第一章 基础知识

本章带有复习性质，目的是帮助读者整理一下学习矩阵理论的一些有用的定理与方法。有些定理当作已知，只讲其应用。要求读者仔细学习本章，为后几章学习打下良好的基础。

第一节 矩阵的乘法

矩阵的乘法是矩阵的重要运算之一。如果有 $AB=C$，除了要知道如何进行乘法运算之外，还需将重点放在矩阵 C 的行向量与列向量的结构上，这点在理论上与计算上都是很重要的。下面矩阵 A 是 $m\times n$ 型，用 $A^{(j)}$表示 A 的第 j 列，而 $A_{(i)}$表示 A 的第 i 行。先从一个简单的结果开始。

1.

$$\begin{bmatrix} a_{11} & a_{12}\cdots a_{1n} \\ \vdots & \\ a_{m1} & a_{m2}\cdots a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}=\begin{bmatrix} a_{11}x_1+\cdots+a_{1n}x_n \\ \vdots \\ a_{m1}x_1+\cdots+a_{mn}x_n \end{bmatrix}$$

$$=x_1\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}+\cdots+x_n\begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}=x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(n)}。$$

上面的结果表明：矩阵乘一个列向量，其结果是将这个矩阵的列向量进行线性组合，以后当看到乘积 Ax 时，便将它看成由 A的的列向量线性组合而成的向量。组合系数是向量 x 的各个分量 $x_1,\cdots,x_n$。同样，如果将一个行向量左乘矩阵，便得出下面结果

$$(y_1\cdots y_m)\begin{bmatrix} a_{11}\cdots a_{1n} \\ \vdots \\ a_{m1}\cdots a_{mn}\end{bmatrix}=y_1A_{(1)}+y_2A_{(2)}+\cdots+y_mA_{(m)}。$$

它表明是将 A 的各个行向量进行线性组合，而组合的系数是行向量 y 的各个分量 $y_1\cdots y_m$。

例 1 (i) $3\begin{bmatrix}1\\-1\\0\end{bmatrix}-2\begin{bmatrix}0\\6\\7\end{bmatrix}+4\begin{bmatrix}5\\-4\\-3\end{bmatrix}=\begin{bmatrix}23\\-31\\-26\end{bmatrix}$，

亦可写成

$$\begin{bmatrix}1 & 0 & 5\\-1 & 6 & -4\\0 & 7 & -3\end{bmatrix}\begin{bmatrix}3\\-2\\4\end{bmatrix}=\begin{bmatrix}23\\-31\\-26\end{bmatrix}。$$

(ii) $(-1\ \ 0\ \ 2)-7(1\ \ 1\ \ 1)+6(1\ \ 2\ \ 3)=(-2\ \ 5\ \ 13)$，

亦可写成

$$(1\ \ -7\ \ 6)\begin{bmatrix}-1 & 0 & 2\\1 & 1 & 1\\1 & 2 & 3\end{bmatrix}=(-2\ \ 5\ \ 13)。$$

2．$AB=C$ 中乘积 C 的行向量与列向量的结构。由矩阵的乘法可得：

$$C^{(j)}=AB^{(j)} \text{ 与 } C_{(i)}=A_{(i)}B。$$

换句话说，矩阵 C 的第 j 个列向量是由 A 的列向量线性组合而成，组合系数由 B 的第 j 个列向量的各个分量给出。C 的第 i 个行向量是由 B 的行向量线性组合而成，组合系数由 A 的第 i 个行向量的各个分量给出。特别，如果用 $e^{(j)}$ 表示第 j 个标准单位列向量，$e_{(i)}$表示第 i 个标准单位行向量，则 $Ae^{(j)}=A^{(j)}$，$e_{(i)}A=A_{(i)}$。知道了矩阵乘积的行向量与列向量的构成以后，有些矩阵的乘积就很容易得出。

例 2 (i) $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 0 & k & 0 \\ 1 & 0 & 0 \end{bmatrix}=\begin{bmatrix} a_{13} & ka_{12} & a_{11} \\ a_{23} & ka_{22} & a_{21} \\ a_{33} & ka_{32} & a_{31} \end{bmatrix}$。

(ii) $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 1 & b & 0 \\ 0 & 1 & -c \end{bmatrix}$

$=\begin{bmatrix} a_{12} & ba_{12}+a_{13} & -ca_{13} \\ a_{22} & ba_{22}+a_{23} & -ca_{23} \\ a_{32} & ba_{32}+a_{33} & -ca_{33} \end{bmatrix}$。

上面两个结果均可由乘积的列向量构成得出。

3．下面举例说明上面方法的一些应用。

例 3 上三角阵的乘积仍是上三角阵。

现有矩阵：

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix},\quad B=\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ 0 & b_{22} & b_{23} \\ 0 & 0 & b_{33} \end{bmatrix}。$$

$AB=C$，只要按 C 的列向量的构成即可得出结果。因为 $C^{(1)}$ 的第二，三个分量均为零，$C^{(2)}$ 第三个分量为零。

例 4 $AB=O$ 的意义。

由于 $AB^{(j)}=O$，说明 B 的各个列向量，是齐次线性方程组 $Ax=O$ 的解向量。或者有 $A_{(i)}B=O$，即 A 的各个行向量是齐次方程组 $B^Ty=O$ 的解向量。

例 5 (i) $Ax=b$ 有解的充要条件。

由方程的等式可知，如果方程有解，表明向量 b 是矩阵 A 的各个列向量的线性组合。反之，如果 b 是矩阵 A 的各个列向量的线性组合，则可知方程一定有解。因此，方程有解的充要条件是：b 是 $A^{(1)}A^{(2)}\cdots A^{(n)}$ 的线性组合。

(ii) $Ax=O$ 有非零解与唯一解的充要条件。

由等式 $Ax=O$ 说明要将矩阵 A 的各个列向量线性组合为零

向量。因此，如果 A 的各个列向量是相关的，则一定有非零解。如果矩阵 A 的列向量是独立的，则它仅可有零解。显然这些都是充要条件。

例 6 矩阵 A 有左逆或右逆的条件。

该矩阵 A 是 $m\times n$ 型矩阵，如果存在一 $n\times m$ 型矩阵 B 有 $AB=I_m$（m 阶单位阵），则称 B 是 A 的右逆。现在来看一下 A 存在右逆的条件是什么？

$AB=I_m$ 说明 I_m 的各个列向量都是 A 的列向量的线性组合。由于 I_m 的秩为 m，因而矩阵 A 的秩至少是 m，但 A 的秩至多是 m。所以如果 A 有右逆，则矩阵 A 的秩（记为 $r(A)$ 或 r）为 m 或者说 A 是行满秩的。反之，如果 $r(A)=m$，说明 A 的独立的列向量有 m 个。所以 I_m 中任一个列向量都可通过 A 的列向量的线性组合来表示，即 $AB=I_m$。由此证明了下面的命题。

命题 1 矩阵 A 有右逆 $\Longleftrightarrow r(A)=m$。

类似可以证明，矩阵 A 有左逆的命题如下：

命题 2 矩阵 A 有左逆 $\Longleftrightarrow r(A)=n$。

上述关于矩阵乘积的行向量与列向量的结构的结论，还可举出很多应用。请读者在学习矩阵理论时留意。

第二节 矩阵的初等变换

1. 关于矩阵的初等变换的重要结论

矩阵的初等变换实际上是线性方程组消去法运算的一种抽象，它包括对一组向量施行下面三种运算。

(i) 交换二个向量。

(ii) 用一个非零的数乘一个向量。

(iii) 用一个数乘一个向量与另一个向量相加。

在矩阵理论中很重要的一点，是将矩阵施行初等变换与矩阵乘初等矩阵联系起来。关于这方面重要的结论有下面四个。

命题 3 矩阵左乘一个初等矩阵，相当于对矩阵施行行初等变换。

对矩阵右乘一个初等矩阵相当于对矩阵施行列初等变换。

命题 4 施行初等变换，不改变矩阵的秩。

定理 1 （初等变换的标准型）

任何一个秩为 r 的 $m\times n$ 型矩阵 A，总存在有非奇异矩阵 P、Q 使得

$$PAQ=\begin{bmatrix} I_r & \\ & O \end{bmatrix}_{m\times n}。\tag{1-1}$$

命题 5 A 是非奇异矩阵$\Longleftrightarrow A$ 是初等矩阵的乘积。

2．初等矩阵

对单位矩阵施行一次行或列的初等变换后所得的矩阵称为初等矩阵，下面是初等矩阵的例子（以三阶矩阵为例）。

$$E_{12}=\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},\quad E_3(k)=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{bmatrix}, E_{23}(k)=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix}$$

E_{12} 表示单位阵交换第一、二两行后所得的矩阵（交换第一、二两列得同样的矩阵）。$E_3(k)$ 表示用一不为零的常数 k 乘单位阵的第三行（列）所得的矩阵。$E_{23}(k)$ 表示用 k 乘单位阵的第三行加到第二行后所得的矩阵（亦为用 k 乘单位阵的第二列加到第三列）。如果 A 是一 3 阶方阵，这样由前面所讨论矩阵的乘法容易知道：

$E_{12}A, E_3(k)A, E_{23}(k)A$ 的结果恰为将原施于单位阵的行初等变换施于矩阵 A 上。同样，AE_{12}，$AE_3(k)$，$AE_{23}(k)$，便是将原施于单位阵的列变换施于矩阵 A 上。

3．初等变换的标准型的应用

这一小节主要通过实例说明初等变换的标准型的应用。另一方面这些实例在矩阵理论里也是一些重要的结果。

例 7 矩阵的满秩分解定理。

设 A 是 $m\times n$ 型矩阵，$r(A)=r$，则一定存在一个 $m\times r$ 型矩阵 L，与一个 $r\times n$ 型矩阵 R，有 $r(L)=r(R)=r$，$A=LR$。

证 由初等变换的标准型可以知道：

$$A=\underset{m\times m}{P}\begin{bmatrix} I_r & \\ & O \end{bmatrix}_{m\times n}\underset{n\times n}{Q}=\underset{m\times m}{P}\underset{m\times r}{\begin{bmatrix} I_r \\ O \end{bmatrix}}\underset{r\times n}{[I_r O]}\underset{n\times n}{Q}=LR。$$

由证明的过程立即可以知道 L 是由 P 的前 r 列构成，而 R 是由 Q 的前 r 行构成。

例 8 矩阵满秩分解的数值例。

现有矩阵：

$$A=\begin{bmatrix} 1 & 2 & 1 & 5 \\ 2 & 5 & 1 & 14 \\ 4 & 9 & 3 & 24 \end{bmatrix},\quad P=\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -2 & -1 & 1 \end{bmatrix},$$

$$Q=\begin{bmatrix} 1 & -2 & -3 & 3 \\ 0 & 1 & 1 & -4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}。$$

则

$$PAQ=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}。$$

所以

$$A=P^{-1}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}Q^{-1}$$

$$=\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 & 5 \\ 0 & 1 & -1 & 4 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$=\begin{bmatrix}1&0&0\\2&1&0\\4&1&1\end{bmatrix}\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\end{bmatrix}\begin{bmatrix}1&2&1&5\\0&1&-1&4\\0&0&1&-1\\0&0&0&1\end{bmatrix}$$

$$=\begin{bmatrix}1&0\\2&1\\4&1\end{bmatrix}\begin{bmatrix}1&2&1&5\\0&1&-1&4\end{bmatrix}。$$

一般而言，一个矩阵的满秩分解，并不是唯一的。如矩阵A还可写成：

$$A=\begin{bmatrix}1&0\\0&1\\2&1\end{bmatrix}\begin{bmatrix}1&2&1&5\\2&5&1&14\end{bmatrix}。$$

例 9 (Sylvester) 设A是$m\times n$型矩阵，而B是$n\times m$型矩阵($m\geqslant n$)。用f_{AB}与f_{BA}分别记AB与BA的特征多项式，则有$f_{AB}=\lambda^{m-n}f_{BA}$。换句话说，矩阵$AB$与$BA$的非零特征值是一样的，而所差仅是零特征值的重数。

证 设$r_A=r$，因而有$PAQ=\begin{bmatrix}I_r&\\&O\end{bmatrix}_{m\times n}$。

$$PABP^{-1}=PAQQ^{-1}BP^{-1}=\begin{bmatrix}I_r&\\&O\end{bmatrix}_{m\times n}C,$$

这里$C=Q^{-1}BP^{-1}$。

同理可知

$$Q^{-1}BAQ=Q^{-1}BP^{-1}PAQ=C\begin{bmatrix}I_r&\\&O\end{bmatrix}_{m\times n}。$$

因此只要证明矩阵$\begin{bmatrix}I_r&\\&O\end{bmatrix}C$与矩阵$C\begin{bmatrix}I_r&\\&O\end{bmatrix}$的特征多项式满足定理的等式即可。

$$f_{AB}=|\lambda I_m-AB|=\left|\lambda I-\begin{bmatrix} I_r & \\ & O\end{bmatrix}C\right|$$

$$=\begin{vmatrix} \lambda-C_{11} & \cdots & C_{1r} & \cdots & C_{1n} \\ \vdots & & & & \\ -C_{r1} & \cdots & \lambda-C_{rr} & \cdots & C_{rn} \\ & & \lambda & & \\ & & & \ddots & \\ & & & & \lambda \end{vmatrix}=\lambda^{m-r}|\lambda I_r-C_r|。$$

这里 C_r 表示矩阵 C 的首 r 行 r 列构成的子阵。同样的方法，可计算得

$$f_{BA}=|\lambda I_n-BA|=\lambda^{n-r}|\lambda I_r-C_r|。$$

比较两个等式，可知定理成立。

这个例子所讲的结果，亦称为特征多项式的降阶计算公式。它对于计算秩小的矩阵是很方便的。

例 10 设 A 是 n 阶方阵，$r(A)=1$。试求它的特征多项式。

解 由于 A 的秩为 1，因而可设 A 有满秩分解如下：

$$A=xy^T=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n\end{bmatrix}[y_1 \quad y_2\cdots y_n]。$$

又由例 9 的结论可知：

$$|\lambda I_n-A|=\left|\lambda I_n-\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n\end{bmatrix}[y_1 \ \ y_2\cdots y_n]\right|$$

$$=\lambda^{n-1}\left|\lambda-[y_1 \ \ y_2\cdots y_n]\begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n\end{bmatrix}\right|$$

$$=\lambda^{n-1}(\lambda-\sum x_iy_i)。$$

下面讨论标准型在估计矩阵秩方面的应用。现先将一些有关矩阵的秩的一些结论罗列如下。

命题 6 A 是 $m\times n$ 型矩阵，且 $AB=O$，则 $r(B)\leqslant n-r(A)$。

命题 7 A 是 $m\times n$ 型矩阵，任取 s 行构成子阵 B，则 $r(B)\geqslant r(A)+s-m$。

命题 8 $r(AB)\leqslant\min\{r(A)\ r(B)\}$。

命题 9 设 A 是 $m\times n$ 型矩阵，B 是 $n\times s$ 型矩阵，则 $r(AB)\geqslant r(A)+r(B)-n$。

上面关于矩阵的秩的命题，可借助于定理 1 来证明，这里仅举二个命题的证明为例。

例 11 A 是 $m\times n$ 型矩阵，任取 s 行构成子阵 B，则 $r(B)\geqslant r(A)+s-m$。

证 因 $A=P\begin{bmatrix} I_r & \\ & O\end{bmatrix}Q$，而 B 是 A 的子阵，所以在求 A 的标准型的同时对 B 亦施行初等变换，还可得出 B 的标准型。所以 B 的标准型中的零行的数目，不会超过 A 的标准型中零行的数目，即 $s-r(B)\leqslant m-r(A)$。

例 12 A 是 $m\times n$ 型矩阵且 $AB=O$，则 $r(B)\leqslant n-r(A)$。

证 设 $r(A)=r$，由标准型可知：

$$A=P\begin{bmatrix} I_r & \\ & O\end{bmatrix}Q。$$

又由 $AB=O$，可得

$$P\begin{bmatrix} I_r & \\ & O\end{bmatrix}QB=O。$$

由此可知矩阵 QB 的前 r 行一定都为零行，所以 QB 的非零行至多是 $n-r$ 行，即 $r(QB)\leqslant n-r$，而 $r(QB)=r(B)$，故有 $r(B)\leqslant n-r(A)$。

其余两个命题的证明希望读者自己完成。

4. Hermite 标准型

(1) Hermite 标准型的定义

前面所讲的标准型是对矩阵施行行与列的初等变换后所获得的。这里所讨论矩阵的另一种标准型,它是对矩阵仅施行行初等变换,(这变换相当于利用矩阵格式解线性方程组)。这种标准型同样在计算上有很广泛的用途。现在先给出定义如下。

定义 一个 $m\times n$ 型矩阵。如果满足下面的条件便称它为 Hermite 梯形阵。

(i) 它的首 r 行是非零行，且每一个非零行的第一个非零元是 1。

(ii) 设第 i 行的第一个非零元出现在 k_i 列上，则应有 $k_1 < k_2 < \cdots < k_r$。

(iii) 在第 k_i 列上除第 i 行外,其余的元均为零。

下面便是 Hermite 梯形阵

$$\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}。$$

任何一个 $m\times n$ 矩阵 A，都可通过施行行变换的方法，变为 Hermite 梯形阵。

例 13 将下面矩阵 A 变为 Hermite 梯形阵，并将初等变换矩阵记录下来。

$$A=\begin{bmatrix} 0 & 1 & 2 & 2 & 3 & 4 \\ 0 & 2 & 4 & -2 & 0 & -4 \\ 0 & 3 & 6 & 0 & 3 & 0 \end{bmatrix}_{3\times 6}。$$

解 为了要记录变换矩阵,要采用如下的格式。即在 A 的右边放一个 3 阶单位阵 $[A|E]$，对整个矩阵施行行变换。只要将 A 变成 Hermite 梯形阵时，在 E 块便会将变换矩阵记录下来。其理由如下：对分块矩阵 $[A|E]$ 施行行初等变换相当于左乘一

非奇异矩阵 Q，即 $Q[A\,|\,E]=[QA\,|\,QE]$，因而当 $QA=H$，E 块处便指出变换矩阵 Q。其变换过程如下：

$$
\left[\begin{array}{cccccc|ccc}
0 & 1 & 2 & 2 & 3 & 4 & 1 & 0 & 0\\
0 & 2 & 4 & -2 & 0 & -4 & 0 & 1 & 0\\
0 & 3 & 6 & 0 & 3 & 0 & 0 & 0 & 1
\end{array}\right]\longrightarrow
$$

$$
\left[\begin{array}{cccccc|ccc}
0 & 1 & 2 & 2 & 3 & 4 & 1 & 0 & 0\\
0 & 0 & 0 & -6 & -6 & -12 & -2 & 1 & 0\\
0 & 0 & 0 & -6 & -6 & -12 & -3 & 0 & 1
\end{array}\right]\longrightarrow
$$

$$
\left[\begin{array}{cccccc|ccc}
0 & 1 & 2 & 2 & 3 & 4 & 1 & 0 & 0\\
0 & 0 & 0 & 1 & 1 & 2 & \frac{1}{3} & -\frac{1}{6} & 0\\
0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1
\end{array}\right]\longrightarrow
$$

$$
\left[\begin{array}{cccccc|ccc}
0 & 1 & 2 & 0 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & 0\\
0 & 0 & 0 & 1 & 1 & 2 & \frac{1}{3} & -\frac{1}{6} & 0\\
0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1
\end{array}\right]
$$

将 A 的 Hermite 梯形阵记为 H_A。上面的例子便是

$$
H_A=\begin{bmatrix}
0 & 1 & 2 & 0 & 1 & 0\\
0 & 0 & 0 & 1 & 1 & 2\\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix},\quad
Q=\begin{bmatrix}
\frac{1}{3} & \frac{1}{3} & 0\\
\frac{1}{3} & -\frac{1}{6} & 0\\
-1 & -1 & 1
\end{bmatrix}.
$$

（2）Hermite 标准型的应用

Hermite 标准型有很多应用，在谈它的应用之前先介绍一个关于 Hermite 标准型的重要定理，再证明一个关于行初等变换的一个重要性质。

定理 2 如果 A 通过行变换变为矩阵 B，则它们的 Hermite 标准型是相同的，即 $H_A=H_B$。

（证略）

命题 10 对一矩阵施行行初等变换，并不改变矩阵列的线性组合关系。

证 设 $QA=B$，因而 $QA^{(i)}=B^{(i)}$。

如果矩阵 A 的某些列向量有某个线性组合的关系

$$C_1A^{(j_1)}+C_2A^{(j_2)}\cdots+C_kA^{(j_k)}=0。$$

便可得

$$C_1QA^{(j_1)}+C_2QA^{(j_2)}+\cdots+C_kQA^{(j_k)}=0。$$

即

$$C_1B^{(j_1)}+C_2B^{(j_2)}+\cdots+C_kB^{(j_k)}=0。$$

由于 Q 是非奇异矩阵，所以同理可得：如果在矩阵 B 的某些列向量，有某个线性组合的关系，则可推知 A 中相对应的列向量亦有相同的线性组合关系。

例 14 $A=\begin{bmatrix} 0 & 1 & 2 & 2 & 3 & 4 \\ 0 & 2 & 4 & -2 & 0 & -4 \\ 0 & 3 & 6 & 0 & 3 & 0 \end{bmatrix}$，试指出矩阵 A 的最大独立列向量组，并将其余的列向量用它们表示出来。

解 因为矩阵 A 通过行初等变换变为 H_A（见例 13），所以它们的列向量有相同的线性组合关系。如在H_A中第二列与第四列是独立的，因而在 A 中 $A^{(2)}=(1\ \ 2\ \ 3)^T$，$A^{(4)}=(2\ \ -2\ \ 0)^T$ 也是独立的。在 H_A 中各个列向量的线性组合关系是容易看出，如 $H_A^{(3)}=2H_A^{(2)}$，$H_A^{(6)}=2H_A^{(4)}$，$H_A^{(5)}=H_A^{(2)}+H_A^{(4)}$。因而在矩阵 A 中亦有 $A^{(3)}=2A^{(2)}$，$A^{(6)}=2A^{(4)}$，$A^{(5)}=A^{(2)}+A^{(4)}$。

例 15 $A=\begin{bmatrix} 0 & 1 & 2 & 2 & 3 & 4 \\ 0 & 2 & 4 & -2 & 0 & -4 \\ 0 & 3 & 6 & 0 & 3 & 0 \end{bmatrix}$，写出矩阵 A 的满秩分解。

解 设 $A=LR$。由矩阵的乘法可知 A 的各个列向量可以由矩阵 L 的列向量线性组合得出。而组合系数，在矩阵 R 中相应的

列向量的分量给出。由例 14 立即可知 L 的列向量可取矩阵 A 的一组最大独立列向量组如 $A^{(2)}$ $A^{(4)}$,而组合系数的矩阵,可取H_A 中非零行即可。故有

$$L=\begin{bmatrix}1 & 2\\ 2 & -2\\ 3 & 0\end{bmatrix},\ R=\begin{bmatrix}0&1&2&0&1&0\\0&0&0&1&1&2\end{bmatrix}。$$

Hermite 标准型的应用,在后面讲到子空间计算的时候,还要叙述到。

5. 分块矩阵的初等变换

在实际计算中,二阶分块矩阵的初等变换是用得最多的,它是二阶矩阵初等变换的直接推广。

现有矩阵 $A=\begin{bmatrix}a & b\\ c & d\end{bmatrix}$,若 $a\neq 0$ 则可通过行初等变换,使位置(2,1)的元变为零。这只要 A 左乘初等矩阵$\begin{bmatrix}1 & 0\\ ca^{-1} & 1\end{bmatrix}$即可。即

$$\begin{bmatrix}1 & 0\\ ca^{-1} & 1\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix}=\begin{bmatrix}a & b\\ 0 & d-ca^{-1}b\end{bmatrix}。$$

如果要计算行列式,在等式的两边取行列式,即有$\begin{vmatrix}a & b\\ c & d\end{vmatrix}=a(d-ca^{-1}b)$。

对于一般的二阶分块阵亦有类似的结果。

如二阶分块阵$\begin{bmatrix}A & B\\ C & D\end{bmatrix}$,其中$|A|\neq 0$,同样可用广义的初等矩阵$\begin{bmatrix}I & 0\\ CA^{-1} & I\end{bmatrix}$左乘上述矩阵。即

$$\begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}=\begin{bmatrix} A & B \\ 0 & D-CA^{-1}B \end{bmatrix}.$$

如果要计算行列式，便有

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix}=|A||D-CAB|。$$

这是行列式的降阶计算公式。

例 16 若 A 是正定阵，则 $|A|\leqslant a_{11}a_{22}\cdots a_{nn}$。

证 设 $A=\begin{bmatrix} A_{n-1} & \alpha \\ \alpha' & a_{nn} \end{bmatrix}$，由于 A 是正定阵，所以 A_{n-1} 亦是正定阵。同样 A_{n-1}^{-1} 也是正定阵。利用分块矩阵的初等变换将位置在(2,1)的块变成零。即

$$\begin{bmatrix} I & 0 \\ -\alpha' A_{n-1}^{-1} & 1 \end{bmatrix}\begin{bmatrix} A_{n-1} & \alpha \\ \alpha' & a_{nn} \end{bmatrix}=\begin{bmatrix} A_{n-1} & \alpha \\ 0 & a_{nn}-\alpha' A_{n-1}^{-1}\alpha \end{bmatrix}。$$

两边取行列式，便有

$$|A|=|A_{n-1}|(a_{nn}-\alpha' A_{n-1}^{-1}\alpha)。$$

但 $|A|$、$|A_{n-1}|$、a_{nn}、$\alpha' A_{n-1}^{-1}\alpha$ 均大于零，因此有

$$|A|\leqslant|A_{n-1}|a_{nn}。$$

同理 $|A_{n-1}|\leqslant|A_{n-2}|a_{n-1n-1}$。按此递推，可证明原命题正确。

第三节 线性空间与线性子空间

线性空间的概念大家已熟悉了。但是为方便起见这里还是把完整的定义写出来。

1. 线性空间的定义

一数域 F（通常指实数域或复数域）上的线性空间 V，是指一非空集合 V，它的元称为向量，并且在 V 上定义向量的加法以及数乘二种运算，这些运算应满足下面的条件。

A_1　对每一对向量 $\alpha,\beta\in V$，对应 V 中唯一的向量，称为它们的和，记为 $\alpha+\beta$。

A_2　加法是可结合的，即 $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$。

A_3　存在零向量，记为 0。它满足等式 $\alpha+0=\alpha,\forall\alpha\in V$。

A_4　对每一个 $\alpha\in V$，存在一向量，记为 $-\alpha$，它满足等式 $\alpha+(-\alpha)=0$。

A_5　加法是可交换的，即 $\alpha+\beta=\beta+\alpha$。

上面是向量加法运算的规定，下面是数乘运算的规定。

B_1　对每一个数 $a\in F$，以及向量 $\alpha\in V$，存在唯一的向量，称为数 a 与向量 α 的乘积（亦简称为数乘），记为 $a\alpha$。

B_2　数乘是可结合的，即 $a(b\alpha)=(ab)\alpha,\forall a,b\in F$。

B_3　数乘关于向量加法的分配律，即 $a(\alpha+\beta)=a\alpha+a\beta$。

B_4　数乘关于数量加法的分配律，即 $(a+b)\alpha=a\alpha+b\alpha$。

B_5　$1\cdot\alpha=\alpha,(1\in F)$。

矩阵理论中最重要的线性空间是下面二类。

(1) 数域 F 的元构成的 n 维向量全体，具体来讲是

$R^n=\{(x_1\cdots x_n)^T|x_i\in R\}$，　n 维实向量空间；

$C^n=\{(x_1\cdots x_n)^T|x_i\in C\}$，　n 维复向量空间。

在没有必要区分是复数或实数的时候用 F^n 表示。

(2) 数域 F 的元构成的 $m\times n$ 型矩阵全体，具体来讲是

$$R^{m\times n}=\{A=[a_{ij}]_{m\times n}|a_{ij}\in R\},$$

$$C^{m\times n}=\{A=[a_{ij}]_{m\times n}|a_{ij}\in C\}。$$

在没有必要区分是复数还是实数的时候用 $F^{m\times n}$ 表示。

2．线性子空间

在研究任何一种系统的时候，一个重要的方法便是研究它的各种子系统，以及探讨如何通过由较简单的子系统去构造整个系统。这是一种分析的方法，它是研究系统的一种常用方法。在线性空间 V 里，子系统便是子空间。

(1) 线性子空间的定义

V_s是V的一个子集，如果按V的加法与数乘运算，V_s构成一个线性空间，则称V_s是V的一个线性子空间，或简称为子空间。

如何判定线性空间V的一个子集V_s是子空间呢？下面的定理便说明了具体的方法。

定理 3 V_s是线性空间V的一个子集，它是子空间

$$\Longleftrightarrow a\alpha+b\beta\in V_s,\ \forall a,b\in F,\ \forall\alpha,\beta\in V_s。$$

（证略）。

换句话说，要确定线性空间V的一个子集V_s，是否是子空间，只要验证V_s在向量加法与数乘这两个运算的闭性（即验证运算的结果是否仍在集合V_s之内）。下面便列举一些重要的子空间的例子。

例 17 给出齐次线性方程组

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=0 \\ \vdots \qquad\qquad \vdots \\ a_{m1}x_1+\cdots+a_{mn}x_n=0 \end{cases} \qquad 或记为\ Ax=0,$$

它的解集S，是F^n的一个子空间，亦称为矩阵A的零空间，并用$N(A)$记之。

例 18 α，β是线性空间V的某二个固定的向量。构造集合V_s为

$$V_s=\{a\alpha+b\beta\,|\,a,b\in F\}。$$

证 由前面的定理可判定V_s是V的一个子空间，称它为由向量α,β生成的子空间，并记为$[\alpha,\beta]$。

一般有$\{\alpha_1\cdots\alpha_s\}$是一个向量组。考察由它们的线性组合全体构成的集合

$$\{a_1\alpha_1+\cdots+a_s\alpha_s\,|\,a_1\cdots a_s\in F\}$$

它亦构成一个子空间，称它为由向量组$\{\alpha_1\ \alpha_2\cdots\alpha_s\}$生成的子空间，并记为$[\alpha_1\cdots\alpha_s]$。

例 19 给出一个 2×3 矩阵 $A=\begin{bmatrix} 0 & 1 & 2 \\ -1 & 1 & 3 \end{bmatrix}$，它的三个列向量生成的子空间为

$$\left\{a_1\begin{bmatrix} 0 \\ -1 \end{bmatrix}+a_2\begin{bmatrix} 1 \\ 1 \end{bmatrix}+a_3\begin{bmatrix} 2 \\ 3 \end{bmatrix} \middle| a_i\in R\right\},$$

称为矩阵 A 的列空间，用记号 $R(A)$ 表示。由矩阵的乘法可知 $R(A)$ 中任一个向量 α 可写成 $\begin{bmatrix} 0 & 1 & 2 \\ -1 & 1 & 3 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$。一般情形也是这样，以后看到矩阵乘积 Ax，便知道它是属于矩阵 A 的列空间的向量。

一般给出一个实矩阵 A，它是 $m\times n$ 型，与 A 相联系有四个重要的子空间。

$R(A)$ ——A 的列空间，

$N(A)$ ——A 的零空间，

$R(A^T)$ ——A 的行空间，

$N(A^T)$ ——A 的左零空间。

易知 $R(A)$，$N(A^T)$ 是 R^m 的子空间，而 $R(A^T)$，$N(A)$ 是 R^n 的子空间，以后详细讨论它们的关系。

(2) 子空间的和与直和

设 V_1，V_2 是线性空间 V 的两个子空间，容易验证它们的交 $V_1\cap V_2$ 是个子空间。但对于它们的并集 $V_1\cup V_2$，一般而言不是子空间。因为 $V_1\cup V_2$ 不会满足运算的闭性。例如，在 R^2 中考察由向量 u 生成的子空间 V_1，和向量 v 生成的子空间 V_2（见图 1-1）。集合 $V_1\cup V_2$ 显然不构成子空间。在实际中，

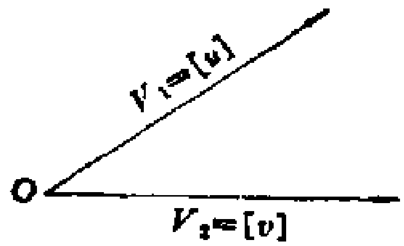

图 1-1

考察由 $V_1 \cup V_2$ 的元所生成的子空间，即考察 $[V_1 \cup V_2]$，并将这个子空间称为子空间 V_1, V_2 的和。下面将定义写出。

定义 V_1, V_2 是线性空间 V 的二个子空间，集合

$$\{a\alpha + b\beta \mid \alpha \in V_1\ \beta \in V_2, a, b \in F\}$$

称为子空间 V_1, V_2 的和，并用记号 $V_1 + V_2$ 表示。

例 20 求由向量$\{\alpha_1, \alpha_2\}$生成的子空间，与由向量$\{\beta_1\ \beta_2\}$生成的子空间的和，并指出它的维数与基。其中

$$\alpha_1 = (1,2,1,0)^T, \beta_1 = (2,-1,0,1)^T,$$

$$\alpha_2 = (-1,1,1,1)^T, \beta_2 = (1,-1,3,7)^T。$$

由子空间和的定义可知$[\alpha_1, \alpha_2] + [\beta_1, \beta_2] = [\alpha_1, \alpha_2, \beta_1, \beta_2]$。因此只要求出向量组$\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$的极大无关组即可，只要将下面矩阵施行行初等变换并将它变为 Hermite 梯形阵即可。

$$\begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & 1 & -1 & 1 \\ 1 & 1 & 0 & 3 \\ 0 & 1 & 1 & 7 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 3 & -5 & -3 \\ 0 & 2 & -2 & 2 \\ 0 & 1 & 1 & 7 \end{bmatrix} \to$$

$$\begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 3 & -5 & -3 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & 1 & 7 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 0 & -2 & -6 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 2 & 6 \end{bmatrix} \to$$

$$\begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}。$$

由此可知向量组 $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$中 $\alpha_1, \alpha_2, \beta_1$ 是独立的，而 $\beta_2 = -\alpha_1 + 4\alpha_2 + 3\beta_1$。因此$[\alpha_1, \alpha_2] + [\beta_1, \beta_2] = [\alpha_1, \alpha_2, \beta_1]$。所以该子空间是三维的，基为 $\alpha_1 \alpha_2 \beta_1$。

在研究子空间的和的时候，最重要的一种情形便是直和。

定义 V_1, V_2 是 V 的两个子空间，如果子空间 $W = V_1 + V_2$

的任意一个向量，可以唯一地写成 $\alpha=\alpha_1+\alpha_2$，其中 $\alpha_i\in V_i$。则称 W 是 V_1,V_2 的直和，记号 $W=V_1\oplus V_2$。

例 21 在 R^3 中，过原点的平面 π，与不在 π 上过原点的直线 L 都是 R^3 的子空间。它们的和是直和并且有

$$R^3=\pi\oplus L。$$

例 22 设 $V_1=\{(x_1\ \ x_2\ \ 0)\,|\,x_1,x_2\in R\}$，$V_2=\{(0\ \ y_2\ \ y_3)\,|\,y_2,y_3\in R\}$。

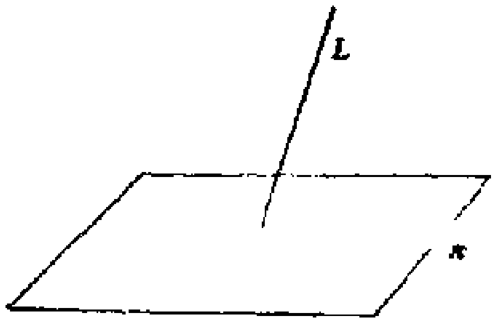

图 1-2

易知 V_1,V_2 都是 R^3 的二维子空间。如果取 $e_1=(1,0,0)$，$e_2=(0,1,0)$为 V_1 的基，$e_2=(0,1,0)$，$e_3=(0,0,1)$为 V_2 的基。这样 V_1+V_2 实际上是 R^3，可是 R^3 的任一个向量在这两个子空间上的分解式并不唯一。例如：

$$(3,7,5)=(3,2,0)+(0,5,5)=(3,-1,0)+(0,8,5),$$

因而 V_1,V_2 的和不是直和。

下面给出判定两个子空间的和为直和的充要条件。

定理 4 V_1+V_2 是直和 $\Longleftrightarrow V_1\cap V_2=\{0\}$

证 $\Longrightarrow$ 因为 V_1+V_2 是直和，所以对任何一个 $\alpha\in V_1+V_2$ 有唯一分解式 $\alpha=\alpha_1+\alpha_2$，$\alpha_i\in V_i$。如果 $V_1\cap V_2$ 中有 $\beta\neq0$ 的向量，这样，α 可有另一分解式 $\alpha=(\alpha_1+\beta)+(\alpha_2-\beta)$，这与 V_1+V_2 是直和矛盾。

$\Longleftarrow$ 设对 $\alpha\in V_1+V_2$ 有两种分解式：$\alpha=\alpha_1+\alpha_2$，$\alpha=\beta_1+\beta_2$，由此可以得出

$$0=(\alpha_1-\beta_1)+(\alpha_2-\beta_2)$$

因此

$$(\alpha_1-\beta_1)=(\beta_2-\alpha_2)。$$

等式两边分属两个子空间，所以 $(\alpha_1-\beta_1),(\beta_2-\alpha_2)\in V_1\cap V_2$。由前提可以知，它们都是零向量，于是 $\alpha_1=\beta_1$，$\alpha_2=\beta_2$。这说明 α 的

分解式是唯一的。☑

命题 11 子空间 V_1, V_2 的和是直和$\Longleftrightarrow\alpha_1+\alpha_2=0$，可得 $\alpha_1=0,\alpha_2=0$

这命题请读者自证。

3. 线性空间扩充基的方法的运用

现已经知道线性空间都有基存在，在矩阵理论里的线性空间是有限维的。该空间存在有限个独立向量构成的向量组 $\{\alpha_1\cdots\alpha_n\}$，使得空间的任一个向量，都可通过这组向量的线性组合唯一表示出来。关于向量组的独立性与基，下面的定理与命题是基本的。

命题 12 $\{\alpha_1\cdots\alpha_s\}$中的向量全不为零向量，则$\{\alpha_1\cdots\alpha_s\}$是相关向量组$\Longleftrightarrow\exists$ 向量 α_t，有 $\alpha_t=\sum\limits_{i<t}a_i\alpha_i$。

由这个命题可以得出下面的定理和命题。

定理 5 $\{\alpha_1\cdots\alpha_s\}$与$\{\beta_1\cdots\beta_t\}$是二个等价的无关向量组(即任一 α_i 可通过向量组$\{\beta_j\}$表示，而任一 β_j 亦可通过向量组 $\{\alpha_i\}$ 表示)，则 $s=t$。

证 因为 β_1 可通过$\{\alpha_1\cdots\alpha_s\}$表示，所以$\{\beta_1\ \alpha_1\cdots\alpha_s\}$是相关的向量组。由上面的命题可以知道：一定有一个向量 α_k，它是$\beta_1\ \alpha_1\cdots\alpha_{k-1}$ 的线性组合。将 α_k 从$\{\beta_1\ \alpha_1\cdots\alpha_s\}$中删去，从而向量组 $\{\beta_1\alpha_1\cdots\alpha_{k-1}\alpha_{k+1}\cdots\alpha_s\}$与$\{\alpha_1\cdots\alpha_s\}$是等价向量组，同样 $\{\beta_2,\ \beta_1\alpha_1\cdots\alpha_{k-1}\ \alpha_{k+1}\cdots\alpha_s\}$ 是相关组，因此有另一个 α_h，它是$\{\beta_2\ \beta_1\cdots\alpha_s\}$ 的线性组合。再将 α_h 从向量组$\{\beta_2\ \beta_1\cdots\alpha_s\}$删去，便又得一新的向量组$\{\beta_2,\beta_1\alpha_1\cdots\alpha_{k-1}\alpha_{k+1}\cdots\alpha_{h-1},\alpha_{h+1}\cdots\alpha_s\}$。它仍与 $\{\alpha_1\ \alpha_2\cdots\alpha_s\}$等价。仿此每加一$\beta_i$ 到新的向量组，又删去一个 α_j，这样替代下去。所得的向量组都与 $\{\alpha_1\cdots\alpha_s\}$ 是等价的。如果 $t>s$ 势必得出$\{\beta_s\ \beta_{s-1}\cdots\beta_1\}$与$\{\alpha_1\cdots\alpha_s\}$等价。因此 $\beta_{s+1}\cdots\beta_t$可通过$\{\beta_s\ \beta_{s-1}\cdots\beta_1\}$表示。这与$\{\beta_1\cdots\beta_t\}$是独立向量组是矛盾的。所以 $t\leqslant s$，用同样的方法，用 α_i 去替代$\{\beta_j\}$ 中的向量，便可证明 $s\leqslant t$，因此便有

$s=t$。 ▨

运用上面定理证明时所用的替代方法，便可证明下面的命题。

命题 13 V 是 n 维空间，$\{\alpha_1\cdots\alpha_s\}$ 是它的一组独立向量，则一定存在向量 $\alpha_{s+1}\cdots\alpha_n$，使得 $\{\alpha_1\cdots\alpha_s,\alpha_{s+1}\cdots\alpha_n\}$ 为 V 的基，用简单的话来说，便是空间的任一组独立向量都可扩充为基。

下面举例说明上面关于"扩充为基"这一命题的应用。

例 23 （关于子空间的维数定理）V_1,V_2 是 V 的子空间，则有（$\dim V$ 表示空间的维数）

$$\dim(V_1+V_2)=\dim V_1+\dim V_2-\dim(V_1\cap V_2)。\qquad(1\text{-}2)$$

证 设 V_1，V_2，$V_1\cap V_2$ 的维数分别为 s,t,n，而 $\alpha_1\cdots\alpha_n$ 是 $V_1\cap V_2$ 的基，将它分别扩充为 V_1 和 V_2 的基，即

$$\alpha_1\cdots\alpha_n,\beta_1\cdots\beta_{s-n}，\text{是 } V_1 \text{ 的基；}$$

$$\alpha_1\cdots\alpha_n,\gamma_1\cdots\gamma_{t-n}。\text{是 } V_2 \text{ 的基。}$$

证明式(1-2)，只要证明 $\alpha_1\cdots\alpha_n,\beta_1\cdots\beta_{s-n},\gamma_1\cdots\gamma_{t-n}$ 构成 V_1+V_2 的基，它们生成 V_1+V_2 是显然的。下面只要证明向量组 $\{\alpha_1\cdots\alpha_n,\beta_1\cdots\beta_{s-n},\gamma_1\cdots\gamma_{t-n}\}$ 是无关的，为此用反证法。设有不全为零的数 $a_1\cdots a_n,b_1\cdots b_{s-n},c_1\cdots c_{t-n}$ 使得

$$\sum_{i=1}^{n}a_i\alpha_i+\sum_{j=1}^{s-n}b_j\beta_j+\sum_{k=1}^{t-n}c_k\gamma_k=0\Longrightarrow\sum_{i=1}^{n}a_i\alpha_i+\sum_{j=1}^{s-n}b_j\beta_j=-\sum_{k=1}^{t-n}c_k\gamma_k。$$

上面等式左边是属于 V_1 的向量，而右边则是属于 V_2 的向量，这说明两边的向量是属于 $V_1\cap V_2$。但因为 $\alpha_1\cdots\alpha_n$ 是它的基，所以有

$$-\sum_{k=1}^{t-n}c_k\gamma_k=l_1\alpha_1+\cdots+l_n\alpha_n,\Longrightarrow\sum_{k=1}^{t-n}c_k\gamma_k+\sum_{h=1}^{n}l_h\alpha_h=0。$$

因为 $\alpha_1\cdots\alpha_n,\gamma_1\cdots\gamma_{t-n}$ 构成了 V_2 的基，所以 $c_1=\cdots=c_{t-n}=l_1=\cdots=l_n=0$。因而有

$$\sum_{i=1}^{n}a_i\alpha_i+\sum_{j=1}^{s-n}b_j\beta_j=0。$$

同理可得

$$a_1=\cdots=a_n=b_1=\cdots=b_{s-n}=0,$$

这便证明了 $\alpha_1\cdots\alpha_n,\beta_1\cdots\beta_{s-n}$, $\gamma_1\cdots\gamma_{t-n}$ 构成了 V_1+V_2 的基。所以它的维数

$$\dim(V_1+V_2)=n+s-n+t-n=s+t-n$$
$$=\dim V_1+\dim V_2-\dim(V_1\cap V_2)。$$

例 24 V_1+V_2 是直和 $\Longleftrightarrow \dim(V_1+V_2)=\dim V_1+\dim V_2$。

这个命题可由上面讲的维数定理得到，因为 V_1+V_2 是直和的充要条件为 $V_1\cap V_2=\{0\}$，换句话说充要条件为 $\dim(V_1\cap V_2)=0$。由维数定理立即可知 $\dim(V_1+V_2)=\dim V_1+\dim V_2$。

例 25 任意 n 阶方阵与上三角形相似：

这是矩阵理论的一个重要命题。对于 $n=1$ 命题显然成立，下面对矩阵的阶 n 施行数学归纳法。现假定对任何 $n-1$ 阶矩阵都相似于上三角阵。设 A 是任一个 n 阶矩阵，证明 A 与上三角阵相似。先任取 A 的一个特征值 λ_1 以及它对应的一个特征向量 X_1，现将 X_1 扩充为 $X_1\ X_2\cdots X_n$，并使这组向量成为 F^n 的基。如果将它们排成一个矩阵 P，$P=[X_1\ X_2\cdots X_n]$，则 P 是非奇异矩阵。考察乘积 AP

$$AP=A[X_1\cdots X_n]=[AX_1\cdots AX_n]=[\lambda_1X_1\ AX_2\cdots AX_n]$$

$$=[X_1\ X_2\cdots X_n]\begin{bmatrix}\lambda_1 & b_{12}\cdots b_{1n}\\ 0 & \\ \vdots & A_1\\ 0 & \end{bmatrix}。$$ 前面最后一个等式成立是因为向量 $\lambda X_1\ AX_2\cdots AX_n$ 都可以经 $X_1\cdots X_n$ 的线性组合表出。因此

$$P^{-1}AP=\begin{bmatrix}\lambda & b_{12}\cdots b_{1n}\\ 0 & A_1\end{bmatrix},$$

由归纳假设，存在一个非奇异矩阵 Q_1 有

$$Q_1^{-1}A_1Q_1=\begin{bmatrix} c_{22} & c_{23}\cdots c_{2n} \\ & c_{33}\cdots c_{3n} \\ & \ddots \\ & & c_{nn} \end{bmatrix},$$

作 n 阶矩阵 $Q=\begin{bmatrix} 1 & 0\cdots 0 \\ 0 & \\ \vdots & Q_1 \\ 0 & \end{bmatrix}$，这样

$$Q^{-1}P^{-1}APQ=\left[\begin{array}{c|c} 1 & 0\cdots 0 \\ \hline 0 & \\ \vdots & Q_1^{-1} \\ 0 & \end{array}\right]\left[\begin{array}{c|c} \lambda_1 & b_{12}\cdots b_{1n} \\ \hline 0 & \\ \vdots & A_1 \\ 0 & \end{array}\right]$$

$$\left[\begin{array}{c|c} 1 & 0\cdots 0 \\ \hline 0 & \\ \vdots & Q_1 \\ 0 & \end{array}\right]=\begin{bmatrix} 1 & 0 \\ 0 & Q_1^{-1} \end{bmatrix}\begin{bmatrix} \lambda_1 & (b_{12}\cdots b_{1n})Q_1 \\ 0 & A_1Q_1 \end{bmatrix}$$

$$=\begin{bmatrix} \lambda_1 & (b_{12}\cdots b_{1n})Q_1 \\ 0 & Q_1^{-1}A_1Q_1 \end{bmatrix}$$ 上三角形矩阵。

在下一章讲特征值、特征向量时，会讲到这个命题的应用。关于“将一独立向量组扩充为基”这一方法，还有很多命题的证明都需要用到的，请读者以后留意第三章关于 Schur 定理的证明。

第四节 线 性 变 换

线性变换亦是大家熟知的，但这里是讨论不同空间之间的线性变换，因此不假定两个空间的维数是一样的。

1. 线性变换及其简单性质

定义 U,V 是域 F 上的两个线性空间，σ 是 U 到 V 的一个映射，它满足条件：

$$\sigma(a\alpha+b\beta)=a\sigma(\alpha)+b\sigma(\beta),$$

则称 σ 为 U 到 V 的线性变换。如果 $U=V$ 则将 σ 称为 V 的一个线性算子。

下面列举一些矩阵理论中的重要的线性变换。

例 26 设 $A=[a_{ij}]$ 是 $m\times n$ 实矩阵。$x=(x_1\cdots x_n)^T\in R^n$，$y=(y_1\cdots y_m)^T\in R^m$。由等式 $y=Ax$，便确定一个由 R^n 到 R^m 的线性变换。有时简单地讲，一个矩阵 A，表示一个由 R^n 至 R^m 的线性变换。

例 27 σ 是 V 的一个映射。定义 $\sigma(\alpha)=\alpha,\forall\alpha\in V$，这样的映射称为 V 上的恒等映射，用 1_V 表示。1_V 显然是线性算子，因为有

$$1_V(a\alpha+b\beta)=a\alpha+b\beta=a\,1_V(\alpha)+b\,1_V(\beta)。$$

例 28 σ 是 V 的一个映射。定义 $\sigma(\alpha)=0,\forall\alpha\in V$，这样的映射称为 V 上的零映射，用 0_V 表示。0_V 亦显然是线性算子。

例 29 由直和 $V=V_1\oplus V_2$ 所确定的投影算子。

如果线性空间 V 有直和 $V=V_1\oplus V_2$。可以用下面的方法确定 V 的一个算子 σ。因 V 中任一个向量 α 都有唯一分解式 $\alpha=\alpha_1+\alpha_2(\alpha_i\in V_i)$，定义 $\sigma(\alpha)=\alpha_1$，下面便证明 σ 是一个线性算子。设另一个向量 $\beta=\beta_1+\beta_2\quad(\beta_i\in V_i)$，则

$$\begin{aligned}\sigma(a\alpha+b\beta)&=\sigma(a\alpha_1+b\beta_1+a\alpha_2+b\beta_2)\ (\text{显然}a\alpha_i+b\beta_i\in V_i)\\&=a\alpha_1+b\beta_1=a\sigma(\alpha)+b\sigma(\beta)。\end{aligned}$$

所以 σ 是线性算子，显然有性质 $\sigma(\alpha)=\alpha,\forall\alpha\in V_1,\sigma(\alpha)=0,\forall\alpha\in V_2$。故称 σ 是 V 沿子空间 V_2 向 V_1 的投影。

读者可仿照这例子所定义的方式，去定义 V 沿子空间 V_1 在 V_2 的投影算子。下面是 R^2 中的几个几何方面的线性变换的例子。

例 30 $\sigma_\theta(x_1,x_2)=(x_1\cos\theta-x_2\sin\theta,x_1\sin\theta+x_2\cos\theta)$。

它是 R^2 的一个旋转 θ 角的变换(如图 1-3 所示)。

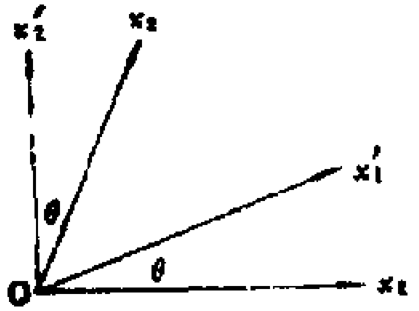

图 1-3

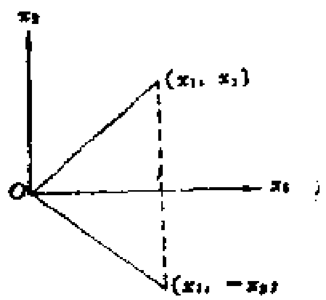

图 1-4

例 31 $\sigma(x_1,x_2)=(x_1,-x_2)$。它实际上是表示关于 x_1 轴上的一个反射(如图 1-4 所示)。

例 32 $\sigma(x_1,x_2)=(x_1,0)$。它实际上是 R^2 沿 x_2 轴在 x_1 轴上的投影(如图 1-5 所示)。

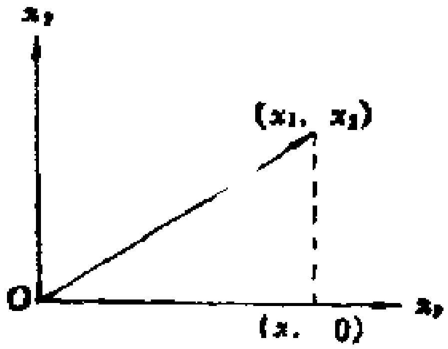

图 1-5

2. 线性变换的基本定理

(1) 线性变换的基本定理

定理说明如何才能确定一个空间 U 至空间 V 的一个线性变换 σ，以后将 U 到 V 的线性变换全体构成的集合，记为 Hom(U,V)。

定理 6 $\alpha_1\cdots\alpha_n$ 是 U 的一组基，$\beta_1\cdots\beta_n$ 是 V 的任意 n 个向量，则存在唯一的一个 $\sigma\in\mathrm{Hom}(U,V)$，并满足 $\sigma(\alpha_i)=\beta_i$，$i=$

$1,2\cdots,n$。

证 设 $\alpha\in U$ 是任一个向量，用基表示便有 $\alpha=x_1\alpha_1+\cdots+x_n\alpha_n$。定义映射

$$\sigma(\alpha)=x_1\beta_1+x_2\beta_2+\cdots+x_n\beta_n, \tag{1-3}$$

证明式(1-3)是所要求的线性变换。设有另一个向量 α'，$\alpha'=x_1'\alpha_1+\cdots+x_n'\alpha_n$。

因而

$$a\alpha+b\alpha'=(ax_1+bx_1')\alpha_1+\cdots+(ax_n+bx_n')\alpha_n,$$

按定义

$$\begin{aligned}\sigma(a\alpha+b\alpha')&=(ax_1+bx_1')\beta_1+\cdots+(ax_n+bx_n')\beta_n\\&=a(x_1\beta_1+\cdots+x_n\beta_n)+b(x_1'\beta_1+\cdots+x_n'\beta_n)\\&=a\sigma(\alpha)+b\sigma(\alpha')。\end{aligned}$$

显然它是满足条件 $\sigma(\alpha_i)=\beta_i$。下面便证明唯一性，假设有另一个线性变换 τ，同样满足 $\tau(\alpha_i)=\beta_i$，则 $\sigma=\tau$ 即 $\sigma(\alpha)=\tau(\alpha)$，$\forall\alpha\in U$，事实上

$$\begin{aligned}\tau(\alpha)&=\tau(x_1\alpha_1+\cdots+x_n\alpha_n)=x_1\tau(\alpha_1)+\cdots+x_n\tau(\alpha_n)\\&=x_1\beta_1+\cdots+x_n\beta_n=\sigma(\alpha)\end{aligned}$$ ▨

这个定理说明，一个线性变换 σ 由该空间的基的像所完全确定。

(2) 线性变换的矩阵

由前一定理知道，一旦一个空间的基的像确定了，则该线性变换便完全确定，所以 $\sigma\in\mathrm{Hom}(U,V)$。$\alpha_1\cdots\alpha_n$是 U 的一组基，$\beta_1\cdots\beta_m$ 是 V的一组基，为了要说明线性变换的实情，应写出下表(即用基 $\beta_1\cdots\beta_m$ 将 $\alpha_1\cdots\alpha_n$ 在 σ 下的像表示出)

$$\begin{aligned}\sigma(\alpha_1)&=a_{11}\beta_1+a_{21}\beta_2+\cdots+a_{m1}\beta_m\\\sigma(\alpha_2)&=a_{12}\beta_1+a_{22}\beta_2+\cdots+a_{m2}\beta_m\\&\vdots\qquad\vdots\qquad\vdots\qquad\qquad\vdots\\\sigma(\alpha_n)&=a_{1n}\beta_1+a_{2n}\beta_2+\cdots+a_{mn}\beta_m。\end{aligned} \tag{1-4}$$

将式(1-4)紧凑地写成一个 $m\times n$ 型矩阵 A 为：

$$A=\begin{bmatrix} a_{11} & a_{12}\cdots a_{1n} \\ a_{21} & a_{22}\cdots a_{2n} \\ \vdots & \vdots \quad \vdots \\ a_{m1} & a_{m2} \quad a_{mn} \end{bmatrix}。$$

它称为 σ 在基$\{\alpha_i\}$与$\{\beta_j\}$ 下的矩阵。特别要注意矩阵 A 的列与式(1-4)行的关系。

例 33 $\sigma\in\mathrm{Hom}(R^2,\ R^3)$，$\sigma(x_1,x_2)=(x_1+3x_2,0,\ 2x_1-4x_2)$。

取 R^2 的基为$(1\ \ 0)^T,(0\ \ 1)^T$；R^3的基为$(1\ \ 0\ \ 0)^T,(0\ \ 1\ \ 0)^T,(0\ \ 0\ \ 1)^T$。

这样 $\sigma(1\ \ 0)=(1\ \ 0\ \ 2)=(1\ \ 0\ \ 0)^T+2(0\ \ 0\ \ 1)^T$，

$\sigma(0\ \ 1)=(3\ \ 0\ \ -2)=3(1\ \ 0\ \ 0)^T-4(0\ \ 0\ \ 1)^T$。

所以在该基下的矩阵 $A=\begin{bmatrix}1 & 3\\ 0 & 0\\ 2 & -4\end{bmatrix}$。

例 34 例 29 给出了由直和 $V=V_1\oplus V_2$ 所确定投影算子 σ，它的定义为对任一 α，$\sigma(\alpha)=\alpha_1$，其中 $\alpha=\alpha_1+\alpha_2$，$\alpha_i\in V_i$。现在可适当选择空间 V 的基，使得 σ 在该组基下的矩阵特别“简单”。如果取 $\alpha_1\cdots\alpha_r$ 为 V_1 的基，$\alpha_{r+1}\cdots\alpha_n$ 为 V_2 的基，由于 V_1、V_2 是直和，所以$\{\alpha_1\ \cdots\alpha_r\ \ \alpha_{r+1}\ \ \cdots\ \ \alpha_n\}$构成 V 的基。因 $\sigma(\alpha_1)=\alpha_1,\cdots,\sigma(\alpha_r)=\alpha_r,\sigma(\alpha_{r+1})=\cdots=\sigma(\alpha_n)=0$，所以$\sigma$在该组基下的矩阵为：

$$\begin{bmatrix} \left.\begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix}\right\} r & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}。$$

例 35 σ：$\mathrm{Hom}(R^{2\times2},P_2(R))$，这里 $P_2(R)$是实系数，且是次不超过 2 的多项式全体所构成的线性空间。σ 的定义如下：

$$\sigma\begin{bmatrix}a & b\\ c & d\end{bmatrix}=(a+b)+2dx+bx^2。$$

选 $R^{2\times 2}$ 的基为 $E_{11}=\begin{bmatrix}1&0\\0&0\end{bmatrix}$ $E_{12}=\begin{bmatrix}0&1\\0&0\end{bmatrix}$ $E_{13}=\begin{bmatrix}0&0\\1&0\end{bmatrix}$，$E_{14}=\begin{bmatrix}0&0\\0&1\end{bmatrix}$，$P_2(R)$的基为 $1\ x\ x^2$。

则 σ 在该基下的矩阵为：

$\sigma(E_{11})=1,\sigma(E_{12})=1+x^2,\sigma(E_{13})=0,E_{14}=x$。

因此线性变换 σ 的矩阵为：$A=\begin{bmatrix}1&1&0&0\\0&0&0&1\\0&1&0&0\end{bmatrix}$。

(3) 用线性变换的矩阵计算变换的像

$\sigma\in \mathrm{Hom}(U,V)$。$\{\alpha_1\cdots\alpha_n\}$ 是 U 的基，$\{\beta_1\cdots\beta_n\}$ 是 V 的基。$\alpha\in U$，有 $\alpha=x_1\alpha_1+\cdots+x_n\alpha_n$，而 $\sigma(\alpha)=y_1\beta_1+\cdots+y_m\beta_m$。在选定上面的基之后，$n$ 维向量$(x_1\cdots x_n)^T$ 便与 α 对应，m 维向量$(y_1\cdots y_m)$便与 $\sigma(\alpha)$对应。现在如果知道 σ 在基$\{\alpha_i\}$，$\{\beta_j\}$的矩阵 A，便可由 $x=(x_1\cdots x_n)^T$ 去求 $y=(y_1\cdots y_m)$，它们是按下面公式进行的。

$$y=Ax \tag{1-5}$$

(证略)

(4) 线性变换的像空间与零空间

给定一个线性变换 $\sigma\in \mathrm{Hom}(U,V)$，有二个重要的子空间与 σ 联系着。一是 σ 的像空间，另一是 σ 的零空间或核。它们分别记为 $R(\sigma)$与 $\mathrm{Ker}(\sigma)$，详细定义如下。

定义 设 $\sigma\in \mathrm{Hom}(U,V)$，像空间 $R(\sigma)$是指V 的子集，即为 $R(\sigma)=\{\sigma(\alpha)\mid\alpha\in U\}$。$\sigma$ 的零空间 $\mathrm{Ker}\sigma=\{\alpha\mid\alpha\in U,\ \sigma(\alpha)=0\}$ 是 U 的一个子集。

容易证明这两个集合都是子空间。$R(\sigma)$的维数称为 σ 的秩，记为 $r(\sigma)$。$\mathrm{Ker}\sigma$ 的维数称为 σ 的零度，记为$\eta(\sigma)$。这是线性变换的两个重要指标，以后在很多的计算场合都要用到这些量。

例 36 $\sigma\in \mathrm{Hom}(R^4,R^3)$，$\sigma(x_1,x_2,x_3,x_4)$

$=(x_1-x_2+x_3+x_4,x_1+2x_3-x_4,x_1+x_2+3x_3-3x_4)$。

求$\mathrm{Ker}\sigma$与$R(\sigma)$的基。

事实上如果将向量写成列向量的形式，上面的变换便可写成矩阵与向量乘积的形式。即

$$\sigma\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}=\begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & 0 & 2 & -1 \\ 1 & 1 & 3 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$即$\sigma(x)=Ax,\forall x\in R^4$。

要求$\mathrm{Ker}\sigma$的基，相当于求齐次线性方程组$Ax=0$的基础解系。为此对系数矩阵进行行初等变换。即

$$A=\begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & 0 & 2 & -1 \\ 1 & 1 & 3 & -3 \end{bmatrix}\longrightarrow\begin{bmatrix} 1 & 0 & 2 & -1 \\ 0 & -1 & -1 & 2 \\ 0 & 1 & 1 & -2 \end{bmatrix}\longrightarrow$$

$$\begin{bmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix}。$$

由此得方程的基础解系为$(-2\ -1\ 1\ 0)^T$ $(1\ 2\ 0\ 1)^T$。为了要求出$R(\sigma)$的基，可采用如下方法：由$\sigma(x)$的表示式便可知道$R(\sigma)$是由矩阵A的列向量线性组合而成，因此$R(\sigma)$就是矩阵A的列空间。为了求出基，只要看上面变换的最后一个矩阵。它实际上是A的Hermite标准型。由于第一、二列是独立的，因此A的第一、二列是独立的，它构成了$R(\sigma)$的基。$R(\sigma)=[(1\ 1\ 1)^T,(-1\ 0\ 1)^T]$。

下面这个定理说明一个线性变换的秩与零度的关系。

定理7 $\sigma\in\mathrm{Hom}(U,V)$，则$r(\sigma)+\eta(\sigma)=\dim U$。

证 用扩充基的方法。设$\dim U=n$，$\dim\mathrm{Ker}\sigma=s$，因而只要证明$\dim R(\sigma)=n-s$。

令$\alpha_1\cdots\alpha_s$为$N(\sigma)$的基，将它扩充为U的基$\alpha_1\cdots\alpha_s,\alpha_{s+1}\cdots\alpha_n$。对任一$\alpha\in U$，有

$$\alpha=x_1\alpha+\cdots+x_s\alpha_s+x_{s+1}\alpha_{s+1}+\cdots+x_n\alpha_n,$$

$$\sigma(\alpha)=x_1\sigma(\alpha_1)+\cdots+x_s\sigma(\alpha_s)+x_{s+1}\sigma(\alpha_{s+1}\sigma(\alpha_{s+1})+\cdots$$
$$+x_n\sigma(\alpha_n)=x_{s+1}\sigma(\alpha_{s+1})+\cdots+x_n\sigma(\alpha_n)。$$

这个结果说明 $R(\sigma)$ 是由向量 $\sigma(\alpha_{s+1})\cdots\sigma(\alpha_n)$ 生成。余下只要证明上面 $n-s$ 个向量是独立的。

如果存在有不全为零的数 $a_1\cdots a_{n-s}$，使得

$$a_1\sigma(\alpha_{s+1})+a_2\sigma(\alpha_{s+2})+\cdots+a_{n-s}\sigma(\alpha_n)=0,$$

这相当于有

$$\sigma(a\alpha_{s+1}+a_2\alpha_{s+2}+\cdots+a_{n-s}\alpha_n)=0。$$

因此

$$a_1\alpha_{s+1}+\cdots+a_{n-s}\alpha_n\in \mathrm{Ker}\sigma。$$

从而

$$a_1\alpha_{s+1}+\cdots+a_{n-s}\alpha_n=b_1\alpha_1+\cdots+b_s\alpha_s,$$

但 $\alpha_1\cdots\alpha_n$ 是独立的，则

$$a_1=\cdots=a_{n-s}=b_1=\cdots=b_s=0,$$

所以 $\sigma(\alpha_{s+1})\cdots\sigma(\alpha_n)$ 是独立的。 ▨

例 37 设 $\sigma\in\mathrm{Hom}(R^3,R^2)$，定义 $\sigma(x_1,x_2,x_3)=(x_3,0)$。求 $\mathrm{Ker}\sigma$ 和 $R(\sigma)$ 的基。

采用上面证明的方法。先求 $\mathrm{Ker}\sigma$ 的基，由 σ 的定义可知，要求 $x_3=0$，所以 $\mathrm{Ker}\sigma=\{(x_1,x_2,0)\mid x_1,x_2\in R\}$。它是一个二维的子空间，选它的基为(1 0 0)，(0 1 0)，然后将它扩充为 R^3 的基，比如加入向量(0 0 1)，σ(0 0 1)便是 $R(\sigma)$ 的基。即(1 0)是 $R(\sigma)$ 的基。

3. 线性变换的基本类型

(1) 在这节里，就前面提出的线性变换 σ 的两个基本量，秩 $r(\sigma)$ 及零度 $\eta(\sigma)$ 的应用，讲解一些基本结果。

首先讨论几种线性变换的基本类型。

定义 $\sigma\in\mathrm{Hom}(U,V)$，$\dim U=n$，$\dim V=m$。

如果 $\alpha_1\neq\alpha_2$，有 $\sigma(\alpha_1)\neq\sigma(\alpha_2)$，$\sigma$ 称为单变换；

如果对任一 $\beta\in V$，$\exists\alpha\in U$，有 $\sigma(\alpha)=\beta$，则称 σ 是满变换；

又如果 σ 又是单变换又是满变换，则称为双变换，亦称为同构。

下面这个定理是用线性变换的 $r(\sigma)$ 与 $\eta(\sigma)$ 去判定线性变换的类型。

定理 8 $\sigma\in \mathrm{Hom}(U,V)$，则

(i) σ 是单变换 $\Longleftrightarrow \eta(\sigma)=0$。

(ii) σ 是满变换 $\Longleftrightarrow r(\sigma)=\dim V$。

证 (i) $\Longrightarrow$ 因为 σ 是单变换。则 $\alpha_1,\alpha_2\in U$，$\alpha_1\neq\alpha_2$ 有 $\sigma(\alpha_1)\neq\sigma(\alpha_2)$，或者可讲 $\alpha\neq 0$ 时有 $\sigma(\alpha)\neq 0$。因而 $\mathrm{Ker}\sigma$ 仅由零向量构成即 $\eta(\sigma)=0$。

$\Longleftarrow \eta(\sigma)=0$ 说明 $\mathrm{Ker}\sigma=\{0\}$。设 $\alpha_1\neq\alpha_2$，如果 $\sigma(\alpha_1)=\sigma(\alpha_2)$，这说明 $\sigma(\alpha_1-\alpha_2)=0$ 或 $\alpha_1-\alpha_2\in\mathrm{Ker}\sigma$。但它应为 $\alpha_1-\alpha_2=0$，所以矛盾。因此 $\alpha_1\neq\alpha_2$ 一定有 $\sigma(\alpha_1)\neq\sigma(\alpha_2)$。即 σ 是单变换。

