线性代数入门习题解答

刘余

清华大学

2023

目录

第零章　预备知识

0.1　逻辑与集合

练习 0.1.1. *画出例 0.1.1 中的所有集合运算律的 Venn 图.*

证明: 略 □

练习 0.1.2. *对任意两个命题 A, B, $A \Rightarrow B$ 可以定义为 $B \vee (\neg A)$. 直观上可以如此理解: $B \vee (\neg A)$ 成立等价于 A 不成立或 B 成立, 于是当 A 成立时必有 B 成立. 运用例 0.1.1 中的逻辑运算律, 证明 $(A \Rightarrow B) \Leftrightarrow ((\neg B) \Rightarrow (\neg A))$.*

证明: 根据定义，$(A \Rightarrow B)$ 等价于 $B \vee (\neg A)$ 且 $((\neg B) \Rightarrow (\neg A))$ 等价于 $(\neg A) \vee (\neg(\neg B)) = (\neg A) \vee B$. □

练习 0.1.3. *对任意三个命题 A, B, C，运用例 0.1。1 中的逻辑运算律，证明逻辑运算符合模律，即当 A 是 C 的必要条件时，$(A \wedge B) \vee C \Leftrightarrow A \wedge (B \vee C)$.*

证明: 当 A 是 C 的必要条件时，$A \vee C = A$。根据逻辑运算律的第 4 条，$(A \wedge B) \vee C = (A \vee C) \wedge (B \vee C) = A \wedge (B \vee C)$。 □

练习 0.1.4. *对任意两个命题 A, B，我们定义运算异或 $A \oplus B$ 为 $(A \wedge (\neg B)) \vee (B \wedge (\neg A))$.*

1. *按例 0.1.1 的条件画出 $A \oplus B$ 对应的 Venn 图.*
2. *证明异或满足交换律和结合律.*

证明:

1. 略
2. 交换律即 $A \oplus B = B \oplus A$; 这是因为

$$A \oplus B = (A \wedge (\neg B)) \vee (B \wedge (\neg A)) = (B \wedge (\neg A)) \vee (A \wedge (\neg B)) = B \oplus A,$$

其中第一个和最后一个等号是定义，第二个等号是逻辑运算律的第 2 条。
结合律即：$(A \oplus B) \oplus C = A \oplus (B \oplus C)$。首先由 $\oplus$ 的定义：

$$\begin{aligned}(A \oplus B) \oplus C &= \Big((A \oplus B) \wedge (\neg C)\Big) \vee \Big(C \wedge (\neg(A \oplus B))\Big) \\ &= \Big(((A \wedge (\neg B)) \vee (B \wedge (\neg A))) \wedge (\neg C)\Big) \vee \Big(C \wedge (\neg((A \wedge (\neg B)) \vee (B \wedge (\neg A)))\Big)\end{aligned}$$

而

$$\begin{aligned}\Big((A \wedge (\neg B)) \vee (B \wedge (\neg A))\Big) \wedge (\neg C) &= \Big((A \wedge (\neg B)) \wedge (\neg C)\Big) \vee \Big((B \wedge (\neg A)) \wedge (\neg C)\Big) \\ &= \Big(A \wedge (\neg(B \vee C))\Big) \vee \Big((\neg A) \wedge (B \wedge (\neg C))\Big)\end{aligned}$$

$$
\begin{aligned}
\neg((A\wedge(\neg B))\vee(B\wedge(\neg A))) &= ((\neg A)\vee B)\wedge((\neg B)\vee A)\\
&= ((\neg A)\vee B)\wedge((\neg B)\vee A)\\
&= ((\neg A)\wedge\neg(B))\vee((\neg A)\wedge A)\vee(B\wedge(\neg B))\vee(B\wedge A)\\
&= ((\neg A)\wedge\neg(B))\vee(B\wedge A)
\end{aligned}
$$

需要说明的是，这一个等式实际上是在说 $\neg(A\oplus B)=(\neg A)\oplus B=A\oplus(\neg B)$. 因此

$$
\begin{aligned}
C\wedge\Big(\neg((A\wedge(\neg B))\vee(B\wedge(\neg A)))\Big) &= C\wedge\Big(((\neg A)\wedge(\neg B))\vee(B\wedge A)\Big)\\
&= \Big(C\wedge((\neg A)\wedge(\neg B))\Big)\vee\Big(C\wedge(B\wedge A)\Big)\\
&= \Big((\neg A)\wedge((\neg B)\wedge C)\Big)\vee\Big(A\wedge(B\wedge C)\Big)
\end{aligned}
$$

结合起来就有

$$
\begin{aligned}
(A\oplus B)\oplus C &= (A\wedge(\neg(B\vee C)))\vee((\neg A)\wedge(B\wedge(\neg C)))\vee((\neg A)\wedge((\neg B)\wedge C))\vee(A\wedge(B\wedge C))\\
&= \Big(A\wedge(((\neg B)\wedge(\neg C))\vee(B\wedge C))\Big)\vee\Big((\neg A)\wedge((B\wedge(\neg C))\vee(C\wedge(\neg B)))\Big)\\
&= \Big(A\wedge(\neg(B\oplus C))\Big)\vee\Big((\neg A)\wedge(B\oplus C)\Big)\\
&= A\oplus(B\oplus C)
\end{aligned}
$$

□

0.2 间接证明法

练习 0.2.1. *请用反证法证明以下命题.*

1. *若 $p>2$ 是素数，则 p 是奇数.*
2. *对正整数 n，若 n^2 是奇数，则 n 是奇数.*
3. *不存在最小的正有理数.*

证明:

1. 假设存在素数 $p>2$ 且 p 是偶数，那么我们可以把 p 写成 $p=2q$ 的形式，其中 q 是整数；因为 $p>2$, 则 $q>1$；从而与 p 是素数矛盾。
2. 假设存在正整数 n 满足 n^2 是奇数但是 n 是偶数, 则我们可以把 n 写成 $n=2q$ 的形式，从而 $n^2=4q^2$ 也是偶数，矛盾。
3. 假设存在最小的正有理数，记为 ϵ。我们考虑 $\delta=\frac{\epsilon}{2}$, 此时 $0<\delta<\epsilon$, 且 δ 也是有理数，这与 ϵ 的最小性矛盾。

□

练习 0.2.2. *经典逻辑的基本公理之一是排中律，即对任意命题 A，A 不能既不真又不假. 反证法可以看作排中律的推论，即如果我们发现 A 不能是错的，那么 A 就只能是对的. 因此，想要证明命题 A 真，不妨挑一个命题 B，先证明 $B\Rightarrow A$，再证明 $(\neg B)\Rightarrow A$. 那么 A 就必须是真的了. 给定命题：存在两个无理数 a,b，满足 a^b 是一个有理数. 请先假设 $\sqrt{2}^{\sqrt{2}}$ 是有理数，再假设 $\sqrt{2}^{\sqrt{2}}$ 是无理数，针对两种情形讨论来证明这个命题.*

证明: 若 $\sqrt{2}^{\sqrt{2}}$ 是有理数，我们取 $a=b=\sqrt{2}$, 则原命题成立。若 $\sqrt{2}^{\sqrt{2}}$ 是无理数，我们取 $a=\sqrt{2}^{\sqrt{2}},b=\sqrt{2}$, 则 $a^b=(\sqrt{2})^2=2$, 是有理数，此时原命题也成立。 □

练习 0.2.3. 证明 $2^{2^{n-1}}-1$ 必然至少有互异的 $n-1$ 个奇素因数.（因此第 n 个素数 p_n 一定小于等于 $2^{2^{n-1}}$.）

证明: 我们先证明：对任意 $m \neq n, 2^{2^m}+1$ 与 $2^{2^n}+1$ 互素。不妨设 $m<n$，注意到：

$$(2^{2^m}+1, 2^{2^n}+1) = (2^{2^m}+1, 2^{2^n}-2^{2^m}) = (2^{2^m}+1, 2^{2^n-2^m}-1).$$

另一方面，当 k 是奇数时，$X^k+1=(X+1)(X^{k-1}-X^{k-2}+X^{k-3}+\cdots+1)$; 取 $X=2^{2^m}, k=2^{n-m}-1$, 我们就知道存在正整数 t 使得

$$2^{2^n-2^m} = 2^{2^m(2^{n-m}-1)} = (2^{2^m}+1)t-1,$$

因此

$$(2^{2^m}+1, 2^{2^n}+1) = (2^{2^m}+1, (2^{2^m}+1)t-2) = (2^{2^m}+1, 2) = 1.$$

下面我们用数学归纳法来证明原命题。当 $n=1$ 时，$2^{2^{n-1}}-1=2^{2^0}-1=1$，素因子个数为 $0=n-1$；因此命题成立。现在假设命题对于 n 成立（此即归纳假设），我们需要证明命题对于 $n+1$ 也成立。注意到

$$2^{2^{(n+1)-1}}-1 = (2^{2^{n-1}}+1)(2^{2^{n-1}}-1),$$

根据归纳假设，$2^{2^{n-1}}-1$ 至少有 $n-1$ 个素因子；而 $2^{2^{n-1}}+1>1$ 至少有 1 个素因子，根据我们之前证明的互素性，$2^{2^{n-1}}+1$ 的素因子和 $2^{2^{n-1}}-1$ 的素因子互不相同，因此 $2^{2^{(n+1)-1}}-1$ 的素因子个数比 $2^{2^{n-1}}-1$ 的素因子个数至少多 1，因此 $2^{2^{(n+1)-1}}-1$ 至少有 $(n-1)+1=(n+1)-1$ 个素因子，即命题对于 $n+1$ 也成立。根据数学归纳法，原命题成立。

□

练习 0.2.4. 假设有一个 $a \times b$ 的巧克力排块，我们需要将它掰成 1×1 的小块，在掰的同时会被打分. 假设掰一次将 $x+y$ 块掰成了 x 块和 y 块，则得分 xy. 证明：当巧克力被完全掰成 1×1 的小块时，总得分永远为 $\frac{1}{2}ab(ab-1)$.

证明: 我们对 $n=ab$ 作归纳，即巧克力的小块个数做归纳。$n=1,2$ 时，直接验证即可。假设命题对于所有 $\leq n$ 的整数成立，我们需要证明他对于 $n+1$ 成立。现在给定一块由 $n+1$ 小块组成的巧克力，我们掰第一次，分成了两块，块数分别为 n_1, n_2，则 $1 \leq n_1, n_2 \leq n, n_1+n_2=n+1$，根据题意，这一步得分为 n_1n_2. 根据归纳假设，我们将得到的两块巧克力完全掰完后，得分分别为 $\frac{1}{2}n_1(n_1-1), \frac{1}{2}n_2(n_2-1)$. 于是总得分为

$$n_1n_2+\frac{1}{2}n_1(n_1-1)+\frac{1}{2}n_2(n_2-1) = \frac{1}{2}(n_1^2-n_1+n_2^2-n_2+2n_1n_2) = \frac{1}{2}(n+1)n.$$

故原命题成立。 □

0.3 映　　射

练习 0.3.1. 判断下列映射是否为单射、满射、双射，并写出双射的逆映射:

1. $f\colon \mathbb{R}\to\mathbb{R},\quad x\mapsto x+1.$ *2.* $f\colon \mathbb{R}\to\mathbb{R},\quad x\mapsto 2x.$ *3.* $f\colon \mathbb{R}\to\mathbb{R},\quad x\mapsto 3.$

4. $f\colon \mathbb{R}\to\mathbb{R},\quad x\mapsto x^2.$ *5.* $f\colon (-\infty,0]\to\mathbb{R},\quad x\mapsto x^2.$ *6.* $f\colon \mathbb{R}\to(0,\infty),\quad x\mapsto e^x.$

7. $f\colon \left[-\dfrac{3\pi}{2},-\dfrac{\pi}{2}\right]\to[-1,1],\quad x\mapsto \sin x.$

证明*:* 1. 双射，逆映射为 $x\mapsto x-1$.

2. 双射，逆映射为 $x\mapsto \frac{x}{2}$.
3. 都不是。
4. 都不是。
5. 单射。
6. 双射，逆映射为: $(0,\infty)\to\mathbb{R}, x\mapsto \ln(x)$.
7. 双射。注意到反正弦函数：$\arcsin:[-1,1]\to[-\frac{\pi}{2},\frac{\pi}{2}]$, f 的逆映射是 $x\mapsto -\arcsin(x)-\pi$

□

练习 0.3.2. 对复合映射 $f=g\circ h$，证明或举出反例.

1. g,h 都是满射，则 f 是满射. *2.* g,h 都是单射，则 f 是单射.

3. h 不是满射，则 f 不是满射. *4.* g 不是满射，则 f 不是满射.

5. g 不是单射，则 f 不是单射. *6.* h 不是单射，则 f 不是单射.

7. g,h 都不是双射，则 f 不是双射.

证明*:* 设 $h:X\to Y, g:Y\to Z$, 则 $f=g\circ h:X\to Z$.

1. f 是满射。对任意 $z\in Z$, 因为 g 满射，存在 $y\in Y$ s.t. $g(y)=z$; 又 h 满，存在 $x\in X$ s.t. $h(x)=y$; 则 $f(x)=z$.
2. f 是单射。假设 $f(x_1)=f(x_2)$, 即 $g(h(x_1))=g(h(x_2))$, 因为 g 单，则 $h(x_1)=h(x_2)$; 再由 h 单，则 $x_1=x_2$.
3. 反例：$X=Y=\mathbb{R}, Z=[0,+\infty), h(x)=x^2, g(y)=|y|$.
4. 我们只需要注意到 f 的像集是 g 的像集的子集，因此 g 不满则 f 不满。
5. 反例：$X=Z=[0,+\infty), Y=\mathbb{R}, h(x)=x, g(y)=|y|$, 但 $f:[0,+\infty)\to[0,+\infty), f(x)=x$ 是单射。
6. h 不单，则存在 $x_1\neq x_2$ 使得 $h(x_1)=h(x_2)$, 则 $f(x_1)=f(x_2)$, 即 f 不单。
7. 反例：$X=Z=[0,+\infty), Y=\mathbb{R}, h(x)=x, g(y)=|y|$, 但 $f:[0,+\infty)\to[0,+\infty), f(x)=x$ 是双射。

□

练习 0.3.3.

1. 在不改变对应法则和定义域的前提下，$\mathbb{R}$ 上的变换 $f(x)=x^2+2x+3$，把陪域换成哪个集合，得到的映射是满射?

2. 证明映射 $f\colon[0,1]\to\mathbb{R}, f(x)=x$ 是单射.

证明*:* 1. 只需将陪域换成值域即可，因此我们把陪域换成 $[2,+\infty)$.

2. 假设 $x_1, x_2 \in [0,1]$ 满足 $f(x_1) = f(x_2)$; 根据定义：$f(x_1) = x_1, f(x_2) = x_2$, 因此 $x_1 = x_2$, 从而 f 是单射。

□

练习 0.3.4. 下列 $\mathbb{R}$ 上的变换，哪些满足交换律 $f \circ g = g \circ f$?

1. $f(x) = x + 1, g(x) = 2x$.
2. $f(x) = x^2, g(x) = x^3$.
3. $f(x) = 2^x, g(x) = 3^x$.
4. $f(x) = 2x + 1, g(x) = 3x + 2$.
5. $f(x) = 2x + 1, g(x) = 3x + 1$.
6. $f(x) = \sin x, g(x) = \cos x$.

证明: 1. $(f \circ g)(x) = 2x + 1, (g \circ f)(x) = 2x + 2$;

2. $(f \circ g)(x) = (x^3)^2 = x^6, (g \circ f)(x) = (x^2)^3 = x^6$;
3. $(f \circ g)(x) = 2^{3^x}, (g \circ f)(x) = 3^{2^x}$，它们不相等：$(f \circ g)(1) = 2^3 = 8, (g \circ f)(1) = 3^2 = 9$;
4. $(f \circ g)(x) = 2(3x + 2) + 1 = 6x + 5, (g \circ f)(x) = 3(2x + 1) + 2 = 6x + 5$;
5. $(f \circ g)(x) = 2(3x + 1) + 1 = 6x + 3, (g \circ f)(x) = 3(2x + 1) + 1 = 6x + 4$;
6. $(f \circ g)(x) = \sin(\cos(x)), (g \circ f)(x) = \cos(\sin x)$.

□

练习 0.3.5. 对 $\mathbb{R}$ 上的变换 $f(x) = 2x + 1$ 和 $g(x) = ax + b$，求实数 a, b，使得 $f \circ g = g \circ f$.

证明: $(f \circ g)(x) = 2(ax + b) + 1 = 2ax + (2b + 1), (g \circ f)(x) = a(2x + 1) + b = 2ax + (a + b)$; $f \circ g = g \circ f \Leftrightarrow 2b + 1 = a + b \Leftrightarrow a = b + 1$. □

练习 0.3.6. 在化简函数 $\arccos \circ \sin\colon \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [0, \pi]$ 的下列步骤中，分别利用了映射的哪些性质?

1. 如果 $y = (\arccos \circ \sin)(x)$，则 $\cos y = (\cos \circ (\arccos \circ \sin))(x)$.
2. $\cos \circ (\arccos \circ \sin) = (\cos \circ \arccos) \circ \sin$.
3. $(\cos \circ \arccos) \circ \sin = \mathrm{id}_{[-1,1]} \circ \sin$.
4. $\mathrm{id}_{[-1,1]} \circ \sin = \sin$.

综上得 $\sin x = \cos y$，由此推断化简结果.

证明: 1. 用了映射的良定性，即给定一个对象，它对应的像是唯一一个。

2. 映射复合的结合律。
3. $\cos : [0, \pi] \to [-1, 1]$ 和 $\arccos : [-1, 1] \to [0, \pi]$ 互逆。
4. 恒同变换和映射的复合不改变映射。

□

练习 0.3.7. 用数学归纳法证明，任取有限个映射 $f_1, f_2, \ldots, f_n$，如果对 $1 \leq i \leq n - 1$，f_{i+1} 的定义域都等于 f_i 的陪域，则复合映射 $f_n \circ \cdots \circ f_2 \circ f_1$ 不因映射复合的计算次序而改变.

证明: $n=1,,2$ 无须证明，$n=3$ 即映射复合的结合律。假设命题对于任意 $k\leq n$ 个映射成立，我们需要证明命题对于 $n+1$ 个映射成立；事实上我们证明任何次序的复合都等于 $f_{n+1}\circ(f_n\circ(\cdots))$。当我们有 $n+1$ 个映射，复合次序形如 $(\cdots f_{m+1})\circ(f_m\cdots)$ 其中 $1\leq m\leq n$; 若则根据归纳假设，复合映射等于

$$\Big(f_{n+1}\circ(f_n\circ(\cdots(\cdots f_{m+1})))\Big)\circ\Big(f_m\circ(f_{m-1}\circ(\cdots(\cdots f_1)))\Big)$$

我们把 $(f_m\circ(f_{m-1}\circ(\cdots(\cdots f_1)))$ 作为一个整体看成一个映射，再利用归纳假设就得到原复合映射等于 $f_{n+1}\circ(f_n\circ(\cdots))$。 □

练习 0.3.8. 证明，任意映射 f 都存在分解 $f=g\circ h$，其中 h 是满射，g 是单射.

证明: 设 $f:X\to Y$, 我们取 $Z=f(X)\subseteq Y$, $h:X\to f(X)$ 定义为 $h(x)=f(x)$; $g:f(X)\to Y$ 定义为 $g(y)=y$. □

练习 0.3.9. 给定映射 h,g 和 $f=g\circ h$，证明，若 f 是双射，则 h 是单射，g 是满射.

证明: 设 $h:X\to Y, g:Y\to Z$, 则 $f=g\circ h:X\to Z$.

若 $h(x_1)=h(x_2)$, 则 $g(h(x_1))=g(h(x_2))$, 即 $f(x_1)=f(x_2)$; 而 f 是单射，则 $x_1=x_2$, 从而 h 是单射。

对于任意 $z\in Z$，f 是满射，于是存在 $x\in X$ 使得 $f(x)=z$. 令 $y=h(x)$, 则 $g(y)=g(h(x))=f(x)=z$, 即 g 是满射。 □

练习 0.3.10. 证明命题 *0.3.2*

证明:

1. 若存在 $g:Y\to X$ 使得 $g\circ f=\mathrm{id}_X$, 则由上一题知 f 是单射。反之，若 f 是单射，我们把 Y 分成两部分：$f(X),(Y\backslash f(X))$. 对任意 $y\in f(X)$, 因为 f 是单射，存在唯一的 $x_y\in X$ 使得 $f(x_y)=y$, 对这样的 y, 我们令 $g(y)=x_y$; 对于 $y\in Y\backslash f(X)$, 我们可以取 $g(y)$ 是 X 中的任意一个元素，例如我们可以固定 $x_0\in X$, 令 $g(y)=x_0,\forall y\in Y\backslash f(X)$. 直接验证可知 $g\circ f=\mathrm{id}_X$.
2. 若存在 $g:Y\to X$ 使得 $f\circ g=\mathrm{id}_Y$, 则由上一题知 f 是满射。反之，若 f 是满射，对任意的 $y\in Y$, 都存在某个 $x\in X$ 使得 $f(x)=y$, 我们任意取定一个即可，记作 x_y, 令 $g(y)=x_y$. 直接验证可知 $f\circ g=\mathrm{id}_Y$.
3. 若 f 可逆，则由上一题知 f 既是单射又是满射。反之，若 f 是双射，对任意 $y\in Y$, 都存在唯一的元素 $x_y\in X$ 使得 $f(x_y)=y$; 我们令 $g(y)=x_y$. 直接验证 $g\circ f=\mathrm{id}_X, f\circ g=\mathrm{id}_Y$ 即可.

□

第一章　线性映射和矩阵

1.1　基 本 概 念

练习 1.1.1. 如图 1.1所示，钟表表盘上对应整点有 12 个向量，其中 12 点对应向量为 $\begin{bmatrix}0\\1\end{bmatrix}$.

1. 计算 12 个向量之和.
2. 不计 2 点方向向量，计算其他 11 个向量之和.
3. 假设这 12 个向量的起点从表盘中心移到 6 点，则 12 点对应向量变为 $\begin{bmatrix}0\\2\end{bmatrix}$. 计算此时 12 个向量之和.

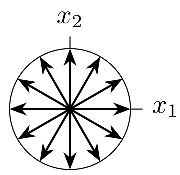


图 1.1: 钟表

证明: 1. 这十二个向量中终点是对径点的两个向量加起来是零向量，因此这十二个向量的和是零向量。

2. 由于其它 11 个向量加上两点钟方向的向量是零向量，因此这 11 个向量的和是 2 点钟方向的向量的加法逆元，即 8 点钟方向的向量，因此和等于 $\begin{bmatrix}-\sqrt{3}/2\\-1/2\end{bmatrix}$

3. 此时，终点互为对径点的两个向量的和是 $\begin{bmatrix}0\\2\end{bmatrix}$，因此这 12 个向量的和是 $\begin{bmatrix}0\\12\end{bmatrix}$.

□

练习 1.1.2. 如果平面上的向量 $\begin{bmatrix}a\\b\end{bmatrix}$ 与 $\begin{bmatrix}c\\d\end{bmatrix}$ 共线，那么 $\begin{bmatrix}a\\c\end{bmatrix}$ 与 $\begin{bmatrix}b\\d\end{bmatrix}$ 是否共线?

若向量 $\boldsymbol{a} \neq 0$, 则向量 $\boldsymbol{a}, \boldsymbol{b}$ 共线，当且仅当存在 λ 使得 $\boldsymbol{b} = \lambda\boldsymbol{a}$.

证明: 我们不妨设 $\begin{bmatrix}a\\b\end{bmatrix} \neq 0$, 则 $\begin{bmatrix}a\\b\end{bmatrix}$ 与 $\begin{bmatrix}c\\d\end{bmatrix}$ 共线当且仅当存在 λ 使得 $\begin{bmatrix}c\\d\end{bmatrix} = \lambda\begin{bmatrix}a\\b\end{bmatrix}$，即 $c = \lambda a, d = \lambda b$. 若 $a \neq 0$, 则 $\begin{bmatrix}b\\d\end{bmatrix} = \frac{b}{a}\begin{bmatrix}a\\c\end{bmatrix}$; 若 $b \neq 0$, 则 $\begin{bmatrix}a\\c\end{bmatrix} = \frac{a}{b}\begin{bmatrix}b\\d\end{bmatrix}$. □

练习 1.1.3. 证明命题 1.1.5.

证明: 设 $\boldsymbol{a} = \begin{bmatrix}a_1\\\vdots\\a_m\end{bmatrix}, \boldsymbol{b} = \begin{bmatrix}b_1\\\vdots\\b_m\end{bmatrix}, \boldsymbol{c} = \begin{bmatrix}c_1\\\vdots\\c_m\end{bmatrix},$

1. $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c} = \begin{bmatrix} a_1+b_1 \\ \vdots \\ a_m+b_m \end{bmatrix} + \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} a_1+b_1+c_1 \\ \vdots \\ a_m+b_m+c_m \end{bmatrix}$, $\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c}) = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} + \begin{bmatrix} b_1+c_1 \\ \vdots \\ b_m+c_m \end{bmatrix} = \begin{bmatrix} a_1+b_1+c_1 \\ \vdots \\ a_m+b_m+c_m \end{bmatrix}$.
2. 由定义直接验证 $\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}=\begin{bmatrix} a_1+b_1 \\ \vdots \\ a_m+b_m \end{bmatrix}$
3. 略
4. 略
5. 略
6. 由定义直接验证 $(kl)\boldsymbol{a}=k(l\boldsymbol{a})=\begin{bmatrix} kla_1 \\ \vdots \\ kla_m \end{bmatrix}$
7. 由定义直接验证 $(k+l)\boldsymbol{a}=k\boldsymbol{a}+l\boldsymbol{a}=\begin{bmatrix} (k+l)a_1 \\ \vdots \\ (k+l)a_m \end{bmatrix}$
8. $k(\boldsymbol{a}+\boldsymbol{b})=k\begin{bmatrix} a_1+b_1 \\ \vdots \\ a_m+b_m \end{bmatrix} = \begin{bmatrix} k(a_1+b_1) \\ \vdots \\ k(a_m+b_m) \end{bmatrix}$, $k\boldsymbol{a}+k\boldsymbol{b} = \begin{bmatrix} ka_1 \\ \vdots \\ ka_m \end{bmatrix} + \begin{bmatrix} kb_1 \\ \vdots \\ kb_m \end{bmatrix} = \begin{bmatrix} ka_1+kb_1 \\ \vdots \\ ka_m+kb_m \end{bmatrix}$

□

练习 1.1.4.

