线性代数 (B1)

习题课讲义

作者：助教-崔平凡、李浩哲、刘玉涵

时间：2024 年秋季学期

版本：4.3

我看见极光撕开夜空，星星从间隙里洒下，洒在我的怀里，洒满我的眼底。

前言

内容介绍

本讲义是 2024 年秋季学期叶郁老师线性代数 (B1) 班级习题课讲义，由课程助教制作，由于助教水平有限，可能有所疏漏，若发现讲义中的错误与不足之处，欢迎与助教联系交流[1]。

线性代数 (B1) 作为一门基础的数学类通修课程，历来饱受科大学子诟病，原因多集中于其教材编写较差、所授内容不足以满足部分专业后续的学习需求等等，因此也导致许多人认为线性代数仅仅是套公式与概念的暴力计算，但实际上线性代数是一门非常优美的学科，许多概念都可以从良好的几何直观出发，再推广到一般意义下。故本份讲义力求在合适范围内补充一些深入且实用的内容，同时尽可能从几何直观角度对一些概念进行解释，以求加深同学们对线性代数本质与作用的理解。

关于课外资料

(1)《线性代数》——李尚志

本书为笔者修读线性代数 (B2) 课程时陈发来老师班级所采用的教材。此教材内容充实，我们的课程内容可被该教材真包含，且大部分内容叙述清晰、例题丰富，比较适合数学专业初学者学习，可以作为较好的课外补充，学有余力者可以参考。此外该教材有一本配套的《线性代数学习指导》，内含该教材的全部课后习题答案。

(2)《线性代数》——李炯生

本书为笔者修读线性代数 (B2) 课程时陈洪佳老师班级所采用的教材，本书难度较大，号称“亚洲第一难书”，整本书大部分采用形式化叙述，几乎没有具体数值算例，部分课后习题颇有难度，因此可能不适合初学者学习。此外在网上可以找到一部分该教材课后习题的解答。

(3)《线性代数的本质》系列视频——3Blue1Brown

该系列视频制作精良，从几何直观角度对线性代数中各个重要的基本概念作出了阐释，非常有利于初学者对线性代数构建一个初步的认知体系，且视频并不长，闲暇时间便可看完（同时也推荐 3Blue1Brown 的其他视频）。

(4)《线性代数与解析几何》——陈发来等

本书为 2023 秋之前线性代数 (B1) 课程所用教材，该教材编写较差，内容与现行讲义基本一致，但其实个人认为对于非数学专业或仅考虑应试的同学，这本书还是足够用的。

笔记 上述相关课外资料已经上传至课程群，大家可以自行查阅

——刘玉涵 于 2024 年 9 月

[1]邮箱:liuyh594@mail.ustc.edu.cn　QQ:1820635673

目录

第 1 章 作业答案

1.1 第二周作业

习题 1.1 (第二章第 2 题)

当 a 为何值时，下列线性方程组有解？有解时求出它的通解.

$$(1)\quad \begin{cases} 3x_1+2x_2+x_3=2 \\ x_1-x_2-2x_3=-3 \\ ax_1-2x_2+2x_3=6 \end{cases} \qquad (2)\quad \begin{cases} x_1-4x_2+2x_3=-1 \\ -x_1+11x_2-x_3=3 \\ 3x_1-5x_2+7x_3=a \end{cases}$$

解 (1). $\begin{pmatrix} 3 & 2 & 1 & 2 \\ 1 & -1 & -2 & -3 \\ a & -2 & 2 & 6 \end{pmatrix} \xrightarrow[-3r_1\to r_2,-ar_1\to r_3]{r_1\leftrightarrow r_2} \begin{pmatrix} 1 & -1 & -2 & -3 \\ 0 & 5 & 7 & 11 \\ 0 & a-2 & 2a+2 & 3a+6 \end{pmatrix} \xrightarrow[(2-a)r_2\to r_3]{\frac{1}{5}r_2} \begin{pmatrix} 1 & -1 & -2 & -3 \\ 0 & 1 & \frac{7}{5} & \frac{11}{5} \\ 0 & 0 & \frac{3a+24}{5} & \frac{4a+52}{5} \end{pmatrix}.$

故当 $\dfrac{3a+24}{5}=0$且$\dfrac{4a+52}{5}=0$, 或 $\dfrac{3a+24}{5}\neq 0$时, 方程组有解, 易知前者不成立. 故 $\dfrac{3a+24}{5}\neq 0$, 即$a\neq -8$, 此时:

$$\begin{pmatrix} 1 & -1 & -2 & -3 \\ 0 & 1 & \frac{7}{5} & \frac{11}{5} \\ 0 & 0 & \frac{3a+24}{5} & \frac{4a+52}{5} \end{pmatrix} \xrightarrow[-\frac{7}{5}r_3\to r_2,2r_3\to r_1]{\frac{5}{3a+24}r_3} \begin{pmatrix} 1 & -1 & 0 & \frac{-a+32}{3a+24} \\ 0 & 1 & 0 & \frac{a-20}{3a+24} \\ 0 & 0 & 1 & \frac{4a+52}{3a+24} \end{pmatrix} \xrightarrow{r_2\to r_1} \begin{pmatrix} 1 & 0 & 0 & \frac{4}{a+8} \\ 0 & 1 & 0 & \frac{a-20}{3a+24} \\ 0 & 0 & 1 & \frac{4a+52}{3a+24} \end{pmatrix}.$$

故方程组的解为 $X=(x_1,x_2,x_3)^T=(\frac{4}{a+8},\frac{a-20}{3a+24},\frac{4a+52}{3a+24})^T\ (a\neq -8)$.

(2). $\begin{pmatrix} 1 & -4 & 2 & -1 \\ -1 & 11 & -1 & 3 \\ 3 & -5 & 7 & a \end{pmatrix} \xrightarrow[-3r_1\to r_3]{r_1\to r_2} \begin{pmatrix} 1 & -4 & 2 & -1 \\ 0 & 7 & 1 & 2 \\ 0 & 7 & 1 & a+3 \end{pmatrix} \xrightarrow{-r_2\to r_3} \begin{pmatrix} 1 & -4 & 2 & -1 \\ 0 & 7 & 1 & 2 \\ 0 & 0 & 0 & a+1 \end{pmatrix}.$

故只有 $a+1=0$ 即 $a=-1$ 时，方程组有解，此时原方程组等价于:

$$\begin{cases} x_1-4x_2+2x_3=-1 \\ 7x_2+x_3=2 \end{cases} \quad 令x_2=t，解得: \quad \begin{cases} x_1=18t-5 \\ x_2=t \\ x_3=-7t+2 \end{cases}$$

故方程组的解为 $X=(x_1,x_2,x_3)^T=(18t-5,t,-7t+2)^T\ (t\in\mathbb{F})$.

习题 1.2 (第二章第 4 题)

求三次多项式 $f(x)=ax^3+bx^2+cx+d$, 使 $y=f(x)$ 的图象经过以下 4 个点:

$$A(1,2),B(-1,3),C(3,0),D(0,2).$$

解 将四个点代入 $y=f(x)$，可得:

$$\begin{cases} a+b+c+d=2 \\ -a+b-c+d=3 \\ 27a+9b+3c+d=0 \\ d=2 \end{cases} \quad 即 \quad \begin{cases} a+b+c=0 \\ -a+b-c=1 \\ 27a+9b+3c=-2 \\ d=2 \end{cases}$$

故解以 a,b,c 为未知数的方程组即可

$$\begin{pmatrix}1&1&1&0\\-1&1&-1&1\\27&9&3&-2\end{pmatrix}\xrightarrow[-27r_1\to r_3]{r_1\to r_2}\begin{pmatrix}1&1&1&0\\0&2&0&1\\0&-18&-24&-2\end{pmatrix}\xrightarrow[\frac{1}{2}r_2,-\frac{1}{24}r_3]{9r_2\to r_3}\begin{pmatrix}1&1&1&0\\0&1&0&\frac{1}{2}\\0&0&1&-\frac{7}{24}\end{pmatrix}\xrightarrow[-r_3\to r_1]{-r_2\to r_1}\begin{pmatrix}1&0&0&-\frac{5}{24}\\0&1&0&\frac{1}{2}\\0&0&1&-\frac{7}{24}\end{pmatrix}$$

故 $f(x)=-\dfrac{5}{24}x^3+\dfrac{1}{2}x^2-\dfrac{7}{24}x+2$.

习题 1.3 (第二章第 5 题)

求三次多项式 $f(x)=ax^3+bx^2+cx+d$ 满足

$$f(0)=1,f(1)=2,f'(0)=1,f'(1)=-1.$$

解 由 $f'(x)=3ax^2+2bx+c$，将四个点代入可得:

$$\begin{cases}a+b+c+d=2\\3a+2b+c=-1\end{cases}\quad 及\ c=d=1\quad 即\quad \begin{cases}a+b=0\\3a+2b=-2\end{cases}\quad 及\ c=d=1$$

故解以 a,b 为未知数的方程组即可，易见 $a=-2,b=2$，故 $f(x)=-2x^3+2x^2+x+1$.

习题 1.4 (第二章第 7 题)

兽医建议某宠物的食谱每天要包含 100 单位的蛋白质，200 单位的糖，50 单位的脂肪. 某宠物商店出售四种食品 A,B,C,D. 这四种食物每千克含蛋白质、糖、脂肪的含量（单位）如下:

食物	蛋白质	糖	脂肪
A	5	20	2
B	4	25	2
C	7	10	10
D	10	5	6

问是否可以适量配备上述四种食品, 满足兽医的建议.

解 根据题意设配置四种食物 A,B,C,D 的份量分别为 x_1,x_2,x_3,x_4 千克，则考虑如下方程组:

$$\begin{cases}5x_1+4x_2+7x_3+10x_4=100\\20x_1+25x_2+10x_3+5x_4=200\\2x_1+2x_2+10x_3+6x_4=50\end{cases}$$

$$\begin{pmatrix}5&4&7&10&100\\20&25&10&5&200\\2&2&10&6&50\end{pmatrix}\xrightarrow[-20r_1\to r_2,-5r_1\to r_3]{r_1\leftrightarrow r_3,\frac{1}{2}r_1}\begin{pmatrix}1&1&5&3&25\\0&5&-90&-55&-300\\0&-1&-18&-5&-25\end{pmatrix}\xrightarrow[r_2\to r_3]{\frac{1}{5}r_2}\begin{pmatrix}1&1&5&3&25\\0&1&-18&-11&-60\\0&0&-36&-16&-85\end{pmatrix}$$

$$\xrightarrow[18r_3\to r_2,-5r_3\to r_1]{-\frac{1}{36}r_3}\begin{pmatrix}1&1&0&\frac{7}{9}&\frac{475}{36}\\0&1&0&-3&-\frac{35}{2}\\0&0&1&\frac{4}{9}&\frac{85}{36}\end{pmatrix}\xrightarrow{-r_2\to r_1}\begin{pmatrix}1&0&0&\frac{34}{9}&\frac{1105}{36}\\0&1&0&-3&-\frac{35}{2}\\0&0&1&\frac{4}{9}&\frac{85}{36}\end{pmatrix}.$$

令 $x_4=t$，代入原方程可解得: $(x_1,x_2,x_3,x_4)=(-\dfrac{34}{9}t+\dfrac{1105}{36},3t-\dfrac{35}{2},-\dfrac{4}{9}t+\dfrac{85}{36},t)$

而由题意可知 $x_i\geqslant 0,(i=1,2,3,4)$，故 $t\leqslant\dfrac{1105}{136},t\geqslant\dfrac{35}{6},t\leqslant\dfrac{85}{16},t\geqslant 0$，矛盾，故满足问题的解不存在。

1.2 第四周作业

习题 1.5 (第三章第 1 题)

计算下列行列式:

(1) $\begin{vmatrix} 1 & 0 & 1 & -4 \\ -1 & -3 & 0 & -2 \\ 2 & -1 & 4 & 0 \\ 0 & 3 & -3 & 2 \end{vmatrix}$ (2) $\begin{vmatrix} 1 & 4 & -1 & -1 \\ 1 & 0 & 0 & 1 \\ -3 & 3 & -4 & 0 \\ 0 & 1 & -1 & -1 \end{vmatrix}$ (3) $\begin{vmatrix} x+a & x+b & x+c \\ y+a & y+b & y+c \\ z+a & z+b & z+c \end{vmatrix}$

(4) $\begin{vmatrix} & & & a_{1n} \\ & & a_{2,n-1} & a_{2n} \\ & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$ (5) $\begin{vmatrix} a^2 & (a+1)^2 & (a+2)^2 \\ b^2 & (b+1)^2 & (b+2)^2 \\ c^2 & (c+1)^2 & (c+2)^2 \end{vmatrix}$ (6) $\begin{vmatrix} a_1 & a_2 & 0 & 0 & 0 \\ b_1 & b_2 & 0 & 0 & 0 \\ c_1 & c_2 & c_3 & 0 & 0 \\ d_1 & d2 & d_3 & d_4 & 0 \\ e_1 & e_2 & e_3 & e_4 & e_5 \end{vmatrix}$

解 (1)

$$\begin{vmatrix} 1 & 0 & 1 & -4 \\ -1 & -3 & 0 & -2 \\ 2 & -1 & 4 & 0 \\ 0 & 3 & -3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 & -4 \\ 0 & -3 & 1 & -6 \\ 2 & -1 & 4 & 0 \\ 0 & 3 & -3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 & -4 \\ 0 & -3 & 1 & -6 \\ 0 & -1 & 2 & 8 \\ 0 & 3 & -3 & 2 \end{vmatrix} = \begin{vmatrix} -3 & 1 & -6 \\ -1 & 2 & 8 \\ 3 & -3 & 2 \end{vmatrix} = -40.$$

(2)

$$\begin{vmatrix} 1 & 4 & -1 & -1 \\ 1 & 0 & 0 & 1 \\ -3 & 3 & -4 & 0 \\ 0 & 1 & -1 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 4 & -1 & -1 \\ 0 & -4 & 1 & 2 \\ -3 & 3 & -4 & 0 \\ 0 & 1 & -1 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 4 & -1 & -1 \\ 0 & -4 & 1 & 2 \\ 0 & 15 & -7 & -3 \\ 0 & 1 & -1 & -1 \end{vmatrix} = \begin{vmatrix} -4 & 1 & 2 \\ 15 & -7 & -3 \\ 1 & -1 & -1 \end{vmatrix} = -20.$$

(3)

(法一) $\begin{vmatrix} x+a & x+b & x+c \\ y+a & y+b & y+c \\ z+a & z+b & z+c \end{vmatrix} = \begin{vmatrix} x-y & x-y & x-y \\ y+a & y+b & y+c \\ z+a & z+b & z+c \end{vmatrix} = (x-y)\begin{vmatrix} 1 & 1 & 1 \\ y+a & y+b & y+c \\ z+a & z+b & z+c \end{vmatrix} = (x-y)\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a & b & c \end{vmatrix} = 0.$

(法二)可以将行列式

$$f(x) \triangleq \begin{vmatrix} x+a & x+b & x+c \\ y+a & y+b & y+c \\ z+a & z+b & z+c \end{vmatrix}$$

看成系数为 $\mathbb{F}[y,z]$ 关于 x 的多项式. 根据行列式的完全展开式知 $f(x)$ 为关于 x 的一次多项式且有两个根 y,z. 故 $f(x)=0$.

(4)

$$\begin{vmatrix} & & & a_{1n} \\ & & a_{2,n-1} & a_{2n} \\ & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = (-1)^{n+1}a_{1n}\begin{vmatrix} & & a_{2,n-1} \\ & & \vdots \\ a_{n1} & \cdots & a_{n-1,n-1} \end{vmatrix} = \cdots = (-1)^{\frac{n(n-1)}{2}}a_{1n}a_{2,n-1}\cdots a_{n1}.$$

(5) 考虑如下多项式函数

$$f(x) \triangleq \begin{vmatrix} x^2 & (x+1)^2 & (x+2)^2 \\ b^2 & (b+1)^2 & (b+2)^2 \\ c^2 & (c+1)^2 & (c+2)^2 \end{vmatrix}$$

由行列式的完全展开式知 $f(x)$ 为关于 x 的二次多项式. 易知 $f(b)=f(c)=0$, 故可设 $f(x)=\lambda(x-b)(x-c)$.

$$\lambda bc = f(0) = \begin{vmatrix} 0 & 1 & 4 \\ b^2 & (b+1)^2 & (b+2)^2 \\ c^2 & (c+1)^2 & (c+2)^2 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 4 \\ b^2 & b^2+b & b^2 \\ c^2 & c^2+c & c^2 \end{vmatrix} = -\begin{vmatrix} b^2 & b^2 \\ c^2 & c^2 \end{vmatrix} + 4\begin{vmatrix} b^2 & b^2+b \\ c^2 & c^2+c \end{vmatrix} = 4bc(b-c)$$

故 $\lambda = 4(b-c)$, 于是 $f(x)=4(b-c)(x-b)(x-c)$. 带入 $x=a$ 有

$$f(a) = \begin{vmatrix} a^2 & (a+1)^2 & (a+2)^2 \\ b^2 & (b+1)^2 & (b+2)^2 \\ c^2 & (c+1)^2 & (c+2)^2 \end{vmatrix} = 4(b-c)(a-b)(a-c).$$

(6) 由行列式的 *Laplace* 展开

$$\begin{vmatrix} a_1 & a_2 & 0 & 0 & 0 \\ b_1 & b_2 & 0 & 0 & 0 \\ c_1 & c_2 & c_3 & 0 & 0 \\ d_1 & d2 & d_3 & d_4 & 0 \\ e_1 & e_2 & e_3 & e_4 & e_5 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \begin{vmatrix} c_3 & 0 & 0 \\ d_3 & d_4 & 0 \\ e_3 & e_4 & e_5 \end{vmatrix} = (a_1b_2-a_2b_1)c_3d_4e_5.$$

习题 1.6 (第三章第 2 题)

在三维直角坐标系中，已知点 A,B,C,D 的坐标分别是（1，1，0)，(3，1，2)，(0，1，3)，(2，2，4）. 求四面体 $ABCD$ 的体积及各个面的面积.

解 $\overrightarrow{AB}=(2,0,2), \overrightarrow{AC}=(-1,0,3), \overrightarrow{AD}=(1,1,4), \overrightarrow{BC}=(-3,0,1), \overrightarrow{BD}=(-1,1,2)$.

$$\overrightarrow{AB}\times\overrightarrow{AC} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 2 & 0 & 2 \\ -1 & 0 & 3 \end{vmatrix} = (0,-8,0)$$

$$\overrightarrow{AB}\times\overrightarrow{AD} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 2 & 0 & 2 \\ 1 & 1 & 4 \end{vmatrix} = (-2,-6,2)$$

$$\overrightarrow{AC}\times\overrightarrow{AD} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -1 & 0 & 3 \\ 1 & 1 & 4 \end{vmatrix} = (-3,7,-1)$$

$$\overrightarrow{BC}\times\overrightarrow{BD} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -3 & 0 & 1 \\ -1 & 1 & 2 \end{vmatrix} = (-1,5,-3)$$

$S_{\triangle ABC} = \frac{1}{2}|\overrightarrow{AB}\times\overrightarrow{AC}| = 4, S_{\triangle ABD} = \frac{1}{2}|\overrightarrow{AB}\times\overrightarrow{AD}| = \sqrt{11}, S_{\triangle ACD} = \frac{1}{2}|\overrightarrow{AC}\times\overrightarrow{AD}| = \frac{\sqrt{59}}{2}, S_{\triangle BCD} =$

$\frac{1}{2}|\overrightarrow{BC}\times\overrightarrow{BD}|=\frac{\sqrt{35}}{2}$.

$$V_{ABCD}=\frac{1}{6}|(\overrightarrow{AB}\times\overrightarrow{AC})\cdot\overrightarrow{AD}|=\frac{1}{6}\text{abs}(\begin{vmatrix}2&0&2\\-1&0&3\\1&1&4\end{vmatrix})=\frac{4}{3}.$$

习题 1.7 (第三章第 3 题)

将行列式

$$\begin{vmatrix}x-2&1&0&0\\1&x-2&1&0\\0&1&x-2&1\\0&0&1&x-2\end{vmatrix}$$

展开为关于 x 的多项式.

解

$$\begin{vmatrix}x-2&1&0&0\\1&x-2&1&0\\0&1&x-2&1\\0&0&1&x-2\end{vmatrix}=\begin{vmatrix}0&-x^2+4x-3&2-x&0\\1&x-2&1&0\\0&1&x-2&1\\0&0&1&x-2\end{vmatrix}=\begin{vmatrix}1&x-2&1&0\\0&x^2-4x+3&x-2&0\\0&1&x-2&1\\0&0&1&x-2\end{vmatrix}$$

$$=\begin{vmatrix}x^2-4x+3&x-2&0\\1&x-2&1\\0&1&x-2\end{vmatrix}=(x^2-4x+3)\begin{vmatrix}x-2&1\\1&x-2\end{vmatrix}-(x-2)\begin{vmatrix}1&1\\0&x-2\end{vmatrix}=(x^2-4x+3)^2-(x-2)^2$$

$=(x^2-3x+1)(x^2-5x+5)$.

习题 1.8 (第三章第 4 题)

A 为 n 阶方阵，λ 为常数. 证明：$\det(\lambda A)=\lambda^n\det(A)$

证明 由矩阵数乘的定义及行列式映射的多重线性性性知 $\det(\lambda A)=\lambda^n\det(A)$.

习题 1.9 (第三章第 5 题)

方阵 A 称为反对称方阵，如果它的转置方阵等于 $-A$. 证明：奇数阶反对称方阵的行列式为零.

证明 由第四题知 $\det(A)=\det(A^T)=\det(-A)=(-1)^n\det(A)$. 由于 n 为奇数，因此 $\det(A)=-\det(A)$. 故 $2\det(A)=0$, 即 $\det(A)=0$.

1.3 第五周作业

1.3.1 作业答案

习题 1.10 (第三章第 6 题)

证明：

$$\begin{vmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}$$

证明 证法一：考虑四阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ -1 & 0 & b_{11} & b_{12} \\ 0 & -1 & b_{21} & b_{22} \end{vmatrix}$$

我们用后面两行将前两行前两列的四个元素消为 0：

$$\begin{vmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ -1 & 0 & b_{11} & b_{12} \\ 0 & -1 & b_{21} & b_{22} \end{vmatrix} \xlongequal[a_{21}r_3+a_{22}r_4\to r_2]{a_{11}r_3+a_{12}r_4\to r_1} \begin{vmatrix} 0 & 0 & a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ 0 & 0 & a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \\ -1 & 0 & b_{11} & b_{12} \\ 0 & -1 & b_{21} & b_{22} \end{vmatrix} \tag{1.1}$$

分别对式 (1.1) 两边的前两行进行拉普拉斯展开得：

$$LHS = (-1)^{1+2+1+2} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} \tag{1.2}$$

$$\begin{aligned} RHS &= (-1)^{1+2+3+4} \begin{vmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{vmatrix} \cdot \begin{vmatrix} -1 & 0 \\ 0 & -1 \end{vmatrix} \\ &= \begin{vmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{vmatrix} \end{aligned} \tag{1.3}$$

证毕.

证法二：将等式两边行列式完全展开，分别对比 8 项得证 (这种方法只适用于二、三阶这样的低阶行列式，高阶行列式展开项数过多，费时费力)

注 本题即是以下性质的二阶形式，感兴趣的同学可自行证明：设 $A, B \in \mathbb{F}^{n\times n}$，有

$$det(AB) = det(A)det(B) \tag{1.4}$$

习题 1.11 (第三章第 7 题)

证明：

$$\begin{vmatrix} a_{11} & a_{12} & 1 & 0 \\ a_{21} & a_{22} & 0 & 1 \\ 1 & 0 & b_{11} & b_{12} \\ 0 & 1 & b_{21} & b_{22} \end{vmatrix} = \begin{vmatrix} a_{11}b_{11}+a_{12}b_{21}-1 & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22}-1 \end{vmatrix}$$

证明 证法一：类似于上题中的证法一，我们将左边四阶行列式的前两行前两列四个元素消为 0：

$$\begin{vmatrix} a_{11} & a_{12} & 1 & 0 \\ a_{21} & a_{22} & 0 & 1 \\ 1 & 0 & b_{11} & b_{12} \\ 0 & 1 & b_{21} & b_{22} \end{vmatrix} \xlongequal[-a_{21}r_3-a_{22}r_4\to r_2]{-a_{11}r_3-a_{12}r_4\to r_1} \begin{vmatrix} 0 & 0 & 1-a_{11}b_{11}-a_{12}b_{21} & -a_{11}b_{12}-a_{12}b_{22} \\ 0 & 0 & -a_{21}b_{11}-a_{22}b_{21} & 1-a_{21}b_{12}-a_{22}b_{22} \\ 1 & 0 & b_{11} & b_{12} \\ 0 & 1 & b_{21} & b_{22} \end{vmatrix} \tag{1.5}$$

对式 (1.5) 的右边进行拉普拉斯展开得：

$$\begin{aligned} RHS &= (-1)^{1+2+3+4}\begin{vmatrix} 1-a_{11}b_{11}-a_{12}b_{21} & -a_{11}b_{12}-a_{12}b_{22} \\ -a_{21}b_{11}-a_{22}b_{21} & 1-a_{21}b_{12}-a_{22}b_{22} \end{vmatrix} \cdot \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \\ &= \begin{vmatrix} a_{11}b_{11}+a_{12}b_{21}-1 & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22}-1 \end{vmatrix} \end{aligned} \tag{1.6}$$

证毕

证法二：这里给出一种笔者能想到的较为简单的展开证法

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & 1 & 0 \\ a_{21} & a_{22} & 0 & 1 \\ 1 & 0 & b_{11} & b_{12} \\ 0 & 1 & b_{21} & b_{22} \end{vmatrix} &= \begin{vmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 1 \\ 1 & 0 & b_{11} & b_{12} \\ 0 & 1 & b_{21} & b_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & 1 & 0 \\ a_{21} & a_{22} & 0 & 1 \\ 1 & 0 & 0 & b_{12} \\ 0 & 1 & 0 & b_{22} \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ 1 & 0 & b_{11} & b_{12} \\ 0 & 1 & b_{21} & b_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 1 \\ 1 & 0 & b_{11} & 0 \\ 0 & 1 & b_{21} & 0 \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & 1 & 0 \\ a_{21} & a_{22} & 0 & 1 \\ 1 & 0 & 0 & b_{12} \\ 0 & 1 & 0 & b_{22} \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & 0 \\ 1 & 0 & b_{11} \\ 0 & 1 & b_{21} \end{vmatrix} + \begin{vmatrix} a_{21} & a_{22} & 1 \\ 1 & 0 & b_{12} \\ 0 & 1 & b_{22} \end{vmatrix} \\ &= \begin{vmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{vmatrix} - a_{11}b_{11}-a_{12}b_{21}+1-a_{21}b_{12}-a_{22}b_{22} \end{aligned} \tag{1.7}$$

$$\begin{aligned} &\begin{vmatrix} a_{11}b_{11}+a_{12}b_{21}-1 & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22}-1 \end{vmatrix} \\ =&\begin{vmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22}-1 \end{vmatrix} + \begin{vmatrix} -1 & a_{11}b_{12}+a_{12}b_{22} \\ 0 & a_{21}b_{12}+a_{22}b_{22}-1 \end{vmatrix} \\ =&\begin{vmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{vmatrix} + \begin{vmatrix} a_{11}b_{11}+a_{12}b_{21} & 0 \\ a_{21}b_{11}+a_{22}b_{21} & -1 \end{vmatrix} + \begin{vmatrix} -1 & a_{11}b_{12}+a_{12}b_{22} \\ 0 & a_{21}b_{12}+a_{22}b_{22}-1 \end{vmatrix} \\ =&\begin{vmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{vmatrix} - a_{11}b_{11}-a_{12}b_{21}+1-a_{21}b_{12}-a_{22}b_{22} \end{aligned} \tag{1.8}$$

对比式 (1.7),(1.8) 可见命题成立

注 本题即是以下性质的二阶形式，感兴趣的同学可自行证明：设 $A\in F^{m\times n}, B\in F^{n\times m}$，有

$$det(I_n-BA)=det\begin{pmatrix} I_m & A \\ B & I_n \end{pmatrix}=det(I_m-AB) \tag{1.9}$$

习题 1.12 (第三章第 8 题)

设 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d}$ 为 4 维数组向量. 证明：$det(2\boldsymbol{a}-\boldsymbol{b},-\boldsymbol{a}+2\boldsymbol{b}-\boldsymbol{c},-\boldsymbol{b}+2\boldsymbol{c}-\boldsymbol{d},-\boldsymbol{c}+2\boldsymbol{d})=5det(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d})$.

证明 本题有多种变换方法，过程合理即可.

$$
\begin{aligned}
&det(2\boldsymbol{a}-\boldsymbol{b},-\boldsymbol{a}+2\boldsymbol{b}-\boldsymbol{c},-\boldsymbol{b}+2\boldsymbol{c}-\boldsymbol{d},-\boldsymbol{c}+2\boldsymbol{d})\\
\xlongequal[c_3\to c_2]{c_2+c_3+c_4\to c_1}&det(\boldsymbol{a}+\boldsymbol{d},-\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}-\boldsymbol{d},-\boldsymbol{b}+2\boldsymbol{c}-\boldsymbol{d},-\boldsymbol{c}+2\boldsymbol{d})\\
\xlongequal{c_1\to c_2}&det(\boldsymbol{a}+\boldsymbol{d},\boldsymbol{b}+\boldsymbol{c},-\boldsymbol{b}+2\boldsymbol{c}-\boldsymbol{d},-\boldsymbol{c}+2\boldsymbol{d})\\
\xlongequal{c_2\to c_3}&det(\boldsymbol{a}+\boldsymbol{d},\boldsymbol{b}+\boldsymbol{c},3\boldsymbol{c}-\boldsymbol{d},-\boldsymbol{c}+2\boldsymbol{d})\\
\xlongequal{3c_4\to c_3}&det(\boldsymbol{a}+\boldsymbol{d},\boldsymbol{b}+\boldsymbol{c},5\boldsymbol{d},-\boldsymbol{c}+2\boldsymbol{d})\\
=&det(\boldsymbol{a},\boldsymbol{b}+\boldsymbol{c},5\boldsymbol{d},-\boldsymbol{c})\\
=&det(\boldsymbol{a},\boldsymbol{b},5\boldsymbol{d},-\boldsymbol{c})\\
=&5det(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d})
\end{aligned}
$$

习题 1.13 (第三章第 9 题)

求以下排列的逆序数，并指出其奇偶性. (1) (6,8,1,4,7,5,3,2,9) (2) (6,4,2,1,9,7,3,5,8) (3) (7,5,2,3,9,8,1,6,4)

解 (1) $\tau(6,8,1,4,7,5,3,2,9)=19$，为奇排列；
(2) $\tau(6,4,2,1,9,7,3,5,8)=15$，为奇排列；
(3) $\tau(7,5,2,3,9,8,1,6,4)=20$，为偶排列.

习题 1.14 (第三章第 13 题)

用 Cramer 法则求解下列线性方程组：

$$
(1)\begin{cases}x_1-x_2+x_3=3\\x_1+2x_2+4x_3=5\\x_1+3x_2+9x_3=7\end{cases}\qquad
(2)\begin{cases}2x_1+x_2-5x_3+x_4=8\\x_1-3x_2-6x_4=9\\2x_2-x_3+2x_4=-5\\x_1+4x_2-7x_3+6x_4=0\end{cases}
$$

解 (1)

$$
\Delta=\begin{vmatrix}1&-1&1\\1&2&4\\1&3&9\end{vmatrix}=12 \tag{1.10}
$$

$$
\Delta_1=\begin{vmatrix}3&-1&1\\5&2&4\\7&3&9\end{vmatrix}=36 \tag{1.11}
$$

$$
\Delta_2=\begin{vmatrix}1&3&1\\1&5&4\\1&7&9\end{vmatrix}=4 \tag{1.12}
$$

$$
\Delta_3=\begin{vmatrix}1&-1&3\\1&2&5\\1&3&7\end{vmatrix}=4 \tag{1.13}
$$

于是方程组的解为

$$
(x_1,x_2,x_3)=(\frac{\Delta_1}{\Delta},\frac{\Delta_2}{\Delta},\frac{\Delta_3}{\Delta})=(3,\frac{1}{3},\frac{1}{3}) \tag{1.14}
$$

(2)

$$\Delta = \begin{vmatrix} 2 & 1 & -5 & 1 \\ 1 & -3 & 0 & -6 \\ 0 & 2 & -1 & 2 \\ 1 & 4 & -7 & 6 \end{vmatrix} = 27 \tag{1.15}$$

$$\Delta_1 = \begin{vmatrix} 8 & 1 & -5 & 1 \\ 9 & -3 & 0 & -6 \\ -5 & 2 & -1 & 2 \\ 0 & 4 & -7 & 6 \end{vmatrix} = 81 \tag{1.16}$$

$$\Delta_2 = \begin{vmatrix} 2 & 8 & -5 & 1 \\ 1 & 9 & 0 & -6 \\ 0 & -5 & -1 & 2 \\ 1 & 0 & -7 & 6 \end{vmatrix} = -108 \tag{1.17}$$

$$\Delta_3 = \begin{vmatrix} 2 & 1 & 8 & 1 \\ 1 & -3 & 9 & -6 \\ 0 & 2 & -5 & 2 \\ 1 & 4 & 0 & 6 \end{vmatrix} = -27 \tag{1.18}$$

$$\Delta_4 = \begin{vmatrix} 2 & 1 & -5 & 8 \\ 1 & -3 & 0 & 9 \\ 0 & 2 & -1 & -5 \\ 1 & 4 & -7 & 0 \end{vmatrix} = 27 \tag{1.19}$$

于是方程组的解为

$$(x_1, x_2, x_3, x_4) = (\frac{\Delta_1}{\Delta}, \frac{\Delta_2}{\Delta}, \frac{\Delta_3}{\Delta}, \frac{\Delta_4}{\Delta}) = (3, -4, -1, 1) \tag{1.20}$$

习题 1.15 (第三章第 14 题)

设 $x_0, x_1, \cdots, x_n$ 及 $y_0, y_1, \cdots, y_n$ 是任给实数，其中 $x_i (0 \le i \le n)$ 两两互不相等. 证明：存在唯一的次数不超过 n 的多项式 $p(x)$ 满足 $p(x_i) = y_i, i = 0, 1, \cdots, n$.

证明 设 $p(x) = a_0 + a_1 x + \cdots + a_n x^n$，命题等价于证明以下 $(n+1)$ 元一次方程组有唯一解 (设 $a_0, a_1, \cdots, a_n$ 为未知数):

$$\begin{cases} a_0 + x_0 a_1 + \cdots + x_0^n a_n = y_0 \\ a_0 + x_1 a_1 + \cdots + x_1^n a_n = y_1 \\ \quad \vdots \\ a_n + x_n a_1 + \cdots + x_n^n a_n = y_n \end{cases} \tag{1.21}$$

注意到方程组 (1.21) 的系数行列式为 (n+1) 阶 Vandermonde 行列式，有

$$\begin{vmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^n \end{vmatrix} = \prod_{0 \le i < j \le n} (x_j - x_i) \ne 0 \tag{1.22}$$

由 Cramer 法则知方程组有唯一解 $a_i = \frac{\Delta_{i+1}}{\Delta} \quad (0 \le i \le n)$.

注 实际上我们由此得到了拉格朗日插值多项式

习题 1.16 (第三章第 16 题)

计算下列 n 阶行列式

$$
(1)\ \begin{vmatrix}
a_1 & & & & & b_1 \\
& \ddots & & & \iddots & \\
& & a_n & b_n & & \\
& & c_n & d_n & & \\
& \iddots & & & \ddots & \\
c_1 & & & & & d_1
\end{vmatrix}
\qquad
(2)\ \begin{vmatrix}
1+a_1 & 1 & \cdots & 1 \\
1 & 1+a_2 & \ddots & \vdots \\
\vdots & \ddots & \ddots & 1 \\
1 & \cdots & 1 & 1+a_n
\end{vmatrix}
$$

$$
(3)\ \begin{vmatrix}
2cos(\theta) & 1 & 0 & \cdots & 0 \\
1 & 2cos(\theta) & 1 & \cdots & 0 \\
0 & 1 & 2cos(\theta) & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & 2cos(\theta)
\end{vmatrix}
$$

解 (1)

$$
\begin{vmatrix}
a_1 & & & & & b_1 \\
& \ddots & & & \iddots & \\
& & a_n & b_n & & \\
& & c_n & d_n & & \\
& \iddots & & & \ddots & \\
c_1 & & & & & d_1
\end{vmatrix}
=(-1)^{2n-2}
\begin{vmatrix}
a_1 & & & & & & & b_1 \\
c_1 & & & & & & & d_1 \\
& a_2 & & & & & b_2 & \\
& & \ddots & & & \iddots & & \\
& & & a_n & b_n & & & \\
& & & c_n & d_n & & & \\
& & \iddots & & & \ddots & & \\
& c_2 & & & & & d_2 &
\end{vmatrix}
$$

$$
=\begin{vmatrix}
a_1 & b_1 & & & & & & \\
c_1 & d_1 & & & & & & \\
& & a_2 & & & & & b_2 \\
& & & \ddots & & & \iddots & \\
& & & & a_n & b_n & & \\
& & & & c_n & d_n & & \\
& & & \iddots & & & \ddots & \\
& & c_2 & & & & & d_2
\end{vmatrix}
=\cdots
$$

$$
=\begin{vmatrix}
a_1 & b_1 & & & & & \\
c_1 & d_1 & & & & & \\
& & a_2 & b_2 & & & \\
& & c_2 & d_2 & & & \\
& & & & \ddots & & \\
& & & & & a_n & b_n \\
& & & & & c_n & d_n
\end{vmatrix}
$$

$$
\begin{aligned}
&= \begin{vmatrix} a_1 & b_1 \\ c_1 & d_1 \end{vmatrix} \times \begin{vmatrix} a_2 & b_2 \\ c_2 & d_2 \end{vmatrix} \times \cdots \times \begin{vmatrix} a_n & b_n \\ c_n & d_n \end{vmatrix} \\
&= \prod_{i=1}^{n}(a_id_i - c_ib_i) \qquad (1.23)
\end{aligned}
$$

(2) 分以下三种情况讨论：

a. 当 $a_1, a_2, \cdots, a_n$ 中存在两个及以上的 0，那么此时行列式有两行及以上相等，行列式的值为 0；

b. 当 $a_1, a_2, \cdots, a_n$ 中有一个 0 (设为 $a_k(1 \leq k \leq n)$)，而其余元素非零，此时行列式第 k 行元素均为 1，有：

$$
\begin{vmatrix}
1+a_1 & 1 & \cdots & \cdots & \cdots & 1 \\
1 & 1+a_2 & \cdots & \cdots & \cdots & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
1 & 1 & 1 & 1 & \cdots & 1 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 1 & 1 & 1 & \cdots & 1+a_n
\end{vmatrix}
\xlongequal{-r_k \to r_1, \cdots, r_{k-1}, r_{k+1}, \cdots, r_n}
\begin{vmatrix}
a_1 & 0 & \cdots & 0 \\
0 & a_2 & \cdots & 0 \\
\vdots & \vdots & & \vdots \\
1 & 1 & \cdots & 1 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & a_n
\end{vmatrix}
$$

$$
= \begin{vmatrix}
a_1 & 0 & \cdots & \cdots & \cdots & 0 \\
0 & a_2 & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & a_n
\end{vmatrix}
= \prod_{i \neq k} a_i
$$

注 a 和 b 两种情况也可以合并讨论

c. 当 $a_1, a_2, \cdots, a_n$ 均非零：

$$
\begin{vmatrix}
1+a_1 & 1 & \cdots & 1 \\
1 & 1+a_2 & \cdots & 1 \\
\vdots & \vdots & \ddots & \vdots \\
1 & 1 & \cdots & 1+a_n
\end{vmatrix}
\xlongequal{-r_1 \to r_2, \cdots, r_n}
\begin{vmatrix}
1+a_1 & 1 & \cdots & 1 \\
-a_1 & a_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
-a_1 & 0 & \cdots & a_n
\end{vmatrix}
$$

$$
\xlongequal{-\frac{1}{a_2}r_2 - \cdots - \frac{1}{a_n}r_n \to r_1}
\begin{vmatrix}
1+a_1+a_1\sum\limits_{i=2}^{n}\frac{1}{a_i} & 0 & \cdots & 0 \\
-a_1 & a_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
-a_1 & 0 & \cdots & a_n
\end{vmatrix}
= (\prod_{i=1}^{n} a_i)(1 + \sum_{i=1}^{n} \frac{1}{a_i})
$$

(3) 我们将 n 阶这样的行列式记为 K_n，易知 $K_1 = 2cos\theta, K_2 = 4cos^2\theta - 1$，当 $n \geq 3$ 时，将 K_n 按照第一行展开，得：

$$K_n = 2cos\theta K_{n-1} - K_{n-2} \qquad (1.24)$$

方法一：这是一个二阶线性递推数列，其特征方程为：

$$\lambda^2 = 2cos\theta\lambda - 1 \qquad (1.25)$$

解特征方程得到特征根：

$$\lambda = e^{\pm i\theta} \tag{1.26}$$

下面分类讨论：

a. $\theta = 2m\pi(m \in Z)$，此时特征根为重根 $\lambda_1 = \lambda_2 = 1$，且 $cos\theta = 1$，数列的递推公式写为：$K_n = 2K_{n-1} - K_{n-2}$，易知 $K_n = n + 1$.

b. $\theta = (2m+1)\pi(m \in Z)$，此时特征根为重根 $\lambda_1 = \lambda_2 = -1$，且 $cos\theta = -1$，数列的递推公式写为：$K_n = -2K_{n-1} - K_{n-2}$，易知 $K_n = (-1)^n(n+1)$.

c. $\theta \neq m\pi(m \in Z)$，此时两个特征根不等：$\lambda_1 = e^{i\theta}, \lambda_2 = e^{-i\theta}, \lambda_1 \neq \lambda_2$

设 $K_n = c_1\lambda_1^n + c_2\lambda_2^n$，将 K_1, K_2 代入解得 $c_1 = \frac{1-e^{i2\theta}}{2-2cos(2\theta)}, c_2 = \frac{1-e^{-i2\theta}}{2-2cos(2\theta)}$，则 $K_n = \frac{cos(n\theta)-cos((n+2)\theta)}{1-\cos(2\theta)} = \frac{2sin((n+1)\theta)sin\theta}{2sin^2\theta} = \frac{sin((n+1)\theta)}{sin\theta}$

方法二：$cos\theta = \pm 1$ 时与方法一类似，不再赘述.

当 $\sin\theta \neq 0$ 时观察可知 $K_1 = \frac{sin(2\theta)}{\sin\theta}, K_2 = \frac{sin(3\theta)}{sin\theta}$，则我们猜测 $K_n = \frac{sin((n+1)\theta)}{sin\theta}$. 下面进行数学归纳法：

假设 $m < n(n \geq 3)$ 时均有 $K_m = \frac{sin((m+1)\theta)}{sin\theta}$ 成立，则

$$\begin{aligned} K_n &= 2cos\theta K_{n-1} - K_{n-2} \\ &= 2cos\theta\frac{sin(n\theta)}{sin\theta} - \frac{sin((n-1)\theta)}{sin\theta} \\ &= \frac{2cos\theta\sin(n\theta) - sin((n-1)\theta)}{sin\theta} \\ &= \frac{sin((n+1)\theta)}{sin\theta} \end{aligned} \tag{1.27}$$

由归纳法原理知命题成立.

注 本题为三对角行列式的一种特殊情况，类似本题使用的方法可以计算一般的三对角行列式：

$$\begin{vmatrix} a & b & & \\ c & \ddots & \ddots & \\ & \ddots & \ddots & b \\ & & c & a \end{vmatrix}_n = \begin{cases} (n+1)(\frac{a}{2})^n & a^2 = 4bc \\ \dfrac{(a+\sqrt{a^2-4bc})^{n+1} - (a-\sqrt{a^2-4bc})^{n+1}}{2^{n+1}\sqrt{a^2-4bc}} & a^2 \neq 4bc \end{cases} \tag{1.28}$$

1.3.2 附录

对于最后一道题用到的知识的一些补充

二阶线性递推数列

一般的二阶线性递推数列 $\{a_n\}$ 的递推关系写为：

$$a_n = pa_{n-1} + qa_{n-2} \quad (n \geq 3, p \neq 0, q \neq 0) \tag{1.29}$$

若已知 a_1, a_2，怎样求通项公式？

我们尝试将递推关系写为等比关系的形式如下：

$$a_n - \lambda_1 a_{n-1} = \lambda_2(a_{n-1} - \lambda_1 a_{n-2}) \tag{1.30}$$

从而可以列出方程组：

$$\begin{cases} \lambda_1 + \lambda_2 = p \\ \lambda_1\lambda_2 = -q \end{cases} \tag{1.31}$$

由韦达定理可知 λ_1, λ_2 是下面二次方程的两个根

$$\lambda^2 - p\lambda - q = 0 \tag{1.32}$$

我们将式 (1.32) 称为特征方程，λ_1, λ_2 称为特征根，可见特征方程与递推关系式类似，只是将 a_n, a_{n-1}, a_{n-2} 分别替换为了 $\lambda^2, \lambda, 1$.

通过特征方程解出特征根后，有

$$a_n - \lambda_1 a_{n-1} = \lambda_2(a_{n-1} - \lambda_1 a_{n-2}) = \lambda_2^2(a_{n-2} - \lambda_1 a_{n-3}) = \cdots = \lambda_2^{n-2}(a_2 - \lambda_1 a_1) \tag{1.33}$$

观察可知式 (1.30) 可变形为

$$a_n - \lambda_2 a_{n-1} = \lambda_1(a_{n-1} - \lambda_2 a_{n-2}) \tag{1.34}$$

则同理有

$$a_n - \lambda_2 a_{n-1} = \lambda_1^{n-2}(a_2 - \lambda_2 a_1) \tag{1.35}$$

当 $\lambda_1 \neq \lambda_2$ 时，联立式 (2.3)(1.35) 得

$$a_n = \frac{a_2 - \lambda_2 a_1}{\lambda_1 - \lambda_2}\lambda_1^{n-1} - \frac{a_2 - \lambda_1 a_1}{\lambda_1 - \lambda_2}\lambda_2^{n-1} \tag{1.36}$$

当 $\lambda_1 = \lambda_2 = \lambda$ 时，有

$$a_n - \lambda a_{n-1} = \lambda(a_{n-1} - \lambda a_{n-2}) = \lambda^2(a_{n-2} - \lambda a_{n-3}) = \cdots = \lambda^{n-2}(a_2 - \lambda a_1) \tag{1.37}$$

此时

$$\begin{aligned} a_n &= (a_n - \lambda a_{n-1}) + \lambda(a_{n-1} - \lambda a_{n-2}) + \lambda^2(a_{n-2} - \lambda a_{n-3}) + \cdots + \lambda^{n-2}(a_2 - \lambda a_1) + \lambda^{n-1}a_1 \\ &= (n-1)\lambda^{n-2}(a_2 - \lambda a_1) + \lambda^{n-1}a_1 \end{aligned} \tag{1.38}$$

观察式 (1.36)(1.38) 可见，通项公式 a_n 是以 $a_1, a_2, \lambda_1, \lambda_2$ 为参数，n 为自变量的函数，我们下面给出求通项公式的一般程式：

(1) 求特征方程 $\lambda^2 - p\lambda - q = 0$ 的根 λ_1, λ_2，在上述证明过程中，我们并未将根限定在实数范围内，实际上，在复数域中，二次方程必有两根.

(2) 若 $\lambda_1 \neq \lambda_2$，通项公式可写为 $a_n = c_1\lambda_1^n + c_2\lambda_2^n$；若 $\lambda_1 = \lambda_2 = \lambda$，通项公式可写为 $(c_1 + c_2 n)\lambda^n$. 其中 c_1, c_2 为待定系数.

(3) 将 $n = 1, 2$ 代入通项公式，得到两个方程构成的方程组，可由此方程组求得 c_1, c_2，其表达式含 $a_1, a_2, \lambda_1, \lambda_2$，即通项公式的参数.

至此，通项公式完全求出.

复数运算

¶ 实系数一元二次方程的复数解

在复数域中，任一实系数二次方程 $ax^2 + bx + c = 0$ 都存在两根，当其判别式 $\Delta = b^2 - 4ac < 0$ 时，根为两共轭复数：

$$x_1 = \frac{-b + i\sqrt{4ac - b^2}}{2a}, \quad x_2 = \frac{-b - i\sqrt{4ac - b^2}}{2a} \tag{1.39}$$

¶ 欧拉公式

我们知道欧拉公式写为：

$$e^{i\theta} = cos(\theta) + isin(\theta) \tag{1.40}$$

从而 $e^{-i\theta} = cos(-\theta) + isin(-\theta) = cos(\theta) - isin(\theta)$，反解得到：

$$cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2}, \quad sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i} \tag{1.41}$$

1.4 第六周作业

习题 1.17 (补充题)

证明：设 $\mathscr{A}:\mathbb{F}^{n\times 1}\to\mathbb{F}^{m\times 1}$ 为集合映射，则：$\mathscr{A}$ 为线性映射 $\Longleftrightarrow \pi_i\circ\mathscr{A}$ 为线性映射，$\forall 1\leqslant i\leqslant m$.

证明 $\Rightarrow$: 设线性映射 $\mathscr{A}:\mathbb{F}^{n\times 1}\to\mathbb{F}^{m\times 1}$ 对应的矩阵为 $A=(a_{ij})_{m\times n}$，则 $\forall i,1\leqslant i\leqslant m$，有：

$$\pi_i\circ\mathscr{A}:\mathbb{F}^{n\times 1}\to\mathbb{F},\pi_i\circ\mathscr{A}(x_1,x_2,\cdots,x_n)^T=a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n$$

则该映射可用 $1\times n$ 矩阵 $A_i=(a_{i1}\ a_{i2}\ \cdots\ a_{in})$ 表示，从而为线性映射。

$\Leftarrow$: 设线性映射 $\pi_i\circ\mathscr{A}:\mathbb{F}^{n\times 1}\to\mathbb{F}$ 对应的矩阵为 $A_i=(a_{i1}\ a_{i2}\ \cdots\ a_{in})$，则 $\forall\boldsymbol{x}=(x_1,\cdots,x_n)^T\in\mathbb{F}^{n\times 1}$，将 $\mathscr{A}\boldsymbol{x}$ 记为 $\boldsymbol{y}=(y_1,\cdots,y_m)^T$，则 $\boldsymbol{y}$ 的第 $i(1\leqslant i\leqslant m)$ 个分量：

$$y_i=\pi_i\circ\mathscr{A}\boldsymbol{x}=a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n$$

由 $\boldsymbol{x}$ 的任意性，该映射可用矩阵 $A=(a_{ij})_{m\times n}$ 来表示，从而为线性映射。

习题 1.18 (补充题)

记 $\mathscr{L}(\mathbb{F}^{n\times 1},\mathbb{F}^{m\times 1})$ 为 $\mathbb{F}^{n\times 1}$ 到 $\mathbb{F}^{m\times 1}$ 的线性映射全体。

对 $\forall A\in\mathbb{F}^{m\times n}$，令 $\ell_A:\mathbb{F}^{n\times 1}\to\mathbb{F}^{m\times 1},\ell_A(X)=AX,\forall X\in\mathbb{F}^{n\times 1}$

(1) 证明 ℓ_A 为线性映射。

(2) 记 $\Phi:\mathscr{L}(\mathbb{F}^{n\times 1},\mathbb{F}^{m\times 1})\to\mathbb{F}^{m\times n},\Psi:\mathbb{F}^{m\times n}\to\mathscr{L}(\mathbb{F}^{n\times 1},\mathbb{F}^{m\times 1})$

$$\Phi(\mathscr{A})=(\mathscr{A}e_1,\mathscr{A}e_2,\cdots,\mathscr{A}e_n),\forall\mathscr{A}\in\mathscr{L}(\mathbb{F}^{n\times 1},\mathbb{F}^{m\times 1});$$

$$\Psi(A)=\ell_A,\forall A\in\mathbb{F}^{m\times n}.$$

证明：Φ,Ψ 为互逆映射。

证明 (1) 任意取定 $A\in\mathbb{F}^{m\times n}$, 设 $A=(a_{ij})_{m\times n}$, 则 $\forall X=(x_1,\cdots,x_n)^T\in\mathbb{F}^{n\times 1}$, 记 $\ell_A(X)=Y=(y_1,\cdots,y_m)^T$，则由矩阵乘法运算规则可知：

$$Y=\begin{pmatrix}y_1\\y_2\\\vdots\\y_m\end{pmatrix}=AX=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\cdots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}a_{11}x_1+\cdots+a_{1n}x_n\\a_{21}x_1+\cdots+a_{2n}x_n\\\vdots\\a_{m1}x_1+\cdots+a_{mn}x_n\end{pmatrix}$$

从而映射 ℓ_A 可用矩阵 A 来表示，故为线性映射。

(2) 这道题其实就是个阅读理解。首先明确一点：映射是两个集合之间的关系，$\mathscr{L}(\mathbb{F}^{n\times 1},\mathbb{F}^{m\times 1})$ 是一族线性映射构成的集合，而 $\mathbb{F}^{m\times n}$ 是 $m\times n$ 的矩阵全体构成的集合，从而 Φ 作用的对象是线性映射，并且它把线性映射映为矩阵，而 Ψ 作用的对象是矩阵，并且它把矩阵映为线性映射。

回到题目中，由第 (1) 问可知，Ψ 即为 $\mathbb{F}^{m\times n}\to\mathscr{L}(\mathbb{F}^{n\times 1},\mathbb{F}^{m\times 1})$ 的一个一一对应，从而 Ψ 可逆，再由 Φ 的值域和 Ψ 的定义域相同，那么我们只需要验证：

$$\forall\mathscr{A}\in\mathscr{L}(\mathbb{F}^{n\times 1},\mathbb{F}^{m\times 1})\text{，都有}\Psi(\Phi(\mathscr{A}))=\mathscr{A}\text{，从而}\Phi(\mathscr{A})=\Psi^{-1}(\mathscr{A})$$

设 $\mathscr{A}$ 对应的矩阵为 $A=(a_{ij})_{m\times n}$，由第 (1) 问 $\mathscr{A}=\ell_A$。且 $\mathscr{A}e_k=(a_{1k},a_{2k},\cdots,a_{mk})^T,1\leqslant k\leqslant n$，则：

$$\Phi(\mathscr{A})=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\cdots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}=A\text{，从而}\Psi(\Phi(\mathscr{A}))=\Psi(A)=\ell_A=\mathscr{A}$$

笔记 以上两题采用的是教材上对线性映射的定义来证明，课上讲到的验证加法与数乘的方法是更本质的做法，不过在数组向量空间中，这两种做法是等价的。

习题 1.19 (补充题)

$\forall A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{n\times p}, C \in \mathbb{F}^{p\times q}$，证明矩阵乘法满足下列性质：

(1) $(AB)C = A(BC)$

(2) $AI_n = I_mA = A, \quad I_n = diag(1, \cdots, 1)$

(3) $(A_1 + A_2)B = A_1B + A_2B, \quad A(B_1 + B_2) = AB_1 + AB_2$

(4) $(\lambda A)B = A(\lambda B) = \lambda(AB)$

证明 只需要验证等式两边矩阵的任意位置元素均相等即可。以下记 $(A)_{ij}$ 为矩阵 A 的 (i,j) 位置的元素。

(1) $(AB)_{ij} = \sum_{k=1}^{n}(A)_{ik}(B)_{kj}$，故 $((AB)C)_{ij} = \sum_{t=1}^{p}(AB)_{it}(C)_{tj} = \sum_{t=1}^{p}\sum_{k=1}^{n}(A)_{ik}(B)_{kt}(C)_{tj}$.

$(BC)_{ij} = \sum_{t=1}^{p}(B)_{it}(C)_{tj}$，故 $(A(BC))_{ij} = \sum_{k=1}^{n}(A)_{ik}(BC)_{kj} = \sum_{k=1}^{n}\sum_{t=1}^{p}(A)_{ik}(B)_{kt}(C)_{tj} = ((AB)C)_{ij}$.

(2) $(AI_n)_{ij} = \sum_{k=1}^{n}(A)_{ik}(I_n)_{kj} = \sum_{k=1}^{n}(A)_{ik}\delta_{kj} = A_{ij}$. (其中 δ_{ij} 为 Kronecker 符号)

$(I_mA)_{ij} = \sum_{k=1}^{m}(I_m)_{ik}(A)_{kj} = \sum_{k=1}^{m}\delta_{ik}(A)_{kj} = A_{ij}$. 故 $AI_n = I_mA = A$.

(3) $((A_1 + A_2)B)_{ij} = \sum_{k=1}^{n}(A_1 + A_2)_{ik}(B)_{kj} = \sum_{k=1}^{n}(A_1)_{ik}(B)_{kj} + \sum_{k=1}^{n}(A_2)_{ik}(B)_{kj} = (A_1B)_{ij} + (A_2B)_{ij}$.

$(A(B_1 + B_2))_{ij} = \sum_{k=1}^{n}(A)_{ik}(B_1 + B_2)_{kj} = \sum_{k=1}^{n}(A)_{ik}(B_1)_{kj} + \sum_{k=1}^{n}(A)_{ik}(B_2)_{kj} = (AB_1)_{ij} + (AB_2)_{ij}$

(4) $((\lambda A)B)_{ij} = \sum_{k=1}^{n}(\lambda A)_{ik}(B)_{kj} = \lambda\sum_{k=1}^{n}(A)_{ik}(B)_{kj} = \sum_{k=1}^{n}(A)_{ik}(\lambda B)_{kj} = (\lambda(AB))_{ij} = (A(\lambda B))_{ij}$.

习题 1.20 (第四章第 3 题)

设 $A = \begin{pmatrix} -3 & -1 & -2 \\ 1 & 3 & 4 \end{pmatrix}, B = \begin{pmatrix} 2 & 2 & -2 \\ 4 & -1 & -4 \\ 4 & 3 & -3 \end{pmatrix}, C = \begin{pmatrix} 1 & 1 \\ -4 & 1 \\ -1 & -2 \end{pmatrix}$, 计算 AB, BC, ABC, B^2, AC, CA.

解

$$AB = \begin{pmatrix} -3 & -1 & -2 \\ 1 & 3 & 4 \end{pmatrix}\begin{pmatrix} 2 & 2 & -2 \\ 4 & -1 & -4 \\ 4 & 3 & -3 \end{pmatrix} = \begin{pmatrix} -18 & -11 & 16 \\ 30 & 11 & -26 \end{pmatrix}.$$

$$BC = \begin{pmatrix} 2 & 2 & -2 \\ 4 & -1 & -4 \\ 4 & 3 & -3 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -4 & 1 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} -4 & 8 \\ 12 & 11 \\ -5 & 13 \end{pmatrix}.$$

$$ABC = \begin{pmatrix} -3 & -1 & -2 \\ 1 & 3 & 4 \end{pmatrix}\begin{pmatrix} 2 & 2 & -2 \\ 4 & -1 & -4 \\ 4 & 3 & -3 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -4 & 1 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} -18 & -11 & 16 \\ 30 & 11 & -26 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -4 & 1 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} 10 & -61 \\ 12 & 93 \end{pmatrix}.$$

$$B^2 = \begin{pmatrix} 2 & 2 & -2 \\ 4 & -1 & -4 \\ 4 & 3 & -3 \end{pmatrix}\begin{pmatrix} 2 & 2 & -2 \\ 4 & -1 & -4 \\ 4 & 3 & -3 \end{pmatrix} = \begin{pmatrix} 4 & -4 & -6 \\ -12 & -3 & 8 \\ 8 & -4 & -11 \end{pmatrix}.$$

$$AC = \begin{pmatrix} -3 & -1 & -2 \\ 1 & 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -4 & 1 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ -15 & -4 \end{pmatrix}.$$

$$CA = \begin{pmatrix} 1 & 1 \\ -4 & 1 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} -3 & -1 & -2 \\ 1 & 3 & 4 \end{pmatrix} = \begin{pmatrix} -2 & 2 & 2 \\ 13 & 7 & 12 \\ 1 & -5 & -6 \end{pmatrix}.$$

习题 1.21 (第四章第 5 题)

计算$(x_1\ x_2\ \cdots\ x_m)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$.

解 原式 $= \left(\sum_{i=1}^m x_i a_{i1} \quad \sum_{i=1}^m x_i a_{i2} \quad \cdots \quad \sum_{i=1}^m x_i a_{in}\right)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{j=1}^n \sum_{i=1}^m x_i a_{ij} y_j.$

习题 1.22 (第四章第 6 题)

举例求满足条件的 2 阶实方阵 A.

(1) $A^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (2) $A^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ (3) $A^3 = I$ 且 $A \neq I$.

解 (1) 设 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，则令 $A^2 = \begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow \begin{cases} a^2+bc = bc+d^2 = 0 \\ b(a+d) = c(a+d) = 1 \end{cases}$

$\Rightarrow \begin{cases} a^2+bc = bc+d^2 = 0 \\ b(a+d) = c(a+d) = 1 \end{cases} \Rightarrow a = d$ 且 b, c 异号 $\Rightarrow a = b = c = d = 0$，矛盾，故不存在满足条件的$A$.