(ii) $\Longrightarrow$ 由于 σ 是满变换，这样 $R(\sigma)=V$，$\dim R(\sigma)=V$，即 $r(\sigma)=\dim V$。

$\Longleftarrow$ 由于 $\dim R(\sigma)=r(\sigma)=\dim V$，又 $R(\sigma)\subseteq V$，因而可取 $R(\sigma)$ 的基 $\beta_1\cdots\beta_m$ 作为 V 的基。设 $\alpha_1\cdots\alpha_m$ 有 $\sigma(\alpha_i)=\beta_i$。对任一 $\beta\in V$ 有 $\beta=y_1\beta_1+\cdots+y_m\beta_m=y_1\sigma(\alpha_1)+\cdots+y_m\sigma(\alpha_m)=\sigma(y_1\alpha_1+\cdots+y_m\alpha_m)$。$y_1\alpha_1+\cdots+y_m\alpha_m\in U$ 说明 σ 是满变换。▨

例 38 设 A 是 $m\times n$ 型矩阵。证明 $r(A)=n$，A 表示 F^n 至 F^m 的一个单变换，而 $r(A)=m$ 时，A 表示 F^n 至 F^m 的一个满变换。

证 $r(A)=n$，要证明 A 是单变换，只要证明 $Ax=0$ 仅有零解。因为这时 A 的诸列向量是线性独立的，所以线性组合为零向量仅在组合系数全为零时才成立，因而命题成立。

$r(A)=m$ 说明 A 有 m 个独立的列向量（这些列向量都是 m 维的）。所以 A 的列向量可以生成 F^m。因此对任何 $b\in F^m$ 都可通过 A 的列向量的线性组合而得，即 $Ax=b$ 成立。所以 A 是 F^n 至 F^m 的一个满变换。

(2) 同构

前面讲到一个线性变换$\sigma \in \mathrm{Hom}(U,V)$，如果是一个双变换，则$\sigma$便称为同构。显然这时$U$与$V$的元素之间存在一一对应。但$\sigma$是线性变换，所以还有

$$\sigma(a\alpha+b\beta)=a\sigma(\alpha)+b\sigma(\beta)。$$

即两个空间元素的运算保持对应关系（见图 1-6）。

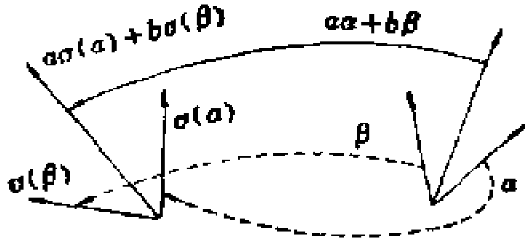

图 1-6

数学里称U，V这两个代数系统为同构，从抽象的角度讲我们并不区别这两个代数系统。下面例举矩阵理论中两个最重要的同构。设V是域F上的n维线性空间。

例 39　V与F^n同构。

选定$\alpha_1\cdots\alpha_n$为V的一组基，而选自然基$e_1=(1\ 0\cdots0)^T e_2=(0\ 1\cdots0)^T\cdots e_n=(0\cdots0\ 1)^T$为$F^n$的一组基。已经知道存在唯一的一个线性变换$\sigma$满足条件$\sigma(\alpha_i)=e_i$。在这个变换之下，$V$中的任一个向量$\alpha=x_1\alpha+x_2\alpha_2+\cdots+x_n\alpha_n$，它的像$\sigma(\alpha)=x_1\sigma(\alpha_1)+\cdots+x_n\sigma(\alpha_n)=x_1e_1+x_2e_2+\cdots+x_ne_n=(x_1\cdots x_n)$。现在证明它是个双变换。首先证明它是个单变换，要$\sigma(\alpha)=0$。由像的表达式可知$(x_1\cdots x_n)=0$即$x_i=0$，所以$\alpha=0$，$\mathrm{Ker}\sigma=\{0\}$，故$\sigma$是单变换。满变换是显然的。所以$\sigma$是双变换，而且$\sigma$是线性变换，因而$V$与$F^n$是同构。

由这个例子说明，研究线性空间F^n是具有普遍意义。但要注意的是，作出这个同构时，需在二个空间分别选定基。

例 40　线性变换与矩阵同构。

前面已经介绍了集合$\mathrm{Hom}(U,V)$，它是n维线性空间U到m维线性空间V的线性变换全体。假定选定了$\alpha_1\cdots\alpha_n$是U的一

组基，而$\beta_1\cdots\beta_m$是V的一组基，在这个前提之下，证明对任一$\sigma\in \mathrm{Hom}(U,V)$对应唯一一个$m\times n$型矩阵$A$，且这对应是同构。事实上由式(1-4)可知$\sigma$完全由矩阵$A=[a_{ij}]_{m\times n}$所确定。便作$\mathrm{Hom}(U,V)$到$F^{m\times n}$的变换$\mathscr{A}$，$\mathscr{A}(\sigma)=A$。首先$\mathscr{A}$是单变换。因为$\sigma_1\neq\sigma_2$必然有某一$i$有$\sigma_1(\alpha_i)\neq\sigma_2(\alpha_i)$（因为如果基中任一个元的像都相同，这两个线性变换是相同的。），所以σ_1的矩阵A_1的第i列与σ_2的矩阵A_2的第i列不同，$\Longrightarrow A_1\neq A_2$ $\Longrightarrow\mathscr{A}$是单变换。$\mathscr{A}$是满变换是显然的。现证明$\mathscr{A}$是线性变换，这相当于要证明$\mathscr{A}(\sigma_1+\sigma_2)=A_1+A_2$，$\mathscr{A}(a\sigma)=aA$，事实上两个线性变换之和仍是一个线性变换。在选定一组基下，线性变换之和的矩阵便是两个矩阵之和，同样$a\sigma$也是一个线性变换，而它的矩阵为aA（详细证明见《工程数学的线性代数》由上海交通大学编写，高等教育出版社出版）。因此在选定一组基下，$\mathrm{Hom}(U,V)$与$F^{m\times n}$是同构的。

由于前面的同构，在讨论抽象的线性空间时可变换为较为具体的F^n空间，并应用矩阵的方法来研究。如果两个空间是同构，那么这两个空间不仅元素之间是一一对应，而且还保持运算关系。因此，独立向量组与独立向量组对应，相关向量组对应相关向量组，且保持组合系数相同。

例 41 线性问题的解法。

给定一$\sigma\in\mathrm{Hom}(U,V)$及$\beta\in V$，求解集$S$

$$S=\{\xi\mid\xi\in U,\sigma(\xi)=\beta\}$$

这称为线性问题。

这个问题的求解，可运用前面讲的同构来解决。事实上选定了U的基$\{\alpha_1\cdots\alpha_n\}$与$V$的基$\{\beta_1\cdots\beta_m\}$，$\sigma$便与矩阵$A$对应，而$U$与$F^n$；$V$与$F^m$同构。在这同构之下（见图1-7），$\xi$与$x$，$\beta$与$b$对应，这样线性问题$\sigma(\xi)=\beta$便变换为解线性方程组$Ax=b$的问题。

例 42 抽象空间向量组相关性的判定。

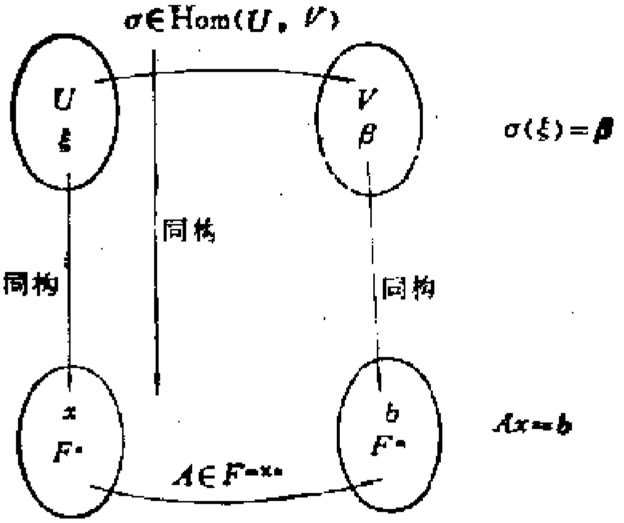

图 1-7

设 V 是n维空间，$\beta_1\cdots\beta_s$是它的一组向量，现要判定它们的相关性。事实上可选定 V 的一组基 $\alpha_1\cdots\alpha_n$。在这组基下 V 与 F^n同构。这时 $\beta_i=a_{i1}\alpha_1+\cdots+a_{in}\alpha_n$，$\beta_i$ 与$(a_{i1}\ \ a_{i2},\cdots a_{in})$对应。这样，向量组$\{\beta_1\cdots\beta_s\}$的相关性判定，便变换为向量组$\{(a_{11}\ \ a_{12}\cdots a_{1n})$ $(a_{21}\ \ a_{22}\cdots a_{2n})\cdots(a_{s1}\ \ a_{s2}\cdots a_{sn})\}$的相关性判定；一个抽象的线性空间向量组的相关性讨论便转变为 F^n 中向量组相关性的讨论。后者可用矩阵的秩来计算。

4. 线性变换的计算问题

这节主要讨论 $\sigma\in\mathrm{Hom}\ (U,V)$，在不同基下 σ 的矩阵的关系。在计算中下面二个问题是基本的。

(i) 线性空间的坐标变换。

(ii) 线性变换用矩阵表示，且用矩阵计算线性变换的像(见式1-5))。

这两个问题在工程数学中已讲过，所以这里只复述一下结果。

(1) 线性空间的坐标变换

设 n 维线性空间 U 有二组基 $\{\alpha_1\cdots\alpha_n\}$与$\{\alpha_1'\cdots\alpha_n'\}$。它们的关系用下面方程组表示：

$$\begin{aligned} \alpha_1' &= p_{11}\alpha_1 + p_{21}\alpha_2 + \cdots + p_{n1}\alpha_n \\ \alpha_2' &= p_{12}\alpha_1 + p_{22}\alpha_2 + \cdots + p_{n2}\alpha_n \\ &\vdots \qquad \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \\ \alpha_n' &= p_{1n}\alpha_1 + p_{2n}\alpha_2 + \cdots + p_{nn}\alpha_n。\end{aligned} \tag{1-6}$$

$P=[p_{ij}]$称为由 α-基到 α'-基的过渡矩阵，如果 $\alpha\in U$，$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n=x_1'\alpha_1'+\cdots+x_n'\alpha_n'$，用 x 记$(x_1\cdots x_n)^T$，x' 记 $(x_1'\cdots x_n')$，则有结果 $x=Px'$。

例 43 $U=R^2$，取 $\alpha_1=(1\ \ 1)^T$，$\alpha_2=(1\ \ -1)^T$，$\alpha_1'=(2\ \ 4)^T$，$\alpha_2'=(3\ \ 1)^T$，求由 α-基到 α'-基的过渡矩阵。

解这题相当于要将 α'-基通过 α-基表示出来。用矩阵的形式表示便是 $[\alpha_1'\ \ \alpha_2']=[\alpha_1\ \ \alpha_2]P$，$P$ 是过渡矩阵。
因此

$$P=[\alpha_1\ \ \alpha_2]^{-1}[\alpha_1'\ \ \alpha_2']$$

如果用 $A=[\alpha_1\ \ \alpha_2]$，$B=[\alpha_1'\ \ \alpha_2']$，解上述结果，可用如下格式：将 AB 合并成一个分块矩阵 $[A|B]$。通过对分块矩阵施行行变换，将块 A 变为单位阵，则原块 B 的部分便是 $A^{-1}B$，即

$$\left[\begin{array}{cc|cc} 1 & 1 & 2 & 3 \\ 1 & -1 & 4 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 1 & 2 & 3 \\ 0 & -2 & 2 & -2 \end{array}\right] \to$$

$$\left[\begin{array}{cc|cc} 1 & 0 & 3 & 2 \\ 0 & -2 & 2 & -2 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 3 & 2 \\ 0 & 1 & -1 & 1 \end{array}\right]$$

故 $$P=\begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix}。$$

(2) 线性变换用矩阵表示，并且用矩阵计算线性变换的像

这问题可参考式(1-4)和(1-5)。

有了这些结果便可讨论 $\sigma\in\mathrm{Hom}(U,V)$，当 U,V 都有坐标变换时，σ 的矩阵关系。先讨论两个简单的情形，然后再由这两种简单的情形合并为一般的情形。

(i) 空间 U 的基为 $\alpha_1\cdots\alpha_n$，空间 V 有坐标变换。它的基分别

为 $\beta_1\cdots\beta_m$；$\beta_1'\cdots\beta_m'$。由 β-基到 β'-基的关系为：

$$\beta_1' = q_{11}\beta_1 + q_{21}\beta_2 + \cdots + q_{m1}\beta_m$$
$$\beta_2' = q_{12}\beta_1 + q_{22}\beta_2 + \cdots + q_{m2}\beta_m$$
$$\vdots$$
$$\beta_m' = q_{1m}\beta_1 + q_{2m}\beta_2 + \cdots + q_{mm}\beta_m。$$

现有 $\sigma\in \mathrm{Hom}(U,V)$，它在 $\{\alpha_i\}$ 基与 $\{\beta_j\}$ 基下的矩阵为 A，在 $\{\alpha_i\}$ 基与 $\{\beta_j'\}$ 基下的矩阵为 B，求 A 与 B 的关系。要解决这个问题先观察图 1-8 所示。

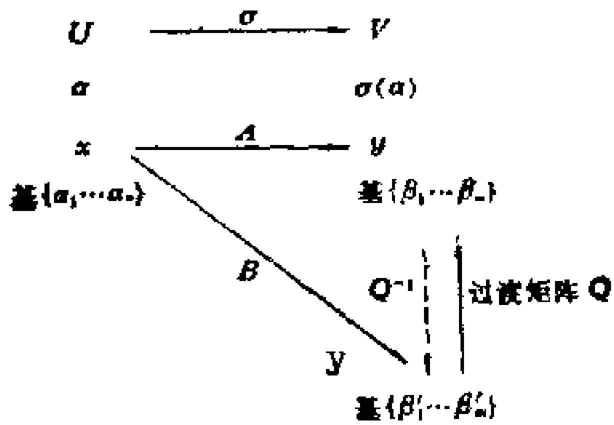

图 1-8

由前面已经得出的结果知道：上面三角形的通路图，箭头一端的量，顺方向要得另一端的量，只要乘上通路上所示的矩阵即可。如果通路可逆转的话，按逆转方向计算，只要乘上原通路所示矩阵的逆阵。由上图可知 $y' = Bx$，$y' = Q^{-1}Ax$。由选定基下 $\mathrm{Hom}(U,V)$ 与 $F^{m\times n}$ 是同构，所以 $B = Q^{-1}A$。

(ii) 空间 U 有坐标变换，它的两组基为 $\{\alpha_1\cdots\alpha_n\}$，$\{\alpha_1'\cdots\alpha_n'\}$，它们的关系为式(1-6)所示，而 V 的基为 $\{\beta_1\cdots\beta_m\}$，这时 σ 在 $\{\alpha_i\}$ 与 $\{\beta_j\}$ 下的矩阵为 A，而在 $\{\alpha_i'\}$ 与 $\{\beta_j\}$ 下的矩阵为 B。它们的关系便可从图 1-9 看出。

由图 1-9 可知

$$y = Ax = APx',\ y = Bx'。$$

因此

$$B = AP。$$

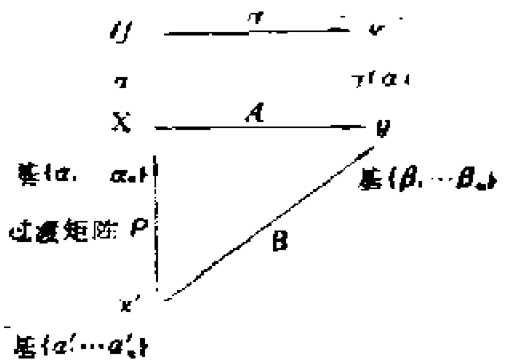

图 1-9

(iii) 现在要讨论是 U,V 两个空间均有坐标变换。在 U 中，过渡矩阵为 P，而在 V 中，过渡矩阵为 Q。又设 σ 在基 $\{\alpha_i\}$ 与 $\{\beta_j\}$ 下的矩阵为 A，在基 $\{\alpha'_i\}$ 与 $\{\beta'_j\}$ 下的矩阵为 B。现求矩阵 A,B 的关系见图 1-10。

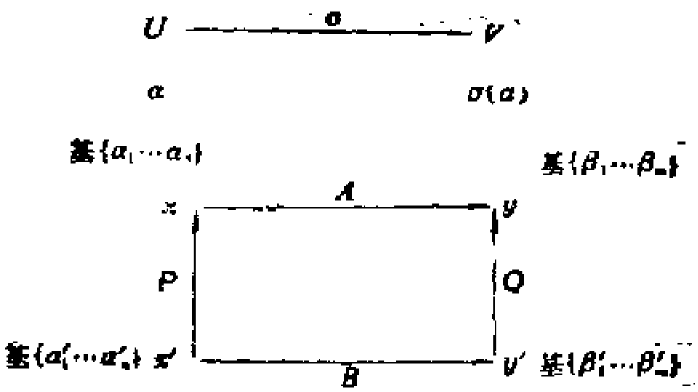

图 1-10

由图 1-10 可知

$$y' = Bx'$$

由另一条通路可知 $y' = Q^{-1}APx'$，

因此
$$B = Q^{-1}AP。$$

例 44　设有 R^3 到 R^2 的一个线性变换，它由矩阵

$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ 定义。试求该线性变换在 R^3 的基 $\alpha_1 = (1\ \ 1\ \ 1)^T$，

$\alpha_2=(1\ \ 1\ \ 0)^T$, $\alpha_3=(1\ \ 0\ \ 0)$，与 R_2 的基 $\beta_1=(1\ \ 1)^T$，$\beta_2=(1\ \ 0)^T$ 下的矩阵。

事实上矩阵 A 可看作为 R^3 的基为 $e_1=(1\ \ 0\ \ 0)^T$，$e_2=(0\ \ 1\ \ 0)^T$，$e_3=(0\ \ 0\ \ 1)^T$，R^2 的基为 $\varepsilon_1=(1\ \ 0)^T$，$\varepsilon_2=(0\ \ 1)^T$ 时的矩阵。这样在 R^3 的过渡矩阵 P 为 $\begin{bmatrix}1&1&1\\1&1&0\\1&0&0\end{bmatrix}$，在 R^2 的过渡矩阵

$Q=\begin{bmatrix}1&1\\1&0\end{bmatrix}$，$Q^{-1}=\begin{bmatrix}0&1\\1&-1\end{bmatrix}$ 因此

$$B=Q^{-1}AP=\begin{bmatrix}0&1\\1&-1\end{bmatrix}\begin{bmatrix}1&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&1&1\\1&1&0\\1&0&0\end{bmatrix}$$

$$=\begin{bmatrix}0&1\\1&-1\end{bmatrix}\begin{bmatrix}2&2&1\\1&0&0\end{bmatrix}=\begin{bmatrix}1&0&0\\1&2&1\end{bmatrix}.$$

第二章　特征值与特征向量

第一节　线性变换的特征值与特征向量

设 $\sigma\in \mathrm{Hom}(U,V)$，由前一章讨论已知道，在空间 U,V 选择不同的基，线性变换的矩阵是不一样的。因此要有效地研究线性变换的性质，适当选择空间的基是相当重要。下面的讨论均假设 $U=V$，换句话说 σ 是空间 U 的一个线性算子。现在举例说明适当选择空间的基，可使 σ 的矩阵表示简单。

1. 实例

例 1　设 $\sigma^2=\sigma,\sigma\in \mathrm{Hom}(V,V)$，试适当选择 V 中的基使 σ 的矩阵表示简单。

证　实际上要证明空间 V 可有如下的分解

$$V=R(\sigma)\oplus \mathrm{Ker}\sigma。$$

先证明对任何 $\alpha\in R(\sigma)$，有 $\sigma(\alpha)=\alpha$。事实上因 $\alpha\in R(\sigma)$，存在 β 有 $\alpha=\sigma(\beta)$，所以 $\sigma(\alpha)=\sigma(\sigma(\beta))=\sigma^2(\beta)=\sigma(\beta)=\alpha$。(这里应用 $\sigma^2=\sigma$)

其次证明 $V=R(\sigma)+\mathrm{Ker}\sigma$。因为 $I=\sigma+(I-\sigma)$，所以对任一个 α 有 $\alpha=I\alpha=\sigma\alpha+(I-\sigma)\alpha$。显然 $\sigma(\alpha)\in R(\sigma)$，$(I-\sigma)\alpha\in \mathrm{Ker}\sigma$。然后证明上面的和是直和。如果有 $\alpha\in R(\sigma)\cap \mathrm{Ker}\sigma$，则 $\alpha=\sigma(\alpha)=0$，所以 $R(\sigma)\cap \mathrm{Ker}\sigma=\{0\}$。这便证明了 V 是 σ 的像空间与核的直和。所以可选择 $R(\sigma)$ 的基是 $\alpha_1\cdots\alpha_r$，$\mathrm{Ker}\sigma$ 的基是 $\alpha_{r+1}\cdots\alpha_n$，并且合并起来为 V 的基，在此基下 σ 的矩阵为

$$\sigma(\alpha_i)=\alpha_i,\ i=1\cdots r,$$
$$\sigma(\alpha_j)=0,\ j=r+1\cdots n,$$

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}。$$

定义 设 $\sigma\in \mathrm{Hom}(U,U)$。如果存在有一 $\alpha\neq 0,\in U$ 且满足条件 $\sigma(\alpha)=\lambda\alpha$，则 α 称为特征向量，λ 称为特征值。

定义 若集合 $V_\lambda=\{\alpha|\sigma(\alpha)=\lambda\alpha,\quad \alpha\in V\}$，则称 V_λ 为对应于 λ 的特征子空间。

显然上面的例中，$R(\sigma)$ 是对应于特征值 $\lambda=1$ 的特征子空间；而 $\mathrm{Ker}\sigma$ 是对应于特征值 $\lambda=0$ 的特征子空间，而在这个例中，整个空间恰为两个特征子空间（σ 仅有这两个特征值）的直和。这时 σ 的矩阵表示为对角阵。这对于一般情况也是对的。

例 2 $\sigma\in \mathrm{Hom}(U,U)$，$\dim U=n$，且 σ 恰有 n 个线性独立的特征向量 $\alpha_1,\cdots,\alpha_n$，则可选择这 n 个向量为 U 的基，在这组基下线性变换 σ 的矩阵便是如下对角形。

$$\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

这里 $\sigma(\alpha_i)=\lambda_i\alpha_i$，$i=1,2,\cdots,n$。

2 特征值与特征向量的计算

要计算特征值与特征向量，就要解方程 $\sigma(\alpha)=\lambda\alpha$。在实际计算的时候，在 U 中取定一组基 $\alpha_1\cdots\alpha_n$。求出 σ 在该组基下的矩阵 A，这时方程 $\sigma(\alpha)=\lambda\alpha$ 便对应线性方程组 $Ax=\lambda x$。所以全部的计算就是解后一个方程。

例 3 $\sigma\in \mathrm{Hom}(R^3,R^3)$，它定义为

$$\sigma(x_1,x_2,x_3)=\begin{bmatrix}2&0&0\\1&2&-3\\1&0&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}。$$

解 (i) 首先适当选择基，并求出线性变换 σ 在所选基下的矩阵。本题最方便是选自然基 $e_1=(1\ 0\ 0)^T, e_2=(0\ 1\ 0)^T, e_3=(0\ 0\ 1)^T$，因为在该组基下 σ 的矩阵是

$$\begin{bmatrix}2&0&0\\1&2&-3\\1&0&-1\end{bmatrix}。$$

(ii) 计算特征多项式：

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

求出特征值为 $\lambda_1=2, \lambda_3=-1$；

(iii) 求特征向量。

$\lambda_1=2$ 时求它所对应的特征向量要解方程

$$[2I-A]x=0\quad 或\begin{bmatrix}0&0&0\\-1&0&3\\-1&0&3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=0。$$

由于 $r(2I-A)=1$，

所以基础解为 2 个，分别解出 $\alpha_1=(3\ 0\ 1)^T, \alpha_2=(0\ 1\ 0)^T$。

同理可解出 $\lambda=-1$ 的特征向量为 $\alpha_3=(0\ 1\ 1)^T$。

这样如果选择 $\alpha_1\ \alpha_2\ \alpha_3$ 为基，σ 在该基下的矩阵便是

$$\begin{bmatrix}2&&\\&2&\\&&-1\end{bmatrix}$$

由自然基 $e_1\ e_2\ e_3$ 过渡到基 $\alpha_1\ \alpha_2\ \alpha_3$ 的过渡矩阵为：

$$P=\begin{bmatrix}3&0&0\\0&1&1\\1&0&1\end{bmatrix}$$

由前面讨论坐标变换的结果可知，应有

$$P^{-1}\begin{bmatrix}2&0&0\\1&2&-3\\1&0&-1\end{bmatrix}P=\begin{bmatrix}2&&\\&2&\\&&-1\end{bmatrix}$$

3. $|\lambda I-A|$ 的展开式及有关的结论

$$|\lambda I-A|=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}。$$

它是一个 n 次多项式，因而亦可写成

$$|\lambda I-A|=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0。$$

其中 n_i 称为特征值 λ_i 的代数重数(代数重数的一个重要性质为：全部特征值的代数重数之和为线性空间的维数)。但是 λ_i 所对应的独立的特征向量的个数不一定是 n_i 个。λ_i 所对应的独立特征向量的个数称为 λ_i 的几何重数，记为 g_i。显然 $g_i=n-r(\lambda_i I-A)$。下面来讨论 $|\lambda I-A|$ 的展开式。

例 4 已知 $A=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}$，求 $|\lambda I-A|$ 的展开式。

解 $$\begin{vmatrix}\lambda-a_{11}&-a_{12}&-a_{13}\\-a_{21}&\lambda-a_{22}&-a_{23}\\-a_{31}&-a_{32}&\lambda-a_{33}\end{vmatrix}=\begin{vmatrix}\lambda-a_{11}&0-a_{12}&0-a_{13}\\0-a_{21}&\lambda-a_{22}&0-a_{23}\\0-a_{31}&0-a_{32}&\lambda-a_{33}\end{vmatrix}$$

$$=\begin{vmatrix}\lambda&0&0\\0&\lambda&0\\0&0&\lambda\end{vmatrix}+\begin{vmatrix}\lambda&0&-a_{13}\\0&\lambda&-a_{23}\\0&0&-a_{33}\end{vmatrix}+\begin{vmatrix}\lambda&-a_{12}&0\\0&-a_{22}&0\\0&-a_{32}&\lambda\end{vmatrix}$$

$$+\begin{vmatrix}-a_{11}&0&0\\-a_{21}&\lambda&0\\-a_{31}&0&\lambda\end{vmatrix}+\begin{vmatrix}\lambda&-a_{12}&-a_{13}\\0&-a_{22}&-a_{23}\\0&-a_{32}&-a_{33}\end{vmatrix}+\begin{vmatrix}-a_{11}&0&-a_{13}\\-a_{21}&\lambda&-a_{23}\\-a_{31}&0&-a_{33}\end{vmatrix}$$

$$+\begin{vmatrix} -a_{11} & -a_{12} & 0 \\ -a_{21} & -a_{22} & 0 \\ -a_{31} & -a_{32} & \lambda \end{vmatrix}+\begin{vmatrix} -a_{11} & -a_{12} & -a_{13} \\ -a_{21} & -a_{22} & -a_{23} \\ -a_{31} & -a_{32} & -a_{33} \end{vmatrix}$$

$$=\lambda^3-(a_{11}+a_{22}+a_{33})\lambda^2+\left(\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}+\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\right.$$

$$\left.+\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}\right)\lambda-|A|。$$

式中，λ^2 的系数是 A 的全部一阶主子式之和，前面变号；

λ 的系数是 A 的全部二阶主子式之和，不变号；

常数项是 $|A|$，变号。

用类似的方法便可得出下面的一般结果。

$$|\lambda I-A|=\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots$$
$$+(-1)^k(\sum S_k)\lambda^{n-k}+\cdots+(-1)^n|A| \qquad (2\text{-}1)$$

式中：$\sum S_k$ 表示 A 的全部 k 阶主子式之和。

由式(2-1)可知如下结论：

(i) 由多项式的根与系数关系可知：

$$\lambda_1+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn},\ |A|=\lambda_1\lambda_2\cdots\lambda_n。$$

(ii) 如果矩阵 A 的行列式为零，则一定有零特征值。反之亦然。如果知道矩阵 A 的秩为 r，这样便可得出关于零特征值的更多知识。因为这时 A 的全部 $r+1$ 阶子式都为零，所以在展开式(2-1)中便有以下结果：

$$|\lambda I-A|=\lambda^n-(a_{11}+\cdots+a_{nn})\lambda^{n-1}+\cdots+(-1)^r(\sum S_r)\lambda^{n-r}$$
$$=\lambda^{n-r}(\lambda^r-(a_{11}+\cdots+a_{nn})\lambda^{r-1}+\cdots+(-1)^r(\sum S_r))。$$

它表明零特征值至少为 $n-r$ 重。同时可知道 $n-r$ 恰是零特征值对应独立特征向量的个数。因而这里表示对零特征值而言，几何重数不会超过它的代数重数。这个结论对一般非零特征值都是对的。

命题1 $n_i\geqslant g_i$（或可写成 $n_i\geqslant n-r(\lambda_i I-A)$）。

证 下面所用的方法是第一章所介绍的"将独立向量组扩充

为基"的方法，实际上设 λ_i 对应的特征向量为 $\alpha_1\cdots\alpha_k$（即 $g_i=k$），这样 $A\alpha_j=\lambda_i\alpha_j$，$j=1,2\cdots k$。将向量 $\alpha_1\cdots\alpha_k$ 扩充为 F^n 的基，设为 $\{\alpha_1\cdots\alpha_k,\alpha_{k+1}\cdots\alpha_n\}$。记 P 为矩阵 $[\alpha_1\cdots\alpha_k,\alpha_{k+1}\cdots\alpha_n]$，则

$$AP=A[\alpha_1\cdots\alpha_k,\alpha_{k+1}\cdots\alpha_n]=[A\alpha_1\cdots A\alpha_k,A\alpha_{k+1}\cdots A\alpha_n]$$
$$=[\lambda_i\alpha_1\cdots\lambda_i\alpha_k,A\alpha_{k+1}\cdots A\alpha_n]$$

$$=[\alpha_1\cdots\alpha_n]\begin{bmatrix}\lambda_i & & & * & \cdots & * \\ & \ddots & k & * & \cdots & * \\ & & \lambda_i & \vdots & & \vdots \\ & & & * & & *\end{bmatrix}=PB。$$

所以有 $P^{-1}AP=B$。矩阵 A 的特征多项式与 B 的相同，$|\lambda I-B|=(\lambda-\lambda_i)^k g(\lambda)$ 说明 A 的特征多项式至少有一个 $(\lambda-\lambda_i)^k$ 的因子，因此 λ_i 的代数重数至少为 k。▨

4. 关于特征值与特征向量的基本命题

命题 2 设 $\lambda_1\ \lambda_2\cdots\lambda_m$ 是算子 σ 的 m 个不同的特征值，它们对应的特征向量依次为 $\alpha_1\ \alpha_2\cdots\alpha_m$，则 $\alpha_1\ \alpha_2\cdots\alpha_m$ 是线性无关向量组。

命题3 设 λ_1，λ_2 是线性算子 σ 的两个不同特征值。$\{\alpha_1\cdots\alpha_{k_1}\}$，$\{\eta_1\cdots\eta_{k_2}\}$ 分别是对应于 λ_1,λ_2 的线性无关的特征向量组，则 $\{\alpha_1\cdots\alpha_{k_1},\eta_1\cdots\eta_{k_2}\}$ 是线性无关向量组。

以上是读者熟悉的二个命题，故证明略去。下面利用"任何矩阵都相似于上三角形矩阵"这个结论，证明二个重要的命题。

命题 4 $f(x)$ 是多项式，A 的 n 个特征值为 $\lambda_1\cdots\lambda_n$，则 $f(A)$ 的 n 个特征值为 $f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)$。

证 由第一章例 25 可知，一定存在一个非奇异矩阵 P，$P^{-1}AP=B$，这里 B 是上三角阵。因此 $P^{-1}A^kP=B^k$，所以有 $P^{-1}f(A)P=f(B)$，这说明 $f(A)$ 的特征值与 $f(B)$ 的相同。B 是上三角阵，因此 B^k 是上三角阵，同理 $f(B)$ 也是上三角阵。另一方面上三角阵的特征值为主对角线的对角元。如

$$B=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & * \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}, \quad B^k=\begin{bmatrix}\lambda_1^k & & & \\ & \lambda_2^k & & * \\ & & \ddots & \\ & & & \lambda_n^k\end{bmatrix}.$$

$$\text{因此}\ f(B)=\begin{bmatrix}f(\lambda_1) & & & \\ & f(\lambda_2) & & * \\ & & \ddots & \\ & & & f(\lambda_n)\end{bmatrix}。$$

（这里运用了上三角阵相乘，乘积的对角元为因子对角元对应相乘。）这便证明了 $f(A)$的 n 个特征值恰为$f(\lambda_1),f(\lambda_2)\cdots f(\lambda_n)$ ▨

上面的命题是一个很有用的命题，它说明如果知道矩阵 A 的各个特征值，便容易计算矩阵多项式的特征值。例如 A 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，则 A^2+2A 的特征值为 $\lambda_1^2+2\lambda_1,\lambda_2^2+2\lambda_2,\cdots,\lambda_n^2+2\lambda_n$，$I+A$ 的特征值为 $1+\lambda_1 1+\lambda_2,\cdots 1+\lambda_n$，等等。

命题 5(Hamilton-Cayley) 设 A 的特征多项式为 $f(\lambda)$，则有 $f(A)=0$。

证 设 $f(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$，

因 $$f(A)=Pf(B)P^{-1},$$

所以有 $$f(A)=Pf(B)P^{-1}=P(B-\lambda_1 I)(B-\lambda_2 I)\cdots(B-\lambda_n I)P^{-1}。$$

要证明 $f(A)=0$，只要证明 $f(B)=0$ 即可。

因有 $$(B-\lambda_1 I)(B-\lambda_2 I)\cdots(B-\lambda_n I)$$

$$=\begin{bmatrix}0 & * & \cdots & \cdots & * \\ 0 & \lambda_2-\lambda_1 & & & * \\ 0 & 0 & \lambda_3-\lambda_1 & & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & \lambda_n-\lambda_1\end{bmatrix}$$

$$
\begin{bmatrix} \lambda_1-\lambda_2 & * & * & \cdots & * \\ 0 & 0 & * & & * \\ 0 & 0 & \lambda_3-\lambda_2 & & * \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & & \lambda_n-\lambda_2 \end{bmatrix}
\begin{bmatrix} \lambda_1-\lambda_3 & * & * & \cdots & * \\ 0 & \lambda_2-\lambda_3 & * & & * \\ 0 & 0 & 0 & & * \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & & \lambda_n-\lambda_3 \end{bmatrix} \cdots
\begin{bmatrix} \lambda_1-\lambda_n & * & * & \cdots & * \\ 0 & \lambda_2-\lambda_n & * & \cdots & * \\ 0 & 0 & \lambda_3-\lambda_n & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & 0 \end{bmatrix}
$$

可自左向右逐个相乘，每乘一个因子至少增加一列零向量，因而确实有 $f(B)=0$，从而 $f(A)=0$。▨

命题 6(Sylvester)　A 是 $m\times n$ 型而 B 是 $n\times m$ 型，$m\geqslant n$。又设 AB 与 BA 的特征多项式分别为 f_{AB} 与 f_{BA}，则

$$f_{AB}=\lambda^{m-n}f_{BA}。$$

这个命题在第一章例 9 已经证过，这里从略。

由这个命题可得如下结论：

(i) 矩阵 AB 与 BA 的非零特征值是相同的；

(ii) 如果 $m=n$，则 AB 与 BA 的特征值相同。

第二节　特殊矩阵的特征值以及命题的应用

1. 一些特殊矩阵的特征值

要确定一个矩阵的特征值，实际上并不简单，但在理论上与应

用上一些重要的特殊矩阵的特征值，还是可从特征值的定义以及特殊矩阵的性质得出。下面便是一些主要的结果。

(1) $A^2=A$（幂等阵），A 的特征值仅可为 1，0。

证 设 λ 是 A 的一个特征值，而 α 是对应的特征向量。则 $A\alpha=\lambda\alpha$ 可得

$$A(A\alpha)=A(\lambda\alpha)\Rightarrow A^2\alpha=\lambda A\alpha\Rightarrow A^2\alpha=\lambda^2\alpha\Rightarrow A\alpha=\lambda^2\alpha。$$

所以有 $$(\lambda^2-\lambda)\alpha=0。$$

由于 $\alpha\neq0$，所以 $\lambda^2-\lambda=0$，即 $\lambda=1$ 或 0。

(2) $A^2=I$（对合阵），则 A 的特征值仅可为 $1,-1$。

用类似 (1) 的方法可证。

(3) $A^p=0$（幂零阵。p 是某一正整数，使该等式成立的最小正整数称为幂零指数），则矩阵 A 的特征值为零。

用类似 (1) 的方法可证。

(4) 实对称阵的特征值是实数。

证 设 A 是实对称阵，λ 是它的一个特征值，对应的特征向量为 v，因此 $Av=\lambda v$。两边取共轭转置（用记号 * 表示），便有 $v^*A=\bar{\lambda}v^*$，所以 $v^*Av=\bar{\lambda}v^*v$。由原式可得 $v^*Av=\lambda v^*v$，因此 $\lambda v^*v=\bar{\lambda}v^*v$，又因 $v^*v\neq0$，便可得 $\lambda=\bar{\lambda}$。▨

(5) 实反对称阵的特征值是纯虚数。

用类似 (4) 的方法证明。

(6) 酉阵的特征值的模为 1（因而正交阵亦有同样性质）。

用类似(4)的方法证明。

2. Hamilton-Cayley 定理的应用

设矩阵 A 的特征多项式为：

$$f(\lambda)=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0。$$

Hamilton-Cayley 定理表明：$A^n+a_nA^{n-1}+\cdots+a_1A+a_0I=0$，即 A^n 可通过 $A^{n-1},\cdots,A,I$ 的线性组合表示。因而 A 的任何幂次都可通过 $A^{n-1},A^{n-2},\cdots,A,I$ 的线性组合表示，下面举例说明它的应用。

例 5 已知 $A=\begin{bmatrix}3 & 2\\1 & 2\end{bmatrix}$，求 A^2, A^3, A^4。

解 A的特征多项式 $f(\lambda)=\lambda^2-5\lambda+4$，

则 $$A^2-5A+4I=0。$$

因而 $A^2=5A-4I, A^3=5A^2-4A=21A-20I,$

$$A^4=21A^2-20A=85A-84I。$$

例 6 已知 $$A=\begin{bmatrix}2 & -1 & -2\\-1 & 2 & 2\\0 & 0 & 1\end{bmatrix},$$

求 $g(A)=A^8-9A^6+A^4-3A^3+4A^2+I$。

解 A 的特征多项式 $f(\lambda)=|\lambda I-A|=(\lambda-1)^2(\lambda-3)$，又

$$g(\lambda)=\lambda^8-9\lambda^6+\lambda^4-3\lambda^3+4\lambda^2-1。$$

将 $g(\lambda)$被 $f(\lambda)$除，设商为 $p(\lambda)$。便可得等式

$$g(\lambda)=p(\lambda)f(\lambda)+32\lambda^2-107\lambda+68。$$

由于 $f(A)=0$，因此

$$g(A)=32A^2-107A+68I$$

$$=32\begin{bmatrix}5 & -4 & -8\\-4 & 5 & 8\\0 & 0 & 1\end{bmatrix}-107\begin{bmatrix}2 & -1 & -2\\-1 & 2 & 2\\0 & 0 & 1\end{bmatrix}$$

$$+68\begin{bmatrix}1 & & \\ & 1 & \\ & & 1\end{bmatrix}=\begin{bmatrix}14 & -21 & -42\\-21 & 14 & 42\\0 & 0 & -7\end{bmatrix}。$$

如果不做多项式的除法，可采用待定系数方法，可先设

$$g(\lambda)=p(\lambda)f(\lambda)+a\lambda^2+b\lambda+c,$$

令 $\lambda=3$， $g(3)=9a+3b+c=35$，

令 $\lambda=1$， $g(1)=a+b+c=-7$，

$$g'(1)=2a+b=-43。$$

解这方程组便可得 $a=32$，$b=-107$，$c=68$。即可得同样的结果。

例 7 设 3 阶方阵 A 的特征值为 $\lambda=1,-1,2$，试将 A^{2n} 表示为 A 的二次多项式。

解 A 的特征多项式为 $f(\lambda)=(\lambda-1)(\lambda+1)(\lambda-2)$。

令 $\lambda^{2n}=p(\lambda)f(\lambda)+a\lambda^2+b\lambda+c$，将 $\lambda=1,-1,2$ 代入上式。依次得

$$1=a+b+c,1=a-b+c,2^{2n}=4a+2b+c,$$

由此解得

$$a=\frac{1}{3}(2^{2n}-1),\ b=0,\ c=\frac{1}{3}(4-2^{2n})。$$

因此

$$A^{2n}=\frac{1}{3}(2^{2n}-1)A^2+\frac{1}{3}(4-2^{2n})I。$$

当一个矩阵A它的特征多项式常数项不为零，说明$|A|\neq0$，因而A是非奇异的，Hamilton-Cayley 定理表明 A^{-1} 亦可用 A^{n-1}，A^{n-2}，$\cdots A$，I的线性组合表示，这也是求逆矩阵的一种方法。

例 8 已知 $A=\begin{bmatrix}-4&5&5\\-5&6&5\\-5&6&6\end{bmatrix}$，求 A^{-1}。

解 因$|\lambda I-A|=\lambda^3-8\lambda^2+13\lambda-6$，因此 A 是可逆的。又

$$A^3-8A^2+13A-6I=0,$$

两边乘 A^{-1}，便可得表示式

$$A^{-1}=\frac{1}{6}(A^2-8A+13I)=\frac{1}{6}\begin{bmatrix}11&-5&-5\\5&1&-5\\5&-5&1\end{bmatrix}。$$

3. 关于特征多项式降阶计算的应用及其他

应用 Sylvester 定理(见前章例 9)时主要将一个 n 阶方阵 A 写成乘积的形式 $A=BC$，它们分别为 $n\times r$ 型与 $r\times n$ 型($n\geq r$)，对于 n 与 r 相差较大时特别有效。

例 9 求 n 阶实镜像阵 $I-2uu^T$ 的特征值及它的迹和行列

式。

解法 1 $|\lambda I-(I-2uu^T)|=|(\lambda-1)I+2uu^T|$

$$=(\lambda-1)^{n-1}|\lambda-1+2u^Tu|=(\lambda-1)^{n-1}(\lambda+1)。$$

因此得 $\lambda=1$ 是 $n-1$ 重根，而 $\lambda=-1$ 是一重根。

解法2 可先求 uu^T 的特征值，然后再按命题 4，求矩阵 $I-2uu^T$ 的特征值。

$$|\lambda I-uu^T|=\lambda^{n-1}|\lambda-u^Tu|=\lambda^{n-1}(\lambda-1)$$

因此 $\lambda=0$ 是 uu^T 的 $n-1$ 重特征值，而 $\lambda=1$ 是它的一重特征值。取多项式 $f(\lambda)=1-2\lambda$，$f(uu^T)$便是实镜像阵。所以 $f(0)=1$ 是 $n-1$ 重特征值，而 $f(1)=-1$ 是一重特征值。并且有

$$t_r(I-2uu^T)=n-2,\ |I-2uu^T|=\lambda_1\ \lambda_2\cdots\lambda_n=-1。$$

例 10 $n+1$ 阶矩阵 A 的特征值是零与 $\lambda_1\cdots\lambda_n$（$\lambda_1\cdots\lambda_n$ 是 n 个单位根），证明 $2I-A$ 是非奇异矩阵。

证 取多项式 $f(\lambda)=2-\lambda$，所证明的矩阵便是 $f(A)$。由命题 4 可知：$2I-A$ 的特征值是 $2, 2-\lambda_1\cdots2-\lambda_n$。只要证明它们均不为零，便证明了 $2I-A$ 是非奇异的。事实上 $|2-\lambda_i|\geqslant|2|-|\lambda_i|=1>0$，所以结论成立。

第三节 最小多项式

1. 矩阵的零化多项式

定义 A 是 n 阶方阵，$f(x)$是多项式。如果有 $f(A)=0$，则称 $f(x)$是 A 的零化多项式。由 Hamilton-Cayley 定理知道一个矩阵的零化多项式是存在的。

2. 方阵 A 的最小多项式

定义 A 的零化多项式中，次数最低的多项式称为 A 的最小多项式。

有关最小多项式的基本命题如下。

命题 7 如果 $m(x)$是矩阵 A 的最小多项式，又 $f(x)$是 A 的

任一零化多项式，则$m(x)\mid f(x)$。

证 利用带余除法可知 $f(x)=q(x)\cdot m(x)+r(x)$，其中$r(x)$的次小于 $m(x)$的次。由 $f(A)=q(A)m(A)+r(A)$，有 $r(A)=0$，因此 $r(x)$也是 A 的一个零化多项式。这与 $m(x)$ 是零化多项式中，次最低相矛盾。因而 $r(x)=0$，即 $m(x)\mid f(x)$。

例 11 试求下列 A 的最小多项式：

$$A=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}_{n\times n}。$$

解 由计算直接可知 $A^n=0$，但 $A^{n-1}\neq 0$，因此 A 的最小多项式为 $m(x)=x^n$，它恰好与特征多项式一样。

例12 试求下面分块矩阵的最小多项式。

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

解 A 的特征多项式为 $f(\lambda)=(\lambda-2)^6$。由命题7可知，$m(\lambda)=(\lambda-2)^k$，所以只要考察$(A-2I)$的乘积便可。但

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

它是由 3 个子块组成的分块对角阵，而每一个子块都是幂零阵。它们的指数分别为 1,2,3 。由分块对角阵的乘法可知，取$(\lambda-2)^3$便可有$(A-2I)^3=0$。

例 13 试求下列 A 的最小多项式：

$$A=\begin{bmatrix} \begin{matrix}2&1\\0&2\end{matrix} & & & \\ & \begin{matrix}2&1&0\\0&2&1\\0&0&2\end{matrix} & & \\ & & 3 & \\ & & & \begin{matrix}3&1\\0&3\end{matrix}\end{bmatrix}。$$

解 A 的特征多项式是 $(\lambda-2)^5(\lambda-3)^3$，因此 A 的最小多项式可选为 $(\lambda-2)^3(\lambda-3)^2$。研究矩阵 $A-2I$ 与 $A-3I$ 的乘积

$$A-2I=\begin{bmatrix} \begin{matrix}0&1\\0&0\end{matrix} & & \\ & \begin{matrix}0&1&0\\0&0&1\\0&0&0\end{matrix} & \\ & & *\end{bmatrix},$$

$$A-3I=\begin{bmatrix} * & & \\ & 0 & \\ & & \begin{matrix}0&1\\0&0\end{matrix}\end{bmatrix}。$$

上面 * 是一个非奇异矩阵。由分块对角阵的乘法可以知道，只要

注意使幂零矩阵的那部分变为零阵所需乘积的次数。$A-2I$ 只须 3 次，即在 $(A-2I)^3$ 中原幂零阵的那部分为零阵。同理 $A-3I$ 只需自乘 2 次便可达到目的。因而 A 的最小多项式为 $m(\lambda)=(\lambda-2)^3(\lambda-3)^2$。

在 Hamilton–Cayley 的应用中，可见特征多项式在计算矩阵多项式时的作用，应用了最小多项式有时会在计算上带来更多的方便。

例 14 一个三阶矩阵 A 它的特征多项式为 $f(\lambda)=\lambda^3-5\lambda^2$，而最小多项式 $m(\lambda)=\lambda^2-5\lambda$，试分别用它们求 A^4 的表达式。

解 由特征多项式有 $A^3-5A^2=0$，可计算得 $A^4=25A^2$；由最小多项式有 $A^2-5A=0$，可计算得 $A^4=125A$。由此可知用最小多项式会使计算简单一些。

命题 8 一个矩阵 A，如果规定最小多项式的首项系数为 1，这样最小多项式是唯一的。

证 设 $m_1(x)$，$m_2(x)$ 是 A 的二个最小多项式。由命题 7 可知 $m_1(x)|m_2(x)$，$m_2(x)|m_1(x)$。由于它们首项系数都是 1，因而 $m_1(x)=m_2(x)$。▨

命题 9 λ_0 是 A 的特征值 $\Longleftrightarrow$ λ_0 是 A 的最小多项式 $m(x)$ 的零点。

证 $\Longleftarrow$ 由命题 7 立即得出。下面仅证明 $\Longrightarrow$。设 λ_0 对应的特征向量为 α，因有 $0=m(A)\alpha=m(\lambda_0)\alpha$；又 $\alpha\neq 0$，所以 $m(\lambda_0)=0$。▨

这个命题可给出由特征多项式检验最小多项式的方法。

例15 求 $A=\left[\begin{array}{cc|cc} 2 & 1 & & \\ 0 & 2 & & \\ \hline & & 1 & 1 \\ & & -2 & 4 \end{array}\right]$ 的最小多项式。

解 A 的特征多项式 $f(\lambda)=(\lambda-3)(\lambda-2)^3$，根据命题 9 使

可知 A 的最小多项式有下面三种可能。

$$(\lambda-3)(\lambda-2),(\lambda-3)(\lambda-2)^2,(\lambda-3)(\lambda-2)^3。$$

$$(A-3I)(A-2I)=\begin{bmatrix}-1&1&0&0\\0&-1&0&0\\0&0&-2&1\\0&0&-2&1\end{bmatrix}\begin{bmatrix}0&1&0&0\\0&0&0&0\\0&0&-1&1\\0&0&-2&2\end{bmatrix}$$

$$=\begin{bmatrix}0&-1&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}\neq 0,$$

而 $(A-3I)(A-2I)^2=0$,因此最小多项式为 $m(x)=(x-3)(x-2)^2$。

命题 10 $P^{-1}AP=B$,则 A 的最小多项式与 B 的最小多项式相同。

证 设 A 的最小多项式为 $m_A(x)$,B 的最小多项式为 $m_B(x)$。

由 $P^{-1}AP=B$,可得 $m_B(B)=P^{-1}m_B(A)P$,但 $m_B(B)=0$,因此 $m_B(A)=0$,所以 $m_B(x)$ 是 A 的零化多项式,因而 $m_A|m_B$。同理 $m_A(x)$ 是 B 的零化多项式,$m_B|m_A$。但都是首项系数为 1 的多项式,因此 $m_A(x)=m_B(x)$。▨

下面的定理说明:由一个方阵的最小多项式,可判断该方阵是否和对角阵相似。

定理 1 n 阶方阵 A 与对角阵相似 $\Longleftrightarrow$ A 的最小多项式没有重根。

证 $\Longrightarrow$ 设 $P^{-1}AP=D$,其中 D 是对角阵。因为 A 的最小多项式与 D 的相同,所以只要证明对角阵 D 的最小多项式是相异的一次因子的乘积。设

$$D=\begin{bmatrix}\lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_1 & & & & & & \\ & & & \lambda_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_2 & & & \\ & & & & & & \ddots & & \\ & & & & & & & \lambda_s & \\ & & & & & & & & \ddots \\ & & & & & & & & & \lambda_s\end{bmatrix},$$

其中 $\lambda_1\ \lambda_2\cdots\lambda_s$ 是 D 相异特征值，作多项式 $(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_s)$。可直接验证 $(D-\lambda_1 I)(D-\lambda_2 I)\cdots(D-\lambda_s I)=0$，因此 $m(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_s)$ 是 D 的最小多项式。

$\Longleftarrow$ 设 A 的最小多项式 $m(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_s)$。要证明 A 与对角阵相似。下面要证明任一个特征值 λ_i 的代数重数 n_i 等于它的几何重数 g_i，或者说要证明

$$n=\sum_{k=1}^{s}n_i=\sum_{i=1}^{s}g_i=\sum_{i=1}^{s}(n-r(\lambda_i I-A))$$ 。如果记 $r_i=r(\lambda_i I-A)$，上式便是：$\sum\limits_{i=1}^{s}r_i=(s-1)n$。因为 $g_i\leqslant n_i$，所以 $\sum\limits_{i=1}^{s}g_i=\sum\limits_{i=1}^{s}(n-r_i)\leqslant\sum\limits_{i=1}^{s}n_i=n$ 或可写成 $\sum\limits_{i=1}^{s}r_i\geqslant(s-1)n$。

下面证明 $\sum\limits_{i=1}^{s}r_i\leqslant(s-1)n$。用 A_i 记矩阵 $A-\lambda_i I$，则 $0=m(A)=(A-\lambda_1 I)(A-\lambda_2 I)\cdots(A-\lambda_s I)=A_1A_2\cdots A_s$，引用结果(第一章命题 8)$r(AB)\geqslant r(A)+r(B)-n$。这样可有

$$\begin{aligned}0=r(m(A))=r(A_1\ A_2\cdots A_{s-1}\ A_s)&\geqslant r(A_1\cdots A_{s-1})+r_s-n\\ &\geqslant r(A_1\cdots A_{s-2})+r_{s-1}+r_s-2n\\ &\geqslant r_1+\cdots+r_s-(s-1)n。\end{aligned}$$

因此有

$$r_1+\cdots+r_s\leqslant(s-1)n。$$

说明对任何一个特征值λ_i，它的代数重数与几何重数是相等的。▨

例 16 矩阵 $A^2=A$ 与对角阵相似，矩阵 $A^2=E$ 与对角阵相似。

证 因为 $A^2=A$，因此多项式 $f(x)=x^2-x$是 A 的零化多项式，但 A 的最小多项式 $m(x)$一定除尽 $f(x)$，可以肯定 $m(x)$ 一定是相异的一次因子乘积，所以 A 与对角阵相似，同理证明如果矩阵满足条件 $A^2=I$ 亦一定与对角阵相似。

第四节 特征值的圆盘定理

计算一个矩阵的特征值并非容易，因此在应用上便有估计特征值的需要。复数域上的 n 阶矩阵 A，它的特征值可用复平面上的点来表示，对于它们位置的估计就是特征值估计问题。这个问题在实用上与理论上都有很大价值。如在自动控制中，常常要估计特征值是否在复平面的左半部，即特征值是否有负实部。又在方程组的迭代法里，常要确定特征值是否在单位圆内。另外实对称阵的特征值的估计也是一个很重要的课题。这一节介绍的圆盘定理是一个便于应用的定理，它是一个估计特征值的重要定理。

1. 圆盘定理(Gershgorin)

设 A 是一个 n 阶的复矩阵，集合 $D_i(A)$

$$D_i(A)=\{x\,|\,|x-a_{ii}|\leqslant\sum_{j\neq i}|a_{ij}|\},i=1,2,\cdots,n。$$

称为矩阵 A 的第 i 个圆盘。

例17 求 $A=\begin{bmatrix}0 & 1 & -1\\ 1.3 & 2 & -0.7\\ 0.5 & 0.5i & 4i\end{bmatrix}$的圆盘。

解 因有 $D_1(A)=\{x\,|\,|x|\leqslant2\}$，$D_2(A)=\{x\,|\,|x-2|\leqslant2\}$，$D_3(A)=\{x\,|\,|x-0.4i|\leqslant1\}$，如果用图表示则为如图 2-1 所示。

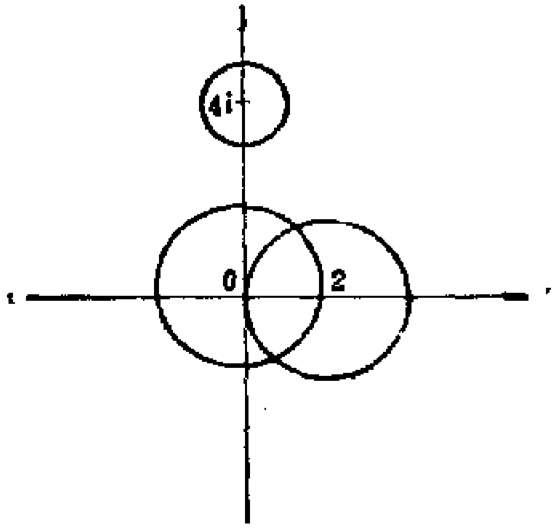

图 2-1

定理 2 A 是 n 阶复方阵，则它的特征值至少满足下列不等式之一。

$$|\lambda - a_{ii}| \leqslant \sum_{\substack{j=1\\j\neq i}}^{n} |a_{ij}|, \quad (i=1,2,\cdots,n)。$$

换句话说 A 的每一个特征值落在 A 的某一个圆盘之内。

证 设 λ 是 A 的一个特征值，$\alpha=(x_1\cdots x_n)^T$ 是它对应的特征向量。因此 $A\alpha=\lambda\alpha$，将它写成方程组的形式便是：

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = \lambda x_1\\
&a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = \lambda x_2,\\
&\vdots\\
&a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = \lambda x_n。
\end{aligned}$$

令 $\max|x_j| = |x_m|$，因而 $x_m \neq 0, j=1,2,\cdots,n$。考察第 m 个方程为

$$(\lambda - a_{mm})x_m = \sum_{\substack{j\neq m\\j=1}}^{n} a_{ij}x_j。$$

两边取绝对值为

$$|\lambda - a_{mm}||x_m| \leqslant \sum_{\substack{j=1\\j\neq m}}^{n} |a_{ij}||x_j| \leqslant |x_m| \sum_{\substack{j=1\\j\neq m}}^{n} |a_{ij}|,$$

所以有

$$|\lambda - a_{mm}| \leqslant \sum_{\substack{j=1\\ j\neq m}}^{n} |a_{ij}|。$$

对这个特征值，它落在第 m 个圆盘内。▨

2. 圆盘定理应用举例

例18 求 $A=\begin{bmatrix} 1 & 0.1 & 0.2 & 0.3 \\ 0.5 & 3 & 0.1 & 0.2 \\ 1 & 0.3 & -1 & 0.5 \\ 0.2 & -0.3 & -0.1 & -4 \end{bmatrix}$ 的圆盘。

解 矩阵 A 的圆盘为

$$|\lambda-1| \leqslant 0.1+0.2+0.3=0.6,$$

$$|\lambda-3| \leqslant 0.1+0.5+0.2=0.8,$$

$$|\lambda+1| \leqslant 1+0.3+0.5=1.8,$$

$$|\lambda+4| \leqslant 0.2+0.3+0.1=0.6。$$

见图 2-2 所示。

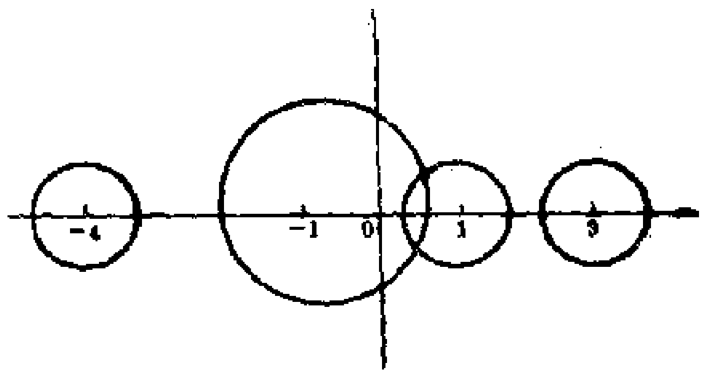

图2-2

例19 n 阶方阵 A 满足对角强优条件

$$|a_{ii}| > \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|,\ (i=1,2,\cdots,n),$$

证明 $|A| \neq 0$。

证 主要证明 A 的特征值全不为零。设 λ 是 A 的一个特征

值，它必然落在某一个圆盘之内。比如第 k 个圆盘，$|\lambda-a_{kk}|\leqslant\sum_{\substack{j=1\\j\neq k}}^{n}|a_{kj}|$。如果 $\lambda=0$，则有 $|a_{kk}|\leqslant\sum_{\substack{j=1\\j\neq k}}^{n}|a_{kj}|$，这与前提相矛盾。所以 A 不可能有零特征值，因此 A 的行列式不为零。

在应用时，要正确理解圆盘定理。它只断言特征值至少满足下述不等式之一：$|\lambda-a_{ii}|\leqslant\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|$，$(i=1,2,\cdots,n)$，但并没有肯定在哪一个，因此实际上有的圆盘，可能并不含特征值，如下例。

例20 求 $A=\begin{bmatrix}1 & -0.8\\0.5 & 0\end{bmatrix}$ 的圆盘。

解 该矩阵的特征多项式为 $\lambda^2-\lambda+0.4$，特征值可直接计算得

$$\lambda_1=\frac{1}{2}+\frac{1}{2}\sqrt{0.6}\,i,\quad \lambda_2=\frac{1}{2}-\frac{1}{2}\sqrt{0.6}\ \mathrm{i}。$$

则
$$|\lambda_1|^2=|\lambda_2|^2=\frac{1}{4}+\frac{0.6}{4}=0.4,$$

因此 $|\lambda_1|=|\lambda_2|=0.632$，$A$ 的圆盘为 $|\lambda-1|\leqslant0.8$，$|\lambda|\leqslant0.5$。

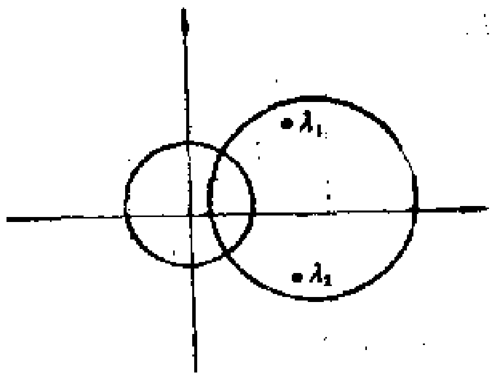

图 2-3

上述两个特征值全部落在 $|\lambda-1|\leqslant0.8$ 内，而在 $|\lambda|\leqslant0.5$ 之外！但它可认为特征值落在这两个圆盘的并集之内，而这两个圆盘恰

构成一个连通部分。

实际上可证明下面的一般命题。

命题11 在圆盘组成的连通部分任取一个。如果它是由 k 个圆盘组成，那么这个连通部分必有且只有 k 个特征值。（证略）

如例 28，A 的圆盘共有四个，它们构成三个连通部分（见图 2-2）：(i) $|\lambda+4|\leqslant 0.6$，(ii) $|\lambda+1|\leqslant 1.8$；$|\lambda-1|\leqslant 0.6$，(iii) $|\lambda-3|\leqslant 0.8$。因此可断言在第二个连通部分含有二个特征值，而在第一、三个由单个圆盘组成的连通部分各含一个特征值。而且由于 A 的特征多项式是实系数。还可进一步断言它们都是实的特征值。

为了要得到较有效的估计，利用相似变换下矩阵的特征值不变这一性质。对矩阵施行相似变换，再对变换后的矩阵应用圆盘定理。

例21 求$A=\begin{bmatrix}0.9 & 0.01 & 0.12\\ 0.01 & 0.8 & 0.13\\ 0.01 & 0.02 & 0.4\end{bmatrix}$的圆盘。

由圆盘定理可知 A 的特征值在下列圆盘之中：$|\lambda-0.9|\leqslant 0.13$，$|\lambda-0.8|\leqslant 0.14$，$|\lambda-0.4|\leqslant 0.03$。用图表示上述三个圆盘可知第一、二两个圆盘构成一连通部分，而第三个圆盘单独构成一个圆盘。如果取矩阵

$$S=\begin{bmatrix}1 & & \\ & 1 & \\ & & \frac{1}{10}\end{bmatrix},$$

作相似变换 $S^{-1}AS=B$，可得矩阵

$$B=\begin{bmatrix}0.9 & 0.01 & 0.012\\ 0.01 & 0.8 & 0.013\\ 0.1 & 0.2 & 0.4\end{bmatrix},$$

B 的圆盘为

$$|\lambda-0.9|\leqslant 0.022,\ |\lambda-0.8|\leqslant 0.023,\ |\lambda-0.4|\leqslant 0.3。$$

这三个圆盘都是孤立的，因而每一个圆盘都有一个特征值。综上所述，矩阵 A 的特征值在下列三个圆盘之中

$$|\lambda-0.9|\leqslant 0.022,\ |\lambda-0.8|\leqslant 0.023,\ |\lambda-0.4|\leqslant 0.03,$$

最后利用圆盘定理对矩阵的谱半径作出估计。首先定义一个矩阵的谱。

定义 A 是 n 阶矩阵，它的特征值全体称为矩阵 A 的谱，用 $\lambda(A)$ 表示该集合。

例22 $A=\begin{bmatrix}1 & 0 & -2\\ 0 & 0 & 0\\ -2 & 0 & 4\end{bmatrix}$，它的特征多项式为 $\lambda^3-5\lambda^2$，因而 $\lambda(A)=\{0,0,5\}$。

定义 设复数域上的 n 阶矩阵 $A=[a_{ij}]$ 的谱 $\lambda(A)=\{\lambda_1\cdots\lambda_n\}$，则称 $\max|\lambda_i|(1\leqslant i\leqslant n)$ 为矩阵 A 的谱半径，用 $\rho(A)$ 表示。

从几何上看，谱半径 $\rho(A)$ 是表示在复平面上，以原点为中心，$\rho(A)$ 为半径的一圆。矩阵 A 的特征值全在这个圆内。谱半径在研究矩阵序列的收敛是一个重要的概念，下面便给出谱半径估计的命题。

命题12 设 $A=[a_{ij}]$ 是复数域上的一个 n 阶矩阵，而令 $\nu=\max\limits_{1\leqslant i\leqslant n}\sum\limits_{j=1}^{n}|a_{ij}|$（即矩阵行的元素的绝对值和的最大者），则 $\rho(A)\leqslant\nu$。

证 设 λ_0 是 A 的任一个特征值，由圆盘定理可知一定有一个 i 使得 $|\lambda_0-a_{ii}|\leqslant\sum\limits_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}|$ 成立。所以 $|\lambda_0|\leqslant|a_{ii}|+\sum\limits_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}|$ $=\sum\limits_{j=1}^{n}|a_{ij}|\leqslant\nu$。自然亦应有 $\rho(A)\leqslant\nu$ 成立。▨

因为 A 与 A^T 有相同的特征值。如果对 A^T 应用上面命题，便可知道用矩阵列的元素的绝对值之和的最大值。亦可对 A 的谱半径作出估计。因此令 $\nu'=\max\limits_{1\leqslant j\leqslant n}\sum\limits_{i=1}|a_{ij}|$，同样有 $\rho(A)\leqslant\nu'$.