1. *如果只用加法交换律，不用加法结合律，那么* $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}$ *有多少种与之相等的表达式？*

2. *如果只用加法结合律，不用加法交换律，那么* $((\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c})+\boldsymbol{d}$ *有多少种与之相等的表达式？*

证明:

1. 只用加法交换律只有 4 种：$(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}$，$(\boldsymbol{b}+\boldsymbol{a})+\boldsymbol{c}$, $\boldsymbol{c}+(\boldsymbol{a}+\boldsymbol{b})$, $\boldsymbol{c}+(\boldsymbol{b}+\boldsymbol{a})$. 我们也可以这么思考：只用加法交换律的话，每个加号处都能改变顺序，只有 2 个加号，因此是 2^2 种。
2. 三个向量相加，只用加法结合律，只有两种相同的表达式。4 个向量相加，若 $\boldsymbol{a}$ 或者 $\boldsymbol{d}$ 最后一步相加，那么各有两种相同的表达式，若 $\boldsymbol{a}$ 和 $\boldsymbol{d}$ 都不是最后一步相加，那么有一种表达式，因此一共是 5 种相同的表达式。

□

练习 1.1.5. *判断下列映射是否为线性映射：*

1. $f\colon \mathbb{R}\to\mathbb{R},\quad x\mapsto x+1.$ *2.* $f\colon \mathbb{R}\to\mathbb{R},\quad x\mapsto 2x.$ *3.* $f\colon \mathbb{R}\to\mathbb{R},\quad x\mapsto 0.$

4. $f\colon \mathbb{R}\to\mathbb{R},\quad x\mapsto 1.$ *5.* $f\colon \mathbb{R}\to\mathbb{R},\quad x\mapsto x^2.$ *6.* $f\colon \mathbb{R}\to\mathbb{R},\quad x\mapsto 2^x.$

7. $f\colon \mathbb{R}^2\to\mathbb{R}^3,\quad \begin{bmatrix}x\\y\end{bmatrix}\mapsto\begin{bmatrix}x+y\\y-x\\2x\end{bmatrix}$ *8.*. $f\colon \mathbb{R}^2\to\mathbb{R}^3,\quad \begin{bmatrix}x\\y\end{bmatrix}\mapsto\begin{bmatrix}x+1\\y-x\\2x\end{bmatrix}.$

证明: 2,3,7 是，1,4,5,6,8 不是。 □

练习 1.1.6. 设 $f\colon \mathbb{R}\to\mathbb{R}$ 是线性映射，证明存在实数 k，使得 $f(x)=kx$.

证明: 令 $k=f(1)$, f 是线性映射，因此对任意 $x\in\mathbb{R}$ $f(x)=f(x\cdot 1)=xf(1)=kx$. □

练习 1.1.7. 设映射 $f\colon \mathbb{R}^3\to\mathbb{R}^2, f\left(\begin{bmatrix}x\\y\\z\end{bmatrix}\right)=\begin{bmatrix}g(x,y,z)\\h(x,y,z)\end{bmatrix}$，其中 $g,h\colon \mathbb{R}^3\to\mathbb{R}$. 证明 f 是线性映射当且仅当 g,h 都是线性映射.

证明:

$$
\begin{aligned}
&f\text{是线性映射}\\
\Leftrightarrow\quad & f\left(\lambda\begin{bmatrix}x_1\\y_1\\z_2\end{bmatrix}+\mu\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix}\right)=\lambda f\left(\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix}\right)+\mu f\left(\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}\right)\\
\Leftrightarrow\quad & \begin{bmatrix}g(\lambda x_1+\mu x_2,\lambda y_1+\mu y_2,\lambda z_1+\mu z_2)\\h(\lambda x_1+\mu x_2,\lambda y_1+\mu y_2,\lambda z_1+\mu z_2)\end{bmatrix}=\lambda\begin{bmatrix}g(x_1,y_1,z_1)\\h(x_1,y_1,z_1)\end{bmatrix}+\mu\begin{bmatrix}g(x_2,y_2,z_2)\\h(x_2,y_2,z_2)\end{bmatrix},\\
\Leftrightarrow\quad & \begin{cases} g(\lambda x_1+\mu x_2,\lambda y_1+\mu y_2,\lambda z_1+\mu z_2) &=\lambda g(x_1,y_1,z_1)+\mu g(x_2,y_2,z_2),\\ h(\lambda x_1+\mu x_2,\lambda y_1+\mu y_2,\lambda z_1+\mu z_2) &=\lambda h(x_1,y_1,z_1)+\mu h(x_2,y_2,z_2)\end{cases}\\
\Leftrightarrow\quad & g,h\text{都是线性映射}
\end{aligned}
$$

□

练习 1.1.8. 设 $\boldsymbol{x}_1=\begin{bmatrix}1\\0\end{bmatrix},\boldsymbol{x}_2=\begin{bmatrix}1\\1\end{bmatrix},\boldsymbol{x}_3=\begin{bmatrix}0\\1\end{bmatrix},\boldsymbol{b}_1=\begin{bmatrix}1\\1\\0\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}1\\0\\1\end{bmatrix},\boldsymbol{b}_3=\begin{bmatrix}0\\1\\1\end{bmatrix}$，是否存在线性映射 $f\colon \mathbb{R}^2\to\mathbb{R}^3$，满足 $f(\boldsymbol{x}_i)=\boldsymbol{b}_i, i=1,2,3$?

证明: 假设存在，注意到 $\boldsymbol{x}_2=\boldsymbol{x}_1+\boldsymbol{x}_3$, 则 $f(\boldsymbol{x}_2)=f(\boldsymbol{x}_1)+f(\boldsymbol{x}_3)$, 即 $\boldsymbol{b}_1+\boldsymbol{b}_3=\boldsymbol{b}_2$, 而 $\boldsymbol{b}_1+\boldsymbol{b}_3=\begin{bmatrix}1\\2\\1\end{bmatrix}\neq\boldsymbol{b}_2$, 矛盾，故不存在。 □

练习 1.1.9. 判断下列映射是否为线性映射:

1. 给定 $\boldsymbol{a}\in\mathbb{R}^m$，$f\colon \mathbb{R}\to\mathbb{R}^m, f(k)=k\boldsymbol{a}$.
2. 给定实数 k，$f\colon \mathbb{R}^m\to\mathbb{R}^m, f(\boldsymbol{a})=k\boldsymbol{a}$.
3. $f\colon \mathbb{R}^{m+1}\to\mathbb{R}^m, f\left(\begin{bmatrix}a_1\\\vdots\\a_m\\a_{m+1}\end{bmatrix}\right)=a_{m+1}\begin{bmatrix}a_1\\\vdots\\a_m\end{bmatrix}.$

证明: 1,2 是线性映射，直接验证定义即可。3 不是线性映射，因为 $f(k\boldsymbol{a}) = k^2 f(\boldsymbol{a}) \neq kf(\boldsymbol{a})$. □

练习 1.1.10. 给定三维向量 $\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$，定义：

1. 二者**点积**为 $\boldsymbol{a} \cdot \boldsymbol{b} := a_1b_1 + a_2b_2 + a_3b_3$.
2. 二者**叉积**为 $\boldsymbol{a} \times \boldsymbol{b} := \begin{bmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{bmatrix}$.

那么给定 $\boldsymbol{a} \in \mathbb{R}^3$，映射 $f\colon \mathbb{R}^3 \to \mathbb{R}, \boldsymbol{b} \mapsto \boldsymbol{a} \cdot \boldsymbol{b}$ 和 $g\colon \mathbb{R}^3 \to \mathbb{R}^3, \boldsymbol{b} \mapsto \boldsymbol{a} \times \boldsymbol{b}$ 是否为线性映射?

证明: 给定 $\boldsymbol{a}$, f, g 都是线性映射。

$$f(k\boldsymbol{b}) = \boldsymbol{a} \cdot (k\boldsymbol{b}) = a_1(kb_1) + a_2(kb_2) + a_3(kb_3) = k(a_1b_1 + a_2b_2 + a_3b_3) = k\boldsymbol{a} \cdot \boldsymbol{b} = kf(\boldsymbol{b}),$$

$$f(\boldsymbol{b} + \boldsymbol{b}') = a_1(b_1 + b_1') + a_2(b_2 + b_2') + a_3(b_3 + b_3') = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{b}' = f(\boldsymbol{b}) + f(\boldsymbol{b}').$$

$$g(k\boldsymbol{b}) = \begin{bmatrix} a_2kb_3 - a_3kb_2 \\ a_3kb_1 - a_1kb_3 \\ a_1kb_2 - a_2kb_1 \end{bmatrix} = k \begin{bmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{bmatrix} = kg(\boldsymbol{b}),$$

$$g(\boldsymbol{b} + \boldsymbol{b}') = \begin{bmatrix} a_2(b_3 + b_3') - a_3(b_2 + b_2') \\ a_3(b_1 + b_1') - a_1(b_3 + b_3') \\ a_1(b_2 + b_2') - a_2(b_1 + b_1') \end{bmatrix} = \begin{bmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{bmatrix} + \begin{bmatrix} a_2b_3' - a_3b_2' \\ a_3b_1' - a_1b_3' \\ a_1b_2' - a_2b_1' \end{bmatrix} = g(\boldsymbol{b}) + g(\boldsymbol{b}').$$

□

练习 1.1.11. 给定平面上任意面积为 1 的三角形，经过下列变换之后，其面积是否确定？如果是，面积为多少？（不需严谨证明，猜测答案即可.）

1. 旋转变换.
2. 反射变换.
3. 对 x_2 投影的投影变换.
4. x_1 方向拉伸 3 倍，x_2 方向不变的伸缩变换.
5. 把 x_1 方向的 3 倍加到 x_2 方向上，保持 x_1 方向不变的错切变换.

证明:

1. 不变
2. 不变
3. 变为零
4. 变为 3
5. 不变

□

练习 1.1.12. 设 $\mathbb{R}^2$ 上的变换 $f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -y + 1 \\ x + 2 \end{bmatrix}$.

1. 证明 f 不是线性变换.

2. 构造分解 $f = g\circ h$，其中 g 是 $\mathbb{R}^2$ 上的平移变换，h 是 $\mathbb{R}^2$ 上的线性变换.（这种平移与线性变换的复合称为仿射变换. 注意，平移变换并不是线性变换.）

证明*:*

证法很多，比如验证 $f\left(\begin{bmatrix}2\\0\end{bmatrix}\right) \neq 2f\left(\begin{bmatrix}1\\0\end{bmatrix}\right)$, $f\left(\begin{bmatrix}1\\0\end{bmatrix}\right) \neq -f\left(\begin{bmatrix}-1\\0\end{bmatrix}\right)$ 等等。

$h\left(\begin{bmatrix}x\\y\end{bmatrix}\right) = \left(\begin{bmatrix}-y\\x\end{bmatrix}\right), g\left(\begin{bmatrix}x\\y\end{bmatrix}\right) = \left(\begin{bmatrix}x+1\\y+2\end{bmatrix}\right)$ □

练习 1.1.13. 设连续函数 $f\colon \mathbb{R}\to\mathbb{R}$ 满足 $f(a+b) = f(a)+f(b)$，f 是否为线性映射?（需要微积分知识）

证明*:* 首先, 取 $a=b=0$, 则 $f(0)=f(0)+f(0)$, 因此 $f(0)=0$. 再取 $b=-a$, 则 $f(a)+f(-a)=f(0)=0$, 因此 $f(-a)=-f(a)$.

断言对任意正整数 n 和任意实数 a, $f(na)=nf(a)$. 事实上 $f(na)=f((n-1)a+a)=f((n-1)a)+f(a)=f((n-2)a+a)+f(a)=f((n-2)a)+2f(a)=\cdots=nf(a)$. 特别的, 对任意正整数 n, 我们有 $f(n)=nf(1)$; 那么对于任意负整数 n, $f(n)=-f(-n)=-(-n)f(1)=nf(1)$.

进一步地，对任意有理数 $q=\frac{m}{n}$, $m=nq$, 则 $f(m)=mf(1)$ 且 $f(nq)=nf(q)$, 因此 $f(q)=\frac{m}{n}f(1)=qf(1)$.

对任意实数 a, 我们取有理数 $a_n\to a$, 那么因为 f 连续，

$$f(a)=\lim_{n\to\infty} f(a_n)=\lim_{n\to\infty} a_n f(1)=af(1).$$

从而 f 是线性映射。 □

1.2 线性映射的表示矩阵

练习 1.2.1. 将下列向量 $\boldsymbol{b}$ 写成矩阵和向量乘积的形式:

1. $\boldsymbol{b} = 2\begin{bmatrix}1\\1\end{bmatrix} + 4\begin{bmatrix}0\\1\end{bmatrix} + 5\begin{bmatrix}1\\0\end{bmatrix}$.

2. $\boldsymbol{b} = 5\begin{bmatrix}1\\2\\3\\4\\5\end{bmatrix} + 4\begin{bmatrix}5\\4\\3\\2\\1\end{bmatrix}$.

3. $\boldsymbol{b} = \begin{bmatrix}2b+a+c\\c-b\\a+b+c\\a+b\end{bmatrix}$，其中 a,b,c 为常数.

4. $\boldsymbol{b} = f\left(\begin{bmatrix}1\\2\\3\\4\\5\end{bmatrix}\right)$，其中 $f\colon\mathbb{R}^5\to\mathbb{R}^5$ 是线性变换，满足 $f(\boldsymbol{e}_k)=k\boldsymbol{e}_{6-k}, k=1,2,\ldots,5$.

5. 假设如果某天下雨，则第二天下雨概率为 0.8；如果当天不下雨，则第二天下雨概率为 0.3. 已知当天有一半的概率会下雨，令 $\boldsymbol{b} \in \mathbb{R}^2$，其两个分量分别是明天下雨和不下雨的概率.

证明: 1. $\boldsymbol{b} = 2\begin{bmatrix}1\\1\end{bmatrix} + 4\begin{bmatrix}0\\1\end{bmatrix} + 5\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}1&0&1\\1&1&0\end{bmatrix}\begin{bmatrix}2\\4\\5\end{bmatrix}$.

2. $\boldsymbol{b} = 5\begin{bmatrix}1\\2\\3\\4\\5\end{bmatrix} + 4\begin{bmatrix}5\\4\\3\\2\\1\end{bmatrix} = \begin{bmatrix}1&5\\2&4\\3&3\\4&2\\5&1\end{bmatrix}\begin{bmatrix}5\\4\end{bmatrix}$.

3. $\boldsymbol{b} = \begin{bmatrix}2b+a+c\\c-b\\a+b+c\\a+b\end{bmatrix} = \begin{bmatrix}a\\0\\a\\a\end{bmatrix} + \begin{bmatrix}2b\\-b\\b\\b\end{bmatrix} + \begin{bmatrix}c\\c\\c\\0\end{bmatrix} = a\begin{bmatrix}1\\0\\1\\1\end{bmatrix} + b\begin{bmatrix}2\\-1\\1\\1\end{bmatrix} + c\begin{bmatrix}1\\1\\1\\0\end{bmatrix} = \begin{bmatrix}1&2&1\\0&-1&1\\1&1&1\\1&1&0\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix}$

4.

$$\begin{aligned}\boldsymbol{b} &= f\left(\begin{bmatrix}1\\2\\3\\4\\5\end{bmatrix}\right)\\ &= f(\boldsymbol{e}_1) + 2f(\boldsymbol{e}_2) + 3f(\boldsymbol{e}_3) + 4f(\boldsymbol{e}_4) + 5f(\boldsymbol{e}_5)\\ &= \boldsymbol{e}_5 + 4\boldsymbol{e}_4 + 9\boldsymbol{e}_3 + 16\boldsymbol{e}_2 + 25\boldsymbol{e}_5\\ &= \begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{e}_4 & \boldsymbol{e}_5\end{bmatrix}\begin{bmatrix}25\\16\\9\\4\\1\end{bmatrix}\end{aligned}$$

5. $b_1 = 0.8\cdot 0.5 + 0.3\cdot 0.5, b_2 = 0.2\cdot 0.5 + 0.7\cdot 0.5$, 故 $\boldsymbol{b} = \begin{bmatrix}0.8&0.3\\0.2&0.7\end{bmatrix}\begin{bmatrix}0.5\\0.5\end{bmatrix}$.

□

练习 1.2.2. 判断下列矩阵和向量的乘积是否良定义[1]. 在可以计算时，先将其写成列向量的线性组合，再进行计算：

[1]所谓良定义，就是说定义不产生矛盾. 这里就是指可以按定义计算.

1. $\begin{bmatrix}1&0\\1&1\\1&1\end{bmatrix}\begin{bmatrix}1\\3\end{bmatrix}$.

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

3. $\begin{bmatrix}1&0\\1&1\end{bmatrix}\begin{bmatrix}1\\3\end{bmatrix}$.

4. $\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}1\\3\\5\end{bmatrix}$.

5. $\begin{bmatrix}1&0&0&0\\1&1&0&0\\1&1&1&0\\1&1&1&1\end{bmatrix}\begin{bmatrix}1\\3\\5\\7\end{bmatrix}$.

6. $\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}\begin{bmatrix}1\\-2\\1\end{bmatrix}$.

7. $\begin{bmatrix}1&4&7\\2&5&8\\3&6&9\end{bmatrix}\begin{bmatrix}1\\-2\\1\end{bmatrix}$.

8. $\begin{bmatrix}1&1&0\\3&2&1\\7&4&3\end{bmatrix}\begin{bmatrix}1\\-1\\-1\end{bmatrix}$.

9. $\begin{bmatrix}1&3&7\\1&2&4\\0&1&3\end{bmatrix}\begin{bmatrix}1\\-1\\-1\end{bmatrix}$.

10. $\begin{bmatrix}1&3&7\\1&2&4\\0&1&3\end{bmatrix}\begin{bmatrix}2\\-3\\1\end{bmatrix}$.

证明*:* 1. 良定义 $\begin{bmatrix}1&0\\1&1\\1&1\end{bmatrix}\begin{bmatrix}1\\3\end{bmatrix}=1\begin{bmatrix}1\\1\\1\end{bmatrix}+3\begin{bmatrix}0\\1\\1\end{bmatrix}=\begin{bmatrix}1\\4\\4\end{bmatrix}$.

2. 非良定义.

3. 良定义 $\begin{bmatrix}1&0\\1&1\end{bmatrix}\begin{bmatrix}1\\3\end{bmatrix}=1\begin{bmatrix}1\\1\end{bmatrix}+3\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}1\\4\end{bmatrix}$.

4. 良定义 $\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}1\\3\\5\end{bmatrix}=\begin{bmatrix}1\\1\\1\end{bmatrix}+3\begin{bmatrix}0\\1\\1\end{bmatrix}+5\begin{bmatrix}0\\0\\1\end{bmatrix}=\begin{bmatrix}1\\4\\9\end{bmatrix}$.

5. 良定义 $\begin{bmatrix}1&0&0&0\\1&1&0&0\\1&1&1&0\\1&1&1&1\end{bmatrix}\begin{bmatrix}1\\3\\5\\7\end{bmatrix}=\begin{bmatrix}1\\1\\1\\1\end{bmatrix}+3\begin{bmatrix}0\\1\\1\\1\end{bmatrix}+5\begin{bmatrix}0\\0\\1\\1\end{bmatrix}+7\begin{bmatrix}0\\0\\0\\1\end{bmatrix}=\begin{bmatrix}1\\4\\9\\16\end{bmatrix}$.

6. 良定义 $\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}\begin{bmatrix}1\\-2\\1\end{bmatrix}=\begin{bmatrix}1\\4\\7\end{bmatrix}-2\begin{bmatrix}2\\5\\8\end{bmatrix}+\begin{bmatrix}3\\6\\9\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$.

7. 良定义 $\begin{bmatrix}1&4&7\\2&5&8\\3&6&9\end{bmatrix}\begin{bmatrix}1\\-2\\1\end{bmatrix}=\begin{bmatrix}1\\2\\3\end{bmatrix}-2\begin{bmatrix}4\\5\\6\end{bmatrix}+\begin{bmatrix}7\\8\\9\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$.

8. 良定义 $\begin{bmatrix}1&1&0\\3&2&1\\7&4&3\end{bmatrix}\begin{bmatrix}1\\-1\\-1\end{bmatrix}=\begin{bmatrix}1\\3\\7\end{bmatrix}-\begin{bmatrix}1\\2\\4\end{bmatrix}-\begin{bmatrix}0\\1\\3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$.

9. 良定义 $\begin{bmatrix}1&3&7\\1&2&4\\0&1&3\end{bmatrix}\begin{bmatrix}1\\-1\\-1\end{bmatrix}=\begin{bmatrix}1\\1\\0\end{bmatrix}-2\begin{bmatrix}3\\2\\1\end{bmatrix}+\begin{bmatrix}7\\4\\3\end{bmatrix}=\begin{bmatrix}-9\\-5\\-4\end{bmatrix}$.

10. 良定义 $\begin{bmatrix}1&3&7\\1&2&4\\0&1&3\end{bmatrix}\begin{bmatrix}2\\-3\\1\end{bmatrix}=2\begin{bmatrix}1\\1\\0\end{bmatrix}-3\begin{bmatrix}3\\2\\1\end{bmatrix}+\begin{bmatrix}7\\4\\3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$.

□

练习 1.2.3. 设 $A=\begin{bmatrix}0.8&0.3\\0.2&0.7\end{bmatrix}, \boldsymbol{u}_0=\begin{bmatrix}1\\0\end{bmatrix}, \boldsymbol{v}_0=\begin{bmatrix}1\\3\end{bmatrix}$. 对任意自然数 i，令 $\boldsymbol{u}_{i+1}=A\boldsymbol{u}_i, \boldsymbol{v}_{i+1}=A\boldsymbol{v}_i$.

1. 对 $i=1,2,3,4$，计算 $\boldsymbol{u}_i,\boldsymbol{v}_i$.

2. 猜测 $\lim\limits_{i\to\infty}\boldsymbol{u}_i, \lim\limits_{i\to\infty}\boldsymbol{v}_i$.

3. 任取初始向量 $\boldsymbol{w}_0$，猜测极限 $\lim\limits_{i\to\infty}\boldsymbol{w}_i$，其中 $\boldsymbol{w}_{i+1}=A\boldsymbol{w}_i$；不同初始向量得到的极限在同一条直线上吗？

证明*:* 1.

$$\begin{aligned}
\boldsymbol{u}_1&=\begin{bmatrix}0.8&0.3\\0.2&0.7\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}0.8\\0.2\end{bmatrix}, & \boldsymbol{v}_1&=\begin{bmatrix}0.8&0.3\\0.2&0.7\end{bmatrix}\begin{bmatrix}1\\3\end{bmatrix}=\begin{bmatrix}1.7\\2.3\end{bmatrix}\\
\boldsymbol{u}_1&=\begin{bmatrix}0.8&0.3\\0.2&0.7\end{bmatrix}\begin{bmatrix}0.8\\0.2\end{bmatrix}=\begin{bmatrix}0.7\\0.3\end{bmatrix}, & \boldsymbol{v}_2&=\begin{bmatrix}0.8&0.3\\0.2&0.7\end{bmatrix}\begin{bmatrix}1.7\\2.3\end{bmatrix}=\begin{bmatrix}2.05\\1.95\end{bmatrix}\\
\boldsymbol{u}_3&=\begin{bmatrix}0.8&0.3\\0.2&0.7\end{bmatrix}\begin{bmatrix}0.7\\0.3\end{bmatrix}=\begin{bmatrix}0.63\\0.37\end{bmatrix}, & \boldsymbol{v}_3&=\begin{bmatrix}0.8&0.3\\0.2&0.7\end{bmatrix}\begin{bmatrix}2.05\\1.95\end{bmatrix}=\begin{bmatrix}2.225\\1.775\end{bmatrix}\\
\boldsymbol{u}_4&=\begin{bmatrix}0.8&0.3\\0.2&0.7\end{bmatrix}\begin{bmatrix}0.63\\0.37\end{bmatrix}=\begin{bmatrix}0.615\\0.385\end{bmatrix} & \boldsymbol{v}_4&=\begin{bmatrix}0.8&0.3\\0.2&0.7\end{bmatrix}\begin{bmatrix}2.225\\1.775\end{bmatrix}=\begin{bmatrix}2.3125\\1.6875\end{bmatrix}
\end{aligned}$$

2. 一个合理的猜测是 $\lim\limits_{i\to\infty}\boldsymbol{u}_i=\begin{bmatrix}0.6\\0.4\end{bmatrix}, \lim\limits_{i\to\infty}\boldsymbol{v}_i=\begin{bmatrix}2.4\\1.6\end{bmatrix}$.

3. 猜测极限 $\lim\limits_{i\to\infty}\boldsymbol{w}_i$ 存在，并且设为 $\boldsymbol{w}=\begin{bmatrix}x\\y\end{bmatrix}$, 那么 $A\boldsymbol{w}=\boldsymbol{w}$ 并且满足 $x+y=x_0+y_0$, 其中 $\boldsymbol{w}_0=\begin{bmatrix}x_0\\y_0\end{bmatrix}$. $A\boldsymbol{w}=\boldsymbol{w}$ 推出 $2x=3y$, 于是我们可以解出 $x=\frac{3}{5}(x_0+y_0), y=\frac{2}{5}(x_0+y_0)$.

□

练习 1.2.4. 设线性变换 f 的表示矩阵为 $A=\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}$.

1. 令 $\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}=\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$，将 y_1,y_2,y_3 分别用 x_1,x_2,x_3 表达出来.

2. 将 x_1,x_2,x_3 分别用 y_1,y_2,y_3 表达出来.

3. 找到矩阵 B，使得 $\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=B\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$.

4. 设 g 是由矩阵 B 决定的线性变换，证明 f,g 互为逆变换.

证明: 1. 直接计算矩阵乘法，就有 $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_1+x_2 \\ x_1+x_2+x_3 \end{bmatrix}$，即 $\begin{cases} y_1 &= x_1 \\ y_2 &= x_1+x_2 \\ y_3 &= x_1+x_2+x_3 \end{cases}$.

2. 直接反解即可：$\begin{cases} x_1 &= y_1 \\ x_2 &= -y_1+y_2 \\ x_3 &= -y_2+y_3 \end{cases}$.

3. $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ -y_1+y_2 \\ -y_2+y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$，故 $B = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

4. 根据定义，$(f\circ g)\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = f\left(B\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}\right) = f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, 同时

$$(g\circ f)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = g\left(\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

故 $f\circ g = g\circ f = \mathrm{id}_{\mathbb{R}^3}$, 从而互逆。

□

练习 1.2.5. 计算下列线性映射 f 的表示矩阵:

1. f 为 xy 平面向 y 轴的投影变换.