这其实是因为一个二维空间中的线性变换最多将平面的定向改变一次，而二维空间中平面定向改变两次后与原来的定向相同，从而若对空间作两次相同的线性变换，定向必然与原来的空间相同。或者也可以通过后续会学到的性质：矩阵行列式的乘积为乘积的行列式，得到 $\det(A^2) = (\det(A))^2 = -1$，但是显然有实矩阵的行列式必须为实数，从而矛盾。

(2) 取 $A = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$ 或 $\begin{pmatrix} -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}$ 即可。(3) 取 $A = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$ 或 $\begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$ 即可。

对于 (2)(3) 两小问，可以考虑几何意义，将矩阵 $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ 作用于二维直角坐标系中的点 (x, y)，等价于将向量 (x, y) 以原点为中心逆时针旋转了 θ 角度，那么 A^n 就代表旋转了 $n\theta$ 角度，从而第 (2) 问中 θ 可取

$-\frac{\pi}{4}$ 或 $\frac{3\pi}{4}$，第 (3) 问中 θ 可取 $\pm\frac{2\pi}{3}$.

现在我们来寻找 (2)(3) 问中所有满足条件的二阶实方阵。

对于第 (2) 问，设 $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$，则令 $A^2=\begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix}=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\Rightarrow\begin{cases} a^2+bc=bc+d^2=0 \\ b(a+d)=1 \\ c(a+d)=-1 \end{cases}$

从而 $a^2=d^2$，而 $a+d\neq 0$，则有 $a=d,b=-c=\frac{1}{2d}$，那么代入第一式有 $a^2=b^2=c^2=d^2=\frac{1}{4d^2}$，从而 $d^2=\frac{1}{2},d=\pm\frac{\sqrt{2}}{2}$，故 $a=\pm\frac{\sqrt{2}}{2},b=\pm\frac{\sqrt{2}}{2},c=\mp\frac{\sqrt{2}}{2}$，即满足要求的矩阵只有上面列出的两个。

对于第 (3) 问，令 $A^3=\begin{pmatrix} a^3+2abc+bcd & a^2b+b^2c+abd+bd^2 \\ a^2c+acd+bc^2+cd^2 & abc+2bcd+d^3 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

则 $\begin{cases} a^3+2abc+bcd=abc+2bcd+d^3=1 \\ b(a^2+bc+ad+d^2)=c(a^2+bc+ad+d^2)=0 \end{cases}$ 若 $b=0$ 或 $c=0$，代入可解得 $A=I$，矛盾，故 b,c 均不为 0，从而方程组等价于 $\begin{cases} a(a^2+bc)=d(bc+d^2)=1-abc-bcd \\ a^2+bc+ad+d^2=0 \end{cases}$，那么所有满足这个方程组的四元数组 (a,b,c,d) 所组成的矩阵均为符合第 (3) 问要求的矩阵。

1.5 第七周作业

习题 1.23 (第四章第 7 题 (2)(3)(4)(5))

计算下列方阵的 k 次方幂，k 为正整数.

$$(2)\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \qquad (3)\begin{pmatrix} 1 & a & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & a \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$(4)\begin{pmatrix} 1 & 1 & & \\ & 1 & \ddots & \\ & & \ddots & 1 \\ & & & 1 \end{pmatrix}_{n\times n} \qquad (5)\begin{pmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & \cdots & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{pmatrix}$$

解 (2) 法一：$a=b=0$ 时，$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}^k = O^k = O$

a,b 不同时为 0 时，$\sqrt{a^2+b^2}\neq 0$，则有

$$\begin{aligned}\begin{pmatrix} a & b \\ -b & a \end{pmatrix}^k &= \left(\sqrt{a^2+b^2}\begin{pmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{pmatrix}\right)^k \\ &= (a^2+b^2)^{k/2}\begin{pmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{pmatrix}^k \\ &= (a^2+b^2)^{k/2}\begin{pmatrix} cosk\theta & sink\theta \\ -sink\theta & cosk\theta \end{pmatrix}\end{aligned} \tag{1.42}$$

其中 $cos\theta=\frac{a}{\sqrt{a^2+b^2}}, sin\theta=\frac{b}{\sqrt{a^2+b^2}}$，最后一步用到了旋转矩阵的 k 次幂 (转 k 次 θ 角相当于转一次 $k\theta$ 角)，即习题 7.(1). 注意：我们上述的讨论是在 a,b 取实数的情况下进行的，当 a,b 取复数时需要对结果进行解析延拓，扩展到复数域，在此不做赘述。

法二：将矩阵进行拆分：$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = aI + b\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

则有

$$\begin{aligned}\begin{pmatrix} a & b \\ -b & a \end{pmatrix}^k &= (aI+b\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix})^k = \sum_{i=0}^{k} C_k^i a^{k-i} b^i \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^i \\ &= \sum_{i=2m}^{0\le i\le k} C_k^i a^{k-i}b^i \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}^m + \sum_{i=2n+1}^{1\le i\le k} C_k^i a^{k-i} b^i \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}^n \\ &= \begin{pmatrix} \sum\limits_{m=0}^{[\frac{k}{2}]} (-1)^m C_k^{2m} a^{k-2m} b^{2m} & \sum\limits_{n=0}^{[\frac{k-1}{2}]} (-1)^n C_k^{2n+1} a^{k-2n-1} b^{2n+1} \\ -\sum\limits_{n=0}^{[\frac{k-1}{2}]} (-1)^n C_k^{2n+1} a^{k-2n-1} b^{2n+1} & \sum\limits_{m=0}^{[\frac{k}{2}]} (-1)^m C_k^{2m} a^{k-2m} b^{2m} \end{pmatrix}\end{aligned} \tag{1.43}$$

注 将矩阵拆分再按二项式展开计算矩阵幂次需满足拆成的两个矩阵可交换：即令 $A=B+C$，需 $BC=CB$

法三：一般的，求矩阵 k 次幂的通法是进行相似对角化 (不要问我怎么化的，后面会学)：

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}\begin{pmatrix} a+bi & 0 \\ 0 & a-bi \end{pmatrix}\begin{pmatrix} 1/2 & -i/2 \\ 1/2 & i/2 \end{pmatrix} \overset{def}{=} A^{-1}BA \tag{1.44}$$

则有

$$\begin{aligned}\begin{pmatrix} a & b \\ -b & a \end{pmatrix}^k &= (A^{-1}BA)^k = A^{-1}B^kA = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}\begin{pmatrix} (a+bi)^k & 0 \\ 0 & (a-bi)^k \end{pmatrix}\begin{pmatrix} 1/2 & -i/2 \\ 1/2 & i/2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{(a+bi)^k+(a-bi)^k}{2} & \frac{i(a-bi)^k-i(a+bi)^k}{2} \\ \frac{i(a+bi)^k-i(a-bi)^k}{2} & \frac{(a+bi)^k+(a-bi)^k}{2} \end{pmatrix}\end{aligned} \tag{1.45}$$

(3) 设 $A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$，有 $A^k = \begin{pmatrix} 1 & ka \\ 0 & 1 \end{pmatrix}$，故

$$\begin{aligned}\begin{pmatrix} 1 & a & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & a \\ 0 & 0 & 0 & 1 \end{pmatrix}^k &= \begin{pmatrix} A & I \\ O & A \end{pmatrix}^k \\ &= \begin{pmatrix} A^k & kA^{k-1} \\ O & A^k \end{pmatrix} \\ &= \begin{pmatrix} 1 & ka & k & k(k-1)a \\ 0 & 1 & 0 & k \\ 0 & 0 & 1 & ka \\ 0 & 0 & 0 & 1 \end{pmatrix}.\end{aligned} \tag{1.46}$$

(4) 将矩阵拆为两部分：$\begin{pmatrix} 1 & 1 & & \\ & 1 & \ddots & \\ & & \ddots & 1 \\ & & & 1 \end{pmatrix}_{n\times n} = I_{n\times n} + J_{n\times n}$，其中 $J_{n\times n} = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}_{n\times n}$，满足 $J^k = \begin{cases} \begin{pmatrix} O & I_{n-k} \\ O & O \end{pmatrix} & k < n \\ O & k \geq n \end{cases}$，则由二项式定理：

$$\begin{pmatrix} 1 & 1 & & \\ & 1 & \ddots & \\ & & \ddots & 1 \\ & & & 1 \end{pmatrix}^k = (I+J)^k = \sum_{i=0}^{k} C_k^i J^i = \begin{pmatrix} 1 & C_k^1 & C_k^2 & \cdots & C_k^{n-1} \\ & 1 & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & C_k^2 \\ & & & \ddots & C_k^1 \\ & & & & 1 \end{pmatrix} \tag{1.47}$$

在上式中，我们并没有讨论 k 和 n 的大小关系，而是使用了更广泛的组合数定义 $C_p^q = 0, \forall q > p$ 来使结果简洁.

(5)

$$\text{注意到}\begin{pmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & \cdots & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{pmatrix} = (a_1, a_2, \cdots, a_n)^T(b_1, b_2, \cdots, b_n) \overset{def}{=} AB,$$

且 $BA = (b_1, b_2, \cdots, b_n)(a_1, a_2, \cdots, a_n)^T = \sum_{i=1}^{n} a_ib_i$，有

$$\begin{pmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & \cdots & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{pmatrix}^k = (AB)^k = ABAB\cdots AB = A(BA)(BA)\cdots(BA)B$$

$$= A(\sum_{i=1}^{n} a_ib_i)^{k-1}B = (\sum_{i=1}^{n} a_ib_i)^{k-1} \begin{pmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & \cdots & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{pmatrix}$$

习题 1.24 (第四章第 8 题)

8. 设 A, B 都是 n 阶对称方阵，且 $AB = BA$. 证明 AB 也是对称方阵.

证明 $(AB)^T = B^TA^T = BA = AB$，即 AB 为对称方阵.

习题 1.25 (第四章第 9 题)

证明：两个 n 阶上 (下) 三角方阵的乘积仍是上 (下) 三角.

证明 我们只证上三角阵，下三角阵同理.

设 A, B 为任意两个 n 阶上三角方阵，则 $\forall 1 \le j < i \le n$ 有 $A_{ij} = B_{ij} = 0$，令 $C = AB$，则 $\forall 1 \le j < i \le n$，有

$$C_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{i-1} a_{ik}b_{kj} + \sum_{k=i}^{n} a_{ik}b_{kj} = \sum_{k=1}^{i-1} 0 \cdot b_{kj} + \sum_{k=i}^{n} a_{ik} \cdot 0 = 0 \tag{1.48}$$

即 C 为上三角阵.

注 上三角阵的等价表述是其行标大于列标的元素均为 0.

习题 1.26 (第四章第 10 题)

证明：与任意 n 阶方阵都乘法可交换的方阵一定是数量阵.

证明 设 A 为与任意 n 阶方阵均乘法可交换的 n 阶方阵，考虑 A 与矩阵 E_{ij}(表示第 i 行第 j 列元素为 1，其余元素均为 0 的 n 阶方阵) 的乘法，有：

$$AE_{ij} = E_{ij}A \implies a_{ii} = a_{jj}, \quad a_{pi} = a_{jq} = 0 \quad (\forall p \ne i, q \ne j) \tag{1.49}$$

其中 a_{st} 表示矩阵 A 的第 s 行第 t 列元素. 在式 (1.49) 中取 $i = 1$, 令 j 遍历 1 到 n 可得：$a_{11} = a_{22} = \cdots = a_{nn}$ 且 $a_{st} = 0 \quad \forall s \ne t$. 即 A 为数量阵.

习题 1.27 (第四章第 17 题)

求所有满足 $A^2 = O, B^2 = I, \overline{C}^T C = I$ 的 2 阶复方阵 A、B、C.

解

(1) 设 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，则由题得到方程组：$a^2 + bc = b(a+d) = c(a+d) = bc + d^2 = 0$，　下面分类讨论：

当 $b=0$ 时 $a=d=0$，c 可取任意复数，此时 $A=\begin{pmatrix}0&0\\c&0\end{pmatrix}$；

当 $b\neq 0$ 时 $a+d=0, bc=-a^2$，此时 $A=\begin{pmatrix}a&b\\-a^2/b&-a\end{pmatrix}$.

(2) 设 $B=\begin{pmatrix}a&b\\c&d\end{pmatrix}$，则由题得到方程组：$a^2+bc=bc+d^2=1$，$b(a+d)=c(a+d)=0$，　　下面分类讨论：

当 $b=0$ 时，$a=\pm1, d=\pm1$，有 $B=\begin{pmatrix}1&0\\0&1\end{pmatrix}$ 或 $\begin{pmatrix}1&0\\c&-1\end{pmatrix}$ 或 $\begin{pmatrix}-1&0\\c&1\end{pmatrix}$ 或 $\begin{pmatrix}-1&0\\0&-1\end{pmatrix}$；

当 $b\neq 0$ 时，$a+d=0, c=\frac{1-a^2}{b}$，即 $B=\begin{pmatrix}a&b\\\frac{1-a^2}{b}&-a\end{pmatrix}$.

(3) 由 $\overline{C}^TC=I$ 易知 $C\overline{C}^T=I$. 设 $C=\begin{pmatrix}ae^{i\alpha_1}&be^{i\alpha_2}\\ce^{i\alpha_3}&de^{i\alpha_4}\end{pmatrix}$ 其中 $a,b,c,d\geq 0$ 为模长. 由 $\overline{C}^TC=I$ 及 $C\overline{C}^T=I$ 可得 $a^2+b^2=a^2+c^2=c^2+d^2=b^2+d^2=1$，设 $a=cos\theta\quad(\theta\in[0,\frac{\pi}{2}])$，有 $b=c=sin\theta, d=cos\theta$. 下面分类讨论：

$\theta=0$ 时 $C=\begin{pmatrix}e^{i\alpha_1}&0\\0&e^{i\alpha_4}\end{pmatrix}$，$\theta=\frac{\pi}{2}$ 时 $C=\begin{pmatrix}0&e^{i\alpha_2}\\e^{i\alpha_3}&0\end{pmatrix}$(这两种情况可以合并到第三种的结果里)

$\theta\in(0,\frac{\pi}{2})$ 时进一步由题有 $e^{i(\alpha_2-\alpha_1)}+e^{i(\alpha_4-\alpha_3)}=0$，得 $\alpha_2=\alpha_1+\alpha_4-\alpha_3-\pi+2k\pi$，故有 $C=\begin{pmatrix}cos\theta e^{i\alpha_1}&-sin\theta e^{i(\alpha_1+\alpha_4-\alpha_3)}\\sin\theta e^{i\alpha_3}&cos\theta e^{i\alpha_4}\end{pmatrix}$，这里将第一行第二列元素的辐角用其他三个角度表示是为了与二阶旋转矩阵形式统一，特别地，当四个辐角均取 0 时 C 即为旋转矩阵.

下面给出丘维声《高等代数》下册对本题的做法：

例 11　求出所有 2 级酉矩阵。

解　设

$$A=\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}$$

是酉矩阵，则 $A^{-1}=A^*$。于是有

$$\frac{1}{|A|}\begin{bmatrix}a_{22}&-a_{12}\\-a_{21}&a_{11}\end{bmatrix}=\begin{bmatrix}\overline{a_{11}}&\overline{a_{21}}\\\overline{a_{12}}&\overline{a_{22}}\end{bmatrix}.$$

由此得出，$a_{22}=|A|\overline{a_{11}}$，　$a_{12}=-|A|\overline{a_{21}}$。

由于 A 的列向量组是 $\mathbf{C}^2$ 的一个标准正交基，因此 $|a_{11}|^2+|a_{21}|^2=1$。从而 $(|a_{11}|,|a_{21}|)$ 是单位圆上的一个点且在第 1 象限或在 x 轴、y 轴的正半轴上，于是存在 $\theta(0\leqslant\theta\leqslant\frac{\pi}{2})$，使得 $|a_{11}|=\cos\theta$，$|a_{21}|=\sin\theta$。因此 $a_{11}=\cos\theta\,e^{i\theta_1}$，$a_{21}=\sin\theta\,e^{i\theta_2}$，其中 $0\leqslant\theta_j<2\pi, j=1,2$。

据本套教材上册习题 4.6 第 18 题，酉矩阵 A 的行列式的模为 1，因此 $|A|=e^{i\theta_3}$，其中 $0\leqslant\theta_3<2\pi$，于是 $a_{22}=e^{i\theta_3}\cos\theta\,e^{-i\theta_1}$，$a_{12}=-e^{i\theta_3}\sin\theta\,e^{-i\theta_2}$。从而

$$A=\begin{bmatrix}\cos\theta\,e^{i\theta_1}&-\sin\theta\,e^{i\theta_3}e^{-i\theta_2}\\\sin\theta\,e^{i\theta_2}&\cos\theta\,e^{i\theta_3}e^{-i\theta_1}\end{bmatrix}.\tag{71}$$

直接验证知道，形如(71)的矩阵都是酉矩阵。于是(71)式给出了所有的 2 级酉矩阵，其中 $0\leqslant\theta\leqslant\frac{\pi}{2}$，$0\leqslant\theta_j<2\pi, j=1,2,3$。

可以把(71)式的 A 写成下述形式：

$$A=\begin{bmatrix}e^{-i\theta_2}&0\\0&e^{-i\theta_1}\end{bmatrix}\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}\begin{bmatrix}e^{i(\theta_1+\theta_2)}&0\\0&e^{i\theta_3}\end{bmatrix}.\tag{72}$$

习题 1.28 (第四章第 30 题)

设 A 为 n 阶方阵，且满足 $A^3 = I_n$. 计算 $\begin{pmatrix} O & I_n \\ A & O \end{pmatrix}^{2024}$.

解

$$\begin{aligned}
\begin{pmatrix} O & I_n \\ A & O \end{pmatrix}^{2024} &= \left(\begin{pmatrix} O & I_n \\ A & O \end{pmatrix}^2\right)^{1012} \\
&= \begin{pmatrix} A & O \\ O & A \end{pmatrix}^{1012} \\
&= \begin{pmatrix} A & O \\ O & A \end{pmatrix}^{1011} \begin{pmatrix} A & O \\ O & A \end{pmatrix} \\
&= \left(\begin{pmatrix} A & O \\ O & A \end{pmatrix}^3\right)^{337} \begin{pmatrix} A & O \\ O & A \end{pmatrix} \\
&= I_{2n} \begin{pmatrix} A & O \\ O & A \end{pmatrix} \\
&= \begin{pmatrix} A & O \\ O & A \end{pmatrix}.
\end{aligned} \tag{1.50}$$

1.6 第八周作业

习题 1.29 (补充题)

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

(1) 求：

$$\begin{pmatrix} O & A \\ B & I_n \end{pmatrix} \cdot \begin{pmatrix} I_n & O \\ -B & I_n \end{pmatrix}$$

(2) 试用 (1) 及 Laplace 展开证明 $\det AB = \det A \cdot \det B$.

解 (1) 利用分块矩阵乘法可得：

$$\begin{pmatrix} O & A \\ B & I_n \end{pmatrix} \cdot \begin{pmatrix} I_n & O \\ -B & I_n \end{pmatrix} = \begin{pmatrix} -AB & A \\ O & I_n \end{pmatrix}.$$

(2) 对 (1) 中的右式做初等行变换可得（第二列右乘 B 加到第一列）：

$$\begin{vmatrix} -AB & A \\ O & I_n \end{vmatrix} = \begin{vmatrix} O & A \\ B & I_n \end{vmatrix}.$$

对前 n 行做 Laplace 展开：$\det(-AB) \cdot \det(I_n) = (-1)^n \cdot \det(A) \cdot \det(B)$，即 $\det AB = \det A \cdot \det B$.

习题 1.30 (补充题)

举例说明存在矩阵 A, B，使得 $\det AB \neq \det BA$. 问：A, B 是否可取为方阵？

解 取 $A = (1, 1, 1), B = A^T$，则 $AB = 3, BA = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, $\det AB = 3 \neq 0 = \det BA$.

A, B 不能取为方阵，否则 A, B 必同阶，则由上题结论，$\det AB = \det A \cdot \det B = \det BA$.

笔记 显然这里的例子是不唯一的，在学习到矩阵的秩后大家就会对这个命题有新的认识了。

习题 1.31 (补充题)

设 $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, 其中 $b_i \neq 0, i = 1, \cdots, n$ $\quad A = \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_n \end{pmatrix}$, 求 $\det\left(A - BB^T\right)$.

解 由题，$(BB^T)_{ij} = b_i b_j$，则：

$$\det\left(A - BB^T\right) = \begin{vmatrix} b_1 - b_1^2 & -b_1 b_2 & \cdots & -b_1 b_n \\ -b_2 b_1 & b_2 - b_2^2 & \cdots & -b_2 b_n \\ \vdots & \vdots & \ddots & \vdots \\ -b_n b_1 & -b_n b_2 & \cdots & b_n - b_n^2 \end{vmatrix} = \prod_{i=1}^{n} (-b_i) \cdot \begin{vmatrix} b_1 - 1 & b_2 & \cdots & b_n \\ b_1 & b_2 - 1 & \cdots & b_n \\ \vdots & \vdots & \ddots & \vdots \\ b_1 & b_2 & \cdots & b_n - 1 \end{vmatrix}$$

$$= \prod_{i=1}^{n} (-b_i) \cdot (-1)^n (1 - \sum_{k=1}^{n} b_k) = (1 - \sum_{k=1}^{n} b_k) \prod_{i=1}^{n} b_i$$

其中第二行的等号用到了习题2.3的结论（再次表明这个类型的行列式很重要）。

或者也可以利用下一题结论，有 $\det\left(A - BB^T\right) = \det(A) \cdot \det\left(I_n - A^{-1}BB^T\right) = \det(A) \cdot \det\left(1 - B^T A^{-1} B\right)$.

习题 1.32 (补充题)

设 $A \in \mathbb{F}^{n\times m}, B \in \mathbb{F}^{m\times n}$，证明：$\lambda^m \det\{(\lambda I_n - AB)\} = \lambda^n \det\{(\lambda I_m - BA)\}$.

证明 $\lambda = 0$ 时，两边均为 0，等式成立，仅考虑 $\lambda \neq 0$ 情形即可。

先证明 $\lambda = 1$ 时等式成立，我们有：

$$\begin{pmatrix} I_m - BA & B \\ O & I_n \end{pmatrix} = \begin{pmatrix} I_m & O \\ A & I_n - AB \end{pmatrix} \begin{pmatrix} I_m & B \\ O & I_n \end{pmatrix} \begin{pmatrix} I_m & O \\ -A & I_n \end{pmatrix}$$

其实这个拆分就是对分块矩阵做初等变换，两边取行列式可得 $\det\{(I_m - BA)\} = \det\{(I_n - AB)\}$.

则 $\lambda^m|\lambda I_n - AB| = \lambda^m\lambda^n|I_n - \lambda^{-1}AB| = \lambda^n|\lambda I_m||I_m - B\lambda^{-1}A| = \lambda^n|\lambda Im(I_m - \lambda^{-1}BA)| = \lambda^n|\lambda I_m - BA|$.

这就得到了我们要证明的结论。

笔记 这个结论十分重要，并且这个证明的过程使用的技巧是很常用的，所以建议大家记住证明过程。并且这个结论告诉我们 AB 与 BA 有相同的非零特征值。

习题 1.33 (第四章第 13 题)

设方阵 A 满足 $A^k = O$，k 为正整数，证明：$I + A$ 可逆，并求 $(I + A)^{-1}$.

证明 注意这里 k 是一个给定的数，而不是对任意自然数都成立。

取 $B = I + \sum_{i=1}^{k-1}(-A)^i$，则 $(I + A)B = B(I + A) = I + \sum_{i=1}^{k-1}(-A)^i + A - \sum_{i=2}^{k}(-A)^i = I - A + A - (-A)^k = I$.

故 $I + A$ 可逆，且 $(I + A)^{-1} = B = I + \sum_{i=1}^{k-1}(-A)^i$.

那为什么这么取呢？我们可以像几何级数一样对 $(I + A)^{-1}$ 做展开，我们知道 $|x| < 1$ 时，$\dfrac{1}{1-x} = \sum_{i=0}^{\infty} x^i$.

由此我们猜测 $(I + A)^{-1} = I + \sum_{i=1}^{\infty}(-A)^i$，代入验证会发现这确实是正确的，再结合 $A^k = O$，便得到了 B.

习题 1.34 (第四章第 14 题)

设方阵 A 满足 $I - 2A - 3A^2 + 4A^3 + 5A^4 - 6A^5 = O$，证明：$I - A$ 可逆，并求 $(I - A)^{-1}$.

证明 这其实是个因式分解，由 $2I - 2A - 3A^2 + 4A^3 + 5A^4 - 6A^5 = I$，则 $(I - A)(2I - 3A^2 + A^3 + 6A^4) = I$.

从而 $I - A$ 可逆，且 $(I - A)^{-1} = 2I - 3A^2 + A^3 + 6A^4$.

习题 1.35 (第四章第 18 题)

证明：不存在 n 阶复方阵 A, B 满足 $AB - BA = I_n$.

证明 假设存在，那么 $tr(AB - BA) = tr(I_n) = n$，但是 $tr(AB - BA) = tr(AB) - tr(BA) = 0$，矛盾！

习题 1.36 (第四章第 19 题)

证明：可逆上 (下) 三角、准对角、对称、反对称方阵的逆矩阵仍然分别是上 (下) 三角、准对角、对称、反对称的。

证明 以下均设原可逆方阵为 A，A_{ij} 为 A 的代数余子式，A^* 为 A 的伴随方阵。

(1) 上 (下) 三角：仅考虑上三角，下三角同理。由 $(A^*)_{ij} = A_{ji}$，则 $i > j$ 时，我们证明 $A_{ji} = 0$，利用数学归纳

法，$n=2$ 时显然，假设 $n=k$ 时命题成立，那么 $n=k+1$ 时，取 $(k+1)$ 阶可逆方阵 $A'=\begin{pmatrix} A & \boldsymbol{\alpha} \\ & a \end{pmatrix}, a\neq 0$ 由归纳假设，$j<i\leqslant k$ 时，A'_{ji} 左上角的 $(k-1)$ 阶子式为 A_{ji}，最后一个对角元为 a，最后一行除对角元外为 0，则将 A'_{ji} 按最后一行展开可得 $A'_{ji}=(-1)^{k+1}aA_{ji}=0$，故只需证 $A'_{j,k+1}=0, j=1,2,\cdots,k$，而这是显然的，因为 $A'_{j,k+1}$ 的最后一行全为 0。至此我们得到了 A^* 为上三角阵，故 A^{-1} 也为上三角阵。

(2) 准对角：易知准对角阵每个对角块都是可逆方阵，设 $A=diag(A_1,\cdots,A_k)$，则取 $B=diag(A_1^{-1},\cdots,A_k^{-1})$，代入验证可知 B 即为 A 的逆矩阵，从而也是准对角方阵。

(3) 对称：由 $A=A^T, AA^{-1}=I$，两边取转置得 $(A^{-1})^TA^T=(A^{-1})^TA=I, A^{-1}=(A^{-1})^T$，即 A^{-1} 对称。

(4) 反对称：由 $-A=A^T, AA^{-1}=I$，同 (3) 有 $(A^{-1})^TA^T=-(A^{-1})^TA=I, A^{-1}=-(A^{-1})^T$，即 A^{-1} 反对称。

习题 1.37 (第四章第 22 题)

设 A,B 是 n 阶方阵，λ 是数。证明：

$$(1)\ (\lambda A)^*=\lambda^{n-1}A^*;\quad (2)\ (AB)^*=B^*A^*;\quad (3)\ \det(A^*)=(\det(A))^{n-1}.$$

证明 (1) $(\lambda A)^*$ 的 (i,j) 位置元素为 λA 的 (j,i) 位置元素对应的代数余子式，为一个 $(n-1)$ 阶行列式，故其值为 A_{ji} 的 λ^{n-1} 倍，故有 $(\lambda A)^*=\lambda^{n-1}A^*$.

(2) 若 A,B 均可逆，则 $A^*=|A|A^{-1}, B^*=|B|B^{-1}$，且 AB 也可逆，则 $(AB)^*=|AB|(AB)^{-1}=|A||B|(B^{-1}A^{-1})$ $=(|B|B^{-1})(|A|A^{-1})=B^*A^*$. 而 A,B 中某个方阵不可逆时，考虑方阵 $tI+A, tI+B$，这两个方阵的行列式都是关于 t 的多项式，只有有限个根，那么必存在原点的去心邻域 $M=(-\delta,0)\cup(0,\delta)$ 满足 $\forall t\in M, tI+A, tI+B$ 均可逆，则由可逆情形时结论 $((tI+A)(tI+B))^*=(tI+B)^*(tI+A)^*$，而该式两边均为 n 阶方阵，且其元素都是关于 t 的多项式，从而关于 t 连续，则令 $t\to 0$，得到 $(AB)^*=B^*A^*$.

(3) 若 A 可逆，则 $A^*=|A|A^{-1}$，两边取行列式可得 $|A^*|=|A|^n|A^{-1}|=|A|^{n-1}$. 而 A 不可逆时，同 (2) 理，存在原点的去心邻域 $M=(-\delta,0)\cup(0,\delta)$ 满足 $\forall t\in M, tI+A$ 可逆，则由可逆情形时结论 $|(tI+A)^*|=|tI+A|^{n-1}$. 而上式两边均为关于 t 的多项式，故连续，则令 $t\to 0$，得到 $|A^*|=|A|^{n-1}$，即 $\det(A^*)=(\det(A))^{n-1}$.

笔记 这道题后两问用到的方法称为“微扰法”，当某个命题在矩阵可逆情形下比较容易证明时，往往采用这种办法来推广至一般矩阵，是很多矩阵证明问题里的常用方法。

习题 1.38 (第四章第 25 题)

设 n 阶方阵 A 的每行、每列元素之和都是 0，证明：A^* 的所有元素都相等。

证明 这等价于证明 A 的所有代数余子式 A_{ij} 相等，我们先证明 A 每行对应的代数余子式相等，不妨取第一行。由于 A 的每行元素之和都是 0，我们把 A_{11} 中其余列均加到列指标为 k 的列 $(k=2,\cdots,n)$，可得：

$$A_{11}=\begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}=\begin{vmatrix} a_{22} & \cdots & a_{2,k-1} & -a_{21} & a_{2,k+1} & \cdots & a_{2n} \\ a_{32} & \cdots & a_{3,k-1} & -a_{31} & a_{3,k+1} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n2} & \cdots & a_{n,k-1} & -a_{n1} & a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}$$

$$=(-1)^{k-2+1}\begin{vmatrix} a_{21} & \cdots & a_{2,k-1} & a_{2,k+1} & \cdots & a_{2n} \\ a_{31} & \cdots & a_{3,k-1} & a_{3,k+1} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}=(-1)^{k-1}M_{1k}=A_{1k}.$$

故 A 第一行对应的代数余子式均相等，同理可知对其他行也成立，再由 A 每列元素之和也都是 0，同理可知 A

每列对应的代数余子式也均相等，那么 A 每个元素对应的代数余子式均相等，从而 A^* 的所有元素相等。

习题 1.39 (第四章第 34 题 (1)(4)(5))

计算下列矩阵的逆矩阵

$$(1)\begin{pmatrix}1&0&1&-4\\-1&-3&-4&-2\\2&-1&4&4\\2&3&-3&2\end{pmatrix}\quad(4)\begin{pmatrix}&&&A_1\\&&A_2&\\&\iddots&&\\A_k&&&\end{pmatrix},A_i\text{可逆}\quad(5)\begin{pmatrix}1+a_1&1&\cdots&1\\1&1+a_2&\ddots&\vdots\\\vdots&\ddots&\ddots&1\\1&\cdots&1&1+a_n\end{pmatrix}$$

解 (1) 逆天计算题，但不排除考试考这种题的可能性，只能小心点别算错了。

$$\left(\begin{array}{cccc|cccc}1&0&1&-4&1&0&0&0\\-1&-3&-4&-2&0&1&0&0\\2&-1&4&4&0&0&1&0\\2&3&-3&2&0&0&0&1\end{array}\right)\to\left(\begin{array}{cccc|cccc}1&0&1&-4&1&0&0&0\\0&-3&-3&-6&1&1&0&0\\0&-1&2&12&-2&0&1&0\\0&3&-5&10&-2&0&0&1\end{array}\right)$$

$$\to\left(\begin{array}{cccc|cccc}1&0&1&-4&1&0&0&0\\0&-3&-3&-6&1&1&0&0\\0&0&3&14&-7/3&-1/3&1&0\\0&0&-8&4&-1&1&0&1\end{array}\right)\to\left(\begin{array}{cccc|cccc}1&0&1&-4&1&0&0&0\\0&1&1&2&-1/3&-1/3&0&0\\0&0&1&14/3&-7/9&-1/9&1/3&0\\0&0&0&124/3&-65/9&1/9&8/3&1\end{array}\right)$$

$$\to\left(\begin{array}{cccc|cccc}1&0&1&-4&1&0&0&0\\0&1&1&2&-1/3&-1/3&0&0\\0&0&1&14/3&-7/9&-1/9&1/3&0\\0&0&0&1&-65/372&1/372&24/372&9/372\end{array}\right)$$

$$\to\left(\begin{array}{cccc|cccc}1&0&1&0&28/93&1/93&8/31&3/31\\0&1&1&0&1/62&-21/62&-4/31&-3/62\\0&0&1&0&7/186&-23/186&1/31&-7/62\\0&0&0&1&-65/372&1/372&24/372&9/372\end{array}\right)$$

$$\to\left(\begin{array}{cccc|cccc}1&0&0&0&49/186&25/186&7/31&13/62\\0&1&0&0&-2/93&-20/93&-5/31&2/31\\0&0&1&0&7/186&-23/186&1/31&-7/62\\0&0&0&1&-65/372&1/372&2/31&3/124\end{array}\right).$$

(4) 观察原矩阵形式，直接取 $B=\begin{pmatrix}&&&A_k^{-1}\\&&A_{k-1}^{-1}&\\&\iddots&&\\A_1^{-1}&&&\end{pmatrix}$，代入验证知 B 即为原矩阵的逆矩阵。

(5) 先证明这样一个引理：

引理 1.1 (Sherman-Morrison 公式)

设 A 是 n 阶可逆阵，$\boldsymbol{\alpha},\boldsymbol{\beta}$ 是 n 维列向量，且 $1+\boldsymbol{\beta}^TA^{-1}\boldsymbol{\alpha}\neq0$，则有：

$$(A+\boldsymbol{\alpha}\boldsymbol{\beta}^T)^{-1}=A^{-1}-\frac{1}{1+\boldsymbol{\beta}^TA^{-1}\boldsymbol{\alpha}}A^{-1}\boldsymbol{\alpha}\boldsymbol{\beta}^TA^{-1}.$$

证明 直接计算可得：

$$
\begin{aligned}
&(A+\boldsymbol{\alpha}\boldsymbol{\beta}^T)\left(A^{-1}-\frac{1}{1+\boldsymbol{\beta}^TA^{-1}\boldsymbol{\alpha}}A^{-1}\boldsymbol{\alpha}\boldsymbol{\beta}^TA^{-1}\right)\\
=&I+\boldsymbol{\alpha}\boldsymbol{\beta}^TA^{-1}-\frac{1}{1+\boldsymbol{\beta}^TA^{-1}\boldsymbol{\alpha}}AA^{-1}\boldsymbol{\alpha}\boldsymbol{\beta}^TA^{-1}-\frac{1}{1+\boldsymbol{\beta}^TA^{-1}\boldsymbol{\alpha}}\boldsymbol{\alpha}\boldsymbol{\beta}^TA^{-1}\boldsymbol{\alpha}\boldsymbol{\beta}^TA^{-1}\\
=&I+\boldsymbol{\alpha}\boldsymbol{\beta}^TA^{-1}-\frac{1}{1+\boldsymbol{\beta}^TA^{-1}\boldsymbol{\alpha}}\boldsymbol{\alpha}\boldsymbol{\beta}^TA^{-1}-\frac{\boldsymbol{\beta}^TA^{-1}\boldsymbol{\alpha}}{1+\boldsymbol{\beta}^TA^{-1}\boldsymbol{\alpha}}\boldsymbol{\alpha}\boldsymbol{\beta}^TA^{-1}\\
=&I+\boldsymbol{\alpha}\boldsymbol{\beta}^TA^{-1}-\frac{1+\boldsymbol{\beta}^TA^{-1}\boldsymbol{\alpha}}{1+\boldsymbol{\beta}^TA^{-1}\boldsymbol{\alpha}}\boldsymbol{\alpha}\boldsymbol{\beta}^TA^{-1}=I.\quad \text{从而引理得证。}
\end{aligned}
$$

接下来回到原题，利用习题2.3的结论先计算该矩阵行列式：

$$
\begin{vmatrix}
1+a_1 & 1 & \cdots & 1\\
1 & 1+a_2 & \ddots & \vdots\\
\vdots & \ddots & \ddots & 1\\
1 & \cdots & 1 & 1+a_n
\end{vmatrix}
=\prod_{i=1}^{n}a_i+\sum_{k=1}^{n}\left(\prod_{i\neq k}a_i\right)\neq 0 \text{ 时}
$$

原矩阵可逆，以下我们均假设此条件成立，容易看出此时 a_i 中至多有一个为 0。

(I) $a_i\neq 0\,(i=1,2,\cdots,n)$ 时：

$$
\text{令 } A=\begin{pmatrix}a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n\end{pmatrix},\boldsymbol{\alpha}=\boldsymbol{\beta}=\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix},\text{ 则 }\begin{pmatrix}
1+a_1 & 1 & \cdots & 1\\
1 & 1+a_2 & \ddots & \vdots\\
\vdots & \ddots & \ddots & 1\\
1 & \cdots & 1 & 1+a_n
\end{pmatrix}=A+\boldsymbol{\alpha}\boldsymbol{\beta}^T
$$

此时 A 是 n 阶可逆方阵，且容易算得：

$$
A^{-1}=\begin{pmatrix}a_1^{-1} & & & \\ & a_2^{-1} & & \\ & & \ddots & \\ & & & a_n^{-1}\end{pmatrix},1+\boldsymbol{\beta}^TA^{-1}\boldsymbol{\alpha}=1+\sum_{i=1}^{n}a_i^{-1}
$$

$$
A^{-1}\boldsymbol{\alpha}\boldsymbol{\beta}^TA^{-1}=\begin{pmatrix}
a_1^{-2} & a_1^{-1}a_2^{-1} & \cdots & a_1^{-1}a_n^{-1}\\
a_2^{-1}a_1^{-1} & a_2^{-2} & \cdots & a_2^{-1}a_n^{-1}\\
\vdots & \vdots & \ddots & \vdots\\
a_n^{-1}a_1^{-1} & a_n^{-1}a_2^{-1} & \cdots & a_n^{-2}
\end{pmatrix}
$$

那么由引理1.1，原矩阵的逆矩阵为：

$$
(A+\boldsymbol{\alpha}\boldsymbol{\beta}^T)^{-1}=A^{-1}-\frac{1}{1+\boldsymbol{\beta}^TA^{-1}\alpha}A^{-1}\boldsymbol{\alpha}\boldsymbol{\beta}^TA^{-1}
$$

$$
=\begin{pmatrix}a_1^{-1} & & & \\ & a_2^{-1} & & \\ & & \ddots & \\ & & & a_n^{-1}\end{pmatrix}-\frac{1}{1+\sum\limits_{i=1}^{n}a_i^{-1}}\begin{pmatrix}
a_1^{-2} & a_1^{-1}a_2^{-1} & \cdots & a_1^{-1}a_n^{-1}\\
a_2^{-1}a_1^{-1} & a_2^{-2} & \cdots & a_2^{-1}a_n^{-1}\\
\vdots & \vdots & \ddots & \vdots\\
a_n^{-1}a_1^{-1} & a_n^{-1}a_2^{-1} & \cdots & a_n^{-2}
\end{pmatrix}
$$

$$
=\begin{pmatrix}
a_1^{-1}-sa_1^{-2} & -sa_1^{-1}a_2^{-1} & \cdots & -sa_1^{-1}a_n^{-1}\\
-sa_2^{-1}a_1^{-1} & a_2^{-1}-sa_2^{-2} & \cdots & -sa_2^{-1}a_n^{-1}\\
\vdots & \vdots & \ddots & \vdots\\
-sa_n^{-1}a_1^{-1} & -sa_n^{-1}a_2^{-1} & \cdots & a_n^{-1}-sa_n^{-2}
\end{pmatrix},\text{其中 } s=\frac{1}{1+\sum\limits_{i=1}^{n}a_i^{-1}}.
$$

(II) $\exists j, a_j=0, a_i\neq 0\ (i\neq j)$ 时：

$$\text{取}a_j=\varepsilon\to 0\text{，则}\lim_{\varepsilon\to 0}s=0,\ \lim_{\varepsilon\to 0}(a_i^{-1}-sa_i^{-2})=a_i^{-1}\ (i\neq j),\ \lim_{\varepsilon\to 0}(-sa_i^{-1}a_k^{-1})=0\ (i,k\neq j)$$

$$\lim_{\varepsilon\to 0}\frac{s}{\varepsilon}=\lim_{\varepsilon\to 0}\frac{s}{a_j}=1,\implies \lim_{\varepsilon\to 0}(-sa_i^{-1}a_j^{-1})=-\frac{1}{a_i}\ (i\neq j)$$

$$\lim_{\varepsilon\to 0}(a_j^{-1}-sa_j^{-2})\xrightarrow{t=s-\frac{1}{a_j}}=\lim_{\varepsilon\to 0}\left(\frac{1}{a_j}-\frac{1}{\left(t+\frac{1}{a_j}\right)a_j^2}\right)=\lim_{\varepsilon\to 0}\frac{ta_j}{a_j(ta_j+1)}=t=1+\sum_{i\neq j}a_i^{-1}.$$

$$\text{综上 }(A+\boldsymbol{\alpha}\boldsymbol{\beta}^T)^{-1}=\begin{pmatrix} a_1^{-1} & & & -a_1^{-1} & & & \\ & \ddots & & \vdots & & & \\ & & a_{j-1}^{-1} & -a_{j-1}^{-1} & & & \\ -a_1^{-1} & \cdots & -a_{j-1}^{-1} & t & -a_{j+1}^{-1} & \cdots & -a_n^{-1} \\ & & & -a_{j+1}^{-1} & a_{j+1}^{-1} & & \\ & & & \vdots & & \ddots & \\ & & & -a_n^{-1} & & & a_n^{-1} \end{pmatrix}\text{，其中}t=1+\sum_{i\neq j}a_i^{-1}.$$

1.7 第九周作业

习题 1.40 (第四章第 37 题)

设 A 是 n 阶可逆方阵，a,b 为 n 维列向量. 证明：$A+ab^T$ 可逆当且仅当 $1+b^TA^{-1}a\neq 0$. 且 $A+ab^T$ 可逆时，$(A+ab^T)^{-1}=A^{-1}-\dfrac{A^{-1}ab^TA^{-1}}{1+b^TA^{-1}a}$.

证明 (1) 先证 $A+ab^T$ 可逆 $\iff 1+b^TA^{-1}a\neq 0$.

$$\begin{aligned} A+ab^T \text{ 可逆} &\iff det(A+ab^T)\neq 0\\ &\iff det(A)det(I+A^{-1}ab^T)\neq 0\\ &\iff det(I+A^{-1}ab^T)\neq 0 \end{aligned}$$

注意到 $det(I_m-PQ)=det(I_n-QP)$，则有 $det(I+A^{-1}ab^T)=det(1+b^TA^{-1}a)$

故 $A+ab^T$ 可逆 $\iff 1+b^TA^{-1}a\neq 0$.

(2) 再求 $A+ab^T$ 的逆.

对这种验证类的题目，可以直接把题目给的结果代入验证，这里不再赘述. 下面给出直接求逆的方法.

$(A+ab^T)^{-1}=(A(I+A^{-1}ab^T))^{-1}=(I+A^{-1}ab^T)^{-1}A^{-1}\overset{def}{=}BA^{-1}$

则有

$$(I+A^{-1}ab^T)B=I\to B+A^{-1}ab^TB=I \tag{1.51}$$

两边同时左乘 $A^{-1}ab^T$ 得

$$\begin{aligned} &A^{-1}ab^TB+A^{-1}ab^TA^{-1}ab^TB=A^{-1}ab^T\to A^{-1}ab^TB+A^{-1}a(b^TA^{-1}a)b^TB=A^{-1}ab^T\\ &\to(1+b^TA^{-1}a)A^{-1}ab^TB=A^{-1}ab^T\to A^{-1}ab^TB=\frac{A^{-1}ab^T}{1+b^TA^{-1}a} \end{aligned} \tag{1.52}$$

上式利用了 $b^TA^{-1}a$ 是数，可以随意交换位置

带回式 (1.51) 得 $B=I-\dfrac{A^{-1}ab^T}{1+b^TA^{-1}a}$，即 $(A+ab^T)^{-1}=BA^{-1}=A^{-1}-\dfrac{A^{-1}ab^TA^{-1}}{1+b^TA^{-1}a}$.

习题 1.41 (第四章第 38 题)

设 $A\in F^{m\times m},B\in F^{n\times n},C\in F^{n\times m}$, 且 A 为对称可逆方阵. 证明：存在可逆方阵 P 使得

$$P\begin{pmatrix}A & C^T\\ C & B\end{pmatrix}P^T\text{为准对角阵.}$$

证明 利用 Schur 公式 $\begin{pmatrix}I & \\ -CA^{-1} & I\end{pmatrix}\begin{pmatrix}A & B\\ C & D\end{pmatrix}\begin{pmatrix}I & -A^{-1}B\\ & I\end{pmatrix}=\begin{pmatrix}A & \\ & D-CA^{-1}B\end{pmatrix}$ 进行对角化：

$$\begin{pmatrix}I & \\ -CA^{-1} & I\end{pmatrix}\begin{pmatrix}A & C^T\\ C & B\end{pmatrix}\begin{pmatrix}I & -A^{-1}C^T\\ & I\end{pmatrix}=\begin{pmatrix}A & \\ & B-CA^{-1}C^T\end{pmatrix} \tag{1.53}$$

易见左式最左边的矩阵即为满足要求的 P：

$det(P)=det(\begin{pmatrix}I & \\ -CA^{-1} & I\end{pmatrix})=1$，故可逆；

$P^T=\begin{pmatrix}I & \\ -CA^{-1} & I\end{pmatrix}^T=\begin{pmatrix}I & (-CA^{-1})^T\\ & I\end{pmatrix}=\begin{pmatrix}I & -A^{-1}C^T\\ & I\end{pmatrix}$.

习题 1.42 (第四章第 40 题 (1)(3))

计算下列矩阵的秩. (1) $\begin{pmatrix} 3 & 2 & -1 & 9 \\ -2 & 1 & -4 & 2 \\ -1 & -2 & 3 & -2 \\ 3 & 2 & -1 & 9 \end{pmatrix}$ (3) $\begin{pmatrix} 1 & 4 & 9 & 16 \\ 4 & 9 & 16 & 25 \\ 9 & 16 & 25 & 36 \\ 16 & 25 & 36 & 49 \end{pmatrix}$

解 (1) 进行初等变换:

$$\begin{pmatrix} 3 & 2 & -1 & 9 \\ -2 & 1 & -4 & 2 \\ -1 & -2 & 3 & -2 \\ 3 & 2 & -1 & 9 \end{pmatrix} \to \begin{pmatrix} 3 & 2 & -1 & 9 \\ 0 & 7/3 & -14/3 & 8 \\ -1 & -2 & 3 & -2 \\ 3 & 2 & -1 & 9 \end{pmatrix} \to \begin{pmatrix} 3 & 2 & -1 & 9 \\ 0 & 7/3 & -14/3 & 8 \\ 0 & -4/3 & 8/3 & 1 \\ 3 & 2 & -1 & 9 \end{pmatrix}$$
$$\to \begin{pmatrix} 3 & 2 & -1 & 9 \\ 0 & 7/3 & -14/3 & 8 \\ 0 & -4/3 & 8/3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 3 & 2 & -1 & 9 \\ 0 & 7/3 & -14/3 & 8 \\ 0 & 0 & 0 & 39/7 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.54}$$

经变换后的矩阵行列式为 0，且容易找到一个非零三阶子式：取第 1、2、3 行和第 1、2、4 列，故矩阵秩为 3

其实也可以直接从原矩阵出发：观察易知第一行和第四行相等，故其行列式为 0，再找到一个非零三阶子式即可说明秩为 3

但进行一定程度的初等变换可以让我们更容易地找到非零子式：找对角元均非零的上三角子式即可.

(2) 进行初等变换：(一般步骤应遵循高斯消元法的过程，由本题的特殊结构给出一种更简单的变换方法)

$$\begin{pmatrix} 1 & 4 & 9 & 16 \\ 4 & 9 & 16 & 25 \\ 9 & 16 & 25 & 36 \\ 16 & 25 & 36 & 49 \end{pmatrix} \xrightarrow{\text{依次减去上一行}} \begin{pmatrix} 1 & 4 & 9 & 16 \\ 3 & 5 & 7 & 9 \\ 5 & 7 & 9 & 11 \\ 7 & 9 & 11 & 13 \end{pmatrix} \to \begin{pmatrix} 1 & 4 & 9 & 16 \\ 3 & 5 & 7 & 9 \\ 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & 4 & 9 & 16 \\ 3 & 5 & 7 & 9 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 4 & 6 \\ 0 & 3 & 8 & 15 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.55}$$

化简后的矩阵行列式为 0，容易找到其左上角的三阶子式不为零，故秩为 3.

习题 1.43 (第四章第 41 题)

对于 a, b 的各种取值，讨论实矩阵 $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & a \\ 3 & b & 9 \end{pmatrix}$ 的秩.

解 进行初等变换:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & a \\ 3 & b & 9 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & a-6 \\ 0 & b-6 & 0 \end{pmatrix} \tag{1.56}$$

对 a,b 的取值进行讨论如下：

(1) a,b 均不为 6 时，秩为 3； (2) a,b 有一个 6 时，秩为 2； (3) a,b 均为 6 时，秩为 1.

习题 1.44 (第四章第 42 题)

设 A 为 n 阶方阵，且 $A^2=I$. 求方阵 $diag(I+A,I-A)$ 的相抵标准型.

解 进行分块初等变换：

$$\begin{pmatrix} I+A & \\ O & I-A \end{pmatrix} \to \begin{pmatrix} I+A & O \\ I+A & I-A \end{pmatrix} \to \begin{pmatrix} I+A & O \\ 2I & I-A \end{pmatrix} \to \begin{pmatrix} O & (I+A)(I-A)/2 \\ 2I & I-A \end{pmatrix}$$
$$\to \begin{pmatrix} O & (I-A^2)/2 \\ 2I & I-A \end{pmatrix} \to \begin{pmatrix} O & O \\ 2I & I-A \end{pmatrix} \to \begin{pmatrix} O & O \\ 2I & O \end{pmatrix} \to \begin{pmatrix} I_n & O \\ O & O \end{pmatrix} \tag{1.57}$$

即相抵标准型为 $\begin{pmatrix} I_n & O \\ O & O \end{pmatrix}$

注 相抵标准型要和原矩阵大小相同，有的同学这道题直接写的 I_n，这是不正确的.

习题 1.45 (第四章第 46 题)

设 n 阶方阵 A 满足 $A^2=I$, 证明：$rank(I+A)+rank(I-A)=n$.

证明 法一：由 $rank(A)+rank(B)=rank(diag(A,B))$ 知本题与 42 题等价.

法二：利用秩不等式：(1) $rank(I+A)+rank(I-A)\geq rank((I+A)+(I-A))=n$

(2) 由 Sylvester 不等式：$rank(I+A)+rank(I-A)\leq rank((I+A)(I-A))+n=rank(O)+n=n$
综合 (1)(2) 知命题成立.

习题 1.46 (第五章第 1 题)

设 $\boldsymbol{a}_1=(1,2,-1),\boldsymbol{a}_2=(2,0,3),\boldsymbol{a}_3=(2,1,0)$ 是三维几何空间中的三个向量. 能否将其中某一个向量写成其它两个向量的线性组合? 这三个向量是否共面?

解 注意到

$$\begin{vmatrix} 1 & 2 & -1 \\ 2 & 0 & 3 \\ 2 & 1 & 0 \end{vmatrix} = 7 \neq 0.$$

故向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 线性无关，因此其中任一向量不能表示为其它两个向量的线性组合，进而这三个向量不共面.

习题 1.47 (第五章第 3 题)

在 $\mathbb{F}^4$ 中，判断向量 $\boldsymbol{b}$ 能否写成 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 的线性组合. 若能，写出一种表达方式.

(1) $\boldsymbol{a}_1=(-1,3,0,-5),\boldsymbol{a}_2=(2,0,7,-3),\boldsymbol{a}_3=(-4,1,-2,6),\boldsymbol{b}=(8,3,-1,-25)$.

(2) $\boldsymbol{a}_1=(3,-5,2,-4)^T,\boldsymbol{a}_2=(-1,7,-3,6)^T,\boldsymbol{a}_3=(3,11,-5,10)^T,\boldsymbol{b}=(2,-30,13,-26)^T$.

解 (1) 这个问题等价于线性方程组 $x_1\alpha_1^T+x_2\alpha_2^T+x_3\alpha_3^T=b^T$ 是否有解，若有解，给出一组解. 考虑增广矩阵

$$\begin{pmatrix} -1 & 2 & -4 & 8 \\ 3 & 0 & 1 & 3 \\ 0 & 7 & -2 & -1 \\ -5 & -3 & 6 & -25 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -2 & 4 & -8 \\ 0 & 1 & -2 & 5 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

解得 $(x_1,x_2,x_3)=(2,-1,-3)$, 故 $b=2\alpha_1-\alpha_2-3\alpha_3$.

(2) 类似于 (1) 考虑线性方程组 $x_1\alpha_1+x_2\alpha_2+x_3\alpha_3=b$. 对增广矩阵作初等变换

$$\begin{pmatrix} 3 & -1 & 3 & 2 \\ -5 & 7 & 11 & -30 \\ 2 & -3 & -5 & 13 \\ -4 & 4 & 10 & -26 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 8 & -11 \\ 0 & 17 & 51 & -85 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

解得 $(x_1,x_2,x_3)=(-3,-8,1)$, 故 $b=-3\alpha_1-8\alpha_2+\alpha_3$.

习题 1.48 (第五章第 4 题)

设 $\boldsymbol{a}_1=(1,0,0,0),\boldsymbol{a}_2=(1,1,0,0),\boldsymbol{a}_3=(1,1,1,0),\boldsymbol{a}_4=(1,1,1,1,1)$. 证明：$\mathbb{F}^4$ 中任何向量可以写成 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3,\boldsymbol{a}_4$ 的线性组合，且表示唯一.

证明 对任意向量 $\boldsymbol{b}=(b_1,b_2,b_3,b_4)\in\mathbb{F}^4$, 考虑线性方程组 $x_1\alpha_1^T+x_2\alpha_2^T+x_3\alpha_3^T+x_4\alpha_4^T=b^T$. 原问题等价于此线性方程组对任意向量 $\boldsymbol{b}$ 是否存在解，且解唯一. 注意到系数矩阵

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

非退化，由 $Cramer$ 法则知该线性方程组存在唯一解.

习题 1.49 (第五章第 5 题)

设 $\boldsymbol{P}_i=(x_i,y_i,z_i),i=1,2,3,4$ 是三维几何空间中的点. 证明：$\boldsymbol{P}_i,i=1,2,3,4$ 共面的条件是

$$\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix}=0.$$

证明 $\boldsymbol{P}_i,i=1,2,3,4$ 共面等价于 $\overrightarrow{P_1P_2},\overrightarrow{P_1P_3},\overrightarrow{P_1P_4}$ 线性相关，也等价于行列式

$$\begin{vmatrix} x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \\ x_4-x_1 & y_4-y_1 & z_4-z_1 \end{vmatrix}=0.$$

即等价于行列式

$$\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix}=0.$$

习题 1.50 (第五章第 7 题)

设 $b_1,b_2,\cdots,b_s$ 中的每一个向量是 n 维向量 $a_1,a_2,\cdots,a_r$ 的线性组合. 证明：$b_1,b_2,\cdots,b_s$ 的任何线性组合都是 $a_1,a_2,\cdots,a_r$ 的线性组合.

证明 由于 $b_1,b_2,\cdots,b_s$ 中的每一个向量是 n 维向量 $a_1,a_2,\cdots,a_r$ 的线性组合，可假设

$$(b_1,b_2,\cdots,b_s)=(a_1,a_2,\cdots,a_r)A, \text{其中} A\in\mathbb{F}^{r\times s}. \quad \text{于是:}$$

$k_1b_1+\cdots+k_sb_s=(b_1,b_2,\cdots,b_s)(k_1,k_2,\cdots,k_s)^T=(a_1,a_2,\cdots,a_r)A(k_1,k_2,\cdots,k_s)^T$, 其中$k_i\in\mathbb{F},\forall i=1,\cdots,s$. 故 $b_1,b_2,\cdots,b_s$ 的任何线性组合都是 $a_1,a_2,\cdots,a_r$ 的线性组合.

习题 1.51 (第五章第 9 题)

判别下列线性方程组是否线性相关

$$(1)\begin{cases}-x_1+2x_2+3x_3=4\\5x_1-x_3=-1\\8x_1-6x_2-10x_3=-13\end{cases}\qquad(2)\begin{cases}2x_1+3x_2+x_3+5x_4=2\\3x_1+2x_2+4x_3+2x_4=3\\x_1+x_2+2x_3+4x_4=1\end{cases}$$

解 (1) 对线性方程组的增广矩阵作如下初等变换

$$\begin{pmatrix}-1&2&3&4\\5&0&-1&-1\\8&-6&-10&-13\end{pmatrix}\xrightarrow{5r_1+r_2,8r_1+r_3}\begin{pmatrix}-1&2&3&4\\0&10&14&19\\0&10&14&19\end{pmatrix}.$$

因此 $5r_1+r_2=8r_1+r_3$, 进而 $r_2=3r_1+r_3$.

(2) 注意到系数矩阵之子式

$$\begin{pmatrix}2&3&1\\3&2&4\\1&1&2\end{pmatrix}$$

的行列式不为零，系数矩阵秩为 3，因此线性无关.

习题 1.52 (第五章第 10 题 (1)(4))

判断下列向量组是否线性相关

(1) $a_1=(1,1,1),a_2=(1,-2,3),a_3=(1,4,9)$;

(4) $a_1=(1,-1,0,0),a_2=(0,1,-1,0),a_3=(0,0,1,-1),a_4=(-1,0,0,1)$.

解 (1) 考虑行列式

$$\begin{vmatrix}1&1&1\\1&-2&3\\1&4&9\end{vmatrix}=-30\neq0.$$

故 a_1,a_2,a_3 线性无关.

(4) 考虑行列式

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

故 a_1,a_2,a_3,a_4 线性相关.

习题 1.53 (第五章第 11 题)

证明：任何一个经过以下两个平面

$$\pi_1:A_1x+B_1y+C_1z+D_1=0,\pi_2:A_2x+B_2y+C_2z+D_2=0$$

交线的平面的方程能写成：

$$\lambda(A_1x+B_1y+C_1z+D_1)+\mu(A_2x+B_2y+C_2z+D_2)=0.$$

其中 λ,μ 为不全为零的常数.

证明 设 π_1,π_2 所确定的直线为 l,$\pi: Ax+By+Cz+D=0$ 为经过 l 的平面方程. 故方程组

$$\begin{cases} A_1x+B_1y+C_1z+D_1=0 \\ A_2x+B_2y+C_2z+D_2=0 \end{cases} \tag{1.58}$$

与

$$\begin{cases} A_1x+B_1y+C_1z+D_1=0 \\ A_2x+B_2y+C_2z+D_2=0 \\ Ax+By+Cz+D=0 \end{cases} \tag{1.59}$$

同解，故 (A,B,C,D) 可由 $(A_1,B_1,C_1,D_1),(A_2,B_2,C_2,D_2)$ 线性表出. 即存在常数 λ,μ, 使得 $Ax+By+Cz+D=\lambda(A_1x+B_1y+C_1z+D_1)+\mu(A_2x+B_2y+C_2z+D_2)$. 故平面 π 的方程即为 $\lambda(A_1x+B_1y+C_1z+D_1)+\mu(A_2x+B_2y+C_2z+D_2)=0$.