或综合起来可得如下命题

命题 13 $\rho(A) \leqslant \min\{\nu, \nu'\}$。

第三章　内积空间与特殊矩阵

在学习线性空间的时候，仅涉及向量的加法、数乘这两种运算，讨论了向量的相关性、独立性、向量组的秩和线性空间的基，并且利用这些概念解决了线性方程组的基本理论。但在很多应用的场合还需要有度量的概念，如矛盾方程的近似解、矩阵级数收敛性的讨论等，都涉及度量的概念。本章要在线性空间中，引入内积的概念。通过它定义向量的长度、正交性、构造空间的标准正交基。并在后半部分讨论与内积有关，且在应用上重要的矩阵。

第一节　内　　积

1. 与三维向量数积有关的若干几何事实

在学习向量代数时，曾学习过二个向量的数积为：$\alpha\cdot\beta=|\alpha||\beta|\cos\phi$。$|\alpha|$表示向量$\alpha$的长度，如果选定了直角系 i, j, k，则 $\alpha=x_1\mathrm{i}+x_2\mathrm{j}+x_3\mathrm{k}$，$\beta=y_1\mathrm{i}+y_2\mathrm{j}+y_3\mathrm{k}$，两个向量$\alpha,\beta$的数积便可由下面的公式计算：$\alpha\cdot\beta=x_1y_1+x_2y_2+x_3y_3$。利用这个公式可进行下面的几何问题的计算。

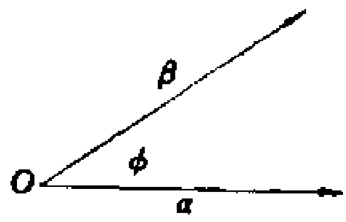

图 3-1

(i) 向量 α 的长度为：$|\alpha|^2=\alpha\cdot\alpha=x_1^2+x_2^2+x_3^2$。

(ii) 向量 α 与 β 间的夹角，$\cos\phi=\dfrac{\alpha\cdot\beta}{|\alpha||\beta|}$，并可知，$\alpha\cdot\beta=0$ 表示两个向量是垂直的。

(iii) 计算向量 β 在向量 α 上的投影与投影向量（如图 3-2、

图 3-3 所示)。

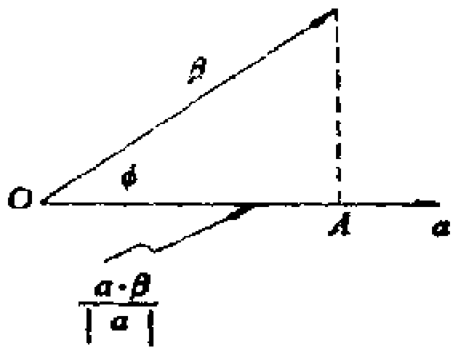

图 3-2

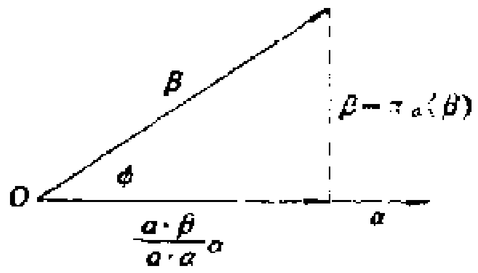

图 3-3

β 在 α 上的投影为 $|\beta|\cos\phi=\frac{\alpha\beta}{|\alpha|}$。

β 在 α 上的投影向量为 $\frac{\alpha\cdot\beta}{|\alpha|}\cdot\frac{\alpha}{|\alpha|}$，或可写成 $\frac{\alpha\cdot\beta}{\alpha\cdot\alpha}\alpha$。

以后用 $\prod_{\alpha}(\beta)$ 表示。

(iv) $\prod_{\alpha}(\beta)$ 有一个重要的性质：$\beta-\prod_{\alpha}(\beta)$ 一定与 α 垂直。

2. 向量的数积

要将几何向量的数积推广到抽象的线性空间的向量。为此，先将上面两个向量的数积在代数运算方面的性质归纳一下，重要的性质有下面四条：

(i) $\alpha\cdot\beta=\beta\cdot\alpha$　　对称性，

(ii) $(\alpha+\beta)\cdot\gamma=\alpha\cdot\gamma+\beta\cdot\alpha$　　对第一个因子有线性性质，

(iii) $(k\alpha)\cdot\beta=k(\alpha\cdot\beta)$　　齐次性，

(iv) $\alpha\cdot\alpha\geqslant 0$ (“$=$”仅在 $\alpha=0$ 时成立)正定性。

现对实线性空间任意两个向量的数积(以后称为内积)定义如下。

定义　设 V 是实域上的线性空间，它的内积是指这样的一个函数，对 V 的任意两个向量 α、β 对应着一个实数，记为 (α,β)，并且这个对应要满足如下四条公理——内积公理。

(i) $(\alpha,\beta)=(\beta,\alpha)$，

(ii) $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$,

(iii) $(k\alpha,\beta)=k(\alpha,\beta)$,

(iv) $(\alpha,\alpha)\geqslant 0$ （“$=$”仅在 $\alpha=0$ 成立）。

满足上面内积公理的函数(α,β)，称为实线性空间的内积。在实线性空间定义了内积之后，称此空间为 Euclid 空间，简称为欧氏空间。

3. 内积的例

例 1 $R^n=\{\alpha=(x_1\cdots x_n)^T|x_i\in R\}$，如果记 $\alpha=(x_1\cdots x_n)^T$，$\beta=(y_1\cdots y_n)^T$。定义 α,β 的内积如下

$$(\alpha,\beta)=\beta^T\alpha=x_1y_1+x_2y_2+\cdots+x_ny_n。\tag{3-1}$$

容易验证由式(3-1)定义的函数(α,β)确实满足内积公理。

例 2 $R^{n\times n}$ 是全体 n 阶实方阵组成的集合，它是实域上的线性空间。如果将每个矩阵视为 n^2 维的向量，便可仿照式(3-1)的方式，定义两个矩阵的内积：$A=[a_{ij}]$，$B=[b_{ij}]$，则

$$(A,B)=\sum_{i=1}^{n}\sum_{k=1}^{n}a_{ik}b_{ik}。$$

如果将它写成紧凑的形式便是 $tr(B^T,A)$，换句话说，可将两个矩阵的内积定义为：

$$(A,B)=tr(B^T,A)。$$

例 3 $C_{[a,b]}$ 是$[a,b]$上由实变量连续函数构成的线性空间。$f(x),g(x)\in C_{[a,b]}$，定义它们的内积为

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

容易验证(f,g)确实满足内积公理。

4. 内积的基本性质

现在讨论一下由内积公理直接得到的一些性质。

命题 1 若 V 是欧氏空间，$\alpha,\beta,\gamma\in V$，$c\in R$，则有

(i) $(\alpha,\beta+\gamma)=(\alpha,\beta)+(\alpha,\gamma)$,

(ii) $(\alpha,c\beta)=C(\alpha,\beta)$（由此可得 $(\alpha,0)=0$），

(iii) $(\alpha,\beta)=(\alpha,\gamma)$对任何 $\alpha\in V$ 均成立，则 $\beta=\gamma$。

证　仅证(iii)。如果 $\beta\neq\gamma$，令向量 $\alpha=\beta-\gamma$，则
$0=(\alpha,\beta)-(\alpha,\gamma)=(\beta-\gamma,\beta)-(\beta-\gamma,\gamma)=(\beta-\gamma,\beta-\gamma)>0$ 自相矛盾，因此 $\beta=\gamma$。▨

有了内积之后，便可定义向量的长度或模。

定义　V 是欧氏空间，$\alpha\in V$，它的长度或模(用 $\|\alpha\|$ 表示)定义为 $\|\alpha\|=(\alpha,\alpha)^{\frac{1}{2}}$。

命题 2　$|(\alpha,\beta)|\leqslant\|\alpha\|\|\beta\|$。(Cauchy)

证　首先可假设 $\alpha\neq0$(因为在 $\alpha=0$ 时，不等式显然成立)，作向量 $\beta-\dfrac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha$(它的意义见图 3-4)。

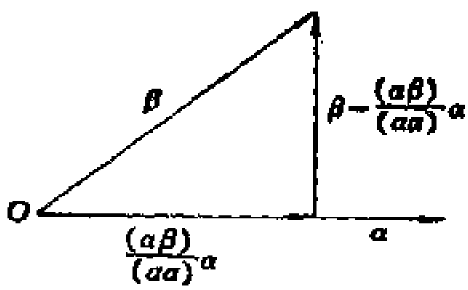

图 3-4

由内积公理可知

$$0\leqslant\left(\beta-\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha,\beta-\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha\right)$$

$$=(\beta,\beta)-\left(\beta,\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha\right)-\left(\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha,\beta\right)$$

$$+\left(\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha,\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha\right)=(\beta,\beta)-\frac{(\alpha,\beta)}{(\alpha,\alpha)}(\beta,\alpha)$$

$$-\frac{(\alpha,\beta)}{(\alpha,\alpha)}(\alpha,\beta)+\frac{(\alpha,\beta)^2}{(\alpha,\alpha)^2}\cdot(\alpha,\alpha)=(\beta,\beta)-\frac{(\alpha,\beta)^2}{(\alpha,\alpha)}$$

因此有

$$(\alpha,\alpha)(\beta,\beta)\geqslant(\alpha,\beta)^2。$$

两边开方即可得出结果。▨

下面的命题说明向量模的基本性质。

命题 3　在欧氏空间里向量模有如下性质：

(i)　$\|c\alpha\|=|c|\|\alpha\|$，

(ii)　$\|\alpha\|>0,\alpha\neq0$，$\|\alpha\|=0$，仅在 $\alpha=0$ 时成立，

(iii)　$\|\alpha+\beta\|\leqslant\|\alpha\|+\|\beta\|$　(三角不等式)，

证 仅证明(iii)。

$$\|\alpha+\beta\|^2=(\alpha+\beta,\alpha+\beta)=(\alpha,\alpha)+(\alpha,\beta)+(\beta,\alpha)+(\beta,\beta)$$
$$=\|\alpha\|^2+2(\alpha,\beta)+\|\beta\|^2\leqslant\|\alpha\|^2+2\|\alpha\|\|\beta\|+\|\beta\|^2$$
$$=(\|\alpha\|+\|\beta\|)^2$$

两边开方便可得出结果。 ▨

Cauchy 不等式是一个重要的不等式。前面三角不等式的成立，就用了 Cauchy 不等式。同时，如果定义二个向量的夹角为 $\cos\phi=\dfrac{(\alpha,\beta)}{\|\alpha\|\|\beta\|}$，这也因 Cauchy 不等式成立，这样定义才是合理的，下面对向量的距离下一个定义，实际上"向量"在应用上有二种几何的形象，其一是通常的几何向量，有长度与方向，常用一个带箭头的线段表示。向量另一个几何形象是通过位置矢量来确定空间的点(如图 3-2)。在空间确定一 O 点。空间任一点 A，与以 O 点为起点，终点在 A 的一个向量对应，因而二个向量的距离，在几何上可理解为两点的距离。

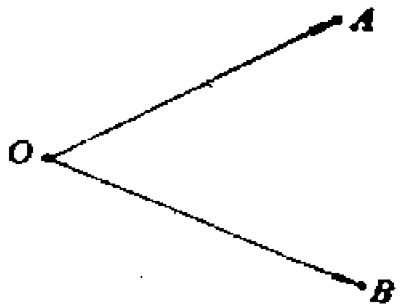

图 3-5

定义 V 是欧氏空间，$\alpha,\beta\in V$，它们之差的模 $\|\alpha-\beta\|$ 称为向量 α 到 β 的距离。

关于距离有如下的命题。

命题 4 在欧氏空间中两个向量的距离有如下性质

(i) $\alpha\neq\beta$，$\|\alpha-\beta\|>0$，

(ii) $\|\alpha-\beta\|=\|\beta-\alpha\|$，

(iii) $\|\alpha-\beta\|+\|\beta-\gamma\|\geqslant\|\alpha-\gamma\|$。

证 (ii) $\|\alpha-\beta\|^2=(\alpha-\beta,\alpha-\beta)=(\beta-\alpha,\beta-\alpha)$
$$=\|\beta-\alpha\|^2。$$

(iii) 如果令 $\alpha'=\alpha-\beta$，$\beta'=\beta-\gamma$，则 $\alpha'+\beta'=\alpha-\gamma$。这样上面(iii)的结果便是 $\|\alpha'\|+\|\beta'\|\geqslant\|\alpha'+\beta'\|$，即是三角不等

式。

第二节　正　交　性

1. 标准正交基的构造

这节主要给出空间存在有标准正交基的构造性证明。它起的作用和三维空间直角坐标系相当。欧氏空间里两个向量正交的定义如下。

定义　V 是欧氏空间，$\alpha,\beta\in V$，如果 $(\alpha,\beta)=0$，则称 α 与 β 正交，记为 $\alpha\perp\beta$。

构造正交基的思想如前所述：α,β 是两个线性无关的向量，先构造 $\prod_\alpha(\beta)=\dfrac{(\beta,\alpha)}{(\alpha,\alpha)}\alpha$（$\beta$ 在 α 上的投影向量），向量 $\beta-\prod_\alpha(\beta)=\beta-\dfrac{(\beta,\alpha)}{(\alpha,\alpha)}\alpha$ 必与 α 正交。

定义　设 $\varepsilon_1\cdots\varepsilon_m$ 是欧氏空间的一组向量。如果它们满足条件 $(\varepsilon_i,\varepsilon_j)=0(i\neq j)$ 便称为正交向量组。又如再有 $(\varepsilon_i,\varepsilon_i)=1$ 称为标准正交向量组。

正交向量的一个重要性质是：任何不含零向量的正交向量组是线性无关的。如果一个 n 阶实矩阵 A 有 $A^TA=I$ 称为正交阵。该等式表明正交阵的列向量构成标准正交向量组。

2. Gram-Schmidt 正交过程

定理 1　设 $\alpha_1\cdots\alpha_m$ 是欧氏空间的一组线性无关向量；则一定可以找到一组正交向量 $\varepsilon_1\cdots\varepsilon_m$，满足条件

$[\alpha_1]=[\varepsilon_1],[\alpha_1,\alpha_2]=[\varepsilon_1,\varepsilon_2],\cdots,[\alpha_1\cdots\alpha_k]=[\varepsilon_1\cdots\varepsilon_k],\cdots,$

$[\alpha_1\cdots\alpha_m]=[\varepsilon_1\cdots\varepsilon_m]$。

证　用数学归纳法，对向量的个数施行归纳证明。

$m=1$ 时结论显然成立。只要选 $\varepsilon_1=\alpha_1$，即有 $[\varepsilon_1]=[\alpha_1]$。假定 $m=k$ 时定理成立。（即有 $[\varepsilon_1]=[\alpha_1],[\varepsilon_1,\varepsilon_2]=[\alpha_1\alpha_2],\cdots,$

$[\alpha_1\cdots\alpha_k]=[\varepsilon_1\cdots\varepsilon_k]$，且 $\varepsilon_1\cdots\varepsilon_k$ 是正交向量组)。现要证明，$m=k+1$时，定理亦成立。实际上先作向量$\gamma=\frac{(\alpha_{k+1},\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1+\frac{(\alpha_{k+1},\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2+\cdots+\frac{(\alpha_{k+1},\varepsilon_k)}{(\varepsilon_k,\varepsilon_k)}\varepsilon_k$，由 γ 的表达式可以看出，它实际上是 α_{k+1} 在 $\varepsilon_1,\cdots,\varepsilon_k$ 各个向量上的投影向量之和，γ 称为 α_{k+1} 在子空间 $[\varepsilon_1\cdots\varepsilon_k]$的投影。然后作向量 $\varepsilon_{k+1}=\alpha_{k+1}-\gamma=\alpha_{k+1}-\frac{(\alpha_{k+1},\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1-\frac{(\alpha_{k+1},\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2-\cdots-\frac{(\alpha_{k+1},\varepsilon_k)}{(\varepsilon_k,\varepsilon_k)}\varepsilon_k$，可证明 ε_{k+1} 与 $\varepsilon_1\cdots\varepsilon_k$ 是正交的，即$(\varepsilon_{k+1},\varepsilon_j)=0,j=1,2,\cdots k$。亦即

$$\left(\alpha_{k+1}-\frac{(\alpha_{k+1},\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1-\frac{(\alpha_{k+1},\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2-\cdots-\frac{(\alpha_{k+1},\varepsilon_k)}{(\varepsilon_k,\varepsilon_k)}\varepsilon_k,\varepsilon_j\right)$$

$$=(\alpha_{k+1})\varepsilon_j)-\frac{(\alpha_{k+1})\varepsilon_j)}{(\varepsilon_j,\varepsilon_j)}\varepsilon_j,\varepsilon_j\Big)=0$$

由$[\alpha_1\cdots\alpha_k]=[\varepsilon_1\cdots\varepsilon_k]$以及 ε_{k+1} 的表达式可知有 $[\alpha_1\cdots\alpha_k\ \alpha_{k+1}]=[\varepsilon_1\cdots\varepsilon_k\ \varepsilon_{k+1}]$成立。 ▨

定理的证明过程是通过$\alpha_1\cdots\alpha_m$去逐步构造正交向量组$\varepsilon_1\cdots\varepsilon_m$，这方法称为 Gram-Schmidt 正交过程。它的基本思想实际上是：一旦作出子空间 $V_k=[\alpha_1\cdots\alpha_k]$的正交基 $\varepsilon_1\cdots\varepsilon_k$，要求出 ε_{k+1}，只要作出 α_{k+1} 在 V_k 上的投影 r，然后再作差 $\alpha_{k+1}-r$。即为 ε_{k+1}(如图 3-6)。

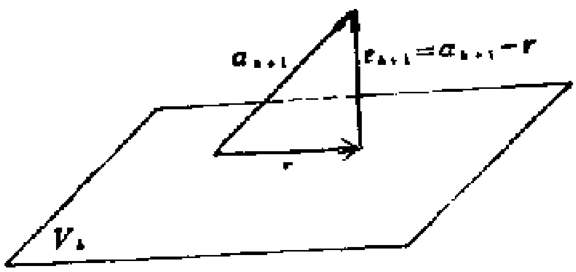

图 3-6

例 4 在欧氏空间 R^4 中求三个向量 $\alpha_1=(1\ \ 0\ \ 1\ \ 1)^T$，$\alpha_2=(2\ \ 1\ \ 0\ \ -3)^T$和 $\alpha_3(1\ \ -1\ \ 1\ \ -1)^T$所生成子空间的标准正交基。

解 所求的标准正交基为：

$$\varepsilon_1=\frac{\alpha_1}{\|\alpha_1\|}=\frac{1}{\sqrt{3}}(1\ \ 0\ \ 1\ \ 1)^T,$$

因有

$$\alpha_2'=\alpha_2-(\alpha,\varepsilon_1)\varepsilon_1=\frac{1}{3}(7\ \ 3\ \ 1\ \ -8)^T,$$

则

$$\varepsilon_2=\frac{\alpha_2'}{\|\alpha_2'\|}=\frac{1}{\sqrt{\frac{41}{3}}}(7\ \ 3\ \ 1\ \ -8)^T=\frac{3}{\sqrt{123}}(7\ \ 3\ \ 1\ \ -8)^T。$$

又因

$$\alpha_3'=\alpha_3-(\alpha_3,\varepsilon_1)\varepsilon_1-(\alpha_3,\varepsilon_2)\varepsilon_2=\frac{1}{41}(-3\ \ -54\ \ -23\ \ -20)^T,$$

则

$$\varepsilon_3=\frac{\alpha_3'}{\|\alpha_3'\|}=\frac{1}{\sqrt{3854}}(-3\ \ -54\ \ -23\ \ -20)^T。$$

且有$[\alpha_1\ \ \alpha_2\ \ \alpha_3]=[\varepsilon_1\ \ \varepsilon_2\ \ \varepsilon_3]$。

例 5 $P_2(x)$ 是表示次不超 2 的实系数多项式全体所构成的线性空间，它的基为 $\alpha_1=1,\alpha_2=x,\alpha_3=x^2$，试求 $P_2[x]$ 的正交基。(假定定义域为 [0 1] 区间)

解 这空间的内积定义为$\int_0^1 P_1(x)P_1(x)\mathrm{d}x$，因而向量的长度或模为 $\|P_1\|=\int_0^1 P_1^2(x)\mathrm{d}x$。所求的正交基为：

$$\varepsilon_1=\alpha_1=1,$$

$$\varepsilon_2=\alpha_2-\frac{(\alpha_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1=x-\frac{\int_0^1 1\cdot x\mathrm{d}x}{\int_0^1 1\,\mathrm{d}x}=x-\frac{1}{2},$$

$$\varepsilon_3 = \alpha_3 - \frac{(\alpha_3, \varepsilon_1)}{(\varepsilon_1, \varepsilon_1)}\varepsilon_1 - \frac{(\alpha_3, \varepsilon_2)}{(\varepsilon_2, \varepsilon_2)}\varepsilon_2 = x^2 - \frac{\int_0^1 1 \cdot x^2 \, dx}{\int_0^1 1 \cdot dx}$$

$$- \frac{\int_0^1 x^2\left(x - \frac{1}{2}\right)dx}{\int_0^1 \left(x - \frac{1}{2}\right)^2 dx}\left(x - \frac{1}{2}\right)$$

$$= x^2 - x + \frac{1}{6}。$$

第三节　Gram-Schmidt 正交过程的应用

这节主要讲一下 Gram-Schmidt 正交过程的应用。由Gram-Schmidt 正交过程可以断言欧氏空间的任一个正交向量组都可扩充为空间的正交基，事实上可设 $\varepsilon_1 \cdots \varepsilon_k$ 是正交向量组．先将它扩充为空间的基 $\varepsilon_1 \cdots \varepsilon_k \alpha_{k+1} \cdots \alpha_n$，然后对上面的基，用 Gram-Schmidt 正交过程去构造正交基。由于 $\varepsilon_1 \cdots \varepsilon_k$ 已是正交向量组，所以构造过程 $\varepsilon_1 \cdots \varepsilon_k$ 便保留下来。因而新向量组必然是 $\varepsilon_1 \cdots \varepsilon_k$ $\varepsilon_{k+1} \cdots \varepsilon_n$。这样便可引入关于一个子空间的正交补空间的概念。

1．正交补空间

定义　V_s 是欧氏空间 V 的一个子空间，下面的集合

$$\{\alpha \mid (\alpha, \beta) = 0, \forall \beta \in V_s, \alpha \in V\} \qquad (3\text{-}2)$$

称为 V_s 的正交补空间，(这名称的合理性可解析如下：集合(3-2)是由与 V_s 中所有向量正交的向量组成，容易证明该集合是 V 的一个子空间。)记为 $V_s^{\perp}$。

首先 $V_s^{\perp}$ 是存在的。因为如果用 $\varepsilon_1 \cdots \varepsilon_k$ 表示 V_s 的一组正交基，只要将它扩充为 V 的正交基 $\varepsilon_1 \cdots \varepsilon_k, \varepsilon_{k+1} \cdots \varepsilon_n$，就可证明如下命题。

命题 5　$V_s^{\perp} = [\varepsilon_{k+1} \cdots \varepsilon_n]$。

证 显然$[\varepsilon_{k+1}\cdots\varepsilon_n]\subseteq V_s^{\perp}$。另一方面空间任一向量$\varepsilon=x_1\varepsilon_1+\cdots+x_k\varepsilon_k+x_{k+1}\varepsilon_{k+1}+\cdots+x_n\varepsilon_n$。如有$\varepsilon\in V_s^{\perp}$，则由$(\varepsilon,\varepsilon_i)=0,i=1\ 2\cdots k$，可得$x_i=0$，$i=1\ 2,\cdots,k$，所以$\varepsilon=x_{k+1}\varepsilon_{k+1}+\cdots+x_n\varepsilon_n$。这说明$V_s^{\perp}\subseteq[\varepsilon_{k+1}\cdots\varepsilon_n]$。

这个命题说明$V_s^{\perp}$的存在性及如何构造，另一方面由$V_s^{\perp}$的构造过程可得下面的重要定理。

定理 2 V是欧氏空间，V_s是它的一个子空间，则

$$V=V_s\oplus V_s^{\perp}。$$

证 由$V_s^{\perp}$的构造可以知道$V=V_s+V_s^{\perp}$，余下只要证明$V_s\cap V_s^{\perp}=\{0\}$。如果有$\alpha\in V_s\cap V_s^{\perp}\Rightarrow(\alpha,\alpha)=0$ $(\alpha\in V_s,\alpha\in V_s^{\perp})\Rightarrow\alpha=0$。这说明二个子空间的和是直和。由于两个子空间的向量是正交的，因而这直和亦称为正交直和。

命题 6 $(V_s^{\perp})^{\perp}=V_s$。

证 由前面的定理可知$V=V_s\oplus V_s^{\perp}$，$V=V_s^{\perp}\oplus(V_s^{\perp})^{\perp}$。所以有$\dim V_s=\dim V-\dim V^{\perp}=\dim(V_s^{\perp})^{\perp}$，另一方面显然有$V_s\subseteq(V_s^{\perp})^{\perp}$，因而$V_s=(V_s^{\perp})^{\perp}$

正交补空间的一个重要应用是下面讲的"最佳近似"定理。现先引入一个向量在一子空间的"最佳近似"的定义。

定义 一个向量在子空间V_s的最佳近似。

设β是欧氏空间的任一向量，它在子空间V_s的最佳近似是指满足下面条件的$\alpha\in V_s$：

$$\|\beta-\alpha\|\leqslant\|\beta-\gamma\|,\ \forall\gamma\in V_s。$$

定理 3 （最佳近似定理） β是欧氏空间V的一个向量，V_s是V的一个子空间。如果$\alpha\in V_s$且$\beta-\alpha\in V_s^{\perp}$，则$\alpha$是$\beta$在$V_s$的最佳近似向量。

证 设γ是V_s的任一个向量，$\beta-\gamma=\beta-\alpha+\alpha-\gamma$，但$\beta-\alpha\in V_s^{\perp}$，$\alpha-\gamma\in V_s$，所以$\|\beta-\gamma\|^2=\|\beta-\alpha\|^2+\|\alpha-\gamma\|^2\geqslant\|\beta-\alpha\|^2$。因此$\alpha$满足条件如下：

$$\|\beta-\alpha\|\leqslant\|\beta-\gamma\|^2,\forall\gamma\in V_s。$$

即 α 是 β 在 V_s 的最佳近似。 ▨

这个定理说明了要求 β 在 V_s 的最佳近似向量，只要将 β 沿正交直和 $V_s\oplus V_s^{\perp}$ 分解，$\beta=\alpha+\gamma, \alpha\in V_s, \gamma\in V_s^{\perp}$，则 α 便是最佳近似向量，但具体的求法可按 Gram-Schmidt 正交过程去做。(见下面例 6)

命题 7 β 是欧氏空间 V 的一个向量，V_s 是 V 的一个子空间，$\varepsilon_1\cdots\varepsilon_k$ 是 V_s 的一组正交基，则 $\gamma=\dfrac{(\beta,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1+\cdots+\dfrac{(\beta,\varepsilon_k)}{(\varepsilon_k,\varepsilon_k)}\varepsilon_k$ 是 β 在 V_s 的最佳近似向量。

证 实际上由 Gram-Schmidt 正交过程可知 $\beta-\gamma\in V_s^{\perp}$，因此 γ 确实是最佳近似向量。由前面知道 γ 是 β 在 V_s 上的投影。

例 6 设 $\beta=(4\ \ -1\ \ -3\ \ 4)^T$，$V_s=[\alpha_1, \alpha_2, \alpha_3]$，$\alpha_1=(1\ \ 1\ \ 1\ \ 1)^T$，$\alpha_2=(1\ \ 2\ \ 2-1)^T$，$\alpha_3=(1\ \ 0\ \ 0\ \ 3)^T$，求 β 沿 $V_s\oplus V_s^{\perp}$ 的分解。

解 先求 V_s 的正交基，然后再求 β 在 V_s 的投影。整个过程是将向量组 $\alpha_1\ \alpha_2\ \alpha_3\ \beta$ 正交化。

选 $\varepsilon_1=\alpha_1=(1\ \ 1\ \ 1\ \ 1)^T$，可得

$$\Pi_{\varepsilon_1}(\alpha_2)=\frac{(\alpha_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1=(1\ \ 1\ \ 1\ \ 1)^T。$$

$$\varepsilon_2=\alpha_2-\pi_{\varepsilon_1}(\alpha_2)=(1\ 2\ 2\ -1)^T-(1\ \ 1\ \ 1\ \ 1)^T=(0\ \ 1\ \ 1-2)^T,$$

又
$$\Pi_{\varepsilon_1}(\alpha_3)=\frac{(\alpha_3,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1=(1\ \ 1\ \ 1\ \ 1)^T,$$

$$\Pi_{\varepsilon_2}(\alpha_3)=\frac{(\alpha_3,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2=-(0\ \ 1\ \ 1\ \ -2)^T。$$

故

$$\varepsilon_3=\alpha_3-\Pi_{\varepsilon_1}(\alpha_3)-\pi_{\varepsilon_2}(\alpha_3)=(1\ \ 0\ \ 0\ \ 3)^T-(1\ \ 1\ \ 1\ \ 1)^T+(0\ \ 1\ \ 1\ \ -2)^T=0。$$

这说明 α_3 不是独立的，它与 α_1, α_2 或 ε_1, ε_2 属同一空间。
因而

$$\pi_{\varepsilon_1}(\beta)=\frac{(\beta_1,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1=(1\ \ 1\ \ 1\ \ 1)^T,$$

$$\prod_{\eta_1}(\beta)=(0\ \ -2\ \ -2\ \ 4),$$

则

$$\gamma=\beta-\prod_{\varepsilon_1}(\beta)-\prod_{\varepsilon_2}(\beta)=(4\ \ -1\ \ -3\ \ 4)^T-(1\ 1\ 1\ 1)^T$$
$$-(0\ \ -2\ \ -2\ \ 4)^T=(3\ 0\ \ -2\ \ -1)^T.$$

$$\beta=\gamma+\prod_{\varepsilon_1}(\beta)+\prod_{\varepsilon_2}(\beta)=(3\ 0\ \ -2\ \ -1)^T+(1\ \ -1\ \ -1\ 5)^T。$$

所以

$$\gamma\in V_s^{\perp},\ \ \pi_{\varepsilon_1}(\beta)+\pi_{\varepsilon_2}(\beta)\in V_s$$

关于最佳近似定理下节讲到矛盾方程的最优解时，还要应用到。

2. 矩阵的 QR 分解

如果将 Gram–Schmidt 正交过程应用于矩阵的列向量，便可得到一个矩阵的 QR 分解。下面以四阶矩阵为例说明。

例 7 设 $A=\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$，且 A 是非奇异的。试写出它的列向量的 Gram-Schmidt 正交过程。

解 用 $\alpha_1\,\alpha_2\,\alpha_3\,\alpha_4$ 表示 A 的各个列向量，则

$$\varepsilon_1=\alpha_1$$

$$\varepsilon_2=\alpha_2-\frac{(\alpha_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1,$$

$$\varepsilon_3=\alpha_3-\frac{(\alpha_3,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1-\frac{(\alpha_3,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2,$$

$$\varepsilon_4=\alpha_4-\frac{(\alpha_4,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1-\frac{(\alpha_4,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2-\frac{(\alpha_4,\varepsilon_3)}{(\varepsilon_3,\varepsilon_3)}\varepsilon_3。$$

这里 $\varepsilon_1\,\varepsilon_2\,\varepsilon_3\,\varepsilon_4$ 是正交向量组。如果要将 $\alpha_1\,\alpha_2\,\alpha_3\,\alpha_4$ 用 $\varepsilon_1\,\varepsilon_2\,\varepsilon_3\,\varepsilon_4$ 表示，且要求 $\varepsilon_1\,\varepsilon_2\,\varepsilon_3\,\varepsilon_4$ 是标准正交向量组，则可写成

$$\alpha_1=\|\varepsilon_1\|\varepsilon_1/\|\varepsilon_1\|=\|\varepsilon_1\|\varepsilon_1',\qquad \varepsilon_1'=\varepsilon_1/\|\varepsilon_1\|,$$

$$\alpha_2=\frac{(\alpha_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\|\varepsilon_1\|\cdot\varepsilon_1/\|\varepsilon_1\|+\|\varepsilon_2\|\varepsilon_2/\|\varepsilon_2\|=\frac{(\alpha_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\|\varepsilon_1\|\varepsilon_1'$$

$$+\|\varepsilon_2\|\varepsilon_2',\quad \varepsilon_2'=\varepsilon_2/\|\varepsilon_2\|,$$

$$\alpha_3=\frac{(\alpha_3,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\|\varepsilon_1\|\cdot\varepsilon_1/\|\varepsilon_1\|+\frac{(\alpha_3,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\|\varepsilon_2\|\cdot\varepsilon_2/\|\varepsilon_2\|+\|\varepsilon_3\|$$

$$\cdot\varepsilon_3/\|\varepsilon_3\|$$

$$=\frac{(\alpha_3,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\|\varepsilon_1\|\varepsilon_1'+\frac{(\alpha_3,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\|\varepsilon_2\|\varepsilon_2'+\|\varepsilon_3\|\varepsilon_3',\quad \varepsilon_3'=\varepsilon_3/\|\varepsilon_3\|,$$

$$\alpha_4=\frac{(\alpha_4,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\|\varepsilon_1\|\cdot\varepsilon_1/\|\varepsilon_1\|+\frac{(\alpha_4,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\|\varepsilon_2\|\cdot\varepsilon_2/\|\varepsilon_2\|$$

$$+\frac{(\alpha_4,\varepsilon_3)}{(\varepsilon_3,\varepsilon_3)}\|\varepsilon_3\|\varepsilon_3/\|\varepsilon_3\|+\|\varepsilon_4\|\cdot\varepsilon_4/\|\varepsilon_4\|$$

$$=\frac{(\alpha_4,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\|\varepsilon_1\|\varepsilon_1'+\frac{(\alpha_4,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\|\varepsilon_2\|\varepsilon_2'+\frac{(\alpha_4,\varepsilon_3)}{(\varepsilon_3,\varepsilon_3)}\|\varepsilon_3\|\varepsilon_3'$$

$$+\|\varepsilon_4\|\varepsilon_4',\quad \varepsilon_4'=\varepsilon_4/\|\varepsilon_4\|。$$

如果将上面四个式子写成矩阵的形式便是

$$[\alpha_1,\alpha_2,\alpha_3,\alpha_4]=[\varepsilon_1',\varepsilon_2',\varepsilon_3',\varepsilon_4']$$

$$\begin{bmatrix}\|\varepsilon_1\| & \frac{(\alpha_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\|\varepsilon_1\| & \frac{(\alpha_3,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\|\varepsilon_1\| & \frac{(\alpha_4,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\|\varepsilon_1\| \\ & \|\varepsilon_2\| & \frac{(\alpha_3,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\|\varepsilon_2\| & \frac{(\alpha_4,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\|\varepsilon_2\| \\ & & \|\varepsilon_3\| & \frac{(\alpha_4,\varepsilon_3)}{(\varepsilon_3,\varepsilon_3)}\|\varepsilon_3\| \\ & & & \|\varepsilon_4\|\end{bmatrix}。$$

即 $A=QR$，其中 R 是对角元为正的上三角阵，Q 是正交阵。这表示式称为 A 的 QR 分解。上面的方法对一般的非奇异 n 阶矩阵都是可行的，这样可得到下面的定理。

定理 4 设 A 是非奇异的 n 阶实方阵，则 A 可分解为乘积

QR（Q 是正交矩阵，R 为对角元为正的上三角阵）。

例 8 已知 $A=\begin{bmatrix}-3 & -2 & 1\\ 0 & 1 & 2\\ 1 & 0 & 3\end{bmatrix}$，试作出 A 的 QR 分解。

解 $\alpha_1=(-3\ \ 0\ \ 1)^T$，$\alpha_2=(-2\ \ 1\ \ 0)^T$，$\alpha_3=(1\ \ 2\ \ 3)^T$，写出它们的正交过程。

有 $\beta_1'=\alpha_1=(-3\ \ 1\ \ 0)^T$，$\beta_1=\dfrac{\beta_1'}{|\beta_1'|}=\dfrac{1}{\sqrt{10}}(-3\ \ 1\ \ 0)^T$。

$$\beta_2'=\alpha_2-\frac{(\alpha_2,\beta_1')}{(\beta_1',\beta_1')}\beta_1'=(-2\ \ 1\ \ 0)^T-\frac{6}{10}(-3\ \ 0\ \ 1)^T$$

$$=\left(-\frac{1}{5}\ \ 1\ \ -\frac{3}{5}\right)^T,$$

$$\beta_2=\frac{\sqrt{5}}{\sqrt{7}}\left(-\frac{1}{5}\ \ 1\ \ -\frac{3}{5}\right)^T=\left(-\frac{1}{\sqrt{35}}\ \ \frac{\sqrt{5}}{\sqrt{7}}\ \ \frac{3}{\sqrt{35}}\right)^T。$$

$$\beta_3'=\alpha_3,\quad \beta_3=\frac{\beta'}{|\beta_3'|}=\left(\frac{1}{\sqrt{14}}\ \ \frac{2}{\sqrt{14}}\ \ \frac{3}{\sqrt{14}}\right)^T.$$

所以

$$A=\begin{bmatrix}-3 & -2 & 1\\ 0 & 1 & 2\\ 1 & 0 & 3\end{bmatrix}=\begin{bmatrix}\frac{-3}{\sqrt{10}} & -\frac{1}{\sqrt{35}} & \frac{1}{\sqrt{14}}\\ 0 & \frac{\sqrt{5}}{\sqrt{7}} & \frac{2}{\sqrt{14}}\\ -\frac{1}{\sqrt{10}} & \frac{3}{\sqrt{3}} & \frac{3}{\sqrt{14}}\end{bmatrix}$$

$$\begin{bmatrix} \sqrt{10} & \frac{6}{\sqrt{10}} & 0 \\ 0 & \frac{\sqrt{7}}{\sqrt{5}} & 0 \\ 0 & 0 & \sqrt{14} \end{bmatrix} = QR。$$

第四节　与矩阵 A 相联系的四个重要子空间

1. 在第一章里曾指出过给出一个矩阵 A 是 $m\times n$ 型，与它联系有四个重要的子空间，它们是

(i) $R(A)$——A 的列空间，

(ii) $N(A)$——A 的零空间，

(iii) $R(A^T)$——A 的行空间(或称为 A^T 的列空间)，

(iv) $N(A^T)$——A 的左零空间(或称为 A^T 的零空间)。

这四个子空间在讨论矩阵理论的很多问题时都常常涉及。为了确定起见，现假定 A 是 $m\times n$ 型实阵，因此 $N(A)$，$R(A^T)$ 是 R^n 的两个子空间；$N(A^T)$，$R(A)$ 是 R^m 的两个子空间。下面的重要定理是讲述它们的关系。

定理 5　$R^n=R(A^T)\oplus N(A)\quad (N(A)=R(A^T)^\perp)$，

$R^m=R(A)\oplus N(A^T)\qquad (N(A^T)=R(A)^\perp)$。

证　只证明第一个分解式，另一个类似可证。因为 $R^n=R(A^T)\oplus R(A^T)^\perp$，所以只要证明 $N(A)=R(A^T)^\perp$ 成立。设 $\alpha=(x_1\cdots x_n)$，且 $\alpha\in N(A)$。这说明 α 是下面线性齐次方程的解

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0 \\ \vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0。\end{cases} \tag{3-3}$$

如果用 β_i 表示向量$(a_{i1}\ a_{i2}\cdots a_{in})$，这样方程组 (3-3) 便可写成

内积的形式

$$(\beta_i,\alpha)=0,\quad i=1,2,\cdots,m。$$

但 $R(A^T)=[\beta_1\cdots\beta_m]$，因此 $\alpha\in R(A^T)^{\perp}$ 即 $N(A)\subseteq(A^T)^{\perp}$。反之，易证 $R(A^T)_{\perp}\subseteq N(A)$，所以有 $R(A^T)^{\perp}=N(A)$。 ▨

例 9 已知 $A=\begin{bmatrix}1&1&2\\0&1&1\\1&3&4\end{bmatrix}$，试求 $N(A)$，$R(A^T)$，$N(A^T)$，$R(A)$的基。

解 解决这类问题最好的方法便是应用第一章 Hermite 标准型。先将 A 通过行初等变换将它化为 Hermite 型

$$\begin{bmatrix}1&1&2\\0&1&1\\1&3&4\end{bmatrix}\longrightarrow\begin{bmatrix}1&1&2\\0&1&1\\0&2&2\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&1\\0&1&1\\0&0&0\end{bmatrix}。$$

因此行空间 $R(A^T)$ 的基为(1 0 1),(0 1 1)，或可写成 $R(A^T)=[(1\ \ 0\ \ 1),(0\ \ 1\ \ 1)]$，$N(A)=R(A^T)^{\perp}$，求 $N(A)$ 便可由解方程

$\begin{cases}x_1+x_3=0\\x_2+x_3=0\end{cases}$ 得出。该方程的基础解系为(-1 -1 1)。

因此 $N(A)=[(-1\ \ -1\ \ 1)]$。

由 Hermite 型中看出 A 中的第一、二列向量是独立的，因而 $R(A)=[(1\ \ 0\ \ 1)^T,\ (1\ \ 1\ \ 3)^T]$。同样由 $N(A)^T=R(A)^{\perp}$ 求 $N(A^T)$。它可由解方程$\begin{cases}x_1+x_3&=0\\x_1+x_2+3x_3&=0\end{cases}$得出，基础解为 $(-1\ \ -2\ \ 1)^T$，所以 $N(A^T)=[(-1\ \ -2\ \ 1)^T]$。

2．矛盾方程的最优解

设有线性方程组

$$
\begin{cases}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \qquad a_{ij}, b_i \in R。 \\
\quad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m,
\end{cases}
$$

现假定向量 $b = (b_1 \cdots b_m)^T \notin R(A)$（这里 A 是上面方程的系数矩阵）。换句话说，对任何 $x = (x_1 \cdots x_n)^T \in R^n$ 都有 $Ax \neq b$。这种方程称为矛盾方程。对这种方程，只能求最优解，即找一向量 $x^0 \in R^n$，有 $\|Ax^0 - b\| \leqslant \|Ax - b\|$，$\forall x \in R^n$。因为向量 Ax 总是属于 $R(A)$，但 $b \notin R(A)$，所以上面的问题就是前面讲的：在子空间 $R(A)$ 上找出向量 b 的最佳近似向量。根据最佳近似定理 3 可知它要求 $Ax^0 - b \in R(A)^\perp$ 或 $Ax^0 - b \in N(A^T)$。所以它应满足方程 $A^T(Ax^0 - b) = 0$，由这方程解出 x^0。方程 $A^TAX = A^Tb$ 称为正规方程。

例 10 求线性近似公式

$$b = a_0 + a_1x_1 + a_2x_2$$

假定上面的公式是反映数量 b, x_1, x_2 的联系的一个式子，但系数尚未能确定，一般通过观察的方法，得出一组关于 b x_1, x_2 的数据，设法由这组数据去求得公式的系数。观察数据如表：

b	x_1	x_2
1	1	0
2	0	1
0	1	1
−1	2	−1

因而有方程组

$$\begin{cases} a_0 + a_1 \qquad\quad = 1 \\ a_0 \qquad\quad + a_2 = 2 \\ a_0 + a_1 \qquad\quad = 0 \\ a_0 + 2a_1 \qquad = 1 \end{cases}$$

用矩阵表示方程为 $Ax = b$，其中

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}, \quad x = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

则

$$A^T A x = \begin{bmatrix} 4 & 4 & 1 \\ 4 & 6 & -1 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} (= A^T b).$$

故

$$a_0 = \frac{17}{6}, \quad a_1 = -\frac{13}{6}, \quad a_2 = -\frac{4}{6},$$

所以 $b = \frac{17}{6} - \frac{13}{6} x_1 - \frac{4}{6} x_2$。

例 11 已知 $A = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 2 & 4 & 2 \\ 0 & 3 & 6 & 1 \end{bmatrix}, b = (1 \ 1 \ 1)^T$，求 $Ax = b$ 的最优解。

解 本例与前例的差别在于解正规方程 $A^T Ax = A^T b$ 时，前例 $A^T A$ 是非奇异的，而本例却是奇异的。所以在求解时，可直接找出 $R(A)$ 中独立的列向量构成新的矩阵 A_1 使得 $R(A_1) = R(A)$，再

解方程 $A_1^TA_1x=A_1^Tb$。如在这题可选 $A_1=\begin{bmatrix}1&3\\2&2\\3&1\end{bmatrix}$，$b=\begin{bmatrix}1\\1\\1\end{bmatrix}$，

则

$$A_1^TA_1u=\begin{bmatrix}14&10\\10&14\end{bmatrix}\begin{bmatrix}u_1\\u_2\end{bmatrix}=A_1^Tb=\begin{bmatrix}6\\6\end{bmatrix},\quad u_1=\frac{1}{4},$$

$$u_2=\frac{1}{4},$$

故原方程的最优解为$\left(0\ \ \frac{1}{4}\ \ 0\ \ \frac{1}{4}\right)$，$b$ 在$R(A)$的最佳近似向量

为

$$A_1u=\begin{bmatrix}1&3\\2&2\\3&1\end{bmatrix}\begin{bmatrix}\frac{1}{4}\\\frac{1}{4}\end{bmatrix}=\begin{bmatrix}1\\1\\1\end{bmatrix}$$

(计算的结果表明 $b\in R(A)$)。

第五节　复内积空间

为了深入讨论一些特殊矩阵的性质，有必要将实内积空间推广至复内积空间。首先对内积公理作适当的修改。

1．复线性空间的内积

设 V 是复数域 C 的内积空间，V 上的内积是指一函数：每一对向量 α,β 对应复数域 C 上的一个数(α,β)，并且这个对应满足下面公理。

(i) $(\alpha,\beta)=\overline{(\beta,\alpha)}$

(ii) $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$，

(iii) $(k\alpha,\beta)=k(\alpha,\beta)$，

(iv) $(\alpha,\alpha)\geqslant 0$，"$=$"仅在 $\alpha=0$ 时成立。

对照前面所讲的实内积公理发现，只是修改第一条对称性。现在要求$(\alpha,\beta)=\overline{(\beta,\alpha)}$，这通常称为 Hermite 对称，它的一个重要推论是$(\alpha,k\beta)=\bar{k}(\alpha,\beta)$。如不作这样的修改，仍沿用对称性，$(\alpha,\beta)=(\beta,\alpha)$ 显见有下面的矛盾

$$0<(i\alpha,i\alpha)=i^2(\alpha,\alpha)=-(\alpha,\alpha),\alpha\neq 0。$$

复内积空间称为酉空间。

2. 酉空间内积的例

例 12 $C^n=\{\alpha|\alpha=(x_1\cdots x_n)^T|x_i\in C\}$。

如果$\alpha=(x_1\cdots x_n)^T$，$\beta=(y_1\cdots y_n)^T$，可定义α，β的内积如下

$(\alpha,\beta)=x_1\bar{y}_1+x_2\bar{y}_2+\cdots+x_n\bar{y}_n=\overline{\beta^T}\alpha=\beta^*\alpha$，（$*$表示运算转置共轭）。特别地$(\alpha,\alpha)=|x_1|^2+|x_2|^2+\cdots+|x_n|^2$。因为这里$x_i$可以是复数。所以$|x_i|$是复数的模。

例 13 $C^{n\times n}$是全体n阶复方阵组成的线性空间。根据上面对C^n的内积的定义，这里可作如下定义：

$$(A,B)=\sum_{i=1}^n\sum_{k=1}^n a_{ik}\bar{b}_{ik}=t_r(B^*,A)。$$

3. Cauchy 不等式（$|(\alpha,\beta)|\leqslant\|\alpha\|\|\beta\|$）仍成立

$|(\alpha,\beta)|$理解为复数的模，同样β在α上的投影向量为$\frac{(\beta,\alpha)}{(\alpha,\alpha)}\alpha\left(不能写成\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha\right)$，因而在证明 Cauchy 不等式时，用

$$\left(\beta-\frac{(\beta,\alpha)}{(\alpha,\alpha)}\alpha,\ \beta-\frac{(\beta,\alpha)}{(\alpha,\alpha)}\alpha\right)\geqslant 0\ 展开即可。$$

4. 向量的模、正交性的定义和欧氏空间一样

与正交阵相对应便是酉阵，$U^*U=I$ 这条件相当于矩阵U的列向量组是一标准正交向量组。当然，U阵同样有$UU^*=I$，这说明它的行向量组亦是标准正交向量组。

例 14 正交阵与酉阵的例。

$$\text{(i)}\quad A=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix},$$

$$\text{(ii)}\ A=\begin{bmatrix}\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\\ 4/3\sqrt{2} & -1/3\sqrt{2} & -1/3\sqrt{2}\end{bmatrix},$$

$$\text{(iii)}\ U=\begin{bmatrix}\frac{1}{\sqrt{3}} & (1-\mathrm{i})/\sqrt{3}\\ (1+\mathrm{i})/\sqrt{3} & -1/\sqrt{3}\end{bmatrix},$$

$$\text{(iv)}\ U=\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\mathrm{i} & \frac{1}{2}-\frac{1}{2}\mathrm{i}\\ \mathrm{i}/\sqrt{2} & -\frac{1}{\sqrt{2}} & 0\\ \frac{1}{2} & -\frac{1}{2}\mathrm{i} & -\frac{1}{2}+\frac{1}{2}\mathrm{i}\end{bmatrix}$$

其中 A 是正交阵，U 是酉阵。

第六节　Schur 定理

Schur 定理是矩阵理论的一条重要定理，它在后面特殊矩阵的讨论，扮演着重要的角色。这一节便是讨论 Schur 定理及有关问题。

1. Schur 定理

定理 6（Schur） 任何一 n 阶方阵 A 酉相似于上三角阵．即存在有酉阵 U 使得

$$U^*AU=B \text{ 成立。}$$

其中 U 是酉阵，B 是上三角阵。

在第一章里曾用"扩充基的方法"证明了如下的命题："任何一个 n 阶方阵相似于上三角阵"。即 $P^{-1}AP=B$，P 是非奇异矩阵。在 Schur 定理中非奇异矩阵还可进一步要求为酉阵，在应用上它有很多方便的地方。如酉阵求逆是 $U^{-1}=U^*$，当然酉阵方便的地方不仅这点。证明 Schur 定理的方法完全和证明"任何一个 n 阶方阵相似于上三角形"的方法相类似。差别在扩充基时，扩充为标准正交基，在内积空间是可以做到的。

Schur 定理的证明：对矩阵 A 的阶施行数学归纳法。$n=1$ 时结论显然成立。现假定对于阶小于 n 的矩阵命题成立，证明命题对 n 阶矩阵亦成立。设 A 是一 n 阶矩阵，任取 A 的一个特征值 λ_1，它相应的一个特征向量为 X_1，则 $AX_1=\lambda_1X_1$。并假定 X^1 是单位向量，将它扩充为 C^n 的标准正交基为：$X_1\ X_2\cdots X_n$。构造矩阵 $U=[X_1\ X_2\cdots X_n]$，它是个酉阵。并考察

$$AU=A[X_1X_2\cdots X_n]=[\lambda_1X_1\ AX_2\cdots AX_n]$$

$$=[X_1X_2\cdots X_n]\begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ 0 & & D & \\ \vdots & & & \\ 0 & & & \end{bmatrix},$$

因此有

$$U^*AU=\begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & D & \\ 0 & & & \end{bmatrix},$$

其中 D 是 $n-1$ 阶方阵。由归纳假设存在一 $n-1$ 阶酉阵 V，有

$V^*DV=C$, C 是上三角阵。构造 n 阶方阵 $Q=\begin{bmatrix}1 & \\ & V\end{bmatrix}$，它是个酉阵。作如下变换

$$Q^*(U^*AU)Q=Q^*\begin{bmatrix}\lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & D & \\ 0 & & & \end{bmatrix}Q=\begin{bmatrix}\lambda_1 & *' & \cdots & *' \\ 0 & & & \\ \vdots & & V^*DV & \\ 0 & & & \end{bmatrix}$$

$$=\begin{bmatrix}\lambda_1 & *' & \cdots & *' \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{bmatrix}。$$

显然 UQ 是酉阵，因此 $(UQ)^*A(UQ)=B$（上三角阵）。 ▨

2. Schur 定理的若干结果

有两个 n 阶方阵 A、B，如果它们存在关系 $A=U^*BU$，那么称 A 与 B 是酉相似，酉相似的一个重要性质是由下面命题给出的。

命题 8 若 $U^*AU=B$，则 $\sum\limits_{i,j}|a_{ij}|^2=\sum\limits_{i,j}|b_{ij}|^2$。文字表达为：如果两个矩阵是酉相似，它们诸元素模的平方和是相同的。（这个性质可视为酉相似的一个必要条件）。

证 若 $U^*AU=B$，则 $U^*A^*U=B^*$，所以 $U^*A^*AU=B^*B$。这说明矩阵 A^*A 与矩阵 B^*B 相似，因而 $t_rA^*A=t_rB^*B$。这便是所求的等式。 ▨

命题 9(Schur) A 是 $n\times n$ 矩阵，$\lambda_1\cdots\lambda_n$ 是 A 的特征值，则

$$\sum_{i=1}^{n}|\lambda_i|^2\leqslant\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2。$$

证 设 $U^*AU=B$，这里 B 是上三角阵，因此它的对角元便是特征值 $\lambda_1\cdots\lambda_n$。由命题便可知

$$\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|=\sum_{i=1}^{n}\sum_{j=1}^{n}|b_{ij}|^2\geqslant\sum_{i=1}^{n}|b_{ii}|^2=\sum_{i=1}^{n}|\lambda_i|^2。$$ ▨

Schur 不等式是矩阵理论里的一个重要不等式，它在特征值的估计中有很多应用。

第七节　正 规 矩 阵

由这节开始讨论应用上与理论上最重要的一种矩阵——正规矩阵。它包括实对称阵、Hermite 阵、正交阵以及酉阵。

定义　A 是 n 阶方阵，如果它有 n 个相互正交的单位向量作为它的特征向量，则称 A 为正规矩阵。

由正规矩阵的定义，立即能得出如下命题。

命题 10　n 阶方阵 A 是正规矩阵 $\Longleftrightarrow A$ 酉相似对角阵。

证　$\Longrightarrow A$ 是 n 阶正规矩阵，则可设 $\varepsilon_1\cdots\varepsilon_n$ 是 A 的特征向量，它们构成标准正交向量组。对应的特征值设为 $\lambda_1\cdots\lambda_n$，因此可有方程组：

$$A\varepsilon_1=\lambda_1\varepsilon_1,\ A\varepsilon_2=\lambda_2\varepsilon_2,\cdots,A\varepsilon_n=\lambda_n\varepsilon_n。$$

将它写成矩阵方程的形式便是

$$[A\varepsilon_1\ A\varepsilon_2\cdots A\varepsilon_n]=[\lambda_1\varepsilon_1\ \lambda_2\varepsilon_2\cdots\lambda_n\varepsilon_n],$$

或

$$A[\varepsilon_1\cdots\varepsilon_n]=[\varepsilon_1\cdots\varepsilon_n]\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}。$$

记 $U=[\varepsilon_1\cdots\varepsilon_n]$，它是 U 阵，便有

$$U^*AU=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}。$$

$\Longleftarrow$ 显然。　■

下面的定理说明了正规矩阵的一个重要性质。在很多场合下，它亦可作为正规矩阵的定义。

定理 7　**A 是 n 阶矩阵，它是正规矩阵 $\Longleftrightarrow A^*A=AA^*$。**

证 $\Longrightarrow$因 A 是正规矩阵，所以有 $A=U\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}U^*$，

因此有 $A^*=U\begin{bmatrix}\overline{\lambda_1} & & \\ & \ddots & \\ & & \overline{\lambda_n}\end{bmatrix}U^*$，

这样

$$A^*A=U\begin{bmatrix}\overline{\lambda_1} & & \\ & \ddots & \\ & & \overline{\lambda_n}\end{bmatrix}U^*U\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}U^*$$

$$=U\begin{bmatrix}|\lambda_1|^2 & & \\ & \ddots & \\ & & |\lambda_n|^2\end{bmatrix}U^*=AA^*。$$

$\Longleftarrow$由 Schur 定理可知 $A=UBU^*$，其中 B 是上三角阵。由 $A^*A=AA^*$ 可得知$B^*B=BB^*$。下面证明 B 是对角阵。

设 $B=[b_{ij}]$，$\begin{bmatrix}\overline{b}_{11} & & & & \\ \overline{b}_{12} & \overline{b}_{22} & & & \\ \overline{b}_{13} & \overline{b}_{23} & \overline{b}_{33} & \ddots & \\ \vdots & \vdots & \vdots & \ddots & \\ \overline{b}_{1n} & \overline{b}_{2n} & \overline{b}_{3n} & & \overline{b}_{nn}\end{bmatrix}\begin{bmatrix}b_{11} & b_{12} & b_{13} & \cdots & b_{1n} \\ & b_{22} & b_{23} & \cdots & b_{2n} \\ & & b_{33} & \cdots & b_{3n} \\ & & & \ddots & \\ & & & & b_{nn}\end{bmatrix}$

$$=\begin{bmatrix}b_{11} & b_{12} & b_{13} & \cdots & b_{1n} \\ & b_{22} & b_{23} & \cdots & b_{2n} \\ & & b_{33} & \cdots & b_{3n} \\ & & & \ddots & \\ & & & & b_{nn}\end{bmatrix}\begin{bmatrix}\overline{b}_{11} & & & & \\ \overline{b}_{12} & \overline{b}_{22} & & & \\ \overline{b}_{13} & \overline{b}_{23} & \overline{b}_{33} & & \\ \vdots & \vdots & \vdots & \ddots & \\ \overline{b}_{1n} & \overline{b}_{2n} & \overline{b}_{3n} & & \overline{b}_{nn}\end{bmatrix}。$$

只要比较两边相乘所得矩阵的对角元，便可知 B 是对角阵。设左边矩阵 $C=[c_{ij}]$，右边矩阵为 $D=[d_{ij}]$，那么

$$c_{11}=|b_{11}|^2,\ d_{11}=|b_{11}|^2+|b_{12}|^2+\cdots+|b_{1n}|^2,$$

因此得 $b_{12}=b_{13}=\cdots=b_{1n}=0$。

$$c_{22}=|b_{22}|^2, d_{22}=|b_{22}|^2+|b_{23}|^2+\cdots+|b_{2n}|^2,$$

因此得 $b_{23}=b_{24}=\cdots=b_{2n}=0$

由上类似可得 $b_{ij}=0(i\neq j)$，因此 A 是酉相似对角阵，所以

A 是正规矩阵。 ▨

由这个定理可以知道实对称阵($A = A^T$)，Hermite 阵 ($A = A^*$)，正交阵 ($A^TA = I$)，酉阵($A^*A = I$)都是正规矩阵，都有 n 个相互正交的特征向量。虽然这样，每种矩阵还是有它的特殊性。下面还是要分别作详细的讨论。

例 15 这里给出两个正规矩阵的例子，但它不是 Hermite 阵与酉阵，

$$A = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}, \quad A = \begin{bmatrix} 1+2i & 1-2i \\ 1-2i & 1+2i \end{bmatrix}。$$

可直接验证它们均满足 $A^*A = AA^*$。

在前一节讲到矩阵 A 的特征值模的平方和有一个 Schur 不等式，下面的命题是指出等号成立的情形。

命题 11 一 n 阶方阵 A 是正规矩阵 $\Longleftrightarrow \sum\limits_{i,j}^{n} |a_{ij}|^2 = \sum\limits_{i=1}^{n} |\lambda_i|^2$。

证 $\Longrightarrow$ 设 $A = [a_{ij}]$。因为 A 是正规矩阵，所以 $A = UDU^*$，这里 D 是对角阵。又由"酉相似的两个矩阵，它们诸元素模的平方和相等，所以 $\sum\limits_{i=1}^{n} |\lambda_i|^2 = \sum\limits_{i,j}^{n} |a_{ij}|^2$

$\Longleftarrow$ 由 Schur 定理可知有 $A = UBU^*$，这里 B 是上三角形，所以

$$\sum_{i,j}^{n} |a_{ij}|^2 = \sum_{i,j}^{n} |b_{ij}|^2 = \sum_{i=1}^{n} |\lambda_i|^2 + \sum_{i,j}^{n} |b_{ij}|^2。$$

这里 $b_{ii} = \lambda_i$)

由前提 $\sum\limits_{i,j} |a_{ij}|^2 = \sum\limits_{i=1}^{n} |\lambda_i|^2$ 可知

$$\sum_{i,j}^{n} |b_{ij}|^2 = 0，即 b_{ij} = 0 (i \neq j)。$$

从而 B 是对角阵，A 酉相似于对角阵，所以 A 是正规矩阵。 ▨

例 16 $A=\begin{bmatrix} 7 & 5i & -5-5i \\ -5i & 7 & 5-5i \\ -5+5i & 5+5i & 2 \end{bmatrix}$，问是否存在酉阵 U 使得 $U^{-1}AU$ 为对角阵。若存在，求这样的 U 和 $U^{-1}AU$。

解 要求 $U^{-1}AU$ 为对角阵，即 $U^{*}AU$ 为对角阵，充要条件为 A 是正规矩阵。因现在给出的是 Hermite 阵，所以这样的酉阵是存在的。由前面关于正规矩阵的性质讨论可知只要求出该矩阵的特征值，以及特征向量，有

$$|\lambda I-A|=\lambda^3-16\lambda^2-48\lambda+1152=(\lambda-12)^2(\lambda+8),$$

求 $\lambda=12$ 的特征向量，要解方程 $(12I-A)x=0$。得特征向量为：

$$\alpha_1=(i\ \ 1\ \ 0)^T,\quad \alpha_2=(1+i\ \ 0\ \ -1)^T。$$

为了要求相互正交的特征向量可将 Gram-Schmidt 正交过程用于 α_1、α_2，则

$$\varepsilon_1=\alpha_1=(i\quad 1\quad 0)^T,$$

上式单位化为：
$$\varepsilon_1'=\frac{\varepsilon}{|\varepsilon_1|}=\frac{1}{\sqrt{2}}(i\ \ 1\ \ 0)^T。$$

又

$$\varepsilon_2=\alpha_2-\frac{(\alpha_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1=\left(\frac{1+i}{2}\quad \frac{-1+i}{2}\quad -1\right)^T,$$

单位化为：

$$\varepsilon_2'=\frac{\varepsilon_2}{\|\varepsilon_2\|}=\frac{1}{\sqrt{2}}\left(\frac{1+i}{2},\ \frac{-1+i}{2}\ -1\right)^T。$$

求 $\lambda=-8$ 的特征向量，解方程 $(-8I-A)x=0$。得特征向量为：

$\alpha_3=(1\quad i\quad 1-i)^T$，单位化为：$\varepsilon_3'=\dfrac{\alpha_3}{|\alpha_3|}=\dfrac{1}{2}(1\ \ i\ \ 1-i)^T$。

由正规矩阵的性质知道 ε_3' 与 ε_1' ε_2' 肯定是正交的。构造 U 阵，$U=[\varepsilon_1'\ \varepsilon_2'\ \varepsilon_3']$，则 $AU=[A\varepsilon_1'\ \ A\varepsilon_2'\ \ A\varepsilon_3']=[12\varepsilon_1'\ \ 12\varepsilon_2'\ \ -8\varepsilon_3']$

$$=[\varepsilon_1' \ \varepsilon_2' \ \varepsilon_3']\begin{bmatrix}12 & & \\ & 12 & \\ & & -8\end{bmatrix}。因此 U^*AU=\begin{bmatrix}12 & & \\ & 12 & \\ & & -8\end{bmatrix},$$

这里 U 写出来为

$$U=\begin{pmatrix}\dfrac{i}{\sqrt{2}} & \dfrac{1+i}{2\sqrt{2}} & \dfrac{1}{2} \\ \dfrac{1}{\sqrt{2}} & \dfrac{-1+i}{2\sqrt{2}} & \dfrac{i}{2} \\ 0 & -\dfrac{1}{\sqrt{2}} & \dfrac{1-i}{2}\end{pmatrix}.$$

第八节　实对称阵与 Hermite 阵

在第二章的第二节里，曾从不同角度讨论过本节所讲的问题，但这里是从正规矩阵出发，且讨论较完整。

1．实对称阵与 Hermite 阵特征值的性质

设 A 是 n 阶 Hermite 阵，因而它是正规矩阵、酉相似对角阵，因为 $A=U\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}U^*$，所以

$$A^*=U\begin{bmatrix}\overline{\lambda_1} & & & \\ & \overline{\lambda_2} & & \\ & & \ddots & \\ & & & \overline{\lambda_n}\end{bmatrix}U^*。因 A=A^*，立即可得\lambda_i=$$

$\overline{\lambda_i}$；这样可得到如下的命题。

命题 12　Hermite 阵的特征值是实数。

命题 13　Hermite 阵对应不同特征值的特征向量是正交的，关于实对称阵还可得出如下的重要定理。

定理 8　A 是 n 阶实对称阵，则它正交相似于对角阵，即

$$Q^TAQ=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},\text{其中 } Q \text{ 是正交阵。}$$

这是因为 A 的特征值均为实数，在求特征向量时，解方程 $(\lambda I-A)x=0$ 是实系数线性齐次方程，这样保证特征向量均为实的向量。因此在酉相似对角阵时，可取实的酉阵，即正交阵。

例 17 已知 $A=\begin{bmatrix}7 & -2 & 1\\ -2 & 10 & -2\\ 1 & -2 & 7\end{bmatrix}$，求正交阵 Q 使 Q^TAQ 为对角阵。

解 由定理 8 可知 Q 是由 A 的特征向量构成，而对角阵的元为其特征值。现求矩阵 A 的特征向量：

$|\lambda I-A|=\lambda^3-24\lambda^2+180\lambda-432=(\lambda-12)(\lambda-6)^2$，

$\lambda=12$ 时特征向量 $\alpha_1=(1\ \ -2\ \ 1)^T$，

$\lambda=6$ 时特征向量 $\alpha_2=(2\ \ 1\ \ 0)^T$，$\alpha_3=(-1\ \ 0\ \ 1)^T$。

由上可知 α_1 与 α_2,α_3 均正交，但 α_2,α_3 并不正交，且 $\alpha_1,\alpha_2,\alpha_3$ 还不是单位向量。下面便做这个工作：

$$\varepsilon_1=\frac{\alpha_1}{|\alpha_1|}=\frac{1}{\sqrt{6}}(1\ \ -2\ \ 1)^T,$$

$$\varepsilon_2=\frac{\alpha_2}{|\alpha_2|}=\frac{1}{\sqrt{5}}(2\ \ 1\ \ 0)^T,$$

$$\varepsilon_3=\alpha_3-(\alpha_3,\varepsilon_2)\varepsilon_2=(-1\ \ 0\ \ 1)^T+\frac{2}{5}(2\ \ 1\ \ 0)=\left(-\frac{1}{5}\ \ \frac{2}{5}\ \ 1\right),$$

再单位化为：

$$\varepsilon_3'=\frac{\varepsilon_3}{|\varepsilon_3|}=\frac{1}{\sqrt{30}}(-1\ \ 2\ \ 5)^T。$$

因此 $$Q=\begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{30}} \\ -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{5}{\sqrt{30}} \end{bmatrix},$$

$$Q^TAQ=\begin{bmatrix} 12 & & \\ & 6 & \\ & & 6 \end{bmatrix}。$$

例 18 将下面二次型

$$f(x)=7x_1^2+10x_2^2+7x_3^2-4x_1x_2+2x_1x_3-4x_2x_3,$$

用正交变换化为标准型。

解 将上面二次型写成矩阵的形式为：

$$f(x)=(x_1\ \ x_2\ \ x_3)\begin{bmatrix} 7 & -2 & 1 \\ -2 & 10 & -2 \\ 1 & -2 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=x^TAx.$$

设 $x=Qy$ (Q 是正交阵)，由例 17 可知

$$f(x)=y^TQ^TAQy=y^TDy$$

$$=(y_1y_2y_3)\begin{bmatrix} 12 & & \\ & 6 & \\ & & 6 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}=12y_1^2+6y_2^2+6y_3^2。$$

从例 17 可看出用正交变换将二次型变成标准型的意义，现叙述如下：

在 n 维 Euclid 空间，取定一组标准正交基 $\varepsilon_1\cdots\varepsilon_n$，如果

$$f(x)=\sum_{ij}a_{ij}x_ix_j=x^TAx$$

是在基 $\varepsilon_1\cdots\varepsilon_n$ 下的二次型(这里 $x=x_1\varepsilon_1+\cdots+x_n\varepsilon_n$)，则必有一个正交变换 $y=Qx$ 将二次型 x^TAx 变成标准型 $\sum_{i=1}^{n}\lambda_iy_i^2$（这里 λ_i 是

矩阵 A 的特征值)。

2. 正定矩阵

在矩阵应用中最重要的，当首推正定矩阵。下面讨论，为确定起见，矩阵 A 均为实对称阵。

定义 矩阵 A 是正定的。如果它对应的二次型 $f(x)=x^TAx$，对任何 $x\neq 0$ 均有 $f(x)>0$，这时用$A>0$ 表示，如果对任何 x 均有 $f(x)\geqslant 0$，则称 A 是半正定的。

下面讨论正定矩阵的基本性质与判定方法。

命题 14 若 $A>0$，则对任何非奇异矩阵 P 有 $P^TAP>0$。

证 现证明 $x^T(P^TAP)x>0, \forall x\neq 0$，

因为 $y=Px$，P 是非奇异矩阵，$x\neq 0$ 总有 $y\neq 0$。再由 $A>0$ $x^T(P^TAP)x=(Px)^TA(Px)=y^TAy>0$。 ▨

命题 15 $A>0 \Longleftrightarrow A$ 的诸特征值均大于零。

证 $\Longrightarrow$设 λ 是 A 的任一个特征值，x 是它相对应的特征向量。

$$x^TAx=x^T(\lambda x)=\lambda x^Tx,\ \lambda=\frac{x^TAx}{x^Tx}>0。$$

$\Longleftarrow$ 由 A 正交相似于对角阵，即可知有正交阵 Q 使得

$$Q^TAQ=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}=B。$$

显然 B 是正定矩阵。因为 $y^TBy=\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2$，在 $y=(y_1\cdots y_n)^T\neq 0$，有 $y^TBy>0$。但 $A=QBQ^T$，由命题 13 可知 $A>0$。 ▨

命题 16 $A>0\Longleftrightarrow$存在一非奇异矩阵 P 有 $A=P^TP$。

证 $\Longleftarrow$ 如 $A=P^TP$，则相应的二次型 $x^TAx=x^TP^TPx=(Px)^T(Px)=\|Px\|^2$。由于 P 是非奇异矩阵，所以 $x\neq 0$ 有 $Px\neq 0$，因此$\|Px\|>0$，$A>0$。

$$\Longrightarrow\quad A>0,\ A=Q\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}Q^T,\ (\lambda_i>0, Q\text{ 是正交阵})$$

故

$$A=Q\begin{bmatrix}\lambda_1^{\frac{1}{2}} & & \\ & \ddots & \\ & & \lambda_n^{\frac{1}{2}}\end{bmatrix}Q^TQ\begin{bmatrix}\lambda_1^{\frac{1}{2}} & & \\ & \ddots & \\ & & \lambda_n^{\frac{1}{2}}\end{bmatrix}Q^T=P\cdot P,$$

$$P=Q\begin{bmatrix}\lambda_1^{\frac{1}{2}} & & \\ & \ddots & \\ & & \lambda_n^{\frac{1}{2}}\end{bmatrix}Q^T。$$

因为 P 是对称阵，因此可写成 $A=P^TP$。 ▨

由证明结果看来上面的命题可改写成为：

$A>0 \Longleftrightarrow$ 当有 $P>0$，有 $A=P^2$。

命题 17 $A>0 \Longleftrightarrow A$ 的顺序子式

$$a_{11}>0,\ \begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}>0,\begin{bmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{bmatrix}>0,\cdots,|A|>0。$$

(这个命题通常称为正定矩阵的行列式判别方法)。

证 $\Longrightarrow$ 主要是利用 $A>0$，则 A 的诸特征值 $\lambda_i>0$，因此 $|A|=\lambda_1\lambda_2\cdots\lambda_n>0$。即一个正定矩阵它的行列式一定大于零。其次第 k 个顺序主子式的矩阵亦是一个正定矩阵。设第 k 个顺序主子式 $|A_k|$，则 $(x_1\cdots x_k)A_k\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_k\end{bmatrix}=(x_1\cdots x_k\ 0\cdots 0)A\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_k\\ 0\\ \vdots\\ 0\end{bmatrix}>0$,

因此有 $|A_k|>0$。

$\Longleftarrow$ 对 n 施行数学归纳法。当 $n=1$ 时定理显然成立。因为这时二次型为 $f(x)=a_{11}x_1^2>0$。假定 $n-1$ 阶对称阵命题成

立，证明对 n 阶对称矩阵定理也成立。由于 A_{n-1} 它的各个顺序主子式均大于零，因而由归纳假设可知，A_{n-1} 是 $n-1$ 阶正定阵，因此存在一正交阵 Q 有 $Q^TA_{n-1}Q=D$，$D=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_{n-1}\end{bmatrix}$。作出正交阵 $G=\begin{bmatrix}Q & \\ & 1\end{bmatrix}$，对 A 作变换：

$$G^TAG=\begin{bmatrix}Q^T & \\ & 1\end{bmatrix}\begin{bmatrix}A_{n-1} & \alpha \\ \alpha^T & a_{nn}\end{bmatrix}\begin{bmatrix}Q & \\ & 1\end{bmatrix}$$

$$=\begin{bmatrix}Q^TA_{n-1}Q & Q^T\alpha \\ \alpha^TQ & a_{nn}\end{bmatrix}=\begin{bmatrix}D & Q^T\alpha \\ \alpha^TQ & a_{nn}\end{bmatrix}。$$

再做一次合同变换，将 G^TAG 化为对角型：

$$\begin{bmatrix}I_{n-1} & 0 \\ -\alpha^TQD^{-1} & 1\end{bmatrix}\begin{bmatrix}D & Q^T\alpha \\ \alpha^TQ & a_{nn}\end{bmatrix}\begin{bmatrix}I_{n-1} & -D^{-1}Q^T\alpha \\ 0 & 1\end{bmatrix}$$

$$=\begin{bmatrix}D & 0 \\ 0 & a_{nn}-\alpha^TGD^{-1}G^T\alpha\end{bmatrix}。$$

等式两边取行列式，因为左边的都为正值，又右边为 $(a_{nn}-\alpha^TGD^{-1}G^T\alpha)|D|$，但 $|D|>0$，所以 $(a_{nn}-\alpha^TGD^{-1}G^T\alpha)>0$。换句话说，矩阵 A 与一个对角元均为正的矩阵合同，因而 $A>0$。▨

3. Hermite 阵的说明

如果要将实的二次型推广至复二次型。Hermite 阵便担任实对称阵的角色，复二次型可写成 $f(x)=x^*Ax$ ($x\in C^n$。$A^*=A$) 称为 Hermite 型，首先有如下命题。

命题 18 Hermite 型的值总是实数。

证 将 (x^*Ax) 取共轭转置（因为 x^*Ax 是个数，所以共轭转置实为共轭），$(x^*Ax)^*=x^*A^*x=x^*Ax$，即 x^*Ax 是实数。▨

例 19 $f(x)=x_1\bar{x}_1+4x_2\bar{x}_2+x_3\bar{x}_3+14ix_1\bar{x}_2-14i\bar{x}_1x_2+11x_1\bar{x}_3$

$$+11\bar{x}_1x_3+14ix_2\bar{x}_3-14i\bar{x}_2x_3。$$

它可写成

$$(\bar{x}_1\ \bar{x}_2\ \bar{x}_3)\begin{bmatrix} 1 & 14i & 11 \\ -14i & 4 & 14i \\ 11 & -14i & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=x^*Ax。$$

对于 Hermite 型或 Hermite 矩阵是正定的判别法则和实的情形一样。

例 20 Hermite 矩阵 $A=\begin{bmatrix} x & -1 & i \\ -1 & 2 & 3i \\ -i & -3i & 5 \end{bmatrix}$，求使 $A>0$ 的实数 x 应满足的条件。

解 $|A_1|=x$，$|A_2|=\begin{vmatrix} x & -1 \\ -1 & 2 \end{vmatrix}=2x-1$，$|A|=x-13$.