2. $f\colon \mathbb{R}^3 \to \mathbb{R}$，$f(\boldsymbol{x}) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \boldsymbol{x}$, 点积的定义见练习 *1.1.10*

3. $f\colon \mathbb{R}^3 \to \mathbb{R}^3$，$f(\boldsymbol{x}) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \times \boldsymbol{x}$. 叉积的定义见练习 *1.1.10*

4. $f\colon \mathbb{R}^4 \to \mathbb{R}^4$，$f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4+x_2 \end{bmatrix}$.

5. $f\colon \mathbb{R}^4 \to \mathbb{R}^4$，$f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} x_4 \\ x_2 \\ x_3 \\ x_1 \end{bmatrix}$.

6. $f\colon \mathbb{R}^m \to \mathbb{R}^m$，$f\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}\right) = \begin{bmatrix} x_m \\ x_{m-1} \\ \vdots \\ x_1 \end{bmatrix}$.

证明: 1. 按照定义，$f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 0 \\ y \end{bmatrix}$, 因此 f 的表示矩阵是 $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

2. $f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = x_1 + 2x_2 + 3x_3 = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, 因此 f 的表示矩阵为 $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$

3. 按照叉积的定义，$f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 2x_3 - 3x_2 \\ 3x_1 - x_3 \\ x_2 - 2x_1 \end{bmatrix} = \begin{bmatrix} 0 & -3 & 2 \\ 3 & 0 & -1 \\ -2 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. 因此 f 的表示矩阵为 $\begin{bmatrix} 0 & -3 & 2 \\ 3 & 0 & -1 \\ -2 & 1 & 0 \end{bmatrix}$.

4. $f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 + x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$, 因此 f 的表示矩阵为 $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$.

5. $f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} x_4 \\ x_2 \\ x_3 \\ x_1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$, 因此 f 的表示矩阵为 $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$.

6. $f\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}\right) = \begin{bmatrix} x_m \\ x_{m-1} \\ \vdots \\ x_1 \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 1 \\ & \iddots & \\ 1 & \cdots & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$. 因此 f 的表示矩阵为 $\begin{bmatrix} 0 & \cdots & 1 \\ & \iddots & \\ 1 & \cdots & 0 \end{bmatrix}$.

□

练习 1.2.6. *考虑桥墩载荷问题，其中 l_1, l_2, d 作为常数.*

1. *映射 f 输入 F_1, F_2 得到输出 f_1, f_2，写出 f 的表示矩阵.*

2. *以桥梁的左端为支点，逆时针方向的力矩为 $df_2 - l_1F_1 - l_2F_2$. 以桥梁的右端为支点或者以桥梁的中点作为支点，都能类似地得到逆时针方向的力矩. 假设映射 f 的输入为 F_1, F_2, f_1, f_2，而输出为桥梁的左端、中点和右端的逆时针方向的力矩，写出 f 的表示矩阵.*

证明: 1. 由于

$$f\left(\begin{bmatrix} F_1 \\ F_2 \end{bmatrix}\right) = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} \frac{d-l_1}{d}F_1 + \frac{d-l_2}{d}F_2 \\ \frac{l_1}{d}F_1 + \frac{l_2}{d}F_2 \end{bmatrix} = \begin{bmatrix} \frac{d-_1}{d} & \frac{d-l_2}{d} \\ \frac{l_1}{d} & \frac{l_2}{d} \end{bmatrix}\begin{bmatrix} F_1 \\ F_2 \end{bmatrix}$$

因此 f 的表示矩阵是 $\begin{bmatrix} \frac{d-_1}{d} & \frac{d-l_2}{d} \\ \frac{l_1}{d} & \frac{l_2}{d} \end{bmatrix}$ 中点为支点时逆时针方向的力矩是 $-\frac{d}{2}f_1+(\frac{d}{2}-l_1)F_1+(\frac{d}{2}-l_2)F_2+\frac{d}{2}f_2$. 右端点为支点时逆时针方向的力矩是 $-df_1+(d-l_1)F_1+(d-l_2)F_2$, 所以

$$f\left(\begin{bmatrix} F_1 \\ F_2 \\ f_1 \\ f_2 \end{bmatrix}\right)=\begin{bmatrix} df_2-l_1F_1-l_2F_2 \\ -\frac{d}{2}f_1+(\frac{d}{2}-l_1)F_1+(\frac{d}{2}-l_2)F_2-\frac{d}{2}f_2 \\ -df_1+(d-l_1)F_1+(d-l_2)F_2 \end{bmatrix}=\begin{bmatrix} -l_1 & -l_2 & 0 & d \\ \frac{d}{2}-l_1 & \frac{d}{2}-l_2 & -\frac{d}{2} & \frac{d}{2} \\ d-l_1 & d-l_2 & -d & 0 \end{bmatrix}\begin{bmatrix} F_1 \\ F_2 \\ f_1 \\ f_2 \end{bmatrix}$$

因此 f 对应的矩阵是 $\begin{bmatrix} -l_1 & -l_2 & 0 & d \\ \frac{d}{2}-l_1 & \frac{d}{2}-l_2 & -\frac{d}{2} & \frac{d}{2} \\ d-l_1 & d-l_2 & -d & 0 \end{bmatrix}$

□

练习 1.2.7. 设 xy 平面 $\mathbb{R}^2$ 上的变换 f 是下列三个变换的复合: 先绕原点逆时针旋转 $\frac{\pi}{6}$; 然后进行一个保持 y 坐标不变, 同时将 y 坐标的两倍加到 x 坐标上的错切; 最后再沿着直线 $x+y=0$ 反射.

1. 证明 f 是线性变换.
2. 计算 $f(\boldsymbol{e}_1)$ 和 $f(\boldsymbol{e}_2)$.
3. 写出 f 的表示矩阵.
4. 计算 $f\left(\begin{bmatrix} 3 \\ 4 \end{bmatrix}\right)$.

证明: 由于 f 是如下三个线性变换的复合:

- $R_{\frac{\pi}{6}}:\mathbb{R}^2\to\mathbb{R}^2, \begin{bmatrix} x \\ y \end{bmatrix}\mapsto\begin{bmatrix} \frac{\sqrt{3}x}{2}-\frac{y}{2} \\ \frac{x}{2}+\frac{\sqrt{3}y}{2} \end{bmatrix}$
- $S_2:\mathbb{R}^2\to\mathbb{R}^2, \begin{bmatrix} x \\ y \end{bmatrix}\mapsto\begin{bmatrix} x+2y \\ y \end{bmatrix}$
- $H_{\frac{3\pi}{4}}:\mathbb{R}^2\to\mathbb{R}^2, \begin{bmatrix} x \\ y \end{bmatrix}\mapsto\begin{bmatrix} -y \\ -x \end{bmatrix}$

复合起来就是

$$\begin{bmatrix} x \\ y \end{bmatrix}\mapsto\begin{bmatrix} \frac{\sqrt{3}x}{2}-\frac{y}{2} \\ \frac{x}{2}+\frac{\sqrt{3}y}{2} \end{bmatrix}\mapsto\begin{bmatrix} \frac{(\sqrt{3}+2)x}{2}+\frac{(2\sqrt{3}-1)y}{2} \\ \frac{x}{2}+\frac{\sqrt{3}y}{2} \end{bmatrix}\mapsto\begin{bmatrix} -\left(\frac{x}{2}+\frac{\sqrt{3}y}{2}\right) \\ -\left(\frac{(\sqrt{3}+2)x}{2}+\frac{(2\sqrt{3}-1)y}{2}\right) \end{bmatrix}$$

$$f(\boldsymbol{e}_1)=f\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)=\begin{bmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}+2}{2} \end{bmatrix}, f(\boldsymbol{e}_2)=f\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)=\begin{bmatrix} -\frac{\sqrt{3}}{2} \\ -\frac{2\sqrt{3}-1}{2} \end{bmatrix}$$

因为

$$f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)=\begin{bmatrix} -\left(\frac{x}{2}+\frac{\sqrt{3}y}{2}\right) \\ -\left(\frac{(\sqrt{3}+2)x}{2}+\frac{(2\sqrt{3}-1)y}{2}\right) \end{bmatrix}=\begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}+2}{2} & -\frac{2\sqrt{3}-1}{2} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

从而 f 的表示矩阵是 $\begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}+2}{2} & -\frac{2\sqrt{3}-1}{2} \end{bmatrix}$.

$$f\left(\begin{bmatrix} 3 \\ 4 \end{bmatrix}\right) = \begin{bmatrix} -\left(\frac{3}{2} + \frac{\sqrt{3}\cdot 4}{2}\right) \\ -\left(\frac{(\sqrt{3}+2)3}{2} + \frac{(2\sqrt{3}-1)4}{2}\right) \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} - 2\sqrt{3} \\ -1 - \frac{11\sqrt{3}}{2} \end{bmatrix} \qquad \square$$

练习 1.2.8.

1. 幻方矩阵是指元素分别是 $1, 2, \ldots, 9$ 的三阶矩阵 M，且每行、每列以及两条对角线上的三个元素之和都相同. 求 $M\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 的所有可能值.

2. 数独矩阵是指 9 阶矩阵 M，从上到下、从左到右依次分成九个 3×3 的子矩阵，且每行、每列以及九个子矩阵中的元素都是 $1, 2, \ldots, 9$. 求 $M\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ 的所有可能值.

证明: 1. 由于 M 的所有元素的和是 45, 因此每行、每列上三个元素相加的和都是 15. 而 $M\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 第 $i(i = 1, 2, 3)$ 个分量就是 M 第 i 行元素的和，因此所有可能的值唯一，就是 15

2. 对于数独矩阵而言，$M\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ 第 $i(i = 1, 2, 3)$ 个分量依然是 M 第 i 行元素的和，因此是 45.

$\square$

练习 1.2.9. 证明，如果 n 阶方阵 A 对任意 n 维列向量 $\boldsymbol{x}$，都有 $A\boldsymbol{x} = \boldsymbol{0}$，则 A 是元素全为 0 的矩阵.

证明: 我们选取 $\boldsymbol{x} = \boldsymbol{e}_i(1 \le i \le n)$, 那么 $A\boldsymbol{e}_i$ 恰好是 A 的第 i 个列向量。根据条件，$A\boldsymbol{e}_i = 0(1 \le i \le n)$, 即 A 的所有列向量都是零向量。因此 A 是元素全为 0 的矩阵。 $\square$

练习 1.2.10. 证明，如果 n 阶方阵 A 对任意 n 维列向量 $\boldsymbol{x}$，都存在依赖于 $\boldsymbol{x}$ 的常数 $c(\boldsymbol{x})$，满足 $A\boldsymbol{x} = c(\boldsymbol{x})\boldsymbol{x}$，则存在常数 c，使得 $A = cI_n$.

证明: 我们选取 $\boldsymbol{x} = \boldsymbol{e}_i(1 \le i \le n)$, 那么根据题意，存在常数 $c_i = c(\boldsymbol{e}_i)$, 满足 $A\boldsymbol{e}_i = c_i\boldsymbol{e}_i$. 另一方面，对 $i \neq j$, 我们取 $\boldsymbol{x} = \boldsymbol{e}_i + \boldsymbol{e}_j$, 那么根据题意，存在常数 $c_{ij} = c(\boldsymbol{e}_i + \boldsymbol{e}_j)$, 满足 $A(\boldsymbol{e}_i + \boldsymbol{e}_j) = c_{ij}(\boldsymbol{e}_i + \boldsymbol{e}_j)$.

另一方面，$A(\boldsymbol{e}_i + \boldsymbol{e}_j) = A\boldsymbol{e}_i + A\boldsymbol{e}_j$, 因此 $c_{ij}(\boldsymbol{e}_i + \boldsymbol{e}_j) = c_i\boldsymbol{e}_i + c_j\boldsymbol{e}_j$, 这就有 $c_i = c_j = c_{ij}$. 这对于任意的 $i \neq j$ 都成立，因此所有的 c_i 都相等，我们统一记作 c.

于是对任意的 i, $A\boldsymbol{e}_i = c\boldsymbol{e}_i, \forall 1 \le i \le n$. 还是注意到 $A\boldsymbol{e}_i$ 恰好是 A 的第 i 个列向量，因此 $A = cI_n$. $\square$

1.3 线性方程组

练习 1.3.1. 把下列矩阵化为行简化阶梯形，并回答问题：

1. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}$，化简后第一列是否为主列？

2. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}$，化简后第二列是否为主列？

3. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}$，化简后第三列是否为主列？

4. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}$，化简后第四列是否为主列？

5. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}$，化简后第五列是否为主列？

作初等行变换老老实实的按照固定的算法去计算是最为保险的，但是往往会数字比较大，略显笨拙。对于具体问题观察数字的规律呢，虽然灵活，但又容易出错，所以需要细致地考虑，掌握其中的诀窍。对于分析某一列是否是主列的问题，其实完全不需要把矩阵作初等行变换变成行（简化）阶梯形矩阵之后再作判断

证明:

1. 第一列有非零元，所以化简后第一列一定是主列。
2. 矩阵的前两行是行阶梯形矩阵，且第二行第二列非零，因此化简后第二列是主列。
3. 矩阵的前两行是行阶梯形矩阵，但第二行第二列是零，第二行第三列非零，虽然最后结果是矩阵的第三列是主列，但是我们要作更多地思考，这一点需要和第 2 小题做一下比较。考虑矩阵的第三行，用第一行消元之后，它可能有如下几种情况 (其中的第一个 $*$ 是非零的，其他的 $*$ 不定)：
   - $\begin{bmatrix} 0 & * & * & * & * \end{bmatrix}$, 此时前三列都是主列；
   - $\begin{bmatrix} 0 & 0 & * & * & * \end{bmatrix}$，此时第一列第三列是主列，还有没有其他主列需要对得到的矩阵的第二三行继续初等行变换。
   - $\begin{bmatrix} 0 & 0 & 0 & * & * \end{bmatrix}$ 此时第一三四列是主列；
   - $\begin{bmatrix} 0 & 0 & 0 & 0 & * \end{bmatrix}$ 此时第一三五列是主列；
   - $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ 此时只有第一三列是主列。
4. 和上题的讨论类数，第四列一定是主列。
5. 和第 3 小题的讨论类似，第五列一定是主列。

□

练习 1.3.2. 下列线性方程组有解时，找到所有解；线性方程组无解时，对线性方程组做初等变换，得到矛盾表达式 $0=1$：

1. $\begin{bmatrix}1&2\\3&4\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}3\\7\end{bmatrix}$.

2. $\begin{bmatrix}1&3&5\\2&4&6\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$.

3. $\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$.

4. $\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}$.

5. $\begin{bmatrix}1&0&1\\2&3&5\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$.

6. $\begin{bmatrix}1&2\\0&3\\1&5\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$.

7. $\begin{bmatrix}1&2\\0&3\\1&5\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}$.

8. $\begin{bmatrix}2&-1&0&0\\-1&2&-1&0\\0&-1&2&-1\\0&0&-1&2\end{bmatrix}\begin{bmatrix}x\\y\\z\\t\end{bmatrix}=\begin{bmatrix}0\\0\\0\\5\end{bmatrix}$.

9. $\begin{bmatrix}2&1&0&0\\1&2&1&0\\0&1&2&1\\0&0&1&2\end{bmatrix}\begin{bmatrix}x\\y\\z\\t\end{bmatrix}=\begin{bmatrix}0\\0\\0\\5\end{bmatrix}$.

这样的问题是最基础的问题，也没有其他方法，就是对增广矩阵做初等行变换，化简成行简化阶梯形矩阵即可。

证明*:* 1.

$$\begin{bmatrix}1&2&3\\3&4&7\end{bmatrix}\longrightarrow\begin{bmatrix}1&2&3\\0&-2&-2\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&1\\0&1&1\end{bmatrix}$$

因此原方程组有唯一解：$x=y=1$.

2. 齐次方程组一定有解：所有未知数都是 0. 我们需要确定这是不是唯一解。

$$\begin{bmatrix}1&3&5&0\\2&4&6&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&3&5&0\\0&-2&-4&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&-1&0\\0&1&2&0\end{bmatrix}$$

因此原方程组所有解为：$x=z, y=-2z$.

3.

$$\begin{bmatrix}1&2&0\\3&4&0\\5&6&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&2&0\\0&-2&0\\0&-4&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$$

因此原方程组有唯一解 $x=y=0$.

4.
$$\begin{bmatrix}1&2&1\\3&4&0\\5&6&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&2&1\\0&-2&-3\\0&-4&-5\end{bmatrix}\longrightarrow\begin{bmatrix}1&2&1\\0&2&3\\0&0&1\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

因此原方程组无解。

5.
$$\begin{bmatrix}1&0&1&0\\2&3&5&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&1&0\\0&3&3&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&1&0\\0&1&1&0\end{bmatrix}$$

因此原方程组所有解为 $x=y=-z$.

6.
$$\begin{bmatrix}1&2&0\\0&3&0\\1&5&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&2&0\\0&1&0\\0&3&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$$

因此原方程组有唯一解 $x=y=0$.

7.
$$\begin{bmatrix}1&2&1\\0&3&0\\1&5&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&2&1\\0&1&0\\0&3&-10\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

因此原方程组无解。

8.
$$\begin{bmatrix}2&-1&0&0&0\\-1&2&-1&0&0\\0&-1&2&-1&0\\0&0&-1&2&5\end{bmatrix}\longrightarrow\begin{bmatrix}2&-1&0&0&0\\0&\frac{3}{2}&-1&0&0\\0&0&\frac{4}{3}&-1&0\\0&0&0&\frac{5}{4}&5\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&0&0&1\\0&1&0&0&2\\0&0&1&0&3\\0&0&0&1&4\end{bmatrix}$$

因此原方程组有唯一解 $x=1,y=2,z=3,t=4$.

9.
$$\begin{bmatrix}2&1&0&0&0\\1&2&1&0&0\\0&1&2&1&0\\0&0&1&2&5\end{bmatrix}\longrightarrow\begin{bmatrix}2&1&0&0&0\\0&\frac{3}{2}&1&0&0\\0&0&\frac{4}{3}&1&0\\0&0&0&\frac{5}{4}&5\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&0&0&-1\\0&1&0&0&2\\0&0&1&0&-3\\0&0&0&1&4\end{bmatrix}$$

因此原方程组有唯一解 $x=-1,y=2,z=-3,t=4$.

□

练习 1.3.3. 将下列问题首先化成 $A\boldsymbol{x}=\boldsymbol{b}$ 的形式，然后求所有解：

1. 对 $\begin{bmatrix}1&2&3&4&5\\6&7&8&9&10\\11&12&13&14&15\end{bmatrix}$，如何将其第三列写成前两列的线性组合？

2. 对 $\begin{bmatrix}1&2&3&4&5\\6&7&8&9&10\\11&12&13&14&15\end{bmatrix}$，如何将其第四列写成前两列的线性组合？

3. 对 $\begin{bmatrix}1&2&3&4&5\\6&7&8&9&10\\11&12&13&14&15\end{bmatrix}$，如何将其第五列写成前两列的线性组合？

证明: 1. 按照题意，我们列出关于未知数 x_1, x_2 的方程组

$$x_1\begin{bmatrix}1\\6\\11\end{bmatrix}+x_2\begin{bmatrix}2\\7\\12\end{bmatrix}=\begin{bmatrix}3\\8\\13\end{bmatrix}，即 \begin{bmatrix}1&2\\6&7\\11&12\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}3\\8\\13\end{bmatrix}$$

解方程组得 $x_1=-1, x_2=2$.

2. 按照题意，我们列出关于未知数 x_1, x_2 的方程组

$$x_1\begin{bmatrix}2\\7\\12\end{bmatrix}+x_2\begin{bmatrix}3\\8\\13\end{bmatrix}=\begin{bmatrix}4\\9\\14\end{bmatrix}，即 \begin{bmatrix}2&3\\7&8\\12&13\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}4\\9\\14\end{bmatrix}$$

解方程组得 $x_1=-1, x_2=2$.

3. 按照题意，我们列出关于未知数 x_1, x_2 的方程组

$$x_1\begin{bmatrix}3\\8\\13\end{bmatrix}+x_2\begin{bmatrix}4\\9\\14\end{bmatrix}=\begin{bmatrix}5\\10\\15\end{bmatrix}，即 \begin{bmatrix}3&4\\8&9\\13&14\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}5\\10\\15\end{bmatrix}$$

解方程组得 $x_1=-1, x_2=2$.

□

练习 1.3.4. 求证：齐次线性方程组

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 = & 0 & ① \\ a_{21}x_1 & +a_{22}x_2 = & 0 & ② \end{cases}$$

有非零解当且仅当 $a_{11}a_{22}-a_{12}a_{21}=0$.

证明: 若 $a_{11}a_{22}-a_{12}a_{21}\neq 0$, ① $\times a_{22}-$ ② $\times a_{12}$ 得 $(a_{11}a_{22}-a_{12}a_{21})x_1=0$，① $\times a_{21}-$ ② $\times a_{11}$ 得 $(a_{11}a_{22}-a_{12}a_{21})x_2=0$，因此方程组只有零解：$x_1=x_2=0$。

若 $a_{11}a_{22}-a_{12}a_{21}=0$, 我们需要证明方程组有非零解。若系数全为零，那么是显然的。若系数不全为零，我们不妨设 $a_{11}\neq 0$, 那么我们取 $x_1=a_{12}, x_2=-a_{11}$ 是方程组的非零解。 □

练习 1.3.5. 求满足下列条件的常数 b, c：

1. 方程组 $\begin{bmatrix}1&2\\3&b\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}3\\7\end{bmatrix}$ 无解.

2. 方程组 $\begin{bmatrix}3&2\\6&4\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}10\\c\end{bmatrix}$ 无解.

3. 方程组 $\begin{bmatrix}2 & 5 & 1\\ 4 & b & 1\\ 0 & 1 & -1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}0\\ 2\\ 3\end{bmatrix}$ 无解.

4. 方程组 $\begin{bmatrix}b & 3\\ 3 & b\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}6\\ -6\end{bmatrix}$ 有无穷多组解.

5. 方程组 $\begin{bmatrix}2 & b\\ 4 & 8\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}16\\ c\end{bmatrix}$ 有无穷多组解.

6. 方程组 $\begin{bmatrix}1 & b & 0\\ 1 & -2 & -1\\ 0 & 1 & 1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$ 有非零解.

7. 方程组 $\begin{bmatrix}b & 2 & 3\\ b & b & 4\\ b & b & b\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$ 有非零解（求三个不同的 b).

证明*:* 1.

$$\begin{bmatrix}1 & 2 & 3\\ 3 & b & 7\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 2 & 3\\ 0 & b-6 & -2\end{bmatrix}$$

方程组无解，当且仅当 $b-6=0$, 因此 $b=6$.

2.

$$\begin{bmatrix}3 & 2 & 10\\ 6 & 4 & c\end{bmatrix}\longrightarrow\begin{bmatrix}3 & 2 & 10\\ 0 & 0 & c-20\end{bmatrix}$$

方程组无解，当且仅当 $c-20\neq 0$, 因此 $c\neq 20$.

3.

$$\begin{bmatrix}2 & 5 & 1 & 0\\ 4 & b & 1 & 2\\ 0 & 1 & -1 & 3\end{bmatrix}\longrightarrow\begin{bmatrix}2 & 5 & 1 & 0\\ 0 & 1 & -1 & 3\\ 0 & b-10 & -1 & 2\end{bmatrix}\longrightarrow\begin{bmatrix}2 & 0 & 6 & -15\\ 0 & 1 & -1 & 3\\ 0 & 0 & b-11 & 32-3b\end{bmatrix}$$

方程组无解，当且仅当 $b-11=0, 32-3b\neq 0$, 因此 $b=11$.

4.

$$\begin{bmatrix}b & 3 & 6\\ 3 & b & -6\end{bmatrix}\longrightarrow\begin{bmatrix}3 & b & -6\\ 0 & 3-\frac{b^2}{3} & 6+2b\end{bmatrix}$$

方程组有无穷多组解，当且仅当 $3-\frac{b^2}{3}=0, 6+2b=0$, 因此 $b=-3$.

5.

$$\begin{bmatrix}2 & b & 16\\ 4 & 8 & c\end{bmatrix}\longrightarrow\begin{bmatrix}4 & 8 & c\\ 0 & b-4 & 16-\frac{c}{2}\end{bmatrix}$$

方程组有无穷多组解，当且仅当 $b-4=0\ 16-\frac{c}{2}=0$, 因此 $b=4, c=32$.

6.

$$\begin{bmatrix}1 & b & 0 & 0\\ 1 & -2 & -1 & 0\\ 0 & 1 & 1 & 0\end{bmatrix}\longrightarrow\begin{bmatrix}1 & b & 0 & 0\\ 0 & 1 & 1 & 0\\ 0 & -2-b & -1 & 0\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 0 & -b & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 1+b & 0\end{bmatrix}$$

方程组有非零解，当且仅当 $1+b=0$, 因此 $b=-1$.

7.

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

显然，当 $b=4$ 时，方程组有非零解。若 $b \neq 4$, 我们还可以继续做初等行变换：

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

当 $b=2$ 是，方程组也有非零解。若 $b \neq 2,4$, 我们还可以继续做初等行变换：

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

此时只有当 $b=0$ 时，方程组才有非零解。

总之，当且仅当 $b=0,2,4$ 时，原方程组有非零解。 □

练习 1.3.6. 设齐次线性方程组

$$\begin{cases} 2x_1 & +x_2 & +x_3 & =0 \\ ax_1 & & -x_3 & =0 \\ -x_1 & & +3x_3 & =0 \end{cases}$$

有非零解，求 a 的值，并求出所有的解.