注 我们对上述证明中“(A,B,C,D) 可由 $(A_1,B_1,C_1,D_1),(A_2,B_2,C_2,D_2)$ 线性表出”作详细解释：
注意到 A_1,B_1,C_1 不全为 0，不妨假设 $A_1\neq 0$，考虑方程组 (1.1) 的增广矩阵，

$$\begin{pmatrix} A_1 & B_1 & C_1 & -D_1 \\ A_2 & B_2 & C_2 & -D_2 \end{pmatrix} \longrightarrow \begin{pmatrix} A_1 & B_1 & C_1 & -D_1 \\ 0 & B_2-\frac{B_1}{A_1} & C_2-\frac{C_1}{A_1} & -D_2+\frac{D_1}{A_1} \end{pmatrix}$$

注意到 $B_2-\frac{B_1}{A_1}$ 与 $C_2-\frac{C_1}{A_1}$ 不能同时为 0. 若同时为 0，当 $-D_2+\frac{D_1}{A_1}=0$ 时，则平面 π_1,π_2 重合；当 $-D_2+\frac{D_1}{A_1}\neq 0$ 时，则方程组 (1.1) 无解，即平面 π_1,π_2 无交 (与交为直线，无数组解相矛盾!). 记 $B_2'=B_2-\frac{B_1}{A_1},C_2'=C_2-\frac{C_1}{A_1},-D_2'=-D_2+\frac{D_1}{A_1}$. 不妨假设 $B_2'\neq 0$, 考虑方程组 (1.2) 的增广矩阵

$$\begin{pmatrix} A_1 & B_1 & C_1 & -D_1 \\ A_2 & B_2 & C_2 & -D_2 \\ A & B & C & -D \end{pmatrix} \longrightarrow \begin{pmatrix} A_1 & B_1 & C_1 & -D_1 \\ 0 & B_2' & C_2' & -D_2' \\ A & B & C & -D \end{pmatrix} \xrightarrow{\text{利用第一行第二行消去第三行前两列元素}} \begin{pmatrix} A_1 & B_1 & C_1 & -D_1 \\ 0 & B_2' & C_2' & -D_2' \\ 0 & 0 & C' & -D' \end{pmatrix}$$

由于方程组 (1.1) 与 (1.2) 同解，又方程组 (1.1) 有无数组解，故方程组 (1.2) 也有无数解，因此 $C'=0$(注意，若 $C'\neq 0$, 根据 *Cramer* 法则，方程组 (1.2) 的解存在且唯一，矛盾!). 若 $D'\neq 0$, 则 (1.2) 无解. 综上，$C'=0$ 且 $D'=0$. 即 $(A,B,C,-D)$ 可由 $(A_1,B_1,C_1,-D_1),(A_2,B_2,C_2,-D_2)$ 线性表出，也即 (A,B,C,D) 可由 $(A_1,B_1,C_1,D_1),(A_2,B_2,C_2,D_2)$ 线性表出.

习题 1.54 (第五章第 12 题)

下列说法是否正确？为什么？(1) 若 $a_1,a_2,\cdots,a_s(s\geq 2)$ 线性相关，则其中每一个向量都可以表示成其他向量的线性组合.
(2) 如果向量组的任何不是它本身的子向量组都线性无关，则该向量组也线性无关.
(3) 若向量组线性无关，则它的任何子向量组都线性无关.
(4)$\mathbb{F}^{n\times n}$ 的 $n+1$ 个向量组成的向量组必线性相关.
(5) 设 $a_1,a_2,\cdots,a_s$ 线性无关，则 $a_1+a_2,a_2+a_3,\cdots,a_s+a_1$ 必线性无关.
(6) 设 $a_1,a_2,\cdots,a_s$ 线性相关，则 $a_1+a_2,a_2+a_3,\cdots,a_s+a_1$ 必线性相关.
(7) 设 $a_1,a_2,\cdots,a_s\in\mathbb{F}^n$ 线性无关，则它们的加长向量组也必线性无关.
(8) 设 $a_1,a_2,\cdots,a_s\in\mathbb{F}^n$ 线性相关，则它们的加长向量组也必线性相关.

解 (1) 不正确，可以考虑向量组 $a_1,a_2,\cdots,a_s$ 中的极大无关组.

(2) 不正确，考虑 $(1,0)^T$ 与 $(2,0)^T$.

(3) 正确. 其逆否命题为“若存在一组子向量组线性相关，则该向量组线性相关”.

(4) 正确，含有 $n+1$ 个未定元的 n 个线性方程组必有非零解.

(5) 不正确，考虑 $s=2$ 时的情况.

(6) 正确，由教材定理 5.3.5 中第 3 条，知 $rank(a_1+a_2, a_2+a_3, \cdots, a_s+a_1) \le rank(a_1, a_2, \cdots, a_s) < s$.

(7) 正确，考虑线性方程组，注意到加长向量组相当于添加了更多的线性方程. 其逆否命题为“若加长向量组线性相关，则本身一定线性相关”.

(8) 不正确，考虑考虑向量组 $(1), (-1)$ 与其加长版本 $(1,0)^T, (-1,1)^T$.

习题 1.55 (第五章第 17 题)

设向量组 $a_1, a_2, \cdots, a_r$ 线性无关，且 $a_1, a_2, \cdots, a_r$ 可由向量组 $b_1, b_2, \cdots, b_r$ 线性表示，则 $b_1, b_2, \cdots, b_r$ 也线性无关.

证明 由教材定理 5.3.5 中第 3 条知向量组 $a_1, a_2, \cdots, a_r$ 的极大无关组个数小于等于向量组 $b_1, b_2, \cdots, b_r$ 的极大无关组个数. 再由向量组 $a_1, a_2, \cdots, a_r$ 线性无关，知向量组 $b_1, b_2, \cdots, b_r$ 也线性无关.

1.8 第十周作业

习题 1.56 (第五章第 19 题)

求下列向量组的极大无关组与秩：

(1) $\boldsymbol{a}_1 = (3, -2, 0)$, $\boldsymbol{a}_2 = (27, -18, 0)$, $\boldsymbol{a}_3 = (-1, 5, 8)$.

(2) $\boldsymbol{a}_1 = (1, -1, 2, 4)$, $\boldsymbol{a}_2 = (0, 3, 1, 2)$, $\boldsymbol{a}_3 = (3, 0, 7, 14)$, $\boldsymbol{a}_4 = (1, -1, 2, 0)$, $\boldsymbol{a}_5 = (2, 1, 5, 6)$.

(3) $\boldsymbol{a}_1 = (0, 1, 2, 3)$, $\boldsymbol{a}_2 = (1, 2, 3, 4)$, $\boldsymbol{a}_3 = (3, 4, 5, 6)$, $\boldsymbol{a}_4 = (4, 3, 2, 1)$, $\boldsymbol{a}_5 = (6, 5, 4, 3)$.

解 (1) 假设 $k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2 + k_3\boldsymbol{a}_3 = 0$，其系数矩阵为：

$$\begin{pmatrix} 3 & 27 & -1 \\ -2 & -18 & 5 \\ 0 & 0 & 8 \end{pmatrix} \text{进行初等行变换，得} \begin{pmatrix} 1 & 9 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

则 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 秩为 2，且有 $\boldsymbol{a}_2 = 9\boldsymbol{a}_1$，则 $\boldsymbol{a}_1, \boldsymbol{a}_3$ 或 $\boldsymbol{a}_2, \boldsymbol{a}_3$ 构成向量组的极大无关组。

(2) 假设 $k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2 + k_3\boldsymbol{a}_3 + k_4\boldsymbol{a}_4 + k_5\boldsymbol{a}_5 = 0$，其系数矩阵为：

$$\begin{pmatrix} 1 & 0 & 3 & 1 & 2 \\ -1 & 3 & 0 & -1 & 1 \\ 2 & 1 & 7 & 2 & 5 \\ 4 & 2 & 14 & 0 & 6 \end{pmatrix} \text{进行初等行变换，得} \begin{pmatrix} 1 & 0 & 3 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

则 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4, \boldsymbol{a}_5$ 秩为 3，且 $\boldsymbol{a}_3 = 3\boldsymbol{a}_1 + \boldsymbol{a}_2, \boldsymbol{a}_5 = \boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_4$，则 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_4$ 可构成向量组的极大无关组。

(3) 假设 $k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2 + k_3\boldsymbol{a}_3 + k_4\boldsymbol{a}_4 + k_5\boldsymbol{a}_5 = 0$，其系数矩阵为：

$$\begin{pmatrix} 0 & 1 & 3 & 4 & 6 \\ 1 & 2 & 4 & 3 & 5 \\ 2 & 3 & 5 & 2 & 4 \\ 3 & 4 & 6 & 1 & 3 \end{pmatrix} \text{进行初等行变换，得} \begin{pmatrix} 1 & 1 & 1 & -1 & -1 \\ 0 & 1 & 3 & 4 & 6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

则 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4, \boldsymbol{a}_5$ 秩为 2，且五个向量中任意两个都线性无关，故任意两个向量均可构成向量组的极大无关组。

习题 1.57 (第五章第 25 题)

求下列矩阵的秩，并求出它的行空间、列空间及零空间的一组基。

$$(1) \begin{pmatrix} 3 & -2 & 0 & 1 \\ -1 & -3 & 2 & 0 \\ 2 & 0 & -4 & 5 \\ 4 & 1 & -2 & 1 \end{pmatrix} \quad (2) \begin{pmatrix} 3 & 6 & 1 & 5 \\ 1 & 4 & -1 & 3 \\ -1 & -10 & 5 & -7 \\ 4 & -2 & 8 & 0 \end{pmatrix}$$

解 在两问中均记矩阵为 A，其四个行向量为 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4$，四个列向量为 $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3, \boldsymbol{b}_4$

(1) 进行初等行变换（不交换行）得：$\begin{pmatrix} 0 & 5 & -6 & 4 \\ 1 & 3 & -2 & 0 \\ 0 & -6 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$，矩阵秩为 3，则 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 可作为行空间的一组基。

进行初等列变换（不交换列）得：$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}$，则 $\boldsymbol{b}_1, \boldsymbol{b}_3, \boldsymbol{b}_4$ 可作为列空间的一组基。

利用行变换后的结果可得 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集为 $(8,30,49,36)^Tt, t\in\mathbb{R}$，则 $(8,30,49,36)^T$ 可作为零空间的一组基。

(2) 进行初等行变换（不交换行）得：$\begin{pmatrix}0&3&-2&2\\1&1&1&1\\0&0&0&0\\0&0&0&0\end{pmatrix}$，矩阵秩为 2，则 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 可作为行空间的一组基。

进行初等列变换（不交换列）得：$\begin{pmatrix}0&0&1&0\\-1&0&0&0\\4&0&1&0\\5&0&3&0\end{pmatrix}$，则 $\boldsymbol{b}_1,\boldsymbol{b}_3$ 可作为列空间的一组基。

利用行变换后的结果可得 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集为 $(-5,2,3,0)^Tt+(-1,-2,0,3)^Ts$，其中 $t,s\in\mathbb{R}$，则 $(-5,2,3,0)^T$ 与 $(-1,-2,0,3)^T$ 可作为零空间的一组基。

习题 1.58 (第五章第 28 题)

证明：n 阶方阵 A 可逆 $\iff \mathrm{rank}(A)=n \iff A$ 的行向量线性无关 $\iff A$ 的列向量线性无关

证明 n 阶方阵 A 可逆 $\iff \det(A)\neq 0 \iff \mathrm{rank}(A)=n$，同时由于矩阵的秩 = 行秩 = 列秩，则 $\mathrm{rank}(A)=n \iff A$ 的行向量线性无关 $\iff A$ 的列向量线性无关，命题得证。

习题 1.59 (第五章第 29 题)

设矩阵 $A\in\mathbb{F}^{m\times n}$ 的秩为 r，则 A 的不等于零的 r 阶子式所在行 (列) 构成 A 的行 (列) 向量的极大无关组。

证明 由对称性，仅考虑行向量情形即可。通过交换行的位置，可以将不为 0 的 r 阶子式所在的行调整到前 r 行，故不妨设 $A=\begin{pmatrix}A_r&B\\C&D\end{pmatrix}$，其中 $\det(A_r)\neq 0$，则 A_r 中 r 个 r 维行向量线性无关，则 $(A_r\ \ B)$ 作为其加长向量组也线性无关，再由矩阵的秩为 r，可知前 r 行即为行向量的极大无关组。

习题 1.60 (第五章第 31 题)

设 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ 是 $\mathbb{F}^n$ 的基，向量组 $\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n$ 与 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ 有关系式

$$(\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n)=(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n)T.$$

证明：$\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n$ 为 $\mathbb{F}^n$ 的基当且仅当 T 为可逆方阵。

证明 设 $T=(t_{ij})_{n\times n}$，则 $\boldsymbol{b}_j=\sum_{i=1}^n t_{ij}\boldsymbol{a}_i$，而 $(\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n)$ 为 $\mathbb{F}^n$ 的基当且仅当它们线性无关，即 $\sum_{j=1}^n\lambda_j\boldsymbol{b}_j=0$ 只有零解，也即 $\sum_{j=1}^n\sum_{i=1}^n\lambda_jt_{ij}\boldsymbol{a}_i=\sum_{i=1}^n\left(\sum_{j=1}^n\lambda_jt_{ij}\right)\boldsymbol{a}_i=0$ 只有零解，由 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 是基，则 $\sum_{i=1}^n\left(\sum_{j=1}^n\lambda_jt_{ij}\right)\boldsymbol{a}_i=0$ 只有 $\sum_{j=1}^n\lambda_jt_{ij}=0, i=1,\cdots,n$，即 $T\boldsymbol{\lambda}=\boldsymbol{0}$，其中 $\boldsymbol{\lambda}=(\lambda_1,\cdots,\lambda_n)^T$，其只有零解当且仅当 T 可逆，命题得证。

1.9 第十一周作业

习题 1.61 (第五章第 34 题)

以向量组 $a_1 = (3, 1, 0), a_2 = (6, 3, 2), a_3 = (1, 3, 5)$ 为基，求 $\beta = (2, -1, 2)$ 的坐标.

解 即求线性表示的系数，做初等变换如下：

$$\begin{pmatrix} 3 & 6 & 1 & 2 \\ 1 & 3 & 3 & -1 \\ 0 & 2 & 5 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 3 & -1 \\ 3 & 6 & 1 & 2 \\ 0 & 2 & 5 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 3 & -1 \\ 0 & -3 & -8 & 5 \\ 0 & 2 & 5 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 3 & -1 \\ 0 & -3 & -8 & 5 \\ 0 & 0 & -1/3 & 16/3 \end{pmatrix} \tag{1.60}$$

故坐标为 $(x_1, x_2, x_3)^T = (-76, 41, -16)^T$

习题 1.62 (第五章第 35 题)

设 $a_1 = (3, 2, -1, 4), a_2 = (2, 3, 0, -1)$.

(1) 将 a_1, a_2 扩充为 R^4 的一组基；

(2) 给出标准基在该组基下的表示；

(3) 求 $\beta = (1, 3, 4, -2)$ 在该组基下的坐标.

解 虽然题目给的向量是行向量，但我们还是习惯性地使用列向量来进行计算，即将所有向量转置.

(1) 注意到 a_1, a_2 的前两个分量线性无关，故可以直接补充 $a_3 = (0, 0, 1, 0)$，$a_4 = (0, 0, 0, 1)$ 得到一组基.

(2) 基变换写为

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

故 $$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}' = \begin{pmatrix} 3 & 2 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 4 & -1 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 4 & -1 & 0 & 1 \end{pmatrix}^{-1}$$

即标准基在这组基下的表示为 $$\begin{pmatrix} 3 & 2 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 4 & -1 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 3/5 & -2/5 & 0 & 0 \\ -2/5 & 3/5 & 0 & 0 \\ 3/5 & -2/5 & 1 & 0 \\ -14/5 & 11/5 & 0 & 1 \end{pmatrix}$$

(3) 由坐标变换公式，$\beta' = T^{-1}\beta = (-3/5, 7/5, 17/5, 9/5)$.

习题 1.63 (第五章第 36 题)

将三维几何空间中的直角坐标系 $[O; e_1, e_2, e_3]$ 绕单位向量 $e = \frac{1}{\sqrt{3}}(1, 1, 1)$ 逆时针旋转 θ 角，求新坐标与原坐标之间的关系.

解 将标准基 1 换为 $e_1' = e$ 及垂直于 e 的平面内互相垂直的两单位向量 $e_2' = \frac{1}{\sqrt{2}}(1, -1, 0), e_3' = \frac{1}{\sqrt{6}}(1, 1, -2)$ 组成的基 2，有过渡矩阵：

$$T_1=\begin{pmatrix}\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}}\end{pmatrix}$$

满足

$$(e_1,e_2,e_3)T_1=(e_1',e_2',e_3'); \tag{1.61}$$

再将基 2 绕 e 逆时针旋转 θ 角得到基 3：(e_1'',e_2'',e_3'')，有过渡矩阵 (即二维旋转)：

$$T_2=\begin{pmatrix}1 & 0 & 0\\ 0 & cos\theta & -sin\theta\\ 0 & sin\theta & cos\theta\end{pmatrix}$$

满足

$$(e_1',e_2',e_3')T_2=(e_1'',e_2'',e_3''); \tag{1.62}$$

由于旋转过程中基 1 与基 2 的相对位置保持不变，故基 1 经旋转得到的基 4：(e_1''',e_2''',e_3''') 与基 3 之间满足和基 1 与基 2 相同的关系：

$$(e_1''',e_2''',e_3''')T_1=(e_1'',e_2'',e_3''); \tag{1.63}$$

即

$$(e_1''',e_2''',e_3''')=(e_1'',e_2'',e_3'')T_1^{-1}. \tag{1.64}$$

综合式 (1.61),(1.62),(1.64) 得到标准基在旋转前后的基变换：

$$(e_1''',e_2''',e_3''')=(e_1,e_2,e_3)T_1T_2T_1^{-1} \tag{1.65}$$

故坐标变换为：

$$\alpha'=R\alpha=T_1T_2^{-1}T_1^{-1}\alpha \tag{1.66}$$

旋转矩阵写为:

$$\begin{aligned}R&=T_1T_2^{-1}T_1^{-1}\\ &=\begin{pmatrix}(1+2cos\theta)/3 & (1-cos\theta+\sqrt{3}sin\theta)/3 & (1-cos\theta-\sqrt{3}sin\theta)/3\\ (1-cos\theta-\sqrt{3}sin\theta)/3 & (1+2cos\theta)/3 & (1-cos\theta+\sqrt{3}sin\theta)/3\\ (1-cos\theta+\sqrt{3}sin\theta)/3 & (1-cos\theta-\sqrt{3}sin\theta)/3 & (1+2cos\theta)/3\end{pmatrix}\end{aligned} \tag{1.67}$$

注 本题也可以直接使用 Rodrigues 转动公式：

$$R(\vec{\psi})=\begin{pmatrix}n_1^2(1-cos\psi)+cos\psi & n_1n_2(1-cos\psi)-n_3sin\psi & n_1n_3(1-cos\psi)+n_2sin\psi\\ n_1n_2(1-cos\psi)+n_3sin\psi & n_2^2(1-cos\psi)+cos\psi & n_2n_3(1-cos\psi)-n_1sin\psi\\ n_1n_3(1-cos\psi)-n_2sin\psi & n_2n_3(1-cos\psi)+n_1sin\psi & n_3^2(1-cos\psi)+cos\psi\end{pmatrix}$$

由于坐标轴绕 e 逆时针转 θ 角，故相对地，坐标绕 e 顺时针转 θ 角，即 ψ 取 $-\theta$，而转轴 e 的三个分量即为 n_1,n_2,n_3，代入得

$$\begin{aligned}R&=\begin{pmatrix}(1-cos\theta)/3+cos\theta & (1-cos\theta)/3+\sqrt{3}sin\theta/3 & (1-cos\theta)/3-\sqrt{3}sin\theta/3\\ (1-cos\theta)/3-\sqrt{3}sin\theta/3 & (1-cos\theta)/3+cos\theta & (1-cos\theta)/3+\sqrt{3}sin\theta/3\\ (1-cos\theta)/3+\sqrt{3}sin\theta/3 & (1-cos\theta)/3-\sqrt{3}sin\theta/3 & (1-cos\theta)/3+cos\theta\end{pmatrix}\\ &=\begin{pmatrix}(1+2cos\theta)/3 & (1-cos\theta+\sqrt{3}sin\theta)/3 & (1-cos\theta-\sqrt{3}sin\theta)/3\\ (1-cos\theta-\sqrt{3}sin\theta)/3 & (1+2cos\theta)/3 & (1-cos\theta+\sqrt{3}sin\theta)/3\\ (1-cos\theta+\sqrt{3}sin\theta)/3 & (1-cos\theta-\sqrt{3}sin\theta)/3 & (1+2cos\theta)/3\end{pmatrix}\end{aligned} \tag{1.68}$$

习题 1.64 (第五章第 40 题)

已知 F^5 中向量 $a_1=(1,2,3,2,1)^T$ 及 $a_2=(1,3,2,1,2)^T$，求找一个齐次线性方程组使得 a_1 与 a_2 为该方程组的基础解系.

解 设系数矩阵为 A，则齐次线性方程组写为 $Ax=O$，a_1 和 a_2 构成方程组的基础解系，即方程组解空间维数等于 2，故 $rankA=5-dimV=5-2=3$，故找一个列数为 5，秩为 3，$Aa_1=O,Aa_2=O$ 的矩阵 A 即可，例如：

$$A=\begin{pmatrix}1&0&-1&1&0\\1&-1&0&0&1\\-4&1&0&1&0\end{pmatrix}\tag{1.69}$$

上式中 A 的每一行是如何确定的？实际上 A 的任意行向量 α 满足 $\begin{pmatrix}a_1^T\\a_2^T\end{pmatrix}\alpha^T=O$，这即是 $A(a_1,a_2)=O$ 转置后的结果.

习题 1.65 (第五章第 41 题)

判断下列集合关于规定的运算是否构成线性空间

(1) V 是所有实数对 (x,y) 的集合，数域 $\mathbb{F}=\mathbb{R}$. 定义

$$(x_1,y_1)+(x_2,y_2)=(x_1+x_2,y_1+y_2),\quad \lambda(x,y)=(x,y).$$

(2) V 是所有满足 $f(-1)=0$ 的实函数集合，数域 $\mathbb{F}=\mathbb{R}$. 定义加法为函数加法，数乘为数与函数的乘法.

(3) V 是所有满足 $f(0)\neq 0$ 的实函数的集合，数域 $\mathbb{F}=\mathbb{R}$. 定义加法为函数加法，数乘为数与函数的乘法.

(4) V 是数域 $\mathbb{F}$ 上所有 n 阶可逆方阵的全体，加法为矩阵的加法，数乘为矩阵的数乘.

解 (1) 不构成线性空间. 因为不满足数乘分配律：

$$(\lambda+\mu)(x,y)=(x,y)\neq(2x,2y)=(x,y)+(x,y)=\lambda(x,y)+\mu(x,y)\tag{1.70}$$

(2) 构成线性空间. 由于实函数在加法和数乘下构成线性空间，我们只需证 V 为实函数空间的子空间，即 V 封闭：设 $f,g\in V$，有 $(f+g)(-1)=f(-1)+g(-1)=0$，故 $f+g\in V$; $\lambda f(-1)=0$，故 $\lambda f\in V$.

(3) 不构成线性空间，因为 V 对加法不封闭：

设 $f\in V$，则 $f(0)\neq 0$，$-f(0)\neq 0$. 故 $-f\in V$，而 $f+(-f)=0$ 是恒为 0 的常量函数，其不在 V 中，故 V 对加法不封闭.

(4) 不构成线性空间，因为 V 对加法不封闭：

设 $A\in V$，则 A 可逆，显然 $-A$ 也可逆，但 $A+(-A)=O$ 不可逆，即零矩阵不在 V 中，V 对加法不封闭.

习题 1.66 (第五章第 42 题)

设 V 是所有实函数全体在实数域上构成的线性空间，判断 V 中下列函数组是否线性相关.

(1)$1,x,\sin x$;

(2)$1,x,e^x$;

(3)$1,\cos 2x,\cos^2 x$;

(4)$1,x^2,(x-1)^3,(x+1)^3$;

(5)$\cos x,\cos 2x,\cdots,\cos nx$.

解 (1) 线性无关. (2) 线性无关. (3) 线性相关，注意到 $\cos 2x=2\cos^2 x-1$.

(4) 线性相关，注意到 $(x+1)^3-(x-1)^3-6x^2-2=0$.

(5) 线性无关，我们先断言 $1, \sin x, \cos x, \sin 2x, \cos 2x, \cdots, \sin nx, \cos nx$ 线性无关. 根据第五章第 12 题的第三条，知 $\cos x, \cos 2x, \cdots, \cos nx$ 线性无关.

以上断言的证明：设

$$f(x) = a + b_1\sin x + c_1\cos x + b_2\sin 2x + c_2\cos 2x + \cdots + b_n\sin nx + c_n\cos nx = 0 \quad \text{其中} a, b_i, c_i \in \mathbb{R}.$$

依次设 $g(x) = 1, \sin x, \cos x, \sin 2x, \cos 2x, \cdots, \sin nx, \cos nx$，分别计算定积分 $\int_0^{2\pi} f(x)g(x)dx$，可得 $a = b_1 = c_1 = \cdots = b_n = c_n = 0$.

习题 1.67 (第五章第 44 题)

设 $\mathbb{F}_n[x]$ 是次数小于或等于 n 的多项式全体构成的线性空间.

(1) 证明：$S = \{1, x-1, (x-1)^2, \cdots, (x-1)^n\}$ 构成 $\mathbb{F}^n$ 的一组基；

(2) 求 S 到基 $T = \{1, x, \cdots, x^n\}$ 的过渡矩阵；

(3) 求多项式 $p(x) = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{F}_n[x]$ 在基 S 下的坐标.

证明 (1) 根据 $x^k = ((x-1)+1)^k (k = 0, 1, 2, \cdots, n)$ 的二项式展开及第五章第 17 题知 S 线性无关，又 $\mathbb{F}_n[x]$ 的维数为 $n+1$，故 S 构成 $\mathbb{F}_n[x]$ 的一组基.

解 (2) 根据二项式展开 $x^k = ((x-1)+1)^k = \sum\limits_{i=1}^{k} C_k^i (x-1)^i$. 故

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

(3)

$$p(x) = a_0 + a_1x + \cdots + a_nx^n = (1, x, \cdots, x^n)\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = (1, x-1, \cdots, (x-1)^n)\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 2 & \cdots & n \\ 0 & 0 & 1 & \cdots & \frac{n(n-1)}{2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

则 $p(x)$ 在 S 下的坐标为

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 2 & \cdots & n \\ 0 & 0 & 1 & \cdots & \frac{n(n-1)}{2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} p(1) \\ \frac{p'(1)}{1!} \\ \vdots \\ \frac{p^{(n)}(1)}{n!} \end{pmatrix}$$

即为 $p(x)$ 在 1 处的 $Taylor$ 展开.

习题 1.68 (第五章第 46 题)

给定矩阵

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

令 V 是与 A 乘法可交换的三阶实方阵全体. 证明: V 在矩阵加法与数乘下构成实数域上的线性空间, 并求 V 的一组基与维数.

解 根据线性空间及矩阵数乘定义知 V 构成线性空间. 注意到下列矩阵

$$A=\begin{pmatrix}0&0&1\\1&0&0\\4&-2&1\end{pmatrix},B=\begin{pmatrix}0&-2&0\\-1&-4&1\\-2&0&0\end{pmatrix},I_3$$

均与矩阵 A 可交换, 易知 A,B,I_3 线性无关. 下面只需说明对任意矩阵 $X=x_{ij}$, 若 X 与 A 乘法可交换, 则 X 可以被 A,B,I_3 线性表示. 比较 XA 与 AX 中各个分量的元素, 知 $X=x_{23}B+x_{11}I_3+x_{13}A$. 故 V 的基为 A,B,I_3, 维数为 3.

习题 1.69 (第五章第 47 题)

$V=\mathbb{F}^{n\times n}$ 是数域 $\mathbb{F}$ 上所有 n 阶矩阵构成的线性空间. 令 W 是数域 $\mathbb{F}$ 上所有满足 $\mathrm{tr}A=0$ 的 n 阶矩阵的全体. 证明: W 是 V 的线性子空间. 并求 W 的一组基和维数.

证明 根据迹的线性性质知 W 是 V 的线性子空间. 易知 $E_{ij}(i\neq j),E_{kk}-E_{nn}\in W$, 其中 $i,j=1,2,\cdots,n,k=1,2,\cdots,n-1$, 且线性无关. 下面只需说明 $E_{ij}(i\neq j),E_{kk}-E_{nn}\in W$ 可以张成 W 即可, 假设 $A\in W$, 即 $\mathrm{tr}A=0$. 则 $a_{11}+a_{22}+\cdots+a_{nn}=0$, 因此 $a_{nn}=-a_{11}-\cdots-a_{n-1,n-1}$. 于是 $A=\sum\limits_{i\neq j}a_{ij}E_{ij}+\sum\limits_{k\neq n}a_{kk}(E_{kk}-E_{nn})$. 因此 $E_{ij}(i\neq j),E_{kk}-E_{nn}\in W$ 构成 W 的一组基, 维数为 n^2-1.

1.10 第十二周作业

习题 1.70 (第六章第 5 题)

证明：$\mathbb{R}^2$ 上的可逆线性变换可以分解为关于坐标轴的伸缩、反射及旋转变换的复合。

证明 显然关于坐标轴的伸缩、反射及旋转变换都是可逆的，且其逆变换仍为关于坐标轴的伸缩、反射及旋转变换。且 $\mathbb{R}^2$ 上任意线性变换都与一个二阶方阵对应。因此我们任取一个可逆线性变换 $\mathscr{A}$ 满足 $\mathscr{A}\boldsymbol{x} = A\boldsymbol{x}$，其中 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 为可逆方阵，那么我们只需要证明可以通过将 A 与关于坐标轴的伸缩、反射、旋转变换对应的矩阵相乘来得到单位阵即可，我们先通过伸缩来使矩阵两列向量长度相等：

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} = \begin{pmatrix} \lambda a & \mu b \\ \lambda c & \mu d \end{pmatrix} \triangleq \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$$

由于 A 可逆，a, c 不同时为 0，b, d 不同时为 0，则必存在 $\lambda, \mu > 0$，满足 $\lambda^2(a^2+c^2) = \mu^2(b^2+d^2)$，这也就是 $a_1^2 + c_1^2 = b_1^2 + d_1^2$，并且由齐次性不妨设其为 1，再通过旋转来使两列向量关于 y 轴对称：

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} = \begin{pmatrix} a_1\cos\theta - c_1\sin\theta & b_1\cos\theta - d_1\sin\theta \\ a_1\sin\theta + c_1\cos\theta & b_1\sin\theta + d_1\cos\theta \end{pmatrix} \triangleq \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$$

设 $a_1 = \cos\alpha, c_1 = \sin\alpha, b_1 = \cos\beta, d_1 = \sin\beta$，那么我们有：

$$a_2 = \cos(\theta+\alpha),\ b_2 = \cos(\theta+\beta),\ c_2 = \sin(\theta+\alpha),\ d_2 = \sin(\theta+\beta)$$

取 $\theta = \dfrac{\pi - \alpha - \beta}{2}$，就有 $a_2 + b_2 = 0, c_2 = d_2$，再通过一次伸缩使两个列向量正交且均化为单位向量：

$$\begin{pmatrix} p & \\ & q \end{pmatrix}\begin{pmatrix} a_2 & -a_2 \\ c_2 & c_2 \end{pmatrix} = \begin{pmatrix} pa_2 & -pa_2 \\ qc_2 & qc_2 \end{pmatrix} \triangleq \begin{pmatrix} a_3 & -a_3 \\ c_3 & c_3 \end{pmatrix}$$

那么必存在 $p, q > 0$，满足 $q^2c_2^2 - p^2a_2^2 = 0$，且由齐次性不妨设 $q^2c_2^2 = p^2a_2^2 = \dfrac{1}{2}$，此时矩阵两列向量互相正交，且关于 y 轴对称，均为单位向量，与坐标轴夹角都是 $\dfrac{\pi}{4}$，所以若两列向量是正定向，只需再做一次旋转便可得到单位阵，若为负定向，则先做一次反射再旋转，也可得到单位阵，至此我们便完成了命题的证明。

习题 1.71 (第六章第 6 题)

在三维几何空间的直角坐标系中，求关于平面 $x + 2y + 3z = 0$ 的对称变换。

解 记所求变换为 T，所给平面为 S，任取 $a = (x, y, z)^T \in \mathbb{R}^3$，设其关于 S 的对称点为 $T(a) = (u, v, w)^T \triangleq b$，则 a, b 中点落在 S 上，且 a, b 之间的向量与 S 垂直，S 法向量为 $\overrightarrow{n} = (1, 2, 3)^T$，从而有：

$$x + u + 2(y+v) + 3(z+w) = 0,\ x - u = \frac{y-v}{2} = \frac{z-w}{3}$$

解得 $T(x, y, z)^T = (u, v, w)^T = \left(\dfrac{6}{7}x - \dfrac{2}{7}y - \dfrac{3}{7}z, -\dfrac{2}{7}x + \dfrac{3}{7}y - \dfrac{6}{7}z, -\dfrac{3}{7}x - \dfrac{6}{7}y - \dfrac{2}{7}z\right)^T$.

习题 1.72 (第六章第 7 题)

在三维几何空间的直角坐标系中，求关于直线 $z = 2y = 3x$ 的对称变换。

解 记所求变换为 T，所给直线为 l，任取 $a = (x, y, z)^T \in \mathbb{R}^3$，设其关于 l 的对称点为 $T(a) = (u, v, w)^T \triangleq b$，则 a, b 中点落在 l 上，且 a, b 之间的向量与 l 垂直，l 的方向向量为 $\overrightarrow{u} = (\frac{1}{3}, \frac{1}{2}, 1)^T$，从而有：

$$3(x+u) = 2(y+v) = z + w,\ \frac{x-u}{3} + \frac{y-v}{2} + z - w = 0$$

解得 $T(x,y,z)^T=(u,v,w)^T=\left(-\frac{41}{49}x+\frac{12}{49}y+\frac{24}{49}z,\frac{12}{49}x-\frac{31}{49}y+\frac{36}{49}z,\frac{24}{49}x+\frac{36}{49}y+\frac{23}{49}z\right)^T$.

习题 1.73 (第六章第 8 题)

在三维几何空间的直角坐标系中，求绕向量 $\boldsymbol{e}=(1,-1,1)^T$ 逆时针旋转 $30°$ 角的变换。

解 记所求变换为 T，任取 $A=(a,b,c)^T\in\mathbb{R}^3$，设 $\boldsymbol{e}$ 所在直线为 l，则过点 A 且与 l 垂直的平面 S 的方程为：

$$(x-a)-(y-b)+(z-c)=0$$

设 l 与 S 交于点 $O'=(\lambda,-\lambda,\lambda)^T$，代入 S 方程得到 $\lambda=\frac{a-b+c}{3}$，所求的 $T(A)$ 即为在平面 S 上，将点 A 绕 O' 逆时针旋转（从 $\boldsymbol{e}$ 的正方向看）$30°$ 角得到的点，记为 B，接下来我们以 O' 为原点取一组正交基，不妨取 $\overrightarrow{u_1}=\overrightarrow{O'A},\overrightarrow{u_3}=\overrightarrow{O'O},\overrightarrow{u_2}=\overrightarrow{u_3}\times\overrightarrow{u_1}$（向量外积），这样得到的 $(\overrightarrow{u_1},\overrightarrow{u_2},\overrightarrow{u_3})$ 构成右手系。再对 $\overrightarrow{u_2}$ 做伸缩使其模长与 $\overrightarrow{u_1}$ 相同，即 $\overrightarrow{u_2}'=\frac{\overrightarrow{u_2}}{|\overrightarrow{u_2}|}|\overrightarrow{u_1}|$，那么 $\overrightarrow{O'B}=\frac{\sqrt{3}}{2}\overrightarrow{u_1}\pm\frac{1}{2}\overrightarrow{u_2}'$ ($\overrightarrow{OA}\cdot\boldsymbol{e}<0$时为$+$,$>0$时为$-$)，上述过程中各个向量均是可求出的，计算得到：

$$\overrightarrow{u_1}=(a-\lambda,b+\lambda,c-\lambda)^T,\overrightarrow{u_2}=\lambda(b+c,c-a,-a-b)^T,\overrightarrow{u_3}=(-\lambda,\lambda,-\lambda)^T$$

而显然 T 是一个线性变换，那么我们只需要计算标准正交基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 的像即可，利用上述过程计算得到：

$$T(\boldsymbol{e}_1)=\left(\frac{\sqrt{3}+1}{3},\frac{\sqrt{3}-1}{3},\frac{1}{3}\right)^T,T(\boldsymbol{e}_2)=\left(-\frac{1}{3},\frac{\sqrt{3}+1}{3},\frac{\sqrt{3}-1}{3}\right)^T,T(\boldsymbol{e}_3)=\left(\frac{-\sqrt{3}+1}{3},-\frac{1}{3},\frac{\sqrt{3}+1}{3}\right)^T$$

据此我们便可得到所有点在该旋转变换下的像点，即 $T(x,y,z)^T=xT(\boldsymbol{e}_1)+yT(\boldsymbol{e}_2)+zT(\boldsymbol{e}_3)$.

笔记 $\overrightarrow{O'B}$ 表达式中正负号的不同是由于在 $\boldsymbol{e}$ 的正方向和负方向看，逆时针是不同的。其实也可以用 $\boldsymbol{e}$ 的方向作为 u_3 轴的方向，就可以避免讨论正负。

习题 1.74 (第六章第 36 题)

求下列线性变换在所指定的基下的矩阵

(1) 在 $F_n[x]$ 中，$\mathscr{A}(P(x))=P'(x)$，在基 $e_0=1,e_1=x,\cdots,e_{n-1}=\frac{x^{n-1}}{(n-1)!}$ 下；

(2) 以四个线性无关的函数

$$\alpha_1=e^{ax}\cos bx\quad\alpha_2=e^{ax}\sin bx\quad\alpha_3=x\,e^{ax}\cos bx\quad\alpha_4=x\,e^{ax}\sin bx$$

为基的四维空间中，线性变换为微分变换；

(3) 给定 2 阶实方阵 A, 求 2 阶实方阵构成的线性空间上的线性变换 $\mathscr{A}(\boldsymbol{x})=A\boldsymbol{x}-\boldsymbol{x}A$ 在基

$$e_1=\begin{pmatrix}1&0\\0&0\end{pmatrix},e_2=\begin{pmatrix}0&1\\0&0\end{pmatrix},e_3=\begin{pmatrix}0&0\\1&0\end{pmatrix},e_4=\begin{pmatrix}0&0\\0&1\end{pmatrix}$$ 下的矩阵。

解 (1) 由题 $\mathscr{A}(e_0,e_1,\cdots,e_{n-1})=(0,e_0,e_1,\cdots,e_{n-2})=(e_0,e_1,\cdots,e_{n-1})\begin{pmatrix}0&1&0&\cdots&0\\&0&1&\ddots&\vdots\\&&\ddots&\ddots&0\\&&&0&1\\&&&&0\end{pmatrix}$.

(2) 由题可知 $\mathscr{A}$ 与 (1) 中相同，那么有：

$$\mathscr{A}(\alpha_1)=(e^{ax}\cos bx)'=e^{ax}(a\cos bx-b\sin bx)=a\alpha_1-b\alpha_2,$$

$$\mathscr{A}(\alpha_2)=(e^{ax}\sin bx)'=e^{ax}(a\sin bx+b\cos bx)=b\alpha_1+a\alpha_2,$$

$$\mathscr{A}(\alpha_3)=(xe^{ax}\cos bx)'=e^{ax}(\cos bx+ax\cos bx-bx\sin bx)=\alpha_1+a\alpha_3-b\alpha_4,$$

$$\mathscr{A}(\alpha_4)=(xe^{ax}\sin bx)'=e^{ax}(\sin bx+ax\sin bx-bx\cos bx)=\alpha_2+b\alpha_3+a\alpha_4.$$

$$\mathscr{A}(\alpha_1,\alpha_2,\alpha_3,\alpha_4)=(\alpha_1,\alpha_2,\alpha_3,\alpha_4)\begin{pmatrix}a&b&1&0\\-b&a&0&1\\0&0&a&b\\0&0&-b&a\end{pmatrix}.$$

(3) 设 $A=\begin{pmatrix}a&b\\c&d\end{pmatrix}$，则 $\mathscr{A}e_1=\begin{pmatrix}0&-b\\c&0\end{pmatrix}=-be_2+ce_3,\mathscr{A}e_2=\begin{pmatrix}-c&a-d\\0&c\end{pmatrix}=-ce_1+(a-d)e_2+ce_4,\mathscr{A}e_3=\begin{pmatrix}b&0\\d-a&-b\end{pmatrix}=be_1+(d-a)e_3-be_4,\mathscr{A}e_4=\begin{pmatrix}0&b\\-c&0\end{pmatrix}=be_2-ce_3$，于是：

$$\mathscr{A}(e_1,e_2,e_3,e_4)=(e_1,e_2,e_3,e_4)\begin{pmatrix}0&-c&b&0\\-b&a-d&0&b\\c&0&d-a&-c\\0&c&-b&0\end{pmatrix}.$$

习题 1.75 (第六章第 37 题)

在 $\mathbb{R}^3$ 中定义线性变换

$$\mathscr{A}(x,y,z)^T=(x+2y,x-3z,2y-z)^T.$$

求 $\mathscr{A}$ 在基 $\boldsymbol{e}_1=(1,0,0)^T,\boldsymbol{e}_2=(0,1,0)^T,\boldsymbol{e}_3=(0,0,1)^T$ 下的矩阵。

解 $\mathscr{A}\boldsymbol{e}_1=(1,1,0)=\boldsymbol{e}_1+\boldsymbol{e}_2,\mathscr{A}\boldsymbol{e}_2=(2,0,2)=2\boldsymbol{e}_1+2\boldsymbol{e}_3,\mathscr{A}\boldsymbol{e}_3=(0,-3,-1)=-3\boldsymbol{e}_2-\boldsymbol{e}_3$，则有：

$$\mathscr{A}(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)=(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)\begin{pmatrix}1&2&0\\1&0&-3\\0&2&-1\end{pmatrix}\triangleq(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)A$$

则 A 即为 $\mathscr{A}$ 在基 $\boldsymbol{e}_1=(1,0,0)^T,\boldsymbol{e}_2=(0,1,0)^T,\boldsymbol{e}_3=(0,0,1)^T$ 下的矩阵。

习题 1.76 (第六章第 38 题)

设 $\mathbb{R}^3$ 中的线性变换 $\mathscr{A}$ 将

$$\alpha_1=(0,0,1)^T,\quad \alpha_2=(0,1,1)^T,\quad \alpha_3=(1,1,1)^T$$

变换到

$$\beta_1=(2,3,5)^T,\quad \beta_2=(1,0,0)^T,\quad \beta_3=(0,1,-1)^T.$$

求 $\mathscr{A}$ 在自然基和 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵。

解 容易看出 $\boldsymbol{e}_1=\alpha_3-\alpha_2,\boldsymbol{e}_2=\alpha_2-\alpha_1,\boldsymbol{e}_3=\alpha_3$，则 $\mathscr{A}\boldsymbol{e}_1=\beta_3-\beta_2=-\boldsymbol{e}_1+\boldsymbol{e}_2-\boldsymbol{e}_3,\mathscr{A}\boldsymbol{e}_2=\beta_2-\beta_1=-\boldsymbol{e}_1-3\boldsymbol{e}_2-5\boldsymbol{e}_3,\mathscr{A}\boldsymbol{e}_3=\beta_1=2\boldsymbol{e}_1+3\boldsymbol{e}_2+5\boldsymbol{e}_3$，则有：

$$\mathscr{A}(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)=(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)\begin{pmatrix}-1&-1&2\\1&-3&3\\-1&-5&5\end{pmatrix}\triangleq(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)A$$

则 A 即为 $\mathscr{A}$ 在自然基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 下的矩阵。

而 $\mathscr{A}\alpha_1=\beta_1=2\alpha_1+\alpha_2+2\alpha_3,\mathscr{A}\alpha_2=\beta_2=-\alpha_2+\alpha_3,\mathscr{A}\alpha_3=\beta_3=-2\alpha_1+\alpha_2$，则有：

$$\mathscr{A}(\alpha_1,\alpha_2,\alpha_3)=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}2&0&-2\\1&-1&1\\2&1&0\end{pmatrix}\triangleq(\alpha_1,\alpha_2,\alpha_3)B$$

则 B 即为 $\mathscr{A}$ 在 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵。

习题 1.77 (第六章第 41 题)

在 $\mathbb{R}^3$ 中给定两组基：

$$\alpha_1=(1,0,1)^T,\alpha_2=(2,1,0)^T,\alpha_3=(1,1,1)^T;\quad \beta_1=(2,3,1)^T,\quad \beta_2=(7,9,5)^T,\quad \beta_3=(3,4,3)^T.$$

定义线性变换 $\mathscr{A}(\alpha_i)=\beta_i\,(i=1,2,3)$.

(1) 求 $\mathscr{A}$ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵；

(2) 求 $\mathscr{A}$ 在基 β_1,β_2,β_3 下的矩阵。

解 (1) 计算得 $\mathscr{A}\alpha_1=\beta_1=-\dfrac{3}{2}\alpha_1+\dfrac{1}{2}\alpha_2+\dfrac{5}{2}\alpha_3,\mathscr{A}\alpha_2=\beta_2=-3\alpha_1+\alpha_2+8\alpha_3,\mathscr{A}\alpha_3=\beta_3=-\alpha_1+4\alpha_3$，则有：

$$\mathscr{A}(\alpha_1,\alpha_2,\alpha_3)=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}-\frac{3}{2}&-3&-1\\\frac{1}{2}&1&0\\\frac{5}{2}&8&4\end{pmatrix}\triangleq(\alpha_1,\alpha_2,\alpha_3)A$$

则 A 即为 $\mathscr{A}$ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵。

(2) 由前一问 $\mathscr{A}(\beta_1,\beta_2,\beta_3)=\mathscr{A}(\mathscr{A}(\alpha_1,\alpha_2,\alpha_3))=\mathscr{A}^2(\alpha_1,\alpha_2,\alpha_3)=(\alpha_1,\alpha_2,\alpha_3)A^2=(\beta_1,\beta_2,\beta_3)A.$

从而 $\mathscr{A}$ 在基 β_1,β_2,β_3 下的矩阵也为 A.

习题 1.78 (第六章第 42 题)

设 V 为 n 维线性空间，$\mathscr{A}:V\to V$ 为线性变换. 若存在 $\alpha\in V$, 使得 $\mathscr{A}^{n-1}\alpha\neq 0$, 但是 $\mathscr{A}^n\alpha=0$, 证明：$\mathscr{A}$ 在某组基下的矩阵为

$$\begin{pmatrix}0&1&0&\cdots&0\\&0&1&\ddots&\vdots\\&&\ddots&\ddots&0\\&&&0&1\\&&&&0\end{pmatrix}$$

证明 先证 $\alpha,\mathscr{A}\alpha,\cdots,\mathscr{A}^{n-1}\alpha$ 为 V 的一组基，只需证它们线性无关即可。设 $\lambda_1\alpha+\lambda_2\mathscr{A}\alpha+\cdots+\lambda_n\mathscr{A}^{n-1}\alpha=0$，由题 $k\geqslant n$ 时，$\mathscr{A}^k\alpha=0$，则等式两边同时作用 $\mathscr{A}^{n-1}$ 可得 $\lambda_1\mathscr{A}^{n-1}\alpha=0$，由于 $\mathscr{A}^{n-1}\alpha\neq 0$，则 $\lambda_1=0$，同理将等式两边同时作用 $\mathscr{A}^{n-m}$ 可得 $\lambda_m=0,m=1,\cdots,n$，从而 $\alpha,\mathscr{A}\alpha,\cdots,\mathscr{A}^{n-1}\alpha$ 为 V 的一组基，于是：

$$\mathscr{A}(\mathscr{A}^{n-1}\alpha,\cdots,\mathscr{A}\alpha,\alpha)=(0,\mathscr{A}^{n-1}\alpha,\cdots,\mathscr{A}\alpha)=(\mathscr{A}^{n-1}\alpha,\cdots,\mathscr{A}\alpha,\alpha)\begin{pmatrix}0&1&0&\cdots&0\\&0&1&\ddots&\vdots\\&&\ddots&\ddots&0\\&&&0&1\\&&&&0\end{pmatrix}$$

故题中所给矩阵为 $\mathscr{A}$ 在基 $(\mathscr{A}^{n-1}\alpha,\cdots,\mathscr{A}\alpha,\alpha)$ 下的矩阵。

笔记 这是个很常见的取法，以后也可能会多次见到，建议记住并掌握证明方法。

习题 1.79 (第六章第 43 题)

设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为线性空间 V 的一组基，证明：对于任意 $\beta_1,\beta_2,\cdots,\beta_n\in V$, 存在线性变换 $\mathscr{A}$, 使得 $\mathscr{A}(\alpha_i)=\beta_i\,(i=1,2,\cdots,n)$.

证明 由于 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 V 的一组基，那么 $\forall\beta_k,k=1,\cdots,n$，存在唯一的一组常数 $\lambda_{1k},\cdots,\lambda_{nk}$ 使得 $\beta_k=\lambda_{1k}\alpha_1+\cdots+\lambda_{nk}\alpha_n$，取方阵 A 满足 $(A)_{ij}=\lambda_{ij}$，由于线性变换与矩阵之间存在着一一对应，那么必存在一个线性变换 $\mathscr{A}$，其在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵为 A，于是有：

$$\mathscr{A}(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A=(\beta_1,\beta_2,\cdots,\beta_n)$$

从而 $\mathscr{A}(\alpha_i)=\beta_i,i=1,\cdots,n$，于是 $\mathscr{A}$ 即为所求的线性变换。

笔记 或者也可以直接定义映射 $\mathscr{A}:\lambda_1\alpha_1+\cdots+\lambda_n\alpha_n\mapsto\lambda_1\beta_1+\cdots+\lambda_n\beta_n$，容易验证线性性。这是因为 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 V 的一组基，所以 V 中任意元素都能写成它们的线性组合，我们把组合系数“迁移”到 $\beta_1,\beta_2,\cdots,\beta_n$ 上，便得到了要求的线性变换。并且我们其实可以把这个结论推广到线性映射，即 $\beta_1,\beta_2,\cdots,\beta_n\in U$ 为另一个线性空间，以同样的方式定义映射，验证线性性即可。

1.11 第十三周作业

习题 1.80 (第六章第 11 题)

给定 $\mathbb{R}^3$ 中线性变换 $\mathscr{A}(x,y,z)^T=(2x+y-z,x+2y+z,4x+5y+z)^T$. 求 $\mathscr{A}$ 的像空间与核空间的维数及一组基

解 由于 $\mathscr{A}e_1=(2,1,4)^T,\mathscr{A}e_2=(1,2,5)^T,\mathscr{A}e_3=(-1,1,1)^T$, 因此

$$\mathscr{A}(e_1,e_2,e_3)=(e_1,e_2,e_3)\begin{pmatrix}2&1&-1\\1&2&1\\4&5&1\end{pmatrix}.$$

作初等列变换 $\begin{pmatrix}2&1&-1\\1&2&1\\4&5&1\end{pmatrix}\to\begin{pmatrix}0&0&-1\\0&1&1\\0&2&1\end{pmatrix}$. 故 $\mathrm{im}\mathscr{A}$ 的基为 $e_2+2e_3,-e_1+e_2+e_3$, 维数为 2.

注意到

$$\mathscr{A}(x_1,x_2,x_3)^T=\mathscr{A}(x_1e_1+x_2e_2+x_3e_3)=\mathscr{A}(e_1,e_2,e_3)\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=(e_1,e_2,e_3)\begin{pmatrix}2&1&-1\\1&2&1\\4&5&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}.$$

由于 e_1,e_2,e_3 作为基是线性无关的，因此 $\mathscr{A}(x_1,x_2,x_3)^T=0$ 当且仅当 $\begin{pmatrix}2&1&-1\\1&2&1\\4&5&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=0$. 可解得 $(x_1,x_2,x_3)=(-1,1,-1)$, 故 $\ker\mathscr{A}$ 的基为 $-e_1+e_2-e_3$, 维数为 1.

习题 1.81 (第六章第 12 题)

证明：从高维数组空间到低维数组空间的投影变换不是可逆变换.

证明 若可逆，这意味着高维数组空间的维数与低维数组空间维数相同，但这显然不可能!

习题 1.82 (补充题)

设 $\mathbb{B}_U=(\alpha_1,\cdots,\alpha_n),\mathbb{B}_V=(\beta_1,\cdots,\beta_m)$ 分别为线性空间 U,V 的基.$\mathbb{B}_{U^*}=(\alpha^1,\cdots,\alpha^n),\mathbb{B}_{V^*}=(\beta^1,\cdots,\beta^m)$ 分别为 $\mathbb{B}_U$ 与 $\mathbb{B}_V$ 的对偶基. 设线性映射 $\mathscr{A}:U\to V$ 在基 $\mathbb{B}_U$ 与 $\mathbb{B}_V$ 下的矩阵为 A, 求 $\mathscr{A}^*:V^*\to U^*$ 在基 $\mathbb{B}_{U^*}$ 与 $\mathbb{B}_{V^*}$ 下的矩阵.

解 设 $\mathscr{A}^*(\beta^1,\cdots,\beta^m)=(\alpha^1,\cdots,\alpha^n)B$, 其中 $B=(b_{ij})$. 故 $\mathscr{A}^*\beta^i=\sum\limits_{j=1}^n b_{ji}\alpha^j$，两边同时作用 α_j，有 $b_{ji}=(\mathscr{A}\beta^i)(\alpha_j)=\beta^i\mathscr{A}(\alpha_j)=\beta^i(\sum\limits_{k=1}^m a_{kj}\beta_j)=a_{ij}$. 故 $B=A^T$.

习题 1.83 (第六章第 21 题)

判断以下矩阵 A,B 是否相似? 并说明理由.

$$(1)A=\begin{pmatrix}1&2&3\\0&1&0\\0&0&1\end{pmatrix},B=\begin{pmatrix}1&2&0\\0&1&3\\0&0&1\end{pmatrix};$$

$$(1)A=\begin{pmatrix}1&2&0&0\\0&1&3&0\\0&0&1&0\\0&0&0&1\end{pmatrix},B=\begin{pmatrix}1&2&0&0\\0&1&0&0\\0&0&1&3\\0&0&0&1\end{pmatrix}.$$

解 (1) 注意到 $r(I-A)\neq r(I-B)$, 因此 A,B 不相似.
(2) 注意到 $r(I-A)^2\neq r(I-B)^2$, 因此 A,B 不相似.

注 后面我们将会证明：两个矩阵是否相似，可以完全由一些矩阵的秩判断.

习题 1.84 (第六章第 22 题)

求三阶可逆方阵 P 使得 $P\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}P^{-1}=\begin{pmatrix}3&0&0\\0&1&0\\0&0&2\end{pmatrix}$.

解 设 $P=(\alpha_1,\alpha_2,\alpha_3)$, 记 $A=\begin{pmatrix}3&0&0\\0&1&0\\0&0&2\end{pmatrix}$. 注意到 $P\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}P^{-1}=\begin{pmatrix}3&0&0\\0&1&0\\0&0&2\end{pmatrix}$ 等价于 $A(\alpha_1,\alpha_2,\alpha_3)=(\alpha_1,2\alpha_2,3\alpha_3)$. 因此问题转化为求矩阵 A 的特征向量.$A\alpha_1=\alpha_1$ 等价于 $(A-I)\alpha_1=0$, 解得 $\alpha_1=(0,1,0)^T(\alpha_1$ 取值不唯一!), 同理 $\alpha_2=(0,0,1)^T,\alpha_3=(1,0,0)^T$. 因此可以取 $P=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$.

注 事实上，以上求解过渡矩阵的过程给出了求解可对角化矩阵到对角矩阵的过渡矩阵的一般方法 (即求特征向量). 具体操作如下：
假设 $P=(\alpha_1,\alpha_2,\cdots,\alpha_n)$, 且 $P^{-1}AP=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$. 于是

$$A(\alpha_1,\alpha_2,\cdots,\alpha_n)=AP=P\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)=(\lambda_1\alpha_1,\lambda_2\alpha_2,\cdots,\lambda_n\alpha_n)$$

因此求 P 即求 α_i, 即求矩阵 A 关于特征值 λ_i 的特征向量，也即求解线性方程组 $(\lambda_i I-A)x=0$ 的解.

习题 1.85 (第六章第 24 题)

设 A 与 B 相似，C 与 D 相似. 证明：$\mathrm{diag}(A,C)$ 与 $\mathrm{diag}(B,D)$ 相似. 反之是否成立?

证明 由于 A 与 B 相似，C 与 D 相似，因此存在可逆矩阵 P,Q，使得 $PAP^{-1}=B,QCQ^{-1}=D$. 故

$$\mathrm{diag}(P,Q)\mathrm{diag}(A,C)\mathrm{diag}(P^{-1},Q^{-1})=\mathrm{diag}(B,D).$$

反之不成立，考虑 $\mathrm{diag}(1,0)$ 与 $\mathrm{diag}(0,1)$.

1.12 第十四周作业

习题 1.86 (第六章第 13 题 (2)(4)(6))

求下列矩阵的全部特征值和特征向量

$$(2)\ \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix} \qquad (4)\ \begin{pmatrix} 3 & -2 & 1 \\ 2 & -2 & 2 \\ 3 & -6 & 5 \end{pmatrix} \qquad (6)\ \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{pmatrix}$$

解 (2) 矩阵的特征方程写为

$$det(\lambda I - A) = (\lambda - cos\theta)^2 + sin^2\theta = 0 \tag{1.71}$$

解得 $\lambda = con\theta \pm isin\theta$.

$sin\theta = 0$ 时矩阵为 $\pm I$, 对应的特征值分别为 ± 1，对应的特征向量为所有非零向量；

$sin\theta \neq 0$ 时：

对于 $\lambda_1 = cos\theta + isin\theta$, 设特征向量为 $(x_1, x_2)^T$, 有

$$\begin{cases} isin\theta x_1 + sin\theta x_2 = 0 \\ -sin\theta x_1 + isin\theta x_2 = 0 \end{cases} \tag{1.72}$$

解得 $x_2 = -ix_1$. 即 $\lambda_1 = cos\theta + isin\theta$ 对应的特征向量为 $c_1(1, -i)^T, c_1 \neq 0$;

对于 $\lambda_2 = cos\theta - isin\theta$, 同理解得 $x_2 = ix_1$, 对应的特征向量为 $c_2(1, i)^T, c_2 \neq 0$.

(4) 矩阵的特征方程写为

$$det(\lambda I - A) = (\lambda - 2)^3 = 0 \tag{1.73}$$

解得 $\lambda = 2$, 为三重根.

设特征向量为 $(x_1, x_2, x_3)^T$, 有 $A(x_1, x_2, x_3)^T = 2(x_1, x_2, x_3)^T$，即

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_1 - 4x_2 + 2x_3 = 0 \\ 3x_1 - 6x_2 + 3x_3 = 0 \end{cases} \tag{1.74}$$

解得特征向量 $c_1(2, 1, 0)^T + c_2(-1, 0, 1)^T$.

(6) 矩阵的特征方程写为

$$det(\lambda I - A) = (\lambda - 2)^2(\lambda^2 - 2) = 0 \tag{1.75}$$

解得 $\lambda = 2, \pm\sqrt{2}$，其中 2 为二重根.

对于 $\lambda_1 = \lambda_2 = 2$, 列方程解得对应的特征向量为 $c_1(1, 0, 1, 0)^T + c_2(1, 0, 0, 1)^T, c_1, c_2 \neq 0$;

对于 $\lambda_3 = \sqrt{2}$, 特征向量为 $c_3(\sqrt{2} - 1, -\sqrt{2} + 1, -\sqrt{2} + 1, 1)^T, c_3 \neq 0$;

对于 $\lambda_4 = -\sqrt{2}$, 特征向量为 $c_4(-\sqrt{2} - 1, \sqrt{2} + 1, \sqrt{2} + 1, 1)^T, c_4 \neq 0$.

习题 1.87 (第六章第 14 题)

给定 $\mathbb{R}^2$ 上的线性变换 $\mathscr{A}(x, y)^T = (3x + 2y, 2x + y)^T$.$\mathscr{A}$ 将单位圆 C 映射为椭圆 C'. 求椭圆 C' 的长短半轴方向与长度, 以及椭圆的面积.

解 易知线性变换对应的矩阵为 $A = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix}$，求得其特征值为 $\lambda_1 = 2 + \sqrt{5}, \lambda_2 = 2 - \sqrt{5}$, 对应的特征向量分别

为 $\alpha = (1, \frac{\sqrt{5}-1}{2})^T, \beta = (1, -\frac{\sqrt{5}+1}{2})^T$. 可以看到 $\alpha \perp \beta$, 则变换相当于沿这两个方向进行伸缩: 在 α 方向拉伸为原来的 $2+\sqrt{5}$ 倍, 即长半轴方向沿 α, 长半轴长度 $a = 2+\sqrt{5}$; 在 β 方向反向压缩为原来的 $\sqrt{5}-2$ 倍, 即短半轴方向沿 β, 短半轴长度 $b = \sqrt{5}-2$. 椭圆面积为 $S = \pi ab = \pi$.