故 $$A>0 \Longleftrightarrow |A_1|>0, |A_2|>0, |A|>0 \Longleftrightarrow x>13。$$

4. 矩阵的奇异值分解

矩阵的奇异值分解，在矩阵的计算与理论中都是很重要的分解。但这里由于篇幅所限只能给出相应的定理。要进一步了解，可参考“*Linear Algebra with Applications*”*by* Steven J. Leon 与“*Matrix analysis*” *by* Roger A. Horn。

(1) 矩阵的奇异值

由前节关于正定矩阵的讨论可知，任何一个正定矩阵 A 可写成正定矩阵 B 的平方：$A=B^2$。实际上用同样的方法亦可证明任何一个半正定的矩阵 A，可写成半正定矩阵 B 的平方。

设 A 是任一个 $m\times n$ 型矩阵，易知 A^*A 是半正定矩阵。因为 $x^*(A^*A)x=(Ax)^*(Ax)=\|Ax\|^2\geqslant 0$，$\forall x\in C^n$，所以它的特征值 $\lambda_1\lambda_2\cdots\lambda_n$ 是非负的，它的平方根 $\sqrt{\lambda_1}\cdots\sqrt{\lambda_n}$ 称为矩阵 A 的 n 个奇异值，通常用 $s_1\ s_2\cdots s_n$ 表示。显然奇异值是非负数。

例 21 求 $A=\begin{bmatrix}1&0\\0&1\\1&0\end{bmatrix}$的奇异值。

解 因为 $A^*A=\begin{bmatrix}1&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}1&0\\0&1\\1&0\end{bmatrix}=\begin{bmatrix}2&0\\0&1\end{bmatrix}$,

因此 A 的奇异值为 $s_1=\sqrt{2}$，$s_2=1$。

命题 19 一个方阵的奇异值在酉变换下是不变的。即矩阵 A,UA,AU 的奇异值是相同的。

证 A 的奇异值为 A^*A 的特征值的平方根，UA 的奇异值为$(UA)^*(UA)=A^*U^*UA=A^*A$ 的特征值的平方根。同理 AU 的奇异值为 $(AU)^*(UA)=U^*(A^*A)U$ 的特征值(即 A^*A 的特征值)的平方根。 ▨

(2) 矩阵的奇异值分解

为了要讲述矩阵的奇异值分解，还要考察 A^*A,AA^* 的特征值与特征向量的关系，首先 A^*A 与 AA^* 的非零特征值相同，所差仅为零特征值的重数。但特征向量的关系又如何？下面命题便回答这个问题。

命题 20 设 $\lambda_1\cdots\lambda_r$ 是 A^*A 的非零特征值。对应的正交的特征向量为 $\alpha_1\ \alpha_2\cdots\alpha_r$，则 AA^* 对应的 $\lambda_1\cdots\lambda_r$ 特征向量为 $A\alpha_1\cdots A\alpha_r$,且亦为正交向量组。

证 因为 $A^*A\alpha_i=\lambda_i\alpha_i$，所以 $AA^*A\alpha_i=\lambda_iA\alpha_i$。余下只要说明 $A\alpha_i\neq0$。如果 $A\alpha_i=0$，就有 $A^*A\alpha_i=0$,它说明 α_i 是对应于 A^*A 的零特征值的特征向量，这与 α_i 的前提相矛盾。下面证明 $A\alpha_i$ 与 $A\alpha_j$ 正交$(i\neq j)$。$(A\alpha_i,A\alpha_j)=(A\alpha_j)^*A\alpha_i=\alpha_j^*A^*A\alpha_i=\lambda_i\alpha_j^*\alpha_i=0$。同时还得出特征向量$A\alpha_i$ 的模 $\|A\alpha_i\|^2=\lambda_i\|\alpha_i\|^2$。▨

定理 9 A 为任一个 $m\times n$ 矩阵，$s_1\cdots s_r$ 是 A 的非零奇异

值，则 A 可表示为如下的形式：

$$A = UDV^*, \tag{3-4}$$

式中 U 是 $m\times m$ 型，V 是 $n\times n$ 型，U、V 都是酉阵，D 是 $m\times n$ 型且在位置 (i,i) 的元是 s_i，其余均为零。或(3-4)称为 A 的奇异值分解。

证 采用命题 19 所用的符号。

A^*A 的特征向量为 $\alpha_1\ \alpha_2\cdots\alpha_r\cdots\alpha_n$，$AA^*$的特征向量为 $\beta_1\ \beta_2\cdots\beta_r\cdots\beta_m$。这里可假定$\{\alpha_i\}\{\beta_j\}$都是标准正交的特征向量系，其中前 r 个为对应于非零特征值的特征向量。由命题 19 可知，$\beta_i = \frac{1}{\sqrt{\lambda_i}}A\alpha_i(i=1\cdots r)$或可写成 $A\alpha_i=\sqrt{\lambda_i}\beta_i\ (i=1\cdots r)$。 $A\alpha_i=0$ $(i=r+1\cdots n)$，因为后者是对应于零特征值的特征向量。构造矩阵 $V=[\alpha_1\cdots\alpha_n]$，

则

$$\begin{aligned}AV &= A[\alpha_1\cdots\alpha_r\ \alpha_{r+1}\cdots\alpha_n]=[A\alpha_1\cdots A\alpha_r\quad A\alpha_{r+1}\cdots A\alpha_n]\\ &=[\sqrt{\lambda_1}\beta_1\cdots\sqrt{\lambda_r}\beta_r\ 0\cdot\beta_{r+1}\cdots 0\cdot\beta_n]\\ &=[\beta_1\cdots\beta_r\ \beta_{r+1}\cdots\beta_m]\begin{bmatrix} s_1 & & & & \\ & \ddots & & & \\ & & s_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0\end{bmatrix}。\end{aligned}$$

用 $U=[\beta_1\cdots\beta_r\cdots\beta_m]$，则上面结果可写成

$$U^*AV=\begin{bmatrix} s_1 & & & & \\ & \ddots & & & \\ & & s_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0\end{bmatrix},$$

这式与式(3-4)本质上是一样的。

例 22 求 $A=\begin{bmatrix}1 & 0\\ 0 & 1\\ 1 & 0\end{bmatrix}$的奇异值分解。

解 前面已知 A 的奇异值为$\sqrt{2}$ 与 1，余下便是求 A^*A 与 AA^* 的正交特征向量，以构成矩阵 U,V。

由于
$$A^*A=\begin{bmatrix}2 & \\ & 1\end{bmatrix},$$
所以它的特征向量可选为 $(1\ \ 0)^T,(0\ \ 1)^T$。

又
$$AA^*=\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 1\end{bmatrix},$$
所以它的特征值为 2，1，0，对应的特征向量为$(1\ \ 0\ \ 1)^T,(0\ \ 1\ \ 0)^T,(1\ \ 0\ \ -1)^T$，它们是相互正交的特征向量。但构成酉阵还要单位化。结果为

$$A=\begin{bmatrix}1 & 0\\ 0 & 1\\ 1 & 0\end{bmatrix}=UDV^*=\begin{bmatrix}\alpha & 0 & \alpha\\ 0 & 1 & 0\\ \alpha & 0 & -\alpha\end{bmatrix}\begin{bmatrix}\sqrt{2} & 0\\ 0 & 1\\ 0 & 0\end{bmatrix}$$
$$\cdot\begin{bmatrix}1 & \\ & 1\end{bmatrix}。$$

第九节 正交阵与酉阵

1．正交阵

一个实阵 A 如果有 $A^TA=I$ 便是正交阵，因而亦是正规阵，所以正交阵亦酉相似于对角阵。但需进一步去探讨正交阵的特征值与特征向量的性质，同时因为正交阵是实阵，所以应注意它正交相似的最简单形式。

命题 21 正交阵的特征值的模等于 1，且对应不同特征值的特征向量是正交的。

证 $A=U\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}U^*$，两边取共轭转置，便有

$$A^T=U\begin{bmatrix}\overline{\lambda_1} & & \\ & \ddots & \\ & & \overline{\lambda_n}\end{bmatrix}U^*。$$

故 $$A^TA=U\begin{bmatrix}\overline{\lambda_1} & & & \\ & \overline{\lambda_2} & & \\ & & \ddots & \\ & & & \overline{\lambda_n}\end{bmatrix}U^*U\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}U^*=I,$$

因此可得 $$\begin{bmatrix}\overline{\lambda_1}\lambda_1 & & & \\ & \overline{\lambda_2}\lambda_2 & & \\ & & \ddots & \\ & & & \overline{\lambda_n}\lambda_n\end{bmatrix}=I,$$

即 $|\lambda_i|=1$。命题第二个结论在正规矩阵时已证。

命题 22 A 是正交阵，$\lambda=a+bi$ 是它的一个复特征值，它对应的特征向量 $v=\alpha+i\beta$，则有 $\|\alpha\|=\|\beta\|$ $(\alpha,\beta)=0$。

证 这个命题给出一个正交阵的复特征向量的性质。由于 A 是实阵，所以 $\overline{\lambda}=a-bi$ 也是特征值，$\bar{v}=\alpha-i\beta$ 是对应的特征向量。由正交阵的性质可知 $(\bar{v},v)=0$。首先可证明 $\|v\|=\|\bar{v}\|$，即

$$\|v\|^2=(v,v)=(\alpha+i\beta,\alpha+i\beta)=(\alpha,\alpha)+(\alpha,i\beta)+(i\beta,\alpha)+(i\beta,i\beta)=(\alpha,\alpha)+(\beta,\beta),$$

$$\|\bar{v}\|^2=(\bar{v},\bar{v})=(\alpha-i\beta,\alpha-i\beta)=(\alpha,\alpha)+(\alpha,-i\beta)+(-i\beta,\alpha)+(-i\beta,-i\beta)=(\alpha,\alpha)+(\beta,\beta)。$$

另一方面 $\alpha=\dfrac{v+\bar{v}}{2}$，$\beta=\dfrac{1}{2i}(v-\bar{v})$，则

$$(\alpha,\beta)=\frac{1}{4i}(v+\bar{v},v-\bar{v})=-\frac{1}{4i}\{(v,v)-(\bar{v},\bar{v})-(v,\bar{v})-(\bar{v},v)\}=0$$

故

$$\|\alpha\|^2=\frac{1}{4}(\|v\|^2+\|\bar{v}\|^2),$$

$$\|\beta\|^2=\frac{1}{4}(\|v\|^2+\|\bar{v}\|)^2\left(=\frac{1}{2}\|v\|^2\right)。$$ ▨

下面讨论正交阵在正交相似下的最简单形式。

命题 23 A 是正交阵，它的一个复特征值 $\lambda=\cos\theta+\mathrm{i}\sin\theta$，对应的特征向量 $v=\alpha+\mathrm{i}\beta$，则一定存在一个正交阵 Q 有

$$Q^TAQ=\left[\begin{array}{cc|c}\cos\theta & \sin\theta & \\ -\sin\theta & \cos\theta & \\ \hline & & A_1\end{array}\right]，$$ 这里 A_1 仍为正交阵。

证 因 $A(\alpha+\mathrm{i}\beta)=(\cos\theta+\mathrm{i}\sin\theta)(\alpha+\mathrm{i}\beta)$，实部与虚部分开便可得到等式

$$A\alpha=\alpha\cdot\cos\theta-\beta\sin\theta,\ A\beta=\beta\cos\theta+\alpha\sin\theta。$$

由于 $\|\alpha\|=\|\beta\|$，且 $(\alpha,\beta)=0$，因此可假定 α、β 都为单位正交向量，并将它扩充为标准正交基 $\alpha,\beta,\alpha_3\cdots\alpha_n$。作矩阵 $Q=[\alpha,\ \beta,\alpha_3\cdots\alpha_n]$，$Q$ 是正交阵。则

$$\begin{aligned}AQ&=A[\alpha_1,\beta,\alpha_3\cdots\alpha_n]=[A\alpha,A\beta,A\alpha_3\cdots A\alpha_n]\\&=[\alpha\cos\theta-\beta\sin\theta,\ \beta\cos\theta+\alpha\sin\theta,\ A\alpha_3\cdots A\alpha_n]\\&=[\alpha,\beta,\alpha_3\cdots\alpha_n]\left[\begin{array}{cc|c}\cos\theta & \sin\theta & \\ -\sin\theta & \cos\theta & C\\ \hline & & A_1\end{array}\right]=QD。\end{aligned}$$

因此 $Q^TAQ=D$，D 亦为正交阵。由 $D^TD=I$ 可得 $C=0$。A_1 也是正交阵。 ▨

由上面的命题，容易得出正交阵的实标准型的定理。

定理 10 A 是正交阵，它有特征值 k_1 个 1。k_2 个 -1，$2k$ 个 $\cos\theta_1\pm\mathrm{i}\sin\theta_1\cdots\cos\theta_k\pm\mathrm{i}\sin\theta_k$，则存在一个正交阵 P，P^TAP 为

如下形式

$$
\begin{pmatrix}
\overbrace{\begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix}}^{k_1} & & & & \\
& \overbrace{\begin{matrix}-1 & & \\ & \ddots & \\ & & -1\end{matrix}}^{k_2} & & & \\
& & \boxed{\begin{matrix}\cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1\end{matrix}} & & \\
& & & \ddots & \\
& & & & \boxed{\begin{matrix}\cos\theta_k & \sin\theta_k \\ -\sin\theta_k & \cos\theta_k\end{matrix}}
\end{pmatrix}。
$$

命题 24 3 阶正交矩阵的标准型不外是下面 6 种：

$$
\begin{bmatrix}1 & & \\ & 1 & \\ & & 1\end{bmatrix},\quad
\begin{bmatrix}1 & & \\ & 1 & \\ & & -1\end{bmatrix},\quad
\begin{bmatrix}1 & & \\ & -1 & \\ & & -1\end{bmatrix},\quad
\begin{bmatrix}-1 & & \\ & -1 & \\ & & -1\end{bmatrix},
$$

$$
\begin{bmatrix}1 & & \\ & \cos\theta & \sin\theta \\ & -\sin\theta & \cos\theta\end{bmatrix}
\begin{bmatrix}-1 & & \\ & \cos\theta & \sin\theta \\ & -\sin\theta & \cos\theta\end{bmatrix}。
$$

例 23 在空间选定直角坐标系x,y,z(如图 3-7)，设变换R_1

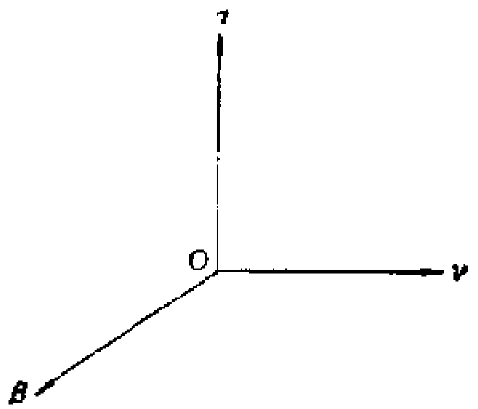

图 3-7

为绕 y 轴旋转 $\frac{\pi}{4}$，变换 R_2 为绕 z 轴旋转 $\frac{\pi}{4}$，试求变换 R_1R_2。

解 首先选定坐标轴上的单位矢量为空间的基，求出 R_1，R_2 在这组标准正交基下的矩阵。R_1 是绕 y 轴的一个旋，这过程 $e_2=(0,1,0)$ 是不动的。而 $e_1=(1,0,0)$ 与 $e_3=(0,0,1)$ 的变化如图(3-8)所示。

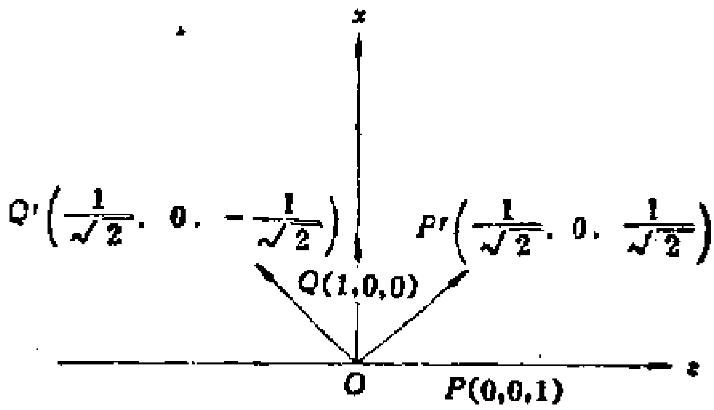

图 3-8

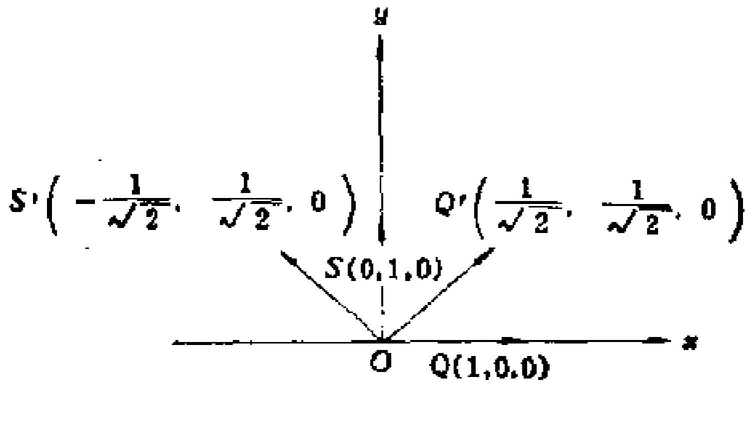

图 3-9

由图 3－8 可知

$$R_1(e_3)=\frac{1}{\sqrt{2}}e_1+\frac{1}{\sqrt{2}}e_3,\quad R_1(e_1)=\frac{1}{\sqrt{2}}e_1-\frac{1}{\sqrt{2}}e_3,$$

因此变换 R_1 的矩阵 $A=\begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$。

由图 3-9 可知

$R_2(e_1)=\frac{1}{\sqrt{2}}e_1+\frac{1}{\sqrt{2}}e_2$，$R_2(e_2)=-\frac{1}{\sqrt{2}}e_1+\frac{1}{\sqrt{2}}e_2$，

$R_2(e_3)=e_3$

因此变换 R_2 的矩阵 $B=\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$。

上面二个矩阵都是正交矩阵。R_1R_2 的矩阵为 AB，也是正交矩阵。

$$AB=\begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$=\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \end{bmatrix}。$$

例 24 假设已知例 23 中 R_1R_2 是一个旋转，试确定它的旋转轴与旋转的角度。

解 设 $R=R_1R_2$，要求出旋转的轴便是要解方程 $R\alpha=\alpha$。

即线性方程$(R-1)\alpha=0$。(1 表示恒等变换。)该方程为

$$\begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}-1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}}-1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=0 \quad 即$$

$$\begin{cases} -x_1-x_2+\sqrt{2}x_3=0 \\ x_1+(1-\sqrt{2})x_2=0, \\ -x_1+x_2+(1-\sqrt{2})x_3=0 \end{cases}$$

解得 $-2x_1+x_3=0$，$x_1+(1-\sqrt{2})x_2=0$。

令$x_1=1$，可得特征向量$\alpha=(1,\sqrt{2}+1,2)$。

求旋转的角度可采用下面的方法。选择新的直角坐标系（如图 3-10），α 为垂直轴，而 β、γ 为其它的两个正交的坐标轴。在这基下，R 的矩阵为

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}。$$

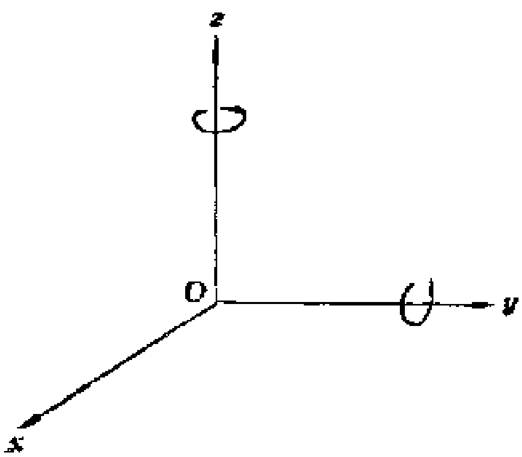

图 3-10

由于这个矩阵的迹与原矩阵是相同的，因而有方程

$$1+2\cos\theta=-\frac{1}{2}+\frac{2}{\sqrt{2}},$$

$$\cos\theta=\frac{1}{2}\left\{\sqrt{2}-\frac{1}{2}\right\}。$$

第十节* 选定基下内积的表示式

本节的目的是在 Euclid 空间里，选定一组基 $\alpha_1\cdots\alpha_n$ 以后，给出任何两个向量内积的表达式，并讨论有关结论。

设 $\alpha=x_1\alpha_1+x_2\alpha_2+\cdots\cdots x_n\alpha_n$，$\beta=y_1\alpha_1+y_2\alpha_2+\cdots+y_n\alpha_n$，则

$$(\alpha,\beta)=(x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n,y_1\alpha_1+y_2\alpha_2+\cdots+y_n\alpha_n)$$

$$=\begin{cases}x_1y_1(\alpha_1,\alpha_1)+x_1y_2(\alpha_1,\alpha_2)+\cdots+x_1y_n(\alpha_1,\alpha_n)\\ x_2y_1(\alpha_2,\alpha_1)+x_2y_2(\alpha_2,\alpha_2)+\cdots+x_2y_n(\alpha_2,\alpha_n)\\ \vdots\\ x_ny_1(\alpha_n,\alpha_1)+x_ny_2(\alpha_n,\alpha_2)+\cdots+x_ny_n(\alpha_n,\alpha_n)。\end{cases}$$

也可写成矩阵乘积的形式

$$(y_1\ \ y_2\cdots y_n)\begin{bmatrix}(\alpha_1,\alpha_1) & (\alpha_2,\alpha_1) & \cdots & (\alpha_n\alpha_1)\\ (\alpha_1,\alpha_2) & (\alpha_2,\alpha_2) & \cdots & (\alpha_n,\alpha_2)\\ \vdots & & & \\ (\alpha_1,\alpha_n) & (\alpha_2,\alpha_n) & \cdots & (\alpha_n,\alpha_n)\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}$$

$$=y^TGx \tag{3-5}$$

这里 G 称为基 $\alpha_1\cdots\alpha_n$ 下的度量矩阵。在 Euclid 空间，它是一个实对称阵，由式(3-5) 可知 $0<(\alpha,\alpha)=x^TGx$，即 G 是一个正定矩阵。同时由式(3-5)还可知道只要 $(\alpha_1,\alpha_1)(\alpha_1,\alpha_2)\cdots(\alpha_1,\alpha_n)$ $(\alpha_2,\alpha_3)\cdots(\alpha_n,\alpha_n)$等确定下来，空间里任何二个向量的内积便可计算。由此可知在选定一组基之后，一个正定矩阵便可确定一个内积，它是由式(3-5)确定。

例25 在 R^2 中定义内积 $(x,y)=x_1y_1+x_1y_2+x_2y_1+2x_2y_2$。试验证之。

解 实际上

$$(x,y)=(y_1,y_2)\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=y^TGx,$$

(x,y)的表达式显然满足内积公理的对称性、可加性、齐次性，余下

只要验证$(x,x)>0$即$G>0$。这由行列式的判别法则即可得出。

由于G是正定矩阵，因此有$\begin{vmatrix}(\alpha_1,\alpha_1) & (\alpha_2,\alpha_1)\\ (\alpha_1,\alpha_2) & (\alpha_2,\alpha_2)\end{vmatrix}>0$，这样亦可导出 Cauchy 不等式。

例 26 $\begin{vmatrix}(\alpha_1,\alpha_1) & (\alpha_2,\alpha_1)\\ (\alpha_1,\alpha_2) & (\alpha_2,\alpha_2)\end{vmatrix}$的几何意义。

解 事实上考察向量α,β所构成的平行四边形的面积S。由图 3-11 可得

$$S^2=\left\|\beta-\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha\right\|^2\|\alpha\|$$

$$=(\alpha,\alpha)\left\{\left(\beta-\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha,\ \beta-\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha\right)\right\}$$

$$=(\alpha,\alpha)\left\{(\beta,\beta)+\left(\beta_1-\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha\right)+\left(-\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha,\beta\right)\right.$$

$$\left.+\left(-\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha,-\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha\right)\right\}$$

$$=(\alpha,\alpha)\left\{(\beta,\beta)-\frac{(\alpha,\beta)(\alpha,\beta)}{(\alpha,\alpha)}\right.$$

$$\left.-\frac{(\alpha,\beta)(\alpha,\beta)}{(\alpha,\alpha)}+\frac{(\alpha,\beta)(\alpha,\beta)(\alpha,\alpha)}{(\alpha,\alpha)(\alpha,\alpha)}\right\}$$

$$=(\alpha,\alpha)(\beta,\beta)-(\alpha,\beta)(\alpha,\beta)。$$

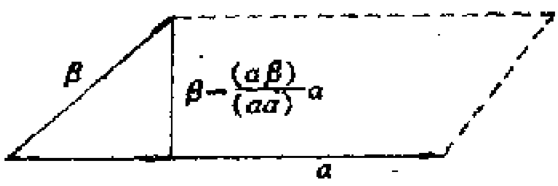

图 3-11

这样二阶的度量矩阵，可解释为由基的向量构成平行四边形的面积。

第四章 Jordan 标准型

第一节 引　　言

前面讨论了在相似变换下,矩阵的最简单形式——对角阵,而且可知并不是所有矩阵都能与对角阵相似的。实际上现所讨论的矩阵是单纯矩阵以及它的一个子集正规矩阵。

$$\text{单纯矩阵}\begin{cases}\text{正规矩阵}\begin{cases}\text{实对称阵, Hermite 阵}\\ \text{正交阵,酉阵}\end{cases}\\ \end{cases}$$

如果用简单的语言来描述这些矩阵便是:

n 阶单纯矩阵——是有完全的特征向量系,

n 阶正规矩阵——是有完全的标准正交特征向量系。

所谓"完全"是指有 n 个独立的向量,它可构成 n 维空间的基。

在讨论正规矩阵时,还得出如下的重要结论。

A 正交相似实对角阵$\Longleftrightarrow A$ 是实对称阵,

A 酉相似实对角阵$\Longleftrightarrow A$ 是 Hermite 阵,

A 是酉相似复对角阵$\Longleftrightarrow A$ 是正规阵,

A 是正规阵,它是酉阵$\Longleftrightarrow |\lambda|=1$。

一个线性变换的性质是与坐标无关的,但讨论或计算时却往往要在一定坐标系之下进行,因而在相似变换下,一个矩阵的"最简单"形式无疑是很重要的。为了说明这个问题下面举一个例子。

例 1　设 A 是 n 阶实阵,它满足条件 $A^2=I$, 求它在相似变换下,所采取的最简形式。

解　事实上,由 $A^2=I$,可得二个结论

(i) $\lambda^2=1$,　(ii) $r(I+A)+r(I-A)=n$。

(i) 表明 A 的特征值为 1 或 -1。设 $r(I+A)=r_1, r(I-A)=r_2$ (ii) 表明 $\{n-r_1\}+\{n-r_2\}=n$，即对应于 1 的特征向量个数与对应于 -1 的特征向量个数之和恰为 n 个。若用 V_1 记对应于 $\lambda=1$ 的特征子空间(见第二章第一节)；用 V_{-1} 记对应于 $\lambda=-1$ 的特征子空间，则有 $R^n=V_1\oplus V_{-1}$。设 $\alpha_1\cdots\alpha_{r_1}$ 是 V_1 的基、$\beta_1\cdots\beta_{r_2}$ 是 V_{-1} 的的基，则 $\{\alpha_1\cdots\alpha_{r_1},\beta_1\cdots\beta_{r_2}\}$ 是 R^n 的基。如果将 A 看成 R^n 到 R^n 的线性变换，在此基下线性变换 A 可用矩阵

$$D=\begin{bmatrix} \overbrace{\begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix}}^{r_1} & \\ & \overbrace{\begin{matrix}-1 & & \\ & \ddots & \\ & & -1\end{matrix}}^{r_2} \end{bmatrix}$$

表示，或讲 A 与 D 相似。

上面的结果说明，要求 A 的最简单形式，先要将空间 R^n 分解为若干个子空间的直和(这里是 A 的特征子空间。另一个例子可参考第二章例 1)，然后在每一个子空间里选择基并成 R^n 的基。将 A 视为 R^n 到 R^n 的线性变换，这样可求得在新基下矩阵 A 的另一表示。

本章的中心定理如下。

如果矩阵 A 有 s 个线性独立的特征向量，则它相似于矩阵 J，即

$$M^{-1}AM=J=\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix}，记为 J=J_1\oplus J_2\oplus\cdots\oplus J_s。$$

这里 J_i 是下面形状的矩阵：

$$J_i=\begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}_{n_i\times n_i}, \quad n_1+n_2+\cdots+n_s=n_\circ$$

J_i 称为 Jordan 块，若完整写时，可写为 $J_{(n_i)}(\lambda_i)$ 或 $J_i(\lambda_i)$。

例 2

$$J=\begin{bmatrix}8&1&0&0&0\\0&8&0&0&0\\0&0&0&1&0\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix}=\begin{bmatrix}8&1\\0&8\end{bmatrix}\oplus\begin{bmatrix}0&1\\0&0\end{bmatrix}\oplus[0]。$$

如果矩阵 A 与 J 相似，有 $P^{-1}AP=J$。设 $P=[\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5]$，由 $AP=PJ$ 可得 P，各列向量应满足的条件为

$A\alpha_1=8\alpha_1, A\alpha_2=\alpha_1+8\alpha_2, A\alpha_3=0\cdot\alpha_3,\ A\alpha_4=0\cdot\alpha_4+\alpha_3,$

$A\alpha_5=0\cdot\alpha_5$。

由此可知独立的特征向量有 3 个 α_1，α_3，α_5，它分别对应 3 个 Jordan 块，但 α_2，α_4 都不是特征向量。下面讨论如何计算一个矩阵 A 的 Jordan 标准型。

第二节　不变子空间与导出算子

1. 不变子空间

定义　设 $\sigma\in\mathrm{Hom}(V,V)$，又 V_s 是 V 的一个子空间，如果对任一 $\alpha\in V_s$ 有 $\sigma(\alpha)\in V_s$，则称 V_s 是线性变换 σ 的不变子空间，或称 σ-不变子空间。

例 3　给定一个 n 阶矩阵 A，有二个重要的 A 的不变子空间，它们是 $R(A)$ 与 $N(A)$。一般 $\sigma\in\mathrm{Hom}(V,V)$，同理 $\mathrm{Im}\,\sigma$ 与 $\mathrm{Ker}\,\sigma$ 也是 σ 的不变子空间。

证　现证明 $\mathrm{Im}\,\sigma$ 与 $\mathrm{Ker}\,\sigma$ 是 σ 的不变子空间。

由 $\alpha\in\mathrm{Im}\,\sigma$，证明 $\sigma(\alpha)\in\mathrm{Im}\,\sigma$。实际上因 $\alpha\in\mathrm{Im}\,\sigma$，所以 $\exists\beta\in V$，$\alpha=\sigma(\beta)$。因而 $\sigma(\alpha)=\sigma(\sigma(\beta))\in\mathrm{Im}\,\sigma$。$\mathrm{Ker}\,\sigma$ 是 σ 的不变子空间易证。

例 4　在前节例 1，V_1，V_{-1} 都是 A 的不变子空间。一般而言，$\sigma\in\mathrm{Hom}(V,V)$，用 V_λ 表示对应于 σ 的特征值 λ 的特征子空间，则 V_λ 是 σ 的不变子空间。这点由特征子空间的定义直接可得。

下面的例子说明不变子空间，在研究线性算子时的作用。

例 5 设 $\sigma \in \mathrm{Hom}(V,V)$，$V$是 5 维线性空间，$V=V_1 \oplus V_2$。$V_1,V_2$ 是二个 σ 的不变子空间，分别为 3 维与 2 维。如果 $\alpha_1,\alpha_2,\alpha_3$ 是 V_1 的基，β_1,β_2 是 V_2 的基，那么 σ 在基$\{\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2\}$下的矩阵是分块对角阵(或称准对角阵)。

解 因

$$
\begin{aligned}
\sigma(\alpha_1) &= b_{11}\alpha_1 + b_{21}\alpha_2 + b_{31}\alpha_3 \\
\sigma(\alpha_2) &= b_{12}\alpha_1 + b_{22}\alpha_2 + b_{32}\alpha_3 \\
\sigma(\alpha_3) &= b_{13}\alpha_1 + b_{23}\alpha_2 + b_{33}\alpha_3 \\
\sigma(\beta_1) &= \qquad\qquad\qquad\qquad c_{11}\beta_1 + c_{21}\beta_2 \\
\sigma(\beta_2) &= \qquad\qquad\qquad\qquad c_{12}\beta_1 + c_{22}\beta_2,
\end{aligned}
$$

很明显变换 σ，在此基下的方程有这种形式是因为这个基由各个不变子空间的基并成的。这时 σ 的矩阵表示为

$$
A=\left[\begin{array}{ccc|cc}
b_{11} & b_{12} & b_{13} & & \\
b_{21} & b_{22} & b_{23} & & \\
b_{31} & b_{32} & b_{33} & & \\
\hline
 & & & c_{11} & c_{12} \\
 & & & c_{21} & c_{22}
\end{array}\right]=\begin{bmatrix} B & \\ & C \end{bmatrix}=B\oplus C.
$$

按上面的方法可得下面的定理。

定理 1 V 是 n 维空间，$\sigma \in \mathrm{Hom}(V,V)$。如果 V 可表示为 σ 不变子空间的直和，即 $V=V_1\oplus V_2\oplus+\cdots\oplus V_s$。并选 V 的基为各个 V_i 的基的并。在该基下σ 的矩阵采取分块对角阵的形式，即

$$A=A_1\oplus A_2\oplus\cdots\oplus A_s.$$

下面取一个 Jordan 块为

$$
J_{(s)}(\lambda_0)=\begin{bmatrix}
\lambda_0 & 1 & & \\
 & \lambda_0 & \ddots & \\
 & & \ddots & 1 \\
 & & & \lambda_0
\end{bmatrix}_{s\times s}.
$$

如果一个 s 阶矩阵 A 与 $J_{(s)}(\lambda_0)$相似，那么 $P^{-1}AP=J_{(s)}(\lambda_0)$，

设 $P=[\alpha_1\cdots\alpha_s]$，它的各个列向量应满足的关系是

$$A\alpha_1=\lambda_0\alpha_1,\quad A\alpha_2=\lambda_0\alpha_2+\alpha_1,\quad A\alpha_3=\lambda_0\alpha_3+\alpha_2,\cdots,$$

$$A\alpha_s=\lambda_0\alpha_s+\alpha_{s-1}。$$

或可写成为

$$(A-\lambda_0I)\alpha_1=0,(A-\lambda_0I)\alpha_2=\alpha_1,(A-\lambda_0I)\alpha_3=\alpha_2,\cdots,$$

$$(A-\lambda_0I)\alpha_s=\alpha_{s-1}。$$

即：$(A-\lambda_0I)\alpha_1=0$，$(A-\lambda_0I)^2\alpha_2=0$，$(A-\lambda_0I)^3\alpha_3=0,\cdots,$

$$(A-\lambda_0I)^s\alpha_s=0。$$

因而 $\alpha_1\cdots\alpha_s$ 都是方程 $(A-\lambda_0I)^sx=0$ 的解。

定义 如果有一向量 $v\neq0$ 数 λ_0 及正整数 k，有 $(A-\lambda_0I)^kv=0$，则称 v 为矩阵 A 的广义特征向量。而 λ_0 实际上是特征值。容易证明，矩阵 A 对应于特征值 λ_0 的广义特征向量构成的集合是一子空间，它称为 λ_0 的广义特征子空间。

命题1 矩阵 A 的 λ_0 的广义特征子空间为 A 的不变子空间。

（证略）

实际上矩阵理论最重要的 A 的不变子空间，就是前面讲的四个：$R(A)$、$N(A)$、A 的 λ_0 特征子空间，A 的 λ_0 广义特征子空间。

2．导出算子

这小节主要讲述前面引入 σ 不变子空间，在分析线性算子 σ 时的作用。

定义 设 $\sigma\in\mathrm{Hom}(V,V)$，又 V_s 是 σ 不变子空间。因为 $\sigma(\alpha)\in V_s$，$\forall\alpha\in V_s$，所以可以定义一个算子 σ'，$\sigma'(\alpha)=\sigma(\alpha)$，$\forall\alpha\in V_s$。$\sigma'$ 本质上是将 σ 限制在不变子空间 V_s 上来考察。它称为算子 σ 在不变子空间 V_s 上的导出算子，用记号 $\sigma|V_s$ 表示。

有了导出算子这概念之后，便可了解将一个线性空间 V 表示为 σ 的不变子空间的直和，在分析算子 σ 时起何作用。

设 $V=V_1\oplus V_2\oplus\cdots\oplus V_s$，引入记号 $\sigma_i=\sigma|V_i$。因此对任何

$\alpha \in V$，有 $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_s, \alpha_i \in V_i$。这样

$$
\begin{aligned}
\sigma(\alpha) &= \sigma(\alpha_1) + \sigma(\alpha_2) + \cdots + \sigma(\alpha_s) \\
&= \sigma_1(\alpha_1) + \sigma_2(\alpha_2) + \cdots + \sigma_s(\alpha_s)。
\end{aligned} \tag{4-1}
$$

式(4-1)可以这样解释：如果给出一个 $\sigma \in \mathrm{Hom}(V,V)$，且 V 可表示为 σ 的不变子空间的直和，则 σ 对空间任一向量 α 的作用可看作为它在各个不变子空间的导出算子，对 α 在该不变子空间分量作用的叠加的结果。这种现象表明，线性算子具有一种"可分裂"性质。现将这种性质写成下面的形式：

$$
\sigma = \sigma_1 \oplus \sigma_2 \oplus \cdots \oplus \sigma_s,
$$

并称它为线性算子是各导出算子的直和。

例 6 如把前面的对合矩阵 $A^2 = I$（例 1）视为 R^n 的一个线性算子，$R^n = V_1 \oplus V_{-1}$，则对任何 $x \in R^n$ 有 $x = y + z$，$y \in V_1$，$z \in V_{-1}$。

$Ax = Ay + Az = y - z$（记住 y, z 分别为特征值 1，-1 的特征向量）。如果引入导出算子 $A|V_1 = A_1, A|V_2 = A_2$，那么 A_1 在 V_1 的作用相当是一个恒等变换，而 A_2 在 V_2 的作用相当是一个"反射"，$A = A_1 \oplus A_2$。这时 A 对 x 的作用便看成一个恒等变换与一个反射变换的叠加。

例 7 第二章的例 1，$\sigma \in \mathrm{Hom}(V,V)$，$\sigma^2 = \sigma$，$V = R(\sigma) \oplus \mathrm{Ker}\sigma$。

$R(\sigma)$ 与 $\mathrm{Ker}\,\sigma$ 都是 σ 不变子空间，且在 $\alpha \in R(\sigma)$ 有 $\sigma(\alpha) = \alpha$，因此定义导出算子 $\sigma_1 = \sigma|_{R(\sigma)}$，它是一个恒等算子。而 $\sigma_2 = \sigma|_{\mathrm{Ker}\sigma}$ 是一个零算子，这样对任何一个 $\alpha \in V$，有 $\alpha = \alpha_1 + \alpha_2$，$\alpha_1 \in R(\sigma)$，$\alpha_2 \in \mathrm{Ker}\,\sigma$。则 $\sigma(\alpha) = \sigma(\alpha_1) + \sigma(\alpha_2) = \sigma_1(\alpha_1) + \sigma_2(\alpha_2) = \alpha_1$，可以看成是一个恒等算子与一个零算子叠加的结果。

例 8 设 $\sigma \in \mathrm{Hom}(R^3, R^3)$，$\sigma(\alpha) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta, z)$，$\alpha = (x, y, z)$（这是一个绕 z 轴的旋转）。

σ 的不变子空间有二个，一个是 x, y 平面，另一个为 z 轴。而 R^3 为这两个不变子空间的直和。这时算子 σ，在这两个不变子空

间的导出算子为：

$$\sigma|_{xy\text{平面}}=\sigma_1(\text{平面绕 }O\text{ 点的旋转}),$$

$$\sigma|_{z\text{轴}}=\sigma_2(z\text{ 轴上的恒等变换}),$$

所以 $\sigma=\sigma_1\oplus\sigma_2$。

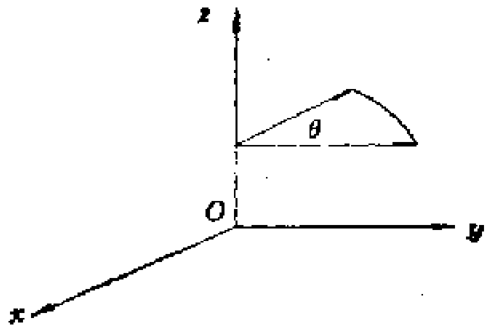

图 4-1

第三节 特征值全为零的矩阵的 Jordan 标准型的计算

1. 幂零矩阵

(1) 幂零矩阵的性质

一个矩阵的特征值全为零，说明它的零化多项式为 $f(x)=x^r$，因此它的最小多项式为 $m(x)=x^m$。换句话说，有 $A^m=0$ 但 $A^{m-1}\neq 0$，A 是指标为 m 的幂零矩阵。事实上有下面的命题。

命题 2 一 n 阶矩阵 A，它是幂零矩阵$\Longleftrightarrow$特征值全为零。

(2) 几个幂零矩阵的例：

例 9 $A_1=\begin{bmatrix}2 & 4\\ -1 & -2\end{bmatrix}$的指标为 2。

例 10

$$A_2=\begin{bmatrix}0&1&1&0&1\\0&0&1&1&1\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix}\text{的指标为 }3$$

例 11

$$N=\begin{bmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&&&&\\0&0&0&\cdots&1\\0&0&0&\cdots&0\end{bmatrix}_{n\times n}\quad \text{的指标为 } n。$$

例 11 是一个最重要的幂零矩阵。如果它是 k 阶方阵，则指标便是 k，可称它是 k 阶幂零块。

2．幂零矩阵的标准型定理

定理 2 设 A 是指标为 p 的 n 阶幂零矩阵，则它一定与下面的分块对角阵相似

$$N=\begin{bmatrix}N_1&&&\\&N_2&&\\&&\ddots&\\&&&N_s\end{bmatrix}(=N_1\oplus N_2\oplus\cdots\oplus N_s)\qquad(4\text{-}2)$$

其中 N_i 是 n_i 阶的幂零块（$n_1+n_2+\cdots+n_s=n$）。另外可断言 n_i 中至少有一个为 p，而其它 $n_i\leqslant p$。幂零块的数目是 A 的零度，而各阶块的数目由 A 确定（N 称为幂零阵在相似变换下的标准型）。（证略）

例 12 设 A_1 为指标 $p=2$ 的矩阵。

由题设可知 A_1 的标准型只有一种可能

$$A_1\sim\begin{bmatrix}0&1\\0&0\end{bmatrix}$$

例 13 设 A_2 是指标 $p=3$ 的矩阵，且零度为 3。

由题设可断言至少有一 3 阶块，且幂零块为 3 块。它的标准型为：

$$N=\left[\begin{array}{ccc|c|c}0&1&0&&\\0&0&1&&\\0&0&0&&\\\hline&&&0&\\\hline&&&&0\end{array}\right]。$$

3. 幂零矩阵标准型的计算

由幂零矩阵的标准型定理可知，$P^{-1}AP=N$(见式(4-2))，因此矩阵 A 的零度、秩以及所论及其它相似不变量都是与 N 相同。下面的讨论请注意以下的事实。

(i) N 的秩是各子块的秩之和，N 的零度是各子块零度之和。

(ii) N 自乘 i 次，是各子块自乘 i 次的直和，即

$$N^i=N_1^i\oplus N_2^i\oplus\cdots\oplus N_s^i。$$

(iii) 各子块 N_j 自乘一次零度增加1。自乘 n_j 次，零度为 n_j(这是 $N_j^{n_j}=0$)。如继续自乘，零度不变。

下面通过各 N^i 的零度或秩的计算(也就是通过 A^i 的零度或秩的计算)，来确定 N 的应具有的形式。

(i) 由于各子块 N_j 的零度均为1，因此 N 的子块的数目 $s=N$ 的零度 $=A$ 的零度。

(ii) 由于 A 的指数 $=p$(即 N 的指数 $=p$)，所以 A 的最小多项式 $m(\lambda)=\lambda^p$，而 $N^p=N_1{}^p\oplus\cdots\oplus N_s^p=0$(零阵)，因此各 N_j 中至少有一个是 p 阶的幂零块，且其余子块的阶均不会超过 p 阶。

(iii) $N^2=N_1^2\oplus N_2^2+\cdots+N_s^2$。

由前面的(iii)可知，当子块的阶为1时，该块在 N^2 中零度不变。但阶≥2时，则零度增加1。

如果用 η_1 表示 N 的零度，η_2 表示 N^2 的零度，则 $\eta_2-\eta_1$ 表示 N 中阶≥2的子块数目。

(iv) 同理考察 $N^3=N_1^3\oplus N_2^3\oplus\cdots\oplus N_s^3$。

在这个乘积中所有阶≤2的子块零度不变，而阶≥3的子块零度增加1。如果用 η_3 表示 N^3 的零度，$\eta_3-\eta_2$ 表示 N 中阶≥3的子块数目。这样可得

$$\begin{aligned}N\text{ 中 2 阶幂零块的数目}&=(\text{阶}\geq 2\text{ 的子块数目})\\&\quad-(\text{阶}\geq 3\text{ 的子块数目})\\&=(\eta_2-\eta_1)-(\eta_3-\eta_2)=2\eta_1-\eta_2-\eta_3。\end{aligned}$$

(v) 按上面的计算方法，可知有下面公式成立。

N 中 k 阶幂零块的数目

=（阶$\geqslant k$的子块数目）-（阶$\geqslant k+1$的子块数目）

$$=(\eta_k-\eta_{k-1})-(\eta_{k+1}-\eta_k)=2\eta_k-\eta_{k-1}-\eta_{k+1}。\qquad (4-3)$$

式中η_k表示N^k的零度。

如果用秩表示上面的结果，则只需用$r_k=n-\eta_k$代入式(4-3)（r_k是N^k的秩）便可有

$$N\text{ 中 }k\text{ 阶幂零块的数目}=r_{k+1}+r_{k-1}-2r_k。\qquad (4-4)$$

上面的公式在阶不高时，计算很方便。

例 14 证明下面矩阵是幂零阵，并确定其标准型。

$$A=\begin{bmatrix}-1 & 0 & -1 & -1\\ 1 & 1 & 0 & 1\\ 1 & 1 & 0 & 1\\ 0 & -1 & 1 & 0\end{bmatrix}$$

解 由直接计算可知$A^2=0$，因此可知A的幂零指标为2。即A的标准型中至少有一个2阶的幂零块。

计算A的零度或秩，容易得知$r(A)=2$，所以零度$=4-2=2$，即A的标准型中有2个幂零块。

综合上面的计算可知A的标准型为

$$N=\left[\begin{array}{cc|cc}0 & 1 & & \\ 0 & 0 & & \\ \hline & & 0 & 1\\ & & 0 & 0\end{array}\right]。$$

例 15 试求例14的变换矩阵P。

解 因为$P^{-1}AP=N$，它等价于$AP=PN$。令$P=[\alpha_1,\alpha_2,\alpha_3,\alpha_4]$便可得

$$A[\alpha_1,\alpha_2,\alpha_3,\alpha_4]=[\alpha_1,\alpha_2,\alpha_3,\alpha_4]\left[\begin{array}{cc|cc} 0 & 1 & & \\ 0 & 0 & & \\ \hline & & 0 & 1 \\ & & 0 & 0 \end{array}\right]。$$

因此可得 P 的各个列向量应满足的方程组为

$$A\alpha_1=0\cdot\alpha_1, A\alpha_2=0\cdot\alpha_2+\alpha_1, A\alpha_3=0\cdot\alpha_3, A\alpha_4=0\cdot\alpha_4+\alpha_3。$$

解方程得

$$A\alpha_1=0, A\alpha_2=\alpha_1, A\alpha_3=0, A\alpha_4=\alpha_3,$$

其中 α_1,α_3 是特征向量，而 α_2，α_4 是广义特征向量。注意到求解的四个方程组的系数都相同。在求解时可采取下面的格式。

$$\left[\begin{array}{cccc|c} -1 & 0 & -1 & -1 & y_1 \\ 1 & 1 & 0 & 1 & y_2 \\ 1 & 1 & 0 & 1 & y_3 \\ 0 & -1 & 1 & 0 & y_4 \end{array}\right] \longrightarrow$$

$$\left[\begin{array}{cccc|c} -1 & 0 & -1 & -1 & y_1 \\ 0 & 1 & -1 & 0 & y_1+y_2 \\ 0 & 0 & 0 & 0 & y_3-y_2 \\ 0 & -1 & 1 & 0 & y_4 \end{array}\right] \longrightarrow$$

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 1 & -y_1 \\ 0 & 1 & -1 & 0 & y_1+y_2 \\ 0 & 0 & 0 & 0 & y_3-y_2 \\ 0 & 0 & 0 & 0 & y_1+y_2+y_4 \end{array}\right]。$$

计算的结果得出 A 的梯形阵以及 $Ax=b$ 相容性条件。

设 $b=(y_1,y_2,y_3,y_4)^T$，要方程 $Ax=b$ 有解，向量 b 满足条件 $y_3-y_2=0, y_1+y_2+y_4=0$。

由 $\begin{cases} x_1+x_3+x_4=0 \\ x_2-x_3=0, \end{cases}$ $\begin{cases} x_1=-x_3-x_4 \\ x_2=x_3, \end{cases}$

可解得

$$\alpha_1=(-1\quad 1\quad 1\quad 0)^T,\alpha_3=(-1\quad 0\quad 0\quad 1)^T。$$

由 $A\alpha_2=\alpha_1$，$\begin{cases}x_1=1-x_3-x_4\\x_2=x_3\end{cases}$ 可得

$$\alpha_2=(-1\quad 1\quad 1\quad 1)^T。$$

由 $A\alpha_4=\alpha_3$，$\begin{cases}x_1=1\ -x_3-x_4\\x_2=-1+x_3\end{cases}$ 可得

$$\alpha_4=(1\quad -1\quad 0\quad 0)^T。$$

在求 α_2,α_4 时，x_3,x_4 是参数，只要给以适当值，即可求得 α_2，α_4。在求解的过程发现，变换矩阵 P 并不是唯一的。

$$P=\begin{bmatrix}-1&-1&-1&1\\1&1&0&-1\\1&1&0&0\\0&1&1&0\end{bmatrix}。$$

第四节　Jordan 标准型的计算(一般情形)

1. Jordan 矩阵的特征多项式与最小多项式

首先讨论 Jordan 块的特征多项式与最小多项式，设 $J_{(k)}(\lambda_0)$

$=\begin{bmatrix}\lambda_0&1&&\\&\lambda_0&\ddots&\\&&\ddots&1\\&&&\lambda_0\end{bmatrix}_k$，显然它的最小多项式 $m(\lambda)=(\lambda-\lambda_0)^r$。由于 $J_{(k)}(\lambda_0)-\lambda_0 I=\begin{bmatrix}0&1&&\\&0&\ddots&\\&&\ddots&1\\&&&0\end{bmatrix}$ 是个 k 阶幂零阵，因此 $m(J_{(k)}(\lambda_0))=0$，即 $r=k$。它的特征多项式亦为 $f(\lambda)=(\lambda-\lambda_0)^k$。由此得出结论：一个Jordan块它的特征多项式与最小多项式相同。

例16 设

$$J=\begin{bmatrix}
2 & 1 & 0 & & & & & & & \\
0 & 2 & 1 & & & & & & & \\
0 & 0 & 2 & & & & & & & \\
 & & & 2 & 1 & & & & & \\
 & & & 0 & 2 & & & & & \\
 & & & & & 1 & 1 & 0 & & \\
 & & & & & 0 & 1 & 1 & & \\
 & & & & & 0 & 0 & 1 & & \\
 & & & & & & & & 0 & \\
 & & & & & & & & & 0
\end{bmatrix}$$

$$=J_{(3)}(2)\oplus J_{(2)}(2)\oplus J_{(3)}(1)\oplus J_{(1)}(0)\oplus J_{(1)}(0)$$

则它们各个 Jordan 块的特征多项式按顺序为：

$$f_1(\lambda)=(\lambda-2)^3,\ f_2(\lambda)=(\lambda-2)^2,\ f_3(\lambda)=(\lambda-1)^3,$$

$$f_4(\lambda)=\lambda, f_5(\lambda)=\lambda。$$

$$J\text{ 的特征多项式}=f_1(\lambda)\cdot f_2(\lambda)\cdot f_3(\lambda)\cdot f_4(\lambda)\cdot f_5(\lambda)$$

$$=(\lambda-2)^5(\lambda-1)^3\lambda^2。$$

将矩阵 J 的各项 Jordan 块的最小多项式称为 J 的初等因子，它们全体便称为矩阵 J 的初等因子组，因此 J 的特征多项式为 J 的全部初等因子的乘积。

现在再看一下最小多项式的情形(可参考第二章的例14，15)。它实际上是各块最小多项式的最小公倍数，因此它是 $m(\lambda)=(\lambda-2)^3(\lambda-1)^3\lambda$。

亦可这样认为：将 J 的初等因子组按特征值分成小组，然后取每个小组幂次最高初等因子的乘积，便是 J 的最小多项式。在这个例子里便是 $(\lambda-2)^3, (\lambda-2)^2; (\lambda-1)^3; \lambda, \lambda$。上面的讨论结果，主要能从一个矩阵的特征多项式与最小多项式大致推断一个矩阵的 Jordan 标准型。因为当 A 与 J 相似时，它们的特征多项

式与最小多项式是一样的。

例 17 一个 6 阶矩阵，它的特征多项式与最小多项式分别为

$$f(\lambda)=(\lambda-2)^4(\lambda-3)^2,\quad m(\lambda)=(\lambda-2)^2(\lambda-3)^2。$$

试写出该矩阵可能的 Jordan 标准型。

解 由最小多项式可知初等因子组里有$(\lambda-2)^2$，$(\lambda-3)^2$。它们是诸 Jordan 块的最小多项式里次数最高的。另一方面诸初等因子的乘积为特征多项式，所以与特征值 $\lambda=3$ 相当的初等因子仅有$(\lambda-3)^2$。而与 $\lambda=2$ 相当的初等因子之一为 $(\lambda-2)^2$。余下，因乘积为$(\lambda-2)^4$，所以只有二种可能：$(\lambda-2)^2$；或$(\lambda-2)$，$(\lambda-2)$。因此这一个 6 阶矩阵它的初等因子组可能有

$(\lambda-2)^2,(\lambda-2)^2,(\lambda-3)^2$，或$(\lambda-2)^2,(\lambda-2),(\lambda-2)$，$(\lambda-3)^2$。

对应的 Jordan 标准型分别为：

$$\begin{bmatrix} 2 & 1 & & & & \\ 0 & 2 & & & & \\ & & 2 & 1 & & \\ & & 0 & 2 & & \\ & & & & 3 & 1 \\ & & & & 0 & 3 \end{bmatrix},\quad \begin{bmatrix} 2 & 1 & & & & \\ 0 & 2 & & & & \\ & & 2 & & & \\ & & & 2 & & \\ & & & & 3 & 1 \\ & & & & 0 & 3 \end{bmatrix}。$$

2. Jordan 标准型的计算

为了要确定对应于特征值 λ_0 的 Jordan 子块的数目，可看前面例 16 的结果。

(i) 对应于 $\lambda=2$ 共有二个子块阶数和为 5 阶（它是 $\lambda=2$ 的代数重数），其中一个块是 3 阶，它是最小多项式中 $\lambda-2$ 的次，

(ii) 块数是 2，它恰为 $\lambda=2$ 的几何重数（类似幂零矩阵，每一个子块有一个特征向量与之对应）。这样，对于 $\lambda=2$ 的 Jordan 子块结构可写成 $\underset{(3,2)}{\lambda=2}$ 类似其它的为 $\underset{(3)}{\lambda=1}$，$\underset{(1,1)}{\lambda=0}$ 或可紧缩写成

为：

$$\lambda = \begin{array}{ccc} 2 & 1 & 0 \\ \{(3,2) & (3) & (1,1)\} \end{array}$$

上面数组可综述如下：特征值 $\lambda=2$ 时 Jordan 子块有 2 块（即 $\lambda=2$ 的几何重数为 2)。其中一个为 3 阶，另一个是 2 阶。类似可解释其它二组。这组数亦称为 Jordan 标准型的特征。显然若能确定特征，一个标准型便完全确定。

例 18

求 $A=\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ -3 & 2 & 1 & 1 \\ 3 & -6 & 1 & 4 \end{bmatrix}$ 的 Jordan 标准型。

解 先求出特征多项式 $p(\lambda)=(\lambda-2)^3(\lambda+1)$。因此 A 的 Jordan 标准形为 $\left[\begin{array}{ccc|c} 2 & * & & \\ & 2 & * & \\ & & 2 & \\ \hline & & & -1 \end{array}\right]$。

实际上只要确定 $\lambda=2$ 时的 Jordan 块的数目。因此只需计算 $A-2I$ 的秩：$A-2I=\begin{bmatrix} -2 & 0 & 0 & 1 \\ 1 & -2 & 0 & 1 \\ -3 & 2 & -1 & 1 \\ 3 & -6 & 1 & 2 \end{bmatrix}$，

经计算可知，$r(A-2I)=3$。即 $\lambda=2$ 的特征向量只有一个，几何重数为 1，所以 $\lambda=2$ 的 Jordan 块只有一块。因而 $A\sim J$

$$J=\left[\begin{array}{ccc|c} 2 & 1 & & \\ & 2 & 1 & \\ & & 2 & \\ \hline & & & -1 \end{array}\right],$$

J 的特征为 $\lambda = \begin{Bmatrix} 2 & -1 \\ (3) & (1) \end{Bmatrix}$。

例 19 一个 5 阶矩阵，它的特征多项式 $p(\lambda)=(\lambda-3)^4(\lambda+2)$，且 $r(A-3I)=2$，试求该矩阵的 Jordan 标准型。

解 只要能写出该矩阵的特征即可。$\lambda=-2$ 的 Jordan 块只有一个 1 阶块。因 $r(A-3I)=2$，所以 $\lambda=3$ 的几何重数 $=5-2=3$。因此对应于 $\lambda=3$ 的 Jordan 块有 3 块。这样可写出 J 的特征为

$$\lambda=\begin{Bmatrix} 3 & -2 \\ (2,1,1) & (1) \end{Bmatrix},$$

$$J=\left[\begin{array}{cc|c|c|c} 3 & 1 & & & \\ & 3 & & & \\ \hline & & 3 & & \\ \hline & & & 3 & \\ \hline & & & & -2 \end{array}\right]。$$

但对于任意矩阵，上面的信息还是不够的。为了要考察对应于 $\lambda=\lambda_0$ 的 Jordan 块，设

$$J=\begin{bmatrix} \lambda_0 & * & & & & \\ & \ddots & \ddots & & & \\ & & \lambda_0 & & & \\ & & & \lambda_1 & * & \\ & & & & \ddots & * \\ & & & & & \lambda_s \end{bmatrix},\lambda_1,\cdots\lambda_s \text{均} \neq \lambda_0$$

考察矩阵 $J-\lambda_0 I=\left[\begin{array}{ccc|c} 0 & * & & \\ & \ddots & * & \\ & & 0 & \\ \hline & & & D \end{array}\right]$，其中 D 是非奇异矩阵。

而原对应于特征值 λ_0 的 Jordan 块都变成幂零块。因此实际上只

要确定矩阵 $J-\lambda_0 I$ 的幂零子块的数目。但 D 是非奇异矩阵，因此计算时所采用的方法和前面确定幂零矩阵的幂零块的方法相同。又 $A\sim J$，所以 $P^{-1}(A-\lambda_0 I)P=J-\lambda_0 I$，因此可通过计算 $A-\lambda_0 I$ 的秩去确定幂零块的数目。前面计算幂零阵时($\lambda=0$)的公式与格式可复述如下：

$$
\begin{aligned}
k \text{ 阶幂零子块的数目} &= (\eta_k-\eta_{k-1})-(\eta_{k+1}-\eta_k)\\
&=2\eta_k-\eta_{k-1}-\eta_{k+1}=(r_{k+1}-r_k)-(r_k-r_{k-1})\\
&=r_{k+1}+r_{k-1}-2r_k。
\end{aligned}
$$

即只要计算矩阵 A，A^2，…，A^k，…相应的秩，并分别记为 r_1，r_2，…，r_k，…，构造下面的差分表为：

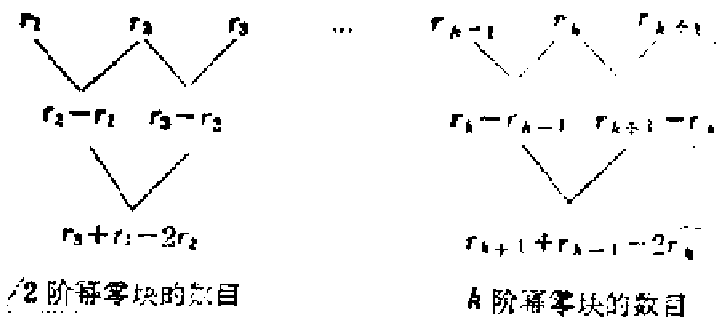

图 4-2

因此要确定对应于 λ_0 的 Jordan 块的数目便计算得出。

$A-\lambda_0 I$，$(A-\lambda_0 I)^2$，…，$(A-\lambda_0 I)^k$，…诸矩阵的秩，仍用 r_1，r_2，…，r_k，…分别表示它们的秩。上面的差分表仍可用在 $\lambda\neq 0$ 的计算中。

例 20 设 $A=\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1\\ 2 & 1 & -1 & -1 & 0 & -1\\ 0 & 0 & 2 & 1 & 0 & 0\\ 0 & 0 & 0 & 2 & 0 & 0\\ 0 & 0 & 0 & 0 & 2 & 0\\ -1 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$，它的特征多项式 $p(\lambda)=(\lambda-1)^3(\lambda-2)^3$。求 A 的 Jordan 标准型。

解　先求 $r(A-I)$。由 $A-I=\begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 1 \\ 2 & 0 & -1 & -1 & 0 & -1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$，

求得 $r=5$。

$\lambda=1$ 的特征向量仅有一个，这表示对应于 $\lambda=1$ 的 Jordan 块的数目是 1。又由于 $r(A-2I)=4$，对应于 $\lambda=2$ 的特征向量是 2 个，因此对应于 $\lambda=2$ 的 Jordan 块是共有 2 块。对于这个问题上面的信息已足够确定矩阵 A 的标准型了。

$$A\sim J=\left[\begin{array}{ccc|cc|c} 1 & 1 & 0 & & & \\ 0 & 1 & 1 & & & \\ 0 & 0 & 1 & & & \\ \hline & & & 2 & 1 & \\ & & & 0 & 2 & \\ \hline & & & & & 2 \end{array}\right]。$$