证明: 我们还是作初等行变换：

$$\begin{bmatrix} 2 & 1 & 1 & 0 \\ a & 0 & -1 & 0 \\ -1 & 0 & 3 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 7 & 0 \\ 0 & 0 & 3a-1 & 0 \end{bmatrix}$$

因此方程组有非零解当且仅当 $3a-1=0$, 即 $a=\frac{1}{3}$. □

练习 1.3.7. 在如下关于 x_1, x_2, x_3 的线性方程组中，讨论在 p 取不同值时方程组是否有解，并在有解时，求出所有的解.

$$\begin{cases} px_1 & +x_2 & +x_3 & =1 \\ x_1 & +px_2 & +x_3 & =p \\ x_1 & +x_2 & +px_3 & =p^2. \end{cases}$$

证明: 我们还是作初等行变换：

$$\begin{bmatrix} p & 1 & 1 & 1 \\ 1 & p & 1 & p \\ 1 & 1 & p & p^2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & p & p^2 \\ 0 & p-1 & 1-p & p-p^2 \\ 0 & 1-p & 1-p^2 & 1-p^3 \end{bmatrix}$$

若 $p=1$, 则原方程组即 $x_1+x_2+x_3=1$。当 $p\neq 1$ 时，我们可以继续做初等行变换

$$\begin{bmatrix}1 & 1 & p & p^2\\ 0 & p-1 & 1-p & p-p^2\\ 0 & 1-p & 1-p^2 & 1-p^3\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 1 & p & p^2\\ 0 & 1 & -1 & -p\\ 0 & 1 & 1+p & 1+p+p^2\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 1 & p & p^2\\ 0 & 1 & -1 & -p\\ 0 & 0 & 2+p & 1+2p+p^2\end{bmatrix}$$

若 $p=-2$, 最后一行为 $\begin{bmatrix}0 & 0 & 0 & 1\end{bmatrix}$, 因此方程组无解。若 $p\neq 1,-2$ 时，我们继续作初等行变换

$$\begin{bmatrix}1 & 1 & p & p^2\\ 0 & 1 & -1 & -p\\ 0 & 0 & 2+p & 1+2p+p^2\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 0 & 0 & -\frac{p+1}{p+2}\\ 0 & 1 & 0 & \frac{1}{p+2}\\ 0 & 0 & 1 & \frac{1+2p+p^2}{p+2}\end{bmatrix}$$

此时原方程组有有唯一解 $x_1=-\frac{p+1}{p+2}, x_2=\frac{1}{p+2}, x_3=\frac{1+2p+p^2}{p+2}$. □

练习 1.3.8. 设 $A=\begin{bmatrix}1 & 1 & 0\\ -1 & 0 & 1\\ 0 & -1 & -1\end{bmatrix}, \boldsymbol{b}=\begin{bmatrix}b_1\\ b_2\\ b_3\end{bmatrix}$. 求证：

1. $A\boldsymbol{x}=\boldsymbol{b}$ 有解当且仅当 $b_1+b_2+b_3=0$.
2. 齐次线性方程组 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集是 $\Big\{k\boldsymbol{x}_1 : k\in\mathbb{R}\Big\}, \boldsymbol{x}_1=\begin{bmatrix}1\\ -1\\ 1\end{bmatrix}$.
3. 当 $A\boldsymbol{x}=\boldsymbol{b}$ 有解时，设 $\boldsymbol{x}_0$ 是一个解，则解集是 $\{\boldsymbol{x}_0+k\boldsymbol{x}_1 k\in\mathbb{R}\}$.

证明: 1. 我们做初等行变换

$$\begin{bmatrix}1 & 1 & 0 & b_1\\ -1 & 0 & 1 & b_2\\ & -1 & -1 & b_3\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 1 & 0 & b_1\\ 0 & 1 & 1 & b_1+b_2\\ 0 & -1 & -1 & b_3\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 1 & 0 & b_1\\ 0 & 1 & 1 & b_1+b_2\\ 0 & 0 & 0 & b_1+b_2+b_3\end{bmatrix}$$

因此方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 有解，当且仅当 $b_1+b_2+b_3=0$.

2. 我们不需要重复做初等行变换了，只需要在上一小题中取 $b_1=b_2=b_3$ 即可，因此我们得到增广矩阵的行简化阶梯形矩阵为 $\begin{bmatrix}1 & 0 & -1 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}$ 因此只需 $x=z, y=-z$ 即可。

3. 对任意解 $\boldsymbol{x}$，我们有 $A(\boldsymbol{x}-\boldsymbol{x}_0)=\boldsymbol{b}-\boldsymbol{b}=0$, 因此 $\boldsymbol{x}-\boldsymbol{x}_0$ 形如 $k\boldsymbol{x}_1$.

□

练习 1.3.9. 求三阶方阵 A,使得线性方程组 $A\boldsymbol{x}=\begin{bmatrix}2\\ 4\\ 2\end{bmatrix}$ 的解集是 $\left\{\begin{bmatrix}2\\ 0\\ 0\end{bmatrix}+c\begin{bmatrix}1\\ 1\\ 0\end{bmatrix}+d\begin{bmatrix}0\\ 0\\ 1\end{bmatrix} c,d\in\mathbb{R}\right\}$.

证明: 首先我们需要判断 A 的尺寸。根据解集，我们知道 $\boldsymbol{x}$ 有三个未知数，因此 A 有 3 列, 根据方程组的形式，我们知道 A 有 3 行，因此 A 是 3-阶方阵。根据解集的形式，我们知道 A 有

一个主元，两个自由变元，那么它的行简化阶梯型矩阵形如：

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

那么我们可以假设 $A=\begin{bmatrix}a & -a & 0\\b & -b & 0\\c & -c & 0\end{bmatrix}$ 再根据 $\begin{bmatrix}2\\0\\0\end{bmatrix}$ 是 $A\boldsymbol{x}=\begin{bmatrix}2\\4\\2\end{bmatrix}$ 的一个特解知 $A\begin{bmatrix}2\\0\\0\end{bmatrix}=\begin{bmatrix}2\\4\\2\end{bmatrix}$ 知：

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

□

练习 1.3.10.

1. 构造三阶方阵，其元素各不相同，且行简化阶梯形有且只有一个主元.

2. 构造 100 阶方阵 A，所有元素非零，且行简化阶梯形恰有 99 个主元. 试描述 $A\boldsymbol{x}=0$ 的解集.

证明*:* 1. 假设这个矩阵是 A, 它的行简化阶梯形矩阵只有几个主元，那么我们可以假设它形如

$$\begin{bmatrix}a & b & c\\ma & mb & mc\\na & nb & nc\end{bmatrix}$$

结合条件 "元素互不相等", 那么 $abc\neq 0$. 我们简单凑一个即可：例如 $a=1,b=2,c=3,m=4,n=10$.

2. 构造方式有很多，以下只是提供一种可能。既然是构造一个 A, 我们不妨先构造一个行简化阶梯形矩阵，然后在通过初等行变换，让所有元素都变长非零的。它有 99 个主元，那么自由变元就只有 1 个；我们就可以令 $x_1,\cdots,x_{99}$ 是主元，x_{100} 是自由变元，这样就得到了一个可行的行简化阶梯形矩阵，

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

$A\boldsymbol{x}=0$ 的解集即 $\left\{k\begin{bmatrix}0\\\vdots\\0\\-1\\1\end{bmatrix}:k\in\mathbb{R}\right\}$ 下一步我们开始做初等行变换，我们先把所有行

都加到最后一行，然后把最后一行加到其他行，这样就得到了一个选择：

$$\begin{bmatrix} 2 & 1 & \cdots & 1 & 1 \\ 1 & 2 & \cdots & 1 & 1 \\ \vdots & \vdots & \ddots & & \vdots \\ 1 & 1 & \cdots & 2 & 2 \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix}$$

□

练习 1.3.11. 给定线性方程组 $A\boldsymbol{x} = \boldsymbol{0}$，其中 A 是 100 阶方阵. 假设 *Gauss* 消元法计算到最后一行得到 $0 = 0$.

1. 消元法是在计算 A 的行的线性组合. 因此 A 的 100 行的某个线性组合是?

2. 计算出 $0 = 0$ 说明方程组有无穷多组解. 因此 A 的 100 列的某个线性组合是?

3. 试说明 *Gauss* 消元法计算出的零行的个数和自由变量的个数相等.

证明*:* 1. 由于最后一行是 A 的行的一个线性组合，因此 A 的 100 行的某个线性组合是 0.

2. 由于方程组有无穷多解，那么我们取一个非零解 $\boldsymbol{x}_0$, 此时 $A\boldsymbol{x}_0$ 是 A 的列的线性组合。因此 A 的 100 列的某个线性组合是 0.

3. Gauss 消元法计算出的零行的个数是指 A 的行简化阶梯形矩阵中的零行的个数，每一个非零行对应一个主元，因此零行的个数和自由变元的个数都等于未知数的个数减去主元的个数。

□

练习 1.3.12. 仅用从上往下的倍加变换（即把上面的行的若干倍加到下面的行上），将下列矩阵化为阶梯形，并分析其主元的规律：

1. $\begin{bmatrix} 1 & 1 & & \\ 1 & 2 & 1 & \\ & 1 & 2 & 1 \\ & & 1 & 2 \end{bmatrix}$. *2.* $\begin{bmatrix} 2 & 1 & & \\ 1 & 2 & 1 & \\ & 1 & 2 & 1 \\ & & 1 & 2 \end{bmatrix}$. *3.* $\begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{bmatrix}$.

这些矩阵都是**三对角矩阵**，即除对角元素及与其相邻的元素外其余元素都是零的方阵.

证明*:* 1.

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

2.

$$\begin{bmatrix}2&1&&\\1&2&1&\\&1&2&1\\&&1&2\end{bmatrix}\longrightarrow\begin{bmatrix}2&1&&\\&\frac{3}{2}&1&\\&1&2&1\\&&1&2\end{bmatrix}\longrightarrow\begin{bmatrix}2&1&&\\&\frac{3}{2}&1&\\&&\frac{4}{3}&1\\&&1&2\end{bmatrix}\longrightarrow\begin{bmatrix}2&1&&\\&\frac{3}{2}&1&\\&&\frac{4}{3}&1\\&&&\frac{5}{4}\end{bmatrix}$$

3.

$$\begin{bmatrix}2&-1&&\\-1&2&-1&\\&-1&2&-1\\&&-1&2\end{bmatrix}\longrightarrow\begin{bmatrix}2&1&&\\&\frac{3}{2}&-1&\\&-1&2&-1\\&&-1&2\end{bmatrix}\longrightarrow\begin{bmatrix}2&1&&\\&\frac{3}{2}&-1&\\&&\frac{4}{3}&-1\\&&-1&2\end{bmatrix}\longrightarrow\begin{bmatrix}2&-1&&\\&\frac{3}{2}&-1&\\&&\frac{4}{3}&-1\\&&&\frac{5}{4}\end{bmatrix}$$

□

练习 1.3.13. 求解

$$\begin{cases}x_1 \qquad\qquad\qquad + x_n = 0,\\ x_1 + x_2 \qquad\qquad\quad = 0,\\ \quad x_2 + x_3 \qquad\quad = 0,\\ \qquad\quad \ddots \qquad\qquad \vdots\\ \qquad\qquad x_{n-1} + x_n = 0.\end{cases}$$

这个问题我们可以简单地把所有 $x_i(1 \leq i \leq n-1)$ 都用 x_n 表示出来，然后再解方程。但是为了演示 Gauss 消元法，我们用初等行变换来做，其本质是一样的。

证明: 由于是齐次方程组，我们对系数做初等行变换：

$$\begin{bmatrix}1&&&&1\\1&1&&&\\&1&1&&\\&&\ddots&&\vdots\\&&&1&1\end{bmatrix}\to\begin{bmatrix}1&&&&1\\&1&&&-1\\&&1&&1\\&&&\ddots&\vdots\\&&&1&1\end{bmatrix}$$

我们看最后一列，找规律，就会发现：

- 若 n 是奇数，那么通过初等行变换，系数矩阵最终转化为

$$\begin{bmatrix}1&&&&1\\&1&&&-1\\&&\ddots&&\vdots\\&&&1&-1\\&&&1&1\end{bmatrix}\to\begin{bmatrix}1&&&&1\\&1&&&-1\\&&\ddots&&\vdots\\&&&1&-1\\&&&&2\end{bmatrix}$$

因此原方程组只有零解。

- 若 n 是偶数，那么通过初等行变换，系数矩阵最终转化为

$$\begin{bmatrix} 1 & & & & 1 \\ & 1 & & & -1 \\ & & \ddots & & \vdots \\ & & & 1 & 1 \\ & & & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & & & & 1 \\ & 1 & & & -1 \\ & & \ddots & & \vdots \\ & & & 1 & 1 \\ & & & & 0 \end{bmatrix}$$

因此原方程组的解集为 $\left\{ k \begin{bmatrix} 1 \\ -1 \\ \vdots \\ -1 \end{bmatrix} : k \in \mathbb{R} \right\}$。

□

练习 1.3.14. 证明用倍加变换与倍乘变换可以实现对换变换.

证明*:* 我们只需考虑两行的矩阵：

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} \to \begin{bmatrix} \alpha \\ \beta+\alpha \end{bmatrix} \to \begin{bmatrix} -\beta \\ \beta+\alpha \end{bmatrix} \to \begin{bmatrix} -\beta \\ \alpha \end{bmatrix} \to \begin{bmatrix} \beta \\ \alpha \end{bmatrix}$$

□

练习 1.3.15. 证明定理 *1.3.7* 的第二部分.

证明*:* 我们已经完成了定理 1.3.7 的第一部分的证明，所以我们只需要证明任何阶梯形矩阵都可以用初等行变换化简为行简化阶梯型矩阵。

阶梯形矩阵和行简化阶梯形矩阵区别有两点：非零行的都是 1；每个主列除了主元以外都是 0。因此我们先通过倍乘变换，把非零行的主元都变成 1，然后通过倍加变换，把主列除主元外的其他元素变成 0. 这样就完成了。 □

练习 1.3.16. 练习 *1.2.8*中的数独矩阵经历哪些行变换或列变换后还是数独矩阵?

证明*:* 这样的行变换或列变换可以分为两类。

- 1-3 行，4-6 行，7-9 行整体作行对换，或者 1-3 列，4-6 列，7-9 列整体作列对换。
- 在 1-3 行，4-6 行，7-9 行内部分别作行的对换，或者 1-3 列，4-6 列，7-9 列内部分别作列对换。

□

练习 1.3.17 (初等列变换在线性方程组上的含义)**.** 考虑线性方程组 $\begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$. 进行下列换元，写出 x', y' 满足的方程组. 对比原方程组，其系数矩阵做了哪种初等变换?

1. $x' = y, y' = x$.

2. $x' = 2x, y' = y$.

3. $x' = x, y' = x + y$.

4. $x' = x+1, y' = y$.

证明: 我们把方程组写出来：

$$\begin{cases} x+2y &= 3 \\ 4x+5y &= 6 \end{cases}$$

1.
$$\begin{cases} 2x'+y' &= 3 \\ 5x'+4y' &= 6 \end{cases} \text{即} \begin{bmatrix} 2 & 1 \\ 5 & 4 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

   因此对系数矩阵作了交换两列的初等列变换。

2.
$$\begin{cases} \frac{1}{2}x'+y' &= 3 \\ 2x'+5y' &= 6 \end{cases} \text{即} \begin{bmatrix} \frac{1}{2} & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

   因此对系数矩阵的第一列作了倍乘变换。

3.
$$\begin{cases} -x'+2y' &= 3 \\ -x'+5y' &= 6 \end{cases} \text{即} \begin{bmatrix} -1 & 2 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

   因此对系数矩阵作了倍加列变换：将第二列的 -1 倍加到第 1 列。

4.
$$\begin{cases} x'+2y' &= 4 \\ 4x'+5y' &= 10 \end{cases} \text{即} \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 4 \\ 10 \end{bmatrix}$$

   因此系数矩阵没有改变，但是对增广矩阵作了倍加列变换：将第一列的 1 倍加到第 3 列。

□

练习 1.3.18 (方程、法向量与超平面)**.** 给定原点 O，则空间中的点 P 与向量 $\overrightarrow{OP}$ 一一对应. 空间中所有点构成的集合称为**点空间**，上述一一对应是点空间到线性空间 $\mathbb{R}^3$ 的双射. 注意，点空间与线性空间并不相同. 特别地，空间中的点 P 也可用向量 $\overrightarrow{OP} \in \mathbb{R}^3$ 表示.

两个 $\mathbb{R}^3$ 中的向量垂直当且仅当其内积为零（练习 1.1.10）. 与空间中某个平面垂直的非零向量称为该平面的**法向量**. 考虑空间中以非零向量 $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ 为法向量的平面.

1. 给定 $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^3$，分别对应该平面上的两个点，证明，$\begin{bmatrix} a & b & c \end{bmatrix} \boldsymbol{p} = \begin{bmatrix} a & b & c \end{bmatrix} \boldsymbol{q}$;
2. 设该平面经过对应于 $\boldsymbol{p} \in \mathbb{R}^3$ 的点，令 $d = \begin{bmatrix} a & b & c \end{bmatrix} \boldsymbol{p}$，证明，平面是方程 $ax+by+cz = d$ 的解集.
3. 设 $a_1x+b_1y+c_1z = d_1$ 和 $a_2x+b_2y+c_2z = d_2$ 的解集平面平行，试分析两个方程的系数间的关系.
4. 设 $\begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}$, $\begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix}$ 为两个不共线的非零向量，则 $a_1x+b_1y+c_1z = d_1$ 与 $a_2x+b_2y+c_2z = d_2$ 的解集平面不平行，因此必相交于一条直线 l.对方程做倍加变换，证明，$a_1x+b_1y+c_1z =$

d_1 与 $(a_1+a_2)x+(b_1+b_2)y+(c_1+c_2)z=d_1+d_2$ 的解集平面的交集还是直线 l.（从几何上看，倍加变换将一个平面沿相交直线旋转，因此不改变解集的交集.）

一个 m 元线性方程的解集是对应于 $\mathbb{R}^m$ 的点空间中的一个**超平面**，因此求解线性方程组等价于求若干超平面的交集.

证明: 1. $\boldsymbol{p},\boldsymbol{q}$ 在以 $\begin{bmatrix}a\\b\\c\end{bmatrix}$ 为法向量的平面上，因此 $\boldsymbol{p}-\boldsymbol{q}$ 与 $\begin{bmatrix}a\\b\\c\end{bmatrix}$ 内积为零。我们只需注意到对任意向量 $\boldsymbol{a}$, $\begin{bmatrix}a\\b\\c\end{bmatrix}\cdot\boldsymbol{a}=\begin{bmatrix}a & b & c\end{bmatrix}\boldsymbol{a}$ 即可。

2. 我们只需注意到 $\boldsymbol{q}$ 在该平面上，当且仅当 $\boldsymbol{p}-\boldsymbol{q}$ 与 $\begin{bmatrix}a\\b\\c\end{bmatrix}$ 内积为零，当且仅当 $\begin{bmatrix}a & b & c\end{bmatrix}\boldsymbol{p}=\begin{bmatrix}a & b & c\end{bmatrix}\boldsymbol{q}$。因此，若 $\boldsymbol{q}$ 坐标为 $\begin{bmatrix}x\\y\\z\end{bmatrix}$, 就有 $ax+by+cz=\begin{bmatrix}a & b & c\end{bmatrix}\boldsymbol{p}=d$.

3. 题中两个方程的解集平面平行，当且仅当则 $a_1x+b_1y+c_1z=0$ 与 $a_2x+b_2y+c_2z=0$ 有相同的解，也就是说，因此 $\begin{bmatrix}a_1\\b_1\\c_1\end{bmatrix}$ 和 $\begin{bmatrix}a_2\\b_2\\c_2\end{bmatrix}$ 共线。

4. l 即方程组

$$\begin{cases} a_1x+b_1y+c_1z &= d_1 \\ a_2x+b_2y+c_2z &= d_2 \end{cases}$$

通过 Gauss 消元法，倍加变换不改变解集，因此这个方程组和方程组

$$\begin{cases} a_1x+b_1y+c_1z &= d_1 \\ (a_1+a_2)x+(b_1+b_2)y+(c_1+c_2)z &= (d_1+d_2) \end{cases}$$

的解集相同，即后一个方程组的解集还是 l.

□

练习 1.3.19. 对下列问题，将其化成 $A\boldsymbol{x}=\boldsymbol{b}$，并找到所有解：

1. 考虑例 *1.1.* 中的桥墩载荷问题. 假设重物的重力分别为 $F_1=2,F_2=3$，放置的位置 l_1,l_2 为未知变量，桥梁长度为 5. 如果两个桥墩的载荷分别为 $f_1=3,f_2=2$，求所有可能的 $\begin{bmatrix}l_1\\l_2\end{bmatrix}$.

2. 笼中有若干鸡兔，每只鸡有一个头两条腿两只翅膀，每只兔有一个头四条腿没有翅膀. 笼中现有 4 个头、12 条腿、8 只翅膀，求鸡兔数目.

3. 设 $A=\begin{bmatrix}a & b\\c & d\end{bmatrix}$ 满足 $A\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}4\\8\end{bmatrix}$，且 A 第一列的元素之和为 2，求所有可能的 A.

4. 求平面上直线 $y=2x+3,y=-x+5$ 的交点.

5. 空间中有三个平面，分别经过点 $\begin{bmatrix}4\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\-1\end{bmatrix}, \begin{bmatrix}2\\0\\2\end{bmatrix}$，具有法向量 $\begin{bmatrix}1\\3\\5\end{bmatrix}, \begin{bmatrix}1\\2\\-3\end{bmatrix}, \begin{bmatrix}2\\5\\2\end{bmatrix}$，求这三个平面的交点（练习 1.3.18）.

6. 空间中有一条经过原点的直线，并且与向量 $\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}$ 垂直，求所有直线上的点.

7. 空间中有一个平面经过点 $\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}$，求所有与该平面垂直的向量.

证明: 1. 根据课本的例题，我们有 $\begin{bmatrix}f_1\\f_2\end{bmatrix} = \begin{bmatrix}\frac{d-l_1}{d}F_1 + \frac{d-l_2}{d}F_2\\ \frac{l_1}{d}F_1 + \frac{l_2}{d}F_2\end{bmatrix}$, 代入我们的具体数值，整理为关于 l_1, l_2 的方程即得：

$$\begin{bmatrix}-2 & -3\\ 2 & 3\end{bmatrix}\begin{bmatrix}l_1\\l_2\end{bmatrix} = \begin{bmatrix}-10\\10\end{bmatrix}.$$

所以 $\begin{bmatrix}l_1\\l_2\end{bmatrix}$ 可能的值为 $\left\{\begin{bmatrix}t\\ \frac{10-2t}{3}\end{bmatrix} : 0 \le t \le 5\right\}$.

2. 设鸡兔的书目分别为 x_1, x_2, 那我们可以列方程：

$$\begin{cases} x_1 & +x_2 & = 4\\ 2x_1 & +4x_2 & = 12\\ 2x_1 & & = 8\end{cases}$$

即：$\begin{bmatrix}1 & 1\\ 2 & 4\\ 2 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}4\\12\\8\end{bmatrix}$，该方程组无解。

3. 根据题意，我们有方程组：

$$\begin{cases} a+b & = 4\\ c+d & = 8\\ a+c & = 2\end{cases}$$

写成矩阵形式就是

$$\begin{bmatrix}1 & 1 & 0 & 0\\ 0 & 0 & 1 & 1\\ 1 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}a\\b\\c\\d\end{bmatrix} = \begin{bmatrix}4\\8\\2\end{bmatrix}$$

解方程组得 $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$ 的解集为 $\left\{ \begin{bmatrix} -6 \\ 10 \\ 8 \\ 0 \end{bmatrix} + k \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} : k \in \mathbb{R} \right\}$, 因此所有可能的 A 为

$$\begin{bmatrix} -6+k & 10-k \\ 8-k & k \end{bmatrix}, \quad k \in \mathbb{R}$$

4. 写成矩阵形式即 $\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$. 解方程组得 $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{13}{3} \end{bmatrix}$. 因此交点坐标为 $\left(\frac{2}{3}, \frac{13}{3}\right)$.

5. 设平面 P_1 经过点 $\begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix}$，并且以 $\begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$ 为法向量，那么 P_1 的方程为 $x+3y+5z=d$; P_1 经过点 $\begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix}$ 可知 $d=4$, 因此 P_1 方程为 $x+3y+5z=4$. 同理可得其它两个平面方程分别为：$x+2y-3z=5$ 和 $2x+5y+2z=8$. 三个平面的交点满足的方程组为

$$\begin{bmatrix} 1 & 3 & 5 \\ 1 & 2 & -3 \\ 2 & 5 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 8 \end{bmatrix}$$

该方程组无解，因此三个平面并不交于一点。

6. 设直线上的点为 $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$，根据题意，则有 $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$. 因此写成矩阵形式为

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

解方程组可得直线上的所有点为 $\left\{ k \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} : k \in \mathbb{R} \right\}$.

7. 向量 $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ 与该平面垂直，当且仅当它和向量 $\begin{bmatrix} 1-0 \\ 1-1 \\ 0-1 \end{bmatrix}$, $\begin{bmatrix} 1-1 \\ 1-0 \\ 0-1 \end{bmatrix}$ 内积为 0。和上一小题类似，我们可以写出矩阵形式的方程

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

解方程组可得所以这样的向量为 $\left\{k\begin{bmatrix}1\\1\\1\end{bmatrix}:k\in\mathbb{R}\right\}$.

□

练习 1.3.20. 如果线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 和 $C\boldsymbol{x}=\boldsymbol{b}$,对任意 $\boldsymbol{b}$ 都有相同的解集,那么 $A=C$ 成立吗?

证明: $A=C$ 成立。首先，A,C 有相同的尺寸：它们的行数都等于 $\boldsymbol{b}$ 的行数，它们的列数都等于未知数的个数。

我们取 $\boldsymbol{b}$ 是 A 的第一列，那么 $\boldsymbol{x}=\boldsymbol{e}_1=\begin{bmatrix}1\\0\\\vdots\end{bmatrix}$ 是方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 的解，由题意，$\boldsymbol{e}_1=\begin{bmatrix}1\\0\\\vdots\end{bmatrix}$ 也是方程组 $C\boldsymbol{x}=\boldsymbol{b}$ 的解。我们注意到 $C\boldsymbol{e}_1$ 是 C 的第一列，因此这就证明了 A 和 C 有相同的第一列。同理，我们再取 $\boldsymbol{b}$ 是 A 的第二、三 $\cdots$ 列，我们可以证明 A 的每一列和 C 的相应的列都相同。因此 $A=C$. □

练习 1.3.21. 设线性映射 $\mathbf{F}\colon\mathbb{R}^{\mathbf{n}}\to\mathbb{R}^{\mathbf{m}}$. 当 $n>m$ 时，$\mathbf{F}$ 是否可能是单射? 当 $n<m$ 时，$\mathbf{F}$ 是否可能是满射?