下面直接给出这个线性变换的图示, 以便于更好地理解线性变换与特征值, 特征向量的联系:

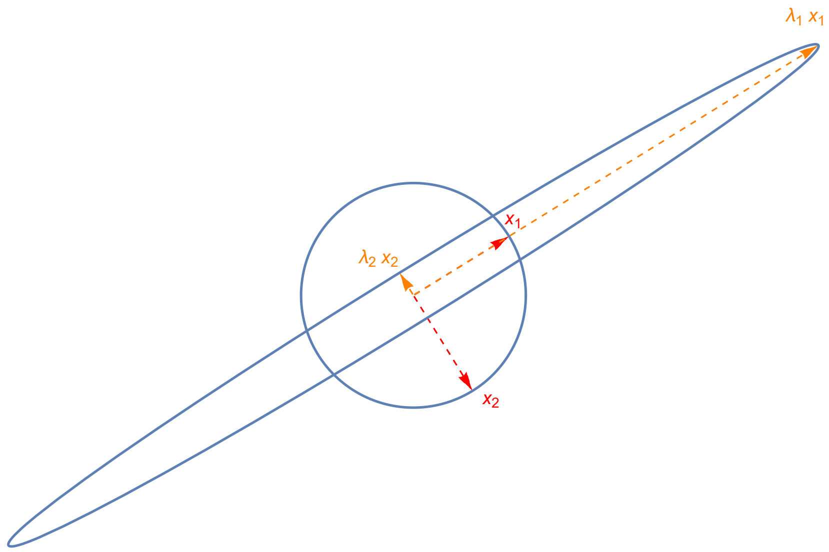


图 1.1: 一般的二维伸缩变换：由两个垂直的特征向量及其特征值可完全确定变换得到的椭圆

习题 1.88 (第六章第 15 题)

设 A 是可逆方阵. 证明: 若 λ 是 A 的一个特征值, 则 λ^{-1} 是 A^{-1} 的特征值, 且对应的特征向量相同.

证明 由于 A 可逆, 故 $detA = \prod_i \lambda_i \neq 0$, 即 A 的特征值均不为零.

设 α 为 A 关于特征值 λ 的特征向量, 则有

$$A\alpha = \lambda\alpha \Rightarrow \lambda^{-1}A\alpha = \alpha \tag{1.76}$$

两边同时左乘 A^{-1}，得到

$$\lambda^{-1}\alpha = A^{-1}\alpha \tag{1.77}$$

则命题成立.

习题 1.89 (第六章第 17 题)

设三阶方阵 A 的特征多项式为 $p_A(\lambda) = (\lambda-1)^2(\lambda-2)$. 求方阵 $3A+2I$ 的特征值与行列式.

解 由特征多项式得 A 的三个特征值:$\lambda_1 = \lambda_2 = 1, \lambda_3 = 2$, 由于单位阵 I 的特征向量为所有非零向量, 故 $3A+2I$ 的特征向量与 A 的特征向量相同, 对应地, 特征值变为 $\lambda_1' = \lambda_2' = 3\times 1+2 = 5, \lambda_3' = 3\times 2+2 = 8$. 行列式为 $det(3A+2I) = \lambda_1'\lambda_2'\lambda_3' = 200$.

注 事实上, 同学们可以自行证明以下性质:

若 α 为可逆矩阵 A 属于 λ 的特征向量, 则 α 也是 $kA, A^2, aA+bI, A^m, f(A), A^{-1}, A^*$ 分别属于 $k\lambda, \lambda^2, a\lambda+b, \lambda^m, f(\lambda), \lambda^{-1}, \frac{|A|}{\lambda}$ 的特征向量.

习题 1.90 (第六章第 18 题)

(1) 若 $A^2=I$, 证明:A 的特征值只能是 ±1;
(2) 设 n 阶实方阵 A 满足 $A^T=-A$, 证明 A 的特征值为零或纯虚数.

证明 (1) 设 A 的特征值为 λ, 其对应的特征向量为 α, 则有

$$A^2\alpha=A\lambda\alpha=\lambda^2\alpha \tag{1.78}$$

而由 $A^2=I$ 得

$$A^2\alpha=I\alpha=\alpha \tag{1.79}$$

故 $\lambda^2=1\Rightarrow\lambda=\pm1$.

(2) 设 A 的特征值为 λ, 其对应的特征向量为 α. 对特征向量做共轭转置得到 $\bar{\alpha}^T$, 将其作用在 $A\alpha=\lambda\alpha$ 两端, 得到

$$\bar{\alpha}^TA\alpha=\lambda\bar{\alpha}^T\alpha=\lambda|\alpha|^2 \tag{1.80}$$

将上式两边同时求共轭转置, 由于 A 为实方阵, 其共轭转置就是转置:

$$\bar{\alpha}^TA^T\alpha=\bar{\lambda}|\alpha|^2\Rightarrow-\bar{\alpha}^TA\alpha=\bar{\lambda}|\alpha|^2 \tag{1.81}$$

将两式相加得

$$(\lambda+\bar{\lambda})|\alpha|^2=0 \tag{1.82}$$

由于 α 为特征向量, 故 $|\alpha|^2\neq0$, 故 $\lambda+\bar{\lambda}=0$, 命题成立.

习题 1.91 (第六章第 26 题)

设矩阵 $A=\begin{pmatrix}0&0&1\\x&1&y\\1&0&0\end{pmatrix}$ 有三个线性无关的特征向量, 求 x 和 y 应满足的条件.

解 A 的特征方程写为

$$det(\lambda I-A)=(\lambda+1)(\lambda-1)^2=0 \tag{1.83}$$

即 A 的三个特征值为 $\lambda_1=-1,\lambda_2=\lambda_3=1$, 其中 1 为二重根, 则命题等价于求特征值 1 对应的特征子空间维数为 2 的条件, 即令矩阵 $(I-A)$ 的秩为 1, 下面对其进行线性变换:

$$I-A=\begin{pmatrix}1&0&-1\\-x&0&-y\\-1&0&1\end{pmatrix}\to\begin{pmatrix}1&0&-1\\-x&0&-y\\0&0&0\end{pmatrix}\to\begin{pmatrix}1&-1&0\\-x&-y&0\\0&0&0\end{pmatrix} \tag{1.84}$$

易见令其秩为 1 等价于 $\begin{vmatrix}1&-1\\-x&-y\end{vmatrix}=0\Rightarrow x+y=0$.

习题 1.92 (第六章第 27 题)

设矩阵 A 和 B 相似, 其中

$$A=\begin{pmatrix}-2&0&0\\2&x&2\\3&1&1\end{pmatrix},\quad B=\begin{pmatrix}-1&0&0\\0&2&0\\0&0&y\end{pmatrix}.$$

(1) 求 x 和 y 的值;

(2) 求可逆矩阵 T, 使得 $T^{-1}AT = B$.

解 (1) $A \overset{s}{\sim} B \Rightarrow f_A(\lambda) = f_B(\lambda)$. 即

$$\lambda^3 + (1-x)\lambda^2 + (-x-4)\lambda + 2x - 4 = \lambda^3 + (-1-y)\lambda^2 + (y-2)\lambda + 2y \tag{1.85}$$

解得 $x = 0, y = -2$.

(2) 三个特征值为 $-1, 2, -2$, 分别求得对应的三个特征向量 $(0,2,-1)^T, (0,1,1)^T, (1,0,-1)^T$, 则 T 可取为 (结果不唯一): $\begin{pmatrix} 0 & 0 & 1 \\ 2 & 1 & 0 \\ -1 & 1 & -1 \end{pmatrix}$

习题 1.93 (第二章第 28 题)

设 n 阶方阵 $A \neq 0$, 满足 $A^m = 0$, 其中 $m \geq 2$ 为正整数.
(1) 求 A 的特征值;
(2) 证明:A 不能相似于对角阵;
(3) 证明:$|I + A| = 1$.

解 (1) 设 A 特征值为 λ, 其对应的特征向量为 α, 则有

$$A^m\alpha = \lambda A^{m-1}\alpha = \cdots = \lambda^m\alpha \tag{1.86}$$

而由题有 $A^m\alpha = 0\,\alpha = 0$, 故 $\lambda = 0$.

(2) 用反证法, 若 A 可相似于对角阵, 则此对角阵为零矩阵, 则 A 为零矩阵, 矛盾.

(3) 存在一个 n 阶可逆方阵 T 使得 $T^{-1}AT$ 为上三角阵, 且其对角元素为 A 的特征值, 即均为 0. 故 $I + T^{-1}AT$ 为对角线元素均为 1 的上三角阵, 即 $|I + T^{-1}AT| = 1$, 则

$$|I + A| = |T(I + T^{-1}AT)T^{-1}| = |T||I + T^{-1}AT||T^{-1}| = 1 \tag{1.87}$$

另解: A 的特征值均为 0, 其为特征多项式的 n 重根, 则特征多项式为 $|\lambda I - A| = \lambda^n$. 取 $\lambda = -1$ 得 $(-1)^n|I + A| = (-1)^n$, 即 $|I + A| = 1$. 由此看到, 利用特征多项式可以简化一些相关行列式的计算.

习题 1.94 (第六章第 29 题 (2)(3))

判断下列矩阵 A 是否可以对角化? 若可以, 试求变换矩阵 T, 使得 $T^{-1}AT$ 为对角阵.

$$(2)\ A = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 4 & -2 \\ -3 & -3 & 5 \end{pmatrix}; \qquad (3)\ A = \begin{pmatrix} 3 & 0 & 1 \\ 4 & -2 & -8 \\ -4 & 0 & -1 \end{pmatrix}.$$

解 (2) 特征多项式为

$$|\lambda I - A| = (\lambda - 2)^2(\lambda - 6) \tag{1.88}$$

即特征值为 $\lambda_1 = \lambda_2 = 2, \lambda_3 = 6$.

对于二重特征值 2, 有 $rank(2I - A) = 1$, 则其特征子空间维数为 2, 即 A 可对角化, 两个基 (解方程组得到) 可写为 $(1,0,1)^T$ 和 $(1,-1,0)^T$.

又特征值 6 的一个特征向量写为 $(1,-2,3)^T$, 故变换矩阵 T(不唯一) 可写为 $\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 1 & 0 & 3 \end{pmatrix}$.

(3) 特征多项式为

$$|\lambda I - A| = (\lambda - 1)^2(\lambda + 2) \tag{1.89}$$

即特征值为 $\lambda_1 = \lambda_2 = 1, \lambda_3 = -2$.

对于二重特征值 1, 有 $rank(I - A) = 2$, 则其特征子空间维数为 1, 即 A 不能对角化.

习题 1.95 (第六章第 30 题)

设 A, B 均为 n 阶方阵,A 有 n 个互异的特征值且 $AB = BA$. 证明 B 相似于对角阵.

证明 (这玩意实际上就是期中考试前那天晚上群里有人问到的往年题.)

由于 A 有 n 个互异特征值, 故 A 可对角化, 有 $T^{-1}AT = diag(\lambda_1, \cdots, \lambda_n) = C$. 记 $T^{-1}BT = D$, 我们下面证明 D 为对角阵.

$AB = BA \Rightarrow T^{-1}ABT = T^{-1}BAT \Rightarrow T^{-1}ATT^{-1}BT = T^{-1}BTT^{-1}AT \Rightarrow CD = DC.$

即 D 与对角阵 C 可交换, 考虑其非对角元素 $d_{ij}, i \neq j$, 有 $\lambda_i d_{ij} = \lambda_j d_{ij}$, 由于 $\lambda_i \neq \lambda_j$, 故 $d_{ij} = 0$, 所以 D 为对角阵, 命题成立.

1.13 第十五周作业

习题 1.96 (第六章第 47 题)

设 $A=\begin{pmatrix}2&-2&3\\10&-4&5\\5&-4&6\end{pmatrix}$，求方阵 T，使 $T^{-1}AT$ 为 Jordan 标准形。

解 A 的特征多项式 $\varphi_A(\lambda)=(\lambda-2)(\lambda-1)^2$，则 A 的特征值为 $2,1,1$.
对于 $\lambda=1$，有 $rank(A-I)=rank(J-I)=2$，故 $J=diag(2,J_2(1))$，设 $T=(T_1,T_2,T_3)$ 满足 $AT=TJ$.
则 $AT_1=2T_1,AT_2=T_2,AT_3=T_2+T_3$，则 $(A-2I)T_1=0,(A-I)T_2=0,(A-I)T_3=T_2,(A-I)^2T_3=0$.
计算得 $T_1=(4t,15t,10t)^T,T_2=(s,5s,3s)^T$，由 $(A-I)^2T_3=0$ 得 $T_3=(2p-q,q,p)^T$，代入 $(A-I)T_3=T_2$
可得满足条件的方阵 T 的通解为 $T=\begin{pmatrix}4t&s&2p-q\\15t&5s&q\\10t&3s&p\end{pmatrix}$，其中 $5p-3q=s$.

习题 1.97 (第六章第 48 题)

设方阵 A 满足 $A^2=A$，用矩阵的 Jordan 标准形证明：$tr(A)=rank(A)$.

证明 任取 A 的一个特征值 λ，以及一个对应的特征向量 $\boldsymbol{x}$，则有 $A\boldsymbol{x}=\lambda\boldsymbol{x}=A^2\boldsymbol{x}=\lambda^2\boldsymbol{x}$，而 $\boldsymbol{x}\neq\boldsymbol{0}$，故 $\lambda=0$ 或 1，即 A 的特征值只能为 0 或 1，则可以设 A 的 Jordan 标准形为 $J=diag(J_{m_1}(1),\cdots,J_{m_p}(1),J_{k_1}(0),\cdots,J_{k_q}(0))$，并且 $P^{-1}AP=J$，其中 P 是可逆方阵，由 $A^2=A$，则 $tr(A)=tr(J)=p,rank(A)=rank(A^2)=rank(J^2)$，且 $J^2=P^{-1}A^2P=P^{-1}AP=J$，而 $J^2=diag(J_{m_1}^2(1),\cdots,J_{m_p}^2(1),J_{k_1}^2(0),\cdots,J_{k_q}^2(0))=J$，那么只能 $k_1=\cdots=k_q=1$，即特征值 0 对应的 Jordan 块都是 1 阶的，也就是 0，那么 $rank(A)=rank(J)=p$.

习题 1.98 (第七章第 1 题)

已知 $\alpha_1=(1,2,-1,1),\alpha_2=(2,3,1,-1),\alpha_3=(-1,-1,-2,2)$.
(1) 求 $\alpha_1,\alpha_2,\alpha_3$ 的长度及彼此之间的夹角。
(2) 求与 $\alpha_1,\alpha_2,\alpha_3$ 都正交的向量。

解 (1) 计算得 $|\alpha_1|=\sqrt{7},|\alpha_2|=\sqrt{15},|\alpha_3|=\sqrt{10}$，以及 $\alpha_1\cdot\alpha_2=6,\alpha_1\cdot\alpha_3=1,\alpha_2\cdot\alpha_3=-9$. 从而有：

$$\theta_{12}=\arccos\left(\frac{6}{\sqrt{105}}\right),\ \theta_{13}=\arccos\left(\frac{1}{\sqrt{70}}\right),\ \theta_{23}=\arccos\left(-\frac{9}{5\sqrt{6}}\right).$$

(2) 这等价于求解方程组：$\begin{cases}x_1+2x_2-x_3+x_4=0\\2x_1+3x_2+x_3-x_4=0\\-x_1-x_2-2x_3+2x_4=0\end{cases}$，通解为：$(-5,3,1,0)^Tt+(5,-3,0,1)^Ts$.

习题 1.99 (第七章第 3 题)

设 $\boldsymbol{x},\boldsymbol{y}$ 是欧氏空间 $\mathbb{R}^n$ 的两个向量，它们之间的夹角为 θ，证明：
(1)（余弦定理）$|\boldsymbol{x}-\boldsymbol{y}|^2=|\boldsymbol{x}|^2+|\boldsymbol{y}|^2-2|\boldsymbol{x}||\boldsymbol{y}|\cos\theta$.
(2)（平行四边形定理）$|\boldsymbol{x}+\boldsymbol{y}|^2+|\boldsymbol{x}-\boldsymbol{y}|^2=2(|\boldsymbol{x}|^2+|\boldsymbol{y}|^2)$

(3)（菱形对角线定理）若 $|\boldsymbol{x}|=|\boldsymbol{y}|$，则 $(\boldsymbol{x}+\boldsymbol{y})\perp(\boldsymbol{x}-\boldsymbol{y})$.

证明 (1) $RHS=|\boldsymbol{x}|^2+|\boldsymbol{y}|^2-2\boldsymbol{x}\cdot\boldsymbol{y}=\boldsymbol{x}\cdot(\boldsymbol{x}-\boldsymbol{y})+(\boldsymbol{y}-\boldsymbol{x})\cdot\boldsymbol{y}=|\boldsymbol{x}-\boldsymbol{y}|^2=LHS$.

(2) $LHS=|\boldsymbol{x}|^2+|\boldsymbol{y}|^2+2\boldsymbol{x}\cdot\boldsymbol{y}+|\boldsymbol{x}|^2+|\boldsymbol{y}|^2-2\boldsymbol{x}\cdot\boldsymbol{y}=2(|\boldsymbol{x}|^2+|\boldsymbol{y}|^2)=RHS$.

(3) $(\boldsymbol{x}+\boldsymbol{y})\cdot(\boldsymbol{x}-\boldsymbol{y})=|\boldsymbol{x}|^2-|\boldsymbol{y}|^2=0$，故 $(\boldsymbol{x}+\boldsymbol{y})\perp(\boldsymbol{x}-\boldsymbol{y})$.

习题 1.100 (第七章第 5 题)

用 Schmidt 正交化方法构造标准正交向量组：

(1) $(0,0,1),(0,1,1),(1,1,1)$;

(2) $(1,1,1,2),(1,1,-5,3),(3,2,8,-7)$.

解 以下均设三个向量为 $\alpha_1,\alpha_2,\alpha_3$.

(1) $|\alpha_1|=1$，则取 $e_1=\alpha_1$，令 $\beta_2=\alpha_2-(\alpha_2,e_1)e_1=(0,1,0),|\beta_2|=1$，则取 $e_2=\beta_2$.

令 $\beta_3=\alpha_3-(\alpha_3,e_1)e_1-(\alpha_3,e_2)e_2=(1,0,0),|\beta_3|=1$，则取 $e_3=\beta_3$.

(2) $|\alpha_1|=\sqrt{7}$，则取 $e_1=\frac{1}{\sqrt{7}}\alpha_1=\frac{1}{\sqrt{7}}(1,1,1,2)$，令 $\beta_2=\alpha_2-(\alpha_2,e_1)e_1=(\frac{4}{7},\frac{4}{7},-\frac{38}{7},\frac{15}{7}),|\beta_2|=\frac{9\sqrt{21}}{7}$，则取 $e_2=\frac{1}{9\sqrt{21}}(4,4,-38,15)$，令 $\beta_3=\alpha_3-(\alpha_3,e_1)e_1-(\alpha_3,e_2)e_2=\alpha_3+\frac{1}{7}\alpha_1+\frac{389\times 7}{1701}\beta_2$,，将 β_3 单位化，得到 $e_3=\frac{1}{\sqrt{13449}}(\frac{493\sqrt{2}}{9},\frac{743\sqrt{2}}{18},-\frac{133\sqrt{2}}{18},-\frac{133\sqrt{2}}{3})$.

习题 1.101 (补充题)

设 $(V,(\cdot,\cdot))$ 为内积空间，$\mathbb{B}_1=(\alpha_1,\cdots,\alpha_n),\mathbb{B}_2=(\beta_1,\cdots,\beta_n)$ 为两组基，$(\beta_1,\cdots,\beta_n)=(\alpha_1,\cdots,\alpha_n)P$，设 $(\cdot,\cdot)$ 在 $\mathbb{B}_1,\mathbb{B}_2$ 下的度量矩阵分别为 G_1,G_2，证明：$G_2=P^TG_1P$.

证明 设 $P=(p_{ij})_{n\times n}$，则 $\beta_k=p_{1k}\alpha_1+\cdots+p_{nk}\alpha_n,k=1,\cdots,n$，记 $(G_1)_{ij}=(\alpha_i,\alpha_j)=g_{ij}$，则 $(G_2)_{ij}=(\beta_i,\beta_j)=(p_{1i}\alpha_1+\cdots+p_{ni}\alpha_n,p_{1j}\alpha_1+\cdots+p_{nj}\alpha_n)=\sum_{m=1}^{n}\sum_{k=1}^{n}p_{mi}g_{mk}p_{kj}=(P^TG_1P)_{ij}$，故 $G_2=P^TG_1P$.

习题 1.102 (第七章第 6 题)

设在 $\mathbb{R}^3$ 中，基 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 的度量矩阵是

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

试求 $\mathbb{R}^3$ 中由 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 表示的一组标准正交基。

解 可以看出 $|\boldsymbol{a}_1|=1,|\boldsymbol{a}_2|=\sqrt{2}$，且 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 正交，故取 $e_1=\boldsymbol{a}_1,e_2=\frac{1}{\sqrt{2}}\boldsymbol{a}_2$，令 $\beta_3=\boldsymbol{a}_3-(\boldsymbol{a}_3,e_1)e_1-(\boldsymbol{a}_3,e_2)e_2=\boldsymbol{a}_3+\boldsymbol{a}_1$，则 $|\beta_3|^2=|\boldsymbol{a}_3|^2+|\boldsymbol{a}_1|^2+2\boldsymbol{a}_1\cdot\boldsymbol{a}_3=1$，则可取 $e_3=\beta_3=\boldsymbol{a}_1+\boldsymbol{a}_3$. e_1,e_2,e_3 即为一组标准正交基。

习题 1.103 (第七章第 10 题)

设 $\boldsymbol{e}_1,\boldsymbol{e}_2.\boldsymbol{e}_3$ 是 $\mathbb{R}^3$ 的一组标准正交基，令 $\boldsymbol{a}_1=\frac{1}{3}(2\boldsymbol{e}_1+2\boldsymbol{e}_2-\boldsymbol{e}_3),\boldsymbol{a}_2=\frac{1}{3}(2\boldsymbol{e}_1-\boldsymbol{e}_2+2\boldsymbol{e}_3)$,

$\boldsymbol{a}_3=\dfrac{1}{3}(\boldsymbol{e}_1-2\boldsymbol{e}_2-2\boldsymbol{e}_3)$ 证明：$\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 也是 $\mathbb{R}^3$ 的一组标准正交基。

证明 由 $\boldsymbol{e}_1,\boldsymbol{e}_2.\boldsymbol{e}_3$ 是标准正交基，则 $|\boldsymbol{a}_1|^2=\dfrac{1}{9}(4|\boldsymbol{e}_1|^2+4|\boldsymbol{e}_2|^2+|\boldsymbol{e}_3|^2)=1$，故 $|\boldsymbol{a}_1|=1$，类似有 $|\boldsymbol{a}_2|=|\boldsymbol{a}_3|=1$. 并且 $\boldsymbol{a}_1\cdot\boldsymbol{a}_2=\dfrac{4}{9}-\dfrac{2}{9}-\dfrac{2}{9}=0$，类似有 $\boldsymbol{a}_1\cdot\boldsymbol{a}_3=\boldsymbol{a}_2\cdot\boldsymbol{a}_3=0$，故 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 也是 $\mathbb{R}^3$ 的一组标准正交基。

习题 1.104 (第七章第 12 题)

设 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ 是 $\mathbb{R}^n$ 的标准正交基，证明：

(1) 对任意 $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^n,(\boldsymbol{a},\boldsymbol{b})=\sum\limits_{i=1}^{n}(\boldsymbol{a},\boldsymbol{a}_i)(\boldsymbol{b},\boldsymbol{a}_i)$.

(2) 对任意 $\boldsymbol{a}\in\mathbb{R}^n,|\boldsymbol{a}|^2=\sum\limits_{i=1}^{n}(\boldsymbol{a},\boldsymbol{a}_i)^2$.

证明 (1) 由于 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ 是标准正交基，可设 $\boldsymbol{a}=\lambda_1\boldsymbol{a}_1+\cdots+\lambda_n\boldsymbol{a}_n,\boldsymbol{b}=\mu_1\boldsymbol{a}_1+\cdots+\mu_n\boldsymbol{a}_n$，则 $(\boldsymbol{a},\boldsymbol{a}_i)=\lambda_i,(\boldsymbol{b},\boldsymbol{a}_i)=\mu_i$，则 $(\boldsymbol{a},\boldsymbol{b})=(\lambda_1\boldsymbol{a}_1+\cdots+\lambda_n\boldsymbol{a}_n,\mu_1\boldsymbol{a}_1+\cdots+\mu_n\boldsymbol{a}_n)=\sum\limits_{i=1}^{n}\lambda_i\mu_i=\sum\limits_{i=1}^{n}(\boldsymbol{a},\boldsymbol{a}_i)(\boldsymbol{b},\boldsymbol{a}_i)$.

(2) 在 (1) 中令 $\boldsymbol{b}=\boldsymbol{a}$ 即得。

习题 1.105 (第七章第 14 题)

写出所有 3 阶正交矩阵，它的元素是 0 或 1.

解 由于正交矩阵每行每列均为单位向量，则每行每列均有且只有一个 1，故满足要求的正交矩阵只能是标准单位向量 $e_1.e_2,e_3$ 的排列，共以下 6 个：

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

习题 1.106 (第七章第 16 题)

若 $\boldsymbol{a}$ 是 $\mathbb{R}^n$ 单位向量，证明：$Q=I_n-2\boldsymbol{a}\boldsymbol{a}^T$ 是一个正交阵。当 $\boldsymbol{a}=\frac{1}{\sqrt{3}}(1,1,1)^T$ 时，具体求出 Q.

证明 由于 $\boldsymbol{a}$ 为单位向量，故 $\boldsymbol{a}^T\boldsymbol{a}=1$，且容易看出 $Q=Q^T$，故：

$$QQ^T=Q^TQ=(I_n-2\boldsymbol{a}\boldsymbol{a}^T)(I_n-2\boldsymbol{a}\boldsymbol{a}^T)^T=I_n-4\boldsymbol{a}\boldsymbol{a}^T+4\boldsymbol{a}\boldsymbol{a}^T\boldsymbol{a}\boldsymbol{a}^T=I_n-4\boldsymbol{a}\boldsymbol{a}^T+4\boldsymbol{a}\boldsymbol{a}^T=I_n.$$

从而 Q 是一个正交矩阵。

解 代入计算得到 $Q=I_3-\dfrac{2}{3}\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}=\begin{pmatrix}\frac{1}{3}&-\frac{2}{3}&-\frac{2}{3}\\-\frac{2}{3}&\frac{1}{3}&-\frac{2}{3}\\-\frac{2}{3}&-\frac{2}{3}&\frac{1}{3}\end{pmatrix}$.

1.14 第十六周作业

习题 1.107 (第七章第 20 题)

设 W 是由 $(1,1,0),(1,-2,1)$ 生成的 $\mathbb{R}^3$ 的子空间.
(1) 求 W 的一组标准正交基.
(2) 求 $W^\perp$ 的标准正交基.
(3) 求向量 $(0,0,2)$ 在 W 的正交投影.

解 (1) 作史密斯正交化：$\beta_1=\alpha_1=(1,1,0)$,$\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}=(\frac{3}{2},-\frac{3}{2},1)$. 最后我们做一次单位化，有 $e_1=\frac{\beta_1}{(\beta_1,\beta_1)}=(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},0),e_2=\frac{\beta_2}{(\beta_2,\beta_2)}=\frac{1}{\sqrt{22}}(3,-3,1)$.

(2) 设 $\alpha=(a_1,a_2,a_3)\in W^\perp$, 故 $(\alpha,\alpha_1)=0,(\alpha,alpha_2)=0$, 因此 $a_1+a_2=0,a_1-2a_2+a_3=0$, 解得 $\alpha=(t,-t,-3t)$. 又 $\mathbb{R}^3=W\oplus W^\perp$, 因此 $dimW^\perp=1$, 故 $W=span\{e_3=(1,-1,-3)\}$.

(3) 记 $x=(0,0,1)$,$Px=(x,e_1)e_1+(x,e_2)e_2=(6,-6,2)$.

习题 1.108 (第七章第 24 题)

设 $\boldsymbol{a}$ 为 $\mathbb{R}^n$ 中的单位向量. 定义线性变换：$\mathscr{A}\boldsymbol{b}=\boldsymbol{b}-2(\boldsymbol{a},\boldsymbol{b})\boldsymbol{a}$. 证明：$\mathscr{A}$ 是正交变换，并找 $\mathbb{R}^n$ 的一组标准正交基，使得 $\mathscr{A}$ 在该组基下的矩阵为对角阵.

证明 $\mathscr{A}\boldsymbol{b},\mathscr{A}\boldsymbol{c}=(\boldsymbol{b}-2(\boldsymbol{a},\boldsymbol{c})\boldsymbol{a},\boldsymbol{c}-2(\boldsymbol{a},\boldsymbol{c})\boldsymbol{a})=(\boldsymbol{b},\boldsymbol{c})$, 因此 $\mathscr{A}$ 是正交变换. 将 $\boldsymbol{a}$ 扩充成 $\mathbb{R}^n$ 中的一组基，在对其进行史密斯正交化，最后在做单位化，$\mathscr{A}$ 在这组基下的矩阵为对角矩阵.

习题 1.109 (第七章第 26 题)

设

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

求正交矩阵 P, 使 $P^{-1}AP$ 为对角阵. 由此计算 A^k，k 是自然数.

解 $\det\{(\lambda I-A)\}=(\lambda-5)(\lambda-2)(\lambda+1)$. 容易算得属于特征值为 5，2，-1 的特征向量分别为：

$$(\frac{1}{3},-\frac{2}{3},\frac{2}{3})^T,(\frac{2}{3},-\frac{1}{3},-\frac{2}{3})^T,(\frac{2}{3},\frac{2}{3},\frac{1}{3})^T$$

于是 $\begin{pmatrix}\frac{2}{3}&\frac{2}{3}&\frac{1}{3}\\\frac{2}{3}&-\frac{1}{3}&-\frac{2}{3}\\\frac{1}{3}&-\frac{2}{3}&\frac{2}{3}\end{pmatrix}$, 故 $P^{-1}AP=diag(-1,2,5)$. 计算得出：

$$A^k=\begin{pmatrix}4(-1)^k+2^{k+2}+5^k&4(-1)^k-2^{k+1}-2\cdot5^k&2(-1)^k-2^{k+2}+2\cdot5^k\\4(-1)^k-2^{k+2}-2\cdot5^k&4(-1)^k+a^{k+2}+4\cdot5^k&2(-1)^k+2^{k+1}-4\cdot5^k\\2(-1)^k-2^{k+2}+2\cdot5^k&2(-1)^k+2^{k+1}-4\cdot5^k&(-1)^k+2^{k+2}+4\cdot5^k\end{pmatrix}$$

.

习题 1.110 (第七章第 27 题)

证明：下列三个条件中只要有两个成立，另一个也必成立.
（1）A 是对称的；（2）A 是正交的；（3）$A^2=I$.

证明 (1)+(2)⇒(3)：$A^2 = AA^T = I$.
(1)+(3)⇒(2):$A^2 = AA^T = I$.
(2)+(3)⇒(1): 由 A 正交，$A^{-1} = A^T$, 又 $A^2 = I$, 且逆矩阵是唯一的，知 $A = A^T$.

习题 1.111 (第七章第 29 题)

设 A 是 n 阶实对称方阵，且 $A^2 = A$. 证明：存在正交方阵 T 使得 $T^{-1}AT = diag(I_r, 0)$, 这里 $r = rank(A)$.

证明 由 A 是对称实矩阵，知存在正交矩阵 P, 使得 $P^{-1}AP = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$, 其中 λ_i 是矩阵 A 的特征值，由矩阵 A 适合多项式 $x^2 - x$, 因此矩阵 A 的特征值为 0，1. 又因为交换矩阵行列对应的矩阵是正交矩阵，于是存在正交方阵 T 使得 $T^{-1}AT = diag(I_r, 0)$.

习题 1.112 (第七章第 30 题)

设 A 是 n 阶实对称方阵. 证明：$max_{0\neq x\in\mathbb{R}^n} \frac{x^TAx}{x^Tx} = \lambda_{max}$，这里 λ_{max} 是 A 的最大特征值.

证明 存在正交矩阵 P, 使得 $P^TAP = diag(\lambda_1, \cdots, \lambda_n)$. 设 $x = Py$，则 $max\frac{x^TAx}{x^Tx} = max\frac{y^TP^TAPy}{y^Ty} = \lambda_{max}$.

习题 1.113 (第七章第 31 题)

设 V 是 n 阶实方阵全体构成的线性空间. 对 $A, B \in V$, 定义 $(A, B) = tr(AB^T)$.
(1) 证明：(A, B) 构成内积，从而 V 在该内积下成为欧氏空间.
(2) 设 $A = (a_{ij})_{n\times n} \in V$，求 A 的模长.
(3) 基本矩阵 E_{ij} 是否构成 V 的一组标准正交基? 请说明理由.

证明 (1) 由矩阵乘法及迹的定义立明.
解 (2)$||A|| = \sqrt{tr(AA^T)} = \sqrt{\sum\limits_{i,j=1}^{n} a_{ij}^2}$.
(3) 由基础矩阵乘法及迹的定义立明.

习题 1.114 (第七章第 32 题)

设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 n 维欧氏空间 V 的一组向量. 定义其 $Gram$ 矩阵 $G = ((\alpha_i, \alpha_j)) \in \mathbb{R}^{n\times n}$.
(1) 证明：$\alpha_1, \alpha_2, \cdots, \alpha_n$ 构成 V 的一组基当且仅当 $\det(G) \neq 0$.
(2) 设 α_i 在一组标准正交基下的坐标为 $\boldsymbol{x_i}, i = 1, 2, \cdots, n$. 记 $\boldsymbol{X} = (\boldsymbol{x_1}, \cdots, \boldsymbol{x_n}) \in \mathbb{R}^{n\times n}$. 则 $\det(G) = \det(X)^2$.

证明 (1) $\alpha_1, \alpha_2, \cdots, \alpha_n$ 构成基等价于 $k_1\alpha_1 + \cdots + k_n\alpha_n = 0$ 只有零解等价于 $(k_1\alpha_1 + \cdots + k_n\alpha_n, \alpha_i) = 0, \forall i$ 等价于 G 可逆 (看成线性方程组，系数矩阵可逆).

(2) 由于 $(\alpha_i, \alpha_j) = \alpha_i^T\alpha_j = \boldsymbol{x}_i^T\boldsymbol{x}_j$, 因此 $G = X^TX$, 故 $\det(G) = \det(X)^2$.

1.15 第十七周作业

习题 1.115 (第八章第 1 题 (2)(4))

将下列二次型表示成矩阵形式

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

(4) $Q(x_1,x_2,\cdots,x_n)=\sum\limits_{i=1}^{n-2}(x_i-x_{i+2})^2$

解 (2) $Q(x_1,x_2,x_3)=x^T\begin{pmatrix}2&1&0\\1&0&-\frac{1}{2}\\0&-\frac{1}{2}&0\end{pmatrix}x$

(4) $n=3$ 时 $Q=x^T\begin{pmatrix}1&0&-1\\0&0&0\\-1&0&1\end{pmatrix}x$，$n\geq 4$ 时 $Q=x^T\begin{pmatrix}1&0&-1&&&&\\0&1&0&\ddots&&&\\-1&0&2&\ddots&\ddots&&\\&\ddots&\ddots&\ddots&\ddots&\ddots&\\&&\ddots&\ddots&2&0&-1\\&&&\ddots&0&1&0\\&&&&-1&0&1\end{pmatrix}x$

习题 1.116 (第八章第 2 题 (1)(2))

写出下列对称矩阵对应的二次型

(1) $\begin{pmatrix}1&3&5\\3&2&6\\5&6&4\end{pmatrix}$ (2) $\begin{pmatrix}a&b&0\\b&a&b\\0&b&a\end{pmatrix}$

解 (1) $Q=x_1^2+2x_2^2+4x_3^2+6x_1x_2+10x_1x_3+12x_2x_3$ (2) $Q=ax_1^2+ax_2^2+ax_3^2+2bx_1x_2+2bx_2x_3$

习题 1.117 (第八章第 3 题 (3)(4))

用配方法将下列二次型化为标准型，并求相应的可逆线性变换

(3) $Q=x_1x_2+x_2x_3+x_3x_4+x_4x_1$

(4) $Q=x_1^2+5x_2^2-4x_3^2+2x_1x_2-4x_2x_3$

解 (3) 令 $\begin{cases}x_1=y_1+y_2\\x_2=y_1-y_2\\x_3=y_3+y_4\\x_4=y_3-y_4\end{cases}$，得 $Q=(y_1+y_3)^2-(y_2+y_4)^2$

再令 $\begin{cases}y_1=z_1-z_3\\y_2=z_2-z_4\\y_3=z_3\\y_4=z_4\end{cases}$，得 $Q=z_1^2-z_2^2$. 相应的变换为 $x=Pz, P=\begin{pmatrix}1&1&-1&-1\\1&-1&-1&1\\&&1&1\\&&1&-1\end{pmatrix}$

(4) 令 $\begin{cases} x_1 = y_1 - y_2 \\ x_2 = y_2 \\ x_3 = y_3 \end{cases}$，得 $Q = y_1^2 + 4y_2^2 - 4y_3^2 - 4y_2y_3$；再令 $\begin{cases} y_1 = z_1 \\ y_2 = z_2 + \frac{z_3}{2} \\ y_3 = z_3 \end{cases}$，

得 $Q = z_1^2 + 4z_2^2 - 5z_3^2$. 相应的变换为 $x = Pz, P = \begin{pmatrix} 1 & -1 & -\frac{1}{2} \\ & 1 & \frac{1}{2} \\ & & 1 \end{pmatrix}$

习题 1.118 (第八章第 4 题 (4))

用初等变换法将下列二次型化为标准型，并求相应的可逆线性变换

(4) $Q = x^T \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix} x$

解 本题第一行和第三行完全相同，第二行和第四行完全相同，可先利用这个特点消去 3、4 行及 3、4 列，再对前两列进行对角化，有：

$$\begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \to \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 1 & & -1 & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \to \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 1 & & -1 & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \to \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & & -1 & \\ & 1 & & -1 \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & & -1 & \\ 1 & 1 & & -1 \\ & & 1 & \\ & & & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -\frac{1}{2} & -1 & \\ 1 & \frac{1}{2} & & -1 \\ & & 1 & \\ & & & 1 \end{pmatrix} \tag{1.90}$$

得标准型 $Q = y^T \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} y$, $x = Py$, $P = \begin{pmatrix} 1 & -\frac{1}{2} & -1 & \\ 1 & \frac{1}{2} & & -1 \\ & & 1 & \\ & & & 1 \end{pmatrix}$

习题 1.119 (第八章第 5 题 (1)(3))

求正交变换化下列实二次型为标准型

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

(3) $Q = 6x_1x_2 + 6x_1x_3 + 6x_2x_3$

解 (1) 二次型对应的矩阵为 $A=\begin{pmatrix}1&-2&2\\-2&4&-4\\2&-4&4\end{pmatrix}$, A 的特征值为 $0,0,9$, 分别对应特征向量 $\alpha_1=(2,1,0)^T,\alpha_2=(-2,0,1)^T,\alpha_1=(1,-2,2)^T$, 注意此时三个特征向量并不正交归一，需要进一步进行 Schmidt 正交化，得到一组标准正交向量 $e_1=(\frac{2}{\sqrt5},\frac{1}{\sqrt5},0)^T,e_2=(-\frac{2}{3\sqrt5},\frac{4}{3\sqrt5},\frac{\sqrt5}{3})^T,e_3=(\frac13,-\frac23,\frac23)^T$, 故得到标准型为 $\begin{pmatrix}0&&\\&0&\\&&9\end{pmatrix}$，相应的正交变换为 $P=\begin{pmatrix}\frac{2}{\sqrt5}&-\frac{2}{3\sqrt5}&\frac13\\\frac{1}{\sqrt5}&\frac{4}{3\sqrt5}&-\frac23\\0&\frac{\sqrt5}{3}&\frac23\end{pmatrix}$

(3) 矩阵对应的特征值为 $-3,-3,6$，一组标准正交向量为 $e_1=(\frac{1}{\sqrt2},-\frac{1}{\sqrt2},0)^T,e_2=(\frac{1}{\sqrt6},\frac{1}{\sqrt6},-\frac{2}{\sqrt6})^T$, $e_3=(\frac{1}{\sqrt3},\frac{1}{\sqrt3},\frac{1}{\sqrt3})^T$，故标准型为 $\begin{pmatrix}-3&&\\&-3&\\&&6\end{pmatrix}$，相应的正交变换为 $P=\begin{pmatrix}\frac{1}{\sqrt2}&\frac{1}{\sqrt6}&\frac{1}{\sqrt3}\\-\frac{1}{\sqrt2}&\frac{1}{\sqrt6}&\frac{1}{\sqrt3}\\0&-\frac{2}{\sqrt6}&\frac{1}{\sqrt3}\end{pmatrix}$

习题 1.120 (第八章第 15 题)

判断下列矩阵是否是正定矩阵

(1) $\begin{pmatrix}2&\frac12&2\\\frac12&2&-2\\2&-2&4\end{pmatrix}$ (2) $\begin{pmatrix}1&-1&1\\-1&2&0\\1&0&1\end{pmatrix}$

解 (1) $\begin{vmatrix}2&\frac12&2\\\frac12&2&-2\\2&-2&4\end{vmatrix}=-5<0$，故矩阵不是正定矩阵.

(2) $\begin{vmatrix}1&-1&1\\-1&2&0\\1&0&1\end{vmatrix}=-1<0$，故矩阵不是正定矩阵.

习题 1.121 (第八章第 16 题)

参数 t 满足什么条件时，下列二次型正定？

(1) $Q(x_1,x_2,x_3)=2x_1^2+x_2^2+x_3^2+2x_1x_2+tx_1x_3$

(2) $Q(x_1,x_2,x_3)=x_1^2+2x_2^2+3x_3^2+tx_1x_2+tx_1x_3+x_2x_3$

解 (1) 二次型对应的矩阵为 $\begin{pmatrix}2&1&t/2\\1&1&0\\t/2&0&1\end{pmatrix}$, 易见 $2>0$, $\begin{vmatrix}2&1\\1&1\end{vmatrix}=1>0$, 则正定 $\iff \begin{vmatrix}2&1&t/2\\1&1&0\\t/2&0&1\end{vmatrix}=1-\frac{t^2}{4}>0$, 即 $-2<t<2$

(2) 二次型对应的矩阵为 $\begin{pmatrix}1&t/2&t/2\\t/2&2&1/2\\t/2&1/2&3\end{pmatrix}$，则正定 $\iff \begin{vmatrix}1&t/2\\t/2&2\end{vmatrix}=2-\frac{t^2}{4}>0$ 且 $\begin{vmatrix}1&t/2&t/2\\t/2&2&1/2\\t/2&1/2&3\end{vmatrix}=\frac{23}{4}-t^2>0$，解得 $-\frac{\sqrt{23}}{2}<t<\frac{\sqrt{23}}{2}$

习题 1.122 (第八章第 20 题)

设 n 元实二次型 $Q(x_1,x_2,\cdots,x_n)=(x_1+a_1x_2)^2+(x_2+a_2x_3)^2+\cdots+(x_{n-1}+a_{n-1}x_n)^2+(x_n+a_nx_1)^2$，其中 $a_i(i=1,\cdots,n)$ 为实数，试问：当 $a_1,\cdots,a_n$ 满足何种条件时，$Q(x_1,\cdots,x_n)$ 为正定二次型？

解 易知 $Q\geq 0$，则 Q 正定 $\iff$ 仅在 $(x_1,x_2,\cdots,x_n)=0$ 时取等号 $\iff$

方程组 $\begin{cases} x_1+a_1x_2=0 \\ x_2+a_2x_3=0 \\ \vdots \\ x_n+a_nx_1=0 \end{cases}$ 无非零解 $\iff$ 系数矩阵行列式 $\begin{vmatrix} 1 & a_1 & & & \\ & 1 & a_2 & & \\ & & \ddots & \ddots & \\ & & & 1 & a_{n-1} \\ a_n & & & & 1 \end{vmatrix}\neq 0$

$\iff 1+(-1)^{n+1}a_1\cdots a_n\neq 0 \iff a_1\cdots a_n\neq(-1)^n$

第 2 章 习题课讲义

2.1 第一次习题课

2.1.1 第一章复习与补充知识

向量的运算

¶ 向量的线性运算

向量的**线性运算**包括向量的加法和数乘运算，满足以下性质：

(1) 加法交换律：$\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$

(2) 加法结合律：$(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$

(3) 分配律 1：$\lambda(\boldsymbol{a}+\boldsymbol{b})=\lambda\boldsymbol{a}+\lambda\boldsymbol{b}$

(4) 分配律 2：$(\lambda+\mu)\boldsymbol{a}=\lambda\boldsymbol{a}+\mu\boldsymbol{a}$

(5) 零向量：$\boldsymbol{a}=\boldsymbol{a}+\boldsymbol{0}=\boldsymbol{0}+\boldsymbol{a}$

(6) 负向量：$\boldsymbol{a}+(-\boldsymbol{a})=\boldsymbol{0}$

(7) 乘法结合律：$\lambda(\mu\boldsymbol{a})=(\lambda\mu)\boldsymbol{a}$

(8) $1\boldsymbol{a}=\boldsymbol{a}$

事实上，不止对于数组空间，对于一般的集合，我们可以更广泛地定义加法和数乘，只要它们满足以上的性质，就将这样的集合 + 运算称为**线性空间**，而我们通过数组空间得到的性质可类推到一般的线性空间，详细内容会在第五章学习到。特别地，我们记 n 维数组空间为 $\mathbb{K}^n$，它包含了分量定义在数域 $\mathbb{F}$ 上的所有 n 维数组向量。有了向量的加法与数乘运算，我们可以便定义线性组合，并引出线性相关、线性无关的定义：

定义 2.1 (线性组合)

给定一个向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...,\boldsymbol{\alpha}_m\in K^n$，任给一组数 $k_1,k_2,...,k_m\in K$，可以得到一个向量 $k_1\boldsymbol{\alpha}_1+k_1\boldsymbol{\alpha}_1+...+k_m\boldsymbol{\alpha}_m$，称这个向量为给定向量组的一个线性组合，其中 $k_1,k_2,...,k_m$ 称为组合系数.

对于 $\boldsymbol{\beta}\in K^n$，如果存在 K 中一组数 $c_1,c_2,...,c_m$ 使得 $\boldsymbol{\beta}=c_1\boldsymbol{\alpha}_1+c_1\boldsymbol{\alpha}_1+...+c_m\boldsymbol{\alpha}_m$，则称 $\boldsymbol{\beta}$ 可以由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...,\boldsymbol{\alpha}_m$ 线性表示.

♣

从线性组合的定义出发，我们可以给出 n 元线性方程组的另一种表达形式：将方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+...+a_{1n}x_n=b_1\\ a_{21}x_1+a_{22}x_2+...+a_{2n}x_n=b_2\\ \quad\vdots\qquad\quad\vdots\qquad\qquad\quad\vdots\qquad\quad\vdots\\ a_{s1}x_1+a_{s2}x_2+...+a_{sn}x_n=b_s \end{cases} \tag{2.1}$$

关于未知数 $x_i(i=1,2,...,n)$ 前的系数组记为列向量 $\boldsymbol{\alpha}_i=(a_{1i},a_{2i},...,a_{si})^T$，记 $\boldsymbol{\beta}=(b_1,b_2,...,b_s)^T$，则方程组改写为

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+...+x_n\boldsymbol{\alpha}_n=\boldsymbol{\beta} \tag{2.2}$$

从而我们得到以下等价命题：

命题 2.1

1. 线性方程组有解 $\iff$ 常数项列向量 $\boldsymbol{\beta}$ 可由系数矩阵的列向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...,\boldsymbol{\alpha}_n$ 线性表示
2. 求线性方程组的解 $\iff$ 求满足线性组合 $\boldsymbol{\beta}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+...+x_n\boldsymbol{\alpha}_n$ 的组合系数

♠

这使得我们可以直接通过方程组的系数和常数项来研究方程组的解，后续我们会由此角度学习张成子空间、秩、解空间、基与维数等概念。

定义 2.2 (线性相关、线性无关)

K^n 中向量组 $\boldsymbol{\alpha}_1,...,\boldsymbol{\alpha}_s(s\geq 1)$ 称为是线性相关的，如果 K 中存在不全为零的数 $k_1,...,k_s$ 使得 $k_1\boldsymbol{\alpha}_1+...+k_s\boldsymbol{\alpha}_s=0$. 否则，称向量组 $\boldsymbol{\alpha}_1,...,\boldsymbol{\alpha}_s$ 线性无关

命题 2.2 (一些性质)

1. 包含零向量的向量组一定线性相关 $(1\cdot\mathbf{0}+0\boldsymbol{\alpha}_2+...+0\boldsymbol{\alpha}_s=\mathbf{0})$

2. 单个向量 $\boldsymbol{\alpha}$ 线性相关的充要条件为 $\boldsymbol{\alpha}=\mathbf{0}$

3. 多于一个向量的向量组线性相关 $\iff$ 其中至少有一个向量可以由其余向量线性表示[a]

证明 必要性：设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...,\boldsymbol{\alpha}_s(s\geq 2)$ 线性相关，则存在不全为零的一组数 $k_1,k_2,...,k_s$ 使得 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+...+k_s\boldsymbol{\alpha}_s=0$. 设 $k_i\neq 0$，则 $\boldsymbol{\alpha}_i=-\frac{k_1}{k_i}\boldsymbol{\alpha}_1-...-\frac{k_{i-1}}{k_i}\boldsymbol{\alpha}_{i-1}-\frac{k_{i+1}}{k_i}\boldsymbol{\alpha}_{i+1}-...-\frac{k_s}{k_i}\boldsymbol{\alpha}_s$.

充分性：设 $\boldsymbol{\alpha}_j$ 可由其他向量线性表示：$\boldsymbol{\alpha}_j=l_1\boldsymbol{\alpha}_1+...+l_{j-1}\boldsymbol{\alpha}_{j-1}+l_{j+1}\boldsymbol{\alpha}_{j+1}+...+l_s\boldsymbol{\alpha}_s$，则移项得 $l_1\boldsymbol{\alpha}_1+...+l_{j-1}\boldsymbol{\alpha}_{j-1}-\boldsymbol{\alpha}_j+l_{j+1}\boldsymbol{\alpha}_{j+1}+...+l_s\boldsymbol{\alpha}_s=\mathbf{0}$，从而向量组线性相关

4. 如果向量组中的一部分向量线性相关，那么整个向量组也线性相关；如果向量组线性无关，那么其中任选的部分向量也线性无关

5. 如果向量组线性无关，则给每个向量对应位置各添上 m 个分量得到的延伸组也线性无关；如果 n 维向量组线性相关，则每个向量对应位置各去掉 $m(m<n)$ 个分量得到的缩短组也线性相关

[a] 我们的教材采用这种方式定义线性相关

例题 2.1（丘维声《高等代数》上册 P87） 设向量组 $\boldsymbol{\alpha}_1,...,\boldsymbol{\alpha}_s$ 线性无关，则向量 $\boldsymbol{\beta}$ 可以由 $\boldsymbol{\alpha}_1,...,\boldsymbol{\alpha}_s$ 线性表示的充分必要条件为 $\boldsymbol{\alpha}_1,...,\boldsymbol{\alpha}_s,\boldsymbol{\beta}$ 线性相关。

证明 必要性由性质 3 是显然的。

充分性：设 $\boldsymbol{\alpha}_1,...,\boldsymbol{\alpha}_s,\boldsymbol{\beta}$ 线性相关，则存在 K 中不全为 0 的数 $k_1,...,k_s,l$ 使得

$$k_1\boldsymbol{\alpha}_1+...+k_s\boldsymbol{\alpha}_s+l\boldsymbol{\beta}=0 \tag{2.3}$$

若 $l=0$，则 $k_1,...,k_s$ 不全为 0，且由式 (2.3) 得 $k_1\boldsymbol{\alpha}_1+...+k_s\boldsymbol{\alpha}_s=0$，从而 $\boldsymbol{\alpha}_1,...,\boldsymbol{\alpha}_s$ 线性相关，与已知条件矛盾，因此 $l\neq 0$，从而由式 (2.3) 得

$$\boldsymbol{\beta}=-\frac{k_1}{l}\boldsymbol{\alpha}_1-...-\frac{k_s}{l}\boldsymbol{\alpha}_s \tag{2.4}$$

证毕。

推论 2.1

设向量组 $\boldsymbol{\alpha}_1,...,\boldsymbol{\alpha}_s$ 线性无关，则向量 $\boldsymbol{\beta}$ 不能由 $\boldsymbol{\alpha}_1,...,\boldsymbol{\alpha}_s$ 线性表示的充分必要条件为 $\boldsymbol{\alpha}_1,...,\boldsymbol{\alpha}_s,\boldsymbol{\beta}$ 线性无关。

例题 2.2（教材第一章习题 4） 证明：三维空间中四个或四个以上的向量一定线性相关

证明 只需证四个向量一定线性相关。

任取四个三维向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$，若其中可以挑出三个向量线性相关，则命题已证，否则，任意三个向量均线性无关，我们选择向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$，由于它们线性无关，则它们不共面，由教材定理 **1.2.1** 得 $\boldsymbol{\alpha}_4$ 可由这三个向量唯一线性表示，即这四个三维向量线性相关。

证毕。

笔记 教材的定理 **1.2.1** 给出了三维空间基的性质，具有推广意义。

定义 2.3

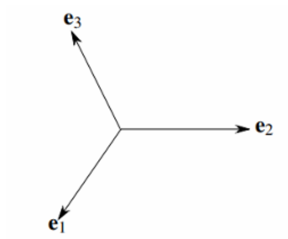
$V \subset K^n$ 为子空间.V 中一组向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_s$ 称为 V 的一组基，如果它满足：

1. 对任意向量 $\boldsymbol{\beta} \in V$，可以表示为 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_s$ 的线性组合 $\boldsymbol{\beta} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + ... + x_s\boldsymbol{\alpha}_s$.
2. $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_s$ 线性无关.

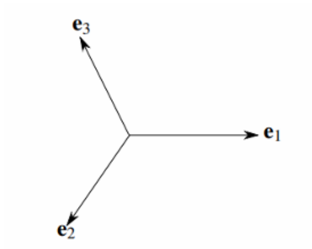
称 $(x_1, x_2, ..., x_s)$ 为向量 $\boldsymbol{\beta}$ 在基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_s$ 下的坐标.

特别地，当我们在三维空间选定了一组基：$\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$，并确定了坐标原点 $\mathbf{O}$，根据它们的空间排列方式可以定义右手系和左手系.

♣



(a) 右手系



(b) 左手系

¶ 三维向量的数量积、向量积

定义 2.4 (数量积)

两向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的数量积为一个实数，记为 $\boldsymbol{a} \cdot \boldsymbol{b}$. 设 $\boldsymbol{a}, \boldsymbol{b}$ 的夹角为 θ，则

$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}|cos\theta \tag{2.5}$$

在直角坐标系下 (基中的向量两两垂直)，两向量的数量积等于对应坐标分量相乘再求和：

$$\boldsymbol{a} \cdot \boldsymbol{b} = (a_1, a_2, a_3) \cdot (b_1, b_2, b_3) = a_1b_1 + a_2b_2 + a_3b_3 \tag{2.6}$$

数量积又称内积、点乘。

♣

内积满足以下性质 (取向量 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ 及实数 λ):

(1) **对称性**：$\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{b} \cdot \boldsymbol{a}$

(2) **线性性**：$(\boldsymbol{a} + \boldsymbol{b}) \cdot \boldsymbol{c} = \boldsymbol{a} \cdot \boldsymbol{c} + \boldsymbol{b} \cdot \boldsymbol{c}$

$(\lambda\boldsymbol{a}) \cdot \boldsymbol{b} = \lambda(\boldsymbol{a} \cdot \boldsymbol{b}) = \boldsymbol{a} \cdot (\lambda\boldsymbol{b})$

(3) **正定性**：$\boldsymbol{a}^2 = \boldsymbol{a} \cdot \boldsymbol{a} \geq 0$，等号在 $\boldsymbol{a} = \mathbf{0}$ 时成立

事实上，除向量的数量积外，只要满足上述三个性质的运算都可以称为内积运算，定义了内积运算的线性空间称为**欧几里得空间**，第七章内积空间将会对其详细研究。

定义 2.5 (外积、叉乘)

两个三维向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的向量积 $\boldsymbol{a} \times \boldsymbol{b}$ 为一个三维向量，它与 $\boldsymbol{a}, \boldsymbol{b}$ 垂直且按序 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a} \times \boldsymbol{b}$ 构成右手系，它的模为以 $\boldsymbol{a}, \boldsymbol{b}$ 为边的平行四边形面积：

$$|\boldsymbol{a} \times \boldsymbol{b}| = |\boldsymbol{a}||\boldsymbol{b}|sin\theta \tag{2.7}$$

在直角坐标系下，两向量的向量积可由坐标如下表示：

$$\boldsymbol{a} \times \boldsymbol{b} = (a_1, a_2, a_3) \times (b_1, b_2, b_3) = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1) \tag{2.8}$$

向量积又称外积、叉乘。

♣

向量积满足以下性质 (取向量 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ 及实数 λ):

(1) **反身性**：$\boldsymbol{a}\times\boldsymbol{b}=-\boldsymbol{b}\times\boldsymbol{a}$

(2) **线性性**：$(\boldsymbol{a}+\boldsymbol{b})\times\boldsymbol{c}=\boldsymbol{a}\times\boldsymbol{c}+\boldsymbol{b}\times\boldsymbol{c}$

$(\lambda\boldsymbol{a})\times\boldsymbol{b}=\lambda(\boldsymbol{a}\times\boldsymbol{b})=\boldsymbol{a}\times(\lambda\boldsymbol{b})$

三个向量的混合积 $(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}$ 和二重外积 $(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{c}$ 由数量积、向量积组合而成，当我们熟悉掌握了数量积和向量积的计算后，关于混合积和二重外积相关性质的推导是容易的。我们将部分性质列举如下：

(1) $(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=(\boldsymbol{b}\times\boldsymbol{c})\cdot\boldsymbol{a}=(\boldsymbol{c}\times\boldsymbol{a})\cdot\boldsymbol{b}$

(2) $(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=-(\boldsymbol{b}\times\boldsymbol{a})\cdot\boldsymbol{c}$

(3) $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 中任意两个平行时，混合积为 0

(4) 在直角坐标系下，有

$$(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=\begin{vmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{vmatrix}$$

(5)$(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{c}=(\boldsymbol{a}\cdot\boldsymbol{c})\boldsymbol{b}-(\boldsymbol{b}\cdot\boldsymbol{c})\boldsymbol{a}$

¶ 求和约定与张量引入（本小节参考潘海俊老师《电动力学》讲义）

上一个小节中的向量运算有的按照坐标写开会有很长一串 (如三阶行列式)，我们是否有办法简化写法？答案是肯定的。

定义 2.6 (Einstein 求和约定)

[a] 若同一指标在同一单项式中重复出现，则默认对其求和 (即可以省略求和符号)，这样的求和指标称为哑指标，其他指标称为自由指标。

例如，我们可以用 a_ib_i 表示向量点乘 $\boldsymbol{a}\cdot\boldsymbol{b}$，这是因为

$$a_ib_i=\sum_{i=1}^{3}a_ib_i=a_1b_1+a_2b_2+a_3b_3=\boldsymbol{a}\cdot\boldsymbol{b} \tag{2.9}$$

可见，i 在例中即为哑指标，我们通过这样的约定将点乘结果的三项求和形式上写为了一项。

[a] 爱因斯坦写论文时用到很多求和公式，他懒得写求和号，所以给出了这项约定

♣

关于指标的一些性质：

(1) 哑指标可替换为任意其他字母，例如：$a_ib_i=\sum_{i=1}^{3}a_ib_i=\sum_{j=1}^{3}a_jb_j=a_jb_j$

(2) 等式两边的自由指标可同时替换为不引起混淆的其他字母，例如：设 A、B 为同阶方阵，则对二者乘积 AB 有 $(AB)_{ij}=A_{ik}B_{kj}$，也可以替换自由指标，写为 $(AB)_{mn}=A_{mk}B_{kn}$

(3) 计算的最终结果不会含有哑指标，但会保留自由指标

定义 2.7 (Kronecker 符号)

$$\delta_{ij}=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases} \tag{2.10}$$

♣

定义 2.8 (Levi-Civita 符号)

$$\epsilon_{ijk}=\begin{cases}+1, & (ijk)=(123),(231)or(312)\\ -1, & (ijk)=(132),(213)or(321)\\ 0, & otherwise\end{cases} \tag{2.11}$$

♣

两个符号满足下面的等式 (感兴趣的同学可以自行验证)：

$$\epsilon_{ijk}\epsilon_{mnk} = \delta_{im}\delta_{jn} - \delta_{in}\delta_{jm} \tag{2.12}$$

有了这两个符号及求和约定，我们便可以告别向量运算的冗长展开式，并为引入张量做好了铺垫。下面来看一些例子：

例题 2.3 设 $\hat{x}_1, \hat{x}_2, \hat{x}_3$ 分别为右手直角坐标系 x, y, z 轴对应的单位向量，它们两两之间点乘和叉乘有什么性质？

容易证明，$\hat{x}_i \cdot \hat{x}_j = \delta_{ij}$，$\hat{x}_i \times \hat{x}_j = \epsilon_{ijk}\hat{x}_k$

例题 2.4 数量积、向量积、混合积与二重外积

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_i b_i = a_i b_j \delta_{ij} \tag{2.13}$$

$$\boldsymbol{a} \times \boldsymbol{b} = a_i \hat{x}_i \times b_j \hat{x}_j = a_i b_j \hat{x}_i \times \hat{x}_j = \epsilon_{ijk} a_i b_j \hat{x}_k \tag{2.14}$$

$$(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = \epsilon_{ijk} a_i b_j \hat{x}_k \cdot c_m \hat{x}_m = \epsilon_{ijk} a_i b_j c_m \delta_{km} = \epsilon_{ijk} a_i b_j c_k \tag{2.15}$$

$$(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} = \epsilon_{ijk} a_i b_j \hat{x}_k \times c_m \hat{x}_m = \epsilon_{ijk}\epsilon_{kmn} a_i b_j c_m \hat{x}_n = \epsilon_{ijk}\epsilon_{mnk} a_i b_j c_m \hat{x}_n \tag{2.16}$$

例题 2.5 证明：$(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{b} \cdot \boldsymbol{c})\boldsymbol{a}$

$$\begin{aligned}
(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} &= \epsilon_{ijk}\epsilon_{mnk} a_i b_j c_m \hat{x}_n \\
&= (\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}) a_i b_j c_m \hat{x}_n \\
&= a_m b_n c_m \hat{x}_n - a_n b_m c_m \hat{x}_n \\
&= (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{b} \cdot \boldsymbol{c})\boldsymbol{a}
\end{aligned} \tag{2.17}$$

例题 2.6（教材例 1.5.2） 证明：$(\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{c} \times \boldsymbol{d}) = (\boldsymbol{a} \cdot \boldsymbol{c})(\boldsymbol{b} \cdot \boldsymbol{d}) - (\boldsymbol{a} \cdot \boldsymbol{d})(\boldsymbol{b} \cdot \boldsymbol{c})$

$$\begin{aligned}
(\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{c} \times \boldsymbol{d}) &= \epsilon_{ijk} a_i b_j \hat{x}_k \cdot \epsilon_{mnp} c_m d_n \hat{x}_p \\
&= \epsilon_{ijk}\epsilon_{mnp} a_i b_j c_m d_n \delta_{kp} \\
&= \epsilon_{ijk}\epsilon_{mnk} a_i b_j c_m d_n \\
&= (\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}) a_i b_j c_m d_n \\
&= a_i b_j c_i d_j - a_i b_j c_j d_i \\
&= (\boldsymbol{a} \cdot \boldsymbol{c})(\boldsymbol{b} \cdot \boldsymbol{d}) - (\boldsymbol{a} \cdot \boldsymbol{d})(\boldsymbol{b} \cdot \boldsymbol{c})
\end{aligned} \tag{2.18}$$

例题 2.7（教材习题 1.15(1)） 证明：$(\boldsymbol{a} \times \boldsymbol{b})^2 = \boldsymbol{a}^2\boldsymbol{b}^2 - (\boldsymbol{a} \cdot \boldsymbol{b})^2$

$$\begin{aligned}
(\boldsymbol{a} \times \boldsymbol{b})^2 &= (\epsilon_{ijk} a_i b_j \hat{x}_k)^2 \\
&= (\epsilon_{ijk} a_i b_j)^2 \\
&= \sum_{i \neq j} a_i^2 b_j^2 \\
&= \sum_{i,j=1}^{3} a_i^2 b_j^2 - a_i^2 b_i^2 \\
&= \boldsymbol{a}^2\boldsymbol{b}^2 - (\boldsymbol{a} \cdot \boldsymbol{b})^2
\end{aligned} \tag{2.19}$$

定义 2.9 (二阶张量)

零阶张量为标量，它可以表示为一个数；一阶张量为矢量，它在选定了坐标系后可以写为列向量的形式，矢量 $\vec{f}$ 的完整表达式写为 $\vec{f} = f_i \hat{x}_i$；类似地，我们可以将二阶张量在选定的坐标系中写为矩阵的形式，张量 $\overleftrightarrow{T}$ 的完整表达式写为 $\overleftrightarrow{T} = T_{ij}\hat{x}_i\hat{x}_j$，其中 T_{ij} 为矩阵 T 的第 i 行第 j 列元素，$\hat{x}_i\hat{x}_j$ 是将两个矢量并写在一起的形式，称为并矢，它表示一个除第 i 行第 j 列元素为 1，其余元素均为 0 的矩阵。

矢量与并矢的点乘、叉乘满足就近法则，即

$$\vec{f}\cdot\vec{g}\vec{h}=(\vec{f}\cdot\vec{g})\vec{h},\qquad \vec{g}\vec{h}\cdot\vec{f}=\vec{g}(\vec{h}\cdot\vec{f}) \tag{2.20}$$

$$\vec{f}\times\vec{g}\vec{h}=(\vec{f}\times\vec{g})\vec{h},\qquad \vec{g}\vec{h}\times\vec{f}=\vec{g}(\vec{h}\times\vec{f}) \tag{2.21}$$

♣

例题 2.8 点乘

$$\vec{f}\cdot\overleftrightarrow{T}=(f_k\hat{x}_k)\cdot(T_{ij}\hat{x}_i\hat{x}_j)=(f_kT_{ij})(\hat{x}_k\cdot\hat{x}_i\hat{x}_j)=(f_iT_{ij})\hat{x}_j \tag{2.22}$$

$$\overleftrightarrow{T}\cdot\vec{f}=(T_{ij}f_j)\hat{x}_i \tag{2.23}$$

$$\overleftrightarrow{T}\cdot\overleftrightarrow{S}=(T_{ij}\hat{x}_i\hat{x}_j)\cdot(S_{kl}\hat{x}_k\hat{x}_l)=(T_{ij}S_{kl})(\hat{x}_i\hat{x}_j\cdot\hat{x}_k\hat{x}_l)=(T_{ij}S_{jl})\hat{x}_i\hat{x}_l \tag{2.24}$$

例题 2.9 叉乘

$$\vec{f}\times\overleftrightarrow{T}=(f_k\hat{x}_k)\times(T_{ij}\hat{x}_i\hat{x}_j)=(f_kT_{ij})(\hat{x}_k\times\hat{x}_i\hat{x}_j)=(\epsilon_{kil}f_kT_{ij})\hat{x}_l\hat{x}_j \tag{2.25}$$

$$\overleftrightarrow{T}\times\vec{f}=(\epsilon_{jkl}T_{ij}f_k)\hat{x}_i\hat{x}_l \tag{2.26}$$

二阶张量之间的叉乘涉及到更多的指标知识，在此不予赘述。

由例题 2.9 我们可以看到，叉乘的更广泛定义并非扩展到高维数组向量 (事实上，我们只在三维向量空间定义了叉乘运算)，而是推广到了高阶张量 (对应于更高维度的矩阵)。

类似二阶张量，我们可以给出更高阶张量的定义，它们对应着更高维度的矩阵。

等价关系与集合分拆

定义 2.10 (等价关系)

集合 A 上，满足以下三个条件的关系 $\sim$ 称为等价关系：

$$\forall a\in A, a\sim a;\quad (\text{自反}) \tag{2.27}$$

$$a\sim b\Rightarrow b\sim a;\quad (\text{对称}) \tag{2.28}$$

$$a\sim b, b\sim c\Rightarrow a\sim c.\quad (\text{传递}) \tag{2.29}$$

♣

例题 2.10 “朋友”、“年长”等关系不是等价关系；“同乡”、“同学”等关系是等价关系。

定义 2.11 (等价类)

集合 A 上定义了等价关系“$\sim$”，设 $a\in A$，称

$$\{b|b\in A, b\sim a\} \tag{2.30}$$

为等价类. 元素 a 称这个等价类的代表元. 由等价关系的传递性知，一个等价类中的任一元素均可称为这个等价类的代表元.