例 21　求 $A=\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ 的 Jordan 标准型。

解　首先注意到 A 的秩为 1，因而它的 Jordan 标准型的秩为 1，故可设想

$$J=\begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}。$$

又因为$A\begin{bmatrix}1\\1\\1\\1\end{bmatrix}=4\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$,因此$\lambda=4$。所以$A\sim J=\begin{bmatrix}4&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}$。

若用计算秩的方法亦可很快确定A的标准型。首先可计算得上面矩阵A的特征多项式为$p(\lambda)=\lambda^3(\lambda-4)$,显然对应于$\lambda=4$的Jordan块只能是一个1阶Jordan块,对应$\lambda=0$的Jordan块的数目只要计算$\lambda=0$的独立的特征向量的个数,$r(A)=1$。因此独立的特征向量有3个,即$\lambda=0$的Jordan块应是3块。

有些场合只要知道一个矩阵A的标准型已足够了。但在某些场合下,还需要知道相似变换的变换矩阵P。

例22 求矩阵A的Jordan标准型,并求变换矩阵P。设

$$A=\begin{bmatrix}0&0&0&4\\1&0&0&-4\\0&1&0&-3\\0&0&1&4\end{bmatrix}。$$

解 矩阵A的特征多项式为$p(\lambda)=\lambda^4-4\lambda^3+3\lambda^2+4\lambda-4=(\lambda-2)^2(\lambda-1)(\lambda+1)$。要确定$A$的Jordan标准型的结构,只要确定对应于特征值$\lambda=2$的Jordan块的数目,计算$r(A-2I)(=3)$。对应于$\lambda=2$的特征向量只有1个,因此对应$\lambda=2$的Jordan块只有1块,所以

$$P^{-1}AP=\begin{bmatrix}2&1&0&0\\0&2&0&0\\0&0&1&0\\0&0&0&-1\end{bmatrix},\quad 即\ AP=P\begin{bmatrix}2&1&0&0\\0&2&0&0\\0&0&1&0\\0&0&0&-1\end{bmatrix}。$$

如果$P=[\alpha_1,\alpha_2,\alpha_3,\alpha_4]$,便可得解$\alpha_1,\alpha_2,\alpha_3,\alpha_4$的方程组为:

$$A\alpha_1=2\alpha_1,A\alpha_2=2\alpha_2+\alpha_1,A\alpha_3=\alpha_3,A\alpha_4=-\alpha_4。$$

则

$$(A-2I)\alpha_1=0,(A-2I)^2\alpha_2=0,(A-I)\alpha_3=0,$$
$$(A+I)\alpha_4=0。$$

计算 $\lambda=2$ 的特征值时，有

$$\begin{bmatrix}-2&0&0&4\\1&-2&0&-4\\0&1&-2&-3\\0&0&1&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=0,\text{则}\begin{cases}-2x_1+4x_2=0\\x_1-2x_2-4x_4=0\\x_3+2x_4=0\end{cases}$$

可解得 $x_1=2x_4,x_2=-x_4,x_3=-2x_4,\alpha_1=(2\ -1\ -2\ \ 1)^T$。

求 $\lambda=2$ 的广义特征向量为

$$\begin{bmatrix}-2&0&0&4\\1&-2&0&-4\\0&1&-2&-3\\0&0&1&2\end{bmatrix}^2\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=0$$

$$\text{或}\begin{bmatrix}4&0&4&0\\-4&4&-4&4\\1&-4&1&-4\\0&1&0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=0,$$

解得 $x_1=-x_3,x_2=-x_4$。

可选 $\alpha_2=(-1\quad 0\quad 1\quad 0)^T$。

解 $\lambda=1$ 的特征向量。

$$\text{由}\begin{bmatrix}-1&0&0&4\\1&-1&0&-4\\0&1&-1&-3\\0&0&1&3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=0\text{ 可得 }\alpha_3=(4\ \ 0\ \ -3\ \ 1)^T。$$

求 $\lambda=-1$ 时的特征向量。

$$\text{由}\begin{bmatrix}1&0&0&4\\1&1&0&-4\\0&1&1&-3\\0&0&1&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=0\text{ 可得 }\alpha_4=(-4\ \ 8\ \ -5\ \ 1)^T。$$

所以可求得P阵为$\begin{bmatrix} -1 & 2 & 4 & -4 \\ 0 & -1 & 0 & 8 \\ 1 & -2 & -3 & -5 \\ 0 & 1 & 1 & 1 \end{bmatrix}$。

上面介绍的方法在阶不高时，是容易求得该矩阵的 Jordan 标准型，但最关键的是求一个矩阵的特征多项式与确定矩阵的秩。上面二个问题在实际计算中是不容易的。

第五节 Jordan 标准型的定理

定理 3 如果一个 n 阶矩阵 A 有 s 个独立的特征向量，则它相似于如下的 Jordan 型矩阵：

$$J = P^{-1}AP = J_1 \oplus J_2 \oplus \cdots \oplus J_s。$$

其中 $J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{bmatrix}$是 n_i 阶，$\sum_{i=1}^{s} n_i = n$。

解释：由相似可得 $PJ = AP$。不妨观察 P 的首 n_1 个列向量

$$A[\alpha_1, \cdots, \alpha_{n_1}, *, \cdots, *] = [\alpha_1, \cdots, \alpha_{n_1}, *, \cdots, *]\left[\begin{array}{cccc|c} \lambda_1 & 1 & & & \\ & \lambda_1 & 1 & & \\ & & \ddots & 1 & \\ & & & \lambda_1 & \\ \hline & & & & \end{array}\right]$$

由比较可得下面的方程

$$A\alpha_1 = \lambda_1\alpha_1, A\alpha_2 = \lambda_1\alpha_2 + \alpha_1, \cdots, A\alpha_{n_1} = \lambda_1\alpha_{n_1} + \alpha_{n_1-1}。$$

其中 $\alpha_1, \alpha_2, \cdots \alpha_{n_1}$称为对应于特征值 λ_1长度为 n_1 的 Jordan 链。特征向量 α_1 称为 Jordan 链的首元，α_{n_1} 称为末元。每一个 Jordan 块便对应一条 Jordan 链。因此要证明上面的定理，就是要证明对

任一个 n 阶矩阵 A，一定存在 A 的 Jordan 链构成的基（亦称为 Jordan 基）。下面对空间的维数施行数学归纳法证明。

证 (i) 构造 Jordan 基的思想。

因 $\dim R(A)+\dim N(A)=n$，

又 $$\dim(R(A)+N(A))=\dim R(A)+\dim N(A)-\dim(R(A)\cap N(A))。$$

所以 $$\dim V=\dim(R(A)+N(A))+\dim(R(A)\cap N(A))。$$

(a) 设 $\dim R(A)=r$（它是矩阵 A 的秩）。先构造 $R(A)$ 的 Jordan 基。在 $r<n$ 时，由归纳假设保证它存在。

(b) 找$R(A)\cap N(A)$的Jordan基，并将其扩充为$N(A)$的基（所扩充的向量必为对应于零特征值的特征向量）。设$\dim(R(A)\cap N(A))=p$，(a)、(b)所得的向量有 $r+(n-r-p)=n-p$ 个。要构成 Jordan 基还差 p 个向量。采用下面的方法找出这 p 个向量。

(c) 因为对任何一个 $\alpha_i\in R(A)\cap N(A)$，它都是对应于零特征值的特征向量，因此必有一以它为首元的 Jordan 链，它的末元记为 w_i，但因 $w_i\ (i=1, 2, \cdots, p)\in R(A)$，所以必有一 β_i 满足 $A\beta_i=w_i$。这样 $\alpha_i\cdots w_i,\beta_i$，也是 Jordan 链，$\beta_1\cdots\beta_p$ 便是要找的 p 个向量。

(ii) 定理证明。

情形(a) $r<n$。

在 $R(A)$中存在 r 个向量 $w_i(i=1,2,\cdots, r)$，它们满足条件 $Aw_i=\lambda_i w_i$ 或 $Aw_i=\lambda_i w_i+w_{i-1}$。对于任一个首元在 $N(A)\cap R(A)$ 的 Jordan 链的末元 w_i，有 $A\beta_i=w_i$（这种向量有 p 个）。同时将属于 $N(A)\cap R(A)$的 w_i 扩充为 $N(A)$基，扩充的元记为 α_i，它共有 $n-r-p$ 个向量。向量组 $\{w_i\},\{\beta_j\},\{\alpha_k\}$。合起来共有 $r+p+(n-r-p)=n$ 个向量。要证明它们构成 Jordan 基，只要证明 $\{w_1\cdots w_r,\beta_1\cdots\beta_p,\alpha_1\cdots\alpha_{n-r-p}\}$是独立向量组。设它们有如下线性组合关系：

$$\sum_{i=1}^{r} a_i w_i + \sum_{j=1}^{p} b_j \beta_j + \sum_{k=1}^{n-r-p} c_k \alpha_k = 0。\qquad (*)$$

两边用矩阵 A 乘之。注意到 α_i 是对应零特征值的特征向量因此有

$$\sum_{i=1}^{r} a_i A_i w_i + \sum_{j=1}^{p} b_j A \beta_j = 0$$

$$\sum_{i=1}^{r} a_i \begin{Bmatrix} \lambda_i w_i \\ \text{或} \\ \lambda_i w_i + w_{i-1} \end{Bmatrix} + \sum_{j=1}^{p} b_j w_j = 0。\qquad (**)$$

w_j 是对应于 $\lambda_i = 0$ 的 Jordan 链的末元，因而它不会出现在 $\sum_{i=1}^{r}$ 中。于是式（**）是 $\{w_i\}$ 的一些元的线性组合，因此 $b_j = 0$。由式（*）便得 $\sum_{i=1}^{r} a_i w_i + \sum c_k \alpha_k = 0$，但 $w_i \in R(A)$，$\alpha_k \in N(A) \setminus R(A)$，因此 $a_i, c_k = 0$。这证明 $\{w_i\}$，$\{\beta_j\}$，$\{\alpha_k\}$ 构成基，由它们构成的方式可知是 Jordan 基。

情形(b) $r = n$。

任取 A 的一个特征值 λ_0，矩阵 $A - \lambda_0 I$ 是奇异矩阵，$r(A - \lambda_0 I) < n$。由情形（a）可知它有 Jordan 基存在。在此基下，有 $P^{-1}(A - \lambda_0 I)P = J$。因此有 $P^{-1}AP = J + \lambda_0 I$，这便是 A 的 Jordan 标准型。所以原 $A - \lambda_0 I$ 的 Jordan 基也是矩阵 A 的 Jordan 基。 ▨

下面的例是用来说明：定理证明中构造 Jordan 基的方法。

例 23 A 是 11 阶准对角阵，设

$$A = A_1 \oplus \cdots \oplus A_5 = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 5 & 3 \\ -2 & -7 & 4 \end{bmatrix} \oplus \begin{bmatrix} 6 & -9 \\ 4 & -6 \end{bmatrix}$$

$$\oplus \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix} \oplus [3] \oplus O_3。$$

其中 O_3 是 3 阶零阵，试求矩阵 A 的 Jordan 基，并写出 A 的

Jordan 标准型。

解 (i) 求 $R(A)$的 Jordan 基。

$R(A)$的维数是矩阵 A 的秩，A 的秩是各子阵秩的和，各子阵的秩分别为：$r(A_1)=2, r(A_2)=1, r(A_3)=2, r(A_4)=1, r(A_5)=0$，所以 $r(A)=6$。先计算 A 的特征向量。

A_1 的特征多项式 $|\lambda I-A_1|=\lambda^3$，特征值为零，特征向量为 $y^{(1)}=(1\ -2\ -3)^T$。求解时，可用下面的格式

$$\begin{cases} x_1+2x_2-x_3=0 \\ -x_1+5x_2+3x_3=1 \\ -2x_1-7x_2+4x_3=0 \end{cases},$$

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & y_1 \\ -1 & 5 & 3 & y_2 \\ -2 & -7 & 4 & y_3 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & y \\ 0 & -3 & 2 & y_1+y_2 \\ 0 & -3 & 2 & 2y_1+y_3 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & y_1 \\ 0 & -3 & 2 & y_1+y_2 \\ 0 & 0 & 0 & y_1-y_2+y_3 \end{array}\right]。$$

A_2 的特征多项式 $|\lambda I-A_2|=\lambda^2$，特征值为零。特征向量 $y^{(2)}=(3,2)^T$。有

$$\begin{cases} 6x_4-9x_5=0 \\ 4x_4-6x_5=0, \end{cases}$$

$$\left[\begin{array}{cc|c} 6 & -9 & y_4 \\ 4 & -6 & y_5 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 12 & -18 & 2y_4 \\ 12 & -18 & 3y_5 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{cc|c} 12 & -18 & 2y_4 \\ 0 & 0 & 3y_5-2y_4 \end{array}\right]$$

A_3 的特征多项式 $|\lambda I-A_3|=(\lambda-2)^2$，特征值为 2。特征向量为 $v^{(1)}=(1,1)^T$。

A_4 的特征值为 3，它的特征向量为 $v^{(3)}=(1)$。现在还要判断一下所求出的特征向量是否属于 $R(A)$。首先非零特征值对应的特征向量一定属于 $R(A)$，余下只要判断 $\lambda=0$ 时的特征向量是否

属于 $R(A)$，只要根据求解时所得的相容性条件便可判别。如 $y^{(1)}=(1,-2,-3)^T$，它满足 $y_1-y_2+y_3=0$，说明方程 $A_1x=y^{(1)}$ 有解，它的解为 $w^{(1)}=(2,-3,-5)^T$，$A_1w^{(1)}=y^{(1)}$。因此 $y^{(1)}\in R(A)$，$y^{(1)}\in R(A)\cap N(A)$。而 $w^{(1)}$ 同样满足 $y_1-y_2+y_3=0$，这样 $w^{(1)}\in R(A)$。同理可得 $y^{(2)}$ 满足条件 $3y_5-2y_4=0$，所以 $y^{(2)}\in R(A)\cap N(A)$。对于 $v^{(1)}$ 可知方程 $(A_3-2I)x=v^{(1)}$ 有解，可得 $v^{(2)}=x=(4,3)^T$，由此可知 $v^{(2)}\in R(A)$。总结上面可知 $R(A)$ 的 Jordan 基为：

$$y^{(1)},w^{(1)},(A_1w^{(1)}=y^{(1)})y^{(2)},v^{(1)},v^{(2)}((A_3-2I)v^{(2)}=v^{(1)})v^{(3)}。$$

(ii) 扩充上面的集合为整个空间的 Jordan 基。

按证明要找出 $R(A)\cap N(A)$ 的元(现在是 $\{y^{(1)}\ y^{(2)}\}$) 为首元的 Jordan 链的末元，一个是 w_1，另一个 $y^{(2)}$，求出满足条件 $A_1x^{(1)}=w^{(1)}$，$A_2x^{(2)}=y^{(2)}$ 的解 $x^{(1)}=(0\ \ 4\ \ 3)^T$，$x^{(2)}=(8\ \ 5)^T$。除此之外，还要扩充 $y^{(1)},y^{(2)}$ 为 $N(A)$的基。由于 $N(A)$的维数为 5，所以要添加 3 个向量，注意到 $A_5=O_3$，所以只要添加 e_9,e_{10},e_{11}。

综上所述 A 的 Jordan 基为

$$
\begin{array}{ll}
y^{(1)},w^{(1)},x^{(1)} & (x^{(1)},Ax^{(1)},A^2x^{(1)}),\\
y^{(2)},x^{(2)}, & (x^{(2)}\quad Ax^{(2)}),\\
y^{(1)},v^{(2)} & (v^{(2)}\quad Av^{(2)}),\\
v^{(3)} & \\
e_9,e_{10},e_{11}。 &
\end{array}
$$

具体写出这些向量时，要记住上面的计算是利用了分块对角阵这一特点，已将向量的维“缩少”，因而正式写出来的时候还要将零分量的部分补上。

有

$$
\begin{aligned}
y^{(1)}&=(1\ \ -2\ \ -3\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0)^T,\\
w^{(1)}&=(2\ \ -3\ \ -5\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0)^T,\\
x^{(1)}&=(0\ \ 4\ \ 3\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0)^T,\\
y^{(2)}&=(0\ \ 0\ \ 0\ \ 3\ \ 2\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0)^T,
\end{aligned}
$$

$$x^{(2)}=(0\quad 0\quad 0\quad 8\quad 5\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0)^T,$$
$$v^{(1)}=(0\quad 0\quad 0\quad 0\quad 0\quad 1\quad 1\quad 0\quad 0\quad 0\quad 0)^T,$$
$$v^{(2)}=(0\quad 0\quad 0\quad 0\quad 0\quad 4\quad 3\quad 0\quad 0\quad 0\quad 0)^T,$$
$$v^{(3)}=(0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 1\quad 0\quad 0\quad 0)^T,$$
$$e_9=(0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 1\quad 0\quad 0)^T,$$
$$e_{10}=(0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 1\quad 0)^T,$$
$$e_{11}=(0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 1)^T。$$

则

$$P=[y^{(1)}\ w^{(1)}\ x^{(1)}\ y^{(2)}\ x^{(2)}\ v^{(1)}\ v^{(2)}v^{(3)}\ e_9\ e_{10}'\ e_{11}]。$$

所以

$$J=J_1\oplus J_2\oplus J_3\oplus\cdots\oplus J_7$$
$$=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}\oplus\begin{bmatrix}0&1\\0&0\end{bmatrix}\oplus\begin{bmatrix}2&1\\0&2\end{bmatrix}\oplus[3]\oplus[0]\oplus[0]\oplus[0]。$$

第五章　矩阵分析初步

本章主要讲述矩阵序列以及矩阵幂级数的有关概念，并由此定义矩阵函数，讨论计算矩阵函数的方法。

第一节　矩 阵 序 列

1. 矩阵序列的定义

定义　假定 $A_1, A_2, \cdots, A_k, \cdots$ 是 $m\times n$ 型的矩阵无穷序列。其中 $A_k=[a_{ij}^{(k)}]$。如果 $\lim\limits_{k\to\infty} a_{ij}^{(k)}=a_{ij}$，对 $i=1,2,\cdots,m$；$j=1,2,\cdots,n$ 均成立，则称矩阵序列 $\{A_k\}$ 收敛。而 $A=[a_{ij}]$ 称为矩阵序列 $\{A_k\}$ 的极限，记号为 $\lim A_k=A$。

例 1　求

$$A_k=\begin{bmatrix}\left(1+\dfrac{1}{k}\right)^k & \dfrac{1+k}{k}\\ 2\,k/k & 1\\ \dfrac{\sin k}{k} & \sqrt[k]{k}\end{bmatrix},$$

当 $k\to\infty$ 时，A_k 的极限。

解　$\{A_k\}$ 是一个 3×2 型的矩阵无穷序列。要求极限，只需将它的每个元取极限。因而它的极限 $A=\begin{bmatrix} e & 1\\ 0 & 1\\ 0 & 1\end{bmatrix}$。

由矩阵序列极限的定义可知，它和数列极限没有本质的差别。它不过是将若干个数列放在一起来考虑它们的极限而已。

2. 矩阵序列的运算法则

讨论这些法则的目的是使矩阵序列的收敛性的讨论变得简单些。

(i) $\lim A_k = A, \lim B_k = B$。而 A_k 是 $m\times n$ 型, B_k 是 $n\times p$ 型。则 $\lim A_kB_k = AB$。

设 $A=[a_{ij}]_{m\times n}, B=[b_{ij}]_{n\times p}$,则有

$$(A_kB_k)_{ij}=a_{i1}^{(k)}b_{1j}^{(k)}+a_{i2}^{(k)}b_{2j}^{(k)}+\cdots+a_{in}^{(k)}b_{nj}^{(k)}。$$

只要逐项取极限便可得,

$$\lim(A_kB_k)_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}=(AB)_{ij}。$$

结论成立。

上面结果有二个有用的推论:

(ii) 如果 $\lim A_k = A$, 则 $\lim A_kB = AB$。

(iii) 如果 $\lim A_k = A$, 则 $\lim PA_kP^{-1} = P(\lim A_k)\ P^{-1} = PAP^{-1}$。

第二节　矩阵的幂级数

1. 矩阵的幂收敛

这里考虑一种特殊的矩阵序列。设 A 是 n 阶方阵, $A, A^2, \cdots, A^k, \cdots$ 是一个矩阵无穷序列。如果 $\{A^k\}$ 收敛, 则称矩阵 A 是幂收敛的。

在考察矩阵 A 的幂收敛时, 采用下面的方法,以简化问题的讨论。首先对 A^k 作相似变换 $P^{-1}A^kP=(P^{-1}AP)^k$, 如果 $P^{-1}AP$ 是 A 的 Jordan 标准型 J, 这样可得 $P^{-1}A^kP=J^k$。因此A的幂收敛与它的 Jordan 标准型的幂收敛是等价的。另一方面因 J 是分块对角阵, $J=J_1\oplus J_2\oplus\cdots\oplus J_s \Longrightarrow J^k=J_1^k\oplus J_2^k\oplus\cdots\oplus J_s^k$。所以 J 的幂收敛便转化为 Jordan 块的幂收敛。下面讨论 Jordan块的幂收敛。现有

$$J_{(n)}(\lambda)=\begin{bmatrix}\lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda\end{bmatrix}=\lambda I+N。$$

式中 I 是 n 阶单位阵，N 是 n 阶幂零阵，由于 I,N 是可交换的。因此下面的运算成立。

$$J_{(n)}^{k}(\lambda)=(\lambda I+N)^{k}=\lambda^{k}I+C_{k}^{1}\lambda^{k-1}N+C_{k}^{2}\lambda^{k-2}N^{2}+\cdots$$
$$+C_{k}^{k-1}\lambda N^{k-1}+N^{k}。$$

因为是研究 $\lim J_{(n)}^{k}(\lambda)$，所以上面的表达式只要研究 k 充分大的情形。N 是幂零阵，当 $i\geqslant n$ 时，$N^{i}=0$，因此对于充分大的 k 有

$$J_{(n)}^{k}(\lambda)=\lambda^{k}I+C_{k}^{1}\lambda^{k-1}N+C_{k}^{2}\lambda^{k-2}N^{2}+\cdots+C_{k}^{n-1}\lambda^{k-n+1}N^{n-1}.$$

写成矩阵的形式便是

$$J_{(n)}^{k}(\lambda)=\begin{pmatrix}\lambda^{k} & C_{k}^{1}\lambda^{k-1} & C_{k}^{2}\lambda^{k-2} & \cdots & C_{k}^{n}\lambda^{k-n+1} \\ & \lambda^{k} & C_{k}^{1}\lambda^{k-1} & \cdots & C_{k}^{n-2}\lambda^{k-n+2} \\ & & \lambda^{k} & C_{k}^{1}\lambda^{k-1} & \\ & & \ddots & \ddots & \\ & & & \ddots & C_{k}^{1}\lambda^{k-1} \\ & & & & \lambda^{k}\end{pmatrix},\quad (5\text{-}1)$$

由 $J_{(n)}^{k}(\lambda)$ 的表达式容易得出一个 Jordan 块幂收敛的条件。下面探讨一下矩阵(5-1)中每一个元在 $k\to\infty$ 时收敛的条件，

(i) 对角元是 λ^{k}。显然，$\lim\lambda^{k}$ 要存在，它只能在 $|\lambda|<1$ 或 $\lambda=1$ 时。

(ii) 一般的元是 $C_{k}^{j}\lambda^{k-j}$，这里 j 是一固定的数。令 $u_{k}=C_{k}^{j}\lambda^{k-j}$，为了研究 $\lim u_{k}$，考察级数 $\sum u_{k}$ 的收敛性。因为

$$\left|\frac{u_{k+1}}{u_{k}}\right|=\left|\frac{C_{k+1}^{j}\lambda^{k+1-j}}{C_{k}^{j}\lambda^{k-j}}\right|$$

$$=\frac{(k+1)!/j!(k+1-j)!}{k!/j!(k-j)!}|\lambda|\longrightarrow|\lambda|\qquad(k\to\infty)$$

因此若 $|\lambda|<1$，该级数收敛，因而一般项 $u_k\to 0$。

综合上面的讨论，可得出关于矩阵幂收敛的条件。

命题 1 $J_{(n)}(\lambda)$ 是幂收敛 $\Longleftrightarrow$ λ 满足条件 (i) $|\lambda|<1$ 或 (ii) $\lambda=1$ 时 $n=1$。

证 $\Longrightarrow$ 如果 $|\lambda|<1$，则由前面的讨论可知 $J_{(n)}^k$ 确实是幂收敛。在 $\lambda=1$ 时，如果 $n\geqslant 2$，则由 $J_{(n)}^k(\lambda)$ 的表达式可知存在有元 C_k^1。当 $k\to\infty$ 有 $C_k^1\to\infty$，所以 $J_{(n)}^k(\lambda)$ 不可能收敛。由于 $J_{(n)}^k(\lambda)$ 是收敛的，所以 $n=1$。

$\Longleftarrow$ 由前面讨论可知定理充分性部分成立。 ▨

现在给出矩阵 A 是幂收敛的定理。

定理 1 A 是幂收敛 $\Longleftrightarrow$ A 的任一特征值 λ 满足条件：(i) $|\lambda|\leqslant 1$ (ii) 如果 $|\lambda|=1$ 时，应是 $\lambda=1$，且对应的 Jordan 块应是一阶 Jordan 块（或者说 $\lambda=1$ 的代数重数与几何重数相等）。

2．矩阵幂级数

(1) 矩阵级数如果给出一个矩阵序列 $A_1,A_2,\cdots,A_k,\cdots$。据此序列可构造和式 $A_1+A_2+\cdots+A_n=S_n$。如果 $\lim S_n=S$，将它形式地记为 $A_1+\cdots+A_n+\cdots=S$，或称无穷级数 $\sum\limits_{k=1}^{\infty}A_k=S$（收敛于 S）。由此可见一个矩阵无穷级数，是用它的部分和序列的极限来定义的。这点和常数项完全一样。由此可导出与常数项级数相类似的性质。

如 (i) 若 $\sum\limits_{k=1}^{\infty}A_k$ 收敛，则对于一般项 A_k 有 $\lim A_k=0$。

(ii) 若 $\sum\limits_{k=1}^{\infty}A_k=A$，$\sum\limits_{k=1}^{\infty}B_k=B$。则

$$\sum_{k=1}^{\infty}(A_k+B_k)=A+B,\ \sum_{k=1}^{\infty}aA_k=aA。$$

(2) 矩阵幂级数

设 A 是给定的 n 阶阵，考察矩阵级数 $\sum\limits_{k=0}\alpha_k A^k$ 的收敛性。本小节要回答级数 $\sum\limits_{k=0}^{\infty}\alpha_k A^k$ 与幂级数 $\sum\limits_{k=0}^{\infty}\alpha_k t^k$ 的收敛性有何关系。级数 $\sum\limits_{k=0}^{\infty}\alpha_k A^k$ 从形式上可看成在幂级数 $\sum\limits_{k=0}^{\infty}\alpha_k t^k$ 中，用矩阵 A 去替代 t，这种级数称为矩阵幂级数。下面是一个最简单的例子——几何级数。

例 2 $\sum\limits_{n=0}^{\infty}Z^n$ 收敛 $\Longleftrightarrow \lim Z^n=0$。(这里 Z 是复数)。

证 （一般数项级数 $\lim u_n=0$，只是收敛的必要条件，但对几何级数，却是充分必要的）。

$\Longrightarrow$ 事实上由部分和 $S_n=1+Z+\cdots+Z^n$，$Z_n=S_n-S_{n-1}$。如果级数收敛，则 $\lim Z_n=\lim(S_n-S_{n-1})=S-S=0$。

$\Longleftarrow$ 因 $(1-Z)S_n=(1-Z)(1+Z+\cdots+Z^n)=1-Z^{n+1}$，又 $\lim Z^n=0$，两边取极限可知有 $(1-Z)\lim S_n=1$，即 $\lim S_n=(1-Z)^{-1}$。

实际上，上面的推理过程用矩阵 A 代替 Z，单位阵 I 代替 1，便可得到下面的命题。

命题 2 $\sum\limits_{k=0}^{\infty}A^k$ 收敛 $\Longleftrightarrow \lim A^n=0$，这时 $\sum\limits_{k=0}^{\infty}A^k=(I-A)^{-1}$。

由前面讨论可知，要矩阵幂级数 $\sum\limits_{k=0}^{\infty}A^k$ 收敛的充要条件是矩阵 A 幂收敛于零阵。再由前面关于幂收敛的讨论可知，一个矩阵 A 幂收敛于零阵的充要条件为 $|\lambda|<1$。这时，A 的特征值 λ 恰好落在数值幂级数 $\sum\limits_{k=1}^{\infty}Z^k$ 的收敛圆 $|Z|<1$ 内。这个结论对一般情形也是对的

为了证明一般的定理，先证明二个引理。

引理 $\sum\limits_{i=0}^{n}\sum\limits_{j=0}^{i}a_{ij}=\sum\limits_{j=0}^{n}\sum\limits_{i=j}^{n}a_{ij}$

证 事实上和式 $\sum\limits_{i=0}^{n}\sum\limits_{j=0}^{i}a_{ij}$ 的各加项可列在下表。

$$
\begin{array}{lllll}
a_{00} & & & & \\
a_{10} & a_{11} & & & \\
a_{20} & a_{21} & a_{22} & & \\
\vdots & \vdots & \vdots & & \\
a_{n0} & a_{n1} & a_{n2} & \cdots & a_{nn}
\end{array}
$$

如果用坐标形象表示加项便是如图 5-1 所示。

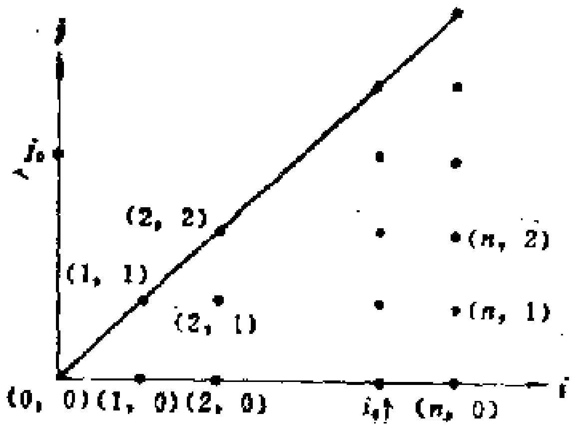

图 5-1

左边的和式相当于是对每一个固定的 i_0 按垂直方向(即 j 的方向)求和，而右边的和式相当于对每一个固定的 j_0 按水平方向求和(即 i 的方向求和)，因此两边是相等的。 ▨

引理 设 $J(\lambda)$ 是 n 阶 Jordan 块，而 $f(t)=\sum a_i t^i$ 是收敛半径为 r 的幂级数。当 $|\lambda|<r$ 时，矩阵幂级数 $\sum\limits_{i=0}^{\infty} a_i J^i$ 是收敛的，且其和为矩阵

$$
\begin{pmatrix}
f(\lambda) & f'(\lambda) & \frac{1}{2!}f''(\lambda) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda) \\
 & f(\lambda) & f'(\lambda) & \cdots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda) \\
 & & \ddots & \ddots & \vdots \\
 & & & & f'(\lambda) \\
 & & & & f(\lambda)
\end{pmatrix}。
$$

证　要证明幂级数收敛，先求部分和 S_m 为：

$$S_m=\sum_{i=0}^{m}\alpha_i J^i=\sum_{i=0}^{m}\alpha_i(\lambda I+N)^i=\sum_{i=0}^{m}\alpha_i\left\{\sum_{j=0}^{i}C_j^i\lambda^{i-j}N^j\right\}$$

$$=\sum_{i=0}^{m}\sum_{j=0}^{i}\alpha_i C_j^i\lambda^{i-j}N^j=\sum_{j=0}^{m}\left\{\sum_{i=j}^{m}\alpha_i C_j^i\lambda^{i-j}\right\}N^j$$

但由于 $j\geqslant n$，有 $N^j=0$，所以标号 $j=0,1,\cdots,n-1$。下面简化系数 $\sum_{i=j}^{m}\alpha_i C_j^i\lambda^{i-j}$，有

$j=0$，$\sum_{i=0}^{m}\alpha_i C_0^i\lambda^i=\sum_{i=0}^{m}\alpha_i\lambda^i=S_m(\lambda)$（幂级数 $f(t)=\sum\alpha_i t^i$ 的部分和）；

$j=1$，$\sum_{i=1}^{m}\alpha_i C_1^i\lambda^{i-1}=\sum_{i=1}^{m}\alpha_i i\lambda^{i-1}=\sum_{i=0}^{m}(\alpha_i\lambda^i)'=S_m'(\lambda)$，

$$j=2,\ \sum_{i=2}^{m}\alpha_i C_2^i\lambda^{i-2}=\sum_{i=2}^{m}\frac{i(i-1)}{2}\alpha_i\lambda^{i-2}$$

$$=\frac{1}{2!}\sum_{i=0}^{m}(\alpha_i\lambda^i)''=\frac{1}{2!}S_m''(\lambda)。$$

$$\text{一般地}\ \alpha_i C_j^i\lambda^{i-j}=\alpha_i\frac{i(i-1)\cdots(i-j+1)}{j!}\lambda^{i-j}$$

$$=\frac{\alpha_i}{j!}(t^i)^{(j)}\Big|_{t=\lambda}。$$

所以 $\sum_{i=j}^{m}\alpha_i C_j^i\lambda^{i-j}=\sum_{i=j}^{m}\frac{\alpha_i}{j!}(t^i)^{(j)}\Big|_{t=\lambda}=\frac{1}{j!}\left\{\sum_{i=0}^{m}\alpha_i t^i\right\}^{(j)}\Big|_{t=\lambda}=$

$\frac{S_m^{(j)}(\lambda)}{j!}$。

因此可归纳得出结果为，

$S_m(J)=\sum_{j=0}^{n-1}\frac{1}{j!}S_m^{(j)}(\lambda)N^j$，写成矩阵的形式为

$$\begin{bmatrix} S_m(\lambda) & S'_m(\lambda) & \frac{1}{2}S''_m(\lambda) & \cdots & \frac{1}{(n-1)!}S_m^{(n-1)}(\lambda) \\ & S_m(\lambda) & S'_m(\lambda) & \cdots & \frac{1}{(n-2)!}S_m^{(n-2)}(\lambda) \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & S'_m(\lambda) \\ & & & & S_m(\lambda) \end{bmatrix}。$$

又 $|\lambda|<r$，所以当 $m\to\infty$，上面矩阵的每一个元都收敛。又由幂级数收敛的性质可知 $\sum\limits_{i=0}^{\infty}\alpha_i t^i=f(t)$，有 $\lim S_m^{(k)}=f^{(k)}(t)$。因此上面部分和矩阵的极限为

$$\begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2}f''(\lambda) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda) \\ & f(\lambda) & f'(\lambda) & \cdots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda) \\ & & & \ddots & \\ & & \ddots & & f'(\lambda) \\ & & & & f(\lambda) \end{bmatrix}。\quad (5\text{-}2)$$

例 3 已知 $f(t)=\sum\limits_{n=0}^{\infty}\left(\frac{t}{6}\right)^n$（由几何级数和，可知 $f(t)=\left(1-\frac{t}{6}\right)^{-1}$，收敛域为 $|t|<6$）。另有

$$J=\begin{bmatrix} 3 & 1 & & \\ & 3 & 1 & \\ & & 3 & 1 \\ & & & 3 \end{bmatrix}，求 f(J)$$

解 由上面引理可知，$\lambda=3$ 落在幂级数的收敛域 $|t|<6$ 内。因此 $f(J)=\sum\limits_{n=0}^{\infty}\left(\frac{J}{6}\right)^n$ 是收敛的。要求出和先求出 $f(t)$ 的 1 至 3 阶的导数

$$f'(t)=\frac{1}{6}\left(1-\frac{t}{6}\right)^{-2},f''(t)=\frac{1}{18}\left(1-\frac{t}{6}\right)^{-3},$$

$$f'''(t)=\frac{1}{36}\left(1-\frac{t}{6}\right)^{-4}。$$

用 $t=3$ 代入便可求得

$$f(3)=2,\ f'(3)=\frac{2}{3},\ f''(3)=\frac{4}{9},$$

$$f'''(3)=\frac{4}{9}。$$

因此，$f(J)=\sum\limits_{n=0}^{\infty}\left(\frac{J}{6}\right)^n=\begin{pmatrix} 2 & \frac{2}{3} & \frac{2}{9} & \frac{2}{27} \\ & 2 & \frac{2}{3} & \frac{2}{9} \\ & & 2 & \frac{2}{3} \\ & & & 2 \end{pmatrix}$。

上面结果可直接与 $f(T)=\left(I-\frac{T}{6}\right)^{-1}$ 计算比较。

有了上面二个引理，便可叙述矩阵幂级数的基本定理。

定理 2 (Lagrange-Sylvester)

(i) 设 $f(t)=\sum\alpha_i t^i$，它的收敛半径为 r。

(ii) 矩阵 A 的诸特征值 λ 满足条件 $|\lambda|<r$，则矩阵幂级数 $\sum\limits_{i=0}^{\infty}\alpha_i A^i$ 收敛。

(iii) 设 $A=PJP^{-1}$，J 是 A 的 Jordan 标准型

$$J=J_1(\lambda_1)\oplus J_2(\lambda_2)\oplus\cdots\oplus J_s(\lambda_s)。$$

其中 $J_i(\lambda_i)$ 是 n_i 阶 Jordan 块，$\sum\limits_{i=1}^{s}n_i=n$。则

$$\begin{aligned}\sum\alpha_k A^k&=\sum\alpha_k(PJP^{-1})^k=P(\sum\alpha_k J^k)P^{-1}\\&=P\{(\sum\alpha_k J_1^k(\lambda_1)\oplus(\sum\alpha_k J_2^k(\lambda_2))\oplus\end{aligned}$$

$$\cdots\oplus\left(\sum \alpha_k J_s^k(\lambda_s)\right)\Big\}P^{-1},$$

$$=P\left\{\begin{bmatrix} f(\lambda_1) & f'(\lambda_1) & \cdots & \dfrac{f^{n_1-1}(\lambda_1)}{(n_1-1)!} \\ & f(\lambda_1) & f'(\lambda_1)\cdots & \dfrac{f^{n_1-2}(\lambda_1)}{(n_1-2)!} \\ & \ddots & \ddots & \vdots \\ & & \ddots & f'(\lambda_1) \\ & & & f(\lambda_1) \end{bmatrix}\right.$$

$$\oplus\begin{bmatrix} f(\lambda_2) & f'(\lambda_2) & \cdots & \dfrac{f^{n_2-1}(\lambda_2)}{(n_2-1)!} \\ & \ddots & \ddots & \dfrac{f^{n_2-2}(\lambda_2)}{(n_2-2)!} \\ & & \ddots & \vdots \\ & & & f'(\lambda_2) \\ & & & f(\lambda_2) \end{bmatrix}\oplus\cdots\oplus$$

$$\left.\oplus\begin{bmatrix} f(\lambda_s) & f'(\lambda_s) & \cdots & \dfrac{f^{n_s-1}(\lambda_s)}{(n_s-1)!} \\ & \ddots & \ddots & \dfrac{f^{n_s-2}(\lambda_s)}{(n_s-2)!} \\ & & \ddots & \cdots \\ & & & f'(\lambda_s) \\ & & & f(\lambda_s) \end{bmatrix}\right\}P^{-1}。\qquad (5\text{-}3)$$

例 4 已知$A=\begin{bmatrix} 3 & 1 & & \\ & 3 & 1 & \\ & & 3 & 1 \\ & & & 3 \end{bmatrix}\oplus\begin{bmatrix} 1 & 1 & \\ & 1 & 1 \\ & & 1 \end{bmatrix}$，又$f(t)=2+t+3t^2$，求$f(A)$。

解 将$f(t)$视为幂级数：$\sum\limits_{k=0}^{\infty}\alpha_k t^k$，$\alpha_0=2$，$\alpha_1=1$，$\alpha_2=3$，$\alpha_j=0(j\geqslant3)$。求出$f(t)$在$t=3$的1至3阶导数。$t=1$的1阶至2阶导数：

$f(3)=32$，$f'(3)=19$，$f''(3)=6$，$f'''(3)=0$。

$f(1)=6$，$f'(1)=7$，$f''(1)=6$。

$$这样\ f(A)=\begin{bmatrix} 32 & 19 & 6 & 0 \\ & 32 & 19 & 6 \\ & & 32 & 19 \\ & & & 32 \end{bmatrix} \oplus \begin{bmatrix} 6 & 7 & 6 \\ & 6 & 7 \\ & & 6 \end{bmatrix}。$$

上面结果可直接与 $2I+A+3A^2$ 计算比较。

第三节　矩阵函数

1．矩阵函数的定义

有了前面的 Lagrange-Sylvester 定理，可由幂级数定义矩阵函数。下面举出重要的矩阵函数。在复变函数论里已知有下面结果：

$$e^z=1+Z+\frac{Z^2}{2!}+\cdots+\frac{Z^n}{n!}+\cdots,$$

$$\sin A=Z-\frac{1}{3!}Z^3+\frac{1}{5!}Z^5+\cdots+\frac{(-1)^{n-1}}{(2n-1)!}Z^{2n-1}+\cdots,$$

$$\cos A=1-\frac{1}{2!}Z^2+\frac{1}{4!}Z^4+\cdots+\frac{(-1)^n}{(2n)!}Z^{2n}+\cdots。$$

它们在整个复平面都是收敛的。因而根据 Lagrange -Sylvester 定理可知对任何方阵 A，级数

$$I+A+\frac{1}{2!}A^2+\cdots+\frac{1}{n!}A^n+\cdots,$$

$$A-\frac{1}{3!}A^3+\frac{1}{5!}A^5-\cdots+\frac{(-1)^{n-1}}{(2n-1)!}A^{2n-1}+\cdots,$$

$$1-\frac{1}{2!}A^2+\frac{1}{4!}A^4-\cdots+\frac{(-1)^n}{(2n)!}A^{2n}+\cdots,$$

都是收敛的。它们的和分别用 e^A，$\sin A$，$\cos A$ 表示。

例 5 设有一 4 阶矩阵，它的特征值为 $\pi,-\pi,0,0$。试求 e^A，$\sin A,\cos A$。

解 已知条件给出了矩阵 A 的特征多项式 $\phi(\lambda)=\lambda^4-\pi^2\lambda^2$。由 Hamilton-Cayley 定理可知 $\phi(A)=0$，即 $A^4=\pi^2A^2$。有了这个等式，上面三个矩阵函数的矩阵幂级数，实际上可用 A 的有限多项式给出。即

$$A^6=\pi^2A^4=\pi^4A^2,A^8=\pi^2A^6=\pi^6A^2,\cdots,A^{2n}=\pi^{2(n-1)}A^2。$$

类似可得 $A^{2n-1}=\pi^{2(n-2)}A^3$。

因此
$$e^A=I+A+\frac{A^2}{2!}+\cdots+\frac{A^n}{n!}+\cdots$$
$$=I+A+\frac{A^2}{2!}+\frac{A^4}{4!}+\cdots+\frac{A^{2n}}{(2n)!}+\cdots$$
$$+\frac{A^3}{3}+\frac{A^5}{5!}+\cdots+\frac{A^{2n+1}}{(2n+1)!}+\cdots$$
$$=I+A+\left(\frac{1}{2!}+\frac{1}{4!}\pi^2+\frac{1}{6}\pi^4+\cdots\right)A^2$$
$$+\left(\frac{1}{3!}+\frac{1}{5!}\pi^2+\frac{1}{7!}\pi^4+\cdots\right)A^3$$
$$=I+A+\frac{p}{\pi^2}A^2+\frac{e^\pi-\pi-1-p}{\pi^3}A^3$$
$$\left(p=\frac{e^\pi+e^{-\pi}}{2}-1\right)。$$

用同样的方法计算得
$$\sin A=A+\frac{\sin\pi-\pi}{\pi^3}A^3=A-\pi^{-2}A^3,$$
$$\cos A=I+\frac{\cos\pi-1}{\pi^2}A^2=I-2\pi^{-2}A^2。$$

例 6 若 A 与对角阵 $D=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}$ 相似，求 $e^A,\sin A,\cos A$。

解 设 $A=PDP^{-1}$,则

$$e^A=\sum_{k=0}^{\infty}\frac{A^k}{k!}=\sum_{k=0}^{\infty}\frac{1}{k!}(PDP^{-1})^k=P\left(\sum_{k=0}^{\infty}\frac{D^k}{k!}\right)P^{-1}$$

$$=P\begin{bmatrix}e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n}\end{bmatrix}P^{-1}。$$

完全类似可得

$$\sin A=P\begin{bmatrix}\sin\lambda_1 & & & \\ & \sin\lambda_2 & & \\ & & \ddots & \\ & & & \sin\lambda_n\end{bmatrix}P^{-1},$$

$$\cos A=P\begin{bmatrix}\cos\lambda_1 & & & \\ & \cos\lambda_2 & & \\ & & \ddots & \\ & & & \cos\lambda_n\end{bmatrix}P^{-1}。$$

由结果可知,

e^A 的特征值为 $e^{\lambda_1},e^{\lambda_2},\cdots,e^{\lambda_n}$,

$\sin A$ 的特征值为 $\sin\lambda_1,\sin\lambda_2,\cdots\sin\lambda_n$,

$\cos A$ 的特征值为 $\cos\lambda_1,\cos\lambda_2,\cdots\cos\lambda_n$。

这个结果一般也是对的。它可由 Lagrange-Sylvestev 定理直接得到。

2 矩阵指数函数

这小节里将讲述矩阵指数函数 e^A 的基本性质。

定理 3 (i) $e^A\cdot e^B=e^{A+B}$ $(AB=BA)$,

(ii) $(e^A)^{-1}=e^{-A}$。

(iii) $\det e^A=e^{\operatorname{tr}A}$。

证 (i) 的证明,主要是应用在条件 $AB=BA$ 之下,二项式定理

$$(A+B)^n=\sum_{k=0}^{n} C_n^k A^{n-k} B^k$$

成立。(这式可用数学归纳法证明)
所以

$$e^A \cdot e^B=\left(\sum_{k=0}^{\infty} \frac{1}{k!} A^k\right)\left(\sum_{i=0}^{\infty} \frac{1}{i!} B^i\right)$$

$$=\sum_{n=0}^{\infty}\left\{\sum_{i=0}^{n} \frac{1}{(n-i)!} A^{n-i} \frac{1}{i!} B^i\right\}$$

$$=\sum_{n=0}^{\infty} \frac{1}{n!}\left\{\sum_{i=0}^{n} \frac{n!}{(n-i)!i!} A^{n-i} B^i\right\}$$

$$=\sum_{n=0}^{\infty} \frac{1}{n!}(A+B)^n=e^{A+B}。$$

(ii) 在 (i) 中,令 $B=-A$,可得 $e^A \cdot e^{-A}=e^0=I$,所以

$$(e^A)^{-1}=e^{-A}$$

(iii) 因为 A 的特征值为 $\lambda_1 \cdots \lambda_n$ 时,e^A 的特征值为 e^{λ_1}, e^{λ_2}, $\cdots, e^{\lambda_n}$,因此 $\det e^A=e^{\lambda_1} \cdot e^{\lambda_2} \cdots e^{\lambda_n}=e^{\sum \lambda_i}=e^{\operatorname{tr} A}$。

3. 矩阵函数的导数与积分

要进一步讨论矩阵指数函数的性质,还应引入矩阵函数的导数与积分。下面讨论的矩阵函数是单变量 t 的矩阵函数 $A(t)=[a_{ij}(t)]_{m\times n}$,其中 $a_{ij}(t)$ 是 t 的实函数。按通常的方式来定义连续、导数、积分。

(i) 若对每一个 $a_{ij}(t)$ 在 $t=t_0$ 处连续,则称 $A(t)$ 在 $t=t_0$ 处连续。

(ii) $$A'(t)=\lim_{\Delta t\to 0} \frac{A(t+\Delta t)-A(t)}{\Delta t}$$

$$=\left[\lim_{\Delta t\to 0} \frac{a_{ij}(t+\Delta t)-a_{ij}(t)}{\Delta t}\right]=[a_{ij}'(t)]$$

即一个单变量 t 的矩阵函数的导数,相当于对矩阵 $A(t)$ 中每一个元求导后所得的矩阵。

(iii) $\int_a^b A(s)\mathrm{d}s = \left[\int_a^b a_{ij}(s)\mathrm{d}s\right]$。

由上面的定义立即可知微积分学很多的运算法则，都可直接应用于矩阵函数。如 $A(t),B(t)$ 是 t 的可导矩阵函数，a,b 是常数，则可有下面的运算法则成立：

$$(aA(t)+bB(t))' = aA'(t)+bB'(t);$$

$$(A(t)B(t))' = A'(t)B(t)+A(t)B'(t);$$

$$\left(\int_a^t A(s)\,\mathrm{d}s\right)' = A(t);$$

$$\int_a^t A'(s)\mathrm{d}s = A(t)-A(a)。$$

$$\int_a^t BA(s)\mathrm{d}s = B\int_a^t A(s)\mathrm{d}s, B \text{ 是常数矩阵。}$$

例 7 设 $A(t)=\begin{bmatrix} \sin t & \cos t & t \\ \dfrac{\sin t}{t} & e^t & t^2 \\ 1 & 0 & t^3 \end{bmatrix}$ $(t\neq 0)$，求 $A'(t)$。

解 $A'(t)=\begin{bmatrix} \cos t & -\sin t & 1 \\ \dfrac{t\cos t-\sin t}{t^2} & e^t & 2t \\ 0 & 0 & 3t^2 \end{bmatrix}$。

例 8 $\int_0^1 \begin{bmatrix} s & \sin \pi s \\ 1 & -s \end{bmatrix} \mathrm{d}s = \begin{bmatrix} \dfrac{1}{2} & \dfrac{2}{\pi} \\ 1 & -\dfrac{1}{2} \end{bmatrix}$。

第四节 e^{At} 的性质与应用

这节专论带参数 t 的矩阵指数函数 $E(t)=e^{At}$ 的性质与应用，其中 t 是实参变量。

1. e^{At} 的性质

(i) $e^{At}\cdot e^{As}=e^{A(t+s)}$。

(ii) $(e^{At})^{-1}=e^{-At}$。

以上二个性质易由前面 e^{A} 的性质直接推得。

(iii) 设 A 是任一实方阵，则 $E(t)$ 满足矩阵微分方程：

$$\dot{X}(t)=AX(t),\quad X(0)=I。$$

证 有 $E(t)=\sum\limits_{k=0}^{\infty}\dfrac{(tA)^k}{k!}$，对该级数逐项求导可得

$$E'(t)=\sum_{k=1}^{\infty}\frac{t^{k-1}A^k}{(k-1)!}=A\sum_{k=0}^{\infty}\frac{(tA)^k}{k!}=AE(t)。$$

同时，还可知 $E(t)$ 还满足另一矩阵微分方程 $E'(t)=E(t)\cdot A$。$E(t)$ 显然满足初始条件 $E(0)=I$。

2. e^{At} 的应用

(1) 本节主要讲矩阵函数 e^{At} 在解常系数微分方程组的应用。

命题 3 线性微分方程 $\dot{x}=Ax(x(0)=x^0)$ 的解为 $x(t)=e^{At}x^0$，且解是唯一的。

(这里 $x(t)=(x_1(t),x_2(t),\cdots,x_n(t))^T,\dot{x}(t)=(\dot{x}_1(t),\dot{x}_2(t),\cdots,\dot{x}_n(t))^T$，$A$ 是 n 阶实方阵，$x^0=(c_1,c_2,\cdots,c_n)$ 是 n 维向量。)

证 对 $\dot{x}(t)$ 求导得

$$\dot{x}(t)=(e^{At}x_0)'=Ae^{At}x_0=Ax(t)。$$

而
$$x(0)=e^0x^0=I\cdot x^0=x^0。\quad ▨$$

唯一性证明：设有另一解 $y=y(t)$ 满足线性微分方程组 $\dot{y}(t)=Ay(t),y(0)=x^0$。要证明 $y(t)=e^{At}x^0$，为此考察函数 $g(t)=e^{-At}y(t)$，求 $\dot{g}(t)$。因 $\dot{g}(t)=-e^{-At}Ay(t)+e^{-At}\dot{g}(t)=-e^{-At}Ay(t)+e^{-At}Ag(t)=0$，即 $g(t)=$ 常向量。这样 $g(t)=g(0)=y(0)=x^0$，即 $e^{-At}y(t)=x^0,y(t)=e^{At}x^0$。 ▨

上述命题将常系数线性微分方程组的解写成一个紧凑的形式。它可看作为一阶微分方程 $y'=ay,y(0)=c$ 的解 $y=e^{ax}c$ 的直接推广，但 $e^{At}x^0$ 的计算并不容易。请读者留意计算的方法。

例 9 求常系数线性齐次方程组

$$\begin{cases}\dot{x}_1(t) = -7x_1 - 7x_2 + 5x_3 & x_1(0) = 3 \\ \dot{x}_2(t) = -8x_1 - 8x_2 - 5x_3 & x_2(0) = -2 \\ \dot{x}_3(t) = \quad -5x_2 & x_3(0) = 1\end{cases}$$ 的特解。

解 可将上面的方程写成矩阵的形式：

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\\ \dot{x}_3(t)\end{bmatrix} = \begin{bmatrix}-7 & -7 & 5\\ -8 & -8 & -5\\ 0 & -5 & 0\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\\ x_3(t)\end{bmatrix}, \begin{bmatrix}x_1(0)\\ x_2(t)\\ x_3(t)\end{bmatrix} = \begin{bmatrix}3\\ -2\\ 1\end{bmatrix}。$$

$$\dot{x} = Ax,\ x(0) = x^0。$$

要计算 e^{At}，先求矩阵 A 的特征值

$$|\lambda I - A| = \begin{vmatrix}\lambda+7 & 7 & -5\\ 8 & \lambda+8 & 5\\ 0 & 5 & \lambda\end{vmatrix} = (\lambda-5)(\lambda+5)(\lambda+15)。$$

因此 $A = P\begin{bmatrix}5 & & \\ & -5 & \\ & & -15\end{bmatrix}P^{-1}$，其中 P 是由特征向量构成，它们分别为

$A\alpha_1 = 5\alpha_1, \alpha_1 = (1\ \ -1\ \ 1)^T; A\alpha_2 = -5\alpha_2, \alpha_2 = (1\ \ -1\ \ -1)^T,$

$A\alpha_3 = -15\alpha_3, \alpha_3 = (2\ \ 3\ \ 1)^T;$

因此 $P = \begin{bmatrix}1 & 1 & 2\\ -1 & -1 & 3\\ 1 & -1 & 1\end{bmatrix}$

则

$$e^{At} = e^{PJP^{-1}t} = Pe^{Jt}P^{-1},$$

这里 J 是对角阵 $\begin{bmatrix}5 & & \\ & -5 & \\ & & -15\end{bmatrix}$，因而

$$e^{Jt}=\begin{bmatrix} e^{5t} & & \\ & e^{-5t} & \\ & & e^{-15t}\end{bmatrix}，（计算可参考例 6）。$$

故

$$e^{At}=Pe^{Jt}P^{-1}=\begin{bmatrix} 1 & 1 & 2 \\ -1 & -1 & 3 \\ 1 & -1 & 1\end{bmatrix}\begin{bmatrix} e^{5t} & & \\ & e^{-5t} & \\ & & e^{-15t}\end{bmatrix}\begin{bmatrix} 1 & 1 & 2 \\ -1 & -1 & 3 \\ 1 & -1 & 1\end{bmatrix}^{-1}$$

$$=\frac{1}{10}\begin{bmatrix} 2e^{5t}+4e^{-5t}+4e^{-15t} & -3e^{5t}-e^{-5t}+4e^{-15t} & 5e^{5t}-5e^{-5t} \\ -2e^{5t}-4e^{-5t}+6e^{-15t} & 3e^{5t}+e^{-5t}+6e^{-15t} & -5e^{5t}+5e^{-5t} \\ 2e^{5t}-4e^{-5t}+2e^{-15t} & -3e^{5t}+e^{-5t}+2e^{-15t} & 5e^{5t}+5e^{-5t}\end{bmatrix}$$

所以特解 $x(t)=e^{At}x^0=\frac{1}{10}\begin{bmatrix} 17e^{5t}+9e^{-5t}+4e^{-15t} \\ -17e^{5t}-9e^{-5t}+6e^{-15t} \\ 17e^{5t}-9e^{5t}+2e^{-15t}\end{bmatrix}$。

（2）e^{At} 的计算

设 $A=PJP^{-1}, J=J_1\oplus J_2\oplus\cdots\oplus J_s$，则

$$e^{At}=\sum_{n=0}^{\infty}\frac{A^nt^n}{n!}=\sum_{n=0}^{\infty}\frac{t^n}{n!}(PJP^{-1})^n$$

$$=\sum_{n=0}^{\infty}\frac{t^n}{n!}PJ^nP^{-1}=P\Big(\sum_{n=0}^{\infty}\frac{(Jt)^n}{n!}\Big)P^{-1}$$

$$=Pe^{Jt}P^{-1}=P\{e^{J_1t}\oplus e^{J_2t}\oplus\cdots\oplus e^{J_st}\}P^{-1}。$$

因此主要是要计算 e^{J_it}（J_i 是 Jordan 块）。

命题 4 设 $J=\begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{bmatrix}_{s\times s}$

则

$$e^{Jt}=e^{\lambda t}\begin{bmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{s-1}}{(s-1)!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{t^2}{2!} \\ & & & \ddots & t \\ & & & & 1 \end{bmatrix}。$$

证 因 $e^{Jt}=e^{(\lambda I+N)t}=e^{\lambda It}\cdot e^{Nt}=e^{\lambda t}\cdot\sum_{k=0}^{\infty}\frac{(Nt)^k}{k!}$。由于 N 是 s 阶幂零阵。$N^s=0$。所以上面无穷级数止于有限项。

$$e^{Jt}=e^{\lambda t}\sum_{k=0}^{s-1}\frac{(Nt)^k}{k!}=e^{\lambda t}\sum_{k=0}^{s-1}\frac{N^k t^k}{k!}。$$

写成单个矩阵的形式，便是所要求的结果。▨

例 10 设 $J=J_1\oplus J_2$，其中 $J_1=\begin{bmatrix}2&1&\\&2&1\\&&2\end{bmatrix}$，$J_2=\begin{bmatrix}-1&1\\&-1\end{bmatrix}$，求 e^{Jt}。

解 因 $e^{Jt}=e^{J_1t}\oplus e^{J_2t}$

而

$$e^{J_1t}=e^{2t}\begin{bmatrix}1&t&\frac{t^2}{2!}\\&1&t\\&&1\end{bmatrix}，\quad e^{J_2t}=e^{-t}\begin{bmatrix}1&t\\&1\end{bmatrix}.$$

所以

$$e^{Jt}=\left[\begin{array}{ccc|cc} e^{2t} & te^{2t} & \frac{t^2}{2!}e^{2t} & & \\ & e^{2t} & te^{2t} & & \\ & & e^{2t} & & \\ \hline & & & e^{-t} & te^{-t} \\ & & & & e^{-t}\end{array}\right]。$$

(3)* 常系数非齐次线性方程组的求解公式

这小节用常数变易来求非齐次线性方程组的求解公式。这方法实际上是一阶微分方程常数变易法的推广。现有 $\dot{x}(t)=Ax+f(t)$，这里 $f(t)=(f_1(t),f_2(t),\cdots,f_n(t))^T$ 是已知的单变量的 n 维向量，A 是 n 阶实阵。

首先已知 $x(t)=e^{At}c$（c 是常值向量），是常系数齐次微分方程 $\dot{x}(t)=Ax$ 的解。现将 c 看成是单变量的 n 维向量函数。设对应的非齐次微分方程组的解为 $x(t)=e^{At}c(t)$。代入微分方程得待定 $c(t)$。因为 $\dot{x}(t)=Ae^{At}c(t)+e^{At}\dot{c}(t)$，

因此 $$Ae^{At}c(t)+e^{At}\dot{c}(t)=Ae^{At}c(t)+f(t)。$$

故 $$e^{At}\dot{c}(t)=f(t),\quad \dot{c}(t)=e^{-At}f(t),$$

所以 $$c(t)=\int_{t_0}^{t}e^{-As}f(s)\mathrm{d}s+c',$$

其中 c' 是常数向量，因此非齐次微分方程的通解可写成为

$$x(t)=e^{At}\left\{\int_{t_0}^{t}e^{-As}f(s)\mathrm{d}s+c'\right\}。\tag{5-4}$$

如果要解初值问题

$$\begin{cases}\dot{x}(t)=Ax+f(t)\\ x(t_0)=x^0\end{cases}\tag{5-5}$$

则只要将初值代入式(5-4)有 $x^0=e^{At_0}c'$，$c'=e^{-At_0}x^0$，所以初值问题式(5-5)的解为

$$x(t)=e^{A(t-t_0)}x_0+\int_{t_0}^{t}e^{A(t-s)}f(s)\mathrm{d}s。\tag{5-6}$$

实际上公式(5-6)真正应用并不容易，但它可用在很多特殊的情况，并可求得解的显式表示。

例 11 如果 $\dot{x}(t)=Ax+c$，$x(t_0)=b$，$-\infty<t<+\infty$。则微分方程的解可写成

$$x(t)=e^{A(t-t_0)}b+\int_{t_0}^{t}e^{A(t-s)}\mathrm{d}s\cdot c。$$

如果上面的矩阵 A 是非奇异的，则还可进一步简化。事实上，

$(e^{A(t-s)})' = -Ae^{A(t-s)}$。因 A 是非奇异的，所 以 $-A^{-1}(e^{A(t-s)})' = e^{A(t-s)}$ 两边取积分。

$$-\int_{t_0}^{t} A^{-1}(e^{A(t-s)})'ds = \int_{t_0}^{t} e^{A(t-s)}ds$$
$$= -A^{-1}\int_{t_0}^{t}(e^{A(t-s)})')ds = A^{-1}[e^{A(t-t_0)} - I]。$$

这时微分方程的解可写成为

$$x(t) = e^{A(t-t_0)}b + A^{-1}[e^{A(t-t_0)} - I]c。$$

第五节　矩阵函数的计算

1. 矩阵函数的计算概念

前面是用幂级数来定义矩阵函数。因此在计算矩阵函数时，直接由矩阵幂级数结合 Hamilton-Cayley 定理进行。如果给出一个 n 阶矩阵 A 的最小多项式

$$m(\lambda) = \lambda^s + a_{s-1}\lambda^{s-1} + \cdots + a_1\lambda + a_0。$$

矩阵 A 的任何次幂都可通过 $I, A, \cdots, A^{s-1}$ 的线性组合表示。因而可设想，一个用幂级数定义的函数 $f(t)$，在计算 $f(A)$ 时，可通过一个次数不超过 $s-1$(s 是最小多项式的次数)的矩阵多项式来计算。现在回顾一下 Jordan 块的情形：设 J 是 n 阶的 Jordan 块。设 $f(t) = \sum \alpha_k t^k$，则

$$f(J) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ & f(\lambda) & f'(\lambda) & \cdots & \frac{f^{(n-2)}(\lambda)}{(n-2)!} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & f'(\lambda) \\ & & & & f(\lambda) \end{pmatrix}。$$

由上面的表达式可知，$f(J)$ 仅涉及函数 $f(t)$ 在 $t = \lambda$ 时的 $f(\lambda)$。$f'(\lambda), \cdots, f^{(n-1)}(\lambda)$（要记住对 Jordan 块而言，$n$ 亦是它的最小多

项式的次)的数值。因此如果能找到一多项式 $p(x)$，在 $x=\lambda$ 处有下面各等式成立。

$$p(\lambda)=f(\lambda),p'(\lambda)=f'(\lambda),\cdots p^{(n-1)}(\lambda)=f^{(n-1)}(\lambda)。$$

则应有 $p(A)=f(A)$。所以，亦可用下面方式来计算矩阵函数。

设 A 的最小多项式为

$$m(\lambda)=(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_s)^{r_s}。\tag{5-7}$$

式中 $\lambda_1\cdots\lambda_s$ 是 A 的相异特征值。如果多项式 $r(\lambda)$ 与函数 $f(\lambda)$ 在下面的数值是一样的。

$$\begin{array}{llll} f(\lambda_1) & f'(\lambda_1) & \cdots & f^{(r_1-1)}(\lambda_1) \\ f(\lambda_2) & f'(\lambda_2) & \cdots & f^{(r_2-1)}(\lambda_2) \\ \vdots & \vdots & & \\ f(\lambda_s) & f'(\lambda_s) & \cdots & f^{(r_s-1)}(\lambda_s)。\end{array}\tag{5-8}$$

则有 $r(A)=f(A)$。一个多项式 $r(\lambda)$，在 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 处和函数 $f(\lambda)$ 有相同的数组(5-8)，称它们在谱上有相同的值。实际上矩阵函数 $f(A)$ 是由 $f(t)$ 在 A 上的谱的数组(5-8)完全确定的。一个最简单的事实是，两个多项式 $p(x)\neq q(x)$，但还可以有 $p(A)=q(A)$。

命题 5 $p_1(\lambda),p_2(\lambda)$ 是 C 上二个多项式，$A\in C^{n\times n}$，它的最小多项式为式(5-7)，则 $p_1(A)=p_2(A)\Longleftrightarrow$ 它们在谱上的值相等，即

$$\begin{array}{llll} p_1(\lambda_1)=p_2(\lambda_1) & p_1'(\lambda_1)=p_2'(\lambda_1) & \cdots & p_1^{(r_1-1)}(\lambda_1)=p_2^{(r_1-1)}(\lambda_1) \\ p_1(\lambda_2)=p_2(\lambda_2) & p_1'(\lambda_2)=p_2'(\lambda_2) & \cdots & p_1^{(r_2-1)}(\lambda_2)=p_2^{(r_2-1)}(\lambda_2) \\ \vdots & & & \\ p_1(\lambda_s)=p_2(\lambda_s) & p_1'(\lambda_s)=p_2'(\lambda_s) & \cdots & p_1^{(r_s-1)}(\lambda_s)=p_2^{(r_s-1)}(\lambda_s)。\end{array}$$

证 $\Longleftarrow$ 由前面的讨论即可知道充分性成立。

$\Longrightarrow$ 设 $p_0(\lambda)=p_1(\lambda)-p_2(\lambda)$，因此 $p_0(A)=0$。$p_0(\lambda)$ 是 A 的一个零化多项式，因此有 $m(\lambda)\mid p_0(\lambda)$。因为 λ_k 是 $m(\lambda)$ 的 r_k 重根。这样 λ_k 亦至少是 $p_0(\lambda)$ 的 r_k 重根，所以 $p_0(\lambda_k)=0$，$p_0'(\lambda_k)=0$，$\cdots$，$p^{(r_k-1)}(\lambda_k)=0$。即 $p_1(\lambda_k)=p_2(\lambda_k)$，$p_1'(\lambda_k)=p_2'(\lambda_k),\cdots,p_1^{(r_k-1)}(\lambda_k)=p_2^{(r_k-1)}(\lambda_k)$。▨

下面通过例子来说明上面的计算。

例 12 已知矩阵 A 的最小多项式是 $m(\lambda)=\lambda^2(\lambda-\pi)(\lambda+\pi)$，证明 $\sin A=A-\frac{1}{\pi^2}A^3$。

解 先作多项式 $f(t)=t-\frac{1}{\pi^2}t^3$，然后证明 $f(t)$ 与 $\sin t$ 在 A 的谱 $\lambda=0$，$\lambda=\pi$，$\lambda=-\pi$ 上的数值相同。即 $f(0)=\sin 0$，$f'(0)=\sin'(0)$，$f(\pi)=\sin\pi$，$f(-\pi)=\sin(-\pi)$。经验证，它们是相等的，所以结论成立。

例 13 已知 $A=\begin{bmatrix}1&4\\3&2\end{bmatrix}$，求 e^A。

解 A 的特征多项式为 $p(\lambda)=(\lambda-5)(\lambda+2)$，因此亦是 A 的最小多项式。现在要找一个一次多项式 $r(t)=\alpha+\beta t$，它在 A 的谱 $\lambda=5,\lambda=-2$ 上，与函数 $f(t)=e^t$ 有相同的数值 $r(5)=e^5$，$r(-2)=e^{-2}$，$\alpha+5\beta=e^5$，$\alpha-2\beta=e^{-2}$，最后解得

$$r(t)=\frac{1}{7}(2e^5+5e^{-2})+\frac{1}{7}(e^5-e^{-2})t,$$

因 $e^A=r(A)$，故

$$e^A=\frac{1}{7}(2e^5+5e^{-2})I+\frac{1}{7}(e^5-e^{-2})A。$$

例 14

已知 $A=\begin{bmatrix}-2&2&-2&4\\-1&2&-1&1\\0&0&1&0\\-2&1&-1&4\end{bmatrix}$，求 $\cos\pi A$。