证明: 令 A 是 $\mathbf{F}$ 对应的矩阵。

1. $\mathbf{F}$ 不能是单射。我们考虑 $\mathbf{0}$ 的原像，即齐次方程组 $A\boldsymbol{x}=\mathbf{0}$ 的解。因为 $n>m$, 那么 A 的主元至多 m 个，比未知数的个数 n 少，因此该齐次方程组有无穷多解，即 $\mathbf{0}$ 的原像有无穷多。因此 $\boldsymbol{F}$ 一定不是单射。
2. $\mathbf{F}$ 不可能是满射。我们只需要证明，一定存在 $\boldsymbol{b}$ 使得 $A\boldsymbol{x}=\boldsymbol{b}$ 无解；那么等价于证明，存在 $\boldsymbol{b}'$ 使得 $\mathrm{rref}(A)\boldsymbol{x}=\boldsymbol{b}'$ 无解。事实上，由于 $n<m$, 那么 $\mathrm{rref}(A)$ 的最后一行一定是零行所以我们取 $\boldsymbol{b}'$ 的最后一个分量等于 1 就能使得 $\mathrm{rref}(A)\boldsymbol{x}=\boldsymbol{b}'$ 无解。

□

1.4 线性映射的运算

练习 1.4.1. 设 A,B,C 分别是 $3\times5,5\times3,3\times1$ 矩阵，则 $BA,AB,BC^{\mathrm{T}},A(B+C)$ 中哪些定义良好?

证明: AB,BA 是良定义的，BC^{T} 不是良定义的，因为 C^{T} 的行数不等于 B 的列数；$A(B+C)$ 也不是良定义的，因为 B,C 行列数并不相同，不能相加。 □

练习 1.4.2. 设 $A=\begin{bmatrix}1&1&-1\\2&0&1\\1&-1&0\end{bmatrix}$, $B=\begin{bmatrix}1&0&1\\2&-1&0\\0&2&-2\end{bmatrix}$，求 $AB,BA,AB-BA$.

这个题仅仅是按照定义计算即可。

证明:

$$AB=\begin{bmatrix}3&-3&3\\2&2&0\\-1&1&1\end{bmatrix},BA=\begin{bmatrix}2&0&-1\\0&2&-3\\2&2&2\end{bmatrix},AB-BA=\begin{bmatrix}1&-3&4\\2&0&3\\-3&-1&-1\end{bmatrix}$$

□

练习 1.4.3. 计算:

1. $\begin{bmatrix}2&3\\5&7\end{bmatrix}\begin{bmatrix}2&0\\0&3\end{bmatrix}\begin{bmatrix}-7&3\\5&-2\end{bmatrix}$. *2.* $\begin{bmatrix}-7&3\\5&-2\end{bmatrix}\begin{bmatrix}2&3\\5&7\end{bmatrix}$. *3.* $\begin{bmatrix}17&-6\\35&-12\end{bmatrix}^5$.

4. $\begin{bmatrix}1&3&1\\2&2&1\\3&4&2\end{bmatrix}\begin{bmatrix}1&&\\&2&\\&&1\end{bmatrix}\begin{bmatrix}0&2&-1\\1&1&-1\\-2&-5&4\end{bmatrix}$. *6.* $\begin{bmatrix}8&6&-6\\4&6&-4\\8&8&-6\end{bmatrix}^6$.

证明: 1. $\begin{bmatrix}2&3\\5&7\end{bmatrix}\begin{bmatrix}2&0\\0&3\end{bmatrix}\begin{bmatrix}-7&3\\5&-2\end{bmatrix}=\begin{bmatrix}17&-6\\35&-12\end{bmatrix}$.

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

3.

$$\begin{aligned}\begin{bmatrix}17&-6\\35&-12\end{bmatrix}^5&=\left(\begin{bmatrix}2&3\\5&7\end{bmatrix}\begin{bmatrix}2&0\\0&3\end{bmatrix}\begin{bmatrix}-7&3\\5&-2\end{bmatrix}\right)^5\\&=\begin{bmatrix}2&3\\5&7\end{bmatrix}\begin{bmatrix}2&0\\0&3\end{bmatrix}^5\begin{bmatrix}-7&3\\5&-2\end{bmatrix}\\&=\begin{bmatrix}2&3\\5&7\end{bmatrix}\begin{bmatrix}2^5&0\\0&3^5\end{bmatrix}\begin{bmatrix}-7&3\\5&-2\end{bmatrix}\\&=\begin{bmatrix}3197&-1266\\7385&-2922\end{bmatrix}\end{aligned}$$

.

4. $\begin{bmatrix}1&3&1\\2&2&1\\3&4&2\end{bmatrix}\begin{bmatrix}1&&\\&2&\\&&1\end{bmatrix}\begin{bmatrix}0&2&-1\\1&1&-1\\-2&-5&4\end{bmatrix}=\begin{bmatrix}4&3&-3\\2&3&-2\\4&-2&-3\end{bmatrix}$.

5. 注意到 $\begin{bmatrix}0&2&-1\\1&1&-1\\-2&-5&4\end{bmatrix}\begin{bmatrix}1&3&1\\2&2&1\\3&4&2\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$, 那么

$$\begin{aligned}\begin{bmatrix}8&6&-6\\4&6&-4\\8&8&-6\end{bmatrix}^6&=2^6\begin{bmatrix}1&3&1\\2&2&1\\3&4&2\end{bmatrix}\begin{bmatrix}1&&\\&2&\\&&1\end{bmatrix}^6\begin{bmatrix}0&2&-1\\1&1&-1\\-2&-5&4\end{bmatrix}\\&=2^6\begin{bmatrix}1&3&1\\2&2&1\\3&4&2\end{bmatrix}\begin{bmatrix}1&&\\&2^6&\\&&1\end{bmatrix}\begin{bmatrix}0&2&-1\\1&1&-1\\-2&-5&4\end{bmatrix}\end{aligned}$$

$\square$

练习 1.4.4. *考虑下列 $\mathbb{R}^2$ 上的线性变换.*

1. 设 $\mathbf{R}_\theta(\boldsymbol{x}) = R_\theta \boldsymbol{x}$，其中 $R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. 求证：

(a) $\mathbf{R}_\theta$ 是绕原点逆时针方向旋转角度 θ 的变换.

(b) 分析当 θ 取何值时，存在常数 λ 和非零向量 $\boldsymbol{x}$，满足 $R_\theta \boldsymbol{x} = \lambda \boldsymbol{x}$.

(c) 计算 $R_\theta^n, n > 0, n \in \mathbb{N}$.

2. 设 $\mathbf{H}_\theta(\boldsymbol{x}) = H_\theta \boldsymbol{x}$，其中 $H_\theta = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$，而 $\boldsymbol{v} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \boldsymbol{w} = \begin{bmatrix} \sin\theta \\ -\cos\theta \end{bmatrix}$ 是 $\mathbb{R}^2$ 中的向量. 求证：

(a) $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w} = 0, \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} = 1; H_\theta \boldsymbol{v} = \boldsymbol{v}, H_\theta \boldsymbol{w} = -\boldsymbol{w}$. 试分析变换 $\mathbf{H}_\theta$ 的几何意义.

(b) $H_\theta^2 = I_2, H_\theta = I_2 - 2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}$.

(c) $R_{-\theta} H_\phi R_\theta = H_{\phi-\theta}, H_\phi R_\theta H_\phi = R_{-\theta}$，并分析其几何意义.

3. 设 $\mathbf{S}(\boldsymbol{x}) = S\boldsymbol{x}$，其中 $S = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. 设 λ 是常数，求证：$S\boldsymbol{x} = \lambda \boldsymbol{x}$ 有非零解当且仅当 $\lambda = 1$，并求出所有的非零解；计算 $S^n, n > 0, n \in \mathbb{N}$.

证明: 1. (a) 由例 1.1.10 直接可得。

(b) $R_\theta \boldsymbol{x} = \lambda \boldsymbol{x}$ 有非零解当且仅当 $(R_\theta - \lambda I_2)\boldsymbol{x}$ 有非零解，由题 1.3.4 知当且仅当 $(\cos\theta - \lambda)^2 + \sin^2\theta = 0$. 整理一下即 $\lambda^2 - 2\cos\theta\lambda + 1 = 0$. 计算判别式知，当且仅当 $\cos\theta = \pm 1$, 即 $\theta = k\pi (k \in \mathbb{Z})$.

(c) 按照题意，R_θ^n 对应的线性变换是绕原点逆时针方向作了 n 次旋转角度 θ 的变换，所以一共旋转了角度 $n\theta$, 因此 $R_\theta^n = \begin{bmatrix} \cos(n\theta) & -\sin(n\theta) \\ \sin(n\theta) & \cos(n\theta) \end{bmatrix}$

2. (a) $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w} = \cos\theta\sin\theta + \sin\theta(-\cos\theta) = 0, \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} = \cos^2\theta + \sin^2\theta = 1, \boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} = \sin^2\theta + (-\cos\theta)^2 = 1$;

$$H_\theta \boldsymbol{v} = \begin{bmatrix} \cos(2\theta)\cos\theta + \sin(2\theta)\sin\theta \\ \sin(2\theta)\cos\theta - \cos(2\theta)\sin\theta \end{bmatrix} = \begin{bmatrix} \cos(2\theta - \theta) \\ \sin(2\theta - \theta) \end{bmatrix} = \boldsymbol{v},$$

$$H_\theta \boldsymbol{w} = \begin{bmatrix} \cos(2\theta)\sin\theta - \sin(2\theta)\cos\theta \\ \sin(2\theta)\sin\theta + \cos(2\theta)\cos\theta \end{bmatrix} = \begin{bmatrix} -\sin(2\theta - \theta) \\ \cos(2\theta - \theta) \end{bmatrix} = -\boldsymbol{w}.$$

$\mathbf{H}_\theta$ 的几何意义即：将 $\boldsymbol{v}$ 方向的向量保持不变，$\boldsymbol{w}$ 方向的向量变成相反方向的向量；或者等价地说，它是以 $\boldsymbol{v}$ 方向为对称轴的反射。

(b) 我们可以从几何意义上得到这两个等式。也可以用矩阵乘法来验算，我们略去详细解答，仅仅演算一下 $I_2 - 2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}$：

$$I_2 - 2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 2\begin{bmatrix} \sin^2\theta & -\sin\theta\cos\theta \\ -\sin\theta\cos\theta & \cos^2\theta \end{bmatrix} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}.$$

(c) 这两个等式的验算我们略去。第一个等式的几何意义是，先旋转一个角度，再反射，然后再反向旋转，等效于将反射的对称轴旋转这个角度；第二个等式的意义是，先反射再旋转一个角度，再以同一个对称轴做反射，则相当于旋转相反方向的同一个角度。

3. $S\boldsymbol{x}=\lambda\boldsymbol{x}$ 有非零解，当且仅当 $(S-\lambda I_2)\boldsymbol{x}=0$ 有非零解。我们对系数矩阵作初等行变换：

$$\begin{bmatrix}1-\lambda & 0\\ 1 & 1-\lambda\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 1-\lambda\\ 0 & 1-\lambda\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 1-\lambda\\ 0 & -(1-\lambda)^2\end{bmatrix}$$

因此 $S\boldsymbol{x}=\lambda\boldsymbol{x}$ 有非零解当且仅当 $\lambda=1$. 要计算 S^n, 我们可以通过找规律猜测 $S^n=\begin{bmatrix}1 & 0\\ n & 1\end{bmatrix}$, 然后通过数学归纳法严格地证明这个等式。我们也有下面的方法：注意到 $S=I_2+J_2$ 以及 $J_2^2=0$, 其中 $J_2=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}$, 由于 I_2,J_2 乘法可交换，我们可以利用二项式展开得到

$$S^n=(I_2+J_2)^n=I_2^n+nI_2^{n-1}J_2=I_2+\begin{bmatrix}0 & 0\\ n & 0\end{bmatrix}=\begin{bmatrix}1 & 0\\ n & 1\end{bmatrix}.$$

□

练习 1.4.5. 计算：

1. $\begin{bmatrix}1 & -1 & -1 & -1\\ -1 & 1 & -1 & -1\\ -1 & -1 & 1 & -1\\ -1 & -1 & -1 & 1\end{bmatrix}^2.$ *2.* $\begin{bmatrix}1 & -1 & -1 & -1\\ -1 & 1 & -1 & -1\\ -1 & -1 & 1 & -1\\ -1 & -1 & -1 & 1\end{bmatrix}^k.$

这样算一个矩阵的高次幂，不要害怕，要有耐心，除了后面章节要学习的对角化的方法以外，所有其他的高次幂的计算本质上都是找规律，否则几乎没有任何手算的可能的。

证明*:*

$$\begin{bmatrix}1 & -1 & -1 & -1\\ -1 & 1 & -1 & -1\\ -1 & -1 & 1 & -1\\ -1 & -1 & -1 & 1\end{bmatrix}=\begin{bmatrix}2 & & & \\ & 2 & & \\ & & 2 & \\ & & & 2\end{bmatrix}+\begin{bmatrix}-1 & -1 & -1 & -1\\ -1 & -1 & -1 & -1\\ -1 & -1 & 1 & -1\\ -1 & -1 & -1 & -1\end{bmatrix}=2I_2+\begin{bmatrix}1\\ 1\\ 1\\ 1\end{bmatrix}\begin{bmatrix}-1 & -1 & -1 & -1\end{bmatrix}$$

我们令 $\boldsymbol{w}=\begin{bmatrix}1/2\\ 1/2\\ 1/2\\ 1/2\end{bmatrix}$, 则

$$\begin{bmatrix}1 & -1 & -1 & -1\\ -1 & 1 & -1 & -1\\ -1 & -1 & 1 & -1\\ -1 & -1 & -1 & 1\end{bmatrix}=2\left(I_4-2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}\right)$$

可见原矩阵本质上一个反射。因此

$$\begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}^k = 2^k \left(I_2 - 2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}\right)^k = \begin{cases} 2^k \left(I_4 - 2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}\right) & \text{若 } k \text{ 是奇数} \\ 2^k I_4 & \text{若 } k \text{ 是偶数} \end{cases}$$

□

练习 1.4.6. 设 E_{ij} 是 (i,j) 元为 1，其余元素为 0 的矩阵，而 n 阶方阵 $J_n = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}$.

1. 计算 $A = I_2 + J_2, B = I_2 + J_2^{\mathrm{T}}, C = J_2 - J_2^{\mathrm{T}}$.

2. 先计算 AB，再计算 $(AB)C$，然后先计算 BC，再计算 $A(BC)$.

3. 先计算 AB, AC，再计算 $AB + AC$，然后先计算 $B + C$，再计算 $A(B + C)$.

4. 计算 $AB + BC, B(A + C), (A + C)B$，三者是否相等？

5. 计算 $(A + B)^2, A^2 + 2AB + B^2, A^2 + AB + BA + B^2$，三者是否相等？

6. 计算 $(AB)^2$ 与 A^2B^2，二者是否相等？

7. 计算 $E_{13} - \boldsymbol{e}_1\boldsymbol{e}_3^{\mathrm{T}}$，$\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}} + \cdots + \boldsymbol{e}_4\boldsymbol{e}_4^{\mathrm{T}}$，$E_{13}E_{32}$，$E_{32}E_{13}$，其中 $\boldsymbol{e}_i \in \mathbb{R}^4$，$E_{ij}$ 是 4 阶方阵.

8. 计算 $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$.

9. 计算 $J_4 \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$，$J_4^{\mathrm{T}} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$，$J_4^2 \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$，$J_4^4$.

10. 计算 J_n^k，其中 k 是正整数.

11. 设 $P = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{bmatrix}$，计算 $PJ_4 + J_4P, (P + I_4)(2P + 3I_4) - (2P + 3I_4)(P + I_4)$.

12. 设 $R = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, S = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}, T = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{bmatrix}$，计算 $SRT, SR^{\mathrm{T}}T$.

证明: 1. $J_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, 因此 $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, C = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

2.

$$AB = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, (AB)C = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ -1 & 1 \end{bmatrix}$$

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

3.

$$AC=\begin{bmatrix}1&1\\0&1\end{bmatrix}\begin{bmatrix}0&1\\-1&0\end{bmatrix}=\begin{bmatrix}-1&1\\-1&0\end{bmatrix}, AB+AC=\begin{bmatrix}2&1\\1&1\end{bmatrix}+\begin{bmatrix}-1&1\\-1&0\end{bmatrix}=\begin{bmatrix}1&2\\0&1\end{bmatrix}$$

$$B+C=\begin{bmatrix}1&0\\1&1\end{bmatrix}+\begin{bmatrix}0&1\\-1&0\end{bmatrix}=\begin{bmatrix}1&1\\0&1\end{bmatrix}\ A(B+C)=\begin{bmatrix}1&1\\0&1\end{bmatrix}\begin{bmatrix}1&1\\0&1\end{bmatrix}=\begin{bmatrix}1&2\\0&1\end{bmatrix}$$

4.

$$AB+BC=\begin{bmatrix}2&1\\1&1\end{bmatrix}+\begin{bmatrix}0&1\\-1&1\end{bmatrix}=\begin{bmatrix}2&2\\0&2\end{bmatrix}$$

$$B(A+C)=\begin{bmatrix}1&0\\1&1\end{bmatrix}\begin{bmatrix}1&2\\-1&1\end{bmatrix}=\begin{bmatrix}1&2\\0&3\end{bmatrix}, (A+C)B=\begin{bmatrix}1&2\\-1&1\end{bmatrix}\begin{bmatrix}1&0\\1&1\end{bmatrix}=\begin{bmatrix}3&2\\0&1\end{bmatrix}$$

因此他们不相等。

5.

$$A+B=\begin{bmatrix}2&1\\1&2\end{bmatrix}, AB=\begin{bmatrix}2&1\\1&1\end{bmatrix}, BA=\begin{bmatrix}1&1\\1&2\end{bmatrix}, (A+B)^2=\begin{bmatrix}5&4\\4&5\end{bmatrix}$$

$$A^2+2AB+B^2=\begin{bmatrix}1&2\\0&1\end{bmatrix}+2\begin{bmatrix}2&1\\1&1\end{bmatrix}+\begin{bmatrix}1&0\\2&1\end{bmatrix}=\begin{bmatrix}6&4\\4&4\end{bmatrix}$$

$$A^2+AB+BA+B^2=\begin{bmatrix}1&2\\0&1\end{bmatrix}+\begin{bmatrix}2&1\\1&1\end{bmatrix}+\begin{bmatrix}1&1\\1&2\end{bmatrix}+\begin{bmatrix}1&0\\2&1\end{bmatrix}=\begin{bmatrix}5&4\\4&5\end{bmatrix}$$

因此 $(A+B)^2=A^2+AB+BA+B^2\neq A^2+2AB+B^2$.

6.

$$(AB)^2=\begin{bmatrix}2&1\\1&1\end{bmatrix}\begin{bmatrix}2&1\\1&1\end{bmatrix}=\begin{bmatrix}5&3\\3&2\end{bmatrix}$$

$$A^2B^2=\begin{bmatrix}1&2\\0&1\end{bmatrix}\begin{bmatrix}1&0\\2&1\end{bmatrix}=\begin{bmatrix}5&2\\2&1\end{bmatrix}$$

因此 $(AB)^2\neq A^2B^2$.

7. $\boldsymbol{e}_1\boldsymbol{e}_3^{\mathrm{T}}=\begin{bmatrix}1\\0\\0\end{bmatrix}\begin{bmatrix}0&0&1\end{bmatrix}=\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix}=E_{13}$，故 $E_{13}-\boldsymbol{e}_1\boldsymbol{e}_3^{\mathrm{T}}=0$. 一般的，我们可以验证 $\boldsymbol{e}_i\boldsymbol{e}_j^{\mathrm{T}}=E_{ij}$, 因此 $\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}+\cdots+\boldsymbol{e}_4\boldsymbol{e}_4^{\mathrm{T}}=I_4$. 于是

$$\begin{aligned}E_{13}E_{32}&=\boldsymbol{e}_1\boldsymbol{e}_3^T\boldsymbol{e}_3\boldsymbol{e}_2^T=\boldsymbol{e}_1\boldsymbol{e}_2^T=E_{12}\\E_{32}E_{13}&=\boldsymbol{e}_3\boldsymbol{e}_2^T\boldsymbol{e}_1\boldsymbol{e}_3^T=\boldsymbol{e}_30\boldsymbol{e}_3^T=\mathbf{0}\end{aligned}$$

8.

$$\begin{bmatrix}0&0&1\end{bmatrix}\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}=\begin{bmatrix}5&6\end{bmatrix}.$$

9.

$$
\begin{aligned}
J_4\begin{bmatrix}1\\2\\3\\4\end{bmatrix} &= \begin{bmatrix}2\\3\\4\\0\end{bmatrix}\\
J_4^T\begin{bmatrix}1\\2\\3\\4\end{bmatrix} &= \begin{bmatrix}0\\1\\2\\3\end{bmatrix}\\
J_4^2\begin{bmatrix}1\\2\\3\\4\end{bmatrix} &= \begin{bmatrix}3\\4\\0\\0\end{bmatrix}\\
J_4^4 &= \mathbf{0}
\end{aligned}.
$$

10. 计算 J_n^k 就是找规律，

$$
J_n^k = \begin{bmatrix}0 & \cdots & 0 & 1 & \cdots & 0\\ \vdots & \cdots & & & \ddots & \vdots\\ 0 & \cdots & & & 0 & 1\\ 0 & \cdots & & & 0 & 0\\ \vdots & & & & & \vdots\end{bmatrix},
$$

这个矩阵从 $(1,k+1)$ 分量到 $(n-k,n)$ 分量这条斜线上是 1，其他位置全是 0.

11.

$$
PJ_4+J_4P=\begin{bmatrix}0&1&1&1\\0&1&2&3\\0&1&3&6\\0&1&4&10\end{bmatrix}+\begin{bmatrix}1&2&3&4\\1&3&6&10\\1&4&10&20\\0&0&0&0\end{bmatrix}=\begin{bmatrix}1&3&4&5\\1&4&8&13\\1&5&13&26\\0&1&4&10\end{bmatrix}
$$

$P+I_4$ 和 $(2P+3I_4)$ 可交换，因此 $(P+I_4)(2P+3I_4)-(2P+3I_4)(P+I_4)=0$.

12.

$$
\begin{aligned}
SRT &= \begin{bmatrix}3&0\\1&1\end{bmatrix}\begin{bmatrix}\frac{1}{3}&0\\0&1\end{bmatrix} = \begin{bmatrix}1&0\\\frac{1}{3}&1\end{bmatrix}\\
SR^{\mathrm{T}}T &= \begin{bmatrix}3&3\\0&1\end{bmatrix}\begin{bmatrix}\frac{1}{3}&0\\0&1\end{bmatrix} = \begin{bmatrix}1&3\\0&1\end{bmatrix}.
\end{aligned}
$$

□

练习 1.4.7. 设二阶矩阵 B 的 $(1,2)$ 元增加 1，对下列矩阵讨论该矩阵的哪行哪列一定不变，并举例说明所有其他行列确实可以变化：

1. $A+B$. *2.* AB. *3.* BA. *4.* B^2.

证明*:* 1. $A+B$ 只有第一行和第二列改变，其他的行、列不改变。

2. AB 的列除了第二列,其他列一定不变;但是所有的行都有可能变化。例如 $A=\begin{bmatrix}1&1\\1&1\end{bmatrix}, B=\begin{bmatrix}0&0\\0&0\end{bmatrix}, B'=\begin{bmatrix}0&1\\0&0\end{bmatrix}, AB=\begin{bmatrix}0&0\\0&0\end{bmatrix}, AB'=\begin{bmatrix}0&1\\0&1\end{bmatrix}$

3. BA 的列除了第一行，其他行一定不变；但是所有的列都有可能变化。例子如上题。

4. 所有行、列都可能变化。例如 $B=\begin{bmatrix}1&1\\1&0\end{bmatrix}, B'=\begin{bmatrix}1&2\\1&0\end{bmatrix}; B^2=\begin{bmatrix}2&1\\1&1\end{bmatrix}, B'^2=\begin{bmatrix}3&2\\1&2\end{bmatrix}$.

□

练习 1.4.8. 判断对错：

1. 如果 B 的第一列等于第三列，则 AB 的第一列等于第三列.

2. 如果 B 的第一列等于第三列，则 BA 的第一列等于第三列.

3. 如果 A 的第一行等于第三行，则 ABC 的第一行等于第三行.

证明*:* 1. 正确

2. 错误

3. 正确

□

练习 1.4.9. 求所有满足条件的矩阵 B：

1. 对任意三阶方阵 A，$BA=4A$.

2. 对任意三阶方阵 A，$BA=4B$.

3. 对任意三阶方阵 A，BA 的每一行都是 A 的第一行.

4. 对任意三阶方阵 A，AB 的每一行的每一个元素都是 A 的对应行的平均值.

5. $B^2\neq 0, B^3=0$，且 B 是三阶上三角矩阵.

6. $B\begin{bmatrix}1&1\\1&1\end{bmatrix}=\begin{bmatrix}1&1\\1&1\end{bmatrix}B$.

证明*:* 1. 取 $A=I_3$ 可得 $B=4I_3$，可以验证满足条件。

2. 取 $A=I_3$ 可得 $B=4I_3$，可以验证满足条件。

3. 取 $A=I_3$ 可得 $B=\begin{bmatrix}1&0&0\\1&0&0\\1&0&0\end{bmatrix}$, 可以验证满足条件。

4. 取 $A=I_3$ 可得 $B=\begin{bmatrix}\frac13&\frac13&\frac13\\\frac13&\frac13&\frac13\\\frac13&\frac13&\frac13\end{bmatrix}$，现在我们验证一下它是否满足条件。设 $A=$

$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, 直接计算可知

$$AB = \begin{bmatrix} \frac{a_{11}+a_{12}+a_{13}}{3} & \frac{a_{11}+a_{12}+a_{13}}{3} & \frac{a_{11}+a_{12}+a_{13}}{3} \\ \frac{a_{21}+a_{22}+a_{23}}{3} & \frac{a_{21}+a_{22}+a_{23}}{3} & \frac{a_{21}+a_{22}+a_{23}}{3} \\ \frac{a_{21}+a_{22}+a_{23}}{3} & \frac{a_{31}+a_{32}+a_{33}}{3} & \frac{a_{31}+a_{32}+a_{33}}{3} \end{bmatrix}$$

所以满足条件。

5. 设 $B = \begin{bmatrix} x_1 & x_2 & x_3 \\ 0 & x_4 & x_5 \\ 0 & 0 & x_6 \end{bmatrix}$. B^3 对角线元素非别为 x_1^3, x_4^4, x_6^3, 从而 $x_1 = x_4 = x_6 = 0$. 于是 $B^2 = \begin{bmatrix} 0 & 0 & x_2x_5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $B^3 = 0$. 因此满足条件的 B 为

$$\left\{ \begin{bmatrix} 0 & x & y \\ 0 & 0 & z \\ 0 & 0 & 0 \end{bmatrix} : x \neq 0, z \neq 0, y \in \mathbb{R} \right\}.$$

6. 设 $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$，则条件转化为 $\begin{bmatrix} a+b & a+b \\ c+d & c+d \end{bmatrix} = \begin{bmatrix} a+c & b+d \\ a+c & b+d \end{bmatrix}$, 等价于 $b = c, a = d$, 因此满足条件的 B 为

$$\left\{ \begin{bmatrix} a & b \\ b & a \end{bmatrix} : a, b \in \mathbb{R} \right\}.$$

□

练习 1.4.10. 给定矩阵 $A = \begin{bmatrix} p & 0 \\ q & r \end{bmatrix}, B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 0 & z \\ 0 & 0 \end{bmatrix}$.

