + xy^*)^{-1} = A^{-1} - \frac{A^{-1}xy^*A^{-1}}{1 + y^*A^{-1}x}.$$

(提示: 可用上题的结论.)

13. 设 n 阶矩阵 A 可逆, B, C, D 分别是 $n\times m, m\times n, m\times m$ 矩阵. 证明

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A||D - CA^{-1}B|.$$

14. (1) 设矩阵 A, C 均可逆, 求分块矩阵 $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ 的逆矩阵.

(2) 设矩阵 A 可逆, $D - CA^{-1}B$ 也可逆, 证明分块矩阵 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 可逆并求其逆.

15. 设矩阵 A 与 $A - BC$ 均可逆, 试用 A, A^{-1}, B, C 表示 $(A - BC)^{-1}$.(提示: 研究分块矩阵 $\begin{pmatrix} A & B \\ C & I \end{pmatrix}$ 的逆矩阵.)

16. 设 $\Omega = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$ 为 $2n$ 阶分块矩阵. 一个 $2n$ 阶复矩阵 M 称为是**辛矩阵** 如果 $M^T\Omega M = \Omega$. 证明:

(1) $2n$ 阶辛矩阵的全体构成一个群, 即辛矩阵的逆矩阵仍是辛矩阵, 两个辛矩阵的乘积仍是辛矩阵;

(2) 任何辛矩阵的行列式均为 1. (提示: 利用分块矩阵.)

---

[13]Joseph Fourier(1768-1830), 著名法国数学家与物理学家, 发现了三角级数, Fourier 变换, 热传导方程, 热传导定律和温室效应.

[14]Ferdinand Georg Frobenius(1849-1917), 德国著名数学家.

[15]Jack Sherman(1927-2007) 和 Winifred Morrison(1910-1961) 均为美国统计学家, 该公式发表于 1949 年.

17. 求下列各矩阵的满秩分解:

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

18. 证明第三种初等矩阵 (即 $I+aE_{ij}$, $i\neq j$, $a\neq 0$) 彼此相似. 又, 第一种初等矩阵是否彼此相似?

19. 设矩阵 A 满足方程 $A^2-A+2I=0$, 问 A 可以对角化吗? 为什么? 将本题一般化.

20. 证明:(1) Hermite 矩阵的特征值均为实数, 且属于不同特征值的特征向量彼此正交.
(2) Hermite 矩阵 A 是正定矩阵 $\Longleftrightarrow$ 存在可逆下三角矩阵 L 使得 $A=LL^*$.

21. 证明 例 1.4.2 中的 $(V,\oplus,\bullet)$ 是 $\mathbb{R}$ 上的线性空间.

22. 设 $V=\mathbb{C}\setminus\{-1\}$. 利用普通加法和普通乘法定义 V 上的加法“$\diamond$”如下:

$$a\diamond b=a+b+ab.$$

证明 $\diamond$ 满足线性空间的加法的全部条件. 进一步, 构造复数与 V 中向量的一个“数乘” $\heartsuit$, 使得 $(V,\diamond,\heartsuit)$ 是 $\mathbb{R}$ 上的线性空间. 请给出该线性空间的一组基.

23. 请将上题的集合 $V=\mathbb{C}\backslash\{-1\}$ 做一适当调整, 使其在加法“$\clubsuit$”下成为加群, 其中 “$a\clubsuit b=a+b+xab$”, x 是某固定的复数. 试设计一个与加法“$\clubsuit$”和谐的数乘运算“$\spadesuit$”, 使得 $(V,\clubsuit,\spadesuit)$ 构成复线性空间. 请给出该线性空间的一组基. 又, 将 $\mathbb{C}$ 换为 $\mathbb{R}$ 结果如何?

24. 设 $A=\{a_1,a_2,\cdots,a_n\}$ 是非空有限集合.
(1) 证明: $\dim_{\mathbb{F}}\mathbb{F}^A=n$;
(2) 求 $\mathbb{F}^A$ 的一组基;
(3) 描述函数空间 $\mathbb{F}^A$ 的结构并推广到 A 是无限集合的情形.

25. 证明线性空间的替换定理: 设 $J=\{\alpha_1,\alpha_2,\cdots,\alpha_s\}$ 与 $K=\{\beta_1,\beta_2,\cdots,\beta_t\}$ 是 n 维线性空间 V 的两个向量组, 其中 J 线性无关. 如果每个 $\alpha_j\in J$ 都可由 K 线性表示, 则 $s\leq t$; 且可将 K 中的某 s 个向量换成 $\alpha_1,\alpha_2,\cdots,\alpha_s$, 使得新的向量组生成的子空间与 K 生成的子空间相同.