解 A 的最小多项式为 $m(t)=t^3-4t^2+5t-2=(t-1)^2(t-2)$。

要找的多项式 $r(t)=\alpha+\beta t+\gamma t^2$，它满足条件

$$r(2)=\cos 2\pi, r(1)=\cos\pi, r'(1)=-\pi\sin\pi$$

可列得方程：$\alpha+2\beta+\gamma=1$，$\alpha+\beta+\gamma=-1$，$\beta+2\gamma=0$。

解得：$r(t)=1-4t+2t^2$，所以

$$\cos\pi A=I-4A+2A^2=\begin{bmatrix}-3&0&0&4\\0&-1&0&0\\0&0&-1&0\\-2&0&0&3\end{bmatrix}。$$

例 15 已知 $A=\begin{bmatrix}0&-1\\4&4\end{bmatrix}$，试计算 $\arcsin\frac{1}{4}A$。

解 实际上特征多项式 $p(\lambda)=(\lambda-2)^2$ 容易验证，它亦是 A 的最小多项式。原先是用 $\arcsin\frac{t}{4}$ 的幂级数来定义 $\arcsin\frac{A}{4}$。幂级数收敛半径为 4，而 A 的特征值 $\lambda=2$，因此 $\arcsin\frac{A}{4}$ 是有定义的。现在要找一个一次多项式 $r(t)=\alpha+\beta t$，它满足条件

$$r(2)=\arcsin\frac{1}{2},\quad r'(2)=\left(\arcsin\frac{t}{4}\right)'_{t=2}$$

$$=\frac{1}{\sqrt{4^2-t^2}}\bigg|_{t=2}=\frac{1}{2\sqrt{3}}。$$

要解方程 $\alpha+2\beta=\arcsin\frac{1}{2}$，$\beta=\frac{1}{2\sqrt{3}}$。解得

$$r(t)=\left(\arcsin\frac{1}{2}-\frac{1}{\sqrt{3}}\right)+\frac{1}{2\sqrt{3}}t,$$

$$r(A)=\left(\arcsin\frac{1}{2}-\frac{1}{\sqrt{3}}\right)I+\frac{1}{2\sqrt{3}}A$$

$$=\begin{bmatrix}\arcsin\frac{1}{2}-\frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}}\\ \frac{2}{\sqrt{3}} & \arcsin\frac{1}{2}+\sqrt{3}\end{bmatrix}。$$

2．关于计算 e^{At} 的若干特殊情形

前面虽然讲了一般计算 e^{At} 的方法，但实际上真正计算并不容易。因为既要求 Jordan 标准型，又要求变换矩阵。为方便起见，将一些特殊情形汇总在下面，希望对读者能有些帮助。

(1) A 是可对角化矩阵

由 $A=PDP^{-1}$ 可得

$$e^{At}=e^{(PDP^{-1})t}=Pe^{Dt}P^{-1}$$

$$=P\begin{bmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ & & & e^{\lambda_n t} \end{bmatrix}P^{-1}。$$

(2) A 有 n 个相异的特征值

这种情形 A 的特征多项式与最小多项式是一样的。设 n 个相异的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$。按前面所讲的，要求一个 $n-1$ 次多项式它在 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 处与函数 $f(\lambda)=e^{\lambda t}$ 有相同的值，即 $n-1$ 次多项式 $p(\lambda)$ 满足条件

$$p(\lambda_1)=e^{\lambda_1 t},p(\lambda_2)=e^{\lambda_2 t},\cdots,p(\lambda_n)=e^{\lambda_n t}。$$

这个多项式可用 Lagrange 内插多项式来求得

$$p(\lambda)=\sum_{i=1}^{n}e^{\lambda_i t}L_i(\lambda),$$

因为

$$L_i(\lambda)=\frac{(\lambda-\lambda_1)\cdots(\lambda-\lambda_{i-1})(\lambda-\lambda_{i+1})\cdots(\lambda-\lambda_n)}{(\lambda_i-\lambda_1)\cdots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\cdots(\lambda_i-\lambda_n)},$$

这样 $$e^{At}=\sum_{i=1}^{n}e^{\lambda_i t}L_i(A),$$

其中

$$L_i(A)=\prod_{\substack{j=1\\ j\neq i}}^{n}\frac{A-\lambda_j I}{\lambda_i-\lambda_j}。$$

(3) A 的特征值全部相同($\lambda=\lambda_0$)

这时 A 的最小多项式 $m(\lambda)=(\lambda-\lambda_0)^s$，因此 $(A-\lambda_0 I)^s=0$，但 $(A-\lambda_0 I)^{s-1}\neq 0$。这时

$$e^{At}=e^{[\lambda_0 I+(A-\lambda_0 I)]t}=e^{\lambda_0 It}\cdot e^{(A-\lambda_0 I)t}$$

$$=e^{\lambda_0 t}e^{(A-\lambda_0 I)t}=e^{\lambda_0 t}\sum_{k=0}^{s-1}\frac{t^k}{k!}(A-\lambda_0 I)^k。$$

命题 6 A 是 n 阶方阵($n\geqslant 3$)，它只有两个相异的特征值，其中之一为 $n-1$ 重特征值。

设 $n-1$ 重特征值为 λ，另一个特征值为 μ，则

$$e^{At}=e^{\lambda t}\sum_{k=0}^{n-2}\frac{t^k}{k!}(A-\lambda I)^k+\left\{\frac{e^{\mu t}}{(\mu-\lambda)^{n-1}}\right.$$

$$\left.-\frac{e^{\lambda t}}{(\mu-\lambda)^{n-1}}\sum_{k=0}^{n-2}(\mu-\lambda)^k\frac{t^k}{k!}\right\}(A-\lambda I)^{n-1}。$$

证 $$e^{At}=e^{(A-\lambda I)t}e^{\lambda It}=e^{\lambda t}\sum_{k=0}^{\infty}\frac{(A-\lambda I)^k}{k!}t^k$$

$$=e^{\lambda t}\left\{\sum_{k=0}^{n-2}\frac{t^k(A-\lambda I)^k}{k!}+\sum_{k=n-1}^{\infty}\frac{(A-\lambda I)^k}{k!}t^k\right\}。\tag{5-9}$$

下面利用矩阵 A 的特征方程 $(x-\lambda)^{n-1}(x-\mu)$ 来简化式 (5-9) 中后一个和式。

$$(x-\lambda)^{n-1}(x-\mu)=(x-\lambda)^{n-1}\{x-\lambda+\lambda-\mu\}$$
$$=(x-\lambda)^n+(\lambda-\mu)(x-\lambda)^{n-1}。$$

由此可得

$$(A-\lambda I)^n=(\mu-\lambda)(A-\lambda I)^{n-1}。$$

连续应用这个关系，便可得出

$$(A-\lambda I)^{n-1+r}=(\mu-\lambda)^r(A-\lambda I)^{n-1}。$$

因此

$$\sum_{k=n-1}^{\infty}\frac{(A-\lambda I)^k}{k!}t^k=\sum_{r=0}^{\infty}\frac{(A-\lambda I)^{n-1+r}}{(n-1+r)!}t^{n-1+r}$$

$$=\sum_{r=0}^{\infty}\frac{(\mu-\lambda)^r(A-\lambda I)^{n-1}}{(n-1+r)!}t^{n-1+r}$$

$$=\frac{(A-\lambda I)^{n-1}}{(\mu-\lambda)^{n-1}}\sum_{r=0}^{\infty}\frac{(\mu-\lambda)^{n-1+r}t^{n-1+r}}{(n-1+r)!}$$

$$=\frac{(A-\lambda I)^{n-1}}{(\mu-\lambda)^{n-1}}\cdot\sum_{k=n-1}^{\infty}\frac{(\mu-\lambda)^k t^k}{k!}$$

$$=\frac{(A-\lambda I)^{n-1}}{(\mu-\lambda)^{n-1}}\cdot\left\{e^{(\mu-\lambda)t}-\sum_{k=0}^{n-2}\frac{(\mu-\lambda)^k t^k}{k!}\right\}$$

只要将这个结果代入式(5-9)，便可得所要求的结果。 ▨

第六章　广义逆矩阵

引言　1955 年 Penrose 建立了下面的命题：对任一个矩阵，$A\in C^{m\times n}$，存在唯一的矩阵 G，同时满足下面四个方程：

$$
\begin{array}{ll}
\text{(i)}\ AGA=A; & \text{(iii)}\ (AG)^*=AG;\\
\text{(ii)}\ GAG=G; & \text{(iv)}\ (GA)^*=GA。
\end{array}
\tag{6-1}
$$

它是在Moore(1922年) 的论文基础上提出的。现将同时满足上面矩阵方程的矩阵 G 称为矩阵 A的Moore–Penrose逆，或简称为 M–P 逆，记为 A^+。上面的问题来源于线性方程组求解。总的一个要求是：对线性方程组 $Ax=b$ 的解能用矩阵的形式给出。它们是

(i) 相容方程的解；

(ii) 相容方程的最小范数解；

(iii) 矛盾方程的最优解；

(iv) 矛盾方程的最小范数的最优解。

要解决上面提出的问题，就要求将方阵的逆阵推广到任意的 $m\times n$ 型矩阵。以后会发现，随着问题(i)～(iv)的逐步解决，所要求的"逆矩阵"G 是满足式(6-1)中某些方程或全部方程。

第一节　矩阵的单边逆

1. 定义

所谓矩阵的单边逆，是指矩阵的左逆或右逆 。设 A 是 $m\times n$ 矩阵。

如果存在一 $n\times m$ 型矩阵 B 有 $AB=I_m$， 称 A 有右逆。

如果存在一 $n\times m$ 型矩阵 B 有 $BA=I_n$， 称 A 有左逆。

第一章的例 6，已经给出一个 $m\times n$ 型矩阵有单边逆的充分

必要条件。现重述如下：

命题 1 A 有右逆$\Longleftrightarrow r_A=m$（行满秩），右逆记号 A_R^{-1}。

A 有左逆$\Longleftrightarrow r_A=n$（列满秩），左逆记号 A_L^{-1}。

2. 利用初等变换的标准型来求得矩阵的单边逆

（1）求 A 的左逆

由于这种情况有$r_A=n(m\geqslant n)$，所以一定可以通过行变换将 A 变为$\begin{bmatrix}I_n\\O\end{bmatrix}$，或者写成 $PA=\begin{bmatrix}I_n\\O\end{bmatrix}$，$P$ 是非奇异矩阵。由矩阵的乘法可知 P 的前 n 行构成的子阵 B 有 $BA=I_n$。

例 1 设 $A=\begin{bmatrix}1 & -2\\0 & 1\\0 & 0\end{bmatrix}$，求 A_L^{-1}。

解 因$\left[\begin{array}{cc|ccc}1 & -2 & 1 & 0 & 0\\0 & 1 & 0 & 1 & 0\\0 & 0 & 0 & 0 & 1\end{array}\right]\longrightarrow\left[\begin{array}{cc|ccc}1 & 0 & 1 & 2 & 0\\0 & 1 & 0 & 1 & 0\\0 & 0 & 0 & 0 & 1\end{array}\right]=\left[\begin{array}{c|c}I_2 & \\ & P\\O & \end{array}\right]$，由$P$的前二行构成的子阵 $B=\begin{bmatrix}1 & 2 & 0\\0 & 1 & 0\end{bmatrix}$，便是$A_L^{-1}$。可直接验证它确实满足条件 $BA=I_2$。

（2）求 A 的右逆

这种情形一定有 $r_A=m(m\leqslant n)$，所以一定可通过列变换，使得 A 变为$[I_m\ \ O]$，或者写成 $AP=[I_m\ \ O]$。这说明 P 的前 m 列与 A 相乘是单位阵 I_m。

例 2 设 $A=\begin{bmatrix}1 & 2 & -1\\0 & -1 & 2\end{bmatrix}$，求 A_R^{-1}。

解 因为$\left[\begin{array}{ccc}1 & 2 & -1\\0 & -1 & 2\\\hline 1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{array}\right]\longrightarrow\left[\begin{array}{ccc}1 & 0 & 0\\0 & -1 & 2\\\hline 1 & -2 & 1\\0 & 1 & 0\\0 & 0 & 1\end{array}\right]\longrightarrow\left[\begin{array}{ccc}1 & 0 & 0\\0 & -1 & 0\\\hline 1 & -2 & -3\\0 & 1 & 2\\0 & 0 & 1\end{array}\right]$

$$\rightarrow \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 1 & 2 & -3 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{bmatrix}, \text{ 所以 } A_R^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ 0 & 0 \end{bmatrix}.$$

3. 矩阵的单边逆一般不是唯一的

如上面的矩阵A,它的另一个右逆为 $A_R^{-1} = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$。

为了给出单边逆的表达式，先证明下面的命题。实际上它也是矩阵理论的一个常用的结论。

命题 2 若 $A \in C^{m\times n}$,则有 $r(A) = r(A^*A) = r(AA^*)$

证 实际上证明 $\operatorname{Ker} A = \operatorname{Ker} A^*A$, $\operatorname{Ker} A \subseteq \operatorname{Ker} A^*A$ 是显然的。下面证明 $\operatorname{Ker} A^*A = \operatorname{Ker} A$。设 $x \in \operatorname{Ker} A^*A$, 因此有 $A^*Ax = 0$, $x^*A^*Ax = 0$, $(Ax)^*Ax = 0$, $|Ax|^2 = 0$。即 $Ax = 0$, $x \in \operatorname{Ker} A$。综上所述有 $\operatorname{Ker} A = \operatorname{Ker}(A^*A)$, $r(A) = n - \dim \operatorname{Ker} A = n - \dim \operatorname{Ker} A^*A = r(A^*A)$。另一个结论可类似证明。 ▨

由上面的命题容易给出矩阵单边逆的表达式与通式。

当 $r(A) = n$, A^*A 是满秩阵 $(r(A^*A) = r(A) = n)$, 所以 $(A^*A)^{-1}$存在，$(A^*A)^{-1}A^*$便是 A 的一个左逆。

当 $r(A) = m$ 时,同理 $(AA^*)^{-1}$ 存在，$A^*(AA^*)^{-1}$ 是 A 的一个右逆。

同时还可给出单边逆的通式：

只要 AVA^* 是可逆,则右逆的通式为$A_R^{-1} = VA^*(AVA^*)^{-1}$。只要 A^*UA 是可逆,则左逆的通式为 $A_L^{-1} = (A^*UA)^{-1}A^*U$。

在证明通式的表达式时，只要注意到：(i) 上面的表达式所

确定的矩阵确实是单边逆。(ii) G 是 A 的任一个右逆，在 G 处的通式中只要选 $V=GA$，$VA^*(AVA^*)^{-1}=GAA^*(AGAA^*)^{-1}=G$。这说明只要适当选取 V，在通式中便可得到所给出的任一个右逆，同理 G' 是 A 的任一个左逆，只要选 $U=AG'$，即可证明左逆的通式成立。

例 3 设$A=\begin{bmatrix}1 & 2 & -1\\ 0 & -1 & 2\end{bmatrix}$，求 A_R^{-1}。

解 易知 $r(A)=2$，则 $A_R^{-1}=A^T(AA^T)^{-1}=\frac{1}{14}\begin{bmatrix}5 & 4\\ 6 & 2\\ 3 & 8\end{bmatrix}$。

下面由通式来求出前面例 2 所求的 A_R^{-1}。

因为 $V=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}$，因此 $AV=\begin{bmatrix}1 & 2 & 0\\ 0 & -1 & 0\end{bmatrix}$，

则
$$AVA^T=\begin{bmatrix}1 & 2 & 0\\ 0 & -1 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\ 2 & -1\\ 0 & 0\end{bmatrix}=\begin{bmatrix}5 & -2\\ -2 & 1\end{bmatrix}。$$

所以$VA^T(AVA^T)^{-1}=\begin{bmatrix}1 & 0\\ 2 & -1\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & 2\\ 2 & 5\end{bmatrix}=\begin{bmatrix}1 & 2\\ 0 & -1\\ 0 & 0\end{bmatrix}$。

如果用 G 表示 $(A^*A)^{-1}A^*$ 或 $A^*(AA^*)^{-1}$，则容易验证 G 满足 Penrose 提出的四个矩阵方程 (6-1)。但要注意由通式表达的右逆或左逆并不满足这四个矩阵方程，这里表明在 A 是列满秩或行满秩时，A^+ 是存在的。即为它的单边逆。

第二节 广义逆$\mathbf{A^-}$及相容方程的解

在这节里总假定线性方程组

$$Ax=b, \ A\in C^{m\times n}。\qquad (E)$$

是相容的，即 $b\in R(A)$。此目的是要找一个矩阵 $G\in C^{n\times m}$，使得方程的解可写成 $x=Gb$。以后讨论方程(E)的解时，常常采用映射的观点，实际上 A 可看成 C^n 至 C^m 的一个线性映射，对方程的(E)求解，变成求向量 b 的像源 x_0 的问题（如图 6-1）。

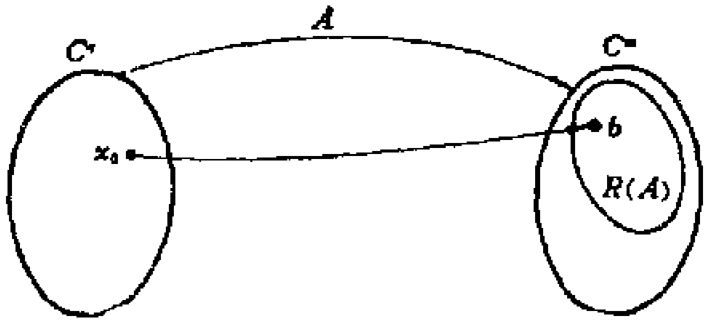

图 6-1

因此如果 $r(A)=m$，这时 $R(A)=C^m$。对任何一个 b 总有解。用 G 记 A 的右逆，则 $x=Gb=A_R^{-1}b$ 就是所要求解的表达式。如果 $r(A)=n$，又 $b\in R(A)$，用 G 记 A 的左逆，则 $x=Gb=A_L^{-1}b$ 亦是解的表达式。所以下面的讨论便假定 $rA)=r<\min(m,n)$。

1．广义逆 A^- 应满足的方程

定理 1 $A\in C^{m\times n}$，$G\in C^{n\times m}$，对任意 $b\in R(A)$，$x=Gb$ 是方程(E)的解$\Longleftrightarrow AGA=A$。

证 这个命题说明：相容方程的解，用矩阵 G 表示，则 G 应满足 Pearose 四个矩阵方程中的第一个方程。

$\Longrightarrow$因为对任何 $b\in R(A)$，方程(E)的解均可写成$x=Gb$。现取 b 为矩阵 A 的各个列向量 $A^{(1)}$，$A^{(2)}$，$\cdots$，$A^{(n)}$。这样，方程 $Ax=A^{(i)}$ 的解便可写成 $x=GA^{(i)}$ $(i=1,2,\cdots,n)$，即 $AGA^{(i)}=A^{(i)}(i=1,2,\cdots,n)$。将这 n 个方程合并为一个矩阵方程，即为

$AGA=A$（矩阵的乘法）。

$\Longleftarrow$因为 $b\in R(A)$，所以 $b=Ay$（即 b 一定有像源）。令 $x=Gb$，则 $AGb=AGAy=Ay=b$。这说明了如果 $AGA=A$时，相容方程的解总可写成 $x=Gb$。 ▨

定义 设 $A\in C^{m\times n}$，如果存在 $G\in C^{n\times m}$ 有

$$AGA=A,$$

则称 G 是 A 的广义逆，简称为 A 的 g 逆，记为 A^-。

前面已经知道当 A 是行满秩或列满秩时，它的 g 逆总是存在的，下面提供一些在一般情形下，g 逆的计算方法，同时说明g逆是存在的。

2．A^-的计算

(1) 用矩阵 A 的满秩分解

设 $A\in C^{m\times n}$，$r(A)=r$。在第一章中已提出，一定存在一个列满秩的矩阵 $C(m\times r$ 型)与一个行满秩的矩阵 $B(r\times n$ 型)有$A=CD$。则

$$A^-=D_R^{-1}C_L^{-1}。$$

这个结论可直接验证。

例 4 设 $A=\begin{bmatrix}1&2&0\\0&0&2\\2&4&0\end{bmatrix}$，求 A^-。

解 先求 A 的满秩分解：

$$\left[\begin{array}{ccc|ccc}1&2&0&1&0&0\\0&0&2&0&1&0\\2&4&0&0&0&1\\\hline 1&0&0&&&\\0&1&0&&&\\0&0&1&&&\end{array}\right]\longrightarrow\left[\begin{array}{ccc|ccc}1&0&0&1&0&0\\0&0&2&0&1&0\\0&0&0&-2&0&1\\\hline 1&-2&0&&&\\0&1&0&&&\\0&0&1&&&\end{array}\right]\longrightarrow$$

$$\begin{array}{cc}
 & \begin{array}{cccccc} & & & & P & \\ & & & & \downarrow & \end{array} \\
\begin{array}{c} \\ \\ \\ \\ Q\rightarrow \\ \\ \end{array} & \left[\begin{array}{ccc|ccc}
1 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & \frac{1}{2} & 0 \\
0 & 0 & 0 & -2 & 0 & 1 \\
\cline{4-6}
1 & 0 & -2 & & & \\
0 & 0 & 1 & & & \\
0 & 1 & 0 & & &
\end{array}\right]
\end{array}。$$

因而有 $PAQ=\begin{bmatrix} I_2 & \\ & 0\end{bmatrix}$，$A=P^{-1}\begin{bmatrix} I_2 & \\ & 0\end{bmatrix}Q^{-1}$。

又因

$$P^{-1}=\begin{bmatrix}1&0&0\\0&2&0\\2&0&1\end{bmatrix},\quad Q^{-1}=\begin{bmatrix}1&2&0\\0&0&1\\0&1&0\end{bmatrix}。$$

则

$$A=\begin{bmatrix}1&0&0\\0&2&0\\2&0&1\end{bmatrix}\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\end{bmatrix}\begin{bmatrix}1&2&0\\0&0&1\\0&1&0\end{bmatrix}=\begin{bmatrix}1&0\\0&2\\2&0\end{bmatrix}\begin{bmatrix}1&2&0\\0&0&1\end{bmatrix}=CD,$$

故

$$C_L^{-1}=\frac{1}{5}\begin{bmatrix}1&0&2\\0&\frac{5}{2}&0\end{bmatrix},D_R^{-1}=\frac{1}{5}\begin{bmatrix}1&0\\2&0\\0&5\end{bmatrix},$$

所以

$$A^-=D_R^{-1}C_L^{-1}=\frac{1}{25}\begin{bmatrix}1&0&2\\2&0&4\\0&\frac{25}{2}&0\end{bmatrix}。$$

由这个例子可以看出，如果用 A 的满秩分解去求 A^-，显然是非常麻烦的。然而，可以验证用这方法求出的广义逆 A^-，它满足 Penrose 的四个矩阵方程(6-1)，即它是 A^+。下面介绍其它的方法以简化计算。

(2) 用矩阵的行交换与列交换的方法

命题 3 设 B_0 是 $m\times n$ 型阵，且 $r(B_0)=r$。它的一个不为零的 r 阶子式 B_1 在矩阵的左上角，

$$B_0=\begin{bmatrix}B_1 & B_2\\ B_3 & B_4\end{bmatrix}, |B_1|\neq 0。$$

则 $B_4=B_3B_1^{-1}B_2$。

证 B_0 的各个列向量是前 r 个列向量的线性组合，因此必存在一矩阵 D(是 $r\times(n-r)$ 型)有$\begin{bmatrix}B_2\\ B_4\end{bmatrix}=\begin{bmatrix}B_1\\ B_3\end{bmatrix}D$，于是有 $B_2=B_1D$，$B_4=B_3D$。由前一式便可得 $D=B_1^{-1}B_2$，代入后式为 $B_4=B_3B_1^{-1}B_2$。 ▨

命题 4 $B_0^-=\begin{bmatrix}B_1^{-1} & 0\\ 0 & 0\end{bmatrix}_{n\times m}$

证 因为 $B_0\begin{bmatrix}B_1^{-1} & 0\\ 0 & 0\end{bmatrix}B_0=\begin{bmatrix}B_1 & B_2\\ B_3 & B_4\end{bmatrix}\begin{bmatrix}B_1^{-1} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}B_1 & B_2\\ B_3 & B_4\end{bmatrix}$

$$=\begin{bmatrix}I & 0\\ B_3B_1^{-1} & 0\end{bmatrix}\begin{bmatrix}B_1 & B_2\\ B_3 & B_4\end{bmatrix}=\begin{bmatrix}B_1 & B_2\\ B_3 & B_3B_1^{-1}B_2\end{bmatrix}$$

$$=\begin{bmatrix}B_1 & B_2\\ B_3 & B_4\end{bmatrix}=B,$$

因此命题成立。 ▨

命题 5 设 $m\times n$ 型矩阵 A 通过行与列的交换变为 B_0(见命题 3)，这时，$PAQ=B_0$，则 $A^-=QB_0^-P$ 。

证 $A(QB_0^-P)A=(P^{-1}B_0Q^{-1})(QB_0^-P)(P^{-1}B_0Q^{-1})=P^{-1}B_0B_0^-B_0Q^{-1}=P^{-1}B_0Q^{-1}=A$。 ▨

例 5 已知 $A=\begin{bmatrix}1 & 1 & 1 & 0\\ -1 & -1 & -1 & 0\\ 1 & 1 & 0 & 0\end{bmatrix}$，求$A^-$。

解 容易看出 $r(A)=2$，将 A 的第二、三行互换，然后再将 A 的第一、三列互换，便有

$$PAQ=\left[\begin{array}{cc|cc}1 & 1 & 1 & 0\\ 0 & 1 & 1 & 0\\ \hline -1 & -1 & -1 & 0\end{array}\right]=B_0,\quad B_1=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix},$$

$$B_1^{-1}=\begin{bmatrix}1 & -1\\ 0 & 1\end{bmatrix}。$$

因此可得

$$B_0^{-1}=\begin{bmatrix}1 & -1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}。$$

这时
$$P=\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix},\quad Q=\begin{bmatrix}0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}。$$

所以 $A^- = QB_0^- P=\begin{bmatrix}0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & -1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix}=\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 1 & 0 & -1\\ 0 & 0 & 0\end{bmatrix}。$

(3) 用初等变换

这个方法是将 A 通过初等交换变为标准型，再求出 A^-。但初等变换的标准型有二种

$$R_1=\begin{bmatrix}I_r & K\\ 0 & 0\end{bmatrix}_{m\times n},\tag{6-2}$$

$$R_2=\begin{bmatrix}I_r & 0\\ 0 & 0\end{bmatrix}_{m\times n}。\tag{6-3}$$

式(6-2)实际上是 Hermite 标准型。如果 Hermite 标准型达不到式(6-2)的要求,则还要通过列的适当交换。否则式(6-2)仅涉及行的变换。

命题 6 若 $R_1^- = \begin{bmatrix} I_r & \\ & L \end{bmatrix}_{n \times m}$，$PAQ = R_1$，则 $A^- = QR_1^- P$。

命题 7 若 $R_2^- = \begin{bmatrix} I_r & \\ & L \end{bmatrix}$，$PAQ = R_2$，则 $A^- = QR_2^- P$。

这两个命题证明留给读者作为练习。

由 A^- 的表达式可知 $r(A^-) = r(A) + r(L)$。因为 L 是任意的，所以 A^- 的秩可取 r 与 $\min(m,n)$ 之间任一正整数。特别如A是方阵，则 A^- 可取非奇异的(尽管 A 可能是奇异的)。

例 6 设 $A = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$，求 A^-。

解

$$\left[\begin{array}{ccc|cccc} 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|cccc} 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 1 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|cccc} 0 & 0 & 0 & 1 & 0 & -2 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 1 \end{array}\right]。$$

连续施行行初等变换到这一步，以后仅用行与列的交换便可以达到标准型(将最后一个矩阵一、三行交换，然后再将前面的子块第一、三列交换即可)。

因为

$$PAQ = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & -2 & 0 \\ 0 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 2 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

所以

$$A^- = Q\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\end{bmatrix}P$$

$$=\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\end{bmatrix}\begin{bmatrix}0&0&1&0\\0&1&0&0\\1&0&-2&0\\0&-1&-1&1\end{bmatrix}=\begin{bmatrix}0&0&0&0\\0&1&0&0\\0&0&1&0\end{bmatrix}。$$

可直接验证 $AA^-A=A$ 确实成立。

3. A^- 的基本性质

(i) $(A^*)^-=(A^-)^*$;

(ii) 设 P,Q 为非奇异矩阵,则 $(PAQ)^-=Q^{-1}A^-P^{-1}$;

(iii) $\operatorname{rank} A\leqslant \operatorname{rank} A^-$。

以上三个性质容易由等式 $AA^-A=A$ 直接证明,留给读者作为练习。

下面的命题说明矩阵 A^- 的另一些重要性质。

命题 8 若 $A\in C^{m\times n}$, 则

(i) AA^- 与 A^-A 都是幂等阵;

(ii) $R(AA^-)=R(A)$, $N(A^-A)=N(A)$。

证 这命题的意思是:如果将矩阵 A^-A 与 AA^- 分别看作为 C^n 与 C^m 的线性变换,则它们都是投影算子。(ii) 是给出了投影算子 AA^- 的像空间,以及投影算子 A^-A 的零空间。

下面证明(i): $AA^-A=A$, 所以有 $A^-(AA^-A)=A^-A$, 即 $(A^-A)^2=(A^-A)$。同理易证 $(AA^-)^2=AA^-$。

(ii) 的证明: 首先,显然有 $R(AA^-)\subseteq R(A)$。下面只要证明 $r(AA^-)\geqslant r(A)$。这是因为 $AA^-A=A$, 所以结果成立。

$N(A^-A)=N(A)$ 的证明: 一个明显的关系是 $N(A)\subseteq N(A^-A)$, 因此只要证明有 A 的零度 $=A^-A$ 的零度。这样便归结为证明 $r(A)=r(A^-A)$。由 $AA^-A=A$ 可立即得出不等式 $r(A^-A)\leqslant r(A)\leqslant r(A^-A)$, 所以命题成立。

由 A^- 的计算可知 A^- 不是唯一的。现将满足 $AGA=A$ 的 G 的全体构成的集合记为 $A\{1\}$。下面讨论集合 $A\{1\}$ 的表达式，即广义逆的通式。

命题 9 A 的 g 逆的全体的集合 $A\{1\}$ 的表达式为

$$G=A^-+V-A^-AVAA^-$$

其中 V 为任一 $n\times m$ 矩阵。

证 首先对任何一个 V 所确定的 G，显然都有 $AGA=A$，即 $G\in A\{1\}$。其次设 G_0 是 $A\{1\}$ 的任一个元，只要令 $V=G_0-A^-$，代入上式，$A^-+(G_0-A^-)-A^-A(G_0-A^-)AA^-=G_0$ 这说明只要适当选择 V 便可得到 $A\{1\}$ 中任一个元。

4. A^- 在解相容线性方程的应用

在开始讨论 g 逆时，可知相容线性方程组(6-2)的解，可写成 $x=A^-b$，然而这还不是解集。这小节主要是通过 g 逆，将 $Ax=0\cdots(E_0)$ 与 $Ax=b\cdots(E)$，这两个方程组的解集表示出来。

命题 10 方程(E_0)的通解可写成 $x=(I-A^-A)y$，$y\in C^n$ 的任意向量。

证 首先要证明上式表示的 x 是解，即

$$Ax=A\{(I-A^-A)y\}=(A-AA^-A)y=0。$$

其次，如果 x_0 是方程(E_0) 的任一解，则只要令 $y=x_0$ 代入表示式，便可得到 $x=x_0$。因此 $x=(I-A^-A)y$ 是方程(E_0) 解集的通式。 ▨

例 7 解齐次方程 $\begin{cases}x_1 \qquad\quad +3x_3=0\\ 2x_1+3x_2 \qquad =0\\ x_1+x_2+x_3=0\end{cases}$。

解 由于 $A=\begin{bmatrix}1&0&3\\2&3&0\\1&1&1\end{bmatrix}$，$r(A)=2$，$A$ 的左上角恰为一非零的

二阶子式 $A_{11}=\begin{bmatrix}1&0\\2&3\end{bmatrix}$，$A_{11}^{-1}=\begin{bmatrix}1&0\\-\frac{2}{3}&\frac{1}{3}\end{bmatrix}$，

所以
$$A^{-}=\begin{bmatrix}1&0&0\\-\frac{2}{3}&\frac{1}{3}&0\\0&0&0\end{bmatrix}。$$

因而齐次方程的通解可写成为：

$$x=(I-A^{-}A)y=\left(I-\begin{bmatrix}1&0&3\\0&1&-2\\0&0&0\end{bmatrix}\right)y=\begin{bmatrix}0&0&3\\0&0&2\\0&0&1\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$$
$$=y_3\begin{bmatrix}3\\2\\1\end{bmatrix}。$$

命题 11 方程(E)的通解可写成 $x=A^{-}b+(I-A^{-}A)y$。

证 上面结果，实际上是"非齐次方程的通解等于它的一个解与齐次方程通解之和"的表示方法。

例 8 解非齐次方程 $\begin{cases}x_1+3x_3=3\\2x_1+3x_2=0\\x_1+x_2+x_3=1\end{cases}$。

解 方程的系数与上题相同，$b=\begin{bmatrix}3\\0\\1\end{bmatrix}$，因而非齐次方程的通解便可写成为

$$x=A^{-}b+(I-A^{-}A)y=\begin{bmatrix}1&0&0\\-\frac{2}{3}&\frac{1}{3}&0\\0&0&0\end{bmatrix}\begin{bmatrix}3\\0\\1\end{bmatrix}+\begin{bmatrix}0&0&3\\0&0&2\\0&0&1\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$$

$$=\begin{bmatrix}3\\-2\\0\end{bmatrix}+y_3\begin{bmatrix}3\\2\\1\end{bmatrix}.$$

第三节　反射广义逆A_r^-

一个矩阵A的广义逆A^-，如果有$(A^-)^-=A$. 这便称它为反射逆。它可看为非奇异矩阵$(A^{-1})^{-1}=A$的推广 。反射逆的要求相当于矩阵G除了满足方程$AGA=A$外，还要满足$GAG=G$（这表示$G^-=A$即$(A^-)^-=A$）。

定义　设$A\in C^{m\times n}$。如果有一矩阵$G\in C^{n\times m}$并满足条件：

$$AGA=A,\ GAG=G。$$

则称G为A的反射广义逆，或反射g逆，记为A_r^-。A的反射广义逆全体构成的集合，记为$A\{1,2\}$（这表示满足 Penrose 四个矩阵方程的头两个方程的解集）。

下面的例说明反射g逆是存在的。

例 9　(i) 设$A^-\in A\{1\}$，则$A^-AA^-\in A\{1,2\}$。

(ii) 设$G_1,G_2\in A\{1\}$，则$G_1AG_2\in A\{1,2\}$。

证　(i) 要验证A^-AA^-满足方程$AGA=A,GAG=G$。则

$A(A^-AA^-)A=AA^-A=A$,

$(A^-AA^-)A(A^-AA^-)=A^-(AA^-A)A^-AA^-=A^-AA^-$。

(ii) $A(G_1AG_2)A=A$，易知成立。故，

$$(G_1AG_2)A(G_1AG_2)=(G_1AG_2),$$

$$G_1(AG_2A)G_1AG_2=G_1AG_2。$$

显然$A\{1,2\}\subseteq A\{1\}$。下面的命题，给出$G\in A\{1\}$又$\in A\{1,2\}$的一个充要条件。

命题 12　设$G\in A\{1\}$,则$G\in A\{1,2\}\Longleftrightarrow r(A)=r(G)$ 。

证　前面在讲A^-计算时，可看到A^-的秩可取$r(A)$至 min(m,n)中的任何一个正整数。这命题又说明只要取秩为$r(A)$的

A^-便是反射 g 逆。

$\Longleftarrow$由 $AGA=A$，便可知 $r(G)\geqslant r(A)$。又由 $GAG=G$，便可知 $r(A)\geqslant r(G)$。综合便得 $r(A)=r(G)$。

$\Longrightarrow$因 $r(A)=r(G)$，又 $r(A)=r(AGA)\leqslant r(GA)\leqslant r(G)$，所以有 $r(GA)=r(G)$。又 $R(GA)\subseteq R(G)$，由此可知 $R(GA)=R(G)$。这说明，G的列向量可通过 GA 的列向量的线性组合而得，写成等式便是 $G=GAD$。因此 $GAG=GAGAD=GAD=G$。▨

例 10 设$A=\begin{bmatrix}1&2&0\\0&0&2\\2&4&0\end{bmatrix}$，求 A_r^-。（见前面例 4）

由前面已计算的结果得的结果便可知：

$$A=P^{-1}\begin{bmatrix}I_r&\\&0\end{bmatrix}Q^{-1}。$$

其中
$$P=\begin{bmatrix}1&0&0\\0&\frac{1}{2}&0\\-2&0&1\end{bmatrix},\ Q=\begin{bmatrix}1&0&-2\\0&0&1\\0&1&0\end{bmatrix}。$$

则
$$A^-=Q\begin{bmatrix}I_2&\\&0\end{bmatrix}^-P。$$

但$\begin{bmatrix}I_2&\\&0\end{bmatrix}^-=\begin{bmatrix}I_2&\\&C\end{bmatrix}$，只要取 $C=0$ 可计算出 A_r^-。因为这时 $r(A^-)=r(A)$，则

$$A^-=\begin{bmatrix}1&0&-2\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}\begin{bmatrix}1&0&0\\0&\frac{1}{2}&0\\-2&0&1\end{bmatrix}$$
$$=\begin{bmatrix}1&0&-2\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}1&0&0\\0&\frac{1}{2}&0\\0&0&0\end{bmatrix}=\begin{bmatrix}1&0&0\\0&0&0\\0&\frac{1}{2}&0\end{bmatrix}=A_r^-。$$

还可直接验证它满足矩阵方程 $GAG=G$。

$$A_r^- A A_r^- = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \\ 2 & 4 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}。$$

第四节　最小范数广义逆 A_m^- 及相容线性方程的最小范数解

这节主要目的是将相容方程(E),在解集为无穷解时,要求它的最小范数的解。如果用 S 表示方程(E)的解集:

$$S=\{x \mid Ax=b, x\in C^n\}。$$

则最小范数解 x^* 为: $x^*\in S$, 且 $\|x^*\| \leqslant \|x\|$, $\forall x\in S$。
这里的范数是用内积定义的向量范数。

1. 最小范数解 x^0 所在的集合

首先要找出最小范数解 x^0 所在的集合。下面将 A 看成 C^n 至 C^m 的映射,则容易找出问题的答案。

引理　设 $A\in C^{m\times n}$。如果将 A 限制在 $R(A^*)$, 则 A 是 $R(A^*)$ 至 $R(A)$ 的一个同构。

证　因为将矩阵 A 看为 C^n 至 C^m 的一个线性映射。为了证

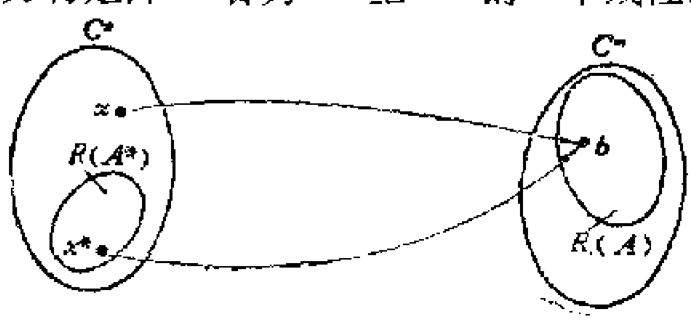

图 6-2

明命题成立，须要证明二点，(i) 当 A 限制在 $R(A^*)$ 时，它是 $R(A^*)$ 至 $R(A)$ 的满映射。(ii) 是单映射。

(i) 的证明：首先对任何一个 $b\in R(A)$，总有一个 $x\in C^n$，使得 $Ax=b$ 成立。又因 $C^n=R(A^*)\oplus N(A)$，因此对任何一个 x 都有分解式 $x=x^*+x^0$，$x^*\in R(A^*)$，$x^0\in N(A)$。所以有 $Ax^*=Ax=b$。它表明对任何一个 $b\in R(A)$，都可在 $R(A^*)$ 中找到它的像源。即 A 是 $R(A^*)$ 至 $R(A)$ 的满映射。

(ii) 的证明：主要证明对任何一个 $b\in R(A)$，在 $R(A^*)$ 中像源只有一个。如果有 $x^*, x^{*\prime}\in R(A^*)$，且 $Ax^*=Ax^{*\prime}$。要证明 $x^*=x^{*\prime}$。事实上 $x^*-x^{*\prime}\in R(A^*)$，又 $A(x^*-x^{*\prime})=0$。因而有 $x^*-x^{*\prime}\in N(A)$，又 $R(A^*)$ 与 $N(A)$ 是直和，所以 $x^*-x^{*\prime}=0$，即 $x^*=x^{*\prime}$。

综上所述 A 是线性变换，且它是 $R(A^*)$ 至 $R(A)$ 的一个双映射(即既是单映射又是满映射)，因此是 $R(A^*)$ 至 $R(A)$ 的一个同构。 ▨

有上述引理容易得出结论：相容方程 $Ax=b$ 的最小范数解 $x^*\in R(A^*)$。这是因为任何一个解 x，可根据分解式 $x=x^*+x^0$ 找到 $Ax^*=Ax=b$，且知道这样的 x^* 只有一个。因此 $\|x\|^2=\|x^*\|^2+\|x^0\|^2$(因为这是正交分解)，$\|x\|\geqslant\|x^*\|$。

2. 下面的定理给出 $x=Gb$ 是相容方程的最小范数解，矩阵 G 应满足的条件

定理 2 设 $Ax=b$ 是相容方程，$x=Gb$ 是最小范数解

$$\Longleftrightarrow\quad AGA=A,\ (GA)^*=GA。$$

证 $\Longleftarrow$ 因为 $AGA=A$，所以 $x=Gb$ 是解。余下只要证明 $Gb\in R(A^*)$。由前面的讨论可知，它是最小范数解。但 $R(A^*)$ 与 $N(A)$ 是正交补空间，因而只要证明 $Gb\in N(A)^\perp$，或对任一个 $u\in N(A)$，有 $(Gb,u)=0$。因 $b\in R(A)$，所以有 $y\in C^n$ 且 $Ay=b$。而

$$(Gb,u)=u^*Gb=u^*GAy=u^*(GA)^*y=u^*A^*G^*y$$

$$=(Au)^*G^*y=0。$$

因此 $Gb\in N(A)^{\perp}=R(A^*)$， $x=Gb$ 是最小范数解。

$\Longrightarrow$因 $x=Gb$ 是最小范数解。因为是解， 所以 $AGA=A$。其次由充分性证明可知，有一个 G_0 满足 $AG_0A=A,(G_0A)^*=G_0A$，使得 $Gb=G_0b$（关于 G_0 存在性， 下面通过构造的方法作出）。由于 $A^{(i)}$（A 的第 i 列）$\in R(A)$。自然有 $GA^{(i)}=G_0A^{(i)}$ $(i=1,2,\cdots,n)$， 即 $GA=G_0A$。 这样 $(GA)^*=(G_0A)^*=G_0A=GA$。 ▨

综上所述，要 $x=Gb$ 是相容方程的最小范数解，则 G 一定要满足 Penrose 的第一、四两个矩阵方程 。将同时满足这两个矩阵方程的解集记为 $A\{1,4\}$。

定义 设 $A\in C^{m\times n}$， $G\in C^{n\times m}$。如果有 $AGA=G,(GA)^*=GA$，则称 G 是 A 的最小范数广义逆，记号为 A_m^-。

下面用构造的方法说明集合 $A\{1,4\}$是非空的。

命题 13 设 $G=A^*(AA^*)^-$，则 $G\in A\{1,4\}$。

证 先证 $AGA=A$ 或 $AGA-A=0$，只要证明

$$(AGA-A)(AGA-A)^*=0。则$$

$$\begin{aligned}
&(AA^*(AA^*)^-A-A)(AA^*(AA^*)^-A-A)^*\\
&=(AA^*(AA^*)^-A-A)(A^*(AA^*)^-AA^*-A^*)\\
&=(AA^*(AA^*)^-A-A)A^*((AA^*)^-AA^*-I)\\
&=(AA^*(AA^*)^-AA^*-AA^*)((AA^*)^-AA^*-I)\\
&=0。
\end{aligned}$$

然后证 $(GA)^*=GA$，

因为 $(A^*(AA^*)^-A)^*=A^*(AA^*)^-A$。 ▨

下面的命题，对于验证一个矩阵 G 是否属于 $A\{1,4\}$ 会有帮助。

命题 14 若 $G\in A\{1\}$，则它是 $A\{1,4\}$的元$\Longleftrightarrow GA=A_m^-A$（$A_m^-$ 是 $A\{1,4\}$ 中任一个元）。

证 $\Longrightarrow GA=GAA_m^-A=(GAA_m^-A)^*=(A_m^-A)^*(GA)^*$

$= A_m^- AGA = A_m^- A$。

$\Longleftarrow$ 显然。 ▨

命题 15 $AGA = A$, $(GA)^* = GA$, $\Longleftrightarrow GAA^* = A^*$。

证 $\Longrightarrow GAA^* = (GA)^* A^* = (AGA)^* = A^*$。

$\Longleftarrow$ 因为 $GAA^* = A^* \Rightarrow GAA^* G^* = A^* G^*$, 因而 $(GA)(GA)^* = (GA)^*$。由 GA 的这个关系式可知它满足 $(GA)^* = GA$,

又 $\quad AGA = A(GA)^* = (GAA^*)^* = (A^*)^* = A$。 ▨

3．对于一个给定的矩阵A,它的最小范数广义逆不是唯一的

下面给出集合 $A\{1,4\}$的通式。

命题 16 设 $A \in C^{m\times n}$, A_m^- 是 $A\{1,4\}$的任一个元。则最小范数广义逆的通式为：

$$G = A_m^- + V(I - AA_m^-)$$

其中 V 是任一 $n \times m$ 矩阵。

证 先要证明 $G \in A\{1,4\}$。由命题 14,只要证明有 $GA = A_m^- A$。$GA = \{A_m^- + V(I - AA_m^-)\}A = A_m^- A + V(A - AA_m^- A) = A_m^- A$。

其次要证明对任一个 $G_0 \in A\{1,4\}$, 都可通过适当选择 V 而得。事实上只要选 $V = G_0 - A_m^-$, 即可得出 $G = G_0$ ▨

例 11 求线性方程组

$$\begin{cases} x_1 + 2x_2 - x_3 = 1 \\ -x_2 + 2x_3 = 2 \end{cases}$$ 的最小范数解

解 由 $A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ (相容性判定略)。

易知矩阵 A 是行满秩,因而 $A_m^- = A_R^{-1} = A^*(AA^*)^{-1}$, 即

$$A_m^- = \frac{1}{14}\begin{bmatrix} 5 & 4 \\ 6 & 2 \\ 3 & 8 \end{bmatrix}$$

最小范数解为

$$x^* = A_m^- b = \frac{1}{14}\begin{bmatrix}5 & 4\\ 6 & 2\\ 3 & 8\end{bmatrix}\begin{bmatrix}1\\ 2\end{bmatrix} = \begin{pmatrix}\frac{13}{14}\\ \frac{10}{14}\\ \frac{19}{14}\end{pmatrix},$$

这时 $\|x^*\| = \frac{1}{14}\sqrt{630}$。

实际上给出方程组的通解为 $x = \begin{bmatrix}5\\ -2\\ 0\end{bmatrix} + k\begin{bmatrix}-3\\ 2\\ 1\end{bmatrix}$，容易验证 x^* 确实是最小范数解。

第五节　最小二乘广义逆 A_L^- 及矛盾方程的最小二乘解

在第三章的第四节是通过解正规方程 $A^*Ax = A^*b$ 来求矛盾方程的最小二乘解(即最优近似解)。前面的计算实际上是假定了 A^*A 是可逆的。如果这个条件不满足，将它转化为相类似的情形(见第三章例 11)。这里重新研究这个问题的目的是：希望将最小二乘解表示为 $x = Gb$ 的形式，且使它仅和矩阵 A 有关。

1．矛盾方程的最佳近似解，以及求解的思想

矛盾方程是 $Ax = b$，$b \in R(A)$。现要求最优近似解 x^* 应满足条件：

$$\|Ax_0 - b\| \leqslant \|Ax - b\|, \quad \forall x \in C^n。$$

向量 Ax 总是在子空间 $R(A)$ 中。因而它又等价于在子空间 $R(A)$ 内找一向量 b^* 满足不等式：

$$\|b^* - b\| \leqslant \|y - b\|, \quad \forall y \in R(A)。$$

这样 b^* 即为 b 在子空间 $R(A)$ 的最佳近似向量。因此解上述矛

盾方程的最佳近似解是(i) 先找出 b 在 $R(A)$ 的正交投影 b^*；(ii) 再解方程 $Ax=b^*$。

命题 17 设 A 是空间 C^m 的投影算子(即满足条件 $A^2=A$)，它是正交投影 $\Longleftrightarrow A^*=A$。

证 因为 A 是空间 C^m 的投影算子，所以 C^m 可分解为 A 的秩空间与核的直和：$C^m=R(A)\oplus N(A)$。所谓正交投影是指 $R(A)$ 与 $N(A)$ 正交。

$\Longrightarrow$ 如果 $R(A)\perp N(A)$，要证明 $A^*=A$。事实上任取 x，$y\in C^m$，作分解 $x=\alpha+\beta$，$y=\alpha'+\beta'$，$\alpha,\alpha'\in R(A)$，β，$\beta'\in N(A)$，则

$$
\begin{aligned}
(Ax,y)&=(A(\alpha+\beta),\alpha'+\beta')=(A\alpha+A\beta,\alpha'+\beta')\\
&=(\alpha,\alpha'+\beta')。
\end{aligned}
$$

(因 $\alpha\in R(A)$，所以有 $A\alpha=\alpha$，又 $\beta\in N(A)$，所以 $A\beta=0$)。

又因 $(\alpha,\alpha'+\beta')=(\alpha,\alpha')+(\alpha,\beta')$，前题是 $R(A)\perp N(A)$，因此

$(\alpha,\alpha'+\beta')=(\alpha,\alpha')+(\beta,\alpha')=(\alpha+\beta,\alpha')=(\alpha+\beta,A(\alpha'+\beta'))$。由计算得出 $(Ax,y)=(x,Ay)$，$\forall x,y\in C^m$。

写成矩阵的乘积，便是

$$y^*Ax=y^*A^*x\quad\forall x,y\in C^m。$$

因此 $A=A^*$

$\Longleftarrow$ 如果 $A^*=A$，要证明 $R(A)\perp N(A)$。即 $(\alpha,\beta)=0$，$\forall\alpha\in R(A)$，$\beta\in N(A)$ 成立。因为 $\alpha\in R(A)$，则 $\alpha=Ax$，所以

$(\alpha,\beta)=(Ax,\beta)=\beta^*Ax=\beta^*A^*x=(A\beta)^*x=0$(因 $\beta\in N(A)$)。 ▨

2. $x=Gb$ 是矛盾方程的最佳近似解，矩阵 G 所应满足的条件

定理 3 对任意 b，$x=Gb$ 是矛盾方程的最小二乘解

$$\Longleftrightarrow AGA=A，\ (AG)^*=AG。$$

证 $\Longleftarrow$ 因 G 满足方程 $AGA=A$，所以 $(AG)^2=AG$。这说明矩阵 AG 是 C^m 的投影算子。其次 $(AG)^*=AG$，根据命题

17 可知，AG 是正交投影。又 $R(AG)=R(A)$，因此 AG 是 C^m 在 $R(A)$ 的正交投影。所以 b 在 $R(A)$ 的最佳近似向量为 AGb，即

$$\|AGb-b\|\leqslant\|Ax-b\|。$$

$x=Gb$ 是矛盾方程的最小二乘解。

$\Longrightarrow$ 对差 $Ax-b$ 作分解。设 P 是 C^m 在 $R(A)$ 的正交投影，则

$$Ax-b=Ax-Pb+Pb-b。$$

(由 P 是正交投影，可知 $Pb-b\perp R(A)$。又 $Ax-Pb\in R(A)$，因此 $(Ax-Pb)\perp(Pb-b)$)。

$$\|Ax-b\|^2=\|Ax-Pb\|^2+\|Pb-b\|^2\geqslant\|Pb-b\|^2。$$

上式说明 $\min\|Ax-b\|=\|Pb-b\|$。现对任何 b，$x=Gb$ 给出方程的最优解，因此要求 $AGb=Pb$，$\forall b\in C^m$，所以 AG 是 C^m 在 $R(A)$ 的正交投影，故 $AGA=A$，$(AG)^*=AG$。 ▨

定义 设 $A\in C^{m\times n}$，$G\in C^{n\times m}$。如果 G 满足条件

$$AGA=A,\quad (AG)^*=AG。$$

则称 G 是 A 的最小二乘广义逆，记为 A_L^-。

A 的最小二乘广义逆的全体构成的集合，记为 $A\{1,3\}$。下面的命题说明 $A\{1,3\}$ 不是空集。

命题 18 若 $G=(A^*A)^-A^*$，则 $G\in A\{1,3\}$。

证 仿前面命题 B 可证。

命题 19 设 $G\in A\{1\}$，则它是 $A\{1,4\}$ 的元 $\Longleftrightarrow$ $AG=AA_L^-$。

证 仿前面命题 14 可证。

命题 20 $AGA=A$，$(AG)^*=AG\Longleftrightarrow A^*AG=A^*$

证 仿前面命题 15 可证。

命题 21 若 $A\in C^{m\times n}$，A_L^- 是 $A\{1,3\}$ 的任一个元，则最小二乘广义逆的通式为：

$$G=A_L^-+(I-A_L^-A)U。$$

证 仿前面命题 16 可证。

第六节 Moore-Penrose 广义逆

1. Moore-Penrose 广义逆的定义及性质

定义 设 $A\in C^{m\times n}$,如果存在一 $n\times m$ 矩阵 G ,满足Penrose 的四个矩阵方程:

(i) $AGA=A$, (ii) $GAG=G$, (iii) $(AG)^*=AG$, (iv) $(GA)^*=GA$。则称 G 是 A 的 Moore-Penrose 广义逆，或称为 M-P 逆,记为 A^+。

由前几节的讨论很容易知道 A^+ 的性质:

(i) $(A^+)^*=(A^*)^+$。

(ii) $(A^+)^+=A$。

(iii) $(AA^*)^+=(A^+)^*A^+$。

A^+ 在解线性方程的应用：$Ax=b$ 是相容时，$x=A^+b$ 是方程的一个解，且是最小范数解。如果 $Ax=b$ 是矛盾方程，则 $x=A^+b$ 是最小二乘解,且是最小范数解,在本章末会证明这点。首先讨论 A^+ 的存在性与唯一性。

定理 4 对于任意矩阵 A, A^+ 必存在,

(i) 如果 A 是非奇异矩阵,则 $A^+=A^{-1}$。

(ii) 如果 A 是行满秩,则 $A^+=A_R^{-1}=A^*(AA^*)^{-1}$。

(iii) 如果 A 是列满秩,则 $A^+=A_L^{-1}=(A^*A)^{-1}A^*$。

(iv) 如果 $r(A)=r<\min\{m,n\}$则作矩阵 A 的满秩分解后,取 $A^+=D_R^{-1}C_L^{-1}=D^*(DD^*)^{-1}(C^*C)^{-1}C^*$。

证 只要直接按 Penrose 的四个矩阵方程去验证，定理可成立。 ▨

命题 22 若 A_m^-, A_L^- 分别为 A 的最小范数逆与最小二乘逆,则 $A^+=A_m^-AA_L^-$

证 见本章练习 16。

命题 23 A^+ 是唯一的。

证 如果 G_1, G_2 是 A 的二个 A^+，则

$G_2 = G_2AG_2 = G_2AG_1AG_2 = G_2(AG_1)^*(AG_2)^* = G_2(AG_2AG_1)^* = G_2(AG_1)^* = G_2AG_1 = G_2AG_1AG_1 = (G_2A)^*(G_1A)^*G_1 = (G_1AG_2A)^*G_1 = (G_1A)^*G_1 = G_1AG_1 = G_1$。

2. A^+ 广义逆的计算

(1) 对 A 作满秩分解，求 A^+

例 12 已知 $A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix}$，求 A^+。

解 因为 $r(A) = 1, A = CD = \begin{bmatrix} 1 \\ 2 \end{bmatrix}(1,1,2)$。

$$C_L^{-1} = (C^*C)^{-1}C^* = \left\{(1,2)\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}^{-1}(1,\ 2) = \frac{1}{5}(1,2),$$

$$D_R^{-1} = D^*(DD^*)^{-1} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}\left\{(1\ 1\ 2)\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}\right\}^{-1} = \frac{1}{6}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}。$$

所以

$$A^+ = D_R^{-1}C_L^{-1} = \frac{1}{30}\begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 2 & 4 \end{bmatrix}。$$

例 13 已知 $A = \begin{bmatrix} 1 & 2 & 1 & 4 & 1 \\ 2 & 4 & 0 & 6 & 6 \\ 1 & 2 & 0 & 3 & 3 \\ 2 & 4 & 0 & 6 & 6 \end{bmatrix}$，求 A^+。

解 要求 A 的满秩分解可通过 Hermite 标准型来进行，即

$$A = \begin{bmatrix} 1 & 2 & 1 & 4 & 1 \\ 2 & 4 & 0 & 6 & 6 \\ 1 & 2 & 0 & 3 & 3 \\ 2 & 4 & 0 & 6 & 6 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 0 & 3 & 3 \\ 0 & 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}。$$

作出 A 的满秩分解为：

$$A=\begin{bmatrix}1 & 1\\ 2 & 0\\ 1 & 0\\ 2 & 0\end{bmatrix}\begin{bmatrix}1 & 2 & 0 & 3 & 3\\ 0 & 0 & 1 & 1 & -2\end{bmatrix}=CD,$$

所以

$$A^{-1}=D_R^{-1}C_L^{-1}=D^*(DD^*)^{-1}(C^*C)^{-1}C^*$$

$$=\begin{bmatrix}1 & 0\\ 2 & 0\\ 0 & 1\\ 3 & 1\\ 3 & -2\end{bmatrix},\ \frac{1}{129}\begin{bmatrix}6 & 3\\ 3 & 23\end{bmatrix}\begin{bmatrix}1 & 2 & 1 & 2\\ 1 & 0 & 0 & 0\end{bmatrix},\frac{1}{9}\begin{bmatrix}1 & -1\\ -1 & 10\end{bmatrix}$$

$$=\frac{1}{1161}\begin{pmatrix}27 & 6 & 3 & 6\\ 54 & 12 & 6 & 12\\ 207 & -40 & 20 & -40\\ 288 & -22 & -11 & -22\\ -333 & 98 & 49 & 98\end{pmatrix},$$

例 14 已知 $R=\begin{bmatrix}R_{11} & 0\\ 0 & 0\end{bmatrix}_{m\times n}$，求 $R^+(|R_{11}|\neq 0)$。

解 $R^+=\begin{bmatrix}R_{11}^{-1} & 0\\ 0 & 0\end{bmatrix}$，只要直接验证即可。

由上面的例自然会提出如下的问题：一个矩阵 A 通过初等变换变成上面形式的矩阵，$PAQ=\begin{bmatrix}A_{11} & 0\\ 0 & 0\end{bmatrix}=R$，这里 $|A_{11}|\neq 0$，是否仍有 $A^+=QR^+P$？（在本章第二节里证明了 $A^-=QR^-P$ 是成立的）。可以直接验证 Penrose 的第一个方程与第二个方程是满足的。现考察 $(AG)^*=AG$ 是否成立。因为 $AG=P^{-1}RQ^{-1}$，$QR^+P=P^{-1}RR^+P$ 两边取共轭转置，便有

$$(AG)^* = P^*(RR^+)^*(P^{-1})^* = P^*(RR^+)^*(P^*)^{-1}。$$

比较两式可知，如果有 $P^* = P^{-1}$，则$(AG)^* = AG$ 可成立。同理，如果 $Q^* = Q^{-1}$，则 $(GA)^* = GA$ 也可成立。换句话说，在求 A^+ 时，如果变换矩阵 P,Q 都是正交矩阵，下面的关系是成立的。

$$PAQ = \begin{bmatrix} A_{11} & D \\ 0 & 0 \end{bmatrix} = R \Rightarrow A^+ = QR^+P。$$

上面的结果表明可用矩阵的奇异值分解（第三章第八节）去求 A^+。

(2) 用奇异值分解求 A^+

命题 24 设 A 的奇异值分解为 $A = UDV^*$，其中

$$D = \begin{pmatrix} s_1 & & & & & \\ & \ddots & & & & \\ & & s_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

则
$$A^+ = VD^+U^*$$

证 前面的分析便可知命题成立。

例 15 设 $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$，用奇异值分解求 A^+。

解 A 的奇异值分解为：

$$A = UDV^* = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix},$$

所以

$$A^{+}=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\frac{1}{2} & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$=\frac{1}{4}\begin{bmatrix}1 & 1 & 0\\ 1 & 1 & 0\end{bmatrix}。$$

3. A^{+}与正交投影

这节主要是讨论 A^{+} 与正交投影的关系，以便完成 $x=A^{+}b$ 是矛盾方程 $Ax=b$ 的最小二乘解的最小范数解的证明。

定理 5 设 A 是 $m\times n$ 矩阵，G 是 $n\times m$ 矩阵，则 G 是 A^{+} $\Longleftrightarrow AG$ 是 C^{m} 在 $R(A)$ 的正交投影；GA 是 C^{n} 在 $R(G)$ 的正交投影。

证 $\Longleftarrow$显然。

$\Longrightarrow AGA=A$，说明 $(AG)A^{(j)}=A^{(j)}$，即 AG 在 $R(A)$上是恒等变换。另一方面 $C^{m}=R(A)\oplus N(A^{*})$，如果 $x\in N(A^{*})$，$AGx=(AG)^{*}x=G^{*}A^{*}x=0$，这便证明了 AG 是 C^{m} 在$R(A)$的正交投影。

$GAG=G$，说明 $(GA)^{2}=GA$，则 GA 是 C^{n} 上的投影。又 $C^{n}=R(A^{*})\oplus N(A)$ 对于任一个 $x\in R(A^{*})$有 $x=A^{*}y$，$GAx=(GA)^{*}A^{*}y=(AGA)^{*}y=A^{*}y=x$，故 GA 在 $R(A^{*})$上是个恒等变换。又 $x\in N(A)$ 有 $GAx=0$，GA 在 $N(A)$ 上是零变换。综合上面证明说明 GA 是 C^{n} 在 $R(A^{*})$ 上的正交投影。此时，$R(GA)=R(A^{*})$，又 $R(GA)=R(G)$，因此定理成立。▨

命题 25 若 $Ax=b$ 是矛盾方程，则 $x=A^{+}b$ 是最小二乘解的最小范数解。

证 $x=A^{+}b$ 是最小二乘解。这是因为 AA^{+} 是 C^{m} 在 $R(A)$ 的正交投影，所以 $AA^{+}b$ 是 b 在 $R(A)$ 的最佳近似。要证明 $x=A^{+}b$ 是最小范数，只要说明 $A^{+}b\in R(A^{*})$。因为$A^{+}b=A^{+}AA^{+}b=$

$(A^{+}A)A^{+}b$，说明 $A^{+}b\in R(A^{+}A)$。由前一定理可知$R(A^{+}A)=R(A^{*})$，即 $A^{+}b\in R(A^{*})$。因而 $x=A^{+}b$ 是最小范数。 ∎

习　题

第一章　基础知识

A 类

1．以下 $AB=C$，试写出矩阵 C 的行向量与列向量，通过它的因子的行向量与列向量的线性组合表达式。

(i)　$A=(1\ 1\ 1\ 1)$，$B=\begin{bmatrix}5 & 2 & -3 & -3\\ -7 & -2 & 4 & 2\\ -1 & 2 & 1 & 1\\ -2 & -2 & -3 & 4\end{bmatrix}$

(ii)　$A=\begin{bmatrix}83 & -29 & -52 & 46\\ -15 & 97 & 78 & -112\\ 38 & -4 & 69 & 85\end{bmatrix}$，$B=\begin{bmatrix}1\\ 0\\ 1\\ 0\end{bmatrix}$；

(iii)　$A=\begin{bmatrix}-2 & 3 & 0 & 1\\ 1 & 1 & 2 & -1\end{bmatrix}$，$B=\begin{bmatrix}2 & 0\\ 1 & -1\\ -1 & 2\\ 1 & 3\end{bmatrix}$。

2．已知

$$A=\begin{bmatrix}1 & 2 & -1 & 2 & 1\\ 2 & 4 & 1 & -2 & 3\\ 3 & 6 & -3 & 6 & 3\end{bmatrix}，$$

求：(i) 矩阵 A 的列向量的最大无关组，以及其余的列向量用这最大无关组的表示式；

(ii) 矩阵 A 的满秩分解；

(iii) 矩阵 A 的行空间的基；

(iv) 构造以 A 的行空间为解集的线性齐次方程组。

3. 求一个五维向量 (x_1,x_2,x_3,x_4,x_5) 是下面向量线性组合的条件　设

$\alpha_1=(1\ \ 0\ 2\ 1\ -1)$，　$\alpha_2=(-1\ 2\ -4\ 2\ 0)$，

$\alpha_3=(2\ -1\ 5\ 2\ \ 1)$，　$\alpha_4=(\ 2\ 1\ \ 3\ 5\ 2)$。

4. 设 V 是下面矩阵 A 的行空间，

$$A=\begin{bmatrix}3 & 21 & 0 & 9 & 0\\ 1 & 7 & -1 & -2 & -1\\ 2 & 14 & 0 & 6 & 1\\ 6 & 42 & -1 & 1.3 & 0\end{bmatrix}。$$

(i) 求 V 的基；

(ii) 写出 $(x_1,\cdots,x_5)\in V$ 的条件；

(iii) 如果 $(x_1,\cdots,x_5)\in V$，写出它关于(a)的基的坐标。

5. 问 $[\alpha_1,\alpha_2,\alpha_3]$ 与 $[\beta_1\beta_2]$ 是否相同。设

$\alpha_1=(1\ 2\ -1\ 3)$，　$\alpha_2=(2\ 4\ 1\ -2)$，　$\alpha_3=(3\ 6\ 3\ -7)$，

$\beta_1=(1\ 2\ -4\ 11)$，　$\beta_2=(2\ 4\ -5\ 14)$。

6. 设 $V_1=[(1\ 2\ 3\ 6),\ (4\ -1\ 3\ 6),\ (5\ 1\ 6\ 12)]$，$V_2=[(1\ -1\ 1\ 1),(2\ -1\ 4\ 5)]$。它们都是 R_4 的空间，

(i) 求 $V_1\cap V_2$ 的基；

(ii) 扩充 $V_1\cap V_2$ 的基使它成为 V_1 的基；

(iii) 扩充 $V_1\cap V_2$ 的基使它成为 V_2 的基；

(iv) 求 V_1+V_2 的基。

7. 设 $V_1=\{(x_1,x_2,x_3,x_4)|x_1+x_2+x_3+x_4=0\}$，$V_2=\{(x_1,x_2,x_3,x_4)|x_1-x_2+x_3-x_4=0\}$。求 $V_1\cap V_2$，V_1+V_2 的维数与基。

8. 设在 R_4 中的向量 $\alpha_1=(1\ 1\ 0\ 0)$，$\alpha_2=(0\ 1\ 1\ 0)$，$\beta_1=(0\ 0\ -1\ -1)$，$\beta_2=(0\ -1\ 0\ 1)$。　令 $V_1=[\alpha_1,\alpha_2]$，$V_2=[\beta_1,\beta_2]$，试问 $W=V_1+V_2$ 是否是直和。

9. 设 $\sigma\in \mathrm{Hom}(U,V)$，证明 $N(\sigma)$ 是 U 的子空间，$R(\sigma)$ 是 V 的子空间。

10. 设 $\sigma\in \mathrm{Hom}(R^3,R^3)$，$\sigma((x_1,x_2,x_3))=(x_1+2x_2-x_3,x_2+x_3,x_1+x_2-2x_3)$，求 σ 的核与像空间的基与维数。