1. p, q, r 取何值时，有 $AB = BA$?
2. z 取何值时，有 $BC = CB$?
3. p, q, r, z 取何值时，有 $ABC = CAB$?

证明: 1. $AB = \begin{bmatrix} p & p \\ q & q+r \end{bmatrix}, BA = \begin{bmatrix} p+q & r \\ q & r \end{bmatrix}$. $AB = BA$ 当且仅当 $q = 0$ 且 $p = r$.

2. 我们直接计算就有 $BC = CB = C$.

3. $ABC = AC = \begin{bmatrix} 0 & pz \\ 0 & qz \end{bmatrix}, CAB = \begin{bmatrix} zq & z(q+r) \\ 0 & 0 \end{bmatrix}$, 因此 $ABC = CAB$ 当且仅当 $qz = 0, pz = z(q+r)$, 当且仅当 $\underline{z = 0}$ 或者 $\underline{q = 0 \text{ 且 } p = r}$.

□

练习 1.4.11. 证明:

1. 设 n 维向量 $\boldsymbol{x}$ 的每个分量都是 1，则 n 阶方阵 A 的各行元素之和为 1 当且仅当 $A\boldsymbol{x} = \boldsymbol{x}$.

2. 若 n 阶方阵 A, B 的各行元素之和均为 1，则 AB 的各行元素之和也均为 1.

3. 若 n 阶方阵 A, B 的各列元素之和均为 1，则 AB 的各列元素之和也均为 1.

证明: 设 $A=[a_{ij}], B=[b_{ij}]$.

1. 我们只需要注意到, 由于 $\boldsymbol{x}$ 的每个分量都是 1, $A\boldsymbol{x}$ 的第 $i(1 \leq i \leq n)$ 个分量就是 A 的第 i 行的和。

2. 由上一小题，$AB\boldsymbol{x}=A(B\boldsymbol{x})=A\boldsymbol{x}=\boldsymbol{x}$, 则 AB 的各行元素之和也均为 1.

3. A, B 的各列元素之和均为 1，当且仅当 $A^{\mathrm{T}}, B^{\mathrm{T}}$ 的各行元素之和均为 1，那么 $B^{\mathrm{T}} A^{\mathrm{T}}=(AB)^{\mathrm{T}}$ 的各行元素之和均为 1, 即 AB 的各列元素之和也均为 1.

□

练习 1.4.12. 证明命题 *1.4.17.*

证明: 我们只证明第一小题,第二小题请自行验证。我们假设 $A=[a_{ij}], B=[b_{ij}] \in M_{m \times n}(\mathbb{R}), C=[c_{kl}] \in M_{n \times r}(\mathbb{R})$, 此时 $(A+B)C, AC, BC$ 均为 $m \times r$ 矩阵；$A+B=[a_{ij}+b_{ij}] \in M_{m \times n}(\mathbb{R})$. 我们验证 $(A+B)C$ 的每一个分量和 $AC+BC$ 的每一个分量相等。

$$
\begin{aligned}
((A+B)C)_{ij} & =\sum_{k=1}^{n}(A+B)_{ik} C_{kj}=\sum_{k=1}^{n}(a_{ik}+b_{ik}) c_{kj} \\
& =\sum_{k=1}^{n} a_{ik} c_{kj}+\sum_{k=1}^{n} b_{ik} c_{kj}=(AC)_{ij}+(BC)_{ij}.
\end{aligned}
$$

□

练习 1.4.13. 证明上三角矩阵对加法、数乘、乘法封闭，即：设 U_1, U_2 是 n 阶上三角矩阵，k 是实数，则 U_1+U_2, kU_1, U_1U_2 都是上三角矩阵. 此外，U_1U_2 的对角元素是 U_1, U_2 对应的对角元素的乘积.

证明: 设 $U_1=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}, U_2=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ & & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix}$, 此时

$$
U_1+U_2=\begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ 0 & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}+b_{nn} \end{bmatrix}, kU_1=\begin{bmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ 0 & ka_{22} & \cdots & ka_{2n} \\ & & \ddots & \vdots \\ 0 & 0 & \cdots & ka_{nn} \end{bmatrix}.
$$

我们可以把矩阵补完整 $U_1=[a_{ij}], U_2=[b_{ij}]$, 但是其中, 当 $i<j$ 时 $a_{ij}=b_{ij}=0$. 我们考虑 U_1U_2 的第 (i, j) 分量：

$$
(U_1U_2)_{ij}=\sum_{k=1}^{n} a_{ik} b_{kj}.
$$

若 $i>j$, 那么对任意的 k, $i<k$ 和 $k<j$ 中至少有一个成立，也就是说 a_{ik} 和 b_{kj} 中至少有一个为 0, 因此 $(U_1U_2)_{ij}=0$.

若 $i=j$, 那么对任意 $k\neq i$, $i<k$ 和 $k<j$ 中至少有一个成立，a_{ik} 和 b_{kj} 中至少有一个为 0, 因此整个求和式里面只有一项有贡献，即 $(U_1U_2)_{ii}=a_{ii}b_{ii}$. □

练习 1.4.14. 设 A 是 n 阶上三角矩阵. 求证：$A^n=0$ 当且仅当 A 是严格上三角矩阵.

我们认证计算过题 1.4.6 的第 10 小题，就有一些经验：指数没增加一个，从左上往右下的那些 1 的位置就往右上角提升一个，同时数量减少一个。那么经过有限次以后就全消失了。本题的关键就在于把这个论证说清楚。

证明: 若 A 不是严格上三角矩阵，那么存在 A 对角线上的某个元素 $a_{ii}\neq 0$, 那么由上一题，A^n 的第 ii 个分量就是 a_{ii}^n 也不是零；因此 $A^n\neq 0$.

现在假设 A 是严格上三角矩阵。我们用归纳法来证明：对 $1\leq k\leq n$, A^k 是如下形式的矩阵：

$$\begin{bmatrix} 0 & \cdots & 0 & * & \cdots & * \\ \vdots & \cdots & & & \ddots & \vdots \\ 0 & \cdots & & & 0 & * \\ 0 & \cdots & & & 0 & 0 \\ \vdots & & & & & \vdots \end{bmatrix},$$

这个矩阵除了从 $(1,k+1)$ 分量到 $(n-k,n)$ 分量这条斜线及其右上方的元素以外的其他位置全是 0. 我们可以严格地描述这种形式的矩阵：$B=[b_{ij}]$ 具有这样的形式，当且仅当 $b_{ij}=0,\forall j\geq i+k-1$. 下面我们开始证明。$k=1$ 时，由严格上三角矩阵的定义显然的。假设这个命题对 k 成立，我们记 $A^k=[c_{ij}]$, 那么我们考虑 $(A^{k+1})_{ij}$:

$$(A^{k+1})_{ij}=(A^kA)_{ij}=\sum_{\ell=1}^{n}c_{i\ell}a_{\ell j}$$

对于 $j\geq i+(k+1)-1=i+k$, 那么 $j\geq \ell+k-1$ 和 $\ell\geq i$ 至少有一个成立，根据归纳假设，$c_{i\ell}$、$a_{\ell j}$ 至少有一个为零。因此当 $j\geq i+(k+1)-1=i+k$ 时，$(A^{k+1})_{ij}=0$。证毕。 □

如果有分块矩阵的知识，我们也可以利用分块矩阵来对方阵的阶做归纳，归纳的关键是利用下面的等式：

$$\text{若 } B=\begin{bmatrix} A & \alpha \\ 0 & 0 \end{bmatrix}, \text{则 } B^k=\begin{bmatrix} A^k & A^{k-1}\alpha \\ 0 & 0 \end{bmatrix}$$

练习 1.4.15. 证明：如果 n 阶方阵 A 满足 $A^n=0$，则 $(I_n-A)(I_n+A+A^2+\cdots+A^{n-1})=I_n$.

证明: $A^n=0$, 则 $I_n=I-A^n=(I_n-A)(I_n+A+A^2+\cdots+A^{n-1})$. □

练习 1.4.16. 对 n 阶方阵 A，考虑集合 $\mathrm{Com}(A)=\Big\{n \text{ 阶方阵 } B:AB=BA\Big\}$.

1. 求证：$\mathrm{Com}(A)$ 是所有 n 阶方阵的集合当且仅当 A 是数量矩阵 kI_n.
2. 设 $A=\mathrm{diag}(d_i)$，d_i 互不相同，求 $\mathrm{Com}(A)$.
3. 求证：任取 $B,C\in\mathrm{Com}(A)$，都有 $I_n, kB+lC, BC\in\mathrm{Com}(A)$.

4. 设 $A=\begin{bmatrix}1&1&0\\0&1&1\\0&0&1\end{bmatrix}$，求证：$J_3=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}\in \mathrm{Com}(A)$，而且

$$\mathrm{Com}(A)=\left\{k_1I_3+k_2J_3+k_3J_3^2: k_1,k_2,k_3\in\mathbb{R}\right\}.$$

证明: 1. 若 $A=kI_n$, 容易证明 $\mathrm{Com}(A)$ 是所有 n 阶方阵的集合。反之，若 $\mathrm{Com}(A)$ 是所有 n 阶方阵的集合，我们考虑 E_{ij}, 它是 (i,j) 分量为 1，其他分量全为零的矩阵，那么 $AE_{ij}=E_{ij}A$, 这个等式左边第 j 列为 A 的第 i 列，其他列全为零; 等式右边第 i 行为 A 的第 j 行，且其他行全为零。比较等式左右两边第 (i,j) 分量，则 $a_{ii}=a_{jj}$；比较其它分量，则得到 A 的第 i 列和第 j 行除了对角线元素以外都是 0. 我们取遍所有 i,j，就得到 A 是数量矩阵。

2. 设 $B=[b_{ij}]\in\mathrm{Com}(A)$. $AB=[d_ib_{ij}], BA=[bijd_j]$, 因此对任意 i,j 有 $d_ib_{ij}=b_{ij}d_j$, 于是 $b_{ij}=0,\forall i\neq j$.

3. $I_nA=AI_n=A$, 因此 $I_n\in\mathrm{Com}(A)$. 若 $B,C\in\mathrm{Com}(A)$, 那么 $(kB+lC)A=kBA+lCA=kaB+lAC=A(kB+lC)$, $BCA=BAC=ABC$, 因此 $kB+lC, BC\in\mathrm{Com}(A)$.

4. 我们首先证明：$\mathrm{Com}(A)=\mathrm{Com}(J_3)$。这是因为 $A=I_3+J_3$, 因此 $AB=BA$ 当且仅当 $B+BJ_3=B+J_3B$ 当且仅当 $BJ_3=J_3B$. 设 $B=\begin{bmatrix}b_{11}&b_{12}&b_{13}\\b_{21}&b_{22}&b_{23}\\b_{31}&b_{32}&b_{33}\end{bmatrix}\in\mathrm{Com}(A)$，则

$$J_3B=\begin{bmatrix}b_{21}&b_{22}&b_{23}\\b_{31}&b_{32}&b_{33}\\0&0&0\end{bmatrix}, BJ_3=\begin{bmatrix}0&b_{11}&b_{12}\\0&b_{21}&b_{22}\\0&b_{31}&b_{32}\end{bmatrix}$$

$J_3B=BJ_3$ 则有 $b_{11}=b_{22}=b_{33}, b_{12}=b_{23}, b_{21}=b_{31}=b_{32}=0$，所以

$$\mathrm{Com}(A)=\left\{k_1I_3+k_2J_3+k_3J_3^2: k_1,k_2,k_3\in\mathbb{R}\right\}.$$

□

练习 1.4.17. 设 $A=\begin{bmatrix}\lambda&1&0\\0&\lambda&1\\0&0&\lambda\end{bmatrix}$，求 A^k，其中 k 是正整数.

证明: 注意到 $A=\lambda I_3+J_3$ 且 $J_3^3=0$, 则

$$A^k=(\lambda I_3+J_3)^k=(\lambda I_3)^k+k(\lambda I_3)^{k-1}J_3+\binom{k}{2}(\lambda I_3)^{k-2}J_3^2=\begin{bmatrix}\lambda^k&k\lambda^{k-1}&\frac{k(k-1)}{2}\lambda^{k-2}\\0&\lambda^k&k\lambda^{k-1}\\0&0&\lambda^k\end{bmatrix}.$$

□

练习 1.4.18. 设 A 是 n 阶方阵，求证：$A+A^{\mathrm{T}}, AA^{\mathrm{T}}, A^{\mathrm{T}}A$ 都是对称矩阵，而 $A-A^{\mathrm{T}}$ 是反对称矩阵.

证明:

$$\begin{aligned}
(A+A^{\mathrm{T}})^{\mathrm{T}} &= A^{\mathrm{T}}+(A^{\mathrm{T}})^{\mathrm{T}} &&= A+A^{\mathrm{T}}\\
(AA^{\mathrm{T}})^{\mathrm{T}} &= (A^{\mathrm{T}})^{\mathrm{T}}A^{\mathrm{T}} &&= AA^{\mathrm{T}}\\
(A^{\mathrm{T}}A)^{\mathrm{T}} &= A^{\mathrm{T}}(A^{\mathrm{T}})^{\mathrm{T}} &&= A^{\mathrm{T}}A\\
(A-A^{\mathrm{T}})^{\mathrm{T}} &= A^{\mathrm{T}}-(A^{\mathrm{T}})^{\mathrm{T}} &&= -(A-A^{\mathrm{T}})
\end{aligned}$$

□

练习 1.4.19. 求证：

1. 任意方阵 A 都可唯一地表示为 $A=B+C$，其中 B 是对称矩阵，C 是反对称矩阵.
2. n 阶方阵 A 是反对称矩阵当且仅当对任意 n 维列向量 $\boldsymbol{x}$，都有 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=0$.
3. 设 A,B 是 n 阶对称矩阵，则 $A=B$ 当且仅当对任意 n 维列向量 $\boldsymbol{x}$，都有 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}B\boldsymbol{x}$.

证明:

1. 假设有这样的表达式，那么 $A^{\mathrm{T}}=(B+C)^{\mathrm{T}}=B^{\mathrm{T}}+C^{\mathrm{T}}=B-C$, 那么就有 $B=\frac{A+A^{\mathrm{T}}}{2}, C=\frac{A-A^{\mathrm{T}}}{2}$. 这证明了唯一性。反之，我们令 $B=\frac{A+A^{\mathrm{T}}}{2}, C=\frac{A-A^{\mathrm{T}}}{2}$，由上一题我们知 B 对称，C 反对称，且 $A=B+C$. 这证明了存在性。
2. 若 A 是反对称矩阵，那么对任意 $\boldsymbol{x}$, $(\boldsymbol{x}^TA\boldsymbol{x})^{\mathrm{T}}=\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{x}=-\boldsymbol{x}^TA\boldsymbol{x}$, 所以 $\boldsymbol{x}^TA\boldsymbol{x}=0$. 反之，若对任意 $\boldsymbol{x}$, 都有 $\boldsymbol{x}^TA\boldsymbol{x}=0$, 考虑 $(\boldsymbol{x}+\boldsymbol{y})^TA(\boldsymbol{x}+\boldsymbol{y})=0$, 于是对任意 $\boldsymbol{x},\boldsymbol{y}$, 都有 $\boldsymbol{x}^TA\boldsymbol{y}=-\boldsymbol{y}^TA\boldsymbol{x}$. 取 $\boldsymbol{x}=\boldsymbol{e}_i,\boldsymbol{y}=\boldsymbol{e}_j$，就有 $a_{ij}=(\boldsymbol{e}_i)^TA\boldsymbol{e}_j=-(\boldsymbol{e}_j)^TA\boldsymbol{e}_i=-a_{ji}$. 所以 A 是反对称矩阵。
3. 有一个方向是显然的。我们只需要证明当对任意 n 维列向量 $\boldsymbol{x}$，都有 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}B\boldsymbol{x}$ 时，$A=B$. 考虑 $A-B$, 我们只需要证明若对任意 n 维列向量 $\boldsymbol{x}$，都有 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=0$, 则 $A=0$. 事实上有上一小题知 A 反对称；而由题意，A 对称。故我们同时有 $A=A^{\mathrm{T}}$ 和 $A=-A^{\mathrm{T}}$. 因此 $A=0$.

□

练习 1.4.20. 设 A,B 是同阶对称矩阵. 求证：AB 是对称矩阵当且仅当 $AB=BA$.

证明: 若 AB 对称，则 $AB=(AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}}=BA$.

若 $AB=BA$，则 $(AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}}=BA=AB$, 因此 AB 是对称矩阵。 □

练习 1.4.21. 设 A 是实对称矩阵，如果 $A^2=0$，求证：$A=0$.

证明: 设 $A=[a_{ij}]$. 因为 A 对称，所以 $a_{ij}=a_{ji}$. 考虑 A^2 的第 $i(1\le i\le n)$ 个对角线元素，它等于 $\sum_{k=1}^n a_{ik}a_{ki}=\sum_{k=1}^n a_{ik}^2$. 而 $A^2=0$, 则 $\sum_{k=1}^n a_{ik}^2=0$. 也就是说，对所有的 i,k, $a_{ik}=0$. 所以 $A=0$. □

练习 1.4.22 (矩阵的迹)**.** 方阵 A 的对角元素的和称为它的**迹**，记作 $\mathrm{trace}(A)$. 验证下列性质：

1. 对任意同阶方阵 A,B，$\mathrm{trace}(A+B)=\mathrm{trace}(A)+\mathrm{trace}(B)$.
2. 对任意方阵 A 与实数 k，$\mathrm{trace}(kA)=k\mathrm{trace}(A)$.
3. 对 m 阶单位矩阵 I_m，$\mathrm{trace}(I_m)=m$.
4. 对任意方阵 A，$\mathrm{trace}(A)=\mathrm{trace}(A^{\mathrm{T}})$.

5. 设 $A=\begin{bmatrix}a_1 & a_2\\ a_3 & a_4\end{bmatrix}, B=\begin{bmatrix}b_1 & b_2\\ b_3 & b_4\end{bmatrix}$，则 $\operatorname{trace}(A^{\mathrm{T}}B)=\sum\limits_{i=1}^{4}a_ib_i$. 如果 A 是 m 阶方阵呢？

6. 设 $\boldsymbol{v},\boldsymbol{w}$ 是 m 维向量，则 $\operatorname{trace}(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w})=\operatorname{trace}(\boldsymbol{w}\boldsymbol{v}^{\mathrm{T}})$.

7. 设 A,B 分别是 $m\times n, n\times m$ 矩阵，则 $\operatorname{trace}(AB)=\operatorname{trace}(BA)$.

8. 设 A,B 是任意 m 阶方阵，则 $AB-BA\neq I_m$.

证明*:* 1. 设 $A=[a_{ij}], B=[b_{ij}]$, 则 $\operatorname{trace}(A)=\sum_{i=1}^m a_{ii}.\operatorname{trace}(A+B)=\sum_{i=1}^m(a_{ii}+b_{ii})=\sum_{i=1}^n a_{ii}+\sum_{i=1}^m b_{ii}=\operatorname{trace}(A)+\operatorname{trace}(B)$.

2. 设 $A=[a_{ij}]$, 则 $\operatorname{trace}(A)=\sum_{i=1}^m a_{ii}.\operatorname{trace}(kA)=\sum_{i=1}^m(ka_{ii})=k\sum_{i=1}^m a_{ii}=k\operatorname{trace}(A)$.

3. $\operatorname{trace}(I_m)=\sum_{i=1}^m 1=m$.

4. 由于 A^T 的对角线元素和 A 是相同的，因此 $\operatorname{trace}(A)=\operatorname{trace}(A^{\mathrm{T}})$.

5. $A^{\mathrm{T}}B=\begin{bmatrix}a_1b_1+a_3b_3 & a_1b_2+a_3b_4\\ a_2b_1+a_4b_3 & a_2b_2+a_4b_4\end{bmatrix}$, 因此 $\operatorname{trace}(A^{\mathrm{T}}B)=\sum_{i=1}^4 a_ib_i$. 若 A,B 是 m 阶方阵，则

$$\operatorname{trace}(A^{\mathrm{T}}B)=\sum_{i=1}^m(A^{\mathrm{T}}B)_{ii}=\sum_{i=1}^m\sum_{j=1}^m(A^{\mathrm{T}})_{ij}B_{ji}=\sum_{i,j}a_{ji}b_{ji},$$

即将 A,B 相应位置的元素相乘，然后全加起来。

6. 设 $\boldsymbol{v}=\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_m\end{bmatrix}, \boldsymbol{w}=\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_m\end{bmatrix}$, 于是 $\boldsymbol{v}^T\boldsymbol{w}=\sum_{i=1}^m x_iy_i$. $\boldsymbol{w}\boldsymbol{v}^T=\begin{bmatrix}y_1x_1 & y_1x_2 & \cdots & y_1x_m\\ y_2x_1 & y_2x_2 & \cdots & y_2x_m\\ \vdots & & \vdots & \vdots\\ y_mx_1 & y_mx_2 & \cdots & y_mx_m\end{bmatrix}$, 所以 $\operatorname{trace}(\boldsymbol{w}\boldsymbol{v}^{\mathrm{T}})=\sum_{i=1}^m x_iy_i=\operatorname{trace}(\boldsymbol{v}^T\boldsymbol{w})$.

7. 设 $A=[a_{ij}]\in M_{m\times n}(\mathbb{R}), B=[b_{kl}]\in M_{n\times m}(\mathbb{R})$, 则

$$\operatorname{trace}(AB)=\sum_{i=1}^m(AB)_{ii}=\sum_{i=1}^m\sum_{j=1}^n a_{ij}b_{ji}=\sum_{j=1}^n\sum_{i=1}^n b_{ji}a_{ij}=\sum_{j=1}^m(BA)_{jj}=\operatorname{trace}(BA)$$

8. 由于 $\operatorname{trace}(AB-BA)=\operatorname{trace}(AB)-\operatorname{trace}(BA)=0\neq\operatorname{trace}(I_m)$, 所以 $AB-BA\neq I_m$.

□

练习 1.4.23 (差分矩阵与求导). 在微积分中，$f(x)$ 的导数 $f'(x)\approx\dfrac{f(x+\delta)-f(x)}{\delta}$. 对数列 $\{a_n\}$ 的相邻两项做减法，可以看作是一种离散导数 $\dfrac{a_{n+1}-a_n}{1}$. 两者具有某种共性. 令 D 为 100 阶向前差分矩阵，而 $\boldsymbol{a}=\begin{bmatrix}a_1\\ a_2\\ \vdots\\ a_{100}\end{bmatrix}$. 求证：

1. 若 a_k 是关于 k 的 3 次多项式，则除第 1 个分量外，$D\boldsymbol{a}$ 的第 k 个分量是关于 k 的 2 次多项式.

2. 若 $a_k=e^k$，则除第 1 个分量外，$\dfrac{e}{e-1}D\boldsymbol{a}$ 的第 k 个分量是 e^k.

证明*:* 差分矩阵的定义见教材 50 页.

1. 设 $a_k=f(k)=ak^3+bk^2+ck+d$, $D\boldsymbol{a}$ 的第 k 个分量为 $-a_{k-1}+a_k=3ak^2+(-3a+2b)k+(a-b+c)$, 是关于 k 的 2 次多项式。
2. 当 $k\geq 2$ 时,$D\boldsymbol{a}$ 的第 k 个分量是 $-e^{k-1}+e^k=e^{k-1}(e-1)$, 因此 $\dfrac{e}{e-1}D\boldsymbol{a}$ 的第 k 个分量是 e^k.

□

练习 1.4.24. 如图 1.2所示的电路包含 5 个顶点,电势分别为 $v_i, i=1,2,\ldots,5$, i,j 两点之间电阻为 $r_{ij}\neq 0$,从顶点 i 到 j 的电流为 c_{ij},又记 $c_{ii}=0$. 令电势矩阵为 $V=\mathrm{diag}(v_1,v_2,\ldots,v_5)$,电流矩阵为 $C=\begin{bmatrix}c_{ij}\end{bmatrix}$,电导矩阵为 $G=\begin{bmatrix}g_{ij}\end{bmatrix}$,其中 $g_{ii}=0, g_{ij}=\dfrac{1}{r_{ij}}, i\neq j$. 求证 $C=VG-GV$.

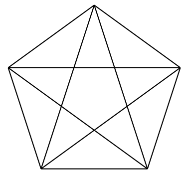


图 1.2

证明: 根据欧姆定律,$v_j-v_i=c_{ij}r_{ij}$,即 $c_{ij}=\frac{v_j-v_i}{r_{ij}}$.

$$(VG-GV)_{ij}=\sum_{k=1}^{n}(V_{ik}G_{kj}-G_{ik}V_{kj})=v_ig_{ij}-v_jg_{ij}$$

因此 C 和 $VG-GV$ 每个分量都相等。 □

练习 1.4.25. 图 1.3中的图含有四个顶点v_1, v_2, v_3, v_4，顶点之间有边连接. 对称矩阵 A 称为该图的**邻接矩阵**：如果 v_i, v_j 之间有边，则 A 的 (i,j) 元是 1；如果 v_i, v_j 之间没有边，则 A 的 (i,j) 元是 0.

从 v_3 出发，三次通过连线，最后回到 v_3 的方法有几种？

求矩阵 A^3 的 $(3,3)$ 元.

分析 A^n 的 (i,j) 元的意义.