♣

定义 2.12 (集合的分拆)

设集合 A 的子集 $X_1, X_2, ..., X_n$ 满足

$$\bigcup_{i=1}^{n} X_i=A, \tag{2.31}$$

$$X_i\cap X_j=\emptyset,\quad for\quad i\neq j\quad and\quad i,j=1,2,...,n. \tag{2.32}$$

则称这些子集构成集合 A 的一个分拆.

♣

命题 2.3

集合的所有等价类构成一个分拆

命题 2.4

(1) 等价关系自反 $\Rightarrow \forall a \in A, \exists$ 等价类 $X_a, a \in X_a \Rightarrow \bigcup_{a\in A} X_a = A$;

(2)$X_a \cap X_b \neq \emptyset \Rightarrow \exists c \in X_a \cap X_b \Rightarrow c \sim a, c \sim b \Rightarrow a \sim b \Rightarrow X_a = X_b$

另一方面，当我们得到了集合的一个分拆 $A = \bigcup_i A_i$，那我们可以很容易定义一种等价关系：$a \sim b$ 当且仅当 a, b 同属一个 A_i，故我们推得：集合的分拆与定义在集合上的等价关系一一对应.

例题 2.11 本课程后续学习到的向量组等价、相抵关系、相似关系、相合关系都是等价关系，故我们可以根据它们分别给出集合的分拆方式。

求和符号的交换性

本节给出二重求和的一些交换性质，在今后关于行列式和矩阵乘法的相关证明中会涉及到

例题 2.12 当两个求和符号的上下限均为定值时，可直接交换求和号次序：

$$\sum_{i=k}^{l}\sum_{j=m}^{n} a_{ij} = \sum_{j=m}^{n}\sum_{i=k}^{l} a_{ij} \tag{2.33}$$

观察下方矩阵：

$$\begin{pmatrix} a_{k,m} & a_{k,m+1} & \cdots & a_{k,n} \\ a_{k+1,m} & a_{k+1,m+1} & \cdots & a_{k+1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{l,m} & a_{l,m+1} & \cdots & a_{l,n} \end{pmatrix} \tag{2.34}$$

式 (2.33) 的左边相当于先对矩阵每行元素求和，再将得到的结果求和；右边相当于先对矩阵每列元素求和，再将得到的结果求和。两种求和方式本质上都是对矩阵中所有元素求和，故二者相等。

有时我们直接使用一个求和符号来表示式 (2.33) 的双重求和：$\sum\limits_{k\le i\le l, m\le j\le n} a_{ij}$

例题 2.13 当一个求和号的上下限含有另一个求和号的指标时，交换求和号的同时需更改求和限：

$$\sum_{j=i}^{n}\sum_{l=j}^{n} a_{jl} = \sum_{l=i}^{n}\sum_{j=i}^{l} a_{jl} \tag{2.35}$$

其中 $i = 1, 2, ..., n$.

取定 $i \in \{1, 2, ..., n\}$，有下方矩阵：

$$\begin{pmatrix} a_{i,i} & a_{i,i+1} & \cdots & a_{i,n} \\ 0 & a_{i+1,i+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n,n} \end{pmatrix} \tag{2.36}$$

这个矩阵是一个上三角阵，式 (2.35) 的左边相当于先对矩阵每行元素求和，再将得到的结果求和；右边相当于先对矩阵每列元素求和，再将得到的结果求和。两种求和方式本质上都是对矩阵中所有元素求和，故二者相等。

有时我们直接使用一个求和符号来表示式 (2.35) 的双重求和：$\sum\limits_{i\le j\le l\le n} a_{jl}$

从上面两个例子我们可以看到，使用矩阵来观察并证明二重求和的相关性质是非常直观的。有了二重求和号的交换性质，我们可以很自然地推导出多重求和的交换性质，在此不做展开。

2.1.2 第二章复习与补充知识

线性方程组的 n,m,r 及解的情况判别

定义 2.13 (线性方程组的初等变换)

(1) 互换两方程位置；
(2) 用一个非零数乘某个方程；
(3) 把一个方程的倍数加到另一个方程上；
容易证明，经过一系列初等变换得到的方程组与原方程组同解。♣

定义 2.14 (线性方程组的系数矩阵与增广矩阵)

将线性方程组各未知数的系数按原本的次序排列为一张表，称为方程组的系数矩阵；系数矩阵最右边加上一列常数项构成一张表，称为方程组的增广矩阵 ♣

类似线性方程组的初等变换，我们可以定义矩阵的初等行变换：

定义 2.15 (矩阵的初等行变换)

(1) 互换矩阵两行；
(2) 用一个非零数乘某一行；
(3) 把一行的倍数加到另一行上. ♣

定义 2.16 (阶梯形矩阵)

满足以下性质的矩阵称阶梯形矩阵：
(1) 元素全为 0 的行 (称为零行) 在下方 (如果存在零行)；
(2) 元素不全为 0 的行 (非零行) 从左起第一个不为 0 的元素称为主元，他们的列指标随着行指标的递增而严格增大。(由此易证主元的列指标大于等于其行指标)
可以证明，任意矩阵可通过初等行变换化为阶梯形矩阵 (即消元法的过程) ♣

有了以上的定义，我们来尝试给出有 $\boldsymbol{m}$ 个方程的 $\boldsymbol{n}$ 元线性方程组解的情况，方程组的增广矩阵可以写为：

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} \tag{2.37}$$

设我们通过一系列初等行变换，将增广矩阵化成了阶梯形矩阵：

$$\begin{pmatrix} 0 & \cdots & a'_{1,p_1} & \cdots & \cdots & \cdots & \cdots & a'_{1n} & a'_{1,n+1} \\ 0 & 0 & \cdots & a'_{2,p_2} & \cdots & \cdots & \cdots & a'_{2n} & a'_{2,n+1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & a'_{r,p_r} & \cdots & a'_{rn} & a'_{r,n+1} \\ 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & & & & & \end{pmatrix} \tag{2.38}$$

这个矩阵有 r 个非零行[1]，从第 $(r+1)$ 行开始矩阵元素均为 0，我们将第 $i(1\le i\le r)$ 个非零行的主元记为 a'_{i,p_i}，此时我们对 p_r 的取值做如下讨论：

(1)$\boldsymbol{p_r=n+1}$，这意味着 a'_{r,p_r} 即为 $a'_{r,n+1}$，此时阶梯矩阵的第 r 行代表着一个 "$0=a'_{r,n+1}$" 的矛盾方程，则此

[1]注意，我们在这里定义的 r 与老师上课定义的 r 稍有区别，老师定义的 r 是系数阶梯矩阵的非零行个数，而这里定义的 r 是增广阶梯矩阵非零行的个数。相应的，老师定义的 r 即是系数矩阵的秩，而这里定义的 r 是增广矩阵的秩

时方程组**无解**。

(2)$\boldsymbol{p_r \le n}$，由于 $r \le p_r$，则 $r \le n$，这时不存在矛盾方程，对应的方程组一定有解，我们分两种情形讨论：

(2.1)$\boldsymbol{r = n}$，此时对第 i 个主元有 $i = p_i$，即主元依次为 $a'_{11}, a'_{22}, \cdots, a'_{nn}$，可以看到，这样的阶梯矩阵对应的方程组没有自由未知量，从第 n 行向上依次求解可以得到方程组的唯一一组解。

(2.2)$\boldsymbol{r < n}$，这样的阶梯矩阵对应的方程组有 n 个未知量和 r 个有效方程，未知量多于有效方程个数，说明部分未知量为自由未知量，可作为参数任意变化，故可以得到无穷多个解。在下一小节我们将介绍这种情况下通解的形式。

注 当第 r 行主元的列指标 p_r 不为 $n+1$ 时，可以看到我们定义的 r 与老师定义的 r 是相等的，这也就是说: 线性方程组有解 $\iff$ 它的增广矩阵的秩等于系数矩阵的秩。

无穷多解情况下方程的通解形式

方程组的增广矩阵化为阶梯形矩阵之后，若为 $r < n$ 的情况，方程组会有无穷多个解，下面我们来求它的通解形式：

我们将每个主元对应的未知数称为**主变量**，其余未知量称为**自由未知量**，可知自由未知量的个数为 $n-r$ 个。在阶梯形方程中，我们将所有自由未知量移项到等号右边，并将它们用参数 $t_1, t_2, \cdots, t_{n-r}$ 表示，这样，我们的 r 个主变量可以用常数和 $n-r$ 个参数表示出来：

$$\begin{pmatrix} x_{p_1} \\ x_{p_2} \\ \vdots \\ x_{p_r} \end{pmatrix} = \sum_{i=1}^{n-r} t_i \begin{pmatrix} \alpha_{p_1,i} \\ \alpha_{p_2,i} \\ \vdots \\ \alpha_{p_r,i} \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_r \end{pmatrix} \tag{2.39}$$

而对于自由未知量，它们的取值就是各自对应的参数，从而我们得到了方程组含有参数 $t_1, t_2, \cdots, t_{n-r}$ 的通解，由于 $n-r$ 个参数可以任意取定，故方程组的解也是无穷多的。

例题 2.14 求解线性方程组

$$\begin{cases} 2x_1 + x_2 - 3x_3 + 5x_4 = 6 \\ -3x_1 + 2x_2 + x_3 - 4x_4 = 5 \\ -x_1 + 3x_2 - 2x_3 + x_4 = 11 \end{cases} \tag{2.40}$$

解 方程组的增广矩阵写为

$$\begin{pmatrix} 2 & 1 & -3 & 5 & 6 \\ -3 & 2 & 1 & -4 & 5 \\ -1 & 3 & -2 & 1 & 11 \end{pmatrix} \tag{2.41}$$

经一系列矩阵初等行变换，它变成了阶梯形矩阵

$$\begin{pmatrix} -1 & 3 & -2 & 1 & 11 \\ 0 & -7 & 7 & -7 & -28 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.42}$$

可见，主变量为 x_1, x_2，自由变量为 x_3, x_4，将它们设为参数 t_1, t_2，则方程组的通解可写为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} t_1 + \begin{pmatrix} -2 \\ -1 \\ 0 \\ 1 \end{pmatrix} t_2 + \begin{pmatrix} 1 \\ 4 \\ 0 \\ 0 \end{pmatrix} \tag{2.43}$$

齐次线性方程组

定义 2.17 (齐次线性方程组)

常数项全为 0 的线性方程组称为齐次线性方程组，容易看出，$(0, 0, \cdots, 0)^T$ 是齐次线性方程组的一组解，称为零解，其余的解称为非零解。♣

由于一个齐次线性方程组必有零解，则它的解的情况只有唯一解和无穷多解两种，不存在无解情况。由上上小节的讨论我们可以看到：齐次线性方程组存在非零解的充要条件是它的系数矩阵经初等行变换化成的阶梯矩阵中非零行数目 $r < n$。

例题 2.15 下列说法正确的是：(D)

A. 线性方程组当未知量个数少于方程个数时无解

B. 齐次线性方程组当未知量个数少于方程个数时无解

C. 线性方程组当未知量个数多于方程个数时有解

D. 齐次线性方程组当未知量个数多于方程个数时有无限多个解

2.2 第二次习题课

首先是作业，这几周的作业其实相对来说都比较简单，所以我就提几个常见的问题，大家以后注意就好了。

2.2.1 第二、三周作业常见问题

虽然只有四道题，但还是有不少同学写错。各种的计算错误就没什么提的必要了，值得一说的是以下几点：

- 最后一题（第二章第 7 题）是一道应用题，所以要结合实际生活来考虑，一些同学没有考虑到解必须均为正数，或者考虑过多，认为食品量必须是整数。以及甚至还有同学觉得三个未知数三个方程就一定有解，方程里的常数项不是 0，所以这里可不是齐次方程组啊！在后续的课程中会学到，必须满足系数矩阵的秩 (rank) 与增广矩阵的秩相等才会有解。
- 同时我改作业时发现大部分同学对于矩阵语言还不够熟悉，还是在用方程做变换代入消元的做法，这种做法本身肯定是没什么问题的，但是线性方程组只不过是线性代数一个小小的应用，是一个具象化的问题，而数学的核心便是一个从具体概括到抽象的过程。从线性方程组来引入矩阵的概念，是想让大家更好地理解“线性性”这一概念是如何产生的。在有了矩阵之后，我们便有了一个更高的视角，比如你可以发现，矩阵既可以用来表达一组线性方程组，也可以用来表示 n 维欧氏空间[2]中的一组向量，以及更多的满足所谓“线性性”的“向量空间”中的一组元素，于是这又引出了两个问题：“什么是线性性？”、“什么是向量？”$\cdots\cdots$ 这就是数学公理化要做的任务，在此不再赘述，感兴趣的同学可以自行学习。说了这么多也只是为了让大家多了解、熟悉并矩阵的语言，具体的过程可以参考第二、三周作业答案的解题过程。

2.2.2 第四周作业常见问题

- 怎么这么多人四面体体积不会算...... 系数是 $\frac{1}{6}$ 啊，有没乘系数的，还有乘 $\frac{1}{4}$ 的、$\frac{1}{3}$ 的或者 $\frac{1}{2}$ 的，以及怎么还有三角形面积不除以 2 的。另外有同学求出来是负数，我只是圈了出来，但给算对了，这种没明确说的话，应该要求的还是正常的体积而不是有向体积。
- 此外依然是很多的计算错误，代数余子式的符号是 -1 的脚标和次幂，如果算行列式的时候不是按第一行或第一列展开的话一定要注意符号啊，以及最好说清楚是按哪行或哪列展开。
- 第一题算具体数值的行列式，我看见了好几个同学硬展开的，这样真的不累吗？而且一不小心就算错了。可以先做做变换再求，既不容易算错也更简便。
- 由于作业不多且不难，具体讲解哪些作业题视到场同学需求决定。
- 还有个小细节是多元多项式如果对称的话，肯定化成一个对称的形式是最好的，同时也比较美观。

2.2.3 拓展内容

集合与映射

集合与映射之间存在着密切的联系，比如我们可以利用双射来定义有限集合与无限集合：（如果不从直观出发，如何去描述集合中元素是有限还是无限？或者说，什么叫有限/无限？这也是一个直观到抽象的过程）

定义 2.18 (有限集合与无限集合)

如果存在自然数 n，使得集合 A 与集合 $1, 2, \cdots, n$ 之间有一一对应（即双射），称 A 为有限集合；反之若这样的 n 不存在，则称 A 为无限集合，如自然数集 $\mathbb{N}$、整数集 $\mathbb{Z}$ 是熟知的无限集合。♣

自然数集 $\mathbb{N}$ 有一个有趣的性质，即 $\mathbb{N}$ 可以和它的每个无限子集一一对应，比如 $x \mapsto x/2$ 是偶数集到 $\mathbb{N}$ 的一个一一对应。通过自然数集我们可以定义可数集与不可数集：

[2] 注意不是欧式空间，这看起来像是搞装修的 (雾)

定义 2.19 (基数与可数集合)

称两个集合 A,B 有相同的基数，是指存在 $A\to B$ 的一一对应。称集合 A 比集合 B 有更大的基数，是指存在 $A\to B$ 的满射，但不存在 A,B 之间的一一对应。一个与 $\mathbb{N}$ 有相同基数的集合称为可数集合。♣

思考：有理数集 $\mathbb{Q}$ 是可数集合吗？

解 我们知道有理数是两个整数之比（分母不为 0），而我们知道整数集是可数的，那么我们可以任取整数集 $\mathbb{Z}$ 的一个排列 $x_1,x_2,\cdots$，再取 $\mathbb{Z}\backslash\{0\}$ 的任意一个排列 $y_1,y_2,\cdots$，记 $z_{ij}=x_i/y_j$，那么 z_{ij} 构成了全部的有理数。

$$\begin{pmatrix} z_{11} & z_{12} & z_{13} & \cdots \\ z_{21} & z_{22} & z_{23} & \cdots \\ z_{31} & z_{32} & z_{33} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

沿着这个无穷矩阵的每一条对角线将元素进行排列可以得到：

$$z_{11}\to z_{21}\to z_{12}\to z_{31}\to z_{22}\to z_{13}\to\cdots$$

由此我们可以得到一个 $\mathbb{N}$ 到 $\mathbb{Q}$ 的满射（注意不是单射），同时很容易构造一个 $\mathbb{Q}$ 到 $\mathbb{N}$ 的满射（因为 $\mathbb{N}$ 包含在 $\mathbb{Q}$ 中），从而 $\mathbb{Q}$ 和 $\mathbb{N}$ 的基数相同，由此得到 $\mathbb{Q}$ 是一个可数集合。

笔记 利用同样的方法可以得到：可数个可数集合的并是可数集合。

思考：那么是否所有的无限集合都有相同的基数呢？直观上来看，我们会感觉实数集 $\mathbb{R}$ 的基数是要比 $\mathbb{N}$ 大的，幸运的是，这个直觉的确是对的。可以先来证明这样一个命题：

命题 2.5

不存在基数比 $\mathbb{N}$ 小的无限集合（任意无限集合 A 的基数大于或等于 $\mathbb{N}$）。♠

证明 考虑一个无限集合 A, 从中取一个元素 x_1, 再从 $A\backslash\{x_1\}$ 中取一个元素 x_2, 重复操作得到 A 中两两不同的一列元素 $x_1,x_2,\cdots$, 从而可以得到一个单射 $f:\mathbb{N}\to A$，由此定义 $g:A\to\mathbb{N}$ 如下：

$$\begin{cases} g(x)=f^{-1}(x), & x\in f(\mathbb{N}), \\ g(x)=1, & x\notin f(\mathbb{N}). \end{cases}$$

从而 g 是 $A\to\mathbb{N}$ 的满射。如果存在 $A\to\mathbb{N}$ 的一一对应，那么 A 与 $\mathbb{N}$ 有相同的基数；如果不存在，那么 A 的基数大于 $\mathbb{N}$ 的基数，从而命题得证。

由此我们可以定义不可数集合：

定义 2.20 (不可数集合)

无限集合称为不可数集合，是指不存在它与 $\mathbb{N}$ 之间的一一对应，即它有比 $\mathbb{N}$ 更大的基数。♣

那么我们如何来说明 $\mathbb{R}$ 的基数比 $\mathbb{N}$ 大呢？先思考一下 $\mathbb{R}$ 中的元素都是什么：**十进制小数**！对于每一个十进制小数，它的小数点后的数位是可数的，那么我们能联想到什么？没错，就是直积（笛卡尔积）。

从下面这个命题的证明，我们就能得到 $\mathbb{R}$ 不可数的结论。

命题 2.6

可数个可数集合的直积是不可数集合。♠

证明 记 $A=\prod_{n=1}^{\infty}A_n$，$A_n$ 均为可数集合，则存在满射 $f_n:A_n\to E_n$，其中 $E_1=\mathbb{Z},E_n=\{0,1,\cdots,9\}(n\geqslant 2)$，记 $E=\prod_{n=1}^{\infty}E_n$，那么存在满射 $f:A\to E$，则 A 具有比 E 更大的基数。而容易看出 E 与 $\mathbb{R}$ 之间存在着一一对应：我们可以把 E 中元素写为十进制小数的形式，任取 E 中的元素，取其第一个分量作为整数部分，第 n 个分量作为小数点后第 $(n-1)$ 位 $(n\geqslant 2)$，由此便可得到一个实数，因此接下来只需证明 $\mathbb{R}$ 不可数。假设 $\mathbb{R}$ 可数，

那么我们可以把 $\mathbb{R}$ 中元素排成一列 $x_1, x_2, \cdots$，取 $x \in \mathbb{R}$ 满足其整数部分与 x_1 不同，小数点后第 n 位与 x_{n+1} 的小数点后第 n 位不同 $(n \geqslant 1)$，那么 x 和排列中的任一个元素都不相同，矛盾！从而 $\mathbb{R}$ 不可数，命题也得证。

接下来，关于课程讲义中提及的集合的一个运算规则（命题 1.7.2）：

命题 2.7 (De Morgan 律)

设 A_i 为某固定集合 U 的子集，则

$$\bigcap_{i\in I} A_i^c = \left(\bigcup_{i\in I} A_i\right)^c \quad \bigcup_{i\in I} A_i^c = \left(\bigcap_{i\in I} A_i\right)^c$$

♠

这个性质是不难验证的，但值得一提的是，这里的下标集 I 可能是一个不可数集合，此时这族集合的交与并不能写成 $\bigcap_{n=1}^{\infty}$ 或 $\bigcup_{n=1}^{\infty}$ 的形式，因而我们才采取这种表示形式，来表明这个运算律对**任意交**、**任意并**都成立，而非仅在有限或可数次交或并运算下成立。

复数

¶ 扩充平面与复数的球面表示

我们知道复数可以和复平面上的点一一对应，从而可以通过复平面给出复数的一个几何表示。而 Riemann 首先引进了复数的球面表示。首先我们在复数域 $\mathbb{C}$ 中引进一个新的数 ∞，这个数的模是 ∞，辐角没有意义，其与其他复数的运算规则规定为：

$$z \pm \infty = \infty,\ z \cdot \infty = \infty (z \neq 0),\ \frac{z}{\infty} = 0,\ \frac{z}{0} = \infty (z \neq 0)$$

而 $0 \cdot \infty$ 和 $\infty \pm \infty$ 都不规定其意义。引入了 ∞ 的复数域记为 $\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\}$，这个时候我们发现，在复平面上没有任何一点和 ∞ 对应，于是我们想象有一个无穷远点和 ∞ 对应，加上无穷远点的复平面称为**扩充平面**，但是我们会感觉这样的 ∞ 和其他复数在复平面上总是有些差别，于是便有了复数的球面表示：

设 $S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$，即 $\mathbb{R}^3$ 中的单位球面，把 $\mathbb{C}$ 看作平面 $\{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\}$，固定 S 的北极点 $N = (0, 0, 1)$，在平面上任取一点 z，其与 N 相连的直线必与 S 交于一点 P，从而 z 与 P 相对应，而 ∞ 与 N 相对应，从而我们得到了一个 $\mathbb{C}_\infty$ 与 S 之间的一一对应，而且此时 ∞ 与其他复数地位对等。

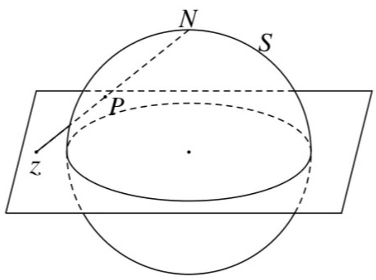


图 2.1: 球极投影

¶ 棣莫弗（De Moivre）公式与因式分解

利用棣莫弗（De Moivre）公式：$(re^{i\theta})^n = (r\cos(\theta) + ir\sin(\theta))^n = r^n e^{in\theta} = r^n(\cos(n\theta) + i\sin(n\theta))$ 我们可以将一些多项式分解为一次复系数因式的乘积，如：

习题（因式分解）

(1) $x^n - C_{2n}^2 x^{n-1} + C_{2n}^4 x^{n-2} + \cdots + (-1)^n C_{2n}^{2n}$

(2) $x^{2n} + C_{2n}^2 x^{2n-2}(x^2 - 1) + C_{2n}^4 x^{2n-4}(x^2 - 1)^2 + \cdots + (x^2 - 1)^n$

(3) $x^{2n+1}+C_{2n+1}^2x^{2n-1}(x^2-1)+C_{2n+1}^4x^{2n-3}(x^2-1)^2+\cdots+x(x^2-1)^n$

解 (1) 在棣莫弗公式中将 r 取为 1，n 替换为 $2n$，两边取实部得：

$$\cos^{2n}(\theta)-C_{2n}^2\cos^{2n-2}(\theta)\sin^2(\theta)+\cdots+(-1)^n\cdot C_{2n}^{2n}\cdot\sin^{2n}(\theta)=\cos(2n\theta)$$

取 $\theta_k=\frac{k\pi+\pi/2}{2n},k=0,1,\cdots,n-1$,(易知 $\theta_k<\dfrac{\pi}{2}$), 此时上式右端为零，而 $\sin(\theta)\neq 0$, 故两端同时除以 $\sin^{2n}(\theta)$:

$$\cot^{2n}(\theta_k)-C_{2n}^2\cot^{2n-2}(\theta_k)+\cdots+(-1)^n\cdot C_{2n}^{2n}=0$$

即 $x_k=\cot^2(\theta_k)$ 是原式的根，又因为它们互不相同，故恰为原式的 n 个不同根，从而得到分解：

$$x^n-C_{2n}^2x^{n-1}+C_{2n}^4x^{n-2}+\cdots+(-1)^nC_{2n}^{2n}=\prod_{k=0}^{n-1}(x-\cot^2(\frac{k\pi+\pi/2}{2n}))$$

(2) 将 $\sin^2(\theta_k)=1-\cos^2(\theta_k)$ 带入第 (1) 问第一个式子，可得 $x_k=\cos(\theta_k),\theta_k=\frac{k\pi+\pi/2}{2n},k=0,1,\cdots,2n-1$ 恰好是第 (2) 问原式的所有根. 从而得到分解：

$$x^{2n}+C_{2n}^2x^{2n-2}(x^2-1)+C_{2n}^4x^{2n-4}(x^2-1)^2+\cdots+(x^2-1)^n=\prod_{k=0}^{2n-1}(x-\cos\left(\frac{k\pi+\pi/2}{2n}\right))$$

(3) 在棣莫弗公式中将 r 取为 1，n 替换为 $2n+1$，两边取实部，类似 (1) (2) 中过程可得分解：

$$x^{2n+1}+C_{2n+1}^2x^{2n-1}(x^2-1)+C_{2n+1}^4x^{2n-3}(x^2-1)^2+\cdots+x(x^2-1)^n=\prod_{k=0}^{2n}(x-\cos\left(\frac{k\pi+\pi/2}{2n+1}\right))$$

行列式的几何意义

行列式在线性代数中有着很重要的地位，然而绝大部分教材对行列式的引入都是从代数视角出发，一串繁杂的代数式经常让初学者十分困惑：为什么是这样一个奇怪的计算方法，算出来的这个数到底有什么作用？其实这些都有着鲜明的几何直观：

命题 2.8 (行列式的几何意义)

对于 n 阶方阵 A，将其看作 n 个行向量（或列向量），以这 n 个向量构成一个新的坐标系中的 n 个基向量，则 A 的行列式即为这 n 个向量张成的高维立方体在 n 维空间中的有向体积（二维情形下即为面积）。

可是...... 这个命题看上去也奇奇怪怪的，难以理解，所以接下来我们先从简单的低维情形入手。

考虑二阶行列式：

$$\begin{vmatrix}a & b\\ c & d\end{vmatrix}=ad-bc$$

从这个我们能想到什么？我们把这个二阶方阵看成两个列向量 $\boldsymbol{x}=(a,c)^T,\boldsymbol{y}=(b,d)^T$，同时记单位向量 $\boldsymbol{i}=(1,0)^T,\boldsymbol{j}=(0,1)^T$，不妨先假设 $\boldsymbol{y}$ 在 $\boldsymbol{x}$ 的逆时针方向，那么行列式的值便是 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 围成的平行四边形的面积，如下图所示：

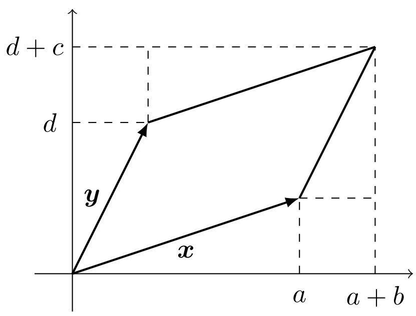


图 2.2: 平行四边形

平行四边形面积 $S=(a+b)(d+c)-ac-bd-2bc=ad-bc$. 为什么命题中说的是“有向体积”呢？这涉及到一个定向的问题，$\boldsymbol{x}$ 在 $\boldsymbol{y}$ 的顺时针方向一侧时，这与 $\boldsymbol{i},\boldsymbol{j}$ 之间的相对位置关系是一样的，两组向量的叉乘 ($\boldsymbol{x}\times\boldsymbol{y}$与$\boldsymbol{i}\times\boldsymbol{j}$) 方向垂直纸面向外，所以我们把它定为正向，而当 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的相对位置关系颠倒后，可以视为以 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 为基的新坐标系把纸面翻转到了背面，从而“体积（这里是面积）”便是负的了。

从上面的过程可以看出这便是二阶行列式交换两列后符号改变的原因，同理，将 $\boldsymbol{x}$ 或 $\boldsymbol{y}$ 变为原来的 λ 倍后，平行四边形沿着 $\boldsymbol{x}$ 或 $\boldsymbol{y}$ 的方向伸缩了 λ 倍，因此面积变为原来的 λ 倍。而将 $\boldsymbol{x}$ 的 λ 倍加到 $\boldsymbol{y}$ 后，相当于 $\boldsymbol{y}$ 沿着 $\boldsymbol{x}$ 的方向做了一个平移，平行四边形的面积自然不会发生改变。而在 $\boldsymbol{x}$ 基础上加一个向量 $\boldsymbol{z}$，那么 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 围成的平行四边形的面积加上 $\boldsymbol{z}$ 与 $\boldsymbol{y}$ 围成的平行四边形的面积就等于 $\boldsymbol{x}+\boldsymbol{z}$ 与 $\boldsymbol{y}$ 围成的平行四边形的面积，也就得到了二阶行列式可将某列拆分的性质，如下图所示：

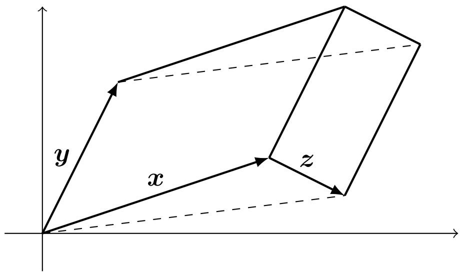


图 2.3: 二阶行列式可按某列拆分

以上也是一般的教科书中可能会给出的一个几何解释。

那么，这时候可能又会有同学问：“我觉得过程里这个 $S=(a+b)(d+c)-ac-bd-2bc=ad-bc$. 还是太‘代数’了，不够直观，并且如何来解释方阵转置后行列式不变呢？还有更高维的情形怎么办呢？”

这当然是个好问题，但想给出严谨的论述比较困难，可以从下面这个视角来看待，仍先考虑二维情形：

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}=\begin{vmatrix} a & 0 \\ c & d \end{vmatrix}+\begin{vmatrix} 0 & b \\ c & d \end{vmatrix}=\begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix}+\begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix}+\begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix}+\begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix}=\begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix}+\begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix}$$

我们可以把这个看作是以 $(a,0)^T$ 与 $(0,d)^T$ 为边的长方形的有向面积加上以 $(0,c)^T$ 与 $(b,0)^T$ 为边的长方形的有向面积，或者再多分一步来看，先将 $\boldsymbol{i}$ 与 $\boldsymbol{j}$ 分别伸缩为原来的 a,d 倍，再将 $\boldsymbol{i}$ 沿 $\boldsymbol{j}$ 的方向平移 c 个单位、$\boldsymbol{j}$ 沿 $\boldsymbol{i}$ 的方向平移 b 个单位，$-bc$ 便是这两次偏移产生的影响。同样对于三维情形也有：

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}=\begin{vmatrix} a_1 & 0 & 0 \\ 0_1 & b_2 & 0 \\ 0 & 0 & c_3 \end{vmatrix}+\begin{vmatrix} 0 & a_2 & 0 \\ 0 & 0 & b_3 \\ c_1 & 0 & 0 \end{vmatrix}+\begin{vmatrix} 0 & 0 & a_3 \\ b_1 & 0 & 0 \\ 0 & c_2 & 0 \end{vmatrix}+\begin{vmatrix} a_1 & 0 & 0 \\ 0 & 0 & b_3 \\ 0 & c_2 & 0 \end{vmatrix}+\begin{vmatrix} 0 & a_2 & 0 \\ b_1 & 0 & 0 \\ 0 & 0 & c_3 \end{vmatrix}+\begin{vmatrix} 0 & 0 & a_3 \\ 0 & b_2 & 0 \\ c_1 & 0 & 0 \end{vmatrix}$$

可以看出这是六个长方体的有向体积和。一般地，取 n 维标准欧氏空间中的基向量 $\boldsymbol{e}_i$（仅第 i 个分量为 1，其余为 0），这 n 个向量有 $n!$ 个排列方式，我们把它们视为 $n!$ 个不同的“定向”，将排列 $(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)$ 定为正向，对于 n 阶方阵 $A=(a_{ij})_{n\times n}$，将行列式按如上方式拆分，可以得到 $n!$ 个高维长方体（各边两两正交）的有向体积和，同时也可以看作每个 $\boldsymbol{e}_i$ 先各自伸缩 a_{ii} 倍，再加上其他 $(n!-1)$ 个“定向”上的偏移的影响（影响可能有正有负，实际上这与排列的逆序数及偏移的正负有关）。而矩阵转置前后，对角线不变，仍然以列向量来看，这些偏移只是在列向量之间“转移”，其和仍是不变的，从而便有转置不改变行列式。

思考一下，你就会发现，其实这就是行列式的完全展开式。

笔记 由于助教水平有限，这段话实际上比较粗糙，只是提供一个相对直观的视角与看法而已。另一方面，其实几何直观在矩阵中的运用往往是用于猜测结论，而非严谨证明结论。

现在，通过以上的过程，我们就可以将结论推广到三维以及更高维度的情形，只需要把平行四边形改为平行六面体/高维平行立方体就可以了。那么行列式为 0 又是什么情形呢？平行四边形面积为 0，那么只有 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 共线了或都为 $\mathbf{0}$ 了，平行六面体的在三维空间中的体积为 0，只能是三个基向量共面或共线或都为 $\mathbf{0}$，更高维情况也类似，概括下来的话，当行列式里的向量线性相关时，这个空间就“降维”了。

笔记 另外，关于行列式的递推定义，其实可以通过将列向量投影到标准欧氏空间来理解，将 1 个 n 维平行立方体的一边在 n 个分量上作正交投影，于是便转化为了 n 个 $(n-1)$ 维平行立方体，实际上与上面的解释是一致的。

一个典型的错误例子 有位同学前几天问过我第三章的第 8 题他为什么证错了，下面我们一起来看一下这位同学写的过程，有了直观的几何理解后，我们很轻易就能发现错误之处。(这里没有针对该同学的意思，只是我感觉这个错误很多初学者都可能会犯)

习题（第三章第 8 题）

设 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d}$ 为 4 维数组向量，证明：$\det(2\boldsymbol{a}-\boldsymbol{b},-\boldsymbol{a}+2\boldsymbol{b}-\boldsymbol{c},-\boldsymbol{b}+2\boldsymbol{c}-\boldsymbol{d},-\boldsymbol{c}+2\boldsymbol{d})=5\det(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d})$.

证明 （错解）

$$\begin{aligned}
\text{原式} &= \det(2\boldsymbol{a}-\boldsymbol{b},\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c},\boldsymbol{b}+\boldsymbol{c}-\boldsymbol{a}-\boldsymbol{d},\boldsymbol{c}+\boldsymbol{d}-\boldsymbol{b})\\
&= \det(3\boldsymbol{a}-\boldsymbol{c},2\boldsymbol{b}-\boldsymbol{d},2\boldsymbol{c}-\boldsymbol{a},\boldsymbol{a}+2\boldsymbol{d}-2\boldsymbol{b})\\
&= \det(3\boldsymbol{a}-\boldsymbol{c},2\boldsymbol{b}-\boldsymbol{d},2\boldsymbol{c}-\boldsymbol{a},\boldsymbol{a}+\boldsymbol{d})\\
&= \det\left(\frac{5}{2}\boldsymbol{a},2\boldsymbol{b}-\boldsymbol{d},2\boldsymbol{c}-\boldsymbol{a},\boldsymbol{a}+\boldsymbol{d}\right)=\frac{5}{2}\det(\boldsymbol{a},2\boldsymbol{b}-\boldsymbol{d},2\boldsymbol{c},\boldsymbol{d})\\
&= 10\det(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d}).
\end{aligned}$$

哪里出了问题呢？显然是第二个等号错了，我们仔细地看这一步，记过程第一行四个向量为 $\boldsymbol{v}_i(i=1,2,3,4)$，我们把 $\boldsymbol{v}_i$ 的常数倍加到 $\boldsymbol{v}_j$ 中时，可以看作是 $\boldsymbol{v}_j$ 沿着 $\boldsymbol{v}_i$ 方向作了一个平移，而第二个等号中，这位同学是把 $\boldsymbol{v}_2$ 加到 $\boldsymbol{v}_1$ 中，把 $\boldsymbol{v}_3$ 加到 $\boldsymbol{v}_2$ 中，把 $\boldsymbol{v}_4$ 加到 $\boldsymbol{v}_3$ 中，这三步都没问题，但是这时候第三个向量已经变成了 $\boldsymbol{v}_3'=2\boldsymbol{c}-\boldsymbol{a}$，如果要想让行列式不变，我们只能把 $\boldsymbol{v}_4$ 沿着 $\boldsymbol{v}_3'$ 的方向作平移，但第二个等号后的第四个向量明显是 $\boldsymbol{v}_4-\boldsymbol{v}_3$ 的结果，即 $\boldsymbol{v}_4$ 沿着 $\boldsymbol{v}_3$ 方向的平移，正是这个错误导致了行列式的变化。

Cramer 法则的几何解释

有了上一部分的理解后，我们是不是也可以给 Cramer 法则一个好的几何解释呢？当然是可以的。这里只讨论二维情形，三维或更高维情形是同理的。

考虑一个二阶线性方程组：$\begin{cases} a_1x+b_1y=c_1\\ a_2x+b_2y=c_2 \end{cases}$ 我们记 $\boldsymbol{a}=(a_1,a_2)^T,\boldsymbol{b}=(b_1,b_2)^T,\boldsymbol{c}=(c_1,c_2)^T,\boldsymbol{x}=(x,y)^T$.

那么上述方程组可以改写为：$x\boldsymbol{a}+y\boldsymbol{b}=\boldsymbol{c}$，那么我们以 $\boldsymbol{a},\boldsymbol{b}$ 作为基向量构成新坐标系，将 $\boldsymbol{a},\boldsymbol{b}$ 分别伸缩为原来的 x,y 倍，从而可以得到 $\boldsymbol{c}$，如下图所示：

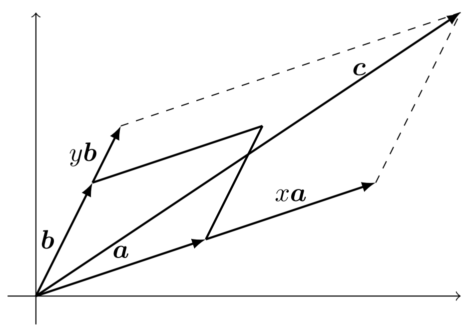


图 2.4: Cramer 法则的几何解释

利用面积比值关系，我们可以得到：$x=\dfrac{|\boldsymbol{c},\boldsymbol{b}|}{|\boldsymbol{a},\boldsymbol{b}|}$，$y=\dfrac{|\boldsymbol{a},\boldsymbol{c}|}{|\boldsymbol{a},\boldsymbol{b}|}$，这便是 Cramer 法则了。

说了这么多，希望大家对行列式能有一个新的认识，能发现线性代数这门学科的优美之处，而非仅仅只是知道了计算方法，把自己搞成了一个矩阵计算器（虽然这门课的考试确实比较偏计算，我也无能为力）。

2.2.4 补充题目

行列式在这门课的期中期末里都是必考题，一些常见的求解行列式的技巧很有必要掌握，所以这里给大家搞了一点点补充题。题目大多数摘录于本门课程的往年考试原题、课外资料中的题目等等。不过其中部分题目难度较高，仅作了解便可，不必深究（除非你有转入数学相关专业的打算 XD）。单看这门课的要求的话，掌握一些基本解题技巧、熟悉基本概念便足以拿到一个不错的成绩了。

笔记 由于我们现在还没学到第四章，所以其实能补充的题目并不是很多。

习题 2.1 (2022 秋期中)

行列式 $\begin{vmatrix} -3 & x & 7 & x \\ 1 & x & 5 & -1 \\ 4 & 3 & -x & 1 \\ 2x & -1 & 1 & x \end{vmatrix}$ 的完全展开式中，x^4 项的系数为：________.

解 沿第一列展开，只有最后一项 $-2x\begin{vmatrix} x & 7 & x \\ x & 5 & -1 \\ 3 & -x & 1 \end{vmatrix}$ 能含有 x^4 项，而 $\begin{vmatrix} x & 7 & x \\ x & 5 & -1 \\ 3 & -x & 1 \end{vmatrix}$ 再沿第一列展开，只有第二项 $-x\begin{vmatrix} 7 & x \\ -x & 1 \end{vmatrix}$ 能含有 x^3 项，且系数为 -1，故原式中 x^4 项的系数为 $-2\times(-1)=2$.

习题 2.2

将 λ 作为变量，$a_{ij}\ (1\leqslant i,j\leqslant n)$ 作为常数，则：

$$f(\lambda)=\begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix}$$

是 λ 的多项式. 求这个多项式的 n 次项和 $n-1$ 次项系数

解 行列式只有 n 个对角元 $\lambda-a_{ii}$，各含有一个 λ, 其余元素都不含 λ. 因此只有这 n 个对角元的乘积 $(\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn})$ 才含 λ 的 n 次项，就是这个乘积的最高次项 λ^n. 因此 $f(\lambda)$ 的 n 次项为 λ^n, 系数为 1。

同时行列式展开式 $f(\lambda)$ 中的 $n-1$ 项只能由 n 个对角元 $\lambda-a_{ii}$ 中的某 $n-1$ 个相乘产生。先选定了这 $n-1$ 个对角元之后，剩下的一列中能够与这 $n-1$ 个对角元相乘（与这 $n-1$ 个对角元位于不同的行）的也只能是对角元，这说明 $n-1$ 次项也只能含于 n 个对角元的乘积 $(\lambda-a_{11})\cdots(\lambda-a_{nn})$，等于这个乘积展开式的 $n-1$ 次项 $-(a_{11}+\cdots+a_{nn})\lambda^{n-1}$，系数为 $-(a_{11}+\cdots+a_{nn})$.

笔记 这个在后续会学到，$f(\lambda)$ 称为方阵 $A=(a_{ij})_{n\times n}$ 的特征多项式，$a_{11}+\cdots+a_{nn}$ 称为 A 的迹 (trace)，如果你还想算一算这个多项式的常数项，你会发现它等于 $(-1)^n\det A$ (取 $\lambda=0$).

习题 2.3 (2022 秋期中)

计算 n 阶行列式 $\begin{vmatrix} x_1-a_1 & x_2 & \cdots & x_n \\ x_1 & x_2-a_2 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \cdots & x_n-a_n \end{vmatrix}$, 其中 $a_1,a_2,\cdots,a_n\neq 0$.

解 容易观察到行列式中每一列都含有一个向量 $(1,1,\cdots,1)^T$ 的常数倍，从而我们可以利用"加边法"求解。

$$\begin{vmatrix} x_1-a_1 & x_2 & \cdots & x_n \\ x_1 & x_2-a_2 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \cdots & x_n-a_n \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & x_1-a_1 & x_2 & \cdots & x_n \\ 1 & x_1 & x_2-a_2 & \cdots & x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_1 & x_2 & \cdots & x_n-a_n \end{vmatrix}$$

$$= \begin{vmatrix} 1 & -x_1 & -x_2 & \cdots & -x_n \\ 1 & -a_1 & 0 & \cdots & 0 \\ 1 & 0 & -a_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & -a_n \end{vmatrix} = \begin{vmatrix} 1-\sum_{i=1}^n \frac{x_i}{a_i} & -x_1 & -x_2 & \cdots & -x_n \\ 0 & -a_1 & 0 & \cdots & 0 \\ 0 & 0 & -a_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -a_n \end{vmatrix} = (-1)^n(1-\sum_{i=1}^n \frac{x_i}{a_i})\prod_{i=1}^n a_i$$

其中倒数第二个等号是因为 $a_1,a_2,\cdots,a_n \neq 0$，那么如果有某个 $a_i=0$ 呢？我们把上式结果换个形式写出来：

$$(-1)^n(1-\sum_{i=1}^n \frac{x_i}{a_i})\prod_{i=1}^n a_i = (-1)^n(\prod_{i=1}^n a_i - \sum_{i=1}^n (x_i \prod_{j\neq i} a_j))$$

我们发现 a_i 并不作分母，其实这本应就是对的，因为我们可以把行列式看作关于 $(a_1,a_2,\cdots,a_n)$ 的多元函数，把 $(x_1,x_2,\cdots,x_n)$ 视为参数，这个多元函数一定是关于 $(a_1,a_2,\cdots,a_n)$ 的多元多项式，故显然连续，那么它在某个 a_i 为 0 时一定有意义。这个方法叫做"微扰法"或者"摄动法"，在后续课程中经常用于某些证明题。一般的过程是先考虑矩阵可逆情形，再通过多项式的连续性推广到一般情形，是一个很重要的思想与方法。

笔记 这道题的结果十分重要，很多求解行列式的题目其实都是这道题的一些特殊情况，比如下面的两个题目：

习题 2.4 (2023 秋期中)

$$A=\begin{pmatrix} 4+x & 4 & 4 & 4 \\ 3 & 3+x & 3 & 3 \\ 2 & 2 & 2+x & 2 \\ 1 & 1 & 1 & 1+x \end{pmatrix}.x>0.\text{ 求 } \det A, A^{-1}.$$

解 （仅求解行列式）：A^T 符合习题2.3的形式，代入得到 $\det A = (1+\frac{10}{x})x^4 = x^4+10x^3$.

习题 2.5 (2015 秋期中)

设 n 阶方阵 $A=\begin{pmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & a \end{pmatrix}$，求：(i) $\det A$; (ii) $\operatorname{rank} A$.

解 （仅求解行列式）：$A=\begin{pmatrix} b-(b-a) & b & \cdots & b \\ b & b-(b-a) & \cdots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & b-(b-a) \end{pmatrix}$ 符合习题2.3的形式，代入得到 $\det A = (-1)^n(1-\frac{nb}{b-a})(b-a)^n = (a-b)^n + nb(a-b)^{n-1}$.

笔记 可以看到这两个题都是习题2.3的简化版，并且往年期中题里类似的题还有很多，就不再一一列举了。

从习题2.3可以看出“加边法”很好用，对吧，嘿嘿，那来看看下面这个题吧！

习题 2.6

已知 $n \geqslant 2, a_1, a_2, \cdots, a_n \neq 0$, 求 n 阶行列式:

$$\begin{vmatrix} 0 & a_1+a_2 & \cdots & a_1+a_n \\ a_2+a_1 & 0 & \cdots & a_2+a_n \\ \vdots & \vdots & & \vdots \\ a_n+a_1 & a_n+a_2 & \cdots & 0 \end{vmatrix}$$

解 这题目一看就很容易让人想到是加边对吧，我们发现每行除了对角元之外，都含有一个 $(a_1, \cdots, a_n)$ 的向量，同时也都含有一个 $(1, \cdots, 1)$ 的常数倍，但是无论我们用这两个里的哪个向量去加边，继续往下做还是无从下手，这个时候怎么办？那就全都要！记原行列式为 Δ，加两次边：

$$\Delta = \begin{vmatrix} 1 & 0 & 1 & 1 & \cdots & 1 \\ 0 & 1 & a_1 & a_2 & \cdots & a_n \\ 0 & 0 & 0 & a_1+a_2 & \cdots & a_1+a_n \\ 0 & 0 & a_2+a_1 & 0 & \cdots & a_2+a_n \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_n+a_1 & a_n+a_2 & \cdots & 0 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 & 1 & \cdots & 1 \\ 0 & 1 & a_1 & a_2 & \cdots & a_n \\ -a_1 & -1 & -2a_1 & 0 & \cdots & 0 \\ -a_2 & -1 & 0 & -2a_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ -a_n & -1 & 0 & 0 & \cdots & -2a_n \end{vmatrix}$$

这个时候是不是就好处理多了，不过稍微有点难算。

$$\Delta = \begin{vmatrix} 1-\dfrac{n}{2} & -\displaystyle\sum_{i=1}^{n}\frac{1}{2a_i} & 1 & 1 & \cdots & 1 \\ -\dfrac{1}{2}\displaystyle\sum_{i=1}^{n} a_i & 1-\dfrac{n}{2} & a_1 & a_2 & \cdots & a_n \\ 0 & 0 & -2a_1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & -2a_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & 0 & \cdots & -2a_n \end{vmatrix}$$

$$= \left[\left(1-\frac{n}{2}\right)^2 - \frac{1}{4}\left(\sum_{i=1}^{n} a_i\right)\left(\sum_{j=1}^{n}\frac{1}{a_j}\right)\right](-2)^n a_1 \cdots a_n$$

$$= (-1)^n 2^{n-2} a_1 \cdots a_n \left[(n-2)^2 - \left(\sum_{i=1}^{n} a_i\right)\left(\sum_{j=1}^{n}\frac{1}{a_j}\right)\right]$$

习题 2.7

证明：偶数阶斜对称方阵的行列式 $|A| = |a_{ij}|_{n\times n}$ 的所有元素的代数余子式 $A_{ij}\ (1 \leqslant i, j \leqslant n)$ 之和等于 0.

证明 将 A 扩充为 $n+1$ 阶斜对称方阵 $B = (b_{ij})_{n+1,n+1}$, 使它的第 1 行 $(b_{11}, \cdots, b_{1,n+1}) = (0, 1, \cdots, 1)$，第 1 列 $(b_{11}, \cdots, b_{n+1,1})^T = (0, -1, \cdots, -1)^T$，即 B 删去第 1 行和第 1 列得到 A，即

$$|B| = \begin{vmatrix} 0 & 1 & \cdots & 1 \\ -1 & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ -1 & a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

由于 B 是奇数阶斜对称方阵，$\det B = \det B^T = (-1)^n \det B = -\det B \Rightarrow \det B = 0$. 另一方面，将 $|B|$ 按第 1 行展开，得到如下等式:

$$0 = |B| = \sum_{j=0}^{n} b_{1,j+1}B_{1,j+1} = \sum_{j=1}^{n}(-1)^{1+j+1}N_{1,j+1}$$

其中 $N_{1,j+1}$ 是 B 的第 $(1,j+1)$ 元素的余子式 ,$B_{1,j+1} = (-1)^{1+j+1}N_{1,j+1}$ 则是相应的代数余子式。余子式 $N_{1,j+1}$ 由 A 的行列式 $|A|$ 删去第 j 列再在左边添上全由 -1 组成的一列得到。将每个余子式

$$N_{1,j+1} = \begin{vmatrix} -1 & a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ -1 & a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$

按第 1 列展开得到：

$$N_{1,j+1} = \sum_{i=1}^{n}(-1)(-1)^{i+1}(N_{1,j+1})_{i1}$$

其中 $(N_{1,j+1})_{i1}$ 是 $N_{1,j+1}$ 的第 $(i,1)$ 元素的余子式，由 $N_{1,j+1}$ 删去第 i 行和第 1 列得到，也就是由 $|A|$ 删去第 i 行和第 j 列得到的 $n-1$ 阶行列式即 $|A|$ 的第 (i,j) 元素的余子式 M_{ij}. 因此等式上式可改写为：

$$N_{1,j+1} = \sum_{i=1}^{n}(-1)^i M_{ij} \quad \text{从而有：} 0 = |B| = \sum_{j=1}^{n}\sum_{i=1}^{n}(-1)^j(-1)^i M_{ij}$$

其中 $(-1)^j(-1)^i M_{ij} = (-1)^{i+j}M_{ij}$ 就是 $|A|$ 的 (i,j) 元素的代数余子式 A_{ij}，从而命题成立。

笔记 这里有一个有用的小引理：奇数阶反对称方阵行列式为 0，所以我们通过加边把偶数阶升为奇数阶，剩下的操作就完全是代数变形了。

习题 2.8 (2019 秋期中)

计算 n 阶行列式 $\Delta_n = \begin{vmatrix} 1-a & -1 & & & \\ a & 1-a & -1 & & \\ & a & \ddots & \ddots & \\ & & \ddots & 1-a & -1 \\ & & & a & 1-a \end{vmatrix}$.

解 这种“三对角”方阵的行列式一般都是用二阶线性递推来做。(并且题上把它写成 Δ_n 也算是一种提示?)

$a \neq -1$ 时，沿第一行展开容易得到：

$$\Delta_n = (1-a)\Delta_{n-1} + a\Delta_{n-2}，\text{特征方程为 } x^2 + (a-1)x - a = 0，\text{解得 } x_1 = -a, x_2 = 1.$$

从而 $\Delta_n = A \times 1^n + B \times (-a)^n$，再由：

$$\Delta_1 = 1-a, \Delta_2 = (1-a)^2 + a = a^2 - a + 1，\text{解得} A = \frac{1}{a+1}, B = \frac{a}{a+1}.$$

从而 $\Delta_n = \dfrac{1}{a+1}(1-(-a)^{n+1})$. 而 $a = -1$ 时，利用习题2.3中的微扰法即可。

笔记 其实往年题做得多了就会发现，考试里求行列式的题用的方法基本就那么几种，像这种很经典的“三对角”矩阵，以及“爪型”矩阵、对角元相等的对称矩阵等等，老师的讲义中都详细地讲了。所以把老师这节的讲义吃透，考试时求行列式就没什么问题了。

行列式拆分的性质在一些习题中也经常会用到。当只有某一个元素不好处理时，就把这个元素改写，拆分成两个行列式的和分别处理，可能就会带来很大的方便，比如看下面的题目：

习题 2.9

计算 n 阶行列式：$\begin{vmatrix} x & a & a & \cdots & a \\ -a & x & a & \cdots & a \\ -a & -a & x & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a \\ -a & -a & \cdots & -a & x \end{vmatrix}$

解 观察发现我们可以把第一行加到其他行消去大部分的 $-a$，但第一列比较麻烦，那没关系，我们可以将第 1 行 $(x, a, \cdots, a)$ 拆分，记原行列式为 Δ_n，则有：

$$\Delta_n = D_1 + D_2 = \begin{vmatrix} a & a & a & \cdots & a \\ -a & x & a & \cdots & a \\ -a & -a & x & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a \\ -a & -a & \cdots & -a & x \end{vmatrix} + \begin{vmatrix} x-a & 0 & 0 & \cdots & 0 \\ -a & x & a & \cdots & a \\ -a & -a & x & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a \\ -a & -a & \cdots & -a & x \end{vmatrix}$$

将 D_1 的第 1 行加到以下各行，得到 $D_1 = a(x+a)^{n-1}$.