11. 设 $\sigma\in \mathrm{Hom}(R^4,R^3)$，　$\sigma((x_1,x_2,x_3,x_4))=(2x_1+x_2-x_4,\ x_1+x_2-x_3,2x_1+2x_2-x_3-x_4)$。求 σ 的核与像空间的基与维数。

12. 设 ξ_0 是$\sigma(\xi)=\beta$ 的一个解，证明此方程的解的全体为$\xi_0+N(\sigma)$。

13. 设 $\sigma\in \mathrm{Hom}(R^2,R^2)$，它定义为：

$$\sigma((3,1))=(2,-4),\quad \sigma((1,1))=(0,2)$$

求 $\sigma((x_1,x_2)),\sigma((7,4))$。

14. 求下列各线性变换 σ 在 R^n 的自然基下的矩阵。

(i) 设 $\sigma\in \mathrm{Hom}(R^3,R^2)$，有

$$\sigma((x_1,x_2,x_3))=(2x_1-4x_2+9x_3,5x_1+3x_2-2x_3);$$

(ii) 设 $\sigma\in \mathrm{Hom}(R^2,R^4)$，有

$$\sigma((x_1,x_2))=(3x_1+4x_2,5x_1-2x_2,x_1+7x_2,4x_1);$$

(iii) 设 $\sigma\in \mathrm{Hom}(R^4,R)$，有

$$\sigma((x_1,x_2,x_3,x_4))=2x_1+3x_2-7x_3-x_4。$$

15. 设 R^2 的两组基为 $\alpha_1=(1,0)$，$\alpha_2=(0,1)$，与 $\beta_1=(1,1)$，$\beta_2=(-1,0)$，求 R^2 中的向量在上述两组基下的坐标变换公式。

16. 设 $\sigma\in \mathrm{Hom}(R^3,R^3)$，有

$$\sigma((x_1,x_2,x_3))=(2x_2+x_3,x_1-4x_2,3x_1)。$$

求 (i) σ 在自然基下的矩阵；

(ii) 在基 $\alpha_1'=(1,1,1)$，$\alpha_2'=(0,1,1)$，$\alpha_3'=(0,0,1)$ 下的矩阵。

17. 设 $\sigma\in \mathrm{Hom}(R^3,R^2)$，有

$$\sigma((x_1,x_2,x_3))=(2x_1+x_2-x_3,2x_1-2x_2+4x_3)。$$

又 R^3，R^2 的基分别为 $\alpha_1=(1,1,1)$，$\alpha_2=(1,1,0)$，$\alpha_3=(1,0,0)$和 $\beta_1=(1,3)$，$\beta_2=(1,4)$。求 σ 在这二组基下的矩阵。

18. 设 $\sigma\in \mathrm{Hom}(R^3,R^3)$，有

$$\sigma((x_1,x_2,x_3))=(3x_1+x_3,-2x_1+x_2,-x_1+2x_2+4x_3)。$$

(i) 求 σ 在 R^3 的标准基下的矩阵；

(ii) 设 R^3 另一组基为 $\alpha_1,\alpha_2,\alpha_3$，它们分别为：

$$\alpha_1=(1,0,1),\quad \alpha_2=(-1,2,1),\quad \alpha_3=(2,1,1)。$$

试求在 $\alpha_1,\alpha_2,\alpha_3$ 下 σ 的矩阵，

(iii) 试证明 σ 是可逆的，并给出 σ 逆的公式。

19. 设 $\sigma,\tau\in \mathrm{Hom}(R^2,R^2)$，它们分别定义为：

$$\sigma((x_1,x_2))=(x_1+x_2,0),$$

$$\tau(x_1,x_2)=(-x_2,x_1)。$$

求变换 $\sigma+\tau,\sigma\tau,\tau\sigma,\sigma^2,\tau^2$。

20. 求线性变换 $\sigma\in \mathrm{Hom}(R^4,R^3)$，它的核由 (1,2,3,4)和(0,1,1,1)

生成。

B 类

1. 如果 A 满足条件 $A^2+2A+I=0$，证明 A 是非奇异的，且 $A^{-1}=-(A+2I)$。

2. 设 D 是对角阵且是非奇异的，证明如果 $D=(I+A)^{-1}A$，则 A 是对角阵。

3. I,A,B 都是 n 阶方阵，证明如果 $I+AB$ 是可逆，则 $I+BA$ 是可逆的，且

$$(I+BA)^{-1}=I-B(I+AB)^{-1}A。$$

4. 证明 Sherman-Morrison 公式

$$(A+uv^T)^{-1}=A^{-1}-\frac{(A^{-1}u)(v^TA^{-1})}{1+v^TA^{-1}u}。$$

5. 设 A 是 $n\times r$ 型矩阵，$r(A)=r$，B 是 r 阶方阵，证明

(i) 若 $AB=O$，则 $B=O$，

(ii) 若 $AB=A$，则 $B=E$，

6. 设 A 是 n 阶方阵，$A-I$ 是可逆，且 $f(A)=(I-A)(I+A)^{-1}$，证明

(i) $(I+f(A))(I+A)=2I$；

(ii) $f[f(A)]=A$。

7. 设 V 是全体 n 阶上三角实阵，p 是 $1\leqslant p\leqslant n$ 的一个固定的数，证明 $\sigma: V\to R$，$\sigma(A)=a_{pp}$，$A=[a_{ij}]$，$A\in V$，是线性变换，且满足条件 $\sigma(AB)=\sigma(A)\sigma(B)$。

8. 设 W 是 $R^{n\times n}$ 的一个子空间，它满足如下条件：

$$AX,XA\in W \qquad \forall X\in W, A\in V。$$

证明 $W=\{0\}$ 或 $R^{n\times n}$。

(*)下面的练习是关于 Cauchy-Binet 定理的。Cauchy-Binet 定理是矩阵理论的重要公式。由它可以得到很多的结论。

命题 (Cauchy-Binet) 设 $A_{m\times n}$，$B_{n\times m}(m\leqslant n)C=AB$，则它的行列式

$$\det C=\sum_{1\leqslant k_1<k_2\cdots<k_m\leqslant n}A\begin{pmatrix}1 & 2\cdots m\\ k_1 & k_2\cdots k_m\end{pmatrix}B\begin{pmatrix}k_1 & k_2\cdots k_m\\ 1 & 2\cdots m\end{pmatrix}$$

这里求和是对集合$\{1,2,\cdots,n\}$里的m个元构成集求和，而$A\begin{pmatrix}1 & 2 & \cdots & m\\ k_1 & k_2 & \cdots & k_m\end{pmatrix}$表示$A$中第$1,2,\cdots,m$与$k_1,k_2,\cdots,k_m$列所构成的$m$阶子式，$B\begin{pmatrix}k_1 & \cdots & k_m\\ 1 & \cdots & m\end{pmatrix}$表示$B$中第$k_1,k_2,\cdots,k_m$行与第$1,2,\cdots,m$列所构成的$m$阶子式。

例 $\underset{A}{\begin{bmatrix}1 & 2 & -1\\ 0 & 1 & -1\end{bmatrix}}\underset{B}{\begin{bmatrix}1 & -1\\ 0 & 1\\ -1 & 0\end{bmatrix}}$，它的行列式应为

$$\det C=A\begin{pmatrix}1 & 2\\ 1 & 2\end{pmatrix}B\begin{pmatrix}1 & 2\\ 1 & 2\end{pmatrix}+A\begin{pmatrix}1 & 2\\ 1 & 3\end{pmatrix}B\begin{pmatrix}1 & 3\\ 1 & 2\end{pmatrix}$$

$$+A\begin{pmatrix}1 & 2\\ 2 & 3\end{pmatrix}B\begin{pmatrix}2 & 3\\ 1 & 2\end{pmatrix}$$

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

$$=1-1-1=-1。$$

Cauchy-Binet 定理的证明。

由$AB=C$，可知

$$\det C=\begin{vmatrix}A_{(1)}B^{(1)} & A_{(1)}B^{(2)} & \cdots & A_{(1)}B^{(m)}\\ A_{(2)}B^{(1)} & A_{(2)}B^{(2)} & \cdots & A_{(2)}B^{(m)}\\ \vdots & & & \\ A_{(m)}B^{(1)} & A_{(m)}B^{(2)} & \cdots & A_{(m)}B^{(m)}\end{vmatrix}$$

$$=\begin{vmatrix}\sum\limits_{k_1=1}^{n}a_{1k_1}b_{k_11} & A_{(1)}B^{(2)} & \cdots & A_{(1)}B^{(m)}\\ \sum\limits_{k_1=1}^{n}a_{2k_1}b_{k_11} & A_{(2)}B^{(2)} & \cdots & A_{(2)}B^{(m)}\\ \vdots & \vdots & & \\ \sum\limits_{k_1=1}^{n}a_{mk_1}b_{k_11} & A_{(m)}B^{(2)} & \cdots & A_{(m)}B^{(m)}\end{vmatrix},$$

由行列式的性质可有

$$\det C=\sum_{k_1=1}^{n}\begin{vmatrix}a_{1k_1}b_{k_11} & A_{(1)}B^{(2)} & \cdots & A_{(1)}B^{(m)}\\ a_{2k_1}b_{k_11} & A_{(2)}B^{(2)} & \cdots & A_{(2)}B^{(m)}\\ \vdots & \vdots & & \\ a_{mk_1}b_{k_11} & A_{(m)}B^{(2)} & \cdots & A_{(m)}B^{(m)}\end{vmatrix},$$

用同样的方法可按各列展开为下面的和式，

$$\det C=\sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\cdots\sum_{k_m=1}^{n}\begin{vmatrix} a_{1k_1}b_{k_11} & a_{1k_2}b_{k_22} & \cdots & a_{1k_m}b_{k_mm} \\ a_{2k_1}b_{k_11} & a_{2k_2}b_{k_22} & \cdots & a_{2k_m}b_{k_mm} \\ \vdots & \vdots & & \\ a_{mk_1}b_{k_11} & a_{mk_2}b_{k_22} & \cdots & a_{mk_m}b_{k_mm} \end{vmatrix}$$

$$=\sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\cdots\sum_{k_m=1}^{n}\begin{vmatrix} a_{1k_1} & a_{1k_2} & \cdots & a_{1k_m} \\ a_{2k_1} & a_{2k_2} & \cdots & a_{2k_m} \\ \vdots & \vdots & & \\ a_{mk_1} & a_{mk_2} & \cdots & a_{mk_m} \end{vmatrix} b_{k_11}b_{k_22}\cdots b_{k_mm}.$$

为了要计算和式(这里 $k_1,k_2,\cdots,k_m$ 各自独立历遍 $1,2,\cdots,n$)，首先注意到，只要各个 k_i 里有二个相同，该项一定为零。所以在归并各项时 只要注意 k_i 取不同值的项。为此可在集合$\{1,2,\cdots,n\}$里取定一个 m 集 $1\leqslant r_1<r_2<\cdots<r_m<n$，对这个 m 集的全部排列求和

$$\sum_{(t_1,t_2,\cdots,t_m)} A\begin{pmatrix}1 & 2 & \cdots & m\\ t_1 & t_2 & \cdots & t_m\end{pmatrix} b_{t_11}b_{t_22}\cdots b_{t_mm}\ (*)$$

通过置换将每一个排列 $(t_1,t_2,\cdots,t_m)$ 换成自 然 顺 序 $r_1<r_2<\cdots<r_m$ 因此

$$A\begin{pmatrix}1 & 2 & \cdots & m\\ t_1 & t_2 & \cdots & t_m\end{pmatrix}=(-1)^{[t_1,t_2,\cdots,t_m]}A\begin{pmatrix}1 & 2 & \cdots & m\\ r_1 & r_2 & \cdots & r_m\end{pmatrix}.$$

因此(*)便可写成

$$\sum_{(t_1t_2\cdots t_m)} A\begin{pmatrix}1 & 2 & \cdots & m\\ r_1 & r_2 & \cdots & r_m\end{pmatrix}(-1)^{[t_1t_2,\cdots t_m]}b_{t_11}b_{t_22}\cdots b_{t_mm}$$

$$=A\begin{pmatrix}1 & 2 & \cdots & m\\ r_1 & r_2 & \cdots & r_m\end{pmatrix}\sum_{(t_1,t_2,\cdots,t_m)}(-1)^{[t_1,\cdots,t_m]}b_{t_11}b_{t_22}\cdots b_{t_mm}$$

$$=A\begin{pmatrix}1 & 2 & \cdots & m\\ r_1 & r_2 & \cdots & m\end{pmatrix}B\begin{pmatrix}r_1 & r_2 & \cdots & r_m\\ 1 & 2 & \cdots & m\end{pmatrix}.$$

这是对一个固定的 m 集求和的结果，全部的和式便是按 $\{1,2,\cdots,n\}$ 的全部 m 集求和。 ▨

由上面推理过程可知，在 $m>n$ 时，在 $\{1,2,\cdots,n\}$ 要取 m 集，一定有某两个 k_i,k_j 是相同的，所以这种情形下 $\det C=0$。

9. 用 Cauchy-Binet 公式证明 Lagrange 恒等式

$$\left(\sum_{i=1}^{n}|a_i|^2\right)\left(\sum_{i=1}^{n}|b_i|^2\right)-\left|\sum_{i=1}^{n}a_i\overline{b_i}\right|^2=\sum_{1\leqslant i<j\leqslant n}|a_ib_j-a_jb_i|^2.$$

10. 证明 Cauchy 不等式

$$\left(\sum_{i=1}^{n}|a_i|^2\right)\left(\sum_{i=1}^{n}|b_i|^2\right)\geqslant|\sum a_ib_i|^2。$$

11. 证明矩阵方程 $Ax=0$（其中 A 是方阵）有非零解$\Longleftrightarrow|A|=0$。

12. 位于有同样号码的行和列的交叉处的子式，称为矩阵的主子式。证明如果 B 是实矩阵，则矩阵 $A=B'B$ 的所有主子式都是非负的。

13. 整数方阵如果它的行列式等于±1，则称为幺模矩阵。证明：整数矩阵有整数逆阵$\Longleftrightarrow$整数矩阵是幺模矩阵。

14 证明对任意实阵 A，有 $r(A^TA)=r(A)$。

15. 在实数域上，线性方程组 $A^TAx=A^Tb$ 总有解。

16. 用数学归纳法证明 $(A+B)^n=\sum_{k=0}^{n}C_n^kA^{n-k}B^k\quad(AB=BA)$。

第二章　特征值与特征向量

A 类

1. 设 λ 为矩阵 A 的特征值，x 为它对应的特征向量。证明：

(i) $a\lambda$ 为矩阵 aA 的特征值，且对应的特征向量仍为 x；

(ii) λ^m 为 A^m 的特征值（m 为正整数），且对应的特征向量仍为 x；

(iii) 当 $\lambda\neq0$ 时，λ^{-1} 为 A^{-1} 的特征值，且对应的特征向量仍为 x；

(iv) λ 也是 A^T 的特征值。

2. 求下列线性算子的特征值与特征向量。

(i) $\sigma((x_1,x_2))=(3x_1+3x_2,x_1+5x_2)$；

(ii) $\sigma((x_1,x_2))=(x_2,x_1)$；

(iii) $\sigma((x_1,x_2,x_3))=(2x_1+x_2,x_2-x_3,2x_3+4x_3)$。

3. 求下列矩阵的特征值与特征向量，指出与对角阵相似的矩阵，并写出相应的对角阵。

(i) $A=\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}$；　　(ii) $A=\begin{bmatrix}-1&1&0\\-4&3&0\\1&0&2\end{bmatrix}$，

(iii) $A=\begin{bmatrix} 0 & 2 & -1 \\ -2 & 5 & -2 \\ -4 & 8 & -3 \end{bmatrix}$； (iv) $A=\begin{bmatrix} 2 & -1 & -1 \\ 2 & -1 & -2 \\ -1 & 1 & 2 \end{bmatrix}$，

(v) $A=\begin{bmatrix} -2 & 1 & 0 \\ -4 & 3 & 0 \\ 4 & -1 & 2 \end{bmatrix}$； (vi) $A=\begin{bmatrix} 1 & 3 & 1 & 2 \\ 0 & -1 & 1 & 3 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 2 \end{bmatrix}$；

(vii) $A=\begin{bmatrix} 1 & 3 & 1 & 3 \\ 0 & -2 & 1 & 3 \\ 0 & 0 & 2 & 3 \\ 0 & & 0 & 1 \end{bmatrix}$。

4. (i) 用 t_rA 表示矩阵 A 的迹，证明 $t_r(AB)=t_r(BA)$；

(ii) 应用(i)证明相似矩阵有相同的迹。

5. 设有分块矩阵 $M=\begin{bmatrix} A_1 & B \\ O & A_2 \end{bmatrix}$，其中 A_1, A_2 是方阵。证明 M 的特征多项式是 A_1 与 A_2 的特征多项式的乘积。

6. 求下列矩阵的零化多项式。

(i) $\begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix}$， (ii) $\begin{bmatrix} 2 & 3 & -2 \\ 0 & 5 & 4 \\ 0 & -2 & -1 \end{bmatrix}$； (iii) $\begin{bmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ 3 & 6 & 1 \end{bmatrix}$；

(iv) $\begin{bmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & 6 & 1 \end{bmatrix}$； (v) $\begin{bmatrix} 1 & 3 & 1 & 2 \\ 0 & -1 & 1 & 3 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 2 \end{bmatrix}$。

7. 试用最小多项式来判定第 6 题中与对角阵相似的矩阵。

8. 求下列矩阵的特征多项式与最小多项式。

(i) $\left[\begin{array}{cc|cc|c} 2 & 5 & & & \\ & 2 & & & \\ \hline & & 4 & 2 & \\ & & 3 & 5 & \\ \hline & & & & 7 \end{array}\right]$； (ii) $\left[\begin{array}{cc|ccc} 3 & 1 & & & \\ & 3 & & & \\ \hline & & 3 & 1 & \\ & & & 3 & 1 \\ & & & & 3 \end{array}\right]$；

(iii) $\begin{bmatrix} a & & & & \\ & a & & & \\ & & a & & \\ & & & a & \\ & & & & a \end{bmatrix}$。

9. 设 $A=\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $B=\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{bmatrix}$.

(i) 求 A 和 B 的特征多项式;

(ii) 求 A 和 B 的最小多项式;

(iii) A 和 B 是否相似。

10. (i) 证明幂零性质是相似不变量;

(ii) 证明幂零指标是相似不变量,

(iii) 证明幂等性是相似不变量。

11. 设 $A=P^{-1}BP$,求证:对任一多项式 $f(x)$ 均有

$$f(A)=P^{-1}f(B)P。$$

12. A 的特征值是 0 1,对应的特征向量是 $(1,2)^T,(2,-1)^T$。你能否断言 A 是对称阵,求 A 的迹,行列式以及矩阵 A。

13. 在第 12 题 A^2 的特征值是什么? 特征向量是什么?

14. 假如 x 是矩阵 A 对应于特征值 λ_0 的特征向量,试求 $P^{-1}AP$ 对应于 λ_0 的特征向量。

15. 设 (i) $A=\begin{bmatrix} 3 & -2 \\ 4 & -3 \end{bmatrix}$, (ii) $A=\begin{bmatrix} -4 & 4 & 0 \\ -5 & 5 & 0 \\ 3 & 6 & 1 \end{bmatrix}$.

求 $A^8-A^5+3A+8I$。

16. 设 A 与 B 是同阶方阵,证明

(i) $AB+B$ 与 $BA+B$ 有相同的特征值;

(ii) 若 $AB=(B-A^T)A$,则 $A=0$。

17. 如果 n 阶对合矩阵 A 不是单位阵,则 $t_rA\leqslant n-2$。

18. 试写出下面矩阵关于特征值估计的圆盘。

(i) $A=\begin{bmatrix} 0.9 & 0.01 & 0.012 \\ 0.01 & 0.8 & 0.013 \\ 0.01 & 0.02 & 0.4 \end{bmatrix}$, (ii) $A=\begin{bmatrix} 3 & 2 & 2 & -4 \\ 2 & 3 & 2 & -1 \\ 1 & 1 & 2 & -1 \\ -2 & 2 & 2 & -1 \end{bmatrix}$。

19. 利用圆盘定理证明矩阵

$$A=\begin{bmatrix}4 & 1 & -2 & 0\\ -2 & 5 & 1 & 1\\ 3 & 1 & 6 & 1\\ 1 & -1 & 2 & 5\end{bmatrix}$$是可逆的。

20. 求下面矩阵 A 的谱半径。

$$A=\begin{bmatrix}1 & 0.1 & 0.2 & 0.3\\ 0.5 & 3 & 0.1 & 0.2\\ 1 & 0.3 & -1 & 0.5\\ 1.2 & -0.3 & -0.1 & -4\end{bmatrix}。$$

21. 应用圆盘定理。估计矩阵

$$A=\begin{pmatrix}1 & -\frac{1}{2} & -\frac{1}{2} & 0\\ -\frac{1}{2} & \frac{3}{2} & i & 0\\ 0 & -\frac{1}{2}i & 5 & \frac{1}{2}i\\ -1 & 0 & 0 & 5i\end{pmatrix}$$的特征值分布范围。

B 类

1. Markor 矩阵是 $n\times n$ 型实阵，它有两个性质

$$0\leqslant a_{ij}\leqslant 1; \qquad \sum_{j=1}^{n} a_{ij}=1。$$

证明 (i) Markov 矩阵与 Markov 矩阵的乘积仍为 Markov 矩阵，

(ii) Markov 矩阵的特征值 λ 满足条件 $|\lambda|\leqslant 1$;

(iii) 任一个 Markov 矩阵必有一特征值 $\lambda=1$。

2. 设 $A=[a_1a_2\cdots a_n]$，

(i) 求矩阵 A^TA 的特征值；

(ii) 求 $A=[1,1,\cdots,1]$ 时 A^TA 的特征值。

3. 不论 A,B 是怎样的两个 n 阶方阵，AB 与 BA 总有相同的特征多项式，但最小多项式可以不一样，试就下面二个矩阵验证之。

$$A=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix};\qquad B=\begin{bmatrix}1&1&1\\1&1&1\\-1&1&0\end{bmatrix}。$$

4. 试求下面矩阵的行列式：

$$A=\begin{bmatrix}x+\lambda & x & \cdots & x & x\\ x & x+\lambda & & x & x\\ \vdots & \vdots & & \vdots & \vdots\\ \vdots & \vdots & & x+\lambda & x\\ x & x & & x & x+\lambda\end{bmatrix},$$

5. 求下面矩阵的特征值

$$A=\begin{bmatrix}a_1^2+1 & a_1a_2 & \cdots & a_1a_n\\ a_2a_1 & v+a_2^2 & \cdots & a_2a_n\\ \vdots & \vdots & & \\ a_na_1 & a_na_2 & \cdots & 1+a_n^2\end{bmatrix}。$$

6. 设 n 阶矩阵 A 的 n 个特征值全同，且不等于零。

(i) 求证矩阵 $nA-t_r(A)I$ 是奇异矩阵；

(ii) 求 A^{-1}。

7. 设 n 阶实方阵 A 的特征值全是实数，且 A 的所有一阶主子式之和和所有二阶主子式之和都等于零，求证 A 是 n 阶幂零阵。

8. 设 A 是n阶奇异阵，求 $\mathrm{adj}A$ 的特征值。

9. 设 n 阶矩阵 A 的主对角元全为 1，且其特征值全是非负数，求证 $|A|\leqslant 1$。

10. $I-A$ 的特征值的模均小于 1，求证 $0<|\det A|<2^n$。

11. 设 A 是 n 阶实阵，若对任何 n 维实向量 $x\neq 0$ 均有 $x^TAx>0$，证明 $|A|>0$。

12. 设 $\lambda_1,\lambda_2,\lambda_3$ 为三阶矩阵 A 的三个特征值，其对应的特征向量为 $(1\ 1\ 1)^T,(0\ 1\ 1)^T,(0\ 0\ 1)^T$，求证：

$$A^n=\begin{bmatrix}\lambda_1^n & 0 & 0\\ \lambda_1^n-\lambda_2^n & \lambda_2^n & 0\\ \lambda_1^n-\lambda_3^n & \lambda_2^n-\lambda_3^n & \lambda_3^n\end{bmatrix}。$$

13. A 是 n 阶实方阵，I 是 n 阶单位阵。证明：若 $A^m=0$，则 $I-A$ 是可逆的。

14. A,B 是实对称阵，证明 $AB-BA$ 的特征值的实部为零。

15. 证明矩阵$\begin{bmatrix} 8 & 2 & 0 & 2 \\ 0 & 5 & 1 & 1 \\ 1 & -1 & 8 & -2 \\ 1 & 1 & 1 & 6 \end{bmatrix}$不是稳定矩阵。

（注：一个矩阵的全部特征值分布在左半平面，称该矩阵为稳定矩阵）

16. A是n阶实方阵，且

$$a_{ii} < -\left(\sum_{j\neq i} |a_{ij}|\right), \qquad i=1,2,\cdots n。$$

证明A是稳定矩阵。

17. $\delta(A)=\min\limits_{j}\left\{|a_{jj}|-\sum\limits_{i\neq j}|a_{ji}|\right\}$，证明对矩阵$A$的任一特征值$\lambda$都满足不等式$|\lambda|\geqslant\delta(A)$。

18. 如果$\delta(A)>0$，A是n阶复矩阵，证明

$$|\det A|\geqslant\delta^n(A)。$$

19. 如A是非奇异的n阶实阵，证明对某个$i\leqslant n$有

$$|a_{ii}|\leqslant\sum_{j\neq i}|a_{ij}|。$$

第三章 内积空间与特殊矩阵

A 类

1. 证明由 Euclid 空间的内积公理可导出下面的结论：

(i) $(\alpha,\beta_1+\beta_2)=(\alpha,\beta_1)+(\alpha,\beta_2)$，

(ii) $(\alpha,k\beta)=k(\alpha,\beta)$；

(iii) $(\alpha-\beta,\gamma)=(\alpha,\gamma)-(\beta,\gamma)$；

(iv) $(0,\alpha)=0$

(v) $\left(\sum\limits_{i=1}^{n}x_i\alpha_i,\ \sum\limits_{j=1}^{m}y_j\beta_j\right)=\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m}x_iy_j(\alpha_i\beta_j)$。

2. 证明：如果$(\alpha\beta)_1$，$(\alpha\beta)_2$是同一个 Euclid 空间的二个不同的内积，则$(\alpha,\beta)=(\alpha,\beta)_1+(\alpha,\beta)_2$亦是空间的一个内积。

3. 在R_4中求与向量$\alpha_1=(1,1,-1,1)$，$\alpha_2=(1,-1,-1,1)$，$\alpha_3=(2,1,1,3)$正交的单位向量α。

4. 下面给出的向量组

(a) $\alpha_1=(1,-2,1,3)$, $\alpha_2=(2,1,-3,1)$;

(b) $\alpha_1=(1,-1,1,-3)$, $\alpha_2=(-4,1,5,0)$

(i) 验证它们是正交向量组;

(ii) 将它扩充为空间的正交基。

5. 求由下列向量组生成的子空间的正交基。

(i) $\alpha_1=(2,3,-4,-6)$, $\alpha_2=(1,8,-2,-16)$,

$\alpha_3=(1,2,5,-14,5)$ $\alpha_4=(3,11,4,-7)$

(ii) $\alpha_1=(1,1,-1,-2)$, $\alpha_2=(-2,1,5,11)$,

$\alpha_3=(0,3,3,7)$, $\alpha_4=(3,-3,-3,-9)$。

6. 求 C_3 中由 $\alpha_1=(1,i,0)$, $\alpha_2=(1,2,1-i)$ 生成子空间的标准正交基。

7. 设 W 是由 $\alpha_1=(1,2,3,-1,2)$, $\alpha_2=(2,4,7,2,-1)$ 生成的 R_5 中的子空间,求 $W^{\perp}$ 的基。

8. R_4 的子空间 W 是下面线性方程组

$$\begin{cases}2x_1+x_2+3x_3-x_4=0\\3x_1+2x_2\qquad-2x_4=0\\3x_1+x_2+9x_3-x_4=0\end{cases}$$

的解空间,试求 $W^{\perp}$ 的向量应满足的线性齐次方程。

9. W_1,W_2 都是 Euclid 空间 V 的子空间。试证

(i) $W_1\supseteq W_2\Longrightarrow W_1^{\perp}\subseteq W_2^{\perp}$,

(ii) $(W_1+W_2)^{\perp}=W_1^{\perp}\cap W_2^{\perp}$;

(iii) $W_1^{\perp}+W_2^{\perp}=(W_1\cap W_2)^{\perp}$。

10. 试证明下列不等式。

(i) $(x_1y_1+\cdots+x_ny_n)^2\leqslant(x_1^2+\cdots+x_n^2)(y_1^2+\cdots+y_n^2)$, $x_i,y_i\in R$;

(ii) $\left(\int_a^b f(x)g(x)\mathrm{d}x\right)^2\leqslant\int_a^b f^2(x)\mathrm{d}x\cdot\int_a^b g^2(x)\mathrm{d}x, f,g\in C[a\ b]$,

(iii) $\sum\limits_{i=1}^{n}|a_i|\leqslant\sqrt{n(a_1^2+\cdots+a_n^2)}$;

(iv) $\int_0^1\sqrt{1+x^3}\,\mathrm{d}x<\sqrt{\frac{5}{4}}$。

11. 求一正交阵,它的第一行为行向量 $\alpha=\left(\frac{1}{3},\frac{2}{3},\frac{2}{3}\right)$。

12. 求一酉阵它的第一行是下列的行向量的单位向量。

(i) $\alpha=(1,1,-i)$

(ii) $(i,-2,2i)$。

13. 求向量 α 到 β 的投影与投影向量。

(i) $\alpha=(2,1,3)^T$, $\beta=(6,3,9)^T$;

(ii) $\alpha=(2,-3)$, $\beta=(3,2)^T$。

14. 在直线 $y=2x+1$ 上,求最近(5,2)的点。

15. 将 Gram-Schmidt 正交过程应用于(i),(ii),并将它的结果写成 $A=QR$。

(i) $\alpha_1=(1,0,0)^T$, $\alpha_2=(0,1,1)^T$, $\alpha_3=(0,-1,2)^T$。

(ii) $\alpha_1=(1,2,2)^\perp$, $\alpha_2=(1,3,1)^\perp$。

16. $A=\begin{bmatrix}1 & 0 & 2\\ 1 & 1 & 4\end{bmatrix}$ 找出 A 的零空间,并求出 $\alpha=(3,3,3)^T$ 在 $N(A)$ 与 $R(A^T)$ 的分解式。

17. 求向量 α 在子空间 L 上的正交投影 α_1 与正交分量 α_2。

$\alpha=(7,-4,-1,2)$, L 是下面方程的解空间

$$\begin{cases}2x_1+x_2+x_3+3x_4=0\\ 3x_1+2x_2+2x_3+x_4=0\\ x_1+2x_2+2x_3-9x_4=0。\end{cases}$$

18. 试找出下面矩阵为正定时,参数 λ 的所有值。

$$\text{(i)}\begin{bmatrix}5 & 2 & -1\\ 2 & 1 & -1\\ -1 & -1 & \lambda\end{bmatrix};\quad \text{(ii)}\begin{bmatrix}2 & \lambda & 1\\ \lambda & 1 & 0\\ 1 & 0 & 3\end{bmatrix}。$$

19. 试证明 A_k 是正规阵,并且求酉阵 U_k 使得 $U_k^H A_k U_k$ 是对角阵。

$$A_1=\begin{bmatrix}-\sqrt{2}i & -4\\ 4 & \sqrt{2}i\end{bmatrix},\quad A_2=\begin{bmatrix}4 & i\\ -i & 4\end{bmatrix}。$$

20. 设 W_1 表示由 $(1,2)^T$ 生成的子空间, W_2 表示由 $(0,1)^T$ 生成的子空间。又 σ 是 R^2 沿 W_2 在 W_1 上的投影,求 σ 在自然基下的矩阵。

21. 设 W 是 $(1,2)^T$ 生成的子空间, σ 是 R^2 在 W 上的正交投影,求 σ 在自然基下的矩阵。

22. 设 $A=\begin{bmatrix}1 & 2 & 1\\ 1 & -1 & 0\\ 2 & 1 & 1\end{bmatrix}$, $\beta=\begin{bmatrix}1\\ 1\\ 0\end{bmatrix}$,求 β 在 $R(A)$ 上的最佳近似。

23. 求矛盾方程 $\begin{cases}x_1+2x_2+x_3=1\\ x_1-x_2=1\\ 2x_1+x_2+x_3=0\end{cases}$ 的最小二乘解。

24. 证明对任意矩阵 $A\neq 0$, A^*A 必不是幂零阵。

25. A 是 n 阶正规阵,证明

(i) $A-\lambda I$ 也是正规阵;

(ii) 对于任何向量 x,向量 Ax 与 A^*x 的长度相同;

(iii) A 的任一特征向量都是 A^* 的特征向量。

26. 用 $\rho(A)$ 表示矩阵 A 的谱半径,证明

$$\rho(A)<\sqrt{t_r(A^*A)}。$$

27. 设 A 是 n 阶实对称阵,证明对于二次型 x^TAx 必有实数 c 使得

$$|x^TAx|\leqslant Cx^Tx。$$

28. 设 A 是正规阵, 其全部特征值为 $\lambda_1\cdots\lambda_n$, 证明 AA^* 与 A^*A 的全部特征值为 $|\lambda_1|^2\cdots\cdots|\lambda_n|^2$。

29. 设 A 是正规阵,证明

(i) A 是 Hermite 阵 $\Longleftrightarrow A$ 的特征值全为实数,

(ii) A 是酉阵 $\Longleftrightarrow A$ 的特征值的模都是 1,

(iii) A 是幂等阵 $\Longleftrightarrow A$ 的特征值只能是 0 与 1。

30. 设 A 是正规阵,证明

(i) 若 A 是幂等阵,则 A 是 Hermite 阵;

(ii) 若 A 是幂零阵,则 $A=0$;

(iii) 若 $A^3=A^2$,则 $A^2=A$,

(iv) 若 A 又是 Hermite 阵,而且也是一个幂幺阵(即有 $A^k=I$),则 A 是对合阵。

B 类

1. 证明下式是 R_2 的内积

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

的充要条件为 $a>0, ac>b^2$。

2. $V=\{c\cos t+c\sin t, c$ 是实数$\}$是二维实线性空间。对于$f,g\in V$，令 $(f,g)=f(0)g(0)+f\left(\frac{\pi}{2}\right)g\left(\frac{\pi}{2}\right)$。试证明$(f,g)$ 满足内积公理。再求 $h(t)=3\cos(t+7)+4\sin(t+9)$ 长度平方。

3. 设 $\alpha_1\cdots\alpha_m$ 是 n 维Euclid 空间的一组向量，令

$$\Delta=\begin{bmatrix}(\alpha_1,\alpha_1) & (\alpha_1,\alpha_2) & \cdots & (\alpha_1,\alpha_m)\\ (\alpha_2,\alpha_1) & (\alpha_2,\alpha_2) & \cdots & (\alpha_2,\alpha_m)\\ \vdots & & & \\ (\alpha_m,\alpha_1) & (\alpha_m,\alpha_2) & \cdots & (\alpha_m,\alpha_m)\end{bmatrix},$$

证明 $|\Delta|\neq 0\Longleftrightarrow \alpha_1\cdots\alpha_m$ 是线性无关。

4. A 是实对称矩阵，且 $A>0$，证明

$$(Ax,y)\leqslant(Ax,x)^{\frac{1}{2}}(Ay,y)^{\frac{1}{2}}.$$

5. A 是正定阵，则二次函数 $p(x)=\frac{1}{2}x^TAx-x^Tb$，且在 $Ax_0=b$ 时，有最小值。

6. 试证明：对任何 A,b（A 是 $m\times n$ 矩阵，b 是 m 维向量），下述两者有且仅有一个方程组有解。

(i) $Ax=b$；　(ii) $A^*y=0,\ y^*b\neq 0$。

（这命题常称为 Fredholm 判则）

7. 设 V 是内积空间，W 是它的一个子空间，α_0是V 的一个向量，如果

$$(\alpha_0,\ \beta)+(\beta,\alpha_0)\leqslant(\beta,\beta),\ \forall\beta\in W,$$

则

$$(\alpha_0,\ \beta)=0,\ \ \forall\beta\in W.$$

8. β 是 Euclid 空间 V 的任一个向量，W 是 V 的一个子空间，则 $\alpha\in W$ 是 β 在 W 的最佳近似$\Longleftrightarrow\beta-\alpha\in W^{\perp}$。

9. 设 $A=\begin{bmatrix}2&1\\1&1\\2&1\end{bmatrix}$，$b=\begin{bmatrix}12\\6\\8\end{bmatrix}$。

(i) 求 $R(A)$ 的标准正交基；

(ii) 写出 A 的 QR 分解；

(iii) 解最小二乘解 $Ax=b$

(iv) 证明 $u_1=(0\ 1\ 0)^T$，$u_2=\left(\frac{1}{\sqrt{2}}\ 0\ \frac{1}{\sqrt{2}}\right)^T$也是 $R(A)$的标准

正交基。

10. 设 A 是 $m\times n$ 实矩阵，$r(A)=n$，令 $P=A(A^TA)^{-1}A^T$；证明

(i) 对所有 $b\in R(A)$ 均有 $Pb=b$ 成立；

(ii) 对于 $b\in R(A)^\perp$ 有 $Pb=0$，

(iii) $P^2=P$，$P^T=P$，

就 R^3，对上面的结果给出几何解释。

11. S 是 Euclid 空间 V 的一个子空间，又 $\{\alpha_1\cdots\alpha_k\}$ 是 S 的一个正交基，$\beta\in S$，证明用 S 中的元素对 β 作最佳的最小二乘逼近为

$$p=\sum_{i=1}^{k}\frac{(\beta,\alpha_i)}{(\alpha_i,\alpha_i)}\alpha_i。$$

12. 设 A,B 都是正定阵，证明 AB 的特征值都大于零。

13. 证明 Hermite 矩阵 A 的特征值全部落在闭区间 $[a,b]$ 上 $\Longleftrightarrow A-aI\geqslant 0$，$bI-A\geqslant 0$。

14. A 是 n 阶方阵，记 $a=\max\limits_{1\leqslant i,j\leqslant n}|a_{ij}|$，证明

$$|\det A|\leqslant\frac{n^{n/2}}{a^{a}}。$$

15. 已知正交矩阵

$$A=\frac{1}{3}\begin{bmatrix}2 & 1 & -2\\ 1 & 2 & 2\\ 2 & -2 & 1\end{bmatrix}$$

表示一个旋转，求其旋转轴与旋转角。

16. 若 3×3 矩阵 S 表示一个反射，则存在一个正交矩阵 C，使得

$$C^{-1}SC=\begin{bmatrix}-1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}。$$

当 $S=\frac{1}{9}\begin{bmatrix}7 & 4 & -1\\ 4 & 1 & 8\\ -4 & 8 & 1\end{bmatrix}$ 时，求这样的矩阵 C。

17. 证明任一正规矩阵 A 必有谱分解

$$A=\lambda_1P_1+\cdots+\lambda_rP_r,\ 1\leqslant r\leqslant n,$$

满足 (i) $\lambda_1\cdots\lambda_r$ 是不同的数；

(ii) $P_i^2=P_i$；

(iii) $P_i^2 = P_i$, $(i \neq j)$;

(iv) $P_i P_j = 0, (i \neq j)$

(v) $P_1 + \cdots P_r = I$。

18. 设 $A = \lambda_1 P_1 + \cdots + \lambda_r P_r$ 是正规矩阵 A 的谱分解，证明对任一多项式 $f(x)$ 有

$$f(A) = f(\lambda_1)P_1 + \cdots + f(\lambda_r)P_r。$$

19. (i) 设 A 是 n 阶正规阵，证明 A 的奇异值 $S_i = |\lambda_i|, i = 1, 2, \cdots, n$, λ_i 是 A 的特征值。

(ii) A 是 n 阶矩阵，证明

$$|\det A| = S_1 \cdot S_2 \cdots S_n,\ S_i \text{ 是 } A \text{ 的奇异值。}$$

20. A 是一个反对称阵，证明 $I + A$ 是可逆的，并且 $U = (I - A)(I + A)^{-1}$ 是正交矩阵。

21. 设 A 是 n 阶阵，它的秩为 r，证明

$$r(A) \geqslant \frac{|t_r A|^2}{t_r(AA^*)} = \frac{\left|\sum\limits_{i=1}^{n} a_{ii}\right|^2}{\left|\sum\limits_{i,j=1}^{n} a_{ij}\right|^2}。$$

利用上面结果求证下面矩阵是非奇异的：

(i) $A = \begin{bmatrix} 4 & 3 & 1 \\ 1 & 4 & 2 \\ 1 & 2 & 4 \end{bmatrix}$,

(ii) $A = \begin{bmatrix} 5 & 0 & 1 & 0 & 1 & -1 \\ 1 & 5 & -1 & -1 & -1 & 0 \\ -1 & -1 & 4 & -1 & -1 & -1 \\ 0 & 1 & 0 & 5 & 1 & -1 \\ 2 & 0 & -2 & -1 & 5 & 1 \\ 0 & -1 & 0 & 1 & 0 & 5 \end{bmatrix}$。

第四章 Jordan 标准型

A 类

1. 设 σ 是 q 维线性空间 V 上的线性算子，V 可分解为 σ 的不变子空

间 V_1, V_2, V_3, V_4 的直和，设它们的维数分别为 2,4,1,1，将 V_i 的基合并为 V 的基，试写出 σ 在这基下的矩阵。

2. 设 σ 是 n 维线性空间 V 上的线性算子。如果 $\sigma^{k-1}\xi \neq 0$ 而 $\sigma^k\xi \neq 0$。求证 $\xi, \sigma(\xi), \cdots \sigma^{k-1}(\xi)$ $(k \geqslant 1)$ 是线性无关。

3. 设 σ 是 n 维线性空间 V 上的线性算子。如果有 $\sigma^n = 0$ 但 $\sigma^{n-1} \neq 0$。证明 σ 在某组基下的矩阵为

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

4. 在 R^2 中 σ 表示平面关于直线 $y = x$ 的反射。

(i) 写出 σ 关于标准基 e_1, e_2 的矩阵；

(ii) 试确定两个 σ 的不变子空间 V, W 有 $R^2 = V \oplus W$，

(iii) 选 V, W 的基并成 R^2 的基，并写出 σ 在该基下的矩阵；

(iv) 指出导出算子 $\sigma_1 = \sigma|V, \sigma_2 = \sigma|W$。

5. 设矩阵 A 是严格的上三角阵（即主对角元都为零），证明 A 是幂零阵。

6. 下列矩阵是幂零阵：

$$A = \begin{bmatrix} 7 & 1 \\ x & y \end{bmatrix}; \quad B = \begin{bmatrix} 0 & 1 \\ x & y \end{bmatrix}, \quad C = \begin{bmatrix} x & 2 & 3 \\ 0 & y & 5 \\ 0 & 0 & z \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 2 & 3 \\ x & 0 & 5 \\ y & 5 & 0 \end{bmatrix}。$$

试求 A, B, C, D。

7. 试证下列矩阵为幂零阵，求出各矩阵的幂零指数，并写出标准型。

$$\text{(i)}\quad A = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \qquad \text{(ii)}\quad A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ -4 & -2 & 0 & 0 \\ 6 & 3 & 0 & 0 \\ 14 & 7 & 0 & 0 \end{bmatrix},$$

$$\text{(iii)}\quad A = \begin{bmatrix} -1 & 0 & -1 & -1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix}。$$

8. 证明:两个三阶的幂零阵相似$\Longleftrightarrow$它们有相同的指标。

9. 已知

$$
J=\left[\begin{array}{ccc|cc|ccc|c|c}
2 & 1 & 0 & & & & & & & \\
0 & 2 & 1 & & & & & & & \\
0 & 0 & 2 & & & & & & & \\
\hline
 & & & 2 & 1 & & & & & \\
 & & & 0 & 2 & & & & & \\
\hline
 & & & & & 1 & 1 & 0 & & \\
 & & & & & 0 & 1 & 1 & & \\
 & & & & & 0 & 0 & 1 & & \\
\hline
 & & & & & & & & 0 & \\
\hline
 & & & & & & & & & 0
\end{array}\right]。
$$

(i) 试写 J 的特征多项式,最小多项式。

(ii) 试写出 J 的特征。

10. 试决定一五阶矩阵的可能的 Jordan 标准型。已知它的最小多项式 $m(\lambda)=(\lambda-2)^2$。

11. 试确定下述情形,矩阵的可能的标准型。

(i) $p(\lambda)=(\lambda-2)^4(\lambda-3)^2$, $m(\lambda)=(\lambda-2)^2(\lambda-3)^3$;

(ii) $p(\lambda)=(\lambda-7)^5$ $m(\lambda)=(\lambda-7)^2$;

(iii) $p(\lambda)=(\lambda-7)^7$, $m(\lambda)=(\lambda-7)^3$;

(iv) $p(\lambda)=(\lambda-3)^4(\lambda-5)^4$, $m(\lambda)=(\lambda-3)^2(\lambda-5)^2$。

12. 求下列矩阵的 Jordan 标准型以及使 $P^{-1}AP=J$ 的 P。

(i) $A=\begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ -2 & -1 & 1 \end{bmatrix}$;　(ii) $A=\begin{bmatrix} 3 & 1 & -3 \\ -7 & -2 & 9 \\ -2 & -1 & 4 \end{bmatrix}$;

(iii) $A=\begin{bmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 7 & 1 & 2 & 1 \\ -17 & -6 & -1 & 0 \end{bmatrix}$;　(iv) $A=\begin{bmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 3 & 1 & 2 & 1 \\ -3 & -1 & -1 & 0 \end{bmatrix}$;

(v) $A=\left[\begin{array}{ccc|cc|cc} 2 & -1 & 2 & & & & \\ & 2 & -1 & & & & \\ & & 2 & & & & \\ \hline & & & 2 & -1 & & \\ & & & & 2 & & \\ \hline & & & & & 2 & -1 \\ & & & & & & 2 \end{array}\right]$.

13. 求下面矩阵的 Jordan 标准型，并求变换矩阵 P。

$$A_1=\begin{bmatrix} 4 & 1 & 1 \\ -7 & -1 & -3 \\ 3 & 1 & 3 \end{bmatrix};\quad A_2=\begin{bmatrix} 3 & 2 & 1 \\ 0 & 4 & 0 \\ -1 & 2 & 5 \end{bmatrix},$$

$$A_3=\begin{bmatrix} 2 & 2 & 2 & 1 \\ -1 & -1 & -3 & -2 \\ 1 & 2 & 5 & 3 \\ -1 & -2 & -4 & -2 \end{bmatrix}。$$

14. $T:p(t)\longrightarrow p(t+3)$ 是线性空间 $P_2(t)$ 的一个线性变换。问适当选择基，T 能否用对角阵表示？

15. 求实矩阵 P 使 $P^{-1}AP$ 是 Jordan 标准型。

$$A=\begin{bmatrix} -9 & -13 & 3 & -17 \\ 0 & 0 & -1 & 1 \\ 7 & 13 & -5 & 17 \\ 7 & 9 & -1 & 13 \end{bmatrix}.$$

16. 设

$$B=\begin{bmatrix} 3 & 5 & 5 & 2 \\ -5 & -7 & -7 & -3 \\ 5 & 8 & 10 & 5 \\ -5 & 9 & -12 & 0 \end{bmatrix},$$

求 B 的 Jordan 标准型以及变换矩阵 P。

17. 证明不为零的幂零阵一定不相似于对角阵。

18. 证明周期矩阵 A（即对某个自然数 k 有 $A^k=I$）相似于对角阵，并写出这种对角阵的形式。

19. 证明单纯算子具有下述性质：

(i) 像空间由对应于非零特征值的特征向量生成，

(ii) 核与像空间的交仅有零向量（因此空间是这两个子空间的直和）。

20. 写出矩阵，它的特征值 2,3,4,5，特征为

$$\lambda = 2 \quad 3 \quad 4 \quad 5$$

$$\{(2\ 1\ 1)(1\ 1\ 1)(3)(2\ 1)\}。$$

B 类

1. 试判断下面 4 个矩阵，哪些是相似的

$$A=\begin{bmatrix}-3 & 3 & -2\\ -7 & 6 & -3\\ 1 & -1 & 2\end{bmatrix};\qquad B=\begin{bmatrix}0 & 1 & -1\\ -4 & 4 & -2\\ -2 & 1 & 1\end{bmatrix};$$

$$C=\begin{bmatrix}0 & -1 & -1\\ -3 & -1 & -2\\ 7 & 5 & 6\end{bmatrix};\qquad D=\begin{bmatrix}0 & 1 & 2\\ 0 & 1 & 1\\ 0 & 0 & 2\end{bmatrix}。$$

2. 设 V 是由函数 e^x, xe^x, x^2e^x, e^{2x} 的线性组合生成的线性空间。定义 V 的一个线性算子如下：$T(f)=f'$，求 T 的Jordan 标准型以及Jordan 基。

3. 设有正整数 k 使得 $(I+A)^k=0$。证明 A 是非奇异的。

4. 证明实矩阵 A,B 在实数域上相似$\Longleftrightarrow A,B$ 在复数域相似。

5. 设 A 是 4×4 矩阵，它的最小多项式为 $(x-1)(x-2)^2$。

(i) 请写出矩阵 A 的可能 Jordan 型；

(ii) 加上什么条件，A 的 Jordan 型唯一确定。

6. 令 V 是实线性空间，σ 是 V 的一个线性变换，它关于 V 的某一个基的矩阵是 $A=\begin{bmatrix}6 & -3 & -2\\ 4 & -1 & -2\\ 10 & -5 & -3\end{bmatrix}$。

(i) 求 σ 的最小多项式 $p(x)$，并把它分解为两个不可约的多项式 p_1, p_2 乘积；

(ii) 令 $W_i=\{\alpha \mid p_i(\sigma)\alpha=0\}$，$(i=1,2)$，证明 W_i 是 σ 的不变子空间，且 $V=W_1\oplus W_2$；

(iii) 在每一个子空间 W_i 选取一基，并成 V 的基，使得 σ 关于这个基的矩阵只出现三个非零元素。

7. 证明：一是秩是1的 n 阶矩阵或可对角化，或是幂零，但这两种情况不可能同时出现。

8. 设 $A=\begin{bmatrix}2 & 0 & 0\\ a & 2 & 0\\ b & c & -1\end{bmatrix}$ 是复矩阵

(i) 求出 A 的一切可能 Jordan 标准型；

(ii) 给出 A 可对角化的条件。

9. 证明：任何一复矩阵 A 可分解为 $A=D+N$，其中 D 是可对角化阵，N 是幂零阵。

10. 设 A 是 Hermite 阵，假定 A 的一阶主子式之和与二阶主子式之和均为零，求证 $A=0$。

11. 计算 $\lim\limits_{n\to\infty}A^n$，其中 $A=\begin{pmatrix}\frac{1}{2} & 1 & 1\\ 0 & \frac{1}{3} & 1\\ 0 & 0 & \frac{1}{5}\end{pmatrix}$.

(矩阵的极限，即它每个元素取极限)

12. 设 $A=\begin{bmatrix}0 & 0 & 1\\ 1 & 0 & -1\\ 0 & 1 & 1\end{bmatrix}$，求 A^{100}。

13. 对于以下情形，请给出三维空间 V_3 的线性变换 σ 的例。

(i) V_3 不是线性变换 σ 的值域与核的直和；

(ii) V_3 是线性变换 A 的值域和核的直和，但不是平行于 $\mathrm{Ker}\sigma$ 在 $\mathrm{Im}\sigma$ 上的投影。

14. A 为幂零矩阵，试求 $A+I$ 的行列式。

15. $A^2=A$，且 $r(A)=r$，试求 $A+I$ 的行列式。

16. 如果矩阵 A 的特征多项式和最小多项式相同，问 A 的 Jordan 标准型有何特点？

第五章　矩阵分析初步

A 类

1. 设 $A_k=\begin{bmatrix}\dfrac{1}{k^2} & \dfrac{k^2+k}{k^2+1}\\ 2 & \left(1-\dfrac{2}{k}\right)^k\end{bmatrix}$，求 $\lim\limits_{k\to\infty}A_k$。

2. 举出一 2×2 型的矩阵序列，该序列中的每一个矩阵都可逆，但其极限是不可逆的。

3. 如果 $\lim\limits_{n\to\infty}A_nB_n$ 存在，是否 $\lim A_n$，$\lim B_n$ 一定存在。若回答肯定，请予以证明。否则请举一反例。

4. 判别下列矩阵的幂收敛性：

(i) $A=\begin{bmatrix}0.8 & 0.3\\ 0.1 & 0.5\end{bmatrix}$；　　(ii) $A=\begin{bmatrix}-0.6 & 1 & 0.8\\ 0 & 0.2 & 0\\ -1.6 & 1 & 0.8\end{bmatrix}$。

5. 如果 $L=\lim\limits_{n\to\infty}A^n$，证明 L 是幂等阵。

6. (i) 设 $A=\begin{bmatrix}C & O\\ D & I\end{bmatrix}$，其中 C，D 是 n 阶方阵，I 为 n 阶单位阵，O 是 n 阶零阵，求 A^k；

(ii) 设 $A=\begin{bmatrix}0.1 & 0.3 & 0 & 0\\ 0.7 & 0.6 & 0 & 0\\ 0.1 & 0 & 1 & 0\\ 0.1 & 0.1 & 0 & 1\end{bmatrix}$，求 $\lim\limits_{k\to\infty}A^k$。

7. 设 $A=\begin{bmatrix}2 & -\dfrac{1}{2}\\ 2 & 0\end{bmatrix}$，求 $\sum\limits_{k=0}^{\infty}\dfrac{A^k}{2^k}$。

8. 设有一 5 阶复矩阵 A 的特征多项式为

$$(\lambda-1)^2\left(\lambda-\frac{1}{2}\right)^2\left(\lambda-\frac{1}{4}\right),$$

证明A是幂收敛$\Longleftrightarrow r(A-I)=3$。

9．设$A=J_3(4)\oplus J_4(-1)$，求$I+A^2$。

10．设$A=P(J_3(2)\oplus J_4(1))P^{-1}$，又$\sum\limits_{k=0}^{\infty} a_k t^k=f(t)$，$(|t|<3)$，求$\sum\limits_{k=0}^{\infty} a_k A^k$。

11．设$A=P\left(J_2\left(\frac{1}{2}\right)\oplus J_1\left(\frac{1}{2}\right)\oplus J_5(-1)\right)P^{-1}$，求$\sum\limits_{k=0}^{\infty}\frac{A^k}{4^k}$。

12．已知$J=\begin{bmatrix}2&0&0&0\\0&1&1&0\\0&0&1&0\\0&0&0&6\end{bmatrix}$，

求$e^J,\sin J,\cos J$。

13．已知$J=J_3(2)\oplus J_2(3)\oplus J_1(-1)$，求$e^J,\sin J,\cos J$。

14．已知4阶矩阵A的特征值为$\pi,-\pi,0,0$求证：

$$\cos A=I-2\pi^{-2}A^2;$$

$$\sin A=A-\pi^{-2}A^3。$$

15．设　$A=\begin{bmatrix}1&1\\0&0\end{bmatrix}$，　$B=\begin{bmatrix}1&-1\\0&0\end{bmatrix}$，求$e^A,e^B,e^{A+B},e^A\cdot e^B$。

16．设　$A=\begin{bmatrix}4&-1\\1&2\end{bmatrix}$，求$e^A,\sin A,\cos A$。

17．设矩阵函数

$$A(t)=\begin{bmatrix}\cos t&\sin t\\-\sin t&\cos t\end{bmatrix},$$

试计算$\frac{d}{dt}A(t)$，　$\frac{d}{dt}A^{-1}(t)$，$\frac{d}{dt}|A(t)|$，　$\left|\frac{d}{dt}A(t)\right|$（其中$|A|$表示矩阵$A$的行列式）。

18．设矩阵函数

$$A(t)=\begin{bmatrix}e^{2t}&te^t&1+t\\e^{-2t}&2e^{2t}&\sin t\\3t&0&t\end{bmatrix}$$

求$\int_0^t A(s)ds$。

19. 设线性空间 V 的线性算子 σ 的特征多项式 $f(\lambda)=(\lambda-2)^7$，最小多项式 $m(\lambda)=(\lambda-2)^3$。

(i) 确定 σ 的所有可能的 Jordan 标准型，并分别指出 $\sigma-2I$ 的零度；

(ii) 求 e^J；

(iii) 求 e^{Jt}。

20. 求下面方程组的通解

$$\frac{dx_1}{dt}=3x_1+x_2-3x_3,$$

$$\frac{dx_2}{dt}=-7x_1-2x_2+9x_3;$$

$$\frac{dx_3}{dt}=-2x_1-x_2+4x_3.$$

B类

1. 设 $L=\lim A^n$,证明 $(A-L)^n=A^n-L$。

2. 证明 $\sum\limits_{n=0}^{\infty}(A^n-L)$收敛$\Longleftrightarrow\lim A^n=L$。

3. 证明 $\sum\limits_{n=1}^{\infty}nT^n=T(I-T)^{-2},\ |T|<1$。

4. 计算 $\sum\limits_{n=0}^{\infty}\dfrac{1}{3^n}\begin{bmatrix}1 & 2\\ 1 & 1\end{bmatrix}^n$。

5. 计算 $\sin A$,已知 $A=\begin{bmatrix}\pi & 1 & 0 & 0 & 0\\ 0 & \pi & 1 & 0 & 0\\ 0 & 0 & \pi & 0 & 0\\ 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$。

6. 已知 $A=\begin{bmatrix}1 & 2 & 1\\ 2 & 4 & 2\\ 1 & 2 & 1\end{bmatrix}$，计算 $\lim\limits_{n\to\infty}\left(\dfrac{A}{\rho(A)}\right)^n$。

7. 设 $Z(t)$是非齐次常系数 $n\to\infty$ 微分方程组 $\dot{x}(t)=A(x)+f(t)$的一个解。证明满足 $x(t_0)=x^0$ 的唯一解为

$$x(t)=Z(t)+e^{(t-t_0)A}(x(t_0)-Z(t_0))。$$

8．求解微分方程 $\dot{x}=\begin{bmatrix}1&0&0\\2&1&-2\\3&2&1\end{bmatrix}x+\begin{bmatrix}1\\0\\0\end{bmatrix}e^{2t}$

9．求初始值问题

$$\dot{x}(t)=\begin{bmatrix}1&12\\3&1\end{bmatrix}x,\quad x(0)=\begin{bmatrix}0\\1\end{bmatrix}。$$

10．(i) 证明微分方程 $\dot{x}(t)=Ax(t)+ce^{\alpha t}$ 有如下形式的解 $x(t)=be^{\alpha t}$ 的充要条件为 $(\alpha I-A)b=c$，(b,c 都是 n 维向量)。

(ii) 解 $\dot{x}(t)=Ax(t)+e^{2t}c$，其中

$$A=\begin{bmatrix}3&1\\2&2\end{bmatrix},\quad c=\begin{bmatrix}-1\\-1\end{bmatrix},\ x(0)=\begin{bmatrix}0\\1\end{bmatrix}。$$

11．A 是 $n\times n$ 常数矩阵，b,c 是 n 维常数向量，m 是正整数。证明非齐次微分方程 $\dot{x}(t)=Ax(t)+t^m c,x(0)=b$ 具有下面形式的特解

$$x(t)=b_0+tb_1+\cdots+t^m b_m$$

的充分必要条件为

$$c=-\frac{1}{m!}A^{m+1}b。$$

12．A 为 $n\times n$ 矩阵，b,c,d 是 n 维常数向量，α 为非零常数。证明常系数微分方程组初始值问题

$$\dot{x}(t)=Ax(t)+(\cos\alpha t)c+(\sin\alpha t)d,\ x(0)=b,$$

具有如下形式的解 $x(t)=(\cos\alpha t)e+(\sin\alpha t)f$ 的充要条件为

$$(A^2+\alpha^2 I)b=-(Ac+\alpha d)。$$

13．已知 $e^{At}=\begin{bmatrix}2e^{2t}-e^t&e^{2t}-e^t&e^t-e^{2t}\\e^{2t}-e^t&2e^{2t}-e^t&e^t-e^{2t}\\3e^{2t}-3e^t&3e^{2t}-3e^t&3e^t-2e^{2t}\end{bmatrix}$，求 A。

14．解非齐微分方程组

$$\dot{x}(t)=\begin{bmatrix}0&1\\-1&0\end{bmatrix}x+\begin{bmatrix}1\\2\end{bmatrix},\ x(0)=\begin{bmatrix}x_1^0\\x_2^0\end{bmatrix}。$$

15. 已知 $A=\begin{bmatrix}0 & -1\\ 4 & 4\end{bmatrix}$，用 A 的最小多项式计算 $\arcsin\frac{A}{4}$。

16. 用谱上有相同值的多项式，计算下面矩阵函数

(i) $A=\begin{bmatrix}0 & 1\\ 0 & -2\end{bmatrix}$，求 e^{At}，

(ii) $A=\begin{bmatrix}2 & 1 & 4\\ 0 & 2 & 0\\ 0 & 3 & 1\end{bmatrix}$，求 e^{At}，$\sin At$。

17. (i) 证明：若 A 是实反对称阵，则 e^A 是正交阵。

(ii) 证明：若 A 为 Hermite，则 e^{iA} 是酉矩阵。

第六章 广义逆矩阵

1. 指出下列矩阵哪些有右逆，哪些有左逆，并求出存在的右逆和左逆。

(i) $\begin{bmatrix}1 & 2 & 5\\ 2 & 4 & 7\end{bmatrix}$，　(ii) $\begin{bmatrix}1 & 3\\ 2 & 4\\ 5 & 6\end{bmatrix}$；

(iii) $\begin{bmatrix}1 & 0 & 1\\ 2 & 0 & 2\\ 1 & 0 & 1\end{bmatrix}$。

2. 设 $r(BC)=r(B)$，证明存在矩阵 D 使 $B=BCD$，且 $C(BC)^-$ 是 B 的一个 g 逆。

3. 设 $r(BC)=r(C)$，证明存在矩阵 D 使 $C=DBC$，且 $(BC)^-B$ 是 C 的一个 g 逆。

4. 证明 (i) 如果 P 是列满秩阵，则 $(PA)^-=A^-P_L^{-1}$；

(ii) 如果 P 是行满秩阵，则 $(AP)^-=P_R^{-1}A^-$；

5. 设 A 是 $m\times n$ 矩阵，B 是 $m\times r$ 矩阵。则等式 $AA^-B=B$ 成立的充要条件是：存在矩阵 D 使得 $B=AD$。

6. 设 A 是 $m\times n$ 矩阵，B 是 $r\times n$ 矩阵，则等式 $BA^-A=B$ 成立的充要条件是存在矩阵 D 使得 $B=DA$。

7. 设 A 是秩为 r 的 $m\times n$ 矩阵，又 $r<\min(m,n)$。经过行初等变换将

A 化为 $B=\left(\frac{C}{D}\right)$，其中 C 是 $r\times n$ 行满秩阵，D 是 $(m-r)\times n$ 矩阵。因而必定存在矩阵 W，使 $D=WC$，证明 $B^-=[C_R^{-1},\ 0]_{n\times m}$。

又如果 P 是使 $PA=B$ 的非奇异矩阵，证明 $A^-=B^-P$。

8. 用第 7 题的行初等变换法计算矩阵 A 的一个 g 逆。

$$A=\begin{bmatrix} 1 & 1 & 1 & 0 \\ -1 & -1 & -1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}。$$

9. 设 A 是秩为 r 的 $m\times n$ 矩阵，又 $r<\min\{m,n\}$，经过列初等变换将 A 化为 $B=C(C\mid D)$，其中 C 是 $m\times r$ 列满秩阵，D 是 $(m\times(n-r))$ 矩阵，证明 $B^-=\begin{bmatrix} C_L^{-1} \\ \hline 0 \end{bmatrix}_{n\times m}$。

又如果 Q 是使 $AQ=B$ 的非奇异矩阵，证明 $A^-=QB^-$。

10. 用第 9 题的列初等变换法计算矩阵 A 的一个 g 逆

$$A=\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \\ 2 & 4 & 0 \end{bmatrix}。$$

11. 计算下列矩阵的 g 逆，并验证所得的结果。

$$A_1=\begin{bmatrix} 1 & 0 & 2 \\ 1 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}; \qquad A_2=\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 2 & 0 & 2 & 1 \end{bmatrix};$$

$$A_3=\begin{bmatrix} 1 & 1 \\ 2 & 2 \\ -1 & -1 \end{bmatrix}, \qquad A_4=\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 0 \end{bmatrix}。$$

12. 取第 11 题中的 A_1,A_2，并设 $b_1=(1,1,0,1)^T$，$b_2=(1,1,2)^T$，分别求出方程 $A_1x=b_1$，$A_2x=b_2$ 的通解。

13. 设矩阵 A 满足 (i) $A^2=A$，(ii) $GAG=G$，(iii) $R(G)\subset R(A)$，证明 $G^2=G$。

14. 设 $A=\begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 0 \end{bmatrix}$，$b=\begin{bmatrix} 2 \\ -1 \end{bmatrix}$，求 $Ax=b$ 的最小范数解。

15. 已知 $A=\begin{bmatrix} 1 & 2 & -1 \\ -1 & -2 & 1 \end{bmatrix}$，$b=\begin{bmatrix} 2 \\ 0 \end{bmatrix}$，求矛盾方程 $Ax=b$ 的最小二乘解。

16. 设 $G_1=A_l^- A A_m^-$，$G_2=A_m^- A A_l^-$，试问它们各是A的哪一类广义逆？

17. 证明 (i) $(A^*)^+=(A^+)^*$，(ii) $(A^+)^+=A$。

18. 求矩阵 A 的 A^+。

(i) $A=\begin{bmatrix} 1 & -1 & 0 & -1 \\ 0 & 0 & 2 & 2 \\ -1 & 1 & 0 & 1 \end{bmatrix}$； (ii) $A=\begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & -2 \\ -1 & 1 & 1 \\ 3 & -1 & -3 \end{bmatrix}$；

(iii) $A=\begin{bmatrix} 1 & 2 & -1 & 2 & 1 \\ 2 & 4 & 1 & -2 & 3 \\ 3 & 6 & -3 & 6 & 3 \end{bmatrix}$。

19. 设 $A=\begin{bmatrix} 2 & 1 \\ 2 & 0 \\ 2 & 0 \end{bmatrix}$，$b_1=\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$，$b_2=\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$，

分别求出 $Ax=b_1$，$Ax=b_2$ 的解，并指出它们各是怎样的解？

20. A 是 $m\times n$ 实阵，且 $r(A)=1$，证明

$$A^+=\frac{1}{\alpha}A^T,\ \left(\ \alpha=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}^2\ \right).$$

习题的提示与答案

第一章　基 础 知 识

A 类

1. (i) $C=AB=(-1,0,-1,4)=(5,2,-3,-3)+(-7,-2,4,2)+$
$(-1,2,1,1)+(2,-2,-3,4)$;

(ii) $C=(31,63,107)^T=(83,-15,38)^T+(-52,78,69)^T$

(iii) $C=0$，$\left.\begin{aligned}&2(-2,1)^T+(3,1)^T-(0,2)^T+(1,-1)^T=(0,0)^T\\&-(3,1)^T+2(0,2)^T+3(1,-1)^T=(0,0)^T\end{aligned}\right\}$ C的列向量

$\left.\begin{aligned}&-2(2,0)+3(1,-1)+(1,3)=(0,0)\\&(2,0)+(1,-1)+2(-1,2)-(1,2)=(0,0)\end{aligned}\right\}$ C的行向量

2. (i) 线性无关的列向量为$(1,2,3)^T,(-1,1,-3)^T$，它们分别记为α_1,α_3，则 $\alpha_2=2\alpha_1,\alpha_4=-2\alpha_3,\alpha_5=\dfrac{4}{3}\alpha_1+\dfrac{1}{2}\alpha_3$。

3. $x_2=x_3$。

4. $x_2=7x_3$，取四个行向量为基，因它们是行满秩的。

5. 同一子空间。

6. $V_1\cap V_2$的基为$(-1,2,1,2)$，V_1是二维的，$(1,2,3,6)$，$(-1,2,1,2)$构成V_1的基，$(1,-1,1,1)(-1,2,1,2)$是V_2的基，$(-1,2,1,2)$，$(1,2,3,6)$，$(1,-1,1,1)$是V_1+V_2的基。

7. $V_1\cap V_2$是2维，它的基为$(-1,0,1,0)$，$(0,-1,0,1)$。V_1+V_2的维是4，因而可取自然基。

8. V_1+V_2不是直和。

9. (略)。

10 $\mathrm{Ker}\sigma$是1维，它的基为$(3,-1,1)$　$\mathrm{Im}\sigma$是2维，它的基为$(1,0,1)^T(2,1,1)^T$。(可仿照定理的证明过程去做)