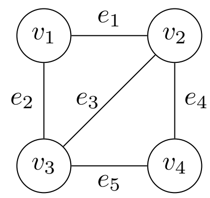


图 1.3

证明:

1. 我们直接数。从 v_3 出发，第一步有三个选择 e_2, e_3, e_5，这三个选择下三次通过连线最后回到 v_3 的方法分别有有 $1, 2, 1$ 种。因此一共有 4 种方法。
2. 按题意，我们写出 A:

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}, \text{则} A^3 = \begin{bmatrix} 2 & 5 & 5 & 2 \\ 5 & 4 & 5 & 5 \\ 5 & 5 & 4 & 5 \\ 2 & 5 & 5 & 2 \end{bmatrix}$$

因此 A^3 的 $(3,3)$ 元就是从 v_3 出发，三次通过连线，最后回到 v_3 的方法数量。

3. A^n 的 (i,j) 元就是从 v_i 出发，n 次通过连线，最后回到 v_j 的方法数量。我们可以通过归纳法来证明这一点，留给同学们思考。

□

练习 1.4.26. 设 $\boldsymbol{v}^{\mathrm{T}}, \boldsymbol{w}^{\mathrm{T}}$ 是 k 维行向量，A 是 $m \times n$ 矩阵.

1. 如果矩阵乘积 $\boldsymbol{v}^{\mathrm{T}} A$ 良定义，需要满足什么条件？

2. 对任意常数 $a, b \in \mathbb{R}$，证明 $(a\boldsymbol{v}^{\mathrm{T}} + b\boldsymbol{w}^{\mathrm{T}})A = a\boldsymbol{v}^{\mathrm{T}}A + b\boldsymbol{w}^{\mathrm{T}}A$.

3. 把 $2\boldsymbol{a}^{\mathrm{T}} + 3\boldsymbol{b}^{\mathrm{T}} + 4\boldsymbol{c}^{\mathrm{T}}$ 写成一个行向量和矩阵的乘积.

4. 求矩阵 A，使得 $\begin{bmatrix} x_1 & x_2 \end{bmatrix} A \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 2x_1y_1 + 3x_1y_2 + 4x_2y_1 + 5x_2y_2$.

5. 求矩阵 A，使得 $\begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} = x^2 + 4xy + 5y^2$. 这样的 A 是否唯一？

6. 求二阶对称矩阵 A，使得对任意非零二维向量 $\boldsymbol{v}$，都有 $\boldsymbol{v}^{\mathrm{T}}A\boldsymbol{v} > 0$.

证明:

1. 矩阵乘积良定义的条件时前一个矩阵的列数等于后一个矩阵的行数，因此 $\boldsymbol{v}^{\mathrm{T}}A$ 良定义 $k = m$.
2. 这是矩阵乘法的分配律；我们也可以用分量写出来验证。略过。
3. $2\boldsymbol{a}^{\mathrm{T}} + 3\boldsymbol{b}^{\mathrm{T}} + 4\boldsymbol{c}^{\mathrm{T}} = \begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} \boldsymbol{a}^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \\ \boldsymbol{c}^{\mathrm{T}} \end{bmatrix}$
4. 设 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$，则 $\begin{bmatrix} x_1 & x_2 \end{bmatrix} A \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = ax_1y_1 + bx_1y_2 + cx_2y_1 + dx_2y_2$. 因此 $a = 2, b =$

$3, c=4, d=5$, 即 $A=\begin{bmatrix}2 & 3\\ 4 & 5\end{bmatrix}$

5. 设 $A=\begin{bmatrix}a & b\\ c & d\end{bmatrix}$, 则 $\begin{bmatrix}x & y\end{bmatrix} A \begin{bmatrix}x\\ y\end{bmatrix} = ax^2+(b+c)xy+dy^2$. 因此 $a=1, d=5, b+c=4$. 这样的 A 不唯一。

6. 设 $A=\begin{bmatrix}a & b\\ b & d\end{bmatrix}$，则条件等价于说，对任意不全为零的 x, y, $ax^2+2bxy+dy^2>0$ 恒成立，当且仅当 $a>0, d>0, b^2-4ad<0$.

□

练习 1.4.27. 由标准坐标向量 $\boldsymbol{e}_i$ 定义 $E_{ij}=\boldsymbol{e}_i\boldsymbol{e}_j^{\mathrm{T}}$，它是 (i,j) 元为 1，其他元素为 0 的矩阵.

1. 证明，当 $j\neq k$ 时，$E_{ij}E_{kl}=0$.

2. 对任意矩阵 A，求向量 $\boldsymbol{v}$ 使得 $A\boldsymbol{v}$ 是 A 的第 i 列.

3. 对任意矩阵 A，求向量 $\boldsymbol{v}$ 使得 $\boldsymbol{v}^{\mathrm{T}}A$ 是 A 的第 i 行.

4. 对任意矩阵 A，证明，$\boldsymbol{e}_i^{\mathrm{T}}A\boldsymbol{e}_j$ 为 A 的 (i,j) 元.

5. 对 $\boldsymbol{e}_k\in\mathbb{R}^m$ 证明，$\sum_{k=1}^{m}\boldsymbol{e}_k\boldsymbol{e}_k^{\mathrm{T}}=I_m$.

6. （阅读）计算矩阵乘积 AB 的 (i,j) 元的另一种方法：设 A 有 m 列，B 有 m 行，则

$$\boldsymbol{e}_i^{\mathrm{T}}AB\boldsymbol{e}_j=\boldsymbol{e}_i^{\mathrm{T}}AI_mB\boldsymbol{e}_j=\boldsymbol{e}_i^{\mathrm{T}}A\left(\sum_{k=1}^{m}\boldsymbol{e}_k\boldsymbol{e}_k^{\mathrm{T}}\right)B\boldsymbol{e}_j=\sum_{k=1}^{m}(\boldsymbol{e}_i^{\mathrm{T}}A\boldsymbol{e}_k)(\boldsymbol{e}_k^{\mathrm{T}}B\boldsymbol{e}_j).$$

证明*:*

1. 当 $j\neq k$ 时，$\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{e}_k=0$, 因此 $E_{ij}E_{kl}=\boldsymbol{e}_i\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{e}_k\boldsymbol{e}_l^{\mathrm{T}}=0$.

2. 设 $\boldsymbol{v}=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$; 我们令 $A=E_{jk}$，则 $E_{jk}\boldsymbol{v}=(\boldsymbol{e}_k^T\boldsymbol{v})\boldsymbol{e}_j=x_k\boldsymbol{e}_j$. 根据条件，$E_{jk}\boldsymbol{v}$ 是 E_{jk} 的第 i 列，因此

$$E_{jk}\boldsymbol{v}=\begin{cases}0 & k\neq i\\ \boldsymbol{e}_j & k=i\end{cases}\quad \text{从而} x_k=\begin{cases}0 & k\neq i\\ 1 & k=i\end{cases}$$

即 $\boldsymbol{v}=\boldsymbol{e}_i$.

3. 这等价于说, 对任意矩阵 A, 向量 $\boldsymbol{v}$ 使得 $A^{\mathrm{T}}\boldsymbol{v}^{\mathrm{T}}$ 是 A^{T} 的第 i 列; 由上一小题, $\boldsymbol{v}=\boldsymbol{e}_i$.

4. 设 $A=[a_{ij}]$, 则 $A\boldsymbol{e}_j=\begin{bmatrix}a_{1j}\\ a_{2j}\\ \cdots\\ a_{nj}\end{bmatrix}$, 所以 $\boldsymbol{e}_i^{\mathrm{T}}A\boldsymbol{e}_j=a_{ij}$.

5. 对 $\boldsymbol{e}_k\in\mathbb{R}^m$, $\boldsymbol{e}_k\boldsymbol{e}_k^{\mathrm{T}}=E_{kk}$, 所以，$\sum_{k=1}^{m}\boldsymbol{e}_k\boldsymbol{e}_k^{\mathrm{T}}=\sum_{k=1}^{m}E_{kk}=I_m$.

□

练习 1.4.28. 注意 m 维向量是 $m\times 1$ 矩阵.

1. 对 $\boldsymbol{v}\in\mathbb{R}^m, k\in\mathbb{R}$，$k\boldsymbol{v}$ 与 $\boldsymbol{v}k$ 是否可以看作矩阵乘法?

2. 对 $\boldsymbol{v}\in\mathbb{R}^m,\boldsymbol{w}\in\mathbb{R}^n$，$\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}$ 是否良定义？如果是，乘积有几行几列？

3. 求 $\begin{bmatrix}1\\2\\\vdots\\100\end{bmatrix}\begin{bmatrix}10 & 9 & \dots & 1\end{bmatrix}$ 的 $(12,7)$ 元.

4. 求 $\boldsymbol{v}\boldsymbol{w}$，使得 $\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}=\begin{bmatrix}(-1)^{i+j}\end{bmatrix}$.

5. 求 $\boldsymbol{v},\boldsymbol{w}$，使得 $\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}=\begin{bmatrix}\frac{i}{j}\end{bmatrix}$.

6. 令 $A=\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_n\end{bmatrix},B=\begin{bmatrix}\boldsymbol{b}_1^{\mathrm{T}}\\\boldsymbol{b}_2^{\mathrm{T}}\\\vdots\\\boldsymbol{b}_n^{\mathrm{T}}\end{bmatrix}$. 证明 $AB=\displaystyle\sum_{i=1}^n\boldsymbol{a}_i\boldsymbol{b}_i^{\mathrm{T}}$.

证明*:* 1. $k\boldsymbol{v}$ 不可以看成矩阵乘法，$\boldsymbol{v}k$ 可以；因为 k 看成矩阵只有 1 列。

2. $\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}$ 良定义，乘积有 m 行 n 列.

3. $\begin{bmatrix}1\\2\\\vdots\\100\end{bmatrix}\begin{bmatrix}10 & 9 & \dots & 1\end{bmatrix}$ 的 $(12,7)$ 元为 $12*4=48$

4. 设 $\boldsymbol{v}=\begin{bmatrix}x_1\\\vdots\\x_m\end{bmatrix},\boldsymbol{w}=\begin{bmatrix}y_1\\\vdots\\y_n\end{bmatrix}$, 于是就是 $x_iy_j=(-1)^{i+j}$ 直接求解可知 $x_i=a(-1)^i,y_j=a^{-1}(-1)^j$, 其中 $a\neq 0$.

5. 设 $\boldsymbol{v}=\begin{bmatrix}x_1\\\vdots\\x_m\end{bmatrix},\boldsymbol{w}=\begin{bmatrix}y_1\\\vdots\\y_n\end{bmatrix}$, 于是就是 $x_iy_j=\frac{i}{j}$, 直接求解，就有 $x_i=ai,y_j=\frac{a}{j}$, 其中 $a\neq 0$.

6. 设 $A=[a_{ij}],B=[b_{ij}]$. 那么由定义，$(AB)_{ij}=\sum_{k=1}^n a_{ik}b_{kj}$. 另一方面，

$$\left(\sum_{k=1}^n\boldsymbol{a}_k\boldsymbol{b}_k^{\mathrm{T}}\right)_{ij}=\sum_{k=1}^n\left(\boldsymbol{a}_k\boldsymbol{b}_k^{\mathrm{T}}\right)_{ij}$$

而 $(\boldsymbol{a}_k\boldsymbol{b}_k^{\mathrm{T}})_{ij}=a_{ik}b_{kj}$, 这就证完了。

□

1.5 矩阵的逆

练习 1.5.1. 计算下列矩阵乘法：

1. $\begin{bmatrix} & & 1\\ & 1 & \\1 & & \end{bmatrix}\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix}\begin{bmatrix} & & 1\\ & 1 & \\1 & & \end{bmatrix}$.

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

3. $\begin{bmatrix} & 1 & & \\ & & 1 & \\ & & & 1 \\ 1 & & & \end{bmatrix}^k$，$k$ 是正整数.　　　　*4.* $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

5. $\begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ -1 & & 1 & & \\ -1 & & & 1 & \\ -1 & & & & 1 \end{bmatrix}\begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 2 & 1 & & \\ 1 & 3 & 3 & 1 & \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$.

证明*:*　1. 这个矩阵乘法是先交换 $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ 的第一行和第三行，然后再对新矩阵交换它的第一列和第三列，因此

$$\begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}\begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} = \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}.$$

2. 这个矩阵乘法是将 $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 1 \\ 1 & 4 & 0 \end{bmatrix}$ 的第一行的 -1 倍分别加到第二行和第三行上，因此

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

3. 我们找规律即可

$$\begin{bmatrix} & 1 & & \\ & & 1 & \\ & & & 1 \\ 1 & & & \end{bmatrix}^k = \begin{cases} \begin{bmatrix} & 1 & & \\ & & 1 & \\ & & & 1 \\ 1 & & & \end{bmatrix} & k \equiv 1 \pmod 4 \\ \begin{bmatrix} & & 1 & \\ & & & 1 \\ 1 & & & \\ & 1 & & \end{bmatrix} & k \equiv 2 \pmod 4 \\ \begin{bmatrix} & & & 1 \\ 1 & & & \\ & 1 & & \\ & & 1 & \end{bmatrix} & k \equiv 3 \pmod 4 \\ \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} & k \equiv 4 \pmod 4 \end{cases}$$

4. 这个无它，直接乘

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix}$$

5. 这个矩阵乘法是将矩阵 $\begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 2 & 1 & & \\ 1 & 3 & 3 & 1 & \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$ 第一行的的 -1 倍分别加到其他行，因此

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

□

练习 1.5.2. 求下列矩阵的逆矩阵：

1. $\begin{bmatrix} 1 & 1 & \\ & 1 & 1 \\ & & 1 \end{bmatrix}$.

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

3. $\begin{bmatrix}1&2&0\\2&1&-1\\3&1&1\end{bmatrix}$.　　*4.* $\begin{bmatrix}1&-1&-1&-1\\-1&1&-1&-1\\-1&-1&1&-1\\-1&-1&-1&1\end{bmatrix}$.

证明*:* 这就是对于 Gauss-Jordan 消去的练习。我们具体计算一下第 3 小题，其他小题仅给出答案。

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

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

3. $\left[\begin{array}{ccc|ccc}1&2&0&1&0&0\\2&1&-1&0&1&0\\3&1&1&0&0&1\end{array}\right]\to\left[\begin{array}{ccc|ccc}1&2&0&1&0&0\\0&-3&-1&-2&1&0\\0&-5&1&-3&0&1\end{array}\right]\to\left[\begin{array}{ccc|ccc}1&2&0&1&0&0\\0&-3&-1&-2&1&0\\0&0&\frac{8}{3}&\frac{1}{3}&-\frac{5}{3}&1\end{array}\right]$
$\to\left[\begin{array}{ccc|ccc}1&2&0&1&0&0\\0&-3&-1&-2&1&0\\0&0&1&\frac{1}{8}&-\frac{5}{8}&\frac{3}{8}\end{array}\right]\to\left[\begin{array}{ccc|ccc}1&2&0&1&0&0\\0&-3&0&-\frac{15}{8}&\frac{3}{8}&\frac{3}{8}\\0&0&1&\frac{1}{8}&-\frac{5}{8}&\frac{3}{8}\end{array}\right]\to\left[\begin{array}{ccc|ccc}1&2&0&1&0&0\\0&1&0&\frac{5}{8}&-\frac{1}{8}&-\frac{1}{8}\\0&0&1&\frac{1}{8}&-\frac{5}{8}&\frac{3}{8}\end{array}\right]$
$\to\left[\begin{array}{ccc|ccc}1&0&0&-\frac{1}{4}&\frac{1}{4}&\frac{1}{4}\\0&1&0&\frac{5}{8}&-\frac{1}{8}&-\frac{1}{8}\\0&0&1&\frac{1}{8}&-\frac{5}{8}&\frac{3}{8}\end{array}\right]$
所以 $\begin{bmatrix}1&2&0\\2&1&-1\\3&1&1\end{bmatrix}^{-1}=\begin{bmatrix}-\frac{1}{4}&\frac{1}{4}&\frac{1}{4}\\\frac{5}{8}&-\frac{1}{8}&-\frac{1}{8}\\\frac{1}{8}&-\frac{5}{8}&\frac{3}{8}\end{bmatrix}$

4. 这一题我们注意到原矩阵本质是一个反射：

$$\begin{bmatrix}1&-1&-1&-1\\-1&1&-1&-1\\-1&-1&1&-1\\-1&-1&-1&1\end{bmatrix}=2\left(I_4-2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}\right)$$

其中 $\boldsymbol{w}=\begin{bmatrix}1/2\\1/2\\1/2\\1/2\end{bmatrix}$, 因此

$$\begin{bmatrix}1&-1&-1&-1\\-1&1&-1&-1\\-1&-1&1&-1\\-1&-1&-1&1\end{bmatrix}^{-1}=\frac{1}{2}(I_4-2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}})=\frac{1}{4}\begin{bmatrix}1&-1&-1&-1\\-1&1&-1&-1\\-1&-1&1&-1\\-1&-1&-1&1\end{bmatrix}$$

$\square$

练习 1.5.3. 设 $A=\begin{bmatrix}1&2&-1\\3&a&-2\\5&-2&1\end{bmatrix}$ 不可逆，求 a.

证明: 我们对 A 矩阵作初等行变换（省略具体步骤）可得: $\begin{bmatrix}1&0&0\\0&1&-\frac{1}{2}\\0&0&\frac{1}{2}a-2\end{bmatrix}$ 因此 A 不可逆，当且仅当 $a=4$. $\square$

练习 1.5.4. 设 $A=\begin{bmatrix}1&&\\2&3&\\4&5&6\end{bmatrix}$，解方程组 $A\boldsymbol{x}=\boldsymbol{e}_i, i=1,2,3$，并求 A^{-1}.

证明: $A\boldsymbol{x}=\boldsymbol{e}_1$ 的解是 $\begin{bmatrix}1\\-\frac{2}{3}\\-\frac{1}{9}\end{bmatrix}$, $A\boldsymbol{x}=\boldsymbol{e}_2$ 的解释 $\begin{bmatrix}0\\\frac{1}{3}\\-\frac{5}{18}\end{bmatrix}$, $A\boldsymbol{x}=\boldsymbol{e}_3$ 的解释 $\begin{bmatrix}0\\0\\\frac{1}{6}\end{bmatrix}$，所以

$$A^{-1}=\begin{bmatrix}1&0&0\\-\frac{2}{3}&\frac{1}{3}&0\\-\frac{1}{9}&-\frac{5}{18}&\frac{1}{6}\end{bmatrix}$$

$\square$

练习 1.5.5. 对 $\begin{bmatrix}A&I_n\end{bmatrix}$ 做初等行变换，求下列矩阵的逆矩阵:

1. $\begin{bmatrix}2&-1&&\\-1&2&-1&\\&-1&2&-1\\&&-1&2\end{bmatrix}$.
2. $\begin{bmatrix}1&-1&&\\-1&2&-1&\\&-1&2&-1\\&&-1&2\end{bmatrix}$.
3. $\begin{bmatrix}0&1&&\\&0&&1\\1&&0&\\&&1&0\end{bmatrix}$.
4. $\begin{bmatrix}&&&1\\&&1&\\&\iddots&&\\1&&&\end{bmatrix}$.
5. $\begin{bmatrix}1&2&&\\3&4&&\\&&1&2\\&&3&4\end{bmatrix}$.
6. $\begin{bmatrix}3&-1&-1\\-1&3&-1\\-1&-1&3\end{bmatrix}$.
7. $\begin{bmatrix}1&-a&&\\&1&-b&\\&&1&-c\\&&&1\end{bmatrix}$.
8. $\begin{bmatrix}1&a&&\\&1&\ddots&\\&&\ddots&a\\&&&1\end{bmatrix}_{n\times n}$.

证明: 1. 我们忽略过程，$A^{-1}=\begin{bmatrix}\frac{4}{5}&\frac{3}{5}&\frac{2}{5}&\frac{1}{5}\\\frac{3}{5}&\frac{6}{5}&\frac{4}{5}&\frac{2}{5}\\\frac{2}{5}&\frac{4}{5}&\frac{6}{5}&\frac{3}{5}\\\frac{1}{5}&\frac{2}{5}&\frac{3}{5}&\frac{4}{5}\end{bmatrix}$

2. $$\begin{bmatrix} 1 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 4 & 3 & 2 & 1 \\ 3 & 3 & 2 & 1 \\ 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

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

4. $$\begin{bmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{bmatrix}^{-1} = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{bmatrix}$$

5. $$\begin{bmatrix} 1 & 2 & & \\ 3 & 4 & & \\ & & 1 & 2 \\ & & 3 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} -2 & 1 & & \\ \frac{3}{2} & -\frac{1}{2} & & \\ & & -2 & 1 \\ & & \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$$

6. $$\begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

7. $$\begin{bmatrix} 1 & -a & & \\ & 1 & -b & \\ & & 1 & -c \\ & & & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & a & ab & abc \\ & 1 & b & bc \\ & & 1 & c \\ & & & 1 \end{bmatrix}$$

8. 要求这个矩阵的逆矩阵，逐次把每一行的 $-a$ 加到上一行即可，

$$\begin{bmatrix} 1 & a & & \\ & 1 & \ddots & \\ & & \ddots & a \\ & & & 1 \end{bmatrix}_{n\times n}^{-1} = \begin{bmatrix} 1 & (-a) & & (-a)^{n-1} \\ & 1 & \ddots & \\ & & \ddots & -a \\ & & & 1 \end{bmatrix}_{n\times n}$$

□

练习 1.5.6. 说明 $\begin{bmatrix} 4 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & 4 \end{bmatrix}$ 可逆，并计算其逆.

证明: 这是个对角占优矩阵，因此可以。要计算它的逆，最朴实的方法还是对 $\begin{bmatrix} A & I_n \end{bmatrix}$ 做初

等行变换，我们忽略。还可以利用 Sherman-Morrison 公式。$\begin{bmatrix} 4 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & 4 \end{bmatrix} = 5I_4 + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 & -1 \end{bmatrix}$, 所以

$$\begin{bmatrix} 4 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & 4 \end{bmatrix}^{-1} = \frac{1}{5}I_4 - \frac{1}{1-\frac{4}{5}}\frac{1}{5}I_4 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 & -1 \end{bmatrix} \frac{1}{5}I_4 = \begin{bmatrix} \frac{2}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{2}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{2}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{2}{5} \end{bmatrix}$$

□

练习 1.5.7. 求 $A = \begin{bmatrix} & a_1 & & \\ & & \ddots & \\ & & & a_{n-1} \\ a_n & & & \end{bmatrix}$ 的逆矩阵，其中 $a_i \neq 0$.

证明*:*

$$\left[\begin{array}{cccc|cccc} & a_1 & & & 1 & & & \\ & & \ddots & & & \ddots & & \\ & & & a_{n-1} & & & 1 & \\ a_n & & & & & & & 1 \end{array}\right] \to \left[\begin{array}{cccc|cccc} a_n & & & & & & & 1 \\ & a_1 & & & 1 & & & \\ & & \ddots & & & \ddots & & \\ & & & a_{n-1} & & & 1 & \end{array}\right]$$

因此 $A^{-1} = \begin{bmatrix} & & & a_n^{-1} \\ a_1^{-1} & & & \\ & \ddots & & \\ & & a_{n-1}^{-1} & \end{bmatrix}$ □

练习 1.5.8. 求下列矩阵方程的解：$\begin{bmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix} X = \begin{bmatrix} 4 & -1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix}$.

证明*:* 我们只需要对

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

作初等行变换把 $\begin{bmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix}$ 转化成 I_3 即可，右侧的矩阵就是我们要解的 X. 我们省略过程，

$$X = \begin{bmatrix} -7 & -5 \\ -9 & -6 \\ 12 & 7 \end{bmatrix}$$

□

练习 1.5.9. 证明二阶方阵 $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ 可逆当且仅当 $d = a_{11}a_{22} - a_{12}a_{21} \neq 0$. 此时，

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

证明: A 可逆当且仅当 $A\boldsymbol{x} = 0$ 只有零解，由题 1.3.4 知当且仅当 $d \neq 0$. A 可逆时，验证 A^{-1} 为题目中的形式是容易的：我们只需要做矩阵乘法即可。 □

练习 1.5.10. 证明：有一列元素（或一行元素）全为零的方阵不可逆.

证明: 设 A 的的第 i 列为零，则 $\boldsymbol{e}_i$ 是齐次方程组 $A\boldsymbol{x} = 0$ 的非零解，因此 A 不可逆。

若 A 的第 i 行为零，则 A^{T} 的第 i 列为零，因此 A^{T} 不可逆，所以 A 也不可逆。 □

练习 1.5.11. 证明：对角元素全非零的上三角矩阵 U 可逆，其逆矩阵 U^{-1} 也是上三角矩阵，且 U^{-1} 的对角元素是 U 的对角元素的倒数.

证明: 我们可以对 U 作倍加初等行变换和数乘初等行变换，把 U 转化成单位矩阵。过程如下：先通过把最后一行的若干倍加往它上面的各行，可以把最后一列除了主元以外全变成零，再通过把倒数第二行的若干倍加往它上面的各行，可以把倒数第二列除了主元以外全变成零，依次类推，这样就把 U 变成对角矩阵，且对角线元素全非零。最后对每一行作数乘行变换即可

在这个过程中，每一个倍加变换都是对角线为 1 的上三角矩阵；对第 i 行作数乘变换是对角矩阵，其第 i 个对角元是 U 第 i 个对角元倒数，其他对角元是 1. 利用题 1.4.13 的结论，即可得到 U^{-1} 是上三角矩阵，且它的对角元素是 U 的对角元素的倒数。 □

我们还可以利用分块矩阵和归纳法来证明这个命题，在此略过。

练习 1.5.12. 是否只有方阵才有可能可逆？设 A 是 $m \times n$ 矩阵.

1. 设 $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$，能否找到 B, C 使得 $AB = I_2, CA = I_3$?

2. 设 $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}$，能否找到 B, C 使得 $AB = I_3, CA = I_2$?

3. 设 $CA = I_n$，证明 $A\boldsymbol{x} = 0$ 只有零解. 此时 m, n 之间有何种关系？

4. 设 $AB = I_m$，证明 $A^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{0}$ 只有零解. 此时 m, n 之间有何种关系？

5. 设 $AB = I_m, CA = I_n$，证明 $m = n$ 且 $B = C$；由此可知，可逆矩阵一定是方阵.

6. 如果 $m \neq n$，那么 $\mathbb{R}^m, \mathbb{R}^n$ 之间是否存在线性双射？

证明*:* 我们先证明第 3,4 小题

3. $A\boldsymbol{x} = 0$, 则 $\boldsymbol{x} = CA\boldsymbol{x} = C0 = 0$, 因此 $A\boldsymbol{x} = 0$ 只有零解。根据 Gauss 消元法，若方程个数比未知数个数少，则齐次方程组一定要非零解，因此有 $m \geq n$.