将 D_2 按第 1 行展开，得到 $D_2 = (x-a)\Delta_{n-1}$. 于是 $\Delta_n = a(x+a)^{n-1} + (x-a)\Delta_{n-1}$. 这时候用数列的递推解起来还是稍显麻烦，但观察到这个矩阵有很好的对称性，所以我们可以在 Δ_n 中用 $-a$ 替换 a, 这样得到的是 Δ_n 的转置的行列式，与原行列式相等，所以有：

$$\Delta_n = \Delta_n^T = -a(x-a)^{n-1} + (x+a)\Delta_{n-1}^T = -a(x-a)^{n-1} + (x+a)\Delta_{n-1}$$

$$\text{则有 } \Delta_n = a(x+a)^{n-1} + (x-a)\Delta_{n-1} = -a(x-a)^{n-1} + (x+a)\Delta_{n-1}$$

当 $a \neq 0$ 时可解出：

$$\Delta_{n-1} = \frac{a(x+a)^{n-1} + a(x-a)^{n-1}}{(x+a)-(x-a)} = \frac{a(x+a)^{n-1} + a(x-a)^{n-1}}{2a} = \frac{(x+a)^{n-1} + (x-a)^{n-1}}{2}$$

从而有：

$$\Delta_n = \frac{(x+a)^n + (x-a)^n}{2}$$

当 $a = 0$ 时易见 $\Delta_n = x^n = \frac{(x+0)^n + (x-0)^n}{2}$, 因此：

$$\Delta_n = \frac{(x+a)^n + (x-a)^n}{2} \text{ 对所有的 } a \text{ 成立}$$

习题 2.10 (2023 秋线性代数 (B2) 期中)

4 阶方阵 A 的第 3 行元素分别为 $-1, 0, 2, 8$，第 4 行元素对应的余子式依次是 $5, 10, a, 2$，则 $a =$________.

解 这题目乍一看非常简单啊，不就是剥蒜嘛，但实际上如果你把第 1,2,4 行的元素全设出来列方程硬求解是极其复杂的（9 个未知数），至少我去年在考场上是放弃了这个做法的。但实际上，只需要稍稍转变一下思路：

$$\text{记 } A = (a_{ij})_{4\times 4},\ b_{ij} = \begin{vmatrix} a_{1i} & a_{1j} \\ a_{2i} & a_{2j} \end{vmatrix},\ \text{则有} \begin{cases} -2b_{24} + 8b_{23} = 5 \\ 2b_{12} - b_{23} = 2 \\ -b_{24} + 8b_{12} = a \end{cases} \Rightarrow a = \frac{21}{2}$$

嗯，其实确实不难，虽然我去年没做出来（逃）。

另外老师讲义上提及的 Vandermonde 行列式也很重要，考试经常考。老师讲义上已经把好几种情况的处理办

法写得很详细了，我就不多写了（但后续还会有求这个矩阵的逆矩阵的题型）。不过我们可以通过 Vandermonde 行列式来求一些类似的矩阵的行列式。

习题 2.11

计算行列式：$\begin{vmatrix} 1+x_1 & 1+x_1^2 & \cdots & 1+x_1^n \\ 1+x_2 & 1+x_2^2 & \cdots & 1+x_2^n \\ \vdots & \vdots & \ddots & \vdots \\ 1+x_n & 1+x_n^2 & \cdots & 1+x_n^n \end{vmatrix}$

解

$$\begin{vmatrix} 1+x_1 & 1+x_1^2 & \cdots & 1+x_1^n \\ 1+x_2 & 1+x_2^2 & \cdots & 1+x_2^n \\ \vdots & \vdots & \ddots & \vdots \\ 1+x_n & 1+x_n^2 & \cdots & 1+x_n^n \end{vmatrix} = \begin{vmatrix} 1 & & & & \\ 1 & 1+x_1 & 1+x_1^2 & \cdots & 1+x_1^n \\ 1 & 1+x_2 & 1+x_2^2 & \cdots & 1+x_2^n \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & 1+x_n & 1+x_n^2 & \cdots & 1+x_n^n \end{vmatrix} = \begin{vmatrix} 1 & -1 & -1 & -1 & -1 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix}$$

$$= -\begin{vmatrix} -1 & 1 & 1 & 1 & 1 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix} = -\begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix} + \begin{vmatrix} 2 & 1 & 1 & 1 & 1 \\ & x_1 & x_1^2 & \cdots & x_1^n \\ & x_2 & x_2^2 & \cdots & x_2^n \\ & \vdots & \vdots & \ddots & \vdots \\ & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix}$$

记 $x_0=1$，则由 Vandermonde 行列式可知：

$$\text{原式} = 2\begin{vmatrix} x_1 & x_1^2 & \cdots & x_1^n \\ x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & & & \vdots \\ x_n & x_n^2 & \cdots & x_n^n \end{vmatrix} - \prod_{0\leqslant i<j\leqslant n}(x_j-x_i) = 2\prod_{i=1}^{n}x_i \cdot \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & & & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix} - \prod_{0\leqslant i<j\leqslant n}(x_j-x_i)$$

$$= 2\prod_{i=1}^{n}x_i\left(\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)\right) - \prod_{0\leqslant j<i\leqslant n}(x_i-x_j) = \left(2\prod_{i=1}^{n}x_i - \prod_{i=1}^{n}(x_i-1)\right)\left(\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)\right).$$

老师课上介绍过的 Laplace 展开在求解行列式的问题中也有着很广泛的应用，不过 Laplace 展开不是我们这门课程的考核要求内容，仅作补充了解即可，该定理的内容如下：

定理 2.1 (Laplace 展开定理)

取定行指标 $i_1,i_2,\cdots,i_p, 1\leqslant i_1<i_2<\cdots<i_p\leqslant n$. 遍取行列式 $\det A$ 中第 $i_1,i_2,\cdots,i_p$ 行上的 p 阶子式，并分别乘以相应的代数余子式，其和即为 $\det A$。具体地说，有：

$$\det A = \sum_{1\leqslant j_1<j_2<\cdots<j_p\leqslant n} A\begin{pmatrix} i_1\ i_2\ \cdots\ i_p \\ j_1\ j_2\ \cdots\ j_p \end{pmatrix}\left((-1)^{i_1+i_2+\cdots+i_p+j_1+j_2+\cdots+j_p} A\begin{pmatrix} i_{p+1}\ \cdots\ i_n \\ j_{p+1}\ \cdots\ j_n \end{pmatrix}\right),$$

其中$i_1i_2\cdots i_pi_{p+1}\cdots i_n$ 和$j_1j_2\cdots j_pj_{p+1}\cdots j_n$ 都是$1,2,\cdots,n$ 的排列, 并且$1\leqslant i_{p+1}<\cdots<i_n\leqslant n, 1\leqslant j_{p+1}<\cdots<j_n\leqslant n$.

笔记 那么，思考一下，Laplace 展开定理有没有什么几何解释呢？当然是有的，如果说行列式的递推定义是 n 维空间到 $n-1$ 维空间的投影再求和，那么 Laplace 展开就是 n 维空间到 $n-p$ 维空间的投影再求和。这里不再过多叙述了。关于定理的严格证明，老师的讲义中有所提及，感兴趣的同学可以自行学习。

以下给出两个 Laplace 定理在解题中的应用：

习题 2.12

利用 Laplace 展开定理计算行列式：$\begin{vmatrix} \lambda & 0 & 0 & \cdots & 0 & a \\ x_1 & c & b & \cdots & b & y_1 \\ x_2 & b & c & \cdots & b & y_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n & b & b & \cdots & c & y_n \\ a & 0 & 0 & \cdots & 0 & \lambda \end{vmatrix}$

解 记原行列式为 $\det A$，观察发现这个行列式最外面一圈和里面格格不入，所以将行列式按第 1 列、第 $n+2$ 列作 Laplace 展开，得到：

$$
\begin{aligned}
\det A &= \sum_{1\leqslant i_1<i_2\leqslant n+2} A\begin{pmatrix} i_1 & i_2 \\ 1 & n+2 \end{pmatrix}\left((-1)^{i_1+i_2+1+(n+2)} A\begin{pmatrix} i_3 & i_4 & \cdots & i_{n+2} \\ 2 & 3 & \cdots & n+1 \end{pmatrix}\right) \\
&= (-1)^{1+(n+2)+1+(n+2)} \begin{vmatrix} \lambda & a \\ a & \lambda \end{vmatrix} \cdot \begin{vmatrix} c & b & \cdots & b \\ b & c & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & c \end{vmatrix} \\
&= \left(\lambda^2 - a^2\right)\left(c+(n-1)b\right)(c-b)^{n-1}.
\end{aligned}
$$

其中倒数第二个等号是因为包含第 1 行或最后一行的子式均为 0，最后一个等号用到了习题2.5的结论。

习题 2.13

计算 $2n$ 阶行列式

$$
\begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1n} & b_{11} & \cdots & b_{1,n-1} & b_{1n} \\
 & a_{22} & \cdots & a_{2n} & b_{21} & \cdots & b_{2,n-1} & \\
 & & \ddots & \vdots & \vdots & \iddots & & \\
 & & & a_{nn} & b_{n1} & & & \\
 & & & c_{1n} & d_{11} & & & \\
 & & \iddots & \vdots & \vdots & \ddots & & \\
 & c_{n-1,2} & \cdots & c_{n-1,n} & d_{n-1,1} & \cdots & d_{n-1,n-1} & \\
c_{n1} & c_{n2} & \cdots & c_{nn} & d_{n1} & \cdots & d_{n,n-1} & d_{nn}
\end{vmatrix}
$$

其中未写出的元素都是零。

解 不断对行列式第一列和最后一列进行 Laplace 展开，我们有：

$$
\begin{aligned}
\text{原式} &= (a_{11}d_{nn} - b_{1n}c_{n1}) \begin{vmatrix}
a_{22} & \cdots & a_{2n} & b_{21} & \cdots & b_{2,n-1} \\
 & \ddots & \vdots & \vdots & \iddots & \\
 & & a_{nn} & b_{n1} & & \\
 & & c_{1n} & d_{11} & & \\
 & \iddots & \vdots & \vdots & \ddots & \\
c_{n-1,2} & \cdots & c_{n-1,n} & d_{n-1,1} & \cdots & d_{n-1,n-1}
\end{vmatrix} \\
&= \cdots = \prod_{k=1}^{n} (a_{kk}d_{n+1-k,n+1-k} - b_{k,n+1-k}c_{n+1-k,k})
\end{aligned}
$$

2.3 第三次习题课

2.3.1 基础矩阵与标准单位向量

定义 2.21

n 维标准列向量是指以下 n 个 n 维列向量

$$e_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},e_2=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\cdots,e_n=\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}.$$

♣

定义 2.22

n 阶基础矩阵是指 n^2 个 n 阶矩阵 $\{E_{ij}(i,j=1,2,\cdots,n)\}$. 其中 $E_{ij}\in\mathbb{F}^{n\times n}$, 它的第 (i,j) 元素是 1，其他元素为 0.

♣

性质 (1) $e_i^Te_j=0, e_i^Te_i=1$. 其中 $i\neq j$;

(2）若 $A\in\mathbb{F}^{m\times n}$, 则 Ae_i 为 A 的第 i 个列向量；e_i^TA 为 A 的第 i 个行向量;

(3) 若 $A\in\mathbb{F}^{m\times n}$, 则 $e_i^TAe_j=a_{ij}$;

(4)$E_{ij}E_{kl}=\delta_{jk}E_{il}$;

(5) 若 $A\in\mathbb{F}^{n\times n}$, 则 $E_{ij}A$ 将 A 的第 j 行变成第 i 行，其他元素变为 0;

(6) 若 $A\in\mathbb{F}^{n\times n}$, 则 AE_{ij} 将 A 的第 i 列变成第 j 列，其他元素变为 0.

习题 2.14

设

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

证明：

$$A^k=\begin{pmatrix}0&I_{n-k}\\I_k&0\end{pmatrix}.$$

♠

证明 将矩阵 A 写成 $A=(e_n,e_1,\cdots,e_{n-1})$. 由分块矩阵的乘法及性质 (2), 有 $A^2=(Ae_n,Ae_1,Ae_2\cdots,Ae_{n-1})=(e_{n-1},e_n,e_1,\cdots,e_{n-2})$. 如此下去便可得到结论.

习题 2.15

具有以下形状的矩阵称为循环矩阵

$$\begin{pmatrix}a_1&a_2&a_3&\cdots&a_n\\a_n&a_1&a_2&\cdots&a_{n-1}\\a_{n-1}&a_n&a_1&\cdots&a_{n-2}\\\cdots&\cdots&\cdots&&\cdots\\a_2&a_3&a_4&\cdots&a_1\end{pmatrix}.$$

证明：同阶循环矩阵的乘积仍然是循环矩阵.

证明 设

$$J = \begin{pmatrix} 0 & I_{n-1} \\ 1 & 0 \end{pmatrix}.$$

注意到任意循环矩阵 A 可以表示为 $A = a_1 I_n + a_2 J + a_3 J^2 + \cdots + a_n J^{n-1}$ 的形式，反之有如前形式的矩阵也一定是循环矩阵. 两个循环矩阵的乘积可以写成关于 J 的两个多项式的乘积，又 $J^n = I_n$, 即可得到结论.

习题 2.16

设 A 是 n 阶上三角矩阵且主对角线上元素全为零，证明：$A^n = 0$.

证明 可以设

$$A = \sum_{i<j} a_{ij} E_{ij}.$$

当 $j \neq k$ 时，$E_{ij}E_{kl} = 0$, 因此在 A^n 的乘法展开式中，可能的非 0 项只能具有形状 $E_{ij_1}E_{j_1j_2}\cdots E_{j_{n-1}j_n}$，且 $1 \le i < j_1 < j_2 < \cdots < j_n \le n$. 显然这是不可能的，因此 $A^n = 0$.

2.3.2 迹及其应用

定义 2.23

设 $A \in \mathbb{F}^{n\times n}$，则 A 上主对角线元素之和称为矩阵 A 的迹，记为 $\mathrm{tr}A$.

性质 若 $A, B \in \mathbb{F}^{n\times n}, k \in \mathbb{F}$，则

(1)$\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$;

(2)$\mathrm{tr}(kA) = k(\mathrm{tr}A)$;

(3)$\mathrm{tr}A^T = \mathrm{tr}A$;

(4)$\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

命题 2.9 (迹的等价刻画)

若映射 $f : \mathbb{F}^{n\times n} \to \mathbb{F}$, 对任意 $A, B \in \mathbb{F}^{n\times n}, k \in \mathbb{F}$，满足

(1)$f(A+B) = f(A) + f(B)$;

(2)$f(kA) = kf(A)$;

(3)$f(AB) = f(BA)$;

(4)$f(I_n) = n$.

则 f 是迹.

证明 由 f 的线性性知，$f(E_{11}) + f(E_{22}) + \cdots + f(E_{nn}) = f(E_{11} + E_{22} + \cdots + E_{nn}) = f(I_n) = n$.
又 $f(E_{ii}) = f(E_{ij}E_{ji}) = f(E_{ji}E_{ij}) = f(E_{jj})$, 故 $f(E_{ii}) = 1$. 当 $i \neq j$ 时，$E_{ij} = E_{i1}E_{1j}$, 则 $f(E_{ij}) = f(E_{i1}E_{1j}) = f(E_{1j}E_{i1}) = f(0) = f(0I_n) = 0f(I_n) = 0$. 设 $A = (a_{ij})$, 则 $f(A) = f(\sum\limits_{i,j=1}^{n} a_{ij}E_{ij}) = \sum\limits_{i,j=1}^{n} a_{ij}f(E_{ij}) = \sum\limits_{i=1}^{n} a_{ii} = \mathrm{tr}A$. 即映射 f 就是迹.

习题 2.17

不存在矩阵 $A, B \in \mathbb{F}^{n\times n}$, 使得 $AB - BA = kI_n$, 其中 $k \neq 0$.

证明 若存在这样的矩阵 A, B, 使得 $AB - BA = kI_n$，且 $k \neq 0$. 同时取迹，有 $0 = \operatorname{tr}(AB) - \operatorname{tr}(BA) = \operatorname{tr}(AB - BA) = \operatorname{tr}(kI_n) = kn$, 矛盾！

习题 2.18

设 $A \in \mathbb{R}^{n\times n}$, 证明：$\operatorname{tr}(AA^T) \geq 0$, 等号成立当且仅当 $A = 0$.

证明 设 $A = (a_{ij})$, 经计算易得 $\operatorname{tr}(AA^T) = \sum\limits_{i,j=1}^{n} a_{ij}^2 \geq 0$. 等号成立当且仅当 $a_{ij} = 0, \forall i, j$, 即 $A = 0$.

习题 2.19

设 $A \in \mathbb{R}^{n\times n}$, 满足 $AA^T = A^2$, 证明：A 是对称矩阵.

证明 只需证明 $A - A^T = 0$. 由上题知只需证明 $\operatorname{tr}((A - A^T)(A - A^T)^T) = 0$ 即可. 由 $AA^T = A^2$, 知 $A^TA = (A^T)^2$. $\operatorname{tr}((A - A^T)(A - A^T)^T) = \operatorname{tr}((A - A^T)(A^T - A)) = \operatorname{tr}(AA^T - A^2 - (A^T)^2 + A^TA) = \operatorname{tr}(A^TA - AA^T) = \operatorname{tr}(A^TA) - \operatorname{tr}(AA^T) = 0$, 从而结论得证.

2.3.3 降阶公式及其应用

命题 2.10

若 $A \in \mathbb{F}^{m\times m}, B \in \mathbb{F}^{m\times n}, C \in \mathbb{F}^{n\times m}, D \in \mathbb{F}^{n\times n}$, 其中 A 可逆，则

$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \cdot \det\big(D - CA^{-1}B\big).$$

证明 注意到

$$\begin{pmatrix} I_m & 0 \\ -CA^{-1} & I_n \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}$$

上述等式两边同时取行列式即可得到结论.

注 当 D 可逆时，有

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} I_m & 0 \\ -D^{-1}C & I_n \end{pmatrix} = \begin{pmatrix} A - BD^{-1}C & B \\ 0 & D \end{pmatrix}.$$

故

$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det D \cdot \det\big(A - BD^{-1}C\big).$$

于是，当矩阵 A, D 同时可逆时，有如下等式

$$\det A \cdot \det\big(D - CA^{-1}B\big) = \det D \cdot \det\big(A - BD^{-1}C\big). \tag{2.44}$$

以上等式我们称为**降阶公式**.

推论 2.2

若 $A \in \mathbb{F}^{n\times m}, B \in \mathbb{F}^{m\times n}$，且 $n \geq m$, 则

$$\det\{(\lambda I_n - AB)\} = \lambda^{n-m}\det\{(\lambda I_m - BA)\}.$$

证明 (法一) 当 $\lambda \neq 0$ 时，考虑矩阵

$$\begin{pmatrix} \lambda I_m & B \\ A & I_n \end{pmatrix}$$

对上述矩阵作同命题2.10及其注记中的处理.

当 $\lambda = 0$ 时，分别考虑 $m = n$ 和 $n > m$ 的情况.

(法二) 先考虑 $m = n$ 的情形，若 A 可逆，由于 $BA = A^{-1}(AB)A$，因此 AB 与 BA 相似，他们的特征多项式相同. 对于一般矩阵 A，可以取一列有理数列 $t_k \to 0$，使得矩阵 $t_k I + A$ 可逆，于是由可逆情形有 $\det\{(\lambda I_n - (t_k I + A)B)\} = \det\{(\lambda I_n - B(t_k I + A))\}$，两边同时取极限，有 $\det\{(\lambda I_n - AB)\} = \det\{(\lambda I_n - BA)\}$.

考虑 $n > m$ 的情形，考虑方阵 $C = (A, 0)$ 与 $D = \begin{pmatrix} B \\ 0 \end{pmatrix}$. 注意到 $CD = AB$,$DC = \begin{pmatrix} BA & 0 \\ 0 & 0 \end{pmatrix}$. 由方阵情形有

$$\det\{(\lambda I_n - AB)\} = \det\{(\lambda I_n - CD)\} = \det\{(\lambda I_n - DC)\} = \begin{vmatrix} \lambda I_m - BA & 0 \\ 0 & \lambda I_{n-m} \end{vmatrix} = \lambda^{n-m} \det\{(\lambda I_m - BA)\}.$$

习题 2.20 (小测第四题)

计算 n 阶行列式

$$|A| = \begin{vmatrix} 0 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 1 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & 1 & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{vmatrix}.$$

解 注意到

$$A = -I_n + \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} (1, 1, \cdots, 1).$$

由推论2.2知

$$\det\{A\} = \det\left\{(-I_n + \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} (1, 1, \cdots, 1))\right\} = (-1)^{n-1} \det\left\{(-1 + (1, 1, \cdots, 1) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix})\right\} = (-1)^{n-1}(n-1).$$

习题 2.21 (第五周作业)

计算下列 n 阶行列式

$$\begin{vmatrix} 1 + a_1 & 1 & \cdots & 1 \\ 1 & 1 + a_2 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 + a_n \end{vmatrix}.$$

其中 $a_i \neq 0, i = 1, 2, \cdots, n$.

解

$$\begin{vmatrix} 1+a_1 & 1 & \cdots & 1 \\ 1 & 1+a_2 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1+a_n \end{vmatrix} = \det\{1\}\cdot\det\left\{(\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} 1^{-1}(1,1,\cdots,1))\right\}$$

$$= \det\left\{\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}\right\}\cdot\det\left\{(1+(1,1,\cdots,1)\begin{pmatrix} \frac{1}{a_1} & & & \\ & \frac{1}{a_2} & & \\ & & \ddots & \\ & & & \frac{1}{a_n} \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix})\right\} = (\prod_{i=1}^n a_i)(1+\sum_{j=1}^n \frac{1}{a_j})$$

$= \prod\limits_{i=1}^n a_i + \sum\limits_{j=1}^n \prod\limits_{k\neq j} a_k.$

注 后面将处理存在 $a_i = 0$ 的情形.

2.3.4 利用矩阵乘法计算行列式

习题 2.22

设 $s_k = x_1^k + x_2^k + \cdots + x_n^k (k \geq 0), s_0 = n$

$$S = \begin{pmatrix} s_0 & s_1 & s_2 & \cdots & s_{n-1} \\ s_1 & s_2 & s_3 & \cdots & s_n \\ s_2 & s_3 & s_4 & \cdots & s_{n+1} \\ \vdots & \vdots & \vdots & & \vdots \\ s_{n-1} & s_n & s_{n+1} & \cdots & s_{2n-2} \end{pmatrix}.$$

求 $\det\{S\}$, 并证明当 $x_i \in \mathbb{R}$ 时，有 $\det\{S\} \geq 0$.

解 设

$$V = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{pmatrix}.$$

则 $S = VV^T$, 因此 $\det\{S\} = (\det\{V\})^2 = \prod\limits_{1\leq i<j\leq n}(x_j - x_i)^2 \geq 0$.

习题 2.23

计算下列矩阵 A 的行列式

$$A = \begin{pmatrix} x & y & -z & w \\ y & -x & -w & -z \\ z & -w & x & y \\ w & z & y & -x \end{pmatrix}.$$

解 注意到 $AA^T = \operatorname{diag}(u,u,u,u)$, 其中 $u = x^2+y^2+z^2+w^2$. 于是

$$(\det\{A\})^2 = (x^2+y^2+z^2+w^2)^4.$$

因此 $\det\{A\} = (x^2+y^2+z^2+w^2)^2$ 或者 $\det\{A\} = -(x^2+y^2+z^2+w^2)^2$，带入 $x=1, y=z=w=0$, 有

$\det\{A\}=1$, 故 $\det\{A\}=(x^2+y^2+z^2+w^2)^2$.

2.3.5 摄动法及其应用

命题 2.11

若 A 是一个 n 阶方阵，则存在一个正数 a, 使得对任意的 $0<t<a$，矩阵 tI_n+A 总是可逆矩阵.

证明 由行列式的展开式，可以假设

$$\det\{(tI_n+A)\}=t^n+a_1t^{n-1}+\cdots+a_{n-1}t+a_n.$$

这是一个关于 t 的 n 次多项式，其上至多有 n 个不同的根. 若上述多项式的根均为零，则可取 $a=1$; 若上述多项式有非零根，则可取 a 为上述多项式模长的最小值. 无论上述哪种情况，当 $0<t_0<a$ 时，t_0 都不会是上述多项式的根，因此 $\det\{(t_0I_n+A)\}\neq 0$, 即矩阵 t_0I_n+A 可逆.

注 这个命题告诉我们对任意的 n 阶矩阵 A, 经过微小的一维摄动后，tI_n+A 总能成为一个可逆矩阵.

笔记 摄动法原理: 设 A 是 n 阶矩阵，由上面命题知，存在一列有理数 $t_k\to 0$，使得 t_kI_n+A 都是可逆矩阵. 如果一个矩阵问题对可逆矩阵成立，特别地对 t_kI_n+A 成立，并且该问题关于 t_k 连续，则可让 $t_k\to 0$，最后得到该问题对一般的方阵 A 也成立. 需要注意的是：摄动法处理矩阵问题时一定要关于 t_k 连续. 这一点非常重要，否则我们将不能用摄动法来归结处理. 一般而言，运用摄动法分为两步：首先处理可逆矩阵情形; 其次再利用摄动以及取极限得到一般情况的证明. 接下来，我们来看一个具体的例子.

习题 2.24

设 $A,B,C,D\in\mathbb{F}^{n\times n}$ 且 $AC=CA$. 证明：

$$\det\left\{\begin{pmatrix}A & B\\ C & D\end{pmatrix}\right\}=\det\{(AD-CB)\}.$$

证明 若矩阵 A 可逆，由命题2.10及矩阵 A,C 的交换性知

$$\det\begin{pmatrix}A & B\\ C & D\end{pmatrix}=\det A\cdot\det\big(D-CA^{-1}B\big)=\det\big\{(AD-ACA^{-1}B)\big\}=\det\{(AD-CB)\}.$$

对于一般的方阵 A，可以取到一列有理数 $t_k\to 0$，使得 t_kI_n+A 是可逆矩阵，且 $(t_kI_n+A)C=C(t_kI_n+A)$. 由前面分析知

$$\det\left\{\begin{pmatrix}t_kI_n+A & B\\ C & D\end{pmatrix}\right\}=\det\{((t_kI_n+A)D-CB)\}.$$

注意到上述等式两边均为关于 t_k 的多项式，从而关于 t_k 连续 (适合摄动法使用条件). 上述等式两边同时取极限，令 $t_k\to 0$，即有

$$\det\left\{\begin{pmatrix}A & B\\ C & D\end{pmatrix}\right\}=\det\{(AD-CB)\}.\quad \text{结论得证.}$$

2.3.6 运用多项式处理行列式

命题 2.12

设多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0.$$

若 $f(x)$ 有 $n+1$ 个不同的根 $b_1,b_2,\cdots,b_{n+1}$, 即 $f(b_1)=f(b_2)=\cdots=f(b_{n+1})=0$, 则 $f(x)$ 是零多项式.

证明 由假设 $x_0=a_0,x_1=a_1,\cdots,x_{n-1}=a_{n-1},x_n=a_n$ 是下列方程组的解

$$\begin{cases} x_0+b_1x_1+\cdots+b_1^{n-1}x_{n-1}+b_1^nx_n=0\\ x_0+b_2x_1+\cdots+b_2^{n-1}x_{n-1}+b_2^nx_n=0\\ \cdots\cdots\\ x_0+b_{n+1}x_1+\cdots+b_{n+1}^{n-1}x_{n-1}+b_{n+1}^nx_n=0 \end{cases}$$

上述线性方程组的系数是一个 $Vander\,Monde$ 行列式，又 $b_1,b_2,\cdots,b_{n+1}$ 互不相同，所以系数行列式不为零. 由 $Cramer$ 法则知上述方程组只有零解，即 $a_n=a_{n-1}=\cdots=a_1=a_0=0$, 故 $f(x)$ 是零多项式.

习题 2.25

设 $f_k(x)(k=1,2,\cdots,n)$ 是次数不超过 $n-2$ 的多项式，证明：对任意 n 个数 $a_1,a_2,\cdots,a_n$ 均有

$$\begin{vmatrix} f_1(a_1) & f_2(a_1) & \cdots & f_n(a_1)\\ f_1(a_2) & f_2(a_2) & \cdots & f_n(a_2)\\ \vdots & \vdots & & \vdots\\ f_1(a_n) & f_2(a_n) & \cdots & f_n(a_n) \end{vmatrix}=0.$$

证明 考虑如下多项式函数

$$g(x)\triangleq\begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x)\\ f_1(a_2) & f_2(a_2) & \cdots & f_n(a_2)\\ \vdots & \vdots & & \vdots\\ f_1(a_n) & f_2(a_n) & \cdots & f_n(a_n) \end{vmatrix}=0.$$

若 $a_i(i=2,\cdots,n)$ 中有相同者，则显然有 $g(x)=0$. 若诸 a_i 互不相同，则由 $g(a_i)=0(i=2,\cdots,n)$, 知 $g(x)$ 有 $n-1$ 个互不相同的根，由命题2.12，知 $g(x)=0$. 综上，无论何种情况，总有 $g(x)=0$，特别 $g(a_1)=0$, 因此原行列式值等于零.

习题 2.26

计算下列 n 阶行列式

$$\begin{vmatrix} 1+a_1 & 1 & \cdots & 1\\ 1 & 1+a_2 & \cdots & 1\\ \vdots & \vdots & & \vdots\\ 1 & 1 & \cdots & 1+a_n \end{vmatrix}.$$

解 考虑如下多项式函数

$$f(x)\triangleq\begin{vmatrix} 1+a_1+x & 1 & \cdots & 1\\ 1 & 1+a_2+x & \cdots & 1\\ \vdots & \vdots & & \vdots\\ 1 & 1 & \cdots & 1+a_n+x \end{vmatrix}.$$

存在自然数 N，当 $x>N$ 时，$a_i+x(i=1,2,\cdots,n)$ 全不为零. 于是由习题2.21知，当 $x>N$ 时 $f(x)=\prod\limits_{i=1}^n(a_i+x)+\sum\limits_{j=1}^n\prod_{k\neq j}(a_k+x)$. 再由命题2.12知 $f(x)=\prod\limits_{i=1}^n(a_i+x)+\sum\limits_{j=1}^n\prod_{k\neq j}(a_k+x)$ 对任意 x 成立. 特别对 $x=0$ 成立. 因此原行列式 $=\prod\limits_{i=1}^n a_i+\sum\limits_{j=1}^n\prod\limits_{k\neq j}a_k$.

2.4 第四次习题课

2.4.1 相关知识补充

分块矩阵初等变换

¶ 初等变换内容

定义 2.24 (分块矩阵的初等行变换)

(1) 把一个块行的左 P 倍加到另一个块行上;

(2) 互换两个块行位置;

(3) 用一个可逆矩阵左乘某一块行. ♣

定义 2.25 (分块矩阵的初等列变换)

(1) 把一个块列的右 P 倍加到另一个块列上;

(2) 互换两个块列位置;

(3) 用一个可逆矩阵右乘某一块列.

可见，这与矩阵的初等变换十分类似，实际上，分块矩阵的初等变换也可以由分块初等矩阵表示. ♣

定义 2.26 (分块初等矩阵)

单位矩阵分块得到的矩阵经一次分块矩阵初等行 (列) 变换得到的矩阵称为分块初等矩阵，共有以下三种:

$$(1)\quad T_{ij}(P)=\begin{pmatrix} I_{11} & & & & & & \\ & \ddots & & & & & \\ & & I_{ii} & & P & & \\ & & & \ddots & & & \\ & & & & I_{jj} & & \\ & & & & & \ddots & \\ & & & & & & I_{nn} \end{pmatrix}.$$

将分块矩阵 A 左乘 $T_{ij}(P)$ 表示将 A 的第 j 行的矩阵左乘 P 加到第 i 行; 右乘 $T_{ij}(P)$ 表示将 A 的第 i 列的矩阵右乘 P 加到第 j 列.

$$例:\begin{pmatrix} I & \\ P & I \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}=\begin{pmatrix} A & B \\ PA+C & PB+D \end{pmatrix};\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} I & Q \\ & I \end{pmatrix}=\begin{pmatrix} A & B+AQ \\ C & D+CQ \end{pmatrix}$$

特别地，当 A 可逆时取 $P=-CA^{-1}, Q=-A^{-1}B$ 即得 Schur 公式:

$$\begin{pmatrix} I & \\ -CA^{-1} & I \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}=\begin{pmatrix} A & B \\ & D-CA^{-1}B \end{pmatrix};\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} I & -A^{-1}B \\ & I \end{pmatrix}=\begin{pmatrix} A & \\ C & D-CA^{-1}B \end{pmatrix}$$

$$(2)\quad S_{ij}=\begin{pmatrix} I_{11} & & & & & & \\ & \ddots & & & & & \\ & & O_{ii} & & I_{ij} & & \\ & & & \ddots & & & \\ & & I_{ji} & & O_{jj} & & \\ & & & & & \ddots & \\ & & & & & & I_{nn} \end{pmatrix}$$

S_{ij} 左 (右) 作用分别对应分块矩阵 i,j 块行 (列) 互换.

$$(3)\quad D_i(P)=\begin{pmatrix} I_{11} & & & & & & \\ & \ddots & & & & & \\ & & I_{i-1,i-1} & & & & \\ & & & P_{ii} & & & \\ & & & & I_{i+1,i+1} & & \\ & & & & & \ddots & \\ & & & & & & I_{nn} \end{pmatrix}$$

$D_i(P)$ 左 (右) 作用分别对应分块矩阵第 i 块行 (列) 左 (右) 乘 P.

¶ 性质

- 三种分块初等矩阵都可以由多个初等方阵相乘得到.
  推论：分块矩阵初等变换就是一次进行了多个矩阵初等变换.
- $T_{ij}(P)^{-1}=T_{ij}(-P)$，$S_{ij}^{-1}=S_{ij}$，$D_i(P)^{-1}=D_i(P^{-1})$，这也是初等变换 (3) 要求 P 可逆的原因.
- $det(T_{ij}(P))=1$，$det(D_i(P))=det(P)$，$det(S_{ij})=(-1)^{(r+t+s-1)r+(s+t-1)s}$，其中 r 为 S_{ij} 的第 i 块行的行数，t 为第 $(i+1)$ 块行到第 $(j-1)$ 块行的总行数，s 为第 j 块行的行数.
  证明：前两个行列式由 Laplace 展开易得；求第三个行列式即求交换两行的次数，把 I_{ij} 的各行从上到下依次换到 I_{ji} 下方，进行了 $(r+t+s-1)r$ 次对换，再将 I_{ji} 的各行从下到上依次换到 $(i+1)$ 块行上方，进行了 $(s+t-1)s$ 次对换.

¶ 应用

(1) 求行列式

例题 2.16 第六次作业 1、4 题

解 以第 4 题为例，$\lambda\neq 0$ 时有

$$\begin{vmatrix}\lambda I_n & -A\\ -B & I_m\end{vmatrix}=\begin{vmatrix}\lambda I_n-AB & \\ -B & I_m\end{vmatrix}=det(\lambda I_n-AB) \tag{2.45}$$

上式是对分块矩阵做了行变换，下面再做列变换：

$$\begin{vmatrix}\lambda I_n & -A\\ -B & I_m\end{vmatrix}=\begin{vmatrix}\lambda I_n & \\ -B & I_m-\frac{BA}{\lambda}\end{vmatrix}=det(\lambda I_n)det(I_m-\frac{BA}{\lambda})=\lambda^{n-m}det(\lambda I_m-BA) \tag{2.46}$$

结合式 (2.45)(2.46) 知命题成立.

注 分块矩阵行列式进行初等变换 (2)(3) 时，并非直接相等，而是相差一个倍数，所以更不容易出错的办法是先将矩阵相乘的等式写出来，再对等式两边求行列式. 用分块初等变换求行列式的目标是将行列式变为准三角阵，之后就可以利用 Laplace 展开进行降阶.

(2) 求逆矩阵：用于分块性质良好的矩阵，步骤与一般初等变换类似.

例题 2.17(2018-2019 第一学期期中 3.) 求矩阵 $A=\begin{pmatrix}1&0&1&2\\0&1&2&3\\1&1&0&0\\0&1&0&0\end{pmatrix}$ 的逆矩阵

解 设 $B=\begin{pmatrix}1&2\\2&3\end{pmatrix},C=\begin{pmatrix}1&1\\0&1\end{pmatrix}$ 易知 $B^{-1}=\begin{pmatrix}-3&2\\2&-1\end{pmatrix},C^{-1}=\begin{pmatrix}1&-1\\0&1\end{pmatrix}$ 则

$$\begin{aligned}&\begin{pmatrix}I&B&I&O\\C&O&O&I\end{pmatrix}\to\begin{pmatrix}I&B&I&O\\O&-CB&-C&I\end{pmatrix}\\&\to\begin{pmatrix}I&O&O&C^{-1}\\O&-CB&-C&I\end{pmatrix}\to\begin{pmatrix}I&O&O&C^{-1}\\O&I&B^{-1}&-B^{-1}C^{-1}\end{pmatrix}\end{aligned}\tag{2.47}$$

故

$$A^{-1}=\begin{pmatrix}O&C^{-1}\\B^{-1}&-B^{-1}C^{-1}\end{pmatrix}=\begin{pmatrix}0&0&1&-1\\0&0&0&1\\-3&2&3&-5\\2&-1&-2&3\end{pmatrix}\tag{2.48}$$

(3) 求矩阵的秩/相抵标准型

由于分块初等变换就是多个初等变换的乘积，故分块初等变换也不改变分块矩阵的秩.

例题 2.18（课本习题三第 42 题） 这是还没截止的作业题，暂不提供解答，可参考后续的作业答案.

例题 2.19（2022-2023 第二学期期中 6.） 设 A,B,C 为 n 阶实方阵，证明：存在矩阵 X,Y 满足 $XA-BY=C\iff$ 分块矩阵 $\begin{pmatrix}A&O\\C&B\end{pmatrix}$ 与 $\begin{pmatrix}A&O\\O&B\end{pmatrix}$ 相抵

证明 (1)$\Rightarrow$

$$\begin{pmatrix}A&O\\O&B\end{pmatrix}\to\begin{pmatrix}A&O\\XA-BY&B\end{pmatrix}\to\begin{pmatrix}A&O\\C&B\end{pmatrix}\tag{2.49}$$

即两矩阵相抵.

(2)$\Leftarrow$

设 $A=P_1\begin{pmatrix}I_r&O\\O&O\end{pmatrix}Q_1,B=P_2\begin{pmatrix}I_s&O\\O&O\end{pmatrix}Q_2$，则有

$$\begin{pmatrix}P_1^{-1}&O\\O&P_2^{-1}\end{pmatrix}\begin{pmatrix}A&O\\C&B\end{pmatrix}\begin{pmatrix}Q_1^{-1}&O\\O&Q_2^{-1}\end{pmatrix}=\begin{pmatrix}P_1^{-1}AQ_1^{-1}&O\\P_2^{-1}CQ_1^{-1}&P_2^{-1}BQ_2^{-1}\end{pmatrix}=\begin{pmatrix}I_r&O&O&O\\O&O&O&O\\C_1&C_2&I_s&O\\C_3&C_4&O&O\end{pmatrix}\tag{2.50}$$

对等式右边初等变换，有

$$\begin{pmatrix}I_r&O&O&O\\O&I_{n-r}&O&O\\-C_1&O&I_s&O\\-C_3&O&O&I_{n-s}\end{pmatrix}\begin{pmatrix}I_r&O&O&O\\O&O&O&O\\C_1&C_2&I_s&O\\C_3&C_4&O&O\end{pmatrix}\begin{pmatrix}I_r&O&O&O\\O&I_{n-r}&O&O\\O&-C_2&I_s&O\\O&O&O&I_{n-s}\end{pmatrix}=\begin{pmatrix}I_r&O&O&O\\O&O&O&O\\O&O&I_s&O\\O&C_4&O&O\end{pmatrix}\tag{2.51}$$

由分块矩阵 $\begin{pmatrix}A&O\\C&B\end{pmatrix}$ 与 $\begin{pmatrix}A&O\\O&B\end{pmatrix}$ 相抵知 $C_4=O$，则 $P_2^{-1}CQ_1^{-1}=\begin{pmatrix}C_1&C_2\\C_3&O\end{pmatrix}$ 进而

$$\begin{aligned}C&=P_2\begin{pmatrix}C_1&C_2\\C_3&O\end{pmatrix}Q_1=P_2\begin{pmatrix}C_1&O\\C_3&O\end{pmatrix}Q_1+P_2\begin{pmatrix}O&C_2\\O&O\end{pmatrix}Q_1\\&=P_2\begin{pmatrix}C_1&O\\C_3&O\end{pmatrix}\begin{pmatrix}I_r&O\\O&O\end{pmatrix}Q_1+P_2\begin{pmatrix}I_s&O\\O&O\end{pmatrix}\begin{pmatrix}O&C_2\\O&O\end{pmatrix}Q_1\end{aligned}$$

$$= P_2 \begin{pmatrix} C_1 & O \\ C_3 & O \end{pmatrix} P_1^{-1} A - B(-Q_2^{-1} \begin{pmatrix} O & C_2 \\ O & O \end{pmatrix} Q_1) \tag{2.52}$$

即存在 $X = P_2 \begin{pmatrix} C_1 & O \\ C_3 & O \end{pmatrix} P_1^{-1}, Y = -Q_2^{-1} \begin{pmatrix} O & C_2 \\ O & O \end{pmatrix} Q_1$ 使得 $XA - BY = C$.

注 一些秩相关结论的证明用到分块初等变换，详见之后的“矩阵秩的相关性质”章节.

求逆矩阵

¶ 初等变换法

对于一般的矩阵求逆，采用初等变换法，即将矩阵与同阶单位阵按左右 (上下) 相邻位置写在一起，同时进行初等行 (列) 变换，当矩阵变为单位阵时，单位阵相应地变为了逆矩阵. 行变换原理即 $P_k P_{k-1} \cdots P_1 A = I \Rightarrow P_k P_{k-1} \cdots P_1 I = A^{-1}$，列变换同理，注意行列变换不能混在一起用. 把矩阵初等变换为单位阵的步骤类似于高斯消元法.

例题 2.20（2018-2019 第二学期期中 1.(3)） 求 $\begin{pmatrix} 1 & -2 & 0 \\ 1 & 1 & -2 \\ 1 & 0 & 1 \end{pmatrix}$ 的逆

解

$$\begin{pmatrix} 1 & -2 & 0 & 1 & 0 & 0 \\ 1 & 1 & -2 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 0 & 1 & 0 & 0 \\ 0 & 3 & -2 & -1 & 1 & 0 \\ 0 & 2 & 1 & -1 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -4/3 & 1/3 & 2/3 & 0 \\ 0 & 3 & -2 & -1 & 1 & 0 \\ 0 & 0 & 7/3 & -1/3 & -2/3 & 1 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & 0 & 0 & 1/7 & 2/7 & 4/7 \\ 0 & 3 & 0 & -9/7 & 3/7 & 6/7 \\ 0 & 0 & 7/3 & -1/3 & -2/3 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 1/7 & 2/7 & 4/7 \\ 0 & 1 & 0 & -3/7 & 1/7 & 2/7 \\ 0 & 0 & 1 & -1/7 & -2/7 & 3/7 \end{pmatrix} \tag{2.53}$$

注 求逆矩阵是必考题，各题型位置均有可能考察，这题是随便挑了一道，大部分题目用初等变换即可解决，有些题目使用其他技巧会更快

¶ 利用伴随矩阵

利用 $A^{-1} = \dfrac{A^*}{detA}$ 来计算逆矩阵，适用于各 $(n-1)$ 阶子式形状接近，只需要求解部分代表就可以得到全部子式的情形.

例题 2.21 作业题 34.(5) 可采用这种方法.

¶ 利用矩阵多项式

例题 2.22 若 $A^l = O, l$ 为整数，则 $I - A$ 可逆，且 $(I-A)^{-1} = I + A + A^2 + \cdots + A^{l-1}$

证明

$$(I-A)(I + A + A^2 + \cdots + A^{l-1}) = I - A^l = I \tag{2.54}$$

注 这类似于淑芬的无穷级数求和，需要矩阵满足 $\lim\limits_{k\to+\infty} A^k \to O$，显然本例的 A 满足条件.

例题 2.23 若 n 阶方阵 A 满足 $b_m A^m + b_{m-1}A^{m-1} + \cdots + b_1 A + b_0 I = 0$ 且 $b_0 \neq 0$，则 A 可逆，且 $A^{-1} = -\dfrac{1}{b_0}(b_m A^{m-1} + b_{m-1}A^{m-2} + \cdots + b_1 I)$

证明 直接相乘验证即可，过程略.

例题 2.24 求 n 阶方阵 A 的逆，$A=\begin{pmatrix}1 & b & b^2 & \cdots & b^{n-1}\\ 0 & 1 & b & \cdots & b^{n-2}\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & b\\ 0 & 0 & 0 & \cdots & 1\end{pmatrix}$

解 记 $J_{n\times n}=\begin{pmatrix}0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1\\ & & & 0\end{pmatrix}_{n\times n}$，满足 $J^k=\begin{cases}\begin{pmatrix}O & I_{n-k}\\ O & O\end{pmatrix} & k<n\\ O & k\geq n\end{cases}$

则易见 $A=I+bJ+b^2J^2+\cdots+b^{n-1}J^{n-1}$，而 $A(I-bJ)=I-b^nJ^n=I$，即

$$A^{-1}=I-bJ=\begin{pmatrix}1 & -b & 0 & \cdots & 0\\ 0 & 1 & -b & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & -b\\ 0 & 0 & 0 & \cdots & 1\end{pmatrix} \tag{2.55}$$

¶ 转化为解方程

$AA^{-1}=I$ 可以写为 n 个线性方程组：$A\alpha_i=e_i \quad (1\leq i\leq n)$，其中 α_i 为 A^{-1} 的第 i 列，e_i 为 n 维向量空间的第 i 个自然基，可以看到，求逆矩阵即求解这 n 个线性方程组.

例题 2.25 求 n 阶方阵 A 的逆 $(n\geq 2)$，$A=\begin{pmatrix}1 & 2 & 3 & \cdots & n\\ n & 1 & 2 & \cdots & n-1\\ \vdots & \vdots & \vdots & & \vdots\\ 2 & 3 & 4 & \cdots & 1\end{pmatrix}$

解 先求解方程组 $Ax=\beta$，其中 $\beta=(b_1,b_2,\cdots,b_n)$. 将 n 个方程相加，得到

$$\frac{n(n+1)}{2}(x_1+x_2+\cdots+x_n)=\sum_{j=1}^{n}b_j \tag{2.56}$$

令 $y=x_1+x_2+\cdots+x_n$，有

$$y=\frac{2}{n(n+1)}\sum_{j=1}^{n}b_j \tag{2.57}$$

用第 i 个方程减去第 $i+1$ 个方程 $(i=1,2,\cdots,n$，规定 $i=n$ 时 $i+1=1)$，得：

$$x_1+x_2+\cdots+x_{i-1}+(1-n)x_i+x_{i+1}+\cdots+x_n=b_i-b_{i+1} \tag{2.58}$$

即

$$x_i=\frac{1}{n}(y-b_i+b_{i+1}) \tag{2.59}$$

记 $s=\frac{2}{n(n+1)}$，依次取 $\beta=e_1,e_2,\cdots,e_n$，得

$$A^{-1}=\frac{1}{n}\begin{pmatrix}s-1 & s+1 & s & \cdots & s\\ s & s-1 & s+1 & \cdots & s\\ s & s & s-1 & \cdots & s\\ \vdots & \vdots & \vdots & & \vdots\\ s+1 & s & s & \cdots & s-1\end{pmatrix} \tag{2.60}$$

¶ 利用一些恒等式

记忆力不好的不推荐，如笔者.

命题 2.13

1.Sherman-Morrison 公式

2. 设 $A \in F^{m\times n}, B \in F^{n\times m}$，则 $I_m - AB$ 可逆等价于 $I_n - BA$ 可逆，且 $(I_m - AB)^{-1} = I_m + A(I_n - BA)^{-1}B$，$(I_n - BA)^{-1} = I_n + B(I_m - AB)^{-1}A$.(适用于主对角线元素明显区别于其他位置元素的情况)

证明 （仅证 2.）由分块矩阵初等变换得：

$$\begin{pmatrix} I_n & -B \\ O & I_m \end{pmatrix}\begin{pmatrix} I_n & O \\ A & I_m \end{pmatrix}\begin{pmatrix} I_n & O \\ O & I_m - AB \end{pmatrix}\begin{pmatrix} I_n & B \\ O & I_m \end{pmatrix}\begin{pmatrix} I_n & O \\ -A & I_m \end{pmatrix} = \begin{pmatrix} I_n - BA & O \\ O & I_m \end{pmatrix} \tag{2.61}$$

两边同时求逆立即得到结果.

这个结果还蛮常用的，很多求逆矩阵的题用它可以大大简化，如下面的例题.

例题 2.26（2022-2023 第二学期期中 4.） 求矩阵 A 的逆 $(x>0)$，$A = \begin{pmatrix} 4+x & 4 & 4 & 4 \\ 3 & 3+x & 3 & 3 \\ 2 & 2 & 2+x & 2 \\ 1 & 1 & 1 & 1+x \end{pmatrix}$

解 由 $A = xI + (4,3,2,1)^T(1,1,1,1)$，则由恒等式 2. 可知

$$\begin{aligned}(\frac{1}{x}A)^{-1} &= I - \frac{1}{x}(4,3,2,1)^T(1+\frac{1}{x}(1,1,1,1)(4,3,2,1)^T)^{-1}(1,1,1,1) \\ &= I - \frac{1}{x+10}(4,3,2,1)^T(1,1,1,1)\end{aligned} \tag{2.62}$$

故

$$A^{-1} = \frac{1}{x}(\frac{1}{x}A)^{-1} = \frac{1}{x(x+10)}\begin{pmatrix} x+6 & -4 & -4 & -4 \\ -3 & x+7 & -3 & -3 \\ -2 & -2 & x+8 & -2 \\ -1 & -1 & -1 & x+9 \end{pmatrix} \tag{2.63}$$

矩阵秩的相关性质

¶ 常用的秩等式与不等式

命题 2.14

1. A 是实矩阵，$rank(A^T) = rank(A) = rank(A^TA) = rank(A^TAA^TA) = \cdots$

证明 需要用到线性方程组解空间的性质[a]

引理 2.1

$rank(AB) = rank(B) \iff$ 方程组 $ABx = O$ 与 $Bx = O$ 同解.

证明 设 B 的列数为 p，AB 和 B 的解空间分别为 V_1 和 V_2，则
$\Rightarrow$: $dim(V_1) = p - rank(AB) = p - rank(B) = dim(V_2)$，而显然 $Bx = O$ 的解均为 $ABx = O$ 的解，则两方程组同解.
$\Leftarrow$: 两方程组同解，则 $dim(V_1) = dim(V_2)$，则 $rank(AB) = rank(B)$.

由此可见，证明 $rank(A) = rank(A^TA)$ 等价于证明 $Ax = O$ 与 $A^TAx = O$ 同解:
(1) 显然 $Ax = O$ 的解必定是 $A^TAx = O$ 的解;
(2) 设 x_1 为 $A^TAx = O$ 的一个解，则 $x_1^TA^TAx_1 = 0$，则 $(Ax_1)^T(Ax_1) = 0$，则 $Ax_1 = O$，即 x_1 也是

$Ax = O$ 的一个解.
结合 (1)(2) 知命题成立.

2. $rank\begin{pmatrix} A & O \\ O & B \end{pmatrix} = rank(A) + rank(B),\quad rank\begin{pmatrix} A & C \\ O & B \end{pmatrix} \geq rank(A) + rank(B)$
3. $rank(A \quad B) \geq max\{rank(A), rank(B)\}$
4. $rank(A+B) \leq rank(A \quad B) \leq rank(A) + rank(B)$
5. $rank(AB) \leq min\{rank(A), rank(B)\}$
6. (Frobenius 秩不等式)$rank(ABC) + rank(B) \geq rank(AB) + rank(BC)$
   特别地，当 B 取 I 时有 (n 为 A 的列数，C 的行数)
   (Sylvester 秩不等式)$rank(AC) \geq rank(A) + rank(C) - n$
7. $A \in F^{m\times n}, B \in F^{n\times m}$，则 $m + rank(I_n - BA) = rank\begin{pmatrix} I_m & A \\ B & I_n \end{pmatrix} = n + rank(I_m - AB)$
8. n 阶方阵 A 幂等 ($A^2 = A$) 等价于 $rank(A) + rank(I - A) = n$
9. n 阶方阵 A 满足 $A^2 = I$ 等价于 $rank(I + A) + rank(I - A) = n$

[a] 2020-2021 第二学期期中 6.

¶ 秩的计算

1. 通过初等变换化为标准型
不同于求逆，求秩时可以初等行、列变换混用.

例题 2.27（2022-2023 第二学期期中 1.(4)） 求矩阵 $A = \begin{pmatrix} 1 & 0 & 1 & -4 \\ -1 & -3 & -4 & -2 \\ 2 & -1 & 4 & 4 \end{pmatrix}$ 的相抵标准型

解 进行初等变换

$$\begin{pmatrix} 1 & 0 & 1 & -4 \\ -1 & -3 & -4 & -2 \\ 2 & -1 & 4 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 & -4 \\ 0 & -3 & -3 & -6 \\ 0 & -1 & 2 & 12 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 & -4 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 3 & 14 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & 0 & 0 & -26/3 \\ 0 & 1 & 0 & -8/3 \\ 0 & 0 & 1 & 14/3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{2.64}$$

2. 求矩阵的最高阶非零子式
通常用于矩阵行列式为 0，但很容易找到一个 $n-1$ 阶子式不为零的情况 (秩比较小时要排除的子式过多)

例题 2.28（讲义例 4.4.5） 计算 n 阶方阵 A 的秩，$A = \begin{pmatrix} 1 & 1 & & \\ & 1 & \ddots & \\ & & \ddots & 1 \\ 1 & & & 1 \end{pmatrix}$

解 A 的行列式按第一列展开易得 $det(A) = 1 + (-1)^{n+1}$
  (1) n 为奇数时 A 的行列式非零，则秩为 n
  (2) n 为偶数时 A 的行列式为 0，但易知其右上角的 $n-1$ 阶子式非零，故秩为 $n-1$

¶ 伴随矩阵 A^* 相关

命题 2.15

A 为 $n(\geq 2)$ 阶方阵，则有

1. $det(A^*) = (det(A))^{n-1}$

2. $rank(A^*) = \begin{cases} n & rank(A) = n \\ 1 & rank(A) = n-1 \\ 0 & rank(A) < n-1 \end{cases}$

证明 （仅证 2.）(1) $rank(A) = n$，则 $detA \neq 0$，则 $det(A^*) \neq 0$，则 $rank(A^*) = n$；

(2) $rank(A) = n-1$，则 A 有不为 0 的 (n-1) 阶子式，则 $A^* \neq 0$，又由 Sylvester 秩不等式：$rank(A^*) \leq rank(A^*A) + n - rank(A) = rank(det(A)I_n) + n - (n-1) = 1$，故 $rank(A^*) = 1$；

(3) $rank(A) < n-1$，则 A 的所有 (n-1) 阶子式均为 0，即 $A^* = O$.

命题 2.16

3. $(A^*)^* = \begin{cases} (det(A))^{n-2}A & n \geq 3 \\ A & n = 2 \end{cases}$

证明 (1) $n \geq 3$ 时，若 $det(A) \neq 0$，则 $(A^*)^{-1}$ 存在，则 $(A^*)^* = det(A^*)(A^*)^{-1} = (det(A))^{n-1}\frac{1}{det(A)}A = (det(A))^{n-2}A$；若 $det(A) = 0$，则 $rank(A^*) \leq 1$，则 $(A^*)^* = O$；

(2) $n = 2$ 时易证.

若已知 $n(\geq 3)$ 阶方阵 A^*，如何求 A?

(1) $rank(A^*) = n$ 时，由 $A^*A = det(A)I$ 得 $A = det(A)(A^*)^{-1} = (det(A^*))^{\frac{1}{n-1}}(A^*)^{-1}$

例题 2.29（2017-2018 第二学期期中 1.(4)） $A^* = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & -1 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{pmatrix}$，求 A

解 易知 A^* 满秩，$det(A^*) = -8$，$(A^*)^{-1} = \begin{pmatrix} 0 & 0 & 0 & 1/4 \\ 0 & 0 & -1 & 0 \\ 0 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$，则

$$A = (det(A^*))^{\frac{1}{n-1}}(A^*)^{-1} = \begin{pmatrix} 0 & 0 & 0 & -1/2 \\ 0 & 0 & 2 & 0 \\ 0 & -1 & 0 & 0 \\ -2 & 0 & 0 & 0 \end{pmatrix} \tag{2.65}$$

注 原矩阵在复数域上不唯一，解答只给出了其中一种实方阵.

(2) $rank(A^*) = 1$ 时，$rank(A) = n-1$，$A^*A = O$，由此可见 A 的 n 个列向量均为方程组 $A^*x = O$ 的解，这个齐次方程组的解空间维数为 $dim(V) = n - rank(A^*) = n-1$，可选出 $n-1$ 个基来进行线性组合得到 A 的各个列向量，并进一步由 $AA^* = O$ 确定组合系数的约束，可最终确定 A 的形式.

(3) $A^* = O$ 时，任意秩小于 $n-1$ 的 n 阶方阵均可作为原矩阵 A.

¶ 满秩分解

习题 2.27 (2014-2015 第二学期期中 8.)

若 $A \in F^{m\times n}$, $rank(A) = r$，则 A 可分解为一个列满秩矩阵 $B \in F^{m\times r}$ 和一个行满秩矩阵 $C \in F^{r\times n}$ 的乘积 $A = BC$. 特别地，当 A 行满秩时，有 $A = I_m A$，当 A 列满秩时，有 $A = AI_n$.

证明 $rank(A) = r \Rightarrow A = P\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}Q$，其中 $P \in F^{m\times m}, Q \in F^{n\times n}$ 均为可逆方阵，则有

$$A = P\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}Q = P\begin{pmatrix} I_r \\ O \end{pmatrix}(I_r \quad O)Q = P'Q' \tag{2.66}$$

其中 $P' = P\begin{pmatrix} I_r \\ O \end{pmatrix}, Q' = (I_r \quad O)Q$ 分别为列满秩矩阵和行满秩矩阵.

线性空间

当我们用线性空间的思想去回看前几章的一些内容，会得到一些新的理解.

1. 线性方程组

线性方程组 $Ax = \beta$ 有解

$\iff x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta$ 有解，其中 α_i 为系数矩阵 A 的第 i 列

$\iff \beta \in <\alpha_1, \alpha_2, \cdots, \alpha_n>$

$\iff <\alpha_1, \alpha_2, \cdots, \alpha_n, \beta> = <\alpha_1, \alpha_2, \cdots, \alpha_n>$

$\iff dim <\alpha_1, \alpha_2, \cdots, \alpha_n, \beta> = dim <\alpha_1, \alpha_2, \cdots, \alpha_n>$

$\iff rank(A \quad \beta) = rank(A)$

也就是说，我们求解线性方程组等价于求解满足 $x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta$ 的线性组合系数，亦即：求某一个向量能否表示成一些向量的线性组合的方法是：把这些向量摆在一起得到一个线性方程组并求解.

2. 矩阵乘法

矩阵等式 $AB = C$ 可以有新的理解角度：

C 的每一行都是 B 的行向量的线性组合，C 的每一列都是 A 的列向量的线性组合

进一步地，B 的第 i 列 b_i 是方程 $Ax = c_i$ 的解，其中 c_i 是 C 的第 i 列，从这个角度，我们可以给出作业题 25. 的另一种解法：

解 由于 A 的每行、每列元素求和都为 0，则 $det(A) = 0$. 下面分类讨论：

(1) $rank(A) < n-1$，有 $rank(A^*) = O$，即 $A^* = O$;

(2) $rank(A) = n-1$，有 $Ax = O$ 的解空间 V 的维数 $dimV = n - rank(A) = 1$，而 $(1, 1, \cdots, 1)^T$ 为 $Ax = O$ 的一个非零解，故 $Ax = O$ 的通解为 $x = \lambda(1, 1, \cdots, 1)^T$.

另一方面，$AA^* = det(A)I = O$，则 A^* 的每一列均为 $Ax = O$ 的解，即正比于 $(1, 1, \cdots, 1)^T$

又注意到 $A^*A = O$，故 $A^T(A^*)^T = O$，即 $(A^*)^T$ 的每一列 (对应 A^* 的每一行) 也正比于 $(1, 1, \cdots, 1)^T$

综上，有 A^* 正比于 $\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$

3. 秩

矩阵的秩等于它的列向量组的秩，也等于它的行向量组的秩

我们可证明以下性质：

命题 2.17

(1) 若向量组 $\{b_1, b_2, \cdots, b_s\}$ 可由 $\{a_1, a_2, \cdots, a_r\}$ 线性表示，则 $rank(b_1, b_2, \cdots, b_s) \leq rank(a_1, a_2, \cdots, a_r)$

(2) 若向量组 $\{b_1, b_2, \cdots, b_s\}$ 与 $\{a_1, a_2, \cdots, a_r\}$ 等价，则 $rank(b_1, b_2, \cdots, b_s) = rank(a_1, a_2, \cdots, a_r)$ ♠

由此，我们可以轻易给出很多秩不等式的证明，例如：

例题 2.30 证明：$rank(AB) \leq min\{rank(A), rank(B)\}$

证明 AB 的每个列向量都是 A 的列向量的线性组合，每个行向量都是 B 的行向量的线性组合，故结论成立.