11. $\mathrm{ker}\sigma$是2维，它的基$(-1,2,1,0)$，$(1,-1,0,1)$。$\mathrm{Im}\sigma$是2维，它

的基是$(2,1,3)^T(1,1,2)^T$

12. (略)。

13. $\sigma(x_1,x_2)=(x_1-x_2,-3x_1+5x_2)$, $\sigma(7,4)=(3,-1)$。

14. (i) $A=\begin{pmatrix}2&-4&9\\5&3&-2\end{pmatrix}$; (ii) $A=\begin{pmatrix}3&4\\5&-2\\1&7\\4&0\end{pmatrix}$;

(iii) $A=(2,2,-7,-1)$。

15. $\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}1&-1\\1&0\end{pmatrix}\begin{pmatrix}x_1'\\x_2'\end{pmatrix}$

16. (i) $A=\begin{pmatrix}0&2&1\\1&-4&0\\3&0&0\end{pmatrix}$; (ii) $P=\begin{pmatrix}1&0&0\\1&1&0\\1&1&1\end{pmatrix}$

$$B=P^{-1}AP=\begin{pmatrix}3&3&1\\-6&-7&-1\\6&4&0\end{pmatrix}。$$

17. $B=\begin{pmatrix}3&11&5\\-1&-8&-3\end{pmatrix}$。

18. (i) $A=\begin{pmatrix}3&0&1\\-2&1&0\\-1&2&4\end{pmatrix}$; (ii) $B=\begin{pmatrix}1&-1&2\\0&2&1\\1&1&1\end{pmatrix}^{-1}\begin{pmatrix}3&0&1\\-2&1&0\\-1&2&4\end{pmatrix}$

$\begin{pmatrix}7&-1&2\\0&2&1\\1&1&1\end{pmatrix}$; (iii) $A^{-1}=\frac{1}{9}\begin{pmatrix}4&2&-1\\8&13&-2\\-3&-6&3\end{pmatrix}$。

19. $(\sigma+\tau)(x_1,x_2)=(x_1,x_1)$, $(\sigma\tau)(x_1,x_2)=(x_1-x_2,0)$,

$(\tau\sigma)(x_1,x_2)=(0,x_1+x_2)$, $\sigma^2(x_1,x_2)=(x_1+x_2,0)$,

$\tau^2(x_1,x_2)=(-x_1,x_2)$。

20. $\begin{cases}x_1+x_2-x_3=0\\2x_1+x_2-x_4=0。\end{cases}$

B 类

1. 因 $A(A+2I)=-I$。

2. 由等式说明 A 也是非奇异的，只要由 D 的表达式解出 A 便可。

3. 因为 $B(I-AB)=(I-BA)B$。则 $B=(I-BA)B(I-AB)^{-1}$，所以 $I=I-BA-BA=I-BA+(I-AB)B(I-AB)^{-1}A$

$$=(I-BA)\{I+B(I-AB)^{-1}A\}。$$

4. 仿 3。因为 $\beta'(I+A^{-1}\alpha\beta')=(1+\beta'A^{-1}\alpha)\beta'$，

所以
$$\beta'=\frac{1}{1+\beta'A^{-1}\alpha}\beta'(I+A^{-1}\alpha\beta')。$$

再仿 3 的方法。

5. (i) 因为 A 的诸列向量是独立的；

(ii) 根据(i)便可得。

6. 直接计算。

7. $\sigma(A+B)=\sigma(A)+\sigma(B)$，$\sigma(kA)=k\sigma(A)$ 易证。

$\sigma(AB)=\sigma(A)\sigma(B)$，因上三角阵乘积的对角元为它的因子对应的对角元的乘积。

8. 若 $W\neq\{0\}$，则必有一不为零的矩阵 $A\in W$，然后利用 W 的性质，证明 $E_{ij}\in W$ 因为$\{E_{ij}\}$构成 $R^{m\times n}$ 的基。

9. $$\left(\sum_{i=1}^{n}|a_i|^2\right)\left(\sum_{i=1}^{m}|b_i|^2\right)-\left|\sum_{i=1}^{n}a_i\overline{b_i}\right|^2=\begin{vmatrix}\sum\limits_{i=1}^{n}a_i\overline{a_i} & \sum a_i\overline{b_i}\\ \sum \overline{a_i}b_i & \sum b_i\overline{b_i}\end{vmatrix}$$

$$=\left|\begin{bmatrix}a_1 & \cdots & a_n\\ b_1 & \cdots & b_n\end{bmatrix}\begin{bmatrix}\overline{a_1} & \overline{b_1}\\ \vdots & \vdots\\ \overline{a_n} & \overline{b_n}\end{bmatrix}\right|=\sum_{1\leqslant i<j\leqslant n}\begin{vmatrix}a_i & a_j\\ b_i & b_j\end{vmatrix}\begin{vmatrix}\overline{a_i} & \overline{b_i}\\ \overline{a_j} & \overline{b_j}\end{vmatrix}$$

$$=\sum_{1\leqslant i<j\leqslant n}(a_ib_j-a_jb_i)(\overline{a_i}\overline{b_j}-\overline{b_i}\overline{a_j})$$

$$=\sum_{1\leqslant i<j\leqslant n}(a_ib_j-a_jb_i)\,(\overline{a_ib_j-a_jb_i})=\sum_{1\leqslant i<j\leqslant n}|a_ib_j-a_jb_i|^2。$$

10. 由 Lagrange 恒等式直接可得。

11. (略)

12. 实际上 Cauchy-Bient 定理更一般的形式为：

A是 $m\times k$型，B是$k\times n$型，又$C=AB$是 $m\times n$ 这样 C 的 $i_1 i_2,\cdots,i_r$行与$j_1,j_2,\cdots,j_r$列交叉处构成的子式$C\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}$有等式

$$C\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}=\sum_{1\leqslant h_1<h_2<\cdots<h_r\leqslant k} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ h_1 & h_2 & \cdots & h_r \end{pmatrix}$$

$$B\begin{pmatrix} h_1 & h_2 & \cdots & h_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}$$

这里$\sum$是对$\{1,2,\cdots,k\}$的全部 r 集求和。按上述公式便立即可证明。

13. 直接按定义以及 $A^{-1}=\dfrac{1}{|A|}\operatorname{adj}A$ 去做。

14. 设 A 为 $m\times n$ 型，只要证明 $N(A^TA)=N(A)$即可。

15. 证明系数矩阵的秩与增广矩阵的积相等。

16. (略)。

第二章　特征值与特征向量

A类

1. (iv) $|\lambda I-A|=0\Longrightarrow|\lambda I-A|^T=0\Longrightarrow|\lambda I-A^T|=0$。

2. (i) $\lambda_1=6$，特征向量为$(1,1)$；$\lambda_1=2$，特征向量为$(3,-1)$

 (ii) $\lambda_1=1,\alpha_1=(1,1)$；$\lambda_2=-1,\alpha_2=(1,-1)$。

 (iii) $\lambda_1=\lambda_2=2$，特征向量仅有一个为$(1,0,0)$，$\lambda_3=3$，特征向量为$(1,1,-2)$。

3. (i) $\lambda_1=\lambda_2=1$，$\lambda_3=-1$；$\lambda=1$对应的特征向量为$(0,1,0)^T$，$(1,0,1)^T$；$\lambda=-1$的特征向量$(1,0,-1)$，与对角阵相似；$A\sim\operatorname{diag}(1,1,-1)$

 (ii) $\lambda_1=\lambda_2=1$，对应特征向量为$(1,2,-1)^T$；$\lambda_3=2$，对应特征向量为$(0,0,1)^T$，独立的特征向量为 2，故不能与对角阵相似。

 (iii) $\lambda_1=0$，$\alpha_1=\left(-\dfrac{1}{2},1,2\right)^T$；$\lambda_2=\lambda_3=1$，$\alpha_2=(1,0,-1)^T$，$\alpha_3=(2,1\ 0)^T$，与对角阵相似；$A\sim\operatorname{diag}(0,1,1)$。

(iv) $\lambda_1=\lambda_2=\lambda_3=1$，对应的特征向量为$(1,1,0)^T,(1,0,1)^T$，故不能与对角阵相似。

(v) $\lambda_1=\lambda_2=2$，$\alpha_1=(1,4,0)^T$，$\alpha_2=(0,0,1)^T$；$\lambda_3=-1$，$\alpha_3=(1,1,-1)^T$，故与对角阵相似：$A\sim \mathrm{diag}(2,2,-1)$。

(vi) $\lambda_1=\lambda_2=2,\lambda_3=1,\lambda_4=-1$；对应的特征向量为：$\alpha_1=(6,1,3,0)^T$，$\alpha_3=(1,0,0,0)^T,\alpha_4=(3,-2,0,0)^T$，故不能与对角阵相似。

(vii) $\lambda_1=\lambda_2=1,\alpha_1=(1,0,0,0)^T,\alpha_2=(0,0,3,-1)^T$；$\lambda_3=2,\alpha_3=(7,1,4,0)^T$；$\lambda_4=-2$，$\alpha_4=(1,-1,0,0)^T$；故与对角阵相似：$A\sim \mathrm{diag}(1,1,2,-2)$。

4．直接计算。

5．$\lambda I-M=\begin{bmatrix}\lambda I_1-A_1 & -B\\ & \lambda I_2-A_2\end{bmatrix}$由 Laplace 展开定理便可知命题成立。

6．由 Hamilton-Cayley 定理可知，求每个矩阵的特征多项式，亦可满足本题要求。

(i) $|\lambda I-A|=\lambda^2-8\lambda+7$；　(ii) $|\lambda I-A|=\lambda^3-6\lambda^2+11\lambda+6$；

(iii) $|\lambda I-A|=(\lambda-1)^2(\lambda+2)$；(iv) $|\lambda I-A|=(\lambda-1)^2(\lambda+2)$；

(v) $|\lambda I-A|=(\lambda-1)(\lambda+1)(\lambda-2)^2$。

7．(i) $A\sim\mathrm{diag}(1,7)$　(ii) $|\lambda I-A|=(\lambda-1)(\lambda-2)(\lambda-3)$，$A\sim\mathrm{diag}(1,2,3)$。

(iii) A 的最小多项式为$(\lambda-1)(\lambda+2)$，因此 $A\sim\mathrm{diag}(1,1,-2)$。

(iv) 最小多项式为$(\lambda-1)^2(\lambda+2)$，因此 A 不能与对角阵相似。

(v) 最小多项式为$(\lambda-1)(\lambda+1)(\lambda-2)^2$，故 A 不能与对角阵相似。

8．(i) 特征多项式为$(\lambda-2)^3(\lambda-7)^2$，最小多项式为$(\lambda-2)^2(\lambda-7)$；

(ii) 特征多项式为$(\lambda-3)^5$，最小多项式为$(\lambda-3)^3$；

(iii) 特征多项式为$(\lambda-a)^5$，最小多项式为$(\lambda-a)$。

9．(i) A 的特征多项式为$(\lambda-1)^2(\lambda-2)$，B 的特征多项式为$(\lambda-1)(\lambda-2)^2$；

(ii) A,B 的最小多项式均为$(\lambda-1)(\lambda-2)$；

(iii) 由于 A,B 的特征多项式不同，所以 A,B 不相似。或者由A,B的迹不相等，因而 A 与 B 不相似。

10．直接验证

11. $A=P^{-1}BP \Longrightarrow A^k=P^{-1}B^kP \Longrightarrow f(A)=P^{-1}f(B)P$.

12. $A=P\begin{bmatrix}0&\\&1\end{bmatrix}P^{-1}$, $P=\begin{bmatrix}1&2\\2&-1\end{bmatrix}$, 但 $P^TP=5I$, 所以 $A=\frac{1}{5}P\begin{bmatrix}0&\\&1\end{bmatrix}P^T$, A 是对称阵, $t_rA=1$, $|A|=0$, $A=\begin{bmatrix}4&-2\\-2&1\end{bmatrix}$。

13. 特征值与特征向量与第 12 题的相同。

14. $P^{-1}AP$ 对应于 λ_0 的特征向量为 $P^{-1}x$。

15. (i) A的特征多项式 λ^2-1, 因此有 $A^2=I$, $g(A)=2A+9I=\begin{bmatrix}15&-4\\8&3\end{bmatrix}$。

(ii) A 的特征多项式 $\lambda(\lambda-1)^2$,同待定系数的方法可求得

$$g(A)=3A^2+8I=\begin{bmatrix}-4&12&0\\-15&23&0\\-117&144&11\end{bmatrix}。$$

16 (i) 因 CD 与 DC 有相同的特征值;

(ii) 由 $A^TA=BA-AB$ 两边取迹。

17. 因对合阵的特征值为 ±1。如果特征值全是 1,则 A 是单位阵。

18. (i) D_1: $|\lambda-0.9|\leqslant0.022$, D_2: $|\lambda-0.8|\leqslant0.023$, D_3: $|\lambda-0.4|\leqslant0.02$;

(ii) D_1: $|\lambda-3|\leqslant8$, D_2: $|\lambda-3|\leqslant5$, D_3: $|\lambda-2|\leqslant3$, D_4: $|\lambda+1|\leqslant6$。

19. 矩阵A满足对角强优条件。

20. A 的行的元绝对值和之最大值为 5.6; A 的列的元绝对值和之最大值为5,因此谱半径为 5。

21. (略)

B 类

1. (i) 设$A=[a_{ij}]$, $B=[b_{ij}]$均是 Markov 矩阵,

$$[AB]_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}(0\leqslant b_{ij}\leqslant1)\leqslant\sum_{k=1}^{n}a_{ik}=1,$$

$$\sum_{j=1}^{n}c_{ij}=\sum_{j=1}^{n}\left(\sum_{k=1}^{n}a_{ik}b_{kj}\right)=\left(\sum_{k=1}^{n}a_{ik}\right)\left(\sum_{j=1}^{n}b_{kj}\right)=1.$$

(ii) 设对应于 λ 的特征向量为 $\alpha=(x_1\cdots x_n)^T$,令 $|x_k|=\max(|x_1|\cdots$

$|x_n|$)，考察线性方程组 $A\alpha=\lambda\alpha$ 的第 k 个方程 $a_{k1}x_1+a_{k2}x_2+\cdots+a_{kn}x_n=\lambda x_k$，则

$$|\lambda||x_k|=|a_{k1}x_1+a_{k2}x_2+\cdots+a_{kn}x_n|\leqslant|a_{k1}||x_1|+\cdots+|a_{kn}||x_n|$$
$$\leqslant(\sum a_{kj})|x_k|,$$

因此 $|\lambda|\leqslant 1$。

(iii) 只要证明 $|A-I|=0$。

2. (i) 若 $A=0$，则 λ 是 n 重零特征值；

(ii) 若 $A\neq 0$，则零是 $n-1$ 重特征值，另一特征值为 $\sum\limits_{i=1}^{n}a_i^2$。

3. 令 $AB=C$，$BA=D$，C 与 D 的特征多项式为 $\lambda^3(\lambda-2)$。因 $C^2=2C$ 可知它的最小多项式为 $\lambda(\lambda-2)$，但 $D^2\neq 2D$，因此 D 的最小多项式为 $\lambda^2(\lambda-2)$。

4. 可写成 $\lambda I+B$，直接用特征多项式的展开式 $\lambda^{n-1}(nx+\lambda)$。

5. $A=I+B$，用降阶公式。$\lambda=1$ 是 $n-1$ 重特征值，另一个特征值为 $\lambda=1+\sum\limits_{i=1}^{n}a_i^2$。

6. (i) 证明行列式 $|nA-t_rAI|=0$。

(ii) 直接用 Hamilton-Caylny 定理，则

$$A^{-1}=\frac{(-1)^n}{\lambda_0^n}\left\{-A^{n-1}+C_n^1\lambda_0A^{n-2}\cdots(-1)^{n-1}C_n^{n-1}\lambda^{n-1}I\right\}.$$

7. 只要证明矩阵 A 的诸特征值为零，因

$$\sum\lambda_i^2=(\sum\lambda_i)^2-2\sum_{i<j}\lambda_i\lambda_j.$$

8. 主要讨论 $r(A)=n-1$ 或 $r(A)<n-1$ 两种情形去求出 $\mathrm{adj}A$ 的秩。

9. $|A|=\lambda_1\cdots\lambda_n$，再用算术平均值不少于几何平均值的不等式。

10. 由本章命题 4 去做。

11. A 是实阵，因此特征值可能情形为零、负数、正数、复数（但这时是共轭复数成对出现），因此只要排除特征值不可能为零或负的。

12. 用相似于对角型做。

13. 证明 $|I-A|\neq 0$。

14. $AB-BA$ 是反对称阵。

15. 直接应用圆盘定理。

16. 直接应用圆盘定理。

17. 由圆盘定理导出$|\lambda|$的下界。

18. 利用 $|\det A| = |\lambda_1\cdots\lambda_n|$。

19. 由圆盘定理的不等式去考察。

第三章　内积空间与特殊矩阵

A 类

1. 略。

2. 略。

3. $\beta = (4,0,1,-3)$，单位化 $\beta' = \frac{1}{\sqrt{26}}(4,0,-1,3)$。

4. (i) 可有很多解，如 $\alpha_3 = (1,1,1,0)$，$\alpha_4 = (-1,1,0,1)$；

(ii) 例如 $\alpha_3 = (2,3,1,0)$，$\alpha_4 = (1,-1,1,1)$。

5. (i) $\beta_1 = \alpha_1 = (2,3,-4,-6)$，$\beta_2 = (-3,2,6,-4)$，$\beta_3 = (6,5,2,3)$；

　(ii) $\beta_1 = \alpha_1 = (1,1,-1,-2)$，$\beta_2 = (2,5,1,3)$，$\beta_3 = (2,-1,1,0)$。

6. $\gamma_1 = \alpha_1 = (1,i,0)$，单位化 $\beta_1 = \left(\frac{1}{\sqrt{2}},\frac{i}{\sqrt{2}},0\right)$，$\beta_2 = \frac{\sqrt{2}}{6}(1+2i,$
$2-i,2-2i)$。

7. $W^\perp$ 的基为 $\alpha_1 = (-2,1,0,0,0)$，$\alpha_2 = (13,0,-4,1,0)$ $\alpha_3 = (-17,0,$
$5,0,1)$

8. $W^\perp$ 的向量应满足的方程为

$$\begin{cases} -6y_1 + 9y_2 + y_3 = 0 \\ y_2 + y_4 = 0. \end{cases}$$

9. 按集合的等式证明方法去做。

10. 用 Cauchy 不等式，(iii)用(i)的结果，(iv)用(ii)的结果。

11. $A = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} & 0 \\ \frac{2}{3\sqrt{5}} & \frac{4}{3\sqrt{5}} & -\frac{5}{3\sqrt{5}} \end{bmatrix}$，但这解不是唯一的。如

$$A=\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$也是解。

12. (i) $A=\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1-i}{\sqrt{3}} \\ \frac{1+i}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{bmatrix}$,

(ii) $A=\begin{pmatrix} \frac{i}{3} & -\frac{2}{3} & \frac{2i}{3} \\ -\frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \\ -\frac{2}{3\sqrt{5}} & \frac{5i}{3\sqrt{5}} & \frac{-4}{3\sqrt{5}} \end{pmatrix}$.

13. (i) 投影为$\sqrt{14}$，投影向量为$(2,1,3)^T$;

(ii) 投影为 0,投影向量为零向量。

14. $p(1,4,3,8)$。

15. (i) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 2 \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{3}{2}\sqrt{2} \end{bmatrix}$,

$A=QR$,

(ii) $\begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 1 \end{bmatrix}=\begin{bmatrix} \frac{1}{3} & 0 \\ \frac{2}{3} & \frac{1}{\sqrt{2}} \\ \frac{2}{3} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 3 & 3 \\ 0 & \sqrt{2} \end{bmatrix}$, $A=QR$ 这里

$Q^TQ=I$。

16. $N(A)$的基为$\varepsilon(1,-2,-2)$，$\alpha=\gamma+\beta$，$\gamma=\frac{(\alpha,\varepsilon)}{(\varepsilon,\varepsilon)}\varepsilon$，$\beta=\alpha-\gamma$，$\gamma=-\varepsilon$，$\beta=(4,1,1)^T$，因 $\beta\in N(A)^{\perp}$，因此 $\beta\in R(A)^T$。

17. $\alpha_1=(5,-5,-2,-1)$，$\alpha_0=(2,1,1,3)$。

18. (i) $\lambda>0$；(ii) $|\lambda|<\sqrt{\frac{5}{3}}$。

19. $U_1=\frac{1}{\sqrt{3}}\begin{bmatrix} i & \sqrt{2} \\ \sqrt{2} & i \end{bmatrix}$，$U_1^*A_1U_1=\begin{bmatrix} 3\sqrt{2}i & \\ & -3\sqrt{2}i \end{bmatrix}$；

$U_2=\frac{1}{\sqrt{2}}\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$，$U_2^*A_2U_2=\begin{bmatrix} 5 & \\ & 3 \end{bmatrix}$。

20. σ 在自然基下的矩阵为$\begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}$。

21. σ 在自然基下的矩阵$\frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$。

22. β 在 $R(A)$上的最佳近似向量为$\frac{1}{3}(1,1,2)^T$。

23. 最小二乘解 $x=\left(c+\frac{1}{3},c,-3c\right)^T$，$t$ 是参数。

24. A^*A 是 Hermite 阵，但特征值全为零。

25. (ii) 只要直接按 C^n 通常的内积去计算 Ax，A^*x 的长；

(iii) 只要由正规阵酉相似对角阵即可看出。

26. 直接用 Schur 不等式。

27. 直接用实对称阵正交相似对角阵计算 x^TAx，并估计所计算的结果。

28. 直接用正规阵酉相似对角阵，计算 A^*A，AA^*。

29. 直接用正规阵酉相似对角阵去证明。

30. (i)、(ii)、(iv)证明方法同第29题，(iii) 证明 A 的特征值为 0 和 1。

B 类

1. 证明 $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ 是正定矩阵。

2. $\|h(t)\|^2 = 25 + 24\sin 2$。

3. 证明用反证法。

4. $A>0$，则有 $B>0$ 且 $A = B^2 = B^T B$。

$$(Axy)^2 = (B^T Bxy)^2 (BxBy)^2 \leqslant (BxBx)(ByBy) = (B^T Bxx)$$
$$= (B^T Byy) = (Axx)(Ayy)$$

5. $p(y) - p(x_0) = \frac{1}{2}(y - x_0)^T A(y - x_0)$。

6. 因有解 $\Longleftrightarrow b \in R(A)$，即 $b \in N(A^*)^{\perp}$。

7. 事实上，如果有 $\beta \in W$ 且 $(\alpha_0, \beta) \neq 0$，那么存在向量 $\gamma = \frac{(\alpha_0, \beta)}{(\beta, \beta)}\beta$。

计算 $(\alpha_0, \gamma) + (\gamma, \alpha_0) \leqslant (\gamma, \gamma)$，便可导出矛盾。

8. $\Longleftarrow$ 是本章定理 3，$\Longrightarrow$ 由最佳近似定义展开，再用第 7 题的结果。

9. (i) $R(A)$ 的标准正交基 $\alpha_1 = \frac{1}{3}(2,1,2)^T$，$\alpha_2 = \frac{\sqrt{2}}{6}(-1,4,1)^T$；

(ii) $Q = \begin{pmatrix} \frac{2}{3} & -\frac{\sqrt{2}}{6} \\ \frac{1}{3} & \frac{2\sqrt{2}}{3} \\ \frac{2}{3} & -\frac{\sqrt{2}}{6} \end{pmatrix}$，$R = \begin{bmatrix} 3 & \frac{5}{3} \\ & \frac{\sqrt{2}}{3} \end{bmatrix}$，

(iii) $C: x = (9 - 3)^T$。

10. $b \in R(A)$，b 可写成 $b = Ay$，再直接计算。

11. 应用第 8 题。

12. 因 $A = P^T P$，$B = Q^T Q$，于是

$$AB = P^T P Q^T Q = Q^{-1} Q P^T P Q^T Q = Q^{-1}(PQ^T)^T PQ^T Q。$$

这说明矩阵 AB 与矩阵 $(PQ^T)^T PQ^T$ 相似（但后者是正定阵），因此它们的特

征值相同。(AB 不一定是对称阵)

13. $A-aI$ 的特征值为 λ_i-a, $bI-A$ 的特征值为 $b-\lambda_i$。

14. $|\det A|^2=|\lambda_1|^2|\lambda_2|^2\cdots|\lambda_n|^2\leqslant\left(\dfrac{|\lambda_1|^2+|\lambda_2|^2+\cdots+|\lambda_n|^2}{n}\right)^n$,

由 Schure 不等式

$$|\det A|^2\leqslant\left(\frac{\sum\sum|a_{ij}|^2}{n}\right)^n\leqslant\left(\frac{n^2a^2}{n}\right)^n=n^na^{2n}.$$

15. 轴过原点与 $p(1,1,0)$,旋转角 α 为 $\cos\alpha=\dfrac{1}{3}$。

16. $C=\dfrac{1}{3}\begin{pmatrix}-1&2&2\\2&2&-1\\-2&1&-2\end{pmatrix}$, 还可以有别的解。

17. $A=U\begin{bmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{bmatrix}U^*$

将对角阵写成如下分块的形式

$$\begin{pmatrix}\lambda_1&&&&&\\&\ddots&&&&\\&&\lambda_1&&&\\&&&\lambda_r&&\\&&&&\ddots&\\&&&&&\lambda_r\end{pmatrix}=\begin{bmatrix}\lambda_1I_1&&\\&\ddots&\\&&\lambda_rI_r\end{bmatrix}$$

令 $D_1=\begin{bmatrix}I_1&&\\&0&\\&&0\end{bmatrix}\cdots D_r=\begin{bmatrix}0&&\\&\ddots&\\&&\lambda_rI_r\end{bmatrix}$,这样 $A=U\{\lambda_1D_1+\cdots+\lambda_rD_r\}$ U^*,

$A=\lambda_1UD_1U^*+\cdots+\lambda_rUD_rU^*$, $P_i=UD_iU^*$ 容易验证 P_i 满足要求

18. 直接计算

19. (i) 用 $A=UDU^*$ 去计算 A 的奇异值;

(ii) 是(i)的一个结果。

20. A 是反对称阵,它的特征值只可能是纯虚数或零, 因此 $|I+A|\neq0$,再直接验证。

21. $r(A)=r$,这样零特征值的代数重至少为 $n-r$, 因此非零的特征值

个数 S，有 $S \leqslant n-(n-r)=r$：

$$\left|\sum_{i=1}^{n}\lambda_i\right|^2=\left|\sum_{i=1}^{s}\lambda_i\right|^2\leqslant\left(\sum_{i=1}^{s}1^2\right)\left(\sum_{i=1}^{s}|\lambda_i|^2\right)=S\sum_{i=1}^{s}|\lambda_i|^2,$$

这式可写成

$$(t_rA)^2\leqslant S\cdot\sum_{i=1}^{s}|\lambda_i|^2\leqslant S\cdot\sum_{i=1}^{n}|\lambda_i|^2\leqslant S\sum_{i,j}|a_{ij}|^2,$$ 由 Schur 不等式，得

$(t_rA)^2 r t_r(AA^*)$。

第四章 Jordan 标 准 型

A 类

1. $A=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}\oplus\begin{bmatrix}b_{11} & b_{12} & b_{13} & b_{14}\\ b_{21} & b_{22} & b_{23} & b_{24}\\ b_{31} & b_{32} & b_{33} & b_{34}\\ b_{41} & b_{42} & b_{43} & b_{44}\end{bmatrix}\oplus[c_{11}]\oplus[d_{11}]$，

2. 用 σ^{k-1} 作用于线性组合 $c_1\xi_1+c_2\sigma(\xi)+\cdots+c_k\sigma^{k-1}(\xi)=0$，证明各系数为零。

3. 用第 2 题的结果。

4. 主要考察 $\sigma(e_1),\sigma(e_2)$。

(i) $A=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$；

(ii) 求出特征向量：

$\lambda_1=1,\alpha_1=(1,1)$；$\lambda_2=-1,\alpha_2=(1,-1)$；$R^2=[\alpha_1]\oplus[\alpha_2]$；

(iii) 选 R^2 的基为 α_1，α_2，矩阵 $B=\begin{bmatrix}1 & \\ & -1\end{bmatrix}$

(iv) $\sigma_1=\sigma|_{[\alpha_1]}$是恒等变换，$\sigma_2=\sigma|_{[\alpha_2]}$是反射。

5. 直接用乘法便可知道其规律。

6. $A=\begin{bmatrix}7 & 1\\ -49 & -7\end{bmatrix}$，$B=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$，$C=\begin{bmatrix}0 & 2 & 3\\ & 0 & 5\\ & & 0\end{bmatrix}$，

$$D=\begin{bmatrix} 0 & 2 & 3\\ 10 & 0 & 5\\ -15 & 5 & 0\end{bmatrix}。$$

7. (i) 指数为 4，因 A 的零度为 1，故标准型为

$$N=\begin{bmatrix} 0&1&0&0\\ 0&0&1&0\\ 0&0&0&1\\ 0&0&0&0\end{bmatrix}。$$

(ii) 指数为 2，A 的零度为 3，$N=\begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & 1 & 0\end{bmatrix}$

(iii) 指数为 2，A 的零度为 2，$N=\left[\begin{array}{cc|cc} 0 & 0 & & \\ 1 & 0 & & \\ \hline & & 0 & 0\\ & & 1 & 0\end{array}\right]$。

8. 必要性易证，主要证明充分性。因指标相同，因而最大的幂零块的阶数确定，讨论变得简单。

9. 特征多项式为 $(\lambda-2)^5(\lambda-1)^3\lambda^2$，最小多项式为 $(\lambda-2)^3(\lambda-1)^3\lambda$，特征为：

$$\begin{array}{llll}\lambda= & 2 & 1 & 0\\ & \left\{\overbrace{(3,2)}^{5}\right. & \overset{3}{(3)} & \left.\overset{2}{(1,1)}\right\}\end{array}$$

10. 只要先写出该矩阵的特征的可能 $5=2\begin{cases}+2+1\\ +1+1+1\end{cases}$，因此特征为

$\overset{5}{(2,2,1)}$ 或 $\overset{5}{(2,1,1,1,1)}$，即

$$\begin{bmatrix} 2 & 1 & & & \\ 0 & 2 & & & \\ & & 2 & 1 & \\ & & 0 & 2 & \\ & & & & 2 \end{bmatrix}, \quad \begin{pmatrix} 2 & 1 & & & \\ 0 & 2 & & & \\ & & 2 & & \\ & & & 2 & \\ & & & & 2 \end{pmatrix}.$$

11. 只要写出相应矩阵的特征。

(i) $\lambda = \underbrace{2}_{4} \quad \underbrace{3}_{2}$

(2,2) (2)

(2,1,1) (2)

$$\begin{pmatrix} 2 & 1 & & & & \\ 0 & 2 & & & & \\ & & 2 & 1 & & \\ & & 0 & 2 & & \\ & & & & 3 & 1 \\ & & & & 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & & & & \\ 0 & 2 & & & & \\ & & 2 & & & \\ & & & 2 & & \\ & & & & 3 & 1 \\ & & & & 0 & 3 \end{pmatrix};$$

(ii) $\lambda = 7$

(2,2,1)

(2,1,1,1)

$$\begin{pmatrix} 7 & 1 & & & \\ 0 & 7 & & & \\ & & 7 & 1 & \\ & & 0 & 7 & \\ & & & & 7 \end{pmatrix}, \quad \begin{pmatrix} 7 & 1 & & & \\ 0 & 7 & & & \\ & & 7 & & \\ & & & 7 & \\ & & & & 7 \end{pmatrix};$$

(iii) $\lambda = 7$

(3,3,1)

(3,2,1,1)

(3,2,2)

(3,1,1,1,1)

$$\begin{pmatrix} 7 & 1 & 0 & & & & \\ 0 & 7 & 1 & & & & \\ 0 & 0 & 7 & & & & \\ & & & 7 & 1 & 0 & \\ & & & 0 & 7 & 1 & \\ & & & 0 & 0 & 7 & \\ & & & & & & 7 \end{pmatrix}, \quad \begin{pmatrix} 7 & 1 & 0 & & & & \\ 0 & 7 & 1 & & & & \\ 0 & 0 & 7 & & & & \\ & & & 7 & 1 & & \\ & & & 0 & 7 & & \\ & & & & & 7 & \\ & & & & & & 7 \end{pmatrix}$$

$$\begin{pmatrix} 7&1&0&&&& \\ 0&7&1&&&& \\ 0&0&7&&&& \\ &&&7&1&& \\ &&&0&7&& \\ &&&&&7&1 \\ &&&&&0&7 \end{pmatrix}, \quad \begin{pmatrix} 7&1&0&&&& \\ 0&7&1&&&& \\ 0&0&7&&&& \\ &&&7&&& \\ &&&&7&& \\ &&&&&7& \\ &&&&&&7 \end{pmatrix}。$$

(iv) $\lambda=3$ $\lambda=5$ 互相搭配共有四种可能

(2,2) (2,2)

(2,1,1) (2,1,1)

$$\begin{array}{cc} \lambda=3 & 5 \\ (2,2) & (2,2) \end{array} \begin{pmatrix} 3&1&&&&&& \\ 0&3&&&&&& \\ &&3&1&&&& \\ &&0&3&&&& \\ &&&&5&1&& \\ &&&&0&5&& \\ &&&&&&5&1 \\ &&&&&&0&5 \end{pmatrix},$$

$$\begin{array}{cc} \lambda=3 & 5 \\ (2,2) & (2,1,1) \end{array} \begin{pmatrix} 3&1&&&&&& \\ 0&3&&&&&& \\ &&3&1&&&& \\ &&0&3&&&& \\ &&&&5&1&& \\ &&&&0&5&& \\ &&&&&&5& \\ &&&&&&&5 \end{pmatrix};$$

$$\begin{array}{ll} \lambda=3 & 5 \\ (2,1,1) & (2,2) \end{array} \quad \begin{pmatrix} 3 & 1 & & & & & & \\ 0 & 3 & & & & & & \\ & & 3 & & & & & \\ & & & 3 & & & & \\ & & & & 5 & 1 & & \\ & & & & 0 & 5 & & \\ & & & & & & 5 & 1 \\ & & & & & & 0 & 5 \end{pmatrix};$$

$$\begin{array}{ll} \lambda=3 & 5 \\ (2,1,1) & (2,1,1) \end{array} \quad \begin{pmatrix} 3 & 1 & & & & & & \\ 0 & 3 & & & & & & \\ & & 3 & & & & & \\ & & & 3 & & & & \\ & & & & 5 & 1 & & \\ & & & & 0 & 5 & & \\ & & & & & & 5 & \\ & & & & & & & 5 \end{pmatrix}.$$

12. (i) $J=\begin{bmatrix} 1 & 1 & \\ 0 & 1 & \\ & & 2 \end{bmatrix}$, $P=\begin{bmatrix} -1 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 2 & -5 \end{bmatrix}$;

(ii) $J=\begin{bmatrix} 2 & 1 & \\ 0 & 2 & \\ & & 1 \end{bmatrix}$, $P=\begin{bmatrix} 0 & 1 & 1 \\ 3 & -3 & 4 \\ 1 & -1 & -1 \end{bmatrix}$;

(iii) $J=\begin{bmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{bmatrix}$, $P=\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -2 & -1 & -2 \\ 0 & 0 & -11 & -5 \\ -6 & -5 & 0 & 0 \end{bmatrix}$;

(iv) $J=\begin{bmatrix}1&0&0&0\\0&1&1&0\\0&0&1&1\\0&0&0&1\end{bmatrix}$, $P=\begin{bmatrix}1&0&1&0\\-2&1&-2&0\\0&-1&0&1\\-1&0&0&-1\end{bmatrix}$;

(v) $J=\begin{bmatrix}2&1&0\\0&2&1\\0&0&2\end{bmatrix}\oplus\begin{bmatrix}2&1\\0&2\end{bmatrix}\oplus\begin{bmatrix}2&1\\0&2\end{bmatrix}$,

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

13. $J_1=\begin{bmatrix}2&1&0\\0&2&1\\0&0&2\end{bmatrix}$, $P_1=\begin{bmatrix}0&1&0\\-1&-3&0\\1&1&1\end{bmatrix}$;

$J_2=\begin{bmatrix}4&1&\\0&4&\\&&4\end{bmatrix}$, $P_2=\begin{bmatrix}-1&1&2\\0&0&1\\-1&0&0\end{bmatrix}$;

$J_3=\begin{bmatrix}1&1&&\\0&1&&\\&&1&1\\&&0&1\end{bmatrix}$, $P_3=\begin{bmatrix}1&1&1&0\\-1&0&-2&0\\1&0&3&0\\-1&0&-3&1\end{bmatrix}$.

14. 选基 $1,t,t^2$, $T(1)=1$, $T(t)=3+t$, $T(t^2)=(t+3)^2=9+6t+t^2$。所在该基之下的矩阵为 $A=\begin{bmatrix}1&3&9\\0&1&6\\0&0&1\end{bmatrix}$,特征多项式为$(t-1)^3$, 因此最小

多项式为$(t-1)$,$(t-1)^2$,$(t-1)^3$,但 $A-I\neq 0$,因此 T 不能对角化。

15. $$J=\begin{bmatrix} -2 & 1 & & \\ 0 & -2 & & \\ & & -2 & \\ & & & 5 \end{bmatrix},\quad P=\begin{bmatrix} 3 & 1 & 1 & -1 \\ -1 & -1 & -1 & 0 \\ -3 & -1 & -2 & 1 \\ -1 & 0 & 0 & 1 \end{bmatrix},$$

16. $$J=\begin{bmatrix} i & 1 & & \\ 0 & i & & \\ & & -i & 1 \\ & & 0 & -i \end{bmatrix},\quad P=\begin{bmatrix} -2 & 3+i & -2 & 3-i \\ 3-i & -4+i & 3+i & -4-i \\ -3+i & 2-i & -3-i & 2+i \\ 3-i & -1 & 3+i & -1 \end{bmatrix},$$

17. 几何重数$\neq$代数重数。

18. 这时 A 的零化多项式为 x^k-1, 在复数域上它可化为相异的一次因子乘积,所以 A 的最小多项式为最相异的一次因子的乘积。如果k是使 $A^k=I$ 成立的最小自然数,x^k-1 便是最小多项式,因此它的 k 个根便是 1,

$w,w^2\cdots w^{k-1}$, 其中 $w=\cos\dfrac{2\pi}{k}+i\sin\dfrac{2\pi}{k}$。对角阵$D=\operatorname{diag}(1,w,w^2,\cdots,$

$w^{k-1})$。

19. 因为单纯算子有完全的特征向量系 $\alpha_1\cdots\alpha_n$。 用 $\alpha_1\cdots\alpha_n$ 做空间的基便可证明(i)、(ii)。

20. $$J=\begin{bmatrix} 2 & 1 & & \\ 0 & 2 & & \\ & & 2 & \\ & & & 2 \end{bmatrix}\oplus\begin{bmatrix} 3 & & \\ & 3 & \\ & & 3 \end{bmatrix}\oplus\begin{bmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix}\oplus\begin{bmatrix} 5 & 1 & \\ 0 & 5 & \\ & & 5 \end{bmatrix}。$$

B 类

1. 首先 A、B、C 都有相同的特征多项式 $-(\lambda-1)(\lambda-2)^2$, 而 D 的迹与 A、B、C 均不相同,因此只要考察 A、B、C 的情形。由计算 $A-2I$, $B-2I$, $C-2I$ 的秩便可知道对应 $\lambda=2$, Jordan 块的情形。可知 A 与 C 相似。

2. Jordan标准型为$\left[\begin{array}{ccc|c}1&1&0&\\0&1&1&\\0&0&1&\\\hline &&&2\end{array}\right]$. Jordan 基为$\{2e^x, 2xe^x, x^2e^x, e^{2x}\}$

3. 讨论$I+A$的特征值与A的特征值关系,证明A的特征值不为零。

4. 证明充分性。设有复阵P,$P^{-1}AP=B$,把P实部与虚部分开,$P=E+iF$。由$AP=PB$,可得$AE=EB$, $AF=FB$。另一方面$|P|=|e+iF|\neq0$,所以$|E+\lambda F|$不是零多项式,因此必有一λ_0使得$|E+\lambda_0F|\neq0$即$E+\lambda_0F$是非奇异矩阵。令$Q=E+\lambda_0F$,则有

$AQ=A(E+\lambda_0F)=AE+\lambda_0AF=EB+\lambda_0FB=(E+\lambda_0F)B=QB$。即$Q^{-1}AQ=B$,$Q$是实阵。

5. 因为A是4阶方阵,且最小多项式为$(x-1)(x-2)^2$,因此它的特征多项式为$(x-1)^2(x-2)^2$或$(x-1)(x-2)^3$,只要$r(A-I)$的秩确定,便可确定A的标准型或知道A的特征多项式亦可以。

6. σ的特征多项式$p(\lambda)=(\lambda-2)(\lambda^2+1)$亦是$\sigma$的最小多项式。这时$p_1(\lambda)=(\lambda-2)$,$p_2(\lambda)=\lambda^2+1$,$W_1=\mathrm{Ker}\,p_1(\sigma)=\{\xi|\xi\in V, (\sigma-2I)\xi=0\}$,它是对应于特征值为2的特征子空间。它是一维的。$W_2=\mathrm{Ker}\,p_2(\sigma)=\{\xi|\xi\in V, (\sigma^2+1)\xi=0\}$,任取一$\alpha\in W_2$且$\alpha\neq0$,所以$\sigma^2\alpha=-\alpha\neq0\Longrightarrow\sigma\alpha\neq0$。下面证明它们是无关的,$a_1\alpha+a_2\sigma\alpha=0$,用$\sigma$作用两边可得$a_1\sigma\alpha+a_2\sigma^2\alpha=0\Longrightarrow-a_2\alpha+a_1\sigma=0$,因此$a_1=a_2=0$。还要证明$W_1\cap W_2=\{0\}$,如$\beta\in W_1\cap W_2$,$\sigma\beta=2\beta$,$\sigma^2\beta=-\beta$,由前一个式便可知$\sigma^2\beta=2\sigma\beta=4\beta$,所以$4\beta=-\beta\Longrightarrow\beta=0$。因为$\dim W_2\geqslant2$。且$W_1+W_2$是直和,$\dim(W_1+W_2)=\dim W_1+\dim W_2\geqslant3$,$\Longrightarrow\dim W_2=2$。现选基为$W_1$的向量$\alpha$,另$W_2$的基$\beta$,$\sigma\beta$,在此基下矩阵为$\begin{bmatrix}2&0&0\\0&0&-1\\0&1&0\end{bmatrix}$

7. A的秩为1,便可求出A的特征值。再讨论特征值的情况(设$A=xy^T$,求A的特征值)。

8. $a \neq 0$, $J = \left[\begin{array}{cc|c} 2 & 1 & \\ 0 & 2 & \\ \hline & & -1 \end{array}\right]$; $a=0$ 时，A 的最小多项式为 $(\lambda-2)(\lambda+1)$，

$J = \begin{bmatrix} 2 & & \\ & 2 & \\ & & -1 \end{bmatrix}$，由此可见 $a=0$ 是 A 可对角化的充要条件。

9. 用 $A = P^{-1}JP$，而 $J = D + N$，其中 D 是对角化矩阵，N 是幂零阵。

10. 设 A 的 n 个特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$，则

$$\lambda_1^2 + \lambda_2^2 + \cdots + \lambda_n^2 = (\lambda_1 + \lambda_2 + \cdots + \lambda_n)^2 - 2\sum_{i \neq j} \lambda_i \lambda_j = 0,$$

上式说明 A^2 的诸特征值之和为零或 $t_r A^2 = t_r(A^* A) = 0 \Longrightarrow A = 0$。

11. 将 A 的 Jordan 标准型求出再做。

12. 将 A 的 Jordan 标准型求出 $A^{100} = PJ^{100}P^{-1}$，$J^{100} = I \Longrightarrow A^{100} = I$。

13. (i) $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 是 V_3 的一组基，定义 $\sigma\varepsilon_1 = \varepsilon_2, \sigma\varepsilon_2 = \varepsilon_1, \sigma\varepsilon_3 = 0$，便满足要求。

(ii) 定义 $\sigma(\varepsilon_1) = 2\varepsilon_1, \sigma(\varepsilon_2) = 2\varepsilon_2, \sigma(\varepsilon_3) = 0$，便满足要求。
解决这个问题只要从 Jordan 标准型去考虑即可。

14. $|A + I| = 1$。

15. $|A + I| = 2^r$，这题与上题都用 Jordan 标准型做。

16. 每一个特征值只对应一个 Jordan 块。

第五章　矩阵分析初步

A 类

1. $\lim\limits_{k \to \infty} A_k = \begin{bmatrix} 0 & 1 \\ 2 & e^{-2} \end{bmatrix}$。

2. $A_k=\begin{bmatrix} 1 & \frac{1}{k} \\ 1 & \frac{1}{k^2} \end{bmatrix}$。

3. 不一定。如 $A_k=\begin{bmatrix} k & \\ & k \end{bmatrix}$，$B_k=\begin{bmatrix} \frac{1}{k} & \\ & \frac{1}{k} \end{bmatrix}$。

4. (i) A 的特征多项式为 $f(\lambda)=\lambda^2-1.3\lambda+0.37$，特征值为

$$\lambda=\frac{1}{2}(1.3\pm\sqrt{1.69-1.48})=\frac{1}{2}(1.3\pm\sqrt{0.21}),|\lambda|$$

<1 故矩阵 A 为幂收敛。

(ii) $f(\lambda)=(\lambda-0.2)^2(\lambda-1)$，$|\lambda|\leqslant 1$，但 $\lambda=1$ 时，Jordan 块是一阶，故矩阵 A 是幂收敛。

5. 因 A 是幂收敛的，$L^2=(\lim A^n)(\lim A^n)=\lim A^{2n}=L$。

6. (i) 用数学归纳法证明 $A^k=\begin{bmatrix} C^k & O \\ D(C^{k-1}+C^{k-2}+\cdots+I) & I \end{bmatrix}$，

(ii) 利用(i)的结果，得

$$C=\begin{bmatrix} 0.1 & 0.3 \\ 0.7 & 0.6 \end{bmatrix},D=\begin{bmatrix} 0.1 & 0 \\ 0.1 & 0.1 \end{bmatrix}.$$

因而

$$|\lambda I-C|=\lambda^2-0.7\lambda-0.15,|\lambda|<1,$$

C 是幂收敛，所以 $\lim C^k=0$，故

$$\lim(C^{k-1}+\cdots+C+I)=(I-C)^{-1}=\begin{bmatrix} 0.9 & -0.3 \\ -0.7 & 0.4 \end{bmatrix}^{-1}$$

$$=\frac{2}{3}\begin{bmatrix} 4 & 3 \\ 7 & 9 \end{bmatrix},$$

$$D(1-C)^{-1}=\frac{1}{15}\begin{bmatrix} 4 & 3 \\ 11 & 12 \end{bmatrix},$$

$$\lim A_k=\begin{bmatrix}0&0&0&0\\0&0&0&0\\\frac{4}{15}&\frac{2}{15}&1&0\\\frac{11}{15}&\frac{12}{15}&0&1\end{bmatrix}。$$

7. 令 $\frac{A}{2}=B$，$(\lambda I-B)=\left(\lambda-\frac{1}{2}\right)^2$ $|\lambda|<1$，因为 B 是幂收敛的。原级数和为 $(I-B)^{-1}=\begin{bmatrix}4&-1\\4&0\end{bmatrix}$。

8. 主要考虑该矩阵 A 的 Jordan 标准型可能的情形。它的特征值分别为 $1,\frac{1}{2},\frac{1}{4}$。进一步考虑 $\lambda=1$ 时 Jordan 块的阶。因 $r(A-I)=3$，所以特征向量有 2 个，因此 $\lambda=1$ 的 Jordan 块都是一阶的，因而是幂收敛。反之亦易证。

9. $$I+A^2=\begin{bmatrix}17&0&0\\8&17&0\\1&8&17\end{bmatrix}\oplus\begin{bmatrix}2&0&0&0\\-2&2&0&0\\1&-2&2&0\\0&1&-2&2\end{bmatrix}.$$

10. $f(A)=$

$$P\left[\begin{bmatrix}f(2)&0&0\\f'(2)&f(2)&0\\\frac{f''(2)}{2!}&f'(2)&f(2)\end{bmatrix}\oplus\begin{bmatrix}f(1)&0&0&0\\f'(1)&f(1)&0&0\\\frac{f''(1)}{2!}&f'(1)&f(1)&0\\\frac{f'''(1)}{3!}&\frac{f''(1)}{2!}&f'(1)&f(1)\end{bmatrix}\right]P^{-1}.$$

11.

$$\sum_{k=0}^{\infty}\frac{A^k}{4^k}=$$

$$P\left(\begin{bmatrix}\frac{8}{7} & 0\\ \frac{16}{49} & \frac{8}{7}\end{bmatrix}\oplus\left[\frac{8}{7}\right]\oplus\begin{pmatrix}\frac{4}{5} & 0 & 0 & 0 & 0\\ \frac{4}{25} & \frac{4}{5} & 0 & 0 & 0\\ \frac{4}{125} & \frac{4}{25} & \frac{4}{5} & 0 & 0\\ \frac{4}{625} & \frac{4}{125} & \frac{4}{25} & \frac{4}{5} & 0\\ \frac{4}{3125} & \frac{4}{625} & \frac{4}{125} & \frac{4}{25} & \frac{4}{5}\end{pmatrix}\right)P^{-1}.$$

12. $e^J=\begin{bmatrix}e^2 & 0 & 0 & 0\\ 0 & e & e & 0\\ 0 & 0 & e & 0\\ 0 & 0 & 0 & e^6\end{bmatrix}$, $\sin J=\begin{bmatrix}\sin 2 & 0 & 0 & 0\\ 0 & \sin 1 & \cos 1 & 0\\ 0 & 0 & \sin 1 & 0\\ 0 & 0 & 0 & \sin 6\end{bmatrix}$,

$$\cos J=\begin{bmatrix}\cos 2 & 0 & 0 & 0\\ 0 & \cos 1 & -\sin 1 & 0\\ 0 & 0 & \cos 1 & 0\\ 0 & 0 & 0 & \cos 6\end{bmatrix}.$$

13. $e^J=\begin{bmatrix}e^2 & e^2 & \frac{e^2}{2!}\\ 0 & e^2 & e^2\\ 0 & 0 & e^2\end{bmatrix}\oplus\begin{bmatrix}e^3 & e^3\\ 0 & e^3\end{bmatrix}\oplus[e^{-1}]$,

$$\sin J=\begin{bmatrix}\sin 2 & \cos 2 & -\frac{\sin 2}{2!}\\ 0 & \sin 2 & \cos 2\\ 0 & 0 & \sin 2\end{bmatrix}\oplus\begin{bmatrix}\sin 3 & \cos 3\\ 0 & \sin 3\end{bmatrix}\oplus[-\sin 1],$$

$$\cos J=\begin{bmatrix}\cos 2 & -\sin 2 & -\frac{\cos 2}{2!}\\ 0 & \cos 2 & -\sin 2\\ 0 & 0 & \cos 2\end{bmatrix}\oplus\begin{bmatrix}\cos 3 & -\sin 3\\ 0 & \cos 3\end{bmatrix}\oplus[\cos 1].$$

14. 仿例5。

$$|\lambda I-A|=\lambda(\lambda-1),A^2=A,e^A=I+(e-1)A=\begin{bmatrix}e & e-1\\ 0 & 1\end{bmatrix},$$

$$e^B=\begin{bmatrix} e & -(e-1) \\ 0 & 1 \end{bmatrix},$$

$$A+B=\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix},\ e^{A+B}=\begin{bmatrix} e^2 & 0 \\ 0 & 1 \end{bmatrix};$$

$$e^A\cdot e^B=\begin{bmatrix} e^2 & -(e-1)^2 \\ 0 & 1 \end{bmatrix}.$$

16. $|\lambda I-A|=(\lambda-3)^2$,

$$e^A=e^3(A-2I)=\begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}\cdot e^3,$$

$$\sin Z=\sin(Z-3+3)=\sin 3\cdot\cos(Z-3)+\cos 3\sin(Z-3)$$

$$=\sin 3\left\{\ 1-\frac{1}{2}(Z-3)^2+\cdots\right\}$$

$$+\cos 3\left\{(Z-3)-\frac{1}{3!}(Z-3)^3+\cdots\right\}$$

注意到特征多项式即可有

$$\sin A=I\cdot\sin 3+(A-3I)\cos 3=\begin{bmatrix} \cos 3+\sin 3 & -\cos 3 \\ \cos 3 & \sin 3-\cos 3 \end{bmatrix}.$$

同理可得 $\cos A=I\cos 3-\sin 3(A-3I)$

$$=\begin{bmatrix} \cos 3-\sin 3 & \sin 3 \\ -\sin 3 & \cos 3+\sin 3 \end{bmatrix}.$$

17. $$\frac{\mathrm{d}A(t)}{\mathrm{d}t}=\begin{bmatrix} -\sin t & \cos t \\ -\cos t & -\sin t \end{bmatrix},$$

$$\frac{\mathrm{d}A^{-1}(t)}{\mathrm{d}t}=\begin{bmatrix} -\sin t & -\cos t \\ \cos t & -\sin t \end{bmatrix},$$

$$\frac{\mathrm{d}}{\mathrm{d}t}|A(t)|=0,\quad \left|\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right|=1.$$

18.

$$\int_0^t A(s)\,ds=\begin{bmatrix}\frac{1}{2}(e^{2t}-1) & t(e^t-1)+1 & \frac{t^2}{2}+t\\ \frac{1}{2}(1-e^{-2t}) & e^{2t}-1 & 1-\cos t\\ \frac{3}{2}t^2 & 0 & \frac{t^2}{2}\end{bmatrix}。$$

19．(i) ① $(\sigma-2I)$的零度为3，Jordan 标准型有二种可能：

$J=J_{(3)}(2)\oplus J_{(3)}(2)\oplus J_{(1)}(2)$；$J=J_{(3)}(2)\oplus J_{(2)}(2)\oplus J_{(1)}(2)$；

② $(\sigma-2I)$的零度为4，Jordan 标准型为：

$J=J_{(3)}(2)\oplus J_{(2)}(2)\oplus J_{(1)}(2)\oplus J_{(1)}(2)$。

③ $(\sigma-2I)$的零度为5，Jordan 标准型为：

$J=J_{(3)}(2)\oplus J_{(1)}(2)\oplus J_{(1)}^{*}(2)\oplus J_{(1)}(2)\oplus J_{(1)}(2)$。

(ii) ① $$e^J=\begin{bmatrix}e^2 & e^2 & \frac{e^2}{2!}\\ 0 & e^2 & e^2\\ 0 & 0 & e^2\end{bmatrix}\oplus\begin{bmatrix}e^2 & e^2 & \frac{e^2}{2!}\\ 0 & e^2 & e^2\\ 0 & 0 & e^2\end{bmatrix}\oplus[e^2];$$

$$e^J=\begin{bmatrix}e^2 & e^2 & \frac{e^2}{2!}\\ 0 & e^2 & e^2\\ 0 & 0 & e^2\end{bmatrix}\oplus\begin{bmatrix}e^2 & e^2\\ 0 & e^2\end{bmatrix}\oplus\begin{bmatrix}e^2 & e^2\\ 0 & e^2\end{bmatrix}.$$

② $$e^J=\begin{bmatrix}e^2 & e^2 & \frac{e^2}{2!}\\ 0 & e^2 & e^2\\ 0 & 0 & e^2\end{bmatrix}\oplus\begin{bmatrix}e^2 & e^2\\ 0 & e^2\end{bmatrix}\oplus[e^2]\oplus[e^2].$$

③ $$e^J=\begin{bmatrix}e^2 & e^2 & \frac{e^2}{2!}\\ 0 & e^2 & e^2\\ 0 & 0 & e^2\end{bmatrix}\oplus[e^2]\oplus[e^2]\oplus[e^2]\oplus[e^2].$$

(iii) e^{Jt} 按本章的命题4公式去做。

20. $$x(t)=e^{At}c=0\begin{bmatrix}0 & -1 & -1\\ 3 & 3 & 4\\ 1 & 1 & 1\end{bmatrix}\begin{bmatrix}e^t & 0 & 0\\ 0 & e^{2t} & te^{2t}\\ 0 & 0 & e^{2t}\end{bmatrix}\begin{bmatrix}1 & 0 & 1\\ -1 & -1 & 3\\ 0 & 1 & -3\end{bmatrix}$$

$$\begin{bmatrix}c_1\\ c_2\\ c_3\end{bmatrix}=\begin{bmatrix}e^{2t}+te^{2t} & te^{2t} & -3te^{2t}\\ 3e^t-3e^{2t}-4te^{2t} & e^{2t}-4te^{2t} & 3e^t-3e^{2t}+12te^{2t}\\ e^t-e^{2t}-te^{2t} & -te^{2t} & e^t+3te^{2t}\end{bmatrix}$$

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}。$$

B 类

1. $LA=L, LA=AL$。(由 $L=\lim A^n$ 即可证明),因此有等式 $(A-L)^n=A^n-L$(可用数学归纳法做)。

2. 由第 1 题的结果即可知 $A^n-L=(A-L)^n$,这样便可得出。

3. 实际上由级数 $\sum\limits_{n=0}^{\infty} Z^n=(1-Z)^{-1}$, 通过两边求导,便可证明等式 $\sum\limits_{n=0}^{\infty} nZ^n=Z(1-Z)^n$,收敛半径仍为 1。

4. 特征值 $\lambda=\dfrac{1-\sqrt{2}}{3}, \dfrac{1+\sqrt{2}}{3}$,则矩阵级数收敛。

$$令 D=\begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix},其和为 (I-D)^{-1}=3\begin{bmatrix} 1 & 1 \\ \frac{1}{2} & 1 \end{bmatrix}。$$

5. $\sin A=\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}\oplus\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$。

6. $A=BB^T=\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}[1\ 2\ 1]$可计算特征值0,0,6。$\rho(A)=6$。由于可通过对角化,因此 $\lim\left(\dfrac{A}{\rho(A)}\right)^n$ 是存在的,将 A 通过正交变换化为对角形,再做 $\lim\left(\dfrac{A}{\rho(A)}\right)^n=\dfrac{1}{6}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$。

7. 设 $y(t)=x(t)-z(t)$,这样 $y(t)$ 是满初始值问题。$\dot{y}(t)=Ay(t)$,$y(t_0)=x(t_0)-z(t_0)$ 的解。$y(t)=e^{A(t-t_0)}(x(t_0)-z(t_0))$。

8. 应用上题结果。一个解为 $x=e^{2t}\ v$, v 是待定向量, 可解得 $v=$

$\left(1, -\frac{4}{5}, \frac{7}{5}\right)^T$

9. $x(t)=C_1e^{7t}v^1+C_2e^{-6t}v^2, v^1=(2,1)^T, v^2=(-2,1)^T$，再用初始值定出常数 C_1, C_2。

10. (i) 只要直接用 $x(t)=e^{\alpha t}b$,代入微分方程即可得出等式；

(ii) 用(i)的结果。

11. $\dot{x}(t)=b_1+2tb_2+3t^2b_3+\cdots+mt^{m-1}b_m$ 代入微分方程,得

$\dot{x}(t)=Ab_0+tAb_1+t^2Ab_2+\cdots+t^m(Ab_m+c)$

由比较系数可得 $b_1=Ab_0, 2b_1=Ab_2, \cdots, mb_m=Ab_{m-1}$,根据初始条件可知 $b_0=b$,便可依次定出各个系数。

12. 亦是用解代入方程,再由初始条件待定。

13. $A=(e^{At})_t{}'_{=0}=\begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ 3 & 3 & -1 \end{bmatrix}$。

14. $x=(x_1(t),x_2(t))^T$(用常数变易法得的公式)。

$x_1(t)=x_1(0)\cos t+x_2(0)\sin t+\sin t-2\cos t+2;$

$x_2(t)=-x_1(0)\sin t+x_2(0)\cos t+\cos t+2\sin t-1$。

15. $\arcsin\frac{A}{4}=\frac{1}{6}\begin{bmatrix} \pi-2\sqrt{3} & -\sqrt{3} \\ 4\sqrt{3} & \pi+2\sqrt{3} \end{bmatrix}$。

16. (i) A 的最小多项式为 $m(\lambda)=\lambda(\lambda+2)$，与 $g(\lambda)=e^{\lambda t}$ 谱上一致的多项式为 $a_0(t)+a_1(t)\lambda$。$\lambda_1=0, \lambda_2=-2$ 代入前式，可解得 $a_0(t)=1$，

$a_1(t)=-\frac{1}{2}(e^{-2t}-1)$,则

$$e^{At}=I-\frac{1}{2}(e^{-2t}-1)A.$$

(ii) A 的最小多项式为 $(\lambda-2)^2(\lambda-1)$,则

$a_0(t)=4e^t-3e^{2t}+2te^{2t}$

$a_1(t)=-4e^t+4e^{2t}-3te^{2t}$,

$a_2(t)=e^t-e^{2t}+te^{2t}$

$e^{At}=a_0(t)I+a_1(t)A+a_2(t)A^2$。

由 $\begin{cases} a_0(t)+2a_1(t)+4a_2(t)=\sin t \\ a_1(t)+4a_2(t)=t\cos 2t, \\ a_0(t)+a_1(t)+a_2(t)=\sin t \end{cases}$ 求出 $a_0(t), a_1(t), a_2(t)$。

故
$$\sin At = a_0(t)I + a_1(t)A + a_2(t)A^2。$$

17. (i) 因 $A=-A'$, $e^A\cdot(e^A)'=e^Ae^{A'}=e^A\cdot e^{-A}=I$。

(ii) 类似(i)证明。

第六章 广义逆矩阵

1. (i) $r(A)=2$，行满秩故有左逆，$A_R^{-1}=\frac{1}{15}\begin{bmatrix} -7 & 5 \\ -14 & 10 \\ 10 & -5 \end{bmatrix}$；

(ii) $r(A)=2$，列满秩故有左逆，$A_L^{-1}=\frac{1}{149}\begin{bmatrix} -62 & -42 & 59 \\ 49 & 38 & -25 \end{bmatrix}$，

(iii) $r(A)=1$，既非行满秩又非列满秩，斯 A 既无右逆又无左逆。

2. $R(BC)\subseteq R(B)$，又 $r(BC)=r(B)$，因此有 $R(BC)=R(B)$，这说明矩阵 B 的列，可通过矩阵 BC_L 的列组合而得，因此 $B=BCD$。第二个结论只须直接验证。

3. 方法同第 2 题。

4. (i) $PA(A^-P_L^{-1})PA=PA$,

(ii) 同(i)验证。

5. 只要直接验证。

6. 只要直接验证。

7. $BB^-B=\left[\frac{C}{D}\right](C_R^{-1}|0)\left[\frac{C}{D}\right]=\left[\frac{C}{DC_R^{-1}C}\right]$。由于 $D=WC$，

所以
$$DC_R^{-1}C=WCC_R^{-1}C=WC=D,$$
故命题成立。后一结论亦只要直接验证。

8. A 的第三行加到第二行去，因为

$$B=\begin{bmatrix}1&1&1&0\\0&0&-1&0\\1&1&0&0\end{bmatrix}=\begin{bmatrix}C\\ \hline D\end{bmatrix},$$ 其中 $C=\begin{bmatrix}1&1&1&0\\0&0&-1&0\end{bmatrix}$，$D=(1,1,0,0)$，

则

$$B^-=\frac{1}{2}\begin{bmatrix}1&1&0\\1&1&0\\0&-2&0\\0&0&0\end{bmatrix},$$

所以

$$A^-=B^-P=\frac{1}{2}\begin{bmatrix}1&1&0\\1&1&0\\0&-2&0\\0&0&0\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&1\\0&0&1\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1&1&1\\1&1&1\\0&-2&-2\\0&0&0\end{bmatrix}.$$

9．用与第 7 题类似的方法。

10．$C=\begin{bmatrix}1&0\\0&2\\2&0\end{bmatrix}$，$D=\begin{bmatrix}2\\0\\-4\end{bmatrix}$，$A^-=\dfrac{1}{10}\begin{bmatrix}2&0&4\\0&0&0\\0&5&0\end{bmatrix}$。

11．A_1，交换二、三行，再用第 7 题的方法

$$A_1^-=\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&0&0&0\end{bmatrix}.$$

A_2，同时用行与列的交换变为 $B=\left[\begin{array}{cc|cc}2&1&2&0\\1&1&1&0\\ \hline 1&1&1&0\end{array}\right]$，所以

$$A_2^-=\begin{bmatrix}0&-1&1\\0&0&0\\0&0&0\\0&2&-1\end{bmatrix},$$

$A_3^-=\begin{bmatrix}1&0&0\\0&0&0\end{bmatrix}$，$A_4^-=\dfrac{1}{3}\begin{bmatrix}2&-1\\-1&2\\0&0\end{bmatrix}$。

12．$A_1x=b_1$ 的通解为 $x=A_1^-b_1+(I-A_1^-A_1)Y$，

$$A_1^- b_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1-2c_3 \\ 0 \\ c_3 \end{bmatrix};$$

$$A_2^- b_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, x = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} 1-c_3 \\ c_2 \\ c_3 \\ 0 \end{bmatrix},$$

13．因 $R(G)=R(A)$，所以有 $G^{(i)}=A\alpha_i(i=1,2,\cdots,n)$，$A\alpha_i=A^2\alpha_i=AG^{(i)}$，$AG=G$，$G^2=(AGAG)=AG=G$。

14．A 是行满秩 $A_m^-=A_R^{-1}=\frac{1}{2}\begin{bmatrix} 1 & 2 \\ 0 & -2 \\ -1 & -2 \end{bmatrix}(A^T(AA^T)^{-1})$，最小范数解为：

$$x=A_m^- b=\frac{1}{2}\begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}=\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}。$$

15．$(A^TA)^-=\begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$，$A^Tb=\begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix}$，

$$(A^TA)^-A^TA=\begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

故最小二乘解为：

$$\begin{aligned} x &= (A^TA)^-A^Tb+(I-(A^TA)^-A^TA)y \\ &= \begin{bmatrix} 1-2y_2-y_3 \\ y_2 \\ y_3 \end{bmatrix}。\end{aligned}$$

16．要逐个验证 Penrose-Moore 的四个矩阵方程。

G_1 是 A 的反射 g 逆，$G_2=A^+$。

17．由 Penrose-Moore 四个矩阵方程去证明。

18．作 A 的满秩分解 $A=CD$。$A^+=D_R^{-1}C_L^{-1}=D^*(DD^*)^{-1}(C^*C)^{-1}C^*$。

(i) $A^{+}=\frac{1}{10}\begin{bmatrix}-2 & 1 & 2\\ 1 & 3 & -1\\ -1 & 2 & 1\end{bmatrix}$；(ii) $A^{+}=\frac{1}{12}\begin{bmatrix}1 & 1 & 0 & 1\\ 6 & 2 & 4 & -2\\ 1 & -1 & 0 & -1\end{bmatrix}$。

$$
\text{(iii)}\ A^{+}=\frac{1}{9300}\begin{bmatrix}54 & 420 & 162\\ 108 & 840 & 324\\ -126 & 570 & -378\\ 252 & -1140 & 756\\ 30 & 750 & 90\end{bmatrix}.
$$

19．$Ax=b_1$ 是矛盾方程，A 是列满秩，所以

$$
A^{+}=A_{L}^{-1}=(A^{T}A)^{-1}A^{T}=\frac{1}{2}\begin{bmatrix}0 & 0 & 1\\ 1 & 1 & -2\end{bmatrix},
$$

$$
x=A^{+}b_1=\frac{1}{2}\begin{pmatrix}1\\ 0\end{pmatrix},
$$

这个解是最小二乘最小范数解。

$Ax=b_2$ 是相容方程，$x=A_m^{-}b_2=\begin{pmatrix}0 & 0 & \frac{1}{2}\\ 1 & 0 & -1\end{pmatrix}\begin{bmatrix}1\\ 1\\ 0\end{bmatrix}=\begin{bmatrix}0\\ 1\end{bmatrix}$.

这是最小范数解。

20．因 A 是 $r(A)=1$，故可作 $A=\begin{bmatrix}a_1\\ a_2\\ \vdots\\ a_m\end{bmatrix}[b_1\ \ b_2\cdots b_n]$，再用公式

$$
A^{+}=D_R^{-1}C_L^{-1}=D^{*}(DD^{*})^{-1}(C^{*}C)^{-1}C^{*}
$$

去计算。