26. 证明有限维线性空间的任意两个基所含向量的个数相同.

27. 证明过渡矩阵必是可逆矩阵.

28. 证明 1, $x-1$, $(x-1)^2,\cdots,(x-1)^n$ 是 $\mathbb{R}[x]_{n+1}$ 的一组基, 并求多项式 $f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$ 在该组基下的坐标.

29. 验证: 若 $(\alpha,\beta)_1$ 与 $(\alpha,\beta)_2$ 是欧氏空间 V 的两个不同的内积, 则 $(\alpha,\beta)=(\alpha,\beta)_1+(\alpha,\beta)_2$ 也是 V 的一个内积. 试创造一种新办法再构造 V 的一种内积.

30. 对 $x=(x_1,x_2)^T$, $y=(y_1,y_2)^T$, 规定

$$(x,y)=ax_1y_1+bx_1y_2+bx_2y_1+cx_2y_2.$$

证明 (x,y) 是 $\mathbb{R}^2$ 的内积 $\Longleftrightarrow$ $a>0,\ ac>b^2$.

31. 设 $V=\{a\cos t+b\sin t,\ 其中\, a,b\, 为任意实数\}$ 是实二维线性空间. 对任意 $f,g\in V$, 定义

$$(f,g)=f(0)g(0)+f(\frac{\pi}{2})g(\frac{\pi}{2}).$$

证明 (f,g) 是 V 上的内积, 并求 $h(t)=3\cos(t+7)+4\sin(t+9)$ 的长度.

32. 设欧氏空间 $\mathbb{R}[x]_2$ 中的内积为

$$(f,g)=\int_{-1}^{1}f(x)g(x)dx.$$

(1) 求基 $1, t, t^2$ 的度量矩阵;

(2) 用矩阵乘法形式计算 $f(x) = 1 - x + x^2$ 与 $g(x) = 1 - 4x - 5x^2$ 的内积.

33. 设 $a_i, 1 \le i \le n$ 是正实数, x_i, y_i 是任意实数, 证明或否定不等式

$$(\sum_{i=1}^{n} a_i x_i y_i)^2 \le (\sum_{i=1}^{n} a_i x_i^2)(\sum_{i=1}^{n} a_i y_i^2).$$

34. (1) 复数域 $\mathbb{C}$ 是实数域 $\mathbb{R}$ 上的 2 维线性空间. 是否存在 $\mathbb{C}$ 上的一个内积, 使得 i 与 $1+\mathrm{i}$ 成为 $\mathbb{C}$ 的一组标准正交基, 为什么?

(2) 试构造实线性空间 $\mathbb{R}^3$ 上的一个内积, 使得向量组 $e_1, e_1+e_2, e_1+e_2+e_3$ 是一组标准正交基. 问此时 e_2 与 e_3 的长度各是多少? 它们的夹角又是多少?

35. 试尽可能一般性地讨论习题 34 中的问题.

36. 证明 Cauchy-Schwarz 不等式与三角不等式.

37. 在欧氏空间 $\mathbb{R}^4$ 中, 求三个向量 $\alpha_1 = (1,0,1,1)^T$, $\alpha_2 = (2,1,0,-3)^T$ 和 $\alpha_3 = (1,-1,1,-1)^T$ 所生成的子空间的一个标准正交基.

38. 定义任意内积空间 V 中两个向量 α 与 β 的距离为

$$d(\alpha, \beta) = \|\alpha - \beta\|.$$

证明如上定义的函数 $d(\alpha, \beta)$ 确实定义了 V 上一个距离, 即满足下列三个条件:

(d1) 对称性: $d(\alpha, \beta) = d(\beta, \alpha)$;

(d2) 非负性: $d(\alpha, \beta) \ge 0$, 且等号成立 $\iff \alpha = \beta$;

(d3) 三角不等式: $d(\alpha, \beta) + d(\beta, \gamma) \ge d(\alpha, \gamma)$.

39. 设 2 维欧氏空间 V 的一组基为 α_1, α_2, 其度量矩阵为

$$A = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}.$$

试求 V 的一个标准正交基到 α_1, α_2 的过渡矩阵.

40. 设 n 维内积空间 V 的一个基为 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 该基的度量矩阵为 A. 设 $\alpha, \beta \in V$ 在该基下的坐标分别为 x 与 y.

(1) 证明 $(\alpha, \beta) = x^T A \bar{y}$. 特别, 当 V 为欧氏空间时, $(\alpha, \beta) = x^T A y$.

(2) 证明 (1) 中内积的矩阵乘法形式与选取的基无关.

41. 设 $V = M_n(\mathbb{R})$ 或 $M_n(\mathbb{C})$. 设 $A = (a_{ij}), B = (b_{ij}) \in V$.