2.4.2 期中考试知识点整理

线性代数的题目具有技巧性，很多时候虽然弄明白了上课讲的知识点，但做题仍然无从下手，这就需要同学们进行一定的题目练习巩固知识点、积累技巧，毫无疑问，课后习题和往年试卷提供了很好的练习样本。总览一遍往年试卷会发现，考试题目的难度普遍不是很大 (不高于作业题和上课例题)，且每年的题型和考点大同小异，因此，在考前刷上几套往年卷来熟悉考试内容，会有很大帮助。为梳理知识点，我在下面按章节列出了一些重要内容供大家参考，可以根据我列出的条目回想具体内容，想不起来的去查看讲义，以此提高复习效率。期中复习建议至少抽出三个半天 (各 4h，周末或周内晚上) 的时间来进行：知识点回顾 → 对照参考答案过一遍作业题 (作业题中有不少很有用的结论，记住它们可以帮助你更快地解题)→ 刷往年题 (每套限制在 2h 之内)。由于 11 月 16 日要考线代和数分两门期中 (加起来 10 学分，比助教本学期所有课加起来还多 (bushi))，所以考试前一天晚上要保证充足睡眠，不要熬夜；考试当天按时起来吃早饭，并至少提前二十分钟到考场 (做准备、查看座位号等)。预祝大家取得理想成绩！

线性方程组

一般会考一道含参的线性方程组求解问题，并依据参数讨论解的情况。

¶ Gauss 消元法

实质就是初等行变换，因为初等变换不改变方程组的解。

一般的步骤是：写出方程组的增广矩阵 → 通过初等行变换化为阶梯形式 → 根据阶梯形矩阵判断方程组有无解 (最后的非零行是否只有最后一列的元素非零)，若有解，从下到上依次回代求解。

¶ 解的情况判别

无解：最后的非零行只有最后一列元素非零；

唯一解：有解，且非零行个数 (即秩) r = 未知数个数 n；

无穷多解：有解且 $r < n$，通解需要设 $(n - r)$ 个参数来表示；

齐次线性方程组：方程右边都为 0 的线性方程组，方程一定有零解，有非零解 $\iff$ 有无穷多解

行列式

一般会有一道大题直接考察计算行列式，且由于行列式在其他很多地方有应用，故可能会出现在算其他东西的过程中

¶ 定义

行列式的行、列数目要相同，运算结果是一个数。

书上采用代数余子式递归定义，也可以采用完全展开式定义，也可以按照函数定义：满足多重线性性、反对称性、规范性的方阵函数。

行列式几何意义：平行多面体的有向体积 (注意四面体体积是平行六面体的 1/6)
直接考察行列式定义的考题不多，可能会考到体积计算

¶ 性质

- 行列等价 $det(A) = det(A^T)$，由于行列等价，下面的性质只描述行
- 交换两行出负号
- 某行的公因子可以提出去
- 若某行可写为两个向量相加，则行列式等于分别把这行替换为这两个向量的行列式相加
- 若有两行成比例 (含相等)，行列式为 0
- 把某行乘一个常数倍加到另一行，行列式不变
- 三角阵 (准三角阵) 的行列式等于对角元素 (的行列式) 相乘

¶ 展开式

- 单行展开：$det(A) = \sum\limits_{k=1}^{n} a_{ik}A_{ik} = \sum\limits_{k=1}^{n} (-1)^{i+k} a_{ik}M_{ik}$
  回忆子式、余子式、代数余子式的概念
  反之，$\sum\limits_{k=1}^{n} a_{jk}A_{ik} = \begin{cases} det(A) & j = i \\ 0 & j \neq i \end{cases}$，这在涉及子式的计算的题可能会用到
- 多行展开 (即 Laplace 展开)，按选定的 $i_1, i_2, \cdots, i_r$ 行展开：
  $$det(A) = \sum_{1\le k_1<k_2<\cdots<k_r\le n} (-1)^{i_1+i_2+\cdots+i_r+k_1+k_2+\cdots+k_r} A\begin{pmatrix} i_1 & \cdots & i_r \\ k_1 & \cdots & k_r \end{pmatrix} A\begin{pmatrix} i_{r+1} & \cdots & i_n \\ k_{r+1} & \cdots & k_n \end{pmatrix}$$
- 完全展开：$det(A) = \sum\limits_{(j_1,j_2,\cdots,j_n)\in S_n} (-1)^{\tau(j_1,j_2,\cdots,j_n)} a_{1j_1}a_{2j_2}\cdots a_{nj_n}$
  回忆全排列、逆序数

¶ 一些计算技巧

- 利用初等变换化为上三角行列式 (注意交换两行出负号)
- 按多重线性拆成多个行列式的求和
- 行和/列和有特殊性时，将其余 $(n-1)$ 行 (列) 加到某一行 (列)
- 利用 Laplace 公式
- 归纳法
- 递推法 (三对角行列式)
- 利用 Binet-Cauchy 公式 (两矩阵相乘的行列式算法)
- 利用分块矩阵降阶 (降阶公式)
- 将行列式视为一些元素的多项式 (例如问你展开式中 x^k 的系数是多少)
- 加边法 (行列式每行/列都有相同的数)
- 利用 Vandermonde 行列式 (熟悉 Vandermonde 行列式求法)

¶ Cramer 法则

对于 n 个方程 n 个未知数的线性方程组，若其系数矩阵的行列式不为 0，则方程组有唯一解，特别地，对齐次方程组系数行列式不为零代表着只有零解。

实际就是 $Ax = \beta$ 在 $detA \neq 0$ 时有 $x = A^{-1}\beta$

矩阵

矩阵这一章有很多知识点，考题也较多样，分布在各个题型中

¶ 定义及一些特殊矩阵

矩阵即由 m 行 n 列共 $(m\times n)$ 个数排成的矩形列表，它可以与线性映射建立一一对应关系。

回忆一些特殊矩阵的定义：零矩阵 O、方阵、单位阵 I、数量阵 aI、(准) 对角阵、(准) 三角阵、对称阵、反对称阵、基础矩阵 E_{ij}

¶ 矩阵的运算

1. 加法：需两个矩阵大小相同，二者对应位置元素相加；
   数乘：矩阵每个位置的元素都乘上相同的数

2. 乘法：需左边矩阵列数等于右边矩阵行数，满足乘法结合律、数乘结合律、分配律，不满足交换律、消去律 ($AB=AC$ 且 $A\neq O$ 不能推出 $B=C$)，矩阵乘积的行列式满足 Binet-Cauchy 公式，矩阵乘法对应于线性映射的复合

3. 由矩阵乘法自然地给出矩阵幂次的概念，一些求幂次的方法：
   - 归纳
   - 分块
   - 拆分为容易求幂且对易的矩阵之和：$A^k=(B+C)^k=\sum\limits_{i=0}^{k}C_k^iB^{k-i}C^i$，其中 $BC=CB$
   - 转化为低阶矩阵乘积
   - 相似对角化：$A=P^{-1}BP$，其中 $B=diag(b_1,b_2,\cdots,b_n)$，则 $A^k=P^{-1}B^kP=P^{-1}diag(b_1^k,b_2^k,\cdots,b_n^k)P$

4. 转置：矩阵的行列互换，记作 A^T，满足：
   - $(A+B)^T=A^T+B^T$
   - $(\lambda A)^T=\lambda A^T$
   - $(AB)^T=B^TA^T$
   - $(A^{-1})^T=(A^T)^{-1}$

5. 共轭：矩阵每个元素求复共轭，记作 $\bar{A}$

6. 迹：方阵对角元的和，记作 $tr(A)$，满足：
   - $tr(A+B)=tr(A)+tr(B)$
   - $tr(\lambda A)=\lambda tr(A)$
   - $tr(AB)=tr(BA)$
   - $tr(A^T)=tr(A),tr(\bar{A})=\overline{tr(A)}$
   - $tr(A\bar{A}^T)=0\to A=O$

7. 分块：对分块矩阵进行运算，可以先将每个矩阵块当做元素运算，再对每个矩阵块进行相同的运算 (有的同学求转置等运算时会忘了这一步)

   分块矩阵可用于求行列式、求逆、求秩等 (具体见“分块矩阵初等变换”章节)

   回顾准三角阵的逆、Schur 公式等

8. 初等变换：熟悉三种初等变换矩阵的形式；左乘代表行变换，右乘代表列变换

初等变换不改变方程组的解，可用于化简方程；与行列式性质相关，可用于求行列式；不改变矩阵的秩，可用于求秩与相抵标准形。

9. 逆矩阵：$A^{-1}A = AA^{-1} = I$，不可逆的方阵又称奇异阵

矩阵可逆的充要条件：可逆 $\Longleftrightarrow det(A) \neq 0 \Longleftrightarrow rank(A) = n \Longleftrightarrow$ A 等于一些初等矩阵相乘 $\Longleftrightarrow$ A 的行 (列) 向量组线性无关

A 的伴随阵 A^* 满足 $A^*A = AA^* = det(A)I$，注意 $A^*_{ij} = A_{ji}$，伴随阵的相关性质见课后习题及本讲义“伴随矩阵 A^* 相关”小节

逆矩阵的求解技巧见本讲义“求逆矩阵”章节

10. 穿脱规则及复合规则

穿脱规则：$(AB)^T = B^TA^T$，　$(AB)^{-1} = B^{-1}A^{-1}$，　$(AB)^* = B^*A^*$

复合规则：$(A^{-1})^T = (A^T)^{-1}$，　$(A^{-1})^* = (A^*)^{-1}$，　$(A^*)^T = (A^T)^*$

¶ 秩与相抵

回顾行秩、列秩、矩阵的秩及满秩的概念，行秩 = 列秩 = 矩阵的秩

矩阵的秩等于其最高阶非零子式的阶数

大小相等的两个矩阵 A, B 若秩相等则称二者相抵，可通过一系列初等变换互相转化，属于同一相抵等价类，具有相同的相抵标准型

相抵标准型：与秩为 r 的矩阵大小相同，且形式为 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 的矩阵，注意一个常见的错误：$A \in F^{m\times n}$ 的秩为 r，它的相抵标准型不为 I_r(只有 $r = m = n$ 成立)

线性方程组 $Ax = \beta$ 有解 $\Longleftrightarrow rank(A) = rank(A \quad \beta) = r$

每个秩为 r 的矩阵都可以写为 r 个秩为 1 的矩阵之和

秩的一些常用等式和不等式及秩的求法详见本讲义“矩阵秩的相关性质”章节

线性空间

¶ 数组空间与一般线性空间

线性空间即数域 $\mathbb{F}$ 上定义了加法和数乘，且满足八条运算性质的集合 V(有的题目让验证集合构成线性空间，即验证这八条运算性质)

由一般线性空间的定义出发，可以给出子空间 (更多的题目其实是在证明线性空间里满足一定性质的元素的集合构成子空间，这即是证明这个集合对数乘和加法封闭)、生成子空间 (由一些元素进行线性组合得到的子空间)、生成元、线性组合、线性相关、线性无关、极大无关组与秩、基与维数的概念

几乎每年都会有一道验证一般线性空间并延续讨论的大题，不会很难，理解好各概念与数组空间中概念的对应关系就很容易求解

¶ 线性相关与线性无关

一组向量线性相关 $\Longleftrightarrow$ 存在一组不全为 0 的组合系数使它们组合出零向量 $\Longleftrightarrow$ 存在其中一个向量可以被其他向量线性表示，反之线性无关

给定向量组判断线性相关/无关一般需要将向量组写在一起形成一个矩阵然后求秩，当然，有很多其他的判断方法

回忆众多的判断方法及相关说法的正确性 (老师上课讲过的及作业题中的)，相关说法的正确性经常在判断题中考察

¶ 极大无关组与秩

极大无关组：向量组的一个子向量组，满足自身线性无关，且外加任何一个子向量组外的向量都线性相关

极大无关组求法：将列向量组摆在一起，经过初等行变换化为更简单的向量组，找出变换后向量组的极大无关组，其对应变换前的极大无关组

秩：一个向量组的极大无关组可以不唯一，但每个极大无关组包含的向量数目相同，记为向量组的秩

m 个向量线性无关 $\iff$ 向量组的秩为 m

回忆行秩和列秩的定义，初等变换不改变秩，矩阵的秩 = 行秩 = 列秩

¶ 基与维数

基：n 维数组空间的子空间 V 中一组向量称为 V 的一组基，如果它们线性无关且 V 中任意向量 a 可以表示成这组向量的线性组合，组合系数称为 a 在这组基下的坐标.

维数：V 的一组基的向量个数称为 V 的维数，有 $dimV = rank(V)$

基变换与坐标变换：向量在不同基下的坐标一般不同，选定两组基 $a_1, a_2, \cdots, a_n$ 和 $b_1, b_2, \cdots, b_n$，两组基可以通过一个矩阵联系：$(b_1, b_2, \cdots, b_n) = (a_1, a_2, \cdots, a_n)T$，称 T 为基 $a_1, a_2, \cdots, a_n$ 到基 $b_1, b_2, \cdots, b_n$ 的过渡矩阵；基线性无关 $\to$ 矩阵 T 可逆，设一个向量在两组基下的坐标分别为 X 和 Y，则有 $(a_1, a_2, \cdots, a_n)X = (b_1, b_2, \cdots, b_n)Y \Rightarrow X = TY \Rightarrow Y = T^{-1}X$，即从老坐标到新坐标的坐标变换公式

求过渡矩阵即是解矩阵方程 $(b_1, b_2, \cdots, b_n) = (a_1, a_2, \cdots, a_n)T$

$$\Rightarrow T = (a_1, a_2, \cdots, a_n)^{-1}(b_1, b_2, \cdots, b_n)$$

定理 2.2 (扩充基定理)

r 维子空间 V 中任意 $s(<r)$ 个线性无关的向量可以加 $r-s$ 个向量扩充为 V 的一组基 (回顾证明)

¶ 线性方程组解的结构

线性方程组 $Ax = \beta$ 有解 $\iff rank(A) = rank(A \quad \beta)$

n 元线性方程组有唯一解 $\iff rank(A) = n$

n 元齐次线性方程组 $Ax = O$ 有非零解 $\iff rank(A) < n$

齐次线性方程组解的结构：$Ax = O$ 的解空间 V 为 $\mathbb{F}^n$ 的子空间，且 $dimV = n - rank(A)$，解空间的一组基 $\alpha_1, \alpha_2, \cdots, \alpha_{n-r}$ 称为线性方程组的一个基础解系，通解可写为基础解系的线性组合形式 $x = t_1\alpha_1 + \cdots + t_{n-r}\alpha_{n-r}$. 基础解系的求法类似于 $Gauss$ 消元法，化为阶梯形矩阵后将 $(n-r)$ 个多余自变量移到等号右边设为参数，由此求得参数解对应的向量组即基础解系.

非齐次线性方程组解的结构：$Ax = \beta$ 的解空间即对应的齐次方程组 $Ax = O$ 的解空间进行了平移的结果，通解即对应的齐次方程组通解 + 任意一个特解

题型与技巧

考试的题型分为填空题 (5 – 6 道)、判断题 (4 – 5 道)、解答题 (一般为 4 道)

¶ 填空题

1. 一般的考点：
   - 矩阵的乘法与方阵的幂
   - 行列式的基本性质与按一行（列）的展开
   - 矩阵的伴随与逆的基本性质
   - 矩阵与向量组的秩，线性相关与线性无关

- 线性方程组解空间的维数
- 基的坐标变换

2. 技巧：填空题只需要答案不需要过程，有些题可以直接使用相关的结论 (不用证明) 快速得到答案，甚至有的题可以直接使用特例来计算结果，例如：

例题 2.31（2018-2019 第一学期 1.(2)） 设 $\alpha_1, \alpha_2, \alpha_3$ 为三维列向量，记矩阵 $A = (\alpha_1, \alpha_2, \alpha_3)$，$B = (\alpha_1 + \alpha_2 + \alpha_3, \alpha_1 + 2\alpha_2 + 4\alpha_3, \alpha_1 + 3\alpha_2 + 9\alpha_3)$. 若 $|A| = 1$，求 $|B|$

解 不妨设 $\alpha_1, \alpha_2, \alpha_3$ 分别为 $(1,0,0)^T, (0,1,0)^T, (0,0,1)^T$，则 $|B| = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{vmatrix} = (2-1)(3-1)(3-2) = 2$

¶ 判断题

1. 一般的考点：
- 矩阵的乘法，转置，共轭，迹等及初等变换
- 行列式及矩阵的秩，相抵关系等
- 向量组的线性关系和极大无关组
- 线性方程组的解的结构相关
- 一般线性空间

2. 判断题一直以来都是比较棘手的题型，一旦判断错误就没有分，并且有的题完全是在玩文字游戏，所以建议同学们谨慎判断严格证明，判断的准确度依赖于平时对各种二级结论的积累，可能某道题正好是之前碰到过的结论，那就非常令人安心了。建议同学们熟悉各种定义、定理和运算的内容及适用的条件，对一些见过的结论留下印象。

判断题的做题步骤一般为：在最开始写下命题正确/错误，若正确则在后面进行证明，若错误则找出反例/进行一定说明即可。

¶ 解答题

1. 一般的考点：
- 线性方程组有解的条件
- 行列式和逆矩阵的计算
- 有关秩的恒等式与不等式
- 线性空间的基与坐标变换
- 线性方程组的解的结构
- 一般线性空间的验证
- 一些性质的证明

2. 解答题将各种知识点综合考察，书上出现过的定理、例题等大概率可以直接使用，老师上课的例题和助教在习题课的补充拿来使用应视情况而定：如果题目本身就是补充的一些结论或只有少许延伸，那就需要进行证明；如果用到的结论只是题目解答的一小部分，那就可以直接使用。

2.5 第五次习题课

2.5.1 作业选讲

首先我挑了几道作业中我个人认为有必要讲的题目来讲解，其他作业题则根据现场同学需求来决定是否讲解，没讲的题可以参考群里发的作业答案。

笔记 由于作业答案都会单独发布，所以下面要讲的几道题就不在讲义里把答案过程抄过来了，而是呈现一个思考的过程，或者是这道题里的某个点值得注意。

习题 2.28 (第五章第 36 题)

将三维几何空间中的直角坐标系 $[O; e_1, e_2, e_3]$ 绕单位向量 $e = \frac{1}{\sqrt{3}}(1,1,1)$ 逆时针旋转 θ 角，求新坐标与原坐标之间的关系。

这道题其实可以不使用基变换的方法来做，可以用一些更偏几何的办法来算出答案。另外本周作业中第六章第 8 题也可以通过这样的方法来做，具体的过程可以参考过两天会发的作业答案。大致的思路就是任取一点，旋转即是其与原点相连的向量绕着旋转轴旋转，这个旋转的轨迹实际上是个圆锥侧面的一部分，那么我们可以只考虑在这个圆锥的底面圆上旋转，以底面圆心为原点建立坐标系，算出三个标准正交基，即 $\boldsymbol{e}_1 = (1,0,0)^T, \boldsymbol{e}_2 = (0,1,0)^T, \boldsymbol{e}_3 = (0,0,1)^T$ 旋转后的像，再由旋转变换的线性性，就可以写出这个旋转变换的表达式了。

习题 2.29 (第五章第 44 题)

设 $\mathbb{F}_n[x]$ 是次数小于或等于 n 的多项式全体构成的线性空间。

(1) 证明 : $S = \{1, x-1, (x-1)^2, \cdots, (x-1)^n\}$ 构成 $\mathbb{F}^n$ 的一组基;

(2) 求 S 到基 $T = \{1, x, \cdots, x^n\}$ 的过渡矩阵;

(3) 求多项式 $p(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{F}_n[x]$ 在基 S 下的坐标

这里要注意这不是一般的数组空间，在数组空间里求过渡矩阵时，基向量可以排列成一个可逆方阵，两边乘上其逆矩阵便能得到过渡矩阵。但这里空间里的元素都是多项式，我们没法把 n 个多项式排列成一个 n 阶可逆矩阵，这种做法也就失效了，所以有同学作业里出现如下这样的：

$$T = (1, x-1, (x-1)^2, \cdots, (x-1)^n)^{-1}(1, x, \cdots, x^n)$$

显然是错误的，甚至是没有意义的写法。此外，更多关于一般线性空间的理解在下一节“拓展内容”中会讲到。

习题 2.30 (第五章第 46 题)

给定矩阵

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

令 V 是与 A 乘法可交换的三阶实方阵全体. 证明 : V 在矩阵加法与数乘下构成实数域上的线性空间，并求 V 的一组基与维数。

这道题参考答案中的方法是因为 A, A^{-1}, I 均与 A 可交换，并且这里 V 恰好维数为 3，对于一般的这类题目，其实通法还是设出九个未知数解方程计算，其实这个矩阵 A 里有很多 0，所以实际上并不算太过复杂。

习题 2.31 (第六章第 5 题)

证明：$\mathbb{R}^2$ 上的可逆线性变换可以分解为关于坐标轴的伸缩、反射及旋转变换的复合。

这道题在群里给过提示，但还是有不少同学在问，思路就是把任意一个可逆方阵 A 通过那三种变换一步步变成单位阵，比如如果有 $P_1P_2\cdots P_nAQ_1Q_2\cdots Q_m=I$，其中 P_i,Q_i 都是伸缩、旋转、反射变换对应的矩阵，那么 $A=P_n^{-1}\cdots P_1^{-1}Q_m^{-1}\cdots Q_1^{-1}$，从而 $\mathscr{A}(x)=Ax$ 就是这三种变换的复合。

那么我们可以先通过右乘伸缩矩阵来使 A 的列向量长度相等，为什么不是左乘呢？因为伸缩矩阵实际上是在拉伸 x 轴和 y 轴，左乘就相当于是先作用变换 $\mathscr{A}$，再复合一个伸缩变换，把矩阵左乘一个伸缩矩阵也能看出来，作用后的结果是原矩阵列向量的 x 分量和 y 分量分别作了伸缩，而不是列向量本身，这个时候想让两个列向量相等的话，我们就不好控制伸缩的比例，甚至还有可能是无解的。但右乘就是先拉伸坐标轴再复合变换 $\mathscr{A}$，就能达到伸缩列向量本身的目的。继续下去就是复合一个旋转，把两个列向量旋转到关于 y 轴对称，这样的话我们伸缩 y 轴，就能很好地控制两个列向量之间的夹角，从而让它们正交，再通过伸缩单位化，最后用旋转和反射就能变为 $(1,0)^T,(0,1)^T$，从而就得到单位阵了。

其实由于矩阵可以分解成初等方阵的乘积，我们也可以仅对初等方阵验证这个命题，但其实没简化多少。

习题 2.32 (第六章第 8 题)

在三维几何空间的直角坐标系中，求绕向量 $\boldsymbol{e}=(1,-1,1)^T$ 逆时针旋转 30° 角的变换。

与第五章第 36 题基本一致，但注意这里的旋转轴不是单位向量（应该是漏了个系数），如果要用旋转公式要先单位化，好几个同学都是直接带的公式，具体过程见后续发布的作业答案。

还有同学想听第六章那两个求对称变换的题目，其实习题课这会儿第六章这次作业还没截止）所以我只能说一下思考的过程（包括上面两个题目），对称这种性质比较好的其实可以直接算，设任意一点 $(x,y,z)^T$，其对称点为 (u,v,w)，我们要做的就是用 x,y,z 线性表示出 u,v,w，由于对称，两个点的中点落在对称轴（或对称平面）上，而且二者连线段对应的向量与对称轴（或对称平面）垂直，这样列出两个方程就能解出关系了。

2.5.2 拓展内容

一般线性空间

一般线性空间，也叫做向量空间，但是这里的“向量”不一定就是通常所指的有方向、有长度的线段，而是一种抽象的概念。线性空间实质上是一个集合 M，只不过是在这个集合上再加上一些运算与结构。考虑一个数域 $\mathbb{F}$，我们在 M 上赋予两个运算：加法与数乘，这里的加法也不一定就是我们通常认为的数字之间的加法，而是我们定义的一种映射关系，这个映射的定义域是 $M\times M$，值域是 M，$f:M\times M\to M,(\alpha,\beta)\mapsto\gamma$，也就是把 M 上的任意两个元素映到 M 上的另外一个元素，我们把这个映射定义为“加法”运算。数乘也同理，它是定义在 $\mathbb{F}\times M$ 上的映射，值域为 M，也就是把 M 上一个元素，及数域 $\mathbb{F}$ 上一个数，映到 M 上另一个元素，我们把这个映射定义为“数乘”运算。

接下来我们需要对这两个映射做一些限制，以保证它们具有我们想要的好的性质，这也就是所谓构成线性空间的八条公理，我们可以随便定义“加法”与“数乘”，只要它们能满足这八条公理，我们就可以说：在定义了这样的“加法”与“数乘”后，M 是数域 $\mathbb{F}$ 上的线性空间（向量空间），并且把 M 中的元素都叫做“向量”。

接下来我们通过几个例子来加深一下理解：

例题 2.32 矩阵空间 $\mathbb{F}^{m\times n}$，取数域为 $\mathbb{F}$，我们定义加法为矩阵加法，数乘为矩阵的数乘，那么不难验证 $\mathbb{F}^{m\times n}$ 就是数域 $\mathbb{F}$ 上的线性空间，这里数域为什么取 $\mathbb{F}$？其实我们也可以取 $\mathbb{F}$ 的任意子域，这是因为我们要保证“加法”与“数乘”的封闭性，即我们定义的映射仍是要映入 $\mathbb{F}^{m\times n}$ 内的。比如考虑实矩阵空间 $\mathbb{R}^{m\times n}$，如果取数域为 $\mathbb{Q}$ 或 $\mathbb{R}$ 等等 $\mathbb{R}$ 的子域，那么数域上的任何数和实矩阵相乘仍然是实矩阵，也就是所谓“封闭”。但如果把数域取为复数域 $\mathbb{C}$，复数与实矩阵的数乘并不都是实矩阵，这时候实矩阵空间在数乘运算下就不封闭，所以 $\mathbb{R}^{m\times n}$ 不是数域 $\mathbb{C}$ 上的线性空间，但是是 $\mathbb{Q}$ 或 $\mathbb{R}$ 上的线性空间。

同时这个向量空间里的“向量”是一个个 $m\times n$ 的矩阵，并且容易得到这个空间的维数是 $m\times n$，一组基为 $\{E_{ij}:1\leqslant i\leqslant m,1\leqslant j\leqslant n\}$，很多同学在刚接触到这个空间时会把它与数组空间搞混，这两个空间里“向

量”的定义都是完全不同的，从本质上来说就是两个完全不同的空间，在理解了这一点后，相信大家就不会再弄混这两个空间了。

笔记 但其实也可以把 $\mathbb{F}^{m\times n}$ 看作是数组空间 $\mathbb{F}^{mn}$，我们可以把矩阵的每一列顺次移动到第一列下面，得到一个 mn 维的列向量，并且这样的操作显然是一一对应的。

例题 2.33 连续函数空间 $C[a,b]$，是定义在区间 $[a,b]$ 上的连续函数全体，加法与数乘就是函数的逐点加法与数乘，这也是一个向量空间，这时候空间里的“向量”就是一个个连续函数。那么这个空间的维数是什么？回想一下上面的第五章第 44 题，多项式都是连续的，所以多项式空间是其一个子空间，并且 $\{1,x,x^2,\cdots\}$ 是一组基，也就是说多项式空间是无穷维的，并且是可数无穷。那么 $C[a,b]$ 也是无穷维向量空间。

笔记 实际上 $C[a,b]$ 的基的“个数”是不可数无穷的，也就是维数是不可数无穷。

例题 2.34 对偶空间 V^* 以及二次对偶空间 V^{**}，其中 V 是线性空间，V^* 指的是定义在 V 上的线性函数全体构成的集合，这也是一个向量空间，并且空间里的“向量”是一个个线性函数，我们取 V 的一组基 $\alpha_1,\cdots,\alpha_n$，那么根据第六章第 43 题的推广，对任意给定的 $1\leqslant i\leqslant n$，取 $\beta_j=\delta_{ij}$ 存在线性函数 $\mathscr{A}_i$ 满足 $\mathscr{A}_i(\alpha_j)=\beta_j, j=1,\cdots,n$，也就是 $\mathscr{A}_i$ 只把 α_i 映为 1，把 $\alpha_j(j\neq i)$ 映为 0，那么我们把 $\mathscr{A}_i$ 记为 α_i^*，这也就得到了对偶基。

为什么要这样取对偶基呢？我们把 V^* 上任意一个线性函数写成对偶基的线性组合，即 $\mathscr{A}=\lambda_1\alpha_1^*+\cdots+\lambda_n\alpha_n^*$，任取 V 中元素 $\alpha=\mu_1\alpha_1+\cdots+\mu_n\alpha_n$，那么 $\mathscr{A}(\alpha)=\lambda_1\alpha_1^*(\alpha)+\cdots+\lambda_n\alpha_n^*(\alpha)=\lambda_1\mu_1+\cdots+\lambda_n\mu_n$，可以看出来我们得到了一种类似向量内积的美观的表示形式！

二次对偶空间 V^{**} 就是对偶空间的对偶空间，也就是定义在 V^* 上的线性函数全体构成的集合，V^{**} 中的“向量”也是函数，但是这些函数作用的对象是 V^* 里的线性函数，任意取定 V 中一个元素 α，定义映射 $\alpha^{**}:V^*\to\mathbb{R}, \alpha^{**}(\mathscr{A})=\mathscr{A}(\alpha)$，容易验证线性性，这样我们得到了一个映射 $i:V\to V^{**}, \alpha\to\alpha^{**}$，称这个映射为 V 到 V^{**} 的自然映射或自然嵌入。

例题 2.35 不寻常的加法与数乘例子：考虑 $\mathbb{R}_+=\{x\in\mathbb{R}:x>0\}$，取数域为 $\mathbb{R}$，定义加法为 $x\oplus y=xy$，数乘为 $\lambda\odot x=x^\lambda$，容易验证 $\mathbb{R}_+$ 是数域 $\mathbb{R}$ 上的向量空间，这时候“向量”指的是一个个正实数，并且这里加法的零元是 1，x 的加法负元是 $1/x$，在这个空间里任意不为 1 的正实数都可以作为一组基，也就是维数为 1。

商空间

商空间也是一种特殊的线性空间，引入商空间需要先回顾一下第一章所涉及过的等价关系的定义：

定义 2.27 (等价关系)

若集合 A 上的一个关系 $\sim$ 满足如下三条性质：

(1) 自反性：对于任意 $x\in A$，均有 $x\sim x$；

(2) 对称性：对于任意 $x,y\in A$，若 $x\sim y$，则 $y\sim x$；

(3) 传递性：对于任意 $x,y,z\in A$，若 $x\sim y, y\sim z$，则 $x\sim z$，

则称 $\sim$ 为 A 上的一个等价关系。

等价关系的例子非常多，比如“同班同学关系”、“性别相同关系（我们假定每个人只有一种性别）”都是常见的等价关系。矩阵的相抵、相似等等也都是等价关系。给定集合上的一个等价关系，我们就能将集合划分为一个个非空且互不相交的“等价类”：

定义 2.28 (等价类、商集、商映射)

设 $\sim$ 是集合 A 上的一个等价关系。

(1) 对于任意的 $x\in A$，称 $[x]\triangleq\{y\in A:x\sim y\}$ 为元素 x 所在的等价类，（注意等价类是一个集合，集合中的元素互相等价）。$[x]$ 中任意一个元素都被称为 $[x]$ 的一个代表元。

(2) 称所有等价类的集合 $A/\sim\triangleq\{[x]:x\in A\}$ 为集合 A 在该等价关系下的商集。

(3) 称映射 $p: A \to A/\sim, x \to [x]$ 为商映射（容易看出 p 一定是满射）。

接下来，考虑数域 $\mathbb{F}$ 上的线性空间 V，任取 V 的一个子空间 W，$\forall \alpha, \beta \in V$，若 $\alpha - \beta \in W$，则称 α 与 β 模 W 同余，记作 $\alpha \equiv \beta(\mathrm{mod}\ W)$，那么不难看出模 W 同余是 V 上的一个等价关系，等价类便是模 W 的同余类。记 $A_\alpha = \{\alpha + \gamma : \gamma \in W\}$，显然 $A_\alpha \subset [\alpha]$，而任取 $\beta \in [\alpha]$，记 $\gamma_0 = \beta - \alpha \in W$，则 $\beta = \alpha + \gamma_0 \in A_\alpha$，从而有 $[\alpha] \subset A_\alpha$，所以 $[\alpha] = A_\alpha$，从而我们得到了在这个等价关系下的等价类的表示。

并且我们定义等价类之间的加法和数乘：$[\alpha] + [\beta] = [\alpha + \beta], \lambda[\alpha] = [\lambda\alpha]$，且由于 W 是 V 的子空间，那么也是线性空间，由此容易验证我们定义的加法和数乘不依赖于代表元的选取，即是良定的。

那么根据上述结果，等价类构成的集合也构成了一个线性空间，我们便可以给出商空间的定义：

定义 2.29 (商空间)

对于数域 $\mathbb{F}$ 上的线性空间 V，W 是 V 的一个子空间，记所有模 W 的同余类（等价类）的集合为 V/W，其在等价类的加法与数乘下也构成数域 $\mathbb{F}$ 上的线性空间，称为 V 关于 W 的商空间。

上面的叙述可能比较抽象，下面我们通过一个简单的例子来辅助理解：
考虑二维平面，即 $\mathbb{R}^2$，记 $W = \{(x, 0) : x \in \mathbb{R}\}$，即为 x 轴，则 W 为 $\mathbb{R}^2$ 的一个子空间，我们考虑 $\mathbb{R}^2/W$ 是什么，容易看出 $\alpha - \beta \in W$ 当且仅当 α 与 β 有相同的纵坐标，那么等价类 $[\alpha]$ 就是过 α 且和 x 轴平行的直线。

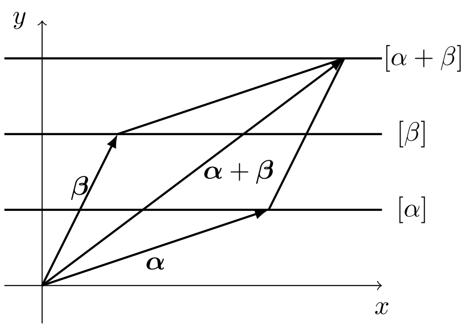


图 2.5: $\mathbb{R}^2/W$

我们在每个等价类中取一个代表元，定义一个映射把每个等价类映回这个代表元，这样就能得到商空间 $\mathbb{R}^2/W$ 到 $\mathbb{R}^2$ 的一个嵌入，比如我们可以取代表元为这些直线与 y 轴的交点，那么我们得到的嵌入映射的像是什么？是 y 轴！而 y 轴与 x 轴 (W) 的直和恰好就是全空间 $\mathbb{R}^2$，这是不是可以给我们一些启发？

事实上，我们有如下的结论：

定理 2.3

设 W 是线性空间 V 的子空间，W' 为其补空间，即 $V = W \oplus W', V/W$ 是 V 模 W 的商空间。定义映射 $\eta : W' \to V/M$ 如下：对任意 $\alpha \in W', \eta(\alpha) = [\alpha]$，则 η 是 W' 到 V/M 的线性同构（一一对应）。

证明 显然 η 是线性映射，任取 $[\alpha] \in V/W$，那么 $\alpha \in V = W \oplus W'$，因此存在唯一的 $\beta \in W, \gamma \in W'$ 满足 $\alpha = \beta + \gamma$，由于 $\alpha - \gamma = \beta \in W$，所以 $\alpha \equiv \gamma\,(\mathrm{mod}\ W)$. 故 $[\gamma] = [\alpha]$. 即对任意 $[\alpha] \in V/W$，存在 $\gamma \in W'$，使得 $\eta(\gamma) = [\gamma] = [\alpha]$，故 η 是满射。

其次，设 $\gamma_1, \gamma_2 \in W'$，且 $\eta(\gamma_1) = \eta(\gamma_2)$，则 $[\gamma_1] = [\gamma_2]$，从而 $\gamma_1 \equiv \gamma_2\,(\mathrm{mod}\ W)$，即 $\gamma_1 - \gamma_2 \in W$，而 $\gamma_1, \gamma_2 \in W'$，故 $\gamma_1 - \gamma_2 \in W'$，因此, $\gamma_1 - \gamma_2 \in W \cap W'$。由于 W' 是 W 的补空间，所以 $W \cap W' = \{0\}$，那么只有 $\gamma_1 - \gamma_2 = 0$，即 $\gamma_1 = \gamma_2$，从而 η 是单射。

因此，η 是线性同构[3]，从而 W' 和 V/W 同构。

由上述定理我们还能得到下面这样一个自然的推论：

[3]两个线性空间线性同构是指存在它们之间的线性双射，将这个双射称为同构映射，简称同构，同构空间具有很多相同的性质。

推论 2.3

设 W 是线性空间 V 的子空间，则 V 关于 W 的商空间 V/W 的维数为 $\dim(V/W) = \dim V - \dim W$.

证明 由定理2.3，$\dim(V/W) = \dim(W') = \dim V - \dim W$，其中 W' 为 W 的补空间。

不变子空间

上次课我们初次学习到了特征值和特征向量的概念，线性映射把它的特征向量映为其常数倍，那么特征向量的任意常数倍也都是特征向量，且相同特征值对应的特征向量的和仍然是该特征值对应的特征向量，那如果我们考虑某个特征值对应的所有特征向量全体构成的集合，这个集合就也是一个线性空间，我们称其为特征子空间。并且特征子空间在映射下的像仍然落在特征子空间里。

那么我们来研究一个更广泛的有类似性质的子空间，即不变子空间：

定义 2.30 (不变子空间)

设 $\mathscr{A}: V \to V$ 是线性变换，U 是 V 的子空间，如果对任意 $\alpha \in U$，都有 $\mathscr{A}(\alpha) \in U$，即 $\mathscr{A}(U) \subset U$，则称 U 为线性变换 $\mathscr{A}$ 的一个不变子空间。

研究不变子空间有什么用处呢？我们取不变子空间上的一组基，并把它们扩充成全空间的基，我们发现线性变换在这组基下的矩阵具有很好的形式：

定理 2.4

设 $\mathscr{A}$ 是 V 上的线性变换，U 是 $\mathscr{A}$ 的不变子空间。设 $\{\alpha_1, \alpha_2, \cdots, \alpha_k\}$ 是 U 的一组基，则 $\mathscr{A}$ 在 V 的基 $\{\alpha_1, \alpha_2, \cdots, \alpha_k, \alpha_{k+1}, \cdots, \alpha_n\}$ 下的方阵是准上三角方阵。

证明 因为 U 是 $\mathscr{A}$ 的不变子空间，所以对 $1 \leqslant i \leqslant k, \mathscr{A}(\alpha_j) \in U$，由于 $\{\alpha_1, \alpha_2, \cdots, \alpha_k\}$ 是 U 的基，因此设 $\mathscr{A}(\alpha_i) = a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + a_{ik}\alpha_k$. 而当 $k+1 \leqslant j \leqslant n$ 时，由于 $\{\alpha_1, \alpha_2, \cdots, \alpha_k, \alpha_{k+1}, \cdots, \alpha_n\}$ 为 V 的基，故设 $\mathscr{A}(\alpha_j) = a_{j1}\alpha_1 + a_{j2}\alpha_2 + \cdots + a_{jn}\alpha_n$，因此：

$$\mathscr{A}(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n) \begin{pmatrix} a_{11} & \cdots & a_{k1} & a_{k+1,1} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ a_{1k} & \cdots & a_{kk} & a_{k+1,k} & \cdots & a_{nk} \\ 0 & \cdots & 0 & a_{k+1,k+1} & \cdots & a_{n,k+1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{k+1,n} & \cdots & a_{nn} \end{pmatrix}, \text{定理得证。}$$

同时反过来也容易发现：如果线性变换 $\mathscr{A}$ 在基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 下的方阵是准上三角的，且其左上角的块是一个 k 阶的方阵，那么由 $\{\alpha_1, \alpha_2, \cdots, \alpha_k\}$ 生成的子空间是 $\mathscr{A}$ 的一个不变子空间。

那么我们就会发现，如果可以把 V 分解成一系列 $\mathscr{A}$ 的不变子空间的直和，那么这些不变子空间的基的并集构成 V 的一组基，并且 $\mathscr{A}$ 在这组基下的矩阵是准对角阵！这个时候，你可能就会对相似标准形的求解产生一些想法，其实求解相似标准形的过程中很重要的一步实质上就是一个不变子空间分解的过程。

2.5.3 补充题目

习题 2.33

将数域 $\mathbb{F}$ 上 n 维 $(n \geqslant 2)$ 数组空间 $\mathbb{F}^n$ 中的每个向量 $\boldsymbol{\alpha} = (a_1, a_2, \cdots, a_n)$ 看作一个具有 n 项的数列，如下集合 W 是否构成 F^n 的一个子空间？如果是，求出它的维数及一组基。

(1) $\mathbb{F}^n$ 中所有的等比数列组成的集合。

(2) $\mathbb{F}^n$ 中所有的等差数列组成的集合。

解 (1) 不是子空间，对加法不封闭，如取等比数列 $(1,q,\cdots,q^{n-1})$ 与 $(-1,q,\cdots,(-1)^n q^{n-1})$ 其和 $(0,2q,0,\cdots)$ 显然不是等比数列，从而不是子空间。

(2) 是子空间，显然任意两个等差数列之和、任意等差数列的常数倍仍是等差数列，且任意等差数列

$$(a,a+d,\cdots,a+(n-1)d)=a(1,1,\cdots,1)+d(0,1,2,\cdots,n-1)$$

都是两个线性无关等差数列 $(1,1,\cdots,1)$ 及 $(0,1,2,\cdots,n-1)$ 的线性组合，从而这两个等差数列构成一组基，那么该子空间的维数是 2.

习题 2.34

设 W_1,W_2 分别是数域 $\mathbb{F}$ 上齐次线性方程组 $x_1+x_2+\cdots+x_n=0$ 与 $x_1=x_2=\cdots=x_n$ 的解空间。求证：$\mathbb{F}^n=W_1\oplus W_2$。

证明 设 $X_0=(x_1,\cdots,x_n)\in W_1\cap W_2$，由 $X_0\in W_2$ 知 $X_0=(x_1,\cdots,x_1)$，由 $X_0\in W_1$ 知 $nx_1=0$，则 $x_1=0$，$X_0=\mathbf{0}$。那么 $W_1\cap W_2=\{0\}$，$W_1+W_2=W_1\oplus W_2$。对每个 $X=(x_1,\cdots,x_n)\in\mathbb{F}^n$，令 $c=\frac{1}{n}(x_1+\cdots+x_n)$，$X_2=(c,\cdots,c)$，$X_1=X-X_2=(x_1-c,\cdots,x_n-c)$，由 $(x_1-c)+\cdots+(x_n-c)=(x_1+\cdots+x_n)-nc=0$ 知 $X_1\in W_1,X_2\in W_2$，而 $X=X_1+X_2\in W_1+W_2$。从而 $\mathbb{F}^n=W_1+W_2=W_1\oplus W_2$.

习题 2.35

设 f 是 $\mathbb{F}^{n\times n}$ 上的线性函数，且 $f(AB)=f(BA)$，对任意 n 阶方阵 A,B 都成立，证明：存在常数 $c\in\mathbb{F}$，使 $f(A)=c\cdot tr(A)$.

证明 取 $\mathbb{F}^{n\times n}$ 上的一组自然基 $\{E_{ij}:1\leqslant i,j\leqslant n\}$，其中 E_{ij} 为 (i,j) 位置元素为 1，其余位置均为 0 的 n 阶方阵。那么对任意的 i,j,k,t，有 $E_{ki}E_{ij}=E_{kj}$，而 $i\neq t$ 时，有 $E_{ki}E_{tj}=O$，因此有：

$$f(E_{ij})=f(E_{ii}E_{ij})=f(E_{ij}E_{ii})=f(O)=O\ (i\neq j);$$

$$f(E_{ii})=f(E_{i1}E_{1i})=f(E_{1i}E_{i1})=f(E_{11})$$

那么对于任意的 n 阶方阵 $A=\sum\limits_{1\leqslant i,j\leqslant n}a_{ij}E_{ij}$，有 $f(A)=\sum\limits_{1\leqslant i,j\leqslant n}a_{ij}f(E_{ij})=(a_{11}+\cdots+a_{nn})f(E_{11})=c\cdot tr(A)$. 其中 $c=f(E_{11})$ 为常数，这样就证明了命题。

习题 2.36

设 $\mathscr{A}:U\to U$ 是线性变换，证明存在线性变换 $\mathscr{B}:U\to U$，使得 $\mathscr{A}\mathscr{B}=0$，且 $rank(\mathscr{A})+rank(\mathscr{B})=\dim U$.

证明 取 U 的一组基 $\{\alpha_1,\cdots,\alpha_n\}$，设 $\mathscr{A}$ 在这组基下的矩阵是 A，那么存在可逆方阵 P,Q 使 $A=P\begin{pmatrix}I_r&O\\O&O\end{pmatrix}Q$，$rank(\mathscr{A})=rank(A)=r$，取 $B=Q^{-1}\begin{pmatrix}O&O\\O&I_{n-r}\end{pmatrix}$，则 $AB=O$，那么由第六章作业第 43 题，可以定义线性变换 $\mathscr{B}$，满足其在基 $\{\alpha_1,\cdots,\alpha_n\}$ 下的矩阵为 B，则 $rank(\mathscr{B})=rank(B)=n-r$，那么有：

$$\mathscr{A}\mathscr{B}(\alpha_1,\cdots,\alpha_n)=\mathscr{A}(\alpha_1,\cdots,\alpha_n)B=(\alpha_1,\cdots,\alpha_n)AB=O$$

于是 $\mathscr{A}\mathscr{B}=0$，且 $rank(\mathscr{A})+rank(\mathscr{B})=r+n-r=n=\dim U$，即证。

笔记 涉及线性变换的秩的题目一般都是转化到矩阵会更好处理。这道题其实就是想办法让 0 和 1 交错开.

习题 2.37

已知线性空间 V 上的线性变换 $\mathscr{A}$ 满足条件 $\mathscr{A}^2=\mathscr{A}$，求证：

(1) $V=\operatorname{Im}\mathscr{A}\oplus\ker A$；

(2) $\mathscr{A}$ 在任何一组基下的矩阵 A 满足条件 $rank(A)=tr(A)$；

(3) 在适当的基下将 V 的向量用坐标表示，可以使 $\mathscr{A}$ 具有投影变换的形式，即：

$$(x_1,\cdots,x_r,x_{r+1},\cdots,x_n)^T\to(x_1,\cdots,x_r,0,\cdots,0)^T$$

证明 (1) 设 $\alpha\in\operatorname{Im}\mathscr{A}\cap\ker\mathscr{A}$，则由 $\alpha\in\operatorname{Im}\mathscr{A},\exists\beta\in V$ 使 $\alpha=\mathscr{A}(\beta)$。再由 $\alpha\in\ker\mathscr{A}$ 知 $\mathscr{A}(\alpha)=0$，于是 $\alpha=\mathscr{A}(\beta)=\mathscr{A}^2(\beta)=\mathscr{A}(\mathscr{A}(\beta))=\mathscr{A}(\alpha)=0$。那么 $\operatorname{Im}\mathscr{A}\cap\ker\mathscr{A}=\{0\}$。故 $\operatorname{Im}\mathscr{A}+\ker\mathscr{A}=\operatorname{Im}\mathscr{A}\oplus\ker\mathscr{A}$。而 $\dim V=\dim\operatorname{Im}\mathscr{A}+\dim\ker\mathscr{A}$，故 $V=\operatorname{Im}\mathscr{A}\oplus\ker\mathscr{A}$.

(2) 由 $V=\operatorname{Im}\mathscr{A}\oplus\ker\mathscr{A}$ 知 $\operatorname{Im}\mathscr{A}$ 的任意一组基 $S_1=\{\alpha_1,\cdots,\alpha_r\}$ 与 $\ker\mathscr{A}$ 的任一组基 $S_0=\{\beta_1,\cdots,\beta_k\}$ 的并 $M_1=\{\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_k\}$ 是 V 的一组基，其中每个 $\beta_j\in\ker\mathscr{A}$ 满足 $\mathscr{A}(\beta_j)=0$，每个 $\alpha_i\in\operatorname{Im}\mathscr{A}$ 能写成 $\mathscr{A}(\gamma_i)$ 的形式 $(\gamma_i\in V)$，于是 $\mathscr{A}(\alpha_i)=\mathscr{A}^2(\gamma_i)=\mathscr{A}(\gamma_i)=\alpha_i$，那么有：

$$\mathscr{A}(\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_k)=(\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_k)\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}$$

即 $\mathscr{A}$ 在这组基下的矩阵为 $A_1=diag(I_r,O)$，其满足 $tr(A_1)=r=rank(A_1)$。而在任意一组基 M 下的矩阵 $A=P^{-1}A_1P$，其中可逆阵 P 是 M_1 到 M 的过渡矩阵。那么有 $rank(A)=rank(P^{-1}A_1P)=rank(A_1)$，且 $tr(A)=tr(P^{-1}A_1P)=tr(A_1)=rank(A_1)=rank(A)$.

(3) 由 (2)，$\mathscr{A}$ 在基 M_1 下的矩阵为 $A_1=diag(1,\cdots,1,0,\cdots,0)$。那么坐标为 $X=(x_1,\cdots,x_n)^T$ 的向量被 $\mathscr{A}$ 映到坐标为 $A_1X=(x_1,\cdots,x_r,0,\cdots,0)^T$ 的向量，即具有所要求的投影变换形式。

2.6 第六次习题课

2.6.1 不可对角化矩阵

习题 2.38

设

$$J_n(\lambda)=\begin{pmatrix}\lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda\end{pmatrix}_{n\times n}.$$

证明：$J_n(\lambda)$ 不可对角化，其中 $n>1$.

证明 注意到 $J_n(\lambda)$ 的特征值为 λ，若 $J_n(\lambda)$ 可对角化，则矩阵 $J_n(\lambda)$ 有 n 个线性无关的特征向量，即方程组 $(\lambda I-J_n(\lambda))x=0$ 的解空间维数为 n, 但注意到 $r(\lambda I-J_n(\lambda))=n-1$, 故解空间维数为 1，矛盾！

注 以上证明过程说明：$J_n(\lambda)$ 可对角化当且仅当 $n=1$.

习题 2.39

设 $A=\begin{pmatrix}0 & 1\\ -2 & 2\end{pmatrix}$, 证明:$A$ 在实数域上不可对角化.

证明 特征多项式 $\det(\lambda I-A)=\lambda^2-2\lambda+2$ 在实数域上无解，因此矩阵 A 在实数域上不可对角化.

注 注意到矩阵 A 在复数域上有两个不同的根，因此 A 在复数域上可以对角化.

笔记 以上两个例子说明，矩阵是否可以对角化不仅与矩阵本身有关，而且和数域也有关.

2.6.2 矩阵多项式，逆矩阵及伴随矩阵的特征值

习题 2.40

设 $A\in\mathbb{F}^{n\times n}, \lambda_1,\lambda_2,\cdots,\lambda_n$ 是矩阵 A 的全部特征值，$f(x)$ 是一个多项式. 证明：$f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)$ 是 $f(A)$ 的全部特征值.

证明 存在一个复可逆矩阵 P, 使得

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

通过矩阵运算不难得到

$$P^{-1}f(A)P=f(P^{-1}AP)=\begin{pmatrix}f(\lambda_1) & * & \cdots & * \\ 0 & f(\lambda_2) & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & f(\lambda_n)\end{pmatrix}.$$

于是 $f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)$ 是 $f(A)$ 的全部特征值.

习题 2.41

设 $A \in \mathbb{F}^{n\times n}$, 且 A 可逆, $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是矩阵 A 的全部特征值. 证明: $\lambda_1^{-1}, \lambda_2^{-1}, \cdots, \lambda_n^{-1}$ 是 A^{-1} 的全部特征值.

证明 存在一个复可逆矩阵 P, 使得

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

注意到 $\det(A) = \lambda_1\lambda_2\cdots\lambda_n \neq 0$, 因此对任意 $\lambda_i \neq 0$. 由于上三角矩阵的逆矩阵仍然是上三角矩阵，于是

$$P^{-1}A^{-1}P = \begin{pmatrix} \lambda_1^{-1} & * & \cdots & * \\ 0 & \lambda_2^{-1} & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^{-1} \end{pmatrix}.$$

故 $\lambda_1^{-1}, \lambda_2^{-1}, \cdots, \lambda_n^{-1}$ 是 A^{-1} 的全部特征值.

习题 2.42

设 $A \in \mathbb{F}^{n\times n}$, $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是矩阵 A 的全部特征值,A^* 是矩阵 A 的伴随矩阵.
证明: $\prod\limits_{i\neq 1} \lambda_i, \prod\limits_{i\neq 2} \lambda_i, \cdots, \prod\limits_{i\neq n} \lambda_i$ 是 A^* 的全部特征值.

证明 存在一个复可逆矩阵 P, 使得

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

上式两边同时取伴随矩阵，有存在一个复可逆矩阵 P, 使得

$$P^*A^*(P^*)^{-1} = (P^{-1}AP)^* = \begin{pmatrix} \prod\limits_{i\neq 1}\lambda_i & * & \cdots & * \\ 0 & \prod\limits_{i\neq 2}\lambda_i & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \prod\limits_{i\neq n}\lambda_i \end{pmatrix}.$$

因此 $\prod\limits_{i\neq 1} \lambda_i, \prod\limits_{i\neq 2} \lambda_i, \cdots, \prod\limits_{i\neq n} \lambda_i$ 是 A^* 的全部特征值.

习题 2.43

设 $A \in \mathbb{F}^{n\times n}$,$g(x)$ 为多项式, λ 为 A 的特征值. 证明: 若 $g(A) = 0$, 则 $g(\lambda) = 0$.

证明 设 x 是 A 的属于 λ 的特征向量, 即 $Ax = \lambda x$. 于是 $g(\lambda)x = g(A)x = 0$, 又 $x \neq 0$, 因此 $g(\lambda) = 0$.

2.6.3 矩阵对角化的应用

习题 2.44

设矩阵 $A=\begin{pmatrix}1&-1&1\\2&4&-2\\-3&-3&5\end{pmatrix}$, 求 A^n.

解 $\det(\lambda I-A)=(\lambda-2)^2(\lambda-6)$, 因此 A 的特征值为 2, 6. 解线性方程组 $(2I-A)x=0$ 与 $(6I-A)x=0$, 分别得到三个线性无关的特征向量 $(-1,1,0)^T,(1,0,1)^T,(1,-2,3)^T$. 因此可设 $P=\begin{pmatrix}-1&1&1\\1&0&-2\\0&1&3\end{pmatrix}$, 于是 $B=P^{-1}AP=\mathrm{diag}(2,2,6)$. 故

$$A^n=PB^nP^{-1}=\frac{1}{4}\begin{pmatrix}5\cdot 2^n-6^n&2^n-6^n&6^n-2^n\\2\cdot 6^n-2^{n+1}&2\cdot 6^n+2^{n+1}&2^{n+1}-2\cdot 6^n\\3\cdot 2^n-3\cdot 6^n&3\cdot 2^n-3\cdot 6^n&2^n+3\cdot 6^n\end{pmatrix}.$$

注 小测第二题可以同上作类似处理.

习题 2.45

设数列 a_n 的递推公式为 $a_{n+2}=a_{n+1}+a_n$, 其中 $a_0=0,a_1=1$. 试求 a_n 的通项公式.

解 注意到

$$\begin{pmatrix}a_{n+1}\\a_n\end{pmatrix}=\begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}a_n\\a_{n-1}\end{pmatrix}.$$

记 $A=\begin{pmatrix}1&1\\1&0\end{pmatrix}$, 于是

$$\begin{pmatrix}a_{n+1}\\a_n\end{pmatrix}=A\begin{pmatrix}a_n\\a_{n-1}\end{pmatrix}=A^2\begin{pmatrix}a_{n-1}\\a_{n-2}\end{pmatrix}=\cdots=A^n\begin{pmatrix}a_1\\a_0\end{pmatrix}=A^n\begin{pmatrix}a_1\\a_0\end{pmatrix}=A^n\begin{pmatrix}1\\0\end{pmatrix}.$$

进而问题转化为求 A^n. 由于 $\det(\lambda I-A)=\lambda^2-\lambda-1$, 故矩阵 A 的两个特征值分别为 $\lambda_1=\frac{1+\sqrt{5}}{2}$,$\lambda_2=\frac{1-\sqrt{5}}{2}$. 对应的特征向量分别是 $(\frac{1+\sqrt{5}}{2},1)^T,(\frac{1-\sqrt{5}}{2},1)^T$ 可以记 $P=\begin{pmatrix}\frac{1+\sqrt{5}}{2}&\frac{1-\sqrt{5}}{2}\\1&1\end{pmatrix}$. 类似于上一题过程，可以算得

$$A^n=\frac{1}{\sqrt{5}}\begin{pmatrix}(\frac{1+\sqrt{5}}{2})^{n+1}-(\frac{1-\sqrt{5}}{2})^{n+1}&(\frac{1+\sqrt{5}}{2})^n-(\frac{1-\sqrt{5}}{2})^n\\(\frac{1+\sqrt{5}}{2})^n-(\frac{1-\sqrt{5}}{2})^n&(\frac{1+\sqrt{5}}{2})^{n-1}-(\frac{1-\sqrt{5}}{2})^{n-1}\end{pmatrix}.$$

于是有 $a_n=\frac{1}{\sqrt{5}}((\frac{1+\sqrt{5}}{2})^n-(\frac{1-\sqrt{5}}{2})^n)$.

注 我们可以用这个方法处理第五周作业第 16 题的第三问（第三章习题第 16 题第三问），注意到 $D_n=2\cos\theta D_{n-1}-D_{n-2}$，且 $D_1=2\cos\theta,D_2=4\cos^2\theta-1$, 同样可以利用矩阵对角化求 D_n 的通项公式.

作为变式，我们给出下面例题，请读者自己完成证明.

习题 2.46

有两个数列 $\{a_n\}$,$\{b_n\}$,$a_1=1,b_1=-1,a_n=a_{n-1}+2b_{n-1},b_n=-a_{n-1}+4b_{n-1}$. 求证：

$$a_n=2^{n+1}-3^n,b_n=2^n-3^n.$$

2.6.4 Jordan 标准形及其应用

首先我们不加证明的给出下面定理

定理 2.5

任何一个复方阵 A 均相似于一个 *Jordan* 矩阵 J, 其中 $J=\mathrm{diag}(J_{r_1}(\lambda_1),\cdots,J_{r_s}(\lambda_s))$. 若不计 *Jordan* 块次序，则 J 是唯一的.(也即在相似意义下唯一)

注 此时 $\lambda_1,\cdots,\lambda_s$ 即为矩阵 A 的全部特征值.

推论 2.4

复方阵 A 可对角化当且仅当他的每个 *Jordan* 块是一阶的.

证明 由矩阵的 *Jordan* 标准形立明.

推论 2.5

设 λ_0 是 n 阶矩阵 A 的特征值. 则对任意的正整数 k，特征值为 λ_0 的 k 阶 *Jordan* 块出现的个数为

$$r((A-\lambda_0 I)^{k-1})+r((A-\lambda_0 I)^{k+1})-2r((A-\lambda_0 I)^{k}),$$

其中 $(A-\lambda_0 I)^0=I$.

证明 存在可逆矩阵 P, 使得 $P^{-1}AP=\mathrm{diag}(J_{r_1}(\lambda_1),\cdots,J_{r_s}(\lambda_s))$. 注意到

$$(A-\lambda_0 I)^k=P\mathrm{diag}(J_{r_1}^k(\lambda_1-\lambda_0),\cdots,J_{r_s}^k(\lambda_s-\lambda_0))P^{-1}.$$

因此 $r((A-\lambda_0 I)^k)=\sum\limits_{i=1}^{s}r(J_{r_i}^k(\lambda_i-\lambda_0))$. 当 $\lambda_i\neq\lambda_0$ 时，$r(J_{r_i}^k(\lambda_i-\lambda_0))=r_i$; 当 $\lambda_i=\lambda_0$ 时，若 $r_i<k$, 则 $r((A-\lambda_0 I)^k)=0$, 若 $r_i\geq k$, 则 $r((A-\lambda_0 I)^k)=r_i-k$. 因此 $r((A-\lambda_0 I)^{k-1})-r((A-\lambda_0 I)^k)$ 等于特征值为 λ_0 且阶数大于等于 k 的 *Jordan* 块个数. 同理 $r((A-\lambda_0 I)^k)-r((A-\lambda_0 I)^{k+1})$ 等于特征值为 λ_0 且阶数大于等于 $k+1$ 的 *Jordan* 块个数. 从而特征值为 λ_0 的 k 阶 *Jordan* 块出现的个数为

$$\begin{aligned}&r((A-\lambda_0 I)^{k-1})-r((A-\lambda_0 I)^k)-(r((A-\lambda_0 I)^k)-r((A-\lambda_0 I)^{k+1}))\\&=r((A-\lambda_0 I)^{k-1})+r((A-\lambda_0 I)^{k+1})-2r((A-\lambda_0 I)^k).\end{aligned}$$

推论 2.6

设 $A,B\in\mathbb{C}^{n\times n}$, 则 A,B 相似等价于对 A,B 的任意特征值 λ_0 及任意正整数 k，有 $r((A-\lambda_0 I)^k)=r((B-\lambda_0 I)^k)$.