(1) 证明 $(A, B) = \mathrm{tr}(AB^*)$ 是 V 的一个内积;

(2) 按 (1) 的内积, 矩阵 A 的长度是多少? 哪些是单位向量?

(3) 证明或否定: 基本矩阵 $E_{ij}, 1 \le i, j \le n$ 是 V 的一组标准正交基;

(4) 求 $M_2(\mathbb{R})$ 的一组由可逆矩阵构成的标准正交基.

42. 设线性空间 $V = \mathbb{R}^2$ 是欧氏空间 (未必是通常的欧氏空间). 设 $\alpha_1 = (1,1)^T$, $\alpha_2 = (1,-1)^T$ 与 $\beta_1 = (0,2)^T, \beta_2 = (6,12)^T$ 是 V 的两组基. 设诸 α_j 与 β_k 的内积分别为

$$(\alpha_1, \beta_1) = 1, (\alpha_1, \beta_2) = 15, (\alpha_2, \beta_1) = -1, (\alpha_2, \beta_2) = 3.$$

(1) 求两组基的度量矩阵;

(2) 求 V 的一个标准正交基.

43. 设 n 维欧氏空间 V 的一组基为 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 证明：存在正定矩阵 C, 使得由

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

确定的向量组 $\beta_1, \beta_2, \cdots, \beta_n$ 是 V 的一个标准正交基.

44. 设 A 是**反对称**实矩阵 (即 $A^T=-A$), 证明:
(1) A 的特征值为 0 或纯虚数;
(2) 设 $\alpha+\beta\mathrm{i}$ 是 A 的属于一个非零特征值的特征向量, 其中 α,β 均为实向量, 则 α 与 β 正交.

45. 设 A 是 Hermite 矩阵. 如果对任意向量 x 均有 $x^*Ax=0$, 则 $A=0$.

46. 设 $A=\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$. 定义 $\mathbb{R}^2$ 上的二元 (向量) 函数 $<x,y>$ 如下:

$$<x,y>=x^TAy.$$

此二元函数与普通内积的差别是什么? 以此二元函数为基础, 建立相应的长度, 角度等概念, 研究其中的正交与平行的定理.

47. 设 $f(x,y)$ 是 $\mathbb{C}^n$ 上的对称**双线性函数** (即 $f(x,y)=f(y,x)$ 且关于两个变元 x 与 y 均是线性的).
(1) 给出 $f(x,y)$ 的一般表达式;
(2) 证明方程 $f(x,x)=0$ 总有非零解;
(3) 设 $f(x,y)$ 非退化 (即若 $f(x,\alpha)=0,\forall x\in\mathbb{C}^n$, 则 $\alpha=0$), 证明存在线性无关的向量 α,β 使得 $f(\alpha,\beta)=f(\beta,\alpha)=0$ 且 $f(\alpha,\beta)=1$.

48. 设 x 是矩阵 A 的一个特征向量 x, 证明相应于 x 的特征值为 x^*Ax/x^*x(此商称为 **Rayleigh**[16] **商**, 是研究特征值的重要工具). 据此研究 n 元二次型 x^*Ax 与 A 的特征值的关系.

49. Vandermonde (范德蒙德)[17] 行列式对应的矩阵称为 Vandermonde 矩阵. 研究 Vandermonde 矩阵是否可以对角化.

50. 设 $I=I_2$, 试求整数矩阵方程 $X^2=\pm I$ 的所有解. 试一般地讨论方程 $X^n=I_n$ 的解 (这样的整数矩阵称为周期矩阵), 其中 n 为某自然数. (提示: 利用特征多项式.)

51. 完整给出 例 1.6.1 的解. (数值计算可以使用相关软件.)

52. 证明 定理 1.6.1.

53. 证明 (1.6.3) 中的第二个等式.

54. 设 (1.6.1) 中的 n 阶消耗矩阵 C 的列和均小于 1, 设 Δx 为满足最终需求 Δd 的产出向量.
(1) 证明: 如果最终需求由 d 变为 $d+\Delta d$, 则新的产出向量必须为 $x+\Delta x$;
(2) 设 $\Delta d=e_1$ 是第一个标准向量, 证明 Δx 恰好是矩阵 $(I-C)^{-1}$ 的第一列. 解释这个结果.

55. 设图 G 的邻接矩阵为

$$A(G)=\begin{pmatrix} 1 & 1 & 0 & 2 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 \end{pmatrix},$$

求图 G.

[16] Rayleigh 男爵, 全名 John William Strutt(1842-1919), 著名英国物理学家, 1904 年诺贝尔物理学奖得主.
[17] Alexandre-Théophile Vandermonde(1735-1796), 法国数学家, 音乐家和化学家.