证明 必要性显然，下面我们证明充分性.
由推论2.5知矩阵 A,B 特征值为 λ_0 的 k 阶 *Jordan* 块出现的个数相同，因此 A,B 相似于同一个 *Jordan* 矩阵，故 A,B 相似.

推论 2.7

设 $A\in\mathbb{C}^{n\times n}$, 若存在正整数 k，使得 $r(A^k)=r(A^{k+1})$. 则 $r(A^k)=r(A^{k+1})=r(A^{k+2})=\cdots$.

证明 (法一) 存在可逆矩阵 P, 使得 $P^{-1}AP=\mathrm{diag}(J_{r_1}(\lambda_1),\cdots,J_{r_s}(\lambda_s))$. 因此 $r(A^k)=\sum\limits_{i=1}^{s}r(J_{r_i}^k(\lambda_i))$. 当 $\lambda_i\neq 0$ 时，$r(J_{r_i}^k(\lambda_i))=r(J_{r_i}^{k+1}(\lambda_i))$; 当 $\lambda_i=0$ 时，若 $k\geq r_i$, 则 $r(J_{r_i}^k(\lambda_i))=r(J_{r_i}^{k+1}(\lambda_i))=0$，若 $k<r_i$ 则 $r(J_{r_i}^k(\lambda_i))>r(J_{r_i}^{k+1}(\lambda_i))$. 因此 $r(A^k)=r(A^{k+1})$ 时，k 一定大于等于特征值为 0 的 *Jordan* 块阶数，分析每个 *Jordan* 块的秩 (考虑特征值是否非零)，自然有 $r(A^k)=r(A^{k+1})=r(A^{k+2})=\cdots$.

(法二) 设 $V_{A^k}=\{x\in\mathbb{F}^{n\times 1}:A^kx=0\}$，同理有 $V_{A^{k+1}}$. 易知 $V_{A^k}\subseteq V_{A^{k+1}}$, 又 $dimV_{A^k}=n-r(A^k)$ 及 $r(A^k)=r(A^{k+1})$, 知 $V_{A^k}=V_{A^{k+1}}$. 因此事实上，证明 $r(A^{k+1})=r(A^{k+2})$ 等价于证明 $V_{A^{k+1}}=V_{A^{k+2}}$. 又

显然有 $V_{A^{k+1}} \subseteq V_{A^{k+2}}$, 下面只需证明 $V_{A^{k+1}} \supseteq V_{A^{k+2}}$. 对任意 $x \in V_{A^{k+2}}$, 有 $A^{k+1}(Ax) = A^{k+2}x = 0$, 因此 $Ax \in V_{A^{k+1}} = V_{A^k}$, 故 $A^{k+1}x = A^k(Ax) = 0$, 因此 $x \in V_{A^{k+1}}$, 于是 $V_{A^{k+1}} \supseteq V_{A^{k+2}}$，命题得证.

推论 2.8

设 $A \in \mathbb{C}^{n\times n}$, 则 A 可对角化等价于对任意 λ，$r(A-\lambda I) = r(A-\lambda I)^2$.

证明 必要性显然，由推论2.5及推论2.7必要性也显然.

习题 2.47

设 $A = \begin{pmatrix} 2 & -2 & 3 \\ 10 & -4 & 5 \\ 5 & -4 & 6 \end{pmatrix}$, 求 A 的 *Jordan* 标准形.

解 $\det(\lambda I - A) = (\lambda-1)^2(\lambda-2)$. 经计算易得 $r(2I-A) = r(2I-A)^2 = 2$，由推论2.5及推论2.7特征值为 2 的 *Jordan* 块只有一个且为 1 阶.$r(I-A) = 2, r(I-A)^2 = r(I-A)^3 = 1$, 由推论2.5 及推论2.7知特征值为 1 的 *Jordan* 块只有一个且为 2 阶. 因此 A 的 *Jordan* 形为 $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

2.7 第七次习题课

2.7.1 正交投影与最小二乘法

正交投影

在三维几何空间中，设 U 是过原点的一个平面 (则其为一个二维子空间)，l 为过原点且垂直于 U 的直线 (其为一个一维子空间)，则 l 中的每个向量与 U 中的每个向量都正交，称 l 为 U 的正交补，类似地，定义一般欧几里得空间的一个子空间的正交补：

定义 2.31 (正交补)
设 V 是一个欧式空间，S 是 V 的一个非空子集，则 V 中与 S 里每个向量都正交的所有向量组成的集合称为 S 的正交补，记作 $S^{\perp}$. 即

$$S^{\perp} \overset{def}{=} \{\alpha \in V | (\alpha, \beta) = 0, \forall \beta \in S\}.$$

♣

容易验证 $S^{\perp}$ 是 V 的一个子空间.

定理 2.6
设 U 是欧式空间 V 的一个有限维子空间，则有 $V = U \oplus U^{\perp}$.

♡

证明 首先说明 $U \cap U^{\perp} = 0$，这是因为：$\forall a \in U \cap U^{\perp}, (a, a) = 0 \Rightarrow a = 0$.

其次说明 $\forall \alpha \in V$，有 $\alpha = \alpha_1 + \alpha_2$，其中 $\alpha_1 \in U, \alpha_2 \in U^{\perp}$：

取 U 的一组标准正交基 $e_1, e_2, \cdots, e_m$，令 $\alpha_1 = \sum\limits_{i=1}^{m} (\alpha, e_i) e_i \in U$，则有 $\alpha_2 = \alpha - \alpha_1$,

$(\alpha_2, e_i) = (\alpha, e_i) - (\alpha_1, e_i) = (\alpha, e_i) - (\alpha, e_i) = 0$, 即 $\alpha_2 \in U^{\perp}$, 得证.

从上述证明过程我们可以看到，V 中的任意向量 α 可以唯一地表示为 $\alpha = \alpha_1 + \alpha_2, \alpha_1 \in U, \alpha_2 \in U^{\perp}$，从而：

定义 2.32 (正交投影)
V 上的线性变换 $P_U : \alpha \to \alpha_1$ 称为 V 在 U 上的正交投影，α_1 称为 α 在 U 上的正交投影.

♣

在三维几何空间中，求一个向量 α 到一个过原点平面的正交投影即从这个向量端点向平面做垂线得到垂足，连接原点与垂足的向量即为正交投影. 同时可以看到，平面内其他向量到 α 的距离都大于正交投影到 α 的距离. 类似地，对于一般欧式空间，有：

定理 2.7
对于 $\alpha \in V, \alpha_1 \in U, \alpha_1$ 是 α 在 U 上的正交投影当且仅当

$$d(\alpha, \alpha_1) \leq d(\alpha, \gamma), \ \forall \gamma \in U.$$

♡

证明 必要性：设 α_1 为 α 在 U 上的正交投影，则 $\alpha - \alpha_1 \in U^{\perp}$. 对 $\forall \gamma \in U$, 有 $\alpha_1 - \gamma \in U$,$(\alpha - \alpha_1, \alpha_1 - \gamma) = 0$, 则有 $|\alpha - \alpha_1|^2 + |\alpha_1 - \gamma|^2 = |\alpha - \gamma|^2 \Rightarrow d(\alpha, \alpha_1) \leq d(\alpha, \gamma)$. 等号成立当且仅当 $\gamma = \alpha_1$.

充分性：设 β 为 α 在 U 上的正交投影，则由必要性知 $d(\alpha, \beta) \leq d(\alpha, \alpha_1)$，又在式 (2.67) 中取 $\gamma = \beta$ 得 $d(\alpha, \alpha_1) \leq d(\alpha, \beta)$, 结合两式得 $d(\alpha, \alpha_1) = d(\alpha, \beta)$, 这等价于 $\alpha_1 = \beta$.

最小二乘法

¶ 离散数据的多项式拟合

对于观察/测量得到的一组离散数据$(x_i, y_i), i = 1, 2, \cdots, m$，我们通常采用插值/拟合的方式来构造它的逼近函数 $\phi(x)$，插值的原则是 $\phi(x_i) = y_i$，在此不做详述；而为给出拟合遵循的原则，首先给出向量范数的定义：

定义 2.33 (范数)

对 $\forall X \in \mathbb{R}^n$，它对应一个实数 $||X||$，满足下面三条性质：
(1) (非负性) $||X|| \geq 0$，等号成立当且仅当 $X = 0$;
(2) (齐次性) $\forall \alpha \in \mathbb{R}$，有 $||\alpha X|| = |\alpha|||X||$;
(3) (三角不等式) $||X+Y|| \leq ||X|| + ||Y||$,
则称 $||X||$ 为 X 的范数.

易见，欧式空间中定义的向量模长即为向量的一种范数，称为**欧几里得范数**，记作 $||X||_2$.

利用不同的范数可以定义出向量 $Z = (\phi(x_1), \phi(x_2), \cdots, \phi(x_m))$ 与 $Y = (y_1, y_2, \cdots, y_m)$ 之间的误差，进而给出拟合方法，特别地，对欧几里得范数 $||X||_2$，有：

定义 2.34

拟合的误差平方和记为 $Q = ||Z - Y||_2^2 = \sum\limits_{i=1}^{m}(\phi(x_i) - y_i)^2$. 按误差平方和达到极小来构造拟合曲线的方法称为最小二乘法.

定义 2.35

多项式拟合：当拟合函数 $\phi(x)$ 形如 $a_0 + a_1 x + \cdots + a_n x^n$ 时，拟合问题称为 n 次多项式拟合，即求 $\alpha = (a_0, a_1, \cdots, a_n)^T$ 使得 $Q(\alpha) = \sum\limits_{i=1}^{m}(a_0 + a_1 x_i + \cdots + a_n x_i^n - y_i)^2$ 取到最小. 我们将会在下一节讨论具体做法.

¶ 矛盾方程组

对于多项式拟合问题，取矩阵 $A = \begin{pmatrix} 1 & x_1 & \cdots & x_1^n \\ 1 & x_2 & \cdots & x_2^n \\ \vdots & \vdots & & \vdots \\ 1 & x_m & \cdots & x_m^n \end{pmatrix}$，有 $Q(\alpha) = ||A\alpha - Y||_2^2$. 当方程组 $A\alpha = Y$ 有解时，拟合函数过所有数据点，误差平方和为 0.

然而，实际的大多数情况数据点或多或少会存在误差 (“噪声”)，这导致方程组 $A\alpha = Y$ 无解，这时我们称它为**矛盾方程组**，使 $||A\alpha - Y||_2$ 最小的 α 称为矛盾方程组的**最小二乘解**.

我们下面讨论对一般的线性方程组，求其最小二乘解：

设 $A \in \mathbb{R}^{m\times n}$，线性方程组 $Ax = \beta$ 的最小二乘解设为 α，A 的列向量组记为 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$，它们张成的子空间记为 U，则有：

$$
\begin{aligned}
& \alpha \text{ 是 } Ax = \beta \text{ 的最小二乘解} \\
\iff & |A\alpha - \beta| \leq |A\alpha' - \beta| \forall \alpha' \in \mathbb{R}^n \\
\iff & |A\alpha - \beta| \leq |\gamma - \beta|, \forall \gamma \in U \\
\iff & A\alpha \text{ 是 } \beta \text{ 在 } U \text{ 上的正交投影} \\
\iff & \alpha_j^T(\beta - A\alpha) = 0,\ j = 1, 2, \cdots, n \\
\iff & A^T(\beta - A\alpha) = 0 \\
\iff & A^T A\alpha = A^T\beta \\
\iff & \alpha \text{ 是 } A^T Ax = A^T\beta (\text{称为方程组 } Ax = \beta \text{ 的法方程}) \text{ 的解.}
\end{aligned}
$$

进一步地，由于 $rank(A^T A) \leq rank(A^T A, A^T\beta) \leq rank(A^T) = rank(A^T A)$，则 $rank(A^T A) = rank(A^T A, A^T\beta)$，即法方程一定有解.

综上，我们将求一个线性方程组的最小二乘解问题等价为了求解法方程的问题.

回到多项式拟合问题，取相应的矩阵 A，可得其对应的法方程为：

$$\begin{pmatrix} m & \sum_{i=1}^m x_i & \cdots & \sum_{i=1}^m x_i^n \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 & \cdots & \sum_{i=1}^m x_i^{n+1} \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^m x_i^n & \sum_{i=1}^m x_i^{n+1} & \cdots & \sum_{i=1}^m x_i^{2n} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i y_i \\ \vdots \\ \sum_{i=1}^m x_i^n y_i \end{pmatrix} \tag{2.67}$$

实际在进行离散数据拟合的时候，数据点的个数 m 一般远大于 n，需要借助计算机来得到矩阵元并进一步通过算法求解法方程组.

例题 2.36 对于给定的一组数据点 $(x_i, y_i), i = 1, 2, \cdots, m$ 进行线性拟合.

解 列出相应的法方程为

$$\begin{pmatrix} m & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i y_i \end{pmatrix} \tag{2.68}$$

解之得 $a_0 = \dfrac{(\sum_{i=1}^m x_i^2)(\sum_{i=1}^m y_i) - (\sum_{i=1}^m x_i)(\sum_{i=1}^m x_i y_i)}{m\sum_{i=1}^m x_i^2 - (\sum_{i=1}^m x_i)^2}, a_1 = \dfrac{m\sum_{i=1}^m x_i y_i - (\sum_{i=1}^m x_i)(\sum_{i=1}^m y_i)}{m\sum_{i=1}^m x_i^2 - (\sum_{i=1}^m x_i)^2}.$

2.7.2 酉空间

酉空间与酉变换

考虑定义在复数域上的线性空间 V，为引进度量，不能单纯地定义内积为：$(\alpha, \beta) = \alpha^T \beta$，因为这样无法保证内积的正定性，从而无法给出长度的概念. 因此，我们舍弃实内积空间关于内积的对称性，改为如下定义：

定义 2.36 (酉空间)

[a]在复数域上线性空间 V 中定义一个二元函数 (α, β)，若它满足下列四条性质:$\forall \alpha, \beta, \gamma \in V, k \in \mathbb{C}$, 有

(1) (Hermite 性) $(\alpha, \beta) = \overline{(\beta, \alpha)}$;

(2) (线性性 1) $(\alpha + \gamma, \beta) = (\alpha, \beta) + (\gamma, \beta)$;

(3) (线性性 2) $(k\alpha, \beta) = k(\alpha, \beta)$;

(4) (正定性) $(\alpha, \alpha) \geq 0$, 取等当且仅当 $\alpha = 0$,

则称这个二元函数为 V 上的一个内积，定义了内积的复线性空间 V 称为酉空间.

[a]此处的定义与课本稍有区别

♣

由内积的 Hermite 性及对第一个变量的线性，可得

$$(\alpha, k_1\beta + k_2\gamma) = \overline{(k_1\beta + k_2\gamma, \alpha)} = \overline{k_1(\beta, \alpha)} + \overline{k_2(\gamma, \alpha)} = \bar{k}_1(\alpha, \beta) + \bar{k}_2(\alpha, \gamma).$$

类似于欧几里得空间，可以在酉空间中给出长度、角度 ($< \alpha, \beta > \overset{def}{=} arccos \frac{|(\alpha,\beta)|}{|\alpha||\beta|}$)、正交、距离等度量概念，酉空间也同样有 Cauchy 不等式、三角不等式、勾股定理等，也可以仿照 Schmidt 正交化过程构造标准正交基.

定义 2.37 (酉矩阵)

复数域上的 n 级矩阵 P 若满足 $P^*P = PP^* = I$，则称 P 为酉矩阵.

其中 $P^* = \bar{P}^T$ 称 P 的共轭转置 (厄米共轭)，有时为了区别于伴随矩阵，也将其记为 $P^H, P^\dagger$.

♣

例题 2.37 我们曾在课本习题三.17. 求得所有的二阶酉矩阵.

定理 2.8

n 级复方阵 P 为酉矩阵 $\Longleftrightarrow$ P 的行 (列) 向量组构成酉空间的一组标准正交基.

命题 2.18

酉矩阵的性质:

(1) 单位阵是酉方阵;

(2) A, B 为同阶酉方阵，则 AB 也是酉方阵;

(3) 设 A 是酉方阵，则 $\bar{A}, A^T, A^{-1}$ 也是酉方阵;

(4) 酉方阵的行列式、特征值模长均为 1.

定义 2.38 (酉变换)

酉空间 V 中保持内积的线性变换 $\mathscr{A}$ 称为酉空间的一个酉变换，即:

$$(\mathscr{A}\alpha, \mathscr{A}\beta) = (\alpha, \beta),\ \forall \alpha, \beta \in V.$$

定理 2.9

下列命题等价:

(1) $\mathscr{A}$ 是酉变换;

(2) $\mathscr{A}$ 保内积;

(3) $\mathscr{A}$ 保模长;

(4) $\mathscr{A}$ 把标准正交基变为标准正交基;

(5) $\mathscr{A}$ 在标准正交基下的矩阵是酉矩阵.

证明 完全类似实内积空间的正交变换.

定理 2.10

设 $\mathscr{A}$ 是酉空间 V 上的酉变换，则 V 中存在一个标准正交基，使得 $\mathscr{A}$ 在此基下的矩阵为对角阵，且主对角元都是模长为 1 的复数.

证明 对酉空间的维数 n 进行归纳:

$n = 1$ 时酉空间的基为 $\mathscr{A}$ 的特征向量，命题成立.

假设对 $n-1$ 维酉空间命题成立，对于 n 维酉空间：取 $\mathscr{A}$ 的一个特征值 λ_1，其模长为 1，设 η_1 是 $\mathscr{A}$ 属于特征值 λ_1 的一个单位特征向量，则 $<\eta_1>$ 是 $\mathscr{A}$ 的一个不变子空间.$V = <\eta_1> \oplus <\eta_1>^\perp$.$<\eta_1>^\perp$ 也是 $\mathscr{A}$ 的一个不变子空间 (证明类似于正交变换)，则 $\mathscr{A}| <\eta_1>^\perp$ 是 $<\eta_1>^\perp$ 上的酉变换，根据归纳假设，存在 $<\eta_1>^\perp$ 中的一个标准正交基 $\eta_2, \cdots, \eta_n$ 使得 $\mathscr{A}| <\eta_1>^\perp$ 在这组基下的矩阵为对角矩阵 $diag\{\lambda_2, \cdots, \lambda_n\}$，且 $\lambda_2, \cdots, \lambda_n$ 模为 1. 于是 $\mathscr{A}$ 在标准正交基 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 下的矩阵为对角阵 $diag\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$，且对角元模长都为 1. 由归纳法原理命题得证.

推论 2.9

n 级酉矩阵 A 一定酉相似于一个对角矩阵，其对角元模长均为 1.

证明 只需取 A 对应的酉变换在定理 2.3 中对应的对角阵即可.

厄米 (Hermite) 变换与厄米矩阵

类似于欧式空间的实对称变换，我们定义厄米变换:

定义 2.39

酉空间 V 上的变换 $\mathscr{A}$ 满足：

$$(\mathscr{A}\alpha, \beta) = (\alpha, \mathscr{A}\beta),\ \forall \alpha, \beta \in V,$$

则称 $\mathscr{A}$ 为 V 上的一个 **Hermite** 变换或自伴随变换.

定义 2.40

n 阶方阵 A 若满足 $A^* = A$，则称其为 **Hermite** 矩阵或自伴随矩阵.

定理 2.11

n 维酉空间 V 上的线性变换 $\mathscr{A}$ 是 Hermite 变换 $\Longleftrightarrow$ $\mathscr{A}$ 在 V 的一个标准正交基下的矩阵 A 为 Hermite 矩阵

定理 2.12

Hermite 变换 (矩阵) 的特征值为实数.

证明 设 λ_1 为 Hermite 变换的一个特征值，其对应的特征向量为 η，则有

$$\lambda_1(\eta, \eta) = (\lambda_1\eta, \eta) = (\mathscr{A}\eta, \eta) = (\eta, \mathscr{A}\eta) = (\eta, \lambda_1\eta) = \bar{\lambda}_1(\eta, \eta). \tag{2.69}$$

故 $\lambda_1 = \bar{\lambda}_1$，即 λ_1 为实数.

定理 2.13

Hermite 变换在 V 中的某个标准正交基下的矩阵为对角阵，且对角元均为实数.

证明 证明过程与定理2.10类似.

推论 2.10

Hermite 矩阵可以酉相似到一个实对角阵.

在许多物理问题中，我们需要在复数空间讨论问题，但得到的物理量又要求是实数，这时，Hermite 矩阵特征值为实数这一特点便能得到很好的应用.

例题 2.38 量子力学第三条基本假设：量子力学体系的力学量由态矢量空间 $\mathbb{H}$ 中的线性厄米算符表示；其本征矢量的全体构成 $\mathbb{H}$ 的一组完备基；在任一量子态下观测此力学量，总是以一定概率得到对应厄米算符的某一个本征值.

用数学语言描述为：$\hat{\mathscr{A}} = \hat{\mathscr{A}}^\dagger$， $\hat{\mathscr{A}}\,|a_i\rangle = a_i\,|a_i\rangle$， $|\Psi\rangle = \sum_i c_i\,|a_i\rangle$.

上述式 1 即 $\hat{\mathscr{A}}$ 为厄米变换，式 2 即变换的特征方程 (由厄米变换知 a_i 为实数)，式 3 为任一态矢量在本征矢量构成的基下展开，其中 $|c_i|^2$ 代表了 $|\Psi\rangle$ 态坍缩到本征态 $|a_i\rangle$ 的概率，亦即观测到力学量 $\hat{\mathscr{A}}$ 的值为 a_i 的概率.

2.7.3 矩阵的分解

将一个矩阵分解为数个矩阵相乘是我们处理实际问题时常用的简化手段，良好的分解方法可以使解决问题事半功倍，本节列举几种分解方法的实际应用.

QR 分解

定理 2.14

QR 分解：n 阶可逆实矩阵 A 可分解为一个正交阵 Q 和一个上三角阵 R 的乘积，$A = QR$.

证明 证明过程即 Schmidt 正交化过程.

QR 分解在计算数学中常用于求矩阵的特征值和特征向量，因其具有数值稳定性高、可通过优化算法加速收敛等特点被广泛使用，下面不加证明地给出求特征值的原理：

例题 2.39 A 为给定的 n 阶实矩阵，记 $A_1 = A$，对 A_1 进行 QR 分解得 $A_1 = Q_1R_1$，令 $A_2 = R_1Q_1$，对 A_2 进行 QR 分解得到 $A_2 = Q_2R_2$，令 $A_3 = R_2Q_2$，依此类推，若 $A_k = Q_kR_k$，则令 $A_{k+1} = R_kQ_k$，这样迭代得到 n 阶矩阵序列 $\{A_k\}$，满足：

(1) $\{A_k\}$ 为正交相似序列. 事实上，有 $Q_kA_{k+1}Q_k^T = Q_kR_kQ_kQ_k^T = Q_kR_k = A_k$.

(2) 记 A_k 的元素为 $(a_{ij}^{(k)})_{n\times n}$，若矩阵序列 $\{A_k\}$ 基本收敛，则有

$$\begin{aligned} a_{ij}^{(k)} &\to 0, \quad k\to\infty, 1\le i<j\le n \\ a_{ii}^{(k)} &\to \lambda_i, \quad k\to\infty, i=1,2,\cdots,n \end{aligned} \tag{2.70}$$

这样便得到了 A 的全部特征值 λ_i.

例题 2.40 除 Schmidt 正交化外，还可利用 Givens 旋转变换、Householder 反射变换等方法实现 QR 分解.

奇异值分解 (SVD)

本小节参考 (https://zhuanlan.zhihu.com/p/480389473)

定理 2.15

奇异值分解 (SVD)：对 $\forall A\in\mathbb{C}^{m\times n}$，存在酉矩阵 $U\in\mathbb{C}^{m\times m}, V\in\mathbb{C}^{n\times n}$，使得 $A=U\Sigma V^H$，其中 $\Sigma = diag\{\sigma_1,\cdots,\sigma_r,0,\cdots,0\}$，其中 $\sigma_i=\sqrt{\lambda_i}$，$\lambda_i$ 为 A^HA 的非零特征值且 $\lambda_1\ge\cdots\ge\lambda_r>0$.

奇异值分解有很多应用，本节以图像压缩为例给出 SVD 在图像处理中的一个应用：

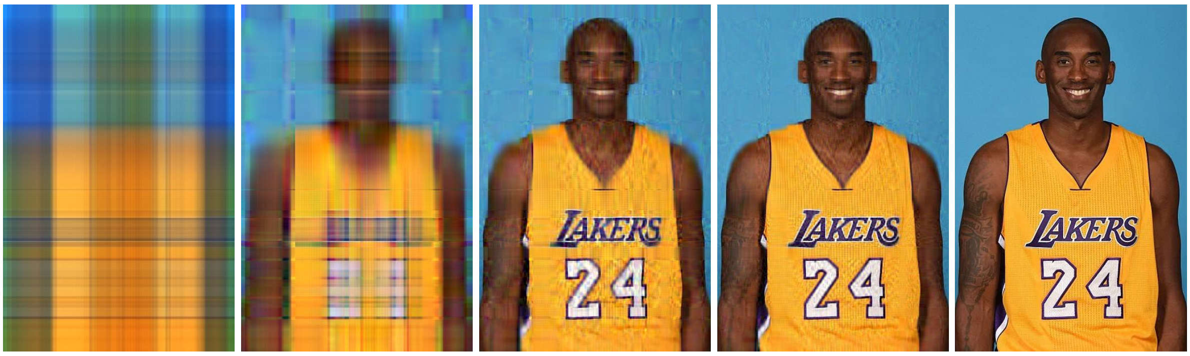
例题 2.41 利用 SVD 对下图进行压缩：

图 2.6: Kobe(612×408)

图像一般被存储为一个 $\mathbb{R}^{m\times n\times 3}$ 的矩阵，前两个指标代表像素点的位置，最后一个指标包含了这个像素点的 RGB 信息，决定了像素的颜色，我们可以分别处理三种颜色，每种颜色对应一个 $\mathbb{R}^{m\times n}$ 的矩阵，可分别进行奇异值分解.

对于一个矩阵 A，较大的奇异值 (位于 Σ 矩阵的左上) 及其对应的行、列向量对 A 的影响比重大，所以我们可以扔掉 0 及靠后的奇异值，仅保留较大的一些奇异值及其对应的行、列向量来实现对图片的压缩. 对于本图，A 有 408 个奇异值，下面给出仅保留前 1，4，20，40，200 个奇异值得到的图像压缩结果：

可以看到，仅取 1 个奇异值时，得到的图片有明显的色块分布，几乎看不出图片内容；取 4 个奇异值时初具人形；取 20 个奇异值时已经能看清面部表情，但图片还是很糊；取 40 个奇异值时效果已经很好了，但还存在一些噪点；取 200 个 (一半) 奇异值时已经和原图十分接近.



(a) num=1　(b) num=4　(c) num=20　(d) num=40　(e) num=200

图 2.7: 图像压缩结果

我们选出取 40 个特征值的情况进行详细计算：

原图所需存储的矩阵大小为 $612\times408\times3=749088$，对三个 612×408 的矩阵分别进行奇异值分解，每个矩阵得到 $U\in\mathbb{R}^{612\times612}, V^T\in\mathbb{R}^{408\times408}$，奇异值有 408 个，选 U 的前 40 列得到 $U'\in\mathbb{R}^{612\times40}$，选 V^T 的前 40 行得到 $V'^T\in\mathbb{R}^{40\times408}$，这样我们构造出新的矩阵 $A'=U'diag\{\lambda_1,\lambda_2,\cdots,\lambda_{40}\}V'^T$，这样压缩后我们仅需要存储 U', V'^T 及 40 个奇异值，相应的存储空间大小为 $40\times(612+408+1)\times3=122520$，为原本的 1/6. 可以看到，我们通过这种方法将图片压缩了六倍，且得到的图片仍较为清晰 (虽然丢掉较小的奇异值不可避免地丢失了信息).

LU 分解

定理 2.16

LU 分解：n 级矩阵 A 可以分解为一个下三角矩阵 L 和一个上三角矩阵 U 的乘积：$A=LU$，这实际上与高斯消元法对应.

不同形式的 LU 分解常用于求解线性方程组，相比于高斯消元法可以节省部分运算量. 其大致的计算思路为：求解线性方程组 $Ax=\beta \iff$ 求解 $Ly=\beta, Ux=y$ (这两个方程都是很容易求解的).

例题 2.42 几种 LU 分解的求解步骤.

(1) Doolittle 分解：
$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{pmatrix}$$

用以下顺序可以逐层求解出 L, U 矩阵的矩阵元：计算 U 的第一行元素，计算 L 的第一列元素，计算 U 的第二行元素，计算 L 的第二列元素，$\cdots$，计算 U 的第 n 行元素，计算 L 的第 n 列元素.

(2) Crout 分解：
$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix} \begin{pmatrix} 1 & u_{12} & \cdots & u_{1n} \\ & 1 & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix}$$

求解方法与 Doolittle 分解完全类似，只是先计算 L 的列，再计算 U 的行.

(3) $\mathrm{LDL^T}$ 分解：对正定阵 A 进行 Doolittle 分解，再提取出 U 的对角元 $D=diag\{u_{11},u_{22}\cdots,u_{nn}\}$，可证 $U=DL^T$，从而 A 被分解为 LDL^T，且 D 为正定阵的标准型，分解步骤与 Doolittle 分解完全相同.

例题 2.43 利用 Doolittle 分解求解方程组

$$\begin{cases} 2x_1 + x_2 + x_3 = 4 \\ x_1 + 3x_2 + 2x_3 = 6 \\ x_1 + 2x_2 + 2x_3 = 5 \end{cases} \tag{2.71}$$

解 系数矩阵经 Doolittle 分解得到：$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0.5 & 0.6 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2.5 & 1.5 \\ 0 & 0 & 0.6 \end{pmatrix}$

解 $Ly = \beta$ 得 $y = (4, 4, 0.6)^T$，解 $Ux = y$ 得 $x = (1, 1, 1)^T$.

2.7.4 利用初等变换将二次型化为标准型

定理 2.17

每一个实对称阵 A 都相合于对角阵，即存在初等阵 $P_1, P_2, \cdots, P_r$ 使得

$$P_r^T P_{r-1}^T \cdots P_1^T A P_1 P_2 \cdots P_r = diag(\mu_1, \mu_2, \cdots, \mu_n). \tag{2.72}$$

这里 $\mu_1, \mu_2, \cdots, \mu_n$ 为实数.

证明 定理的证明即初等变换法的具体操作过程：

(1) 若 $a_{11} \neq 0$ 则利用 a_{11} 将第一行、第一列的其他位置消为零，得 $\begin{pmatrix} a_{11} & 0 \\ 0 & A_{n-1} \end{pmatrix}$，继续对 A_{n-1} 做类似操作即可.

(2) 当 $a_{11} \neq 0$，第一行其他位置一定有一个非零元素，设为 a_{1i}，则将 A 的第 i 列加到第一行，再将第 i 行加到第一行，即令 $A' = T_{1i}(1)AT_{1i}(1)^T$，则转化为第一种情况.

例题 2.44 利用初等变换将二次型 $x_1^2 + 4x_2^2 + x_3^2 - 4x_1x_2 - 8x_1x_3 - 4x_2x_3$ 化为标准型，并求变换矩阵.

解 写出二次型对应的矩阵 $A = \begin{pmatrix} 1 & -2 & -4 \\ -2 & 4 & -2 \\ -4 & -2 & 1 \end{pmatrix}$，将其与单位阵 I 上下写在一起，对 A 做一对对行、列变换，同时只对 I 做列变换，有：

$$\begin{pmatrix} 1 & -2 & -4 \\ -2 & 4 & -2 \\ -4 & -2 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -4 \\ 0 & 0 & -10 \\ -4 & -10 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -10 \\ 0 & -10 & -15 \\ 1 & 2 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & -35 & -25 \\ 0 & -25 & -15 \\ 1 & 6 & 4 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & -35 & 0 \\ 0 & 0 & 20/7 \\ 1 & 6 & -2/7 \\ 0 & 1 & -5/7 \\ 0 & 1 & 2/7 \end{pmatrix}$$

即变换矩阵 P 写为 $\begin{pmatrix} 1 & 6 & -2/7 \\ 0 & 1 & -5/7 \\ 0 & 1 & 2/7 \end{pmatrix}$，令 $(y_1, y_2, y_3)^T = P^{-1}(x_1, x_2, x_3)^T$，二次型化为标准型 $y_1^2 - 35y_2^2 + \frac{20}{7}y_3^2$.

注 上面变换的原理是：对 A 做成对的行、列变换，A 变为 P^TAP；仅对 I 做列变换，I 变为 P，这样，我们同时求得了二次型的标准型及其变换矩阵. 同样的，也可以将 A 和 I 左右写在一起，在对 A 做成对变换的同时只对 I 做行变换，区别只在于这样得到的 I 变换后的矩阵为 P^T.

注 从上例可以看出，由于同一个二次型对应的标准型不唯一，且经过不同初等变换步骤得到的标准型也不相同，这是我们不希望看到的. 因此，在大部分考题中，我们要求使用正交方阵将其对角化，这样得到的结果在对角元相差顺序的情况下唯一，且变换过程不会对二次曲面有拉伸，良好保持了原有形状.

另解：利用正交变换将上述二次型化为标准型[4]：

解 矩阵 A 对应特征值为 $\lambda_1 = \lambda_2 = 5, \lambda_3 = -4$，对应的特征向量为 $\alpha_1 = (-1,0,1)^T, \alpha_2 = (-1,2,0)^T, \alpha_3 = (2,1,2)^T$，经 Schmidt 正交化得到一组标准正交向量 $e_1 = (-1/\sqrt{2}, 0, 1/\sqrt{2})^T, e_2 = (-\sqrt{2}/6, 2\sqrt{2}/3, -\sqrt{2}/6)^T, e_3 = (2/3, 1/3, 2/3)^T$，记 $P = (e_1, e_2, e_3)$，则有 P 为正交阵，且 $P^TAP = diag\{5,5,-4\}$，对应的正交变换为 $y = P^{-1}x$，将二次型化为了标准型 $5y_1^2 + 5y_2^2 - 4y_3^2$. 特别地，$5y_1^2 + 5y_2^2 - 4y_3^2 = 1$ 对应单叶双曲面.

[4] 2021-2022 第二学期期末 3.

2.8 第八次习题课

2.8.1 期末复习知识点整理

期末考试主要考察后三章（即六、七、八）三章，期中前的内容大概率不会单独考察，但实际上后半学期的内容也是建立在前面章节的基础上的，所以打牢基础才是最重要的。在理解和掌握基本概念的基础上再适当做几套往年考试题，就能拿到一个比较理想的成绩了。接下来我们按照书上章节的顺序梳理一下重要的知识点，并适当结合一些往年的真题来加深理解。

注 助教考试前也是看不到卷子的，所以我们仅是在个人理解范围内，尽可能帮助大家建立一个知识框架，可用于查漏补缺，但不能代表考试重点和考试题型，如有错误或疏漏欢迎指正。

注 标注 * 的表示不要求掌握，但我觉得值得了解的内容。

线性变换

¶ 数组空间上的线性映射

- 数组空间上线性映射（变换）可与矩阵一一对应
- 由上条推出的线性映射（变换）的秩与矩阵的秩的对应关系
- 验证是否是线性映射（变换）：保加法、保数乘（* 注意数域的选取）
- 像空间与核空间维数和为 n

¶ 线性变换的特征值与特征向量

- 特征值与特征向量的定义（注意特征向量是非零向量）
- 特征值与特征向量的几何意义：一般考虑二维情形，即椭圆长短轴，就足够了。但是这只对线性变换可对角化的情形成立，变换不可对角化时不能通过求特征值来得到长短轴长度，因为二维情形下，不可对角化说明变换只有一个特征值，且对应的特征子空间是一维的，自然无法求出两个对称轴。

  例题 2.45（2023-2024 第二学期期末填空 6.） 设平面 $\mathbb{R}^2$ 上的线性变换 $\mathscr{B}:(x,y)^T\mapsto(x+y,y)$ 将单位圆 C 变为椭圆 E，则 E 的长半轴的长度为：

  解 这个变换对应的矩阵是不能对角化的，所以我们采用另外一种方式：单位圆上一点 $(\cos\theta,\sin\theta)$ 在变换后的像为 $(\cos\theta+\sin\theta,\sin\theta)$，由线性性，原点仍然是椭圆的中心，那么我们只需要最大化这个点到原点的距离，即为长半轴长度，$(\cos\theta+\sin\theta)^2+(\sin\theta)^2=1+\sin^2\theta+\sin 2\theta=\dfrac{3}{2}-\dfrac{1}{2}\cos 2\theta+\sin 2\theta$，最大值为 $\dfrac{3+\sqrt{5}}{2}$，故 E 的长半轴的长度为 $\dfrac{1+\sqrt{5}}{2}$（实际上根据取最大值时的 θ 值，也可定出长半轴方向）。

- 特征子空间，* 为了使其成为子空间，将特征向量全体再并上零向量。
- 特征值与特征向量的求解：先求出特征多项式，再对代入每个特征值求解线性方程组得到特征向量。
- 注意特征值相同不能推出特征向量相同，比如 A 与 A^T（很多人作业里犯过这个错误）。
- 特征值与矩阵迹、行列式的关系：$tr(A)=\lambda_1+\cdots+\lambda_n,\det(A)=\lambda_1\cdots\lambda_n$.

¶ 矩阵的相似

- 相似的定义。* 看清楚在什么数域上相似（如实相似和复相似是不同的）。
- 相似的矩阵有相同的特征多项式和特征值，但不一定有相同的特征向量。
- 相似的矩阵有相同的迹与行列式，秩也相同，这是上面一条与上节最后一条的推论。
- n 阶方阵 A 可对角化的充要条件：

  (1) A 有 n 个线性无关的特征向量

(2) A 每个特征值的代数重数与几何重数相等（回顾代数重数和几何重数的定义：特征多项式中的幂次与特征子空间的维数。实际上很容易推出与 (1) 等价）。

- 不同特征值对应的特征向量线性无关，利用这个可以推出特征值两两不同的方阵一定可以对角化。同时这个命题的证明方法也值得看一看（利用归纳法）。
- 实对称方阵的特征值都是实数，且不同特征值对应的特征向量是正交的，且一定可以正交相似对角化。（教材命题 7.3.4，命题 7.3.5，定理 7.3.6）这个很重要，最好要熟悉证明过程。

例题 2.46（2023-2024 第二学期期末 6.） 设 A,B 均为 n 阶实对称方阵，满足 $AB=BA$.
(1) 证明：存在非零列向量 $v\in\mathbb{R}^n$，使得 v 同时为 A 和 B 的特征向量。
(2) 证明：存在 n 阶正交方阵 P，使得 $P^{-1}AP$ 与 $P^{-1}BP$ 均为对角矩阵。

证明 任取 (2) 中的 P 的一个列向量作为 v，满足 (1) 中条件，只证明 (2) 即可。
将 A 正交相似到对角阵：$T^{-1}AT=\Lambda$，由题有 $T^{-1}ATT^{-1}BT=T^{-1}ABT=T^{-1}BAT=T^{-1}BTT^{-1}AT$
即 $\tilde{B}=T^{-1}BT$ 与 Λ 可交换。由上面的命题可将 Λ 与 $\tilde{B}$ 写为分块形式：

$$\Lambda=\begin{pmatrix}\lambda_1 I_{r_1} & & 0\\ & \ddots & \\ 0 & & \lambda_s I_{r_s}\end{pmatrix}\qquad \tilde{B}=\begin{pmatrix}B_{11} & \cdots & B_{1s}\\ \vdots & \ddots & \vdots\\ B_{s1} & \cdots & B_{ss}\end{pmatrix}$$

其中 λ_i 是 A 的互不相同的特征值。考察 $\Lambda\tilde{B}$ 与 $\tilde{B}\Lambda$ 的第 (i,j) 个分量得到 $\lambda_i B_{ij}=\lambda_j B_{ij}$，且由于 $\lambda_i\neq\lambda_j, i\neq j$，所以 $B_{ij}=0, i\neq j$。所以 $\tilde{B}$ 是准对角阵。注意到 $\tilde{B}$ 是实对称阵，所以 B_{ii} 是实对称阵，从而对每一个 B_{ii}，均存在一个 Q_i，使得 $Q_i^{-1}B_{ii}Q_i$ 为对角阵。记

$$Q=\begin{pmatrix}Q_1 & & 0\\ & \ddots & \\ 0 & & Q_s\end{pmatrix}$$

取 $P=TQ$，则 P 满足题意。

- 实对称矩阵正交相似到对角阵的具体过程，可参考教材例 7.3.3，这个计算过程一定要会。

例题 2.47（教材例 7.3.3） 设 $A=\begin{pmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 2 & 2 & 1\end{pmatrix}$，求正交阵 T，使 $T^{-1}AT$ 为对角阵

解 具体过程参见教材，一定要自己动手算一下，这类题目既有计算特征多项式，又涉及 Schmidt 正交化，不注意就可能会算错。

- 任意一个复方阵都可以相似于一个上三角阵，且主对角线元素都是特征值（用数学归纳法证明）。

¶ 一般空间上的线性变换

- 一般线性空间上的线性映射（变换）的定义：保加法和数乘。这里我们不能直接以矩阵的形式 ($\boldsymbol{y}=A\boldsymbol{x}$) 来表示线性映射（变换），数组空间可以这么做是因为矩阵与数组向量有一个良定的乘法运算，但在一般的线性空间中，我们甚至不知道这里“向量”的具体形式，因而 $A\boldsymbol{x}$ 也是没有定义的。但我们仍然可以用线性映射（变换）在基下的矩阵来表示它，注意这二者是有区别的。

例题 2.48 定义 $\Lambda: C[a,b]\to\mathbb{R}, \Lambda(f)=f(0)$，显然这是一个线性映射，但是无法用 $\boldsymbol{y}=A\boldsymbol{x}$ 的形式来表示。即使我们取 $C[a,b]$ 的一个有限维子空间 $P_n[a,b]$，即次数不超过 n 次的多项式全体构成的空间，仍然不能以这种形式来表示，但这个映射在基 $\{1,x,\cdots,x^n\}$ 下的矩阵为 $diag(1,0,\cdots,0)$

- 线性相关的向量在线性映射作用后仍然线性相关
- （根据个人经验，本课程在涉及线性映射在基下的矩阵时，一般只考虑线性变换）线性变换在一组基下的矩阵实际上是每个变换后的基向量在这组基下的坐标。
- 线性变换在基下的矩阵导出坐标变换 $\boldsymbol{y}=A\boldsymbol{x}$，这里 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 分别是 α 与 $\mathscr{A}\alpha$ 在基下的坐标（定理 6.4.3）。

- 线性变换在不同基下的矩阵相似，相差的矩阵即为两组基之间的过渡矩阵。

例题 2.49（2024-2025 第一学期期末判断 2.） 设 V 是 n 维线性空间，$\mathscr{A}$ 是 V 上的线性变换，若 $\mathscr{A}$ 在 V 的任意一组基下的矩阵都相等，则 $\mathscr{A}$ 是 V 上的数乘变换。

解 正确，由不同基下矩阵之间相似可以得到 $A = PAP^{-1}$ 对任意可逆矩阵 P 都成立，即 $AP = PA$，从而只能是 $A = \lambda I$，也就是 $\mathscr{A}$ 为数乘变换。

- 线性变换与其在任一组基下的矩阵的有相同的特征值。

例题 2.50（教材例 6.4.10） 设 $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}, V = F^{2\times 2}$ 为线性空间，定义 V 上的线性变换 $\mathscr{A} : \mathscr{A}M = AM, M \in V$，求 $\mathscr{A}$ 的特征值和特征向量。

解 具体过程参见教材，这类题考试还是常考的，只要对线性映射理解透彻的话就很容易做。首先 A 是用来定义这个映射的，并不是这个映射在基下的矩阵，这二者是完全不同的，空间 V 的维数是 4，所以变换在基下的矩阵也一定是 4 阶方阵。并且要记住线性变换在基下的矩阵是坐标，所以这个矩阵的特征向量也是在基下的坐标，所以 $\mathscr{A}$ 的特征向量是这组基乘上坐标，而不是矩阵的特征向量（V 中的“向量”是二阶方阵，而不是数组向量）。

- 线性变换 $\mathscr{A}$ 可对角化的定义：其在某组基下的矩阵 A 为对角阵。同时变换的特征多项式、秩、迹、行列式也是类似定义，* 它们良定因为这些都是相似不变量。

内积空间及变换

¶ 数组内积空间的定义与性质

- 用度量矩阵来定义的数组空间内积，三条性质：对称性、线性性、正定性。
- 用内积来诱导度量（或说是向量的模长），Cauchy-Schwarz 不等式。
- 度量（距离）的三条性质：对称性、正定性、三角不等式，向量夹角，勾股定理（教材命题 7.1.2）

¶ 标准正交基

- 标准正交基：标准：模长都是 1；正交：两两正交；基：构成空间的一组基。
- Schmidt 正交化（基本是每年必考的）：一定要熟悉正交化的过程，实质上是一个正交分解的过程，顺次在前面每个已经构建好的两两正交的向量上做投影，减去这些投影的和，就得到与前面都正交的一个向量，再单位化，不断重复。理解了这个几何视角就会好记忆很多，但是计算时也要小心不要出错。

例题 2.51（2024-2025 第一学期期末 4.） 记 V 是所有 2 阶实对称方阵构成的实线性空间。对于 V 中任意两个矩阵，定义 $\langle A, B\rangle = tr(AB)$，则 $(V, \langle\cdot,\cdot\rangle)$ 构成一个欧式空间。

(1) 考虑 V 中向量 $A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$。将 A_1, A_2 扩充为 V 的一组基 Γ_1。

(2) 利用 Schmidt 正交化，从 Γ_1 构造 V 的一组标准正交基 Γ_2。

解 (1) 显然 V 的维数是 3，我们可以先考虑 A_1, A_2 生成的子空间，我们发现这两个向量的线性组合可以让两个对角元取到任意值，但是另外两个元素受到组合系数的约束，所以我们可以取 $A_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

(2) 按照题目上内积的定义来计算向量模长。$|A_1|^2 = tr(A \cdot A) = 3, |A_1| = \sqrt{3}$，故取 $e_1 = \frac{1}{\sqrt{3}}A_1$. $\beta_2 = A_2 - tr(A_2, e_1)e_1 = \frac{1}{3}\begin{pmatrix} -2 & 1 \\ 1 & 3 \end{pmatrix}, e_2 = \frac{1}{|\beta_2|}\beta_2 = \frac{1}{\sqrt{15}}\begin{pmatrix} -2 & 1 \\ 1 & 3 \end{pmatrix}, \beta_3 = A_3 - tr(A_3, e_1)e_1 - tr(A_3, e_2)e_2 = \frac{1}{5}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, e_3 = \frac{1}{|\beta_3|}\beta_3 = \frac{1}{\sqrt{10}}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$，$e_1, e_2, e_3$ 即为一组标准正交基。

- 正交方阵的定义：$QQ^T=Q^TQ=I$，等价的条件：Q 的行（或列）向量构成 $\mathbb{R}^n$ 的一组标准正交基。前者常用于验证方阵为正交方阵，或在某些问题中将单位阵 I 写为 QQ^T 来简化处理；后者则常常是已知是正交方阵情况下，用这个性质来处理问题。

例题 2.52（2024-2025 第一学期期末判断 3.） 设 A,B 是 n 阶正交矩阵，满足 $|A|+|B|=0$，则 $|A+B|=0$.

解 正确，这里显然是用正交方阵的定义来处理，我们有 $|A+B|=|BB^TA+BA^TA|=|B||B^T+A^T||A|=-|A|^2|A+B|=-|A+B|$，故有 $|A+B|=0$.

例题 2.53（2024-2025 第一学期期末 6.） 考虑实对角阵 $A=diag(a_1,\cdots,a_n)$，其中 $a_1>\cdots>a_n>0$，设 P,Q 是 n 阶正交矩阵，且 $PA=AQ$，证明 $P=Q$.

证明 这里给出了我们 A 中元素的形式，这提示我们用上述正交矩阵的第二个性质，设 $P=(p_{ij}),Q=(q_{ij})$，则 $PA=AQ$ 即：

$$\begin{pmatrix} a_1p_{11} & a_2p_{12} & \cdots & a_np_{1n} \\ a_1p_{21} & a_2p_{22} & \cdots & a_np_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_1p_{n1} & a_2p_{n2} & \cdots & a_np_{nn} \end{pmatrix} = \begin{pmatrix} a_1q_{11} & a_1q_{12} & \cdots & a_1q_{1n} \\ a_2q_{21} & a_2q_{22} & \cdots & a_2q_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_nq_{n1} & a_nq_{n2} & \cdots & a_nq_{nn} \end{pmatrix}$$

由于 Q 是正交方阵，AQ 的第 1 个行向量的模长平方是 a_1^2，再由 PA 的形式，$a_1^2=\sum_{k=1}^n a_k^2p_{1k}^2\leqslant a_1^2$，说明等号成立，那么只有 $|p_{11}|=1,p_{1j}=0,j=2,\cdots,n$，类似地可得 P,Q 均为对角阵，而 $a_i>0$，故 $P=Q$.

- 正交方阵的一些性质（教材定理 7.2.4），上面的例题也有用到这些性质。
  (1) 单位阵是正交方阵
  (2) 若 A,B 是同阶正交方阵，则 AB 也是正交方阵
  (3) 若 A 是正交方阵，则 $A^T=A^{-1}$ 也是正交方阵
  (4) 正交方阵的行列式是 1 或 -1
  (5) 正交方阵的特征值模长为 1
- 正交子空间、正交投影。正交投影在处理最佳逼近元问题时可以在一定程度上简化计算。

例题 2.54（2021-2022 第二学期期末 4.） 在数域 $\mathbb{R}$ 上次数不超过 3 的多项式全体构成的线性空间 $R_3[x]$ 上定义内积 $\langle f,g\rangle=\displaystyle\int_{-1}^1 f(x)g(x)dx$.

(1) 将向量组 $1,x,x^2$ 按顺序使用 Schmidt 正交化得到单位正交向量组。

(2) 令 $W=\langle 1,x,x^2\rangle$。求多项式 $p(x)\in W$ 使得 $|x^3-p(x)|$ 达到最小。

解 (1) 按正交化过程计算即可，单位正交向量组为 $\left\{\dfrac{\sqrt{2}}{2},\dfrac{\sqrt{6}}{2}x,\dfrac{\sqrt{10}}{4}(3x^2-1)\right\}$，记为 $\{e_1,e_2,e_3\}$

(2) 我们用正交投影的方法来处理这个问题，要求的 $p(x)$ 即为 x^3 在子空间 W 上的正交投影。

即 $p(x)=\langle x^3,e_1\rangle e_1+\langle x^3,e_2\rangle e_2+\langle x^3,e_3\rangle e_3=\dfrac{\sqrt{6}}{5}\cdot\dfrac{\sqrt{6}}{2}x=\dfrac{3}{5}x$（前后两项由于是奇函数，积分为 0）。

2.8.2 同时对角化

习题 2.48 (公共特征向量)

设 $\mathscr{A},\mathscr{B}$ 是线性空间 V 上的线性变换，若 $\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$ 且 $\mathscr{A},\mathscr{B}$ 的特征值均在 $\mathbb{F}$ 中. 则 $\mathscr{A},\mathscr{B}$ 有公共特征向量.

证明 由 $\mathscr{A}$ 的特征值在 $\mathbb{F}$ 中，知存在 $\lambda_0\in\mathbb{F}$, 使得 $V_{\lambda_0}=\{x\in V:\mathscr{A}x=\lambda_0x\}\neq\emptyset$. 对任意 $x\in V_{\lambda_0}$, 有 $\mathscr{A}\mathscr{B}x=\mathscr{B}\mathscr{A}x=\lambda_0\mathscr{B}x$, 故 $\mathscr{B}x\in V_{\lambda_0}$，因此 V_{λ_0} 是 $\mathscr{B}$-不变子空间. 由于 $\mathscr{B}$ 的特征值均在 $\mathbb{F}$ 中，因此 $\mathscr{B}|_{V_{\lambda_0}}$ 的特征值也均在 $\mathbb{F}$ 中，故存在 $x_0\in V_{\lambda_0}$ 及 $\mu_0\in\mathbb{F}$, 使得 $\mathscr{B}x_0=\mu_0x_0.x_0$ 及为所求.

注[矩阵语言] 设矩阵 $A,B\in\mathbb{F}^{n\times n}$,$AB=BA$ 且 A,B 特征值均在 $\mathbb{F}$ 中. 则矩阵 A,B 有公共特征向量. 考虑线性

变换 $\mathscr{A}:\mathbb{F}^n\to\mathbb{F}^n\ x\mapsto Ax;\mathscr{B}:\mathbb{F}^n\to\mathbb{F}^n\ x\mapsto Bx$, 易知 $\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$ 由上例立明.

习题 2.49

设 $\mathscr{A},\mathscr{B}$ 是线性空间 V 上的线性变换，若 $\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$，$\mathscr{A},\mathscr{B}$ 的特征值均在 $\mathbb{F}$ 中且 $\mathscr{A}$ 有 $n=dimV$ 个不同的特征值. 则存在 V 中的一组基，使得 $\mathscr{A},\mathscr{B}$ 在这组基下的矩阵均为对角矩阵.

证明 设 $\mathscr{A}$ 的特征值为 $\lambda_1,\cdots,\lambda_n$, 对应的特征向量为 $\alpha_1,\cdots,\alpha_n$. 则 $V=V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_n}$ 且 $V_{\lambda_i}=span\{\alpha\}$. 同上例知 V_{λ_i} 是 $\mathscr{B}$-不变子空间. 特别地，存在 $\mu_i\in\mathbb{F}$，使得 $\mathscr{B}\alpha_i=\mu_i\alpha_i$. 因此线性变换 $\mathscr{A},\mathscr{B}$ 在基 $\{\alpha_1,\cdots,\alpha_n\}$ 下对应的矩阵均为对角矩阵.

注[矩阵语言] 设 $A,B\in\mathbb{F}^{n\times n}$,$AB=BA$,$A,B$ 的特征值均在 $\mathbb{F}$ 中且 A 有 n 个不同的特征值. 则存在可逆矩阵 P，使得 $P^{-1}AP,P^{-1}BP$ 同时为对角矩阵.

证明 作业题.

习题 2.50

设 $\mathscr{A},\mathscr{B}$ 是 V 上的线性变换，若 $\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$ 且存在两组基使得 $\mathscr{A},\mathscr{B}$ 分别在这两组基下对应的矩阵是对角矩阵. 则存在一组基，使得 $\mathscr{A},\mathscr{B}$ 在这组基下对应的矩阵同时为对角矩阵.

证明 由于 $\mathscr{A}$ 可对角化，因此 V 可以分解成 $\mathscr{A}$ 的特征子空间的直和, 即 $V=V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_k}$，其中 $V_{\lambda_i}=\{x\in V:\mathscr{A}x=\lambda_i x\}$. 且 $\lambda_i\neq\lambda_j$. 同上，知 V_{λ_i} 是 $\mathscr{B}$-不变子空间. 下面对 $n=dimV$ 进行归纳. 当 $n=1$ 时，显然. 假设命题对小于 n 时成立，下面需说明命题对 n 成立. 当 $k=1$ 时，$\mathscr{A}=\lambda_1 id_V$, 又 $\mathscr{B}$ 可以对角化，因此 $\mathscr{A},\mathscr{B}$ 可以同时对角化. 当 $k>1$ 时，$dimV_{\lambda_i}<n$, 又 $\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$，故 $\mathscr{A}|_{V_{\lambda_i}}\mathscr{B}|_{V_{\lambda_i}}=\mathscr{B}|_{V_{\lambda_i}}\mathscr{A}|_{V_{\lambda_i}}$,$\mathscr{A}|_{V_{\lambda_i}},\mathscr{B}|_{V_{\lambda_i}}$ 在 V_{λ_i} 上可对角化. 由归纳假设，知存在 V_{λ_i} 上的一组基 $e_{i_1},\cdots,e_{i_{s_i}}$ 使得 $\mathscr{A}|_{V_{\lambda_i}},\mathscr{B}|_{V_{\lambda_i}}$ 在这组基下为对角矩阵，将诸如此类基拼成 V 的基，易知 $\mathscr{A},\mathscr{B}$ 在这组基下的矩阵为对角矩阵.

注[矩阵语言] 设 $A,B\in\mathbb{F}^{n\times n}$, 若 $AB=BA$ 且 A,B 均可对角化. 则 A,B 可以同时对角化.

证明 对矩阵阶数进行归纳，当 $n=1$ 时，显然. 假设命题对于小于 n 时成立，现需证明命题对于 n 成立. 存在可逆矩阵 P, 使得 $P^{-1}AP=diag(\lambda_1 I_{l_1},\cdots,\lambda_s I_{l_s})$, 记 $P^{-1}BP=\begin{pmatrix}B_{11}&\cdots&B_{1k}\\ \vdots&&\vdots\\ B_{k1}&\cdots&B_{kk}\end{pmatrix}$, 其中 $\lambda_i\neq\lambda_j(i\neq j)$. 由 $AB=BA$ 知矩阵 $P^{-1}AP$ 与矩阵 $P^{-1}BP$ 可交换. 故 $\lambda_i B_{ij}=\lambda_j B_{ij}(i\neq j)$, 故 $B_{ij}=0(i\neq j)$. 当 $k=1$ 时，A 为数量矩阵，结论显然. 当 $k>1$ 时，B_{ii} 阶数小于 n，存在可逆矩阵 Q_{ii}, 使得 $Q_{ii}^{-1}B_{ii}Q_{ii}$ 为对角矩阵. 令 $Q=diag(Q_{11},\cdots,Q_{kk})$, 令 $T=QP$，则有 $T^{-1}AT$ 与 $T^{-1}BT$ 同时为对角矩阵.

笔记 还可以证明:(1) 若 $A_1,\cdots,A_m\in\mathbb{F}^{n\times n}$ 两两可交换，且它们的特征值均在 $\mathbb{F}$ 中，则 $A_1,\cdots,A_m$ 有公共特征向量.
(2) 若 $A_1,\cdots,A_m\in\mathbb{F}^{n\times n}$ 两两可交换，且它们在 $\mathbb{F}$ 中均可对角化，则 $A_1,\cdots,A_m$ 可同时对角化.

2.8.3 半正定矩阵开 k 次方

习题 2.51

设 B 是 n 阶半正定矩阵，$\mu_1,\cdots,\mu_n$ 是 B 的 n 个特征值. 则对任意给定的正整数 k，存在一个只和 μ_i^k 有关的实系数多项式 $f(x)$，满足 $B=f(B^k)$.

证明 设 $\mu_1,\cdots,\mu_n$ 中互不相同的数为 $\mu_1,\cdots,\mu_s$, 记 $\lambda_i=\mu_i^k$, 令 $f(x)=\sum\limits_{i=1}^{s}\frac{(x-\lambda_1)\cdots(x-\lambda_{i-1})(x-\lambda_{i+1})\cdots(x-\lambda_s)}{\lambda_i-\lambda_1\cdots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\cdots(\lambda_i-\lambda_s)}$. 存在正交矩阵 Q, 使得 $Q^TBQ=diag(\mu_1,\cdots,\mu_n)$, 易验证 $B=f(B^k)$.

习题 2.52

设 A 是 n 阶半正定矩阵. 则对任意正整数 k，存在唯一的 n 阶半正定矩阵 B, 使得 $A = B^k$.

证明 （存在性）存在正交矩阵 P，使 $P^TAP = diag(\lambda_i, \cdots, \lambda_n)$, 其中 $\lambda_i \geq 0$. 令 $B = Pdiag(\lambda_1^{\frac{1}{k}}, \cdots, \lambda_n^{\frac{1}{k}})P^T$, 则 B 半正定且 $B^k = A$.

（唯一性）若 $B^k = A$, 设 B 的特征值为 μ_i，则 A 的特征值为 $\lambda_i^{\frac{1}{k}}$, 由上例，存在一个只与 μ_i 有关的多项式 $f(x)$ 使得 $f(A) = f(B^k) = B$. 若还有 C, 满足 $C^k = A$, 同上分析有 $C = f(C^k) = f(A)$, 从而 $B = C$.

笔记 我们一般将上例中 B 定义为 $A^{\frac{1}{k}}$, 由以上两个例子知，这个定义是合理的 (存在且唯一).

注 第十六次作业中的补充题对正定矩阵同时开根号即可.

一学期的助教工作已经走向尾声了，在这里送给大家一句话作为本学期课程的结束：

我们是孩子，但是我们精力充沛，勇往直前……

——1831 年 6 月 16 日伽罗瓦在法庭上的辩护词

最后，祝大家期末考试取得好成绩！

参考文献

[1] 线性代数教材

[2]《高等代数》——丘维声

[3]《电动力学讲义》——潘海俊

[4]《线性代数（数学专业用）》——李尚志

[5]《线性代数》——李炯生

[6]《数学分析讲义（第三册）》——程艺，陈卿，李平，许斌

[7]《数值计算方法与算法》——张韵华等