4. 若 $AB = I_m$，则 $B^T A^T = I_m$，由 3, $A^{\mathrm{T}}\boldsymbol{x} = 0$ 只有零解，此时 $n \geq m$.

1. 设 $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$，令 $B = \begin{bmatrix} -\frac{5}{3} & -\frac{2}{3} \\ \frac{4}{3} & \frac{1}{3} \\ 0 & 0 \end{bmatrix}$, 则有 $AB = I_2$. 由第 3 小题知不存在 C 使得 $CA = I_3$.

2. 都不存在。因为 $A\boldsymbol{x} = 0$ 有非零解 $\boldsymbol{x} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $A^{\mathrm{T}}\boldsymbol{x} = 0$ 有非零解 $\boldsymbol{x} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$.

5. 首先由第 3,4 小题知 $m \geq n$ 且 $n \geq m$, 因此 $m = n$. 此时 $C = C(AB) = (CA)B = B$.

3. 如果 $m \neq n$，则 $\mathbb{R}^m, \mathbb{R}^n$ 不存在线性双射。设 $m \neq n$ 且 $f : \mathbb{R}^m \to \mathbb{R}^n$ 是线性映射，再由题 1.3.21 知：若 $m > n$, 则 f 不是单射，若 $n > m$ 则 f 不是满射；总之 f 不是双射。

$\square$

练习 1.5.13. 如果 n 阶方阵 A, B 满足 $AB = I_n$，判断 A, B 是否可逆.

证明*:* A, B 可逆。我们需要证明 $BA = I_n$.

$AB = I_n$, 则 $B\boldsymbol{x} = 0$ 只有零解，根据 Gauss 消元法，B 的行简化阶梯形矩阵为 I_n，因此对任意的 $\boldsymbol{b} \in \mathbb{R}^n$, 增广矩阵 $[B|\boldsymbol{b}]$ 的秩等于系数矩阵 B 的秩，都等于 n, 因此 $B\boldsymbol{x} = \boldsymbol{b}$ 都有解。令 $\boldsymbol{c}_i$ 是 $B\boldsymbol{x} = \boldsymbol{e}_i$ 的解 $C = \begin{bmatrix} \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_n \end{bmatrix}$, 则有 $BC = I_n$, 此时 $A = C$. $\square$

练习 1.5.14. 设

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

1. 求上述矩阵对应的初等行、列变换.

2. 从行变换的角度看，是否一定有 $AB = BA$? 如果否，$AB - BA$ 从行变换的角度意味着什么？

3. 从行变换的角度看，是否一定有 $AC = CA$? 如果否，$AC - CA$ 从行变换的角度意味着什么？

4. 从行变换的角度看，是否一定有 $BC = CB$? 如果否，$BC - CB$ 从行变换的角度意味着什么？

5. 从行变换的角度看，D 是否和 A, B, C 可交换？计算 $DAD^{-1}, DBD^{-1}, DCD^{-1}$.

6. 从行变换的角度看，P 是否和 A, B, C 可交换？计算 $PAP^{-1}, PBP^{-1}, PCP^{-1}$.

7. 对三阶方阵 X，$(AX)B$ 和 $A(XB)$ 对 X 分别做了何种行、列变换？
8. 对任意矩阵 X，先做初等行变换，再做初等列变换，其结果是否等于先做该初等列变换，再做该初等行变换？这对应着矩阵乘法的什么性质？

证明: 1. 一个矩阵的左边乘 A 即将它的第 2 行加到第 1 行；左边乘 B 即将它的第 3 行加到第 1 行; 左边乘 C 即将它的第 3 行加到第 2 行; 左边乘 D 即将它的第 1 行乘以 2; 左边乘 P 将它的第 2 行移到第 1 行，第 3 行移到第 2 行，第 1 行移到第 3 行。一个矩阵的右边乘 A 即将它的第 1 列加到第 2 列；右边乘 B 即将它的第 1 列加到第 3 列; 右边乘 C 即将它的第 2 列加到第 3 列; 右边乘 D 即将它的第 1 列乘以 2; 右边乘 P 即将它的第 1 列移到第 2 列，第 2 列移到第 3 列，第 3 列移到第 1 列。

2. 一定有 $AB = BA$, 从行变换的角度看，结果都是把矩阵的第 2 行和第 3 行加到第 1 行.
3. $AC \neq CA$. AC 是将一个矩阵的第 3 行加到第 2 行，再将新矩阵的第 2 行加到第一行，即将原矩阵的第 2、3 行同时加到第一行；类似的考虑，CA 仅仅是将原矩阵的第 2 行加到第 1 行，第 3 行加到第 2 行；因此 $AC - CA$ 是将一个矩阵的第 2 行挪到第一行，同时其他行全都变成 0 了。
4. $BC = CB$, 从行变换的角度看，结果都是把矩阵的第 3 行分别加到第 1 行和第 2 行.
5. D 和 A, B 不可交换，和 C 可交换。用 DAD^{-1} 是将 A 的第一行乘以 2，第一列乘以 1/2，因此 $DAD^{-1} = \begin{bmatrix}1&2&0\\0&1&0\\0&0&1\end{bmatrix}$, 同理 $DBD^{-1} = \begin{bmatrix}1&0&2\\0&1&0\\0&0&1\end{bmatrix}$, $DCD^{-1} = \begin{bmatrix}1&0&0\\0&1&1\\0&0&1\end{bmatrix}$.
6. 都不可交换。PAP^{-1} 是对 A 的行做了置换，再对列做置换，因此
$$PAP^{-1} = \begin{bmatrix}1&0&0\\0&1&0\\1&0&1\end{bmatrix}, PBP^{-1} = \begin{bmatrix}1&0&0\\0&1&0\\0&1&1\end{bmatrix}, PCP^{-1} = \begin{bmatrix}1&1&0\\0&1&0\\0&0&1\end{bmatrix}$$
7. 对三阶方阵 X，$(AX)B$ 先将 X 的第二行加到第一行，再将新矩阵的第一列加到第三列；$A(XB)$ 先将 X 的第一列加到第三列，再将新矩阵的第二行加到第一行。
8. 等于。这对应了矩阵乘法的结合律。

□

练习 1.5.15. 设矩阵 A 和 B 左相抵. 求证：

1. 如果 A 的第一列全是零，则 B 的第一列全是零.
2. 如果 A 的所有列都相同，则 B 的所有列都相同.
3. 如果 A 的第一列是第二列与第三列的和，则 B 的第一列也是第二列与第三列的和.
4. 如果 A 的第一列和第二列不成比例，则 B 的第一列和第二列也不成比例.

证明: A 和 B 左相抵等价于说对 A 做初等行变换可以得到 B，也等价于存在可逆矩阵 P, 使得 $B = PA$.

1. A 的第一列全是零等价于说 $A\boldsymbol{e}_1 = 0$, 因此 $B\boldsymbol{e}_1 = PA\boldsymbol{e}_1 = 0$

2. A 的所有列都相同，等价于说，对任意 i,j 都有 $A\boldsymbol{e}_i = A\boldsymbol{e}_j$, 因此

$$B\boldsymbol{e}_i = PA\boldsymbol{e}_i = PA\boldsymbol{e}_j = B\boldsymbol{e}_j$$

即 B 的所有列都相同。

3. A 的第一列是第二列与第三列的和，等价于说 $A\boldsymbol{e}_1 = A\boldsymbol{e}_2 + A\boldsymbol{e}_3$, 因此 $B\boldsymbol{e}_1 = PA\boldsymbol{e}_1 = P(A\boldsymbol{e}_2 + A\boldsymbol{e}_3) = B\boldsymbol{e}_2 + B\boldsymbol{e}_3$, 即 B 的第一列也是第二列与第三列的和.

4. A 的第一列和第二列不成比例，等价于说对任意不全为零的 λ, μ, $\lambda A\boldsymbol{e}_1 \neq \mu A\boldsymbol{e}_2$, 由于 P 可逆，则对任意不全为零的 λ, μ, $\lambda PA\boldsymbol{e}_1 \neq \mu PA\boldsymbol{e}_2$，所以 B 的第一列和第二列也不成比例。

□

练习 1.5.16. 设有 M_1, M_2, M_3 三个城市，城市之间有人口迁移. 定义矩阵 A，其 (i,j) 元 a_{ij} 为人在一年中从 A_j 迁移到 A_i 的概率. 注意，A 的每个元素都在 $0,1$ 之间，且 A 的每一列的元素之和都是 1.

1. 设今年城市 M_1, M_2, M_3 的人口分别是 x_1, x_2, x_3，证明明年它们的预期人口分别是 $A\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ 的三个分量.

2. 人口迁移满足什么条件时，A 为列对角占优矩阵 (行对角占优矩阵的转置)? 人口迁移满足什么条件时，A 为行对角占优矩阵? 考虑现实生活中的情形，讨论这些假设是否合理.

3. 如果把矩阵对角占优定义中的大于号换成大于等于号，则称该矩阵为**弱对角占优矩阵**. 证明 $A = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$ 弱对角占优，且可逆，并求其逆矩阵.

4. 对任意 n 阶方阵，设它有 $n-1$ 行都是对角占优，仅有 1 行弱对角占优，该矩阵可逆吗?

证明*:* 1. 设明年城市 M_1, M_2, M_3 的预期人口分别是 y_1, y_2, y_3. 则 y_j 等于 M_j 不迁移的人口加上其他两个城市迁移进来的人口。因此 $y_j = a_{j1}x_1 + a_{j2}x_2 + a_{j3}x_3$, 写成矩阵形式，即

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

2. 列对角占优即 $a_{11} > a_{21} + a_{31}, a_{22} > a_{12} + a_{13}, a_{33} > a_{13} + a_{23}$; 行对角占优即 $a_{11} > a_{12} + a_{13}, a_{22} > a_{21} + a_{23}, a_{33} > a_{31} + a_{32}$. 考虑实际意义，列对角占优意味着每个城市不迁移的比例比其他城市往这个城市迁移的比例要大，行对角占优意味着每个城市不迁移的比例比它往外迁移的比例要大。这两个假设都是合理的。

3. A 弱对角占优是明显的，要证明它可逆并求逆矩阵，只需要用 Gauss-Jordan 消去，我们略过。逆矩阵为 $\begin{bmatrix} 2 & -1 & -1 \\ 0 & 1 & 0 \\ -1 & -1 & 2 \end{bmatrix}$

4. 这样的若对角占优矩阵未必可逆。例如：我们取一个 n 阶严格对角占优矩阵，然后把它的最后一行替换成 0，得到矩阵 A. 那么显然 A 的前 $n-1$ 行严格对角占优，最后一行弱对角占优，所以符合题意。但是显然 A 不可逆。

□

练习 1.5.17. *设 A 是 n 阶方阵.*

1. *对任意两个多项式 $p(x), q(x)$，是否一定有 $p(A)q(A) = q(A)p(A)$?*

2. *证明所有形如 $\begin{bmatrix} a & b & c \\ & a & b \\ & & a \end{bmatrix}$ 的矩阵全都彼此交换.*

3. *设 $f(x) = p(x) + q(x), g(x) = p(x)q(x)$，证明：$f(A) = p(A) + q(A), g(A) = p(A)q(A)$.*

4. *求 A 使得 $A + I_n, A - I_n$ 均不为零矩阵，但是 $A^2 - I_n = 0$.*

5. *设 $A^2 - I_n = 0$，证明：任意 n 维向量 $\boldsymbol{v}$ 都存在分解式 $\boldsymbol{v} = \boldsymbol{x} + \boldsymbol{y}$，其中 $\boldsymbol{x}, \boldsymbol{y}$ 满足*

$$(A - I_n)\boldsymbol{x} = (A + I_n)\boldsymbol{y} = 0.$$

6. *设 $A^3 = 0$，证明 $A + I_n$ 与 $A - I_n$ 都可逆，并求 $p(x), q(x)$ 使得 $p(A), q(A)$ 分别为其逆.*

证明*:* 1. 一定相等。在承认第 3 小题的前提下（第 3 小题将独立证明，因此没有逻辑上的问题），我们利用分配律，只需要考虑 $p = ax^m, q = bx^n$ 是单项式的情形，即只需证明 $(aA^m)(bA^n) = (bA^n)(aA^m)$。这是显然的。

2. 令 $J_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, 如题形式的矩阵为 $cJ_3^2 + bJ_3 + cI_n$, 他们都是某个多项式在 J_3 上的取值，因此彼此交换。

3. 通过补一些系数是 0 的项，我们可以假设 $p(x) = a_0 + a_1x + \cdots + a_mx^m, q(x) = b_0 + b_1x + \cdots + b_mx^m$, 则 $f(x) = \sum_{i=0}^{m}(a_i + b_i)x^i$. 此时 $f(A) = \sum_{i=0}^{m}(a_i + b_i)A^i, p(A) = \sum_{i=0}^{m} a_iA^i, q(A) = \sum_{i=0}^{m} b_iA^i$, 因此由矩阵加法的交换律，直接就有 $f(A) = p(A) + q(A)$. $g(x) = \sum_{i=0}^{2m}\left(\sum_j a_jb_{i-j}\right)x^i$ (若 $i < j$, 则 $b_{i-j} = 0$). 因此

$$p(A)q(A) = \left(\sum_{i=0}^{m} a_iA^i\right)\left(\sum_{i=0}^{m} b_iA^i\right) = \sum_{i=0}^{2m}\left(\sum_j a_jb_{i-j}\right)A^i = g(A).$$

4. A 一定共轭于形如 $\operatorname{diag}(1, \cdots, 1, -1, \cdots, -1)$ 的矩阵。在学习了相似对角化以后再给完整的证明。

5. 若存在满足题意的分解式，则 $A\boldsymbol{x} = \boldsymbol{x}, A\boldsymbol{y} = -\boldsymbol{y}$, 因此 $A\boldsymbol{v} = A\boldsymbol{x} + A\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{y}$, 所以 $\boldsymbol{x} = \dfrac{\boldsymbol{v} + A\boldsymbol{v}}{2}, \boldsymbol{y} = \dfrac{\boldsymbol{v} - A\boldsymbol{v}}{2}$. 事实上，我们令 $\boldsymbol{x} = \dfrac{\boldsymbol{v} + A\boldsymbol{v}}{2}, \boldsymbol{y} = \dfrac{\boldsymbol{v} - A\boldsymbol{v}}{2}$ 就满足题目中的条件。

6. $A^3=0$, 则

$$I_n = A^3 + I_n = (A+I_n)(A^2 - A + I_n), \quad I_n = I_n - A^3 = (I_n - A)(A^2 + A + I_n)$$

因此 $(A+I_n)^{-1} = A^2 - A + I_n, (I_n - A)^{-1} = A^2 + A + I_n$, 因此我们令 $p(x) = x^2 - x + 1, q(x) = x^2 + x + 1$.

□

练习 1.5.18. 证明，如果 n 阶方阵 A 满足 $A^2 = A$，则 $I_n - 2A$ 可逆.

证明: 我们只需证明齐次方程组 $(I_n - 2A)\boldsymbol{x} = 0$ 没有非零解。假设 $(I_n - 2A)\boldsymbol{x} = 0$, 则 $\boldsymbol{x} = 2A\boldsymbol{x}$, $A\boldsymbol{x} = 2A^2\boldsymbol{x} = 2A\boldsymbol{x}$, 所以 $A\boldsymbol{x} = 0$, $\boldsymbol{x} = 0$. □

练习 1.5.19. 如果一个 n 阶方阵从 $(1,n)$ 元到 $(n,1)$ 元的对角线下的所有元素均为零，则称为西北矩阵. 类似地，可以定义东南矩阵. 如果 B 是西北矩阵，那么 $B^{\mathrm{T}}, B^2, B^{-1}$ 是什么矩阵? 西北矩阵和东南矩阵的乘积是什么矩阵?

证明: 令 $P = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}$, 矩阵 X 是西北矩阵，当且仅当 PX 是下三角矩阵，当且仅当 XP 是上三角矩阵。矩阵 X 是东南矩阵，当且仅当 PX 是上三角矩阵，当且仅当 XP 是下三角矩阵。我们需要注意到，$P = P^T = P^{-1}$. 设 $A = PB$.

若 B 是西北矩阵，则 A 是下三角矩阵，因此 A^{T} 是上三角矩阵，而 $B^{\mathrm{T}}P = A^{\mathrm{T}}$, 因此 B^{T} 也是西北矩阵。B^2 的形状不确定，我们举简单的例子就可以发现这一点，例如 $B = \begin{bmatrix} a & b \\ c & 0 \end{bmatrix}$ 其中 $a,b,c \in \{1,-1,0\}$. $B^{-1}P = A^{-1}$, 而 A^{-1} 是下三角矩阵，因此 B^{-1} 是东南矩阵。

若 C 是东南矩阵，则 $BC = BP^2C = (BP)(PC)$; 此时，BP, PC 都是上三角矩阵，因此 BC 是上三角矩阵。 □

练习 1.5.20.

1. 求一对可逆矩阵 A, B，使得 $A+B$ 不可逆.

2. 求一组不可逆矩阵 A, B，使得 $A+B$ 可逆.

3. 求一个三阶不可逆矩阵 A，使得对任意 $k>0$，$A + kI_3$ 都对角占优.

1. 例如 $A = I_n, B = -I_n$.
2. 例如 $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$，使得 $A+B$ 可逆.
3. 例如 A 是零矩阵。

练习 1.5.21. 求所有三阶矩阵 A，满足 $A^2 = I_3$，且 A 的每个元素只能是 0 或 1.

证明: 我们用最朴实的方法：死算。设 $A = \begin{bmatrix} x & y & z \\ u & v & w \\ r & s & t \end{bmatrix}$, 则

$$A^2 = \begin{bmatrix} x^2+yu+zr & xy+yv+zs & xz+yw+zt \\ ux+vu+wr & uy+v^2+ws & uz+vw+wt \\ rx+su+rt & ry+sv+ts & rz+ws+t^2 \end{bmatrix}$$

因为所有元素只能是 0,1, 所以 A^2 的非对角元中出现的所有单项都等于 0, A^2 的每一个对角元中只有一个单项是 1. 我们分类讨论。

- $x=1$, 则考虑 A^2 的第一行和第一列元素，既有 $y=z=u=r=0$. 因此 $v^2+ws=ws+t^2=1, vw=wt=sv=ts=0$. 若 $v=1$, 则 $t=1, w=s=0$. 若 $v=0$, 则 $w=s=1, t=0$. 此时一共有两种可能 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$。
- $x=0$, 则 $yu+zr=1$. 若 $yu=1$, 则 $y=u=1$, 考虑 A^2 非对角元的所有单项，则 $r=v=w=s=z=0$, 因此 $t=1$; 若 $zr=1$, 则 $z=r=1$. 考虑 A^2 非对角元的所有单项，则 $y=s=t=w=u=0$, 因此 $v=1$. 此时一共有两种可能 $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$。

□

练习 1.5.22. 设 $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ 是对称矩阵，通过“对称化简”求其逆矩阵：

1. 将 A 的第一行的 2 倍从第二行中减去，第一行的 3 倍从第三行中减去. 这对应哪个可逆矩阵 E_1? E_1^{T} 对应的列变换是什么？计算 E_1A 和 $A_1 = E_1AE_1^{\mathrm{T}}$.
2. 将 A_1 的第二行与第三行调换. 这对应哪个初等矩阵 E_2? E_2^{T} 对应的列变换是什么？计算 E_2A_1 和 $A_2 = E_2A_1E_2^{\mathrm{T}}$.
3. 将 A_2 的第二行的 $\dfrac{1}{3}$ 倍从第三行中减去. 这对应哪个初等矩阵 E_3? E_3^{T} 对应的列变换是什么？计算 E_3A_2 和 $A_3 = E_3A_2E_3^{\mathrm{T}}$.
4. 综上，$A_3 = E_3E_2E_1AE_1^{\mathrm{T}}E_2^{\mathrm{T}}E_3^{\mathrm{T}}$，由此求 A 的逆.

证明:

1. $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$ E_1^{T} 对应的列变换是将 A 的第一列的 2 倍从第二列中减去，第一列的 3 倍从第三列中减去.

$$E_1A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & -1 \\ 0 & -1 & -3 \end{bmatrix}, \quad A_1 = E_1AE_1^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & -3 \end{bmatrix}.$$

2. $E_2=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}$ E_2^{T} 对应的列变换是交换第二列和第三列。

$$E_2A_1=\begin{bmatrix}1&0&0\\0&-1&-3\\0&0&-1\end{bmatrix},A_2=E_2A_1E_2^{\mathrm{T}}=\begin{bmatrix}1&0&0\\0&-3&-1\\0&-1&0\end{bmatrix}$$

3. $E_3=\begin{bmatrix}1&0&0\\0&1&0\\0&-\frac{1}{3}&1\end{bmatrix}$; E_3^{T} 对应的列变换是将 A_2 的第二列的 $\dfrac{1}{3}$ 倍从第三列中减去.

$$E_3A_2=\begin{bmatrix}1&0&0\\0&-3&-1\\0&0&\frac{1}{3}\end{bmatrix},\quad A_3=E_3A_2E_3^{\mathrm{T}}=\begin{bmatrix}1&0&0\\0&-3&0\\0&0&\frac{1}{3}\end{bmatrix}$$

4. 综上，$A_3=E_3E_2E_1AE_1^{\mathrm{T}}E_2^{\mathrm{T}}E_3^{\mathrm{T}}$，

$$A^{-1}=E_1^{-1}E_2^{-1}E_3^{-1}A_3(E_3^{\mathrm{T}})^{-1}(E_2^{\mathrm{T}})^{-1}(E_1^{\mathrm{T}})^{-1}=\begin{bmatrix}1&-3&2\\-3&-2&-1\\2&-1&0\end{bmatrix}$$

□

练习 1.5.23. 求证：可逆对称矩阵的逆矩阵也是对称矩阵；可逆反对称矩阵的逆矩阵也是反对称矩阵.

证明*:*

$$(A^{-1})^T=(A^T)^{-1}=\begin{cases}A^{-1} & A\text{ 对称}\\-A^{-1} & A\text{ 反对称}\end{cases}$$

故命题得证。 □

练习 1.5.24. 给定 n 阶实反对称矩阵 A，求证：I_n-A 可逆.

证明*:* 我们只需证明：齐次方程组 $(I_n-A)\boldsymbol{x}=0$ 只有零解。设 $(I_n-A)\boldsymbol{x}=0$, 则 $A\boldsymbol{x}=\boldsymbol{x}$.

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=(A\boldsymbol{x})^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}(-A)\boldsymbol{x}=-\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=-\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x},$$

因此 $\boldsymbol{x}=0$. □

1.6 分块矩阵

练习 1.6.1. 设 A 的行简化阶梯形矩阵为 R，行变换对应的可逆矩阵是 P.

1. 求 $\begin{bmatrix}A & 2A\end{bmatrix}$ 的行简化阶梯形矩阵，以及行变换对应的可逆矩阵.

2. 求 $\begin{bmatrix} A \\ 2A \end{bmatrix}$ 的行简化阶梯形矩阵，以及行变换对应的可逆矩阵.

证明: 1. $\begin{bmatrix} A & 2A \end{bmatrix}$ 的行简化阶梯形矩阵是 $\begin{bmatrix} R & 2R \end{bmatrix}$，行变换对应的可逆矩阵是 P.

2. $\begin{bmatrix} A \\ 2A \end{bmatrix}$ 的行简化阶梯形矩阵是 $\begin{bmatrix} R \\ 0 \end{bmatrix}$，行变换对应的可逆矩阵是 $\begin{bmatrix} P & 0 \\ -2P & P \end{bmatrix}$.

□

练习 1.6.2. 求下列矩阵的逆矩阵：

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

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

证明: 1. $\begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}$, $\begin{bmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -4 & -3 \\ 1 & -5 & -3 \\ -1 & 6 & 4 \end{bmatrix}$, 所以

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 3 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -4 & -3 \\ 0 & 0 & 1 & -5 & -3 \\ 0 & 0 & -1 & 6 & 4 \end{bmatrix}$$

2. 令原矩阵为 A, $P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, 则 $AP = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. 因为

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix},$$

所以

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

□

练习 1.6.3. 计算下列分块矩阵：

1. $\begin{bmatrix} I_n & 0 \\ A & I_m \end{bmatrix}^{-1}$.

2. $\begin{bmatrix} 0 & I_m \\ I_n & A \end{bmatrix}^{-1}$.

证明: 1.

$$\begin{bmatrix} I_n & 0 \\ A & I_m \end{bmatrix} \begin{bmatrix} I_n & 0 \\ -A & I_m \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & I_m \end{bmatrix}$$

所以

$$\begin{bmatrix} I_n & 0 \\ A & I_m \end{bmatrix}^{-1} = \begin{bmatrix} I_n & 0 \\ -A & I_m \end{bmatrix}$$

2.

$$\begin{bmatrix} 0 & I_m \\ I_n & A \end{bmatrix} \begin{bmatrix} -A & I_n \\ I_m & 0 \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ 0 & I_n \end{bmatrix}$$

所以

$$\begin{bmatrix} 0 & I_m \\ I_n & A \end{bmatrix}^{-1} = \begin{bmatrix} -A & I_n \\ I_m & 0 \end{bmatrix}$$

□

练习 1.6.4. 设分块矩阵 $X = \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix}$，其中 A, B 可逆. 试证 X 可逆，并求其逆.

证明: 令 $Y = \begin{bmatrix} 0 & B^{-1} \\ A^{-1} & 0 \end{bmatrix}$, 则 $XY = YX = I_{m+n}$, 因此 $X^{-1} = Y$. □

练习 1.6.5. 设分块矩阵 $U = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$，其中 A, B 可逆. 试证 U 可逆，并求其逆.

证明: 根据上三角矩阵的性质,我们可以猜测分块上三角矩阵 U 的逆矩阵形如 $V = \begin{bmatrix} A^{-1} & X \\ 0 & B^{-1} \end{bmatrix}$。我们利用 $UV = I_{m+n}$ 来解 X:

$$UV = \begin{bmatrix} I_n & AX + CB^{-1} \\ 0 & I_m \end{bmatrix}$$

于是 $X = -A^{-1}CB^{-1}$. 所以 U 可逆，且 $U^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}CB^{-1} \\ 0 & B^{-1} \end{bmatrix}$ □

练习 1.6.6. 设 A_1, A_2 分别为 m, n 阶方阵，且存在可逆方阵 T_1, T_2，使得 $T_1^{-1}A_1T_1$ 和 $T_2^{-1}A_2T_2$ 都是对角矩阵. 求证：存在 $m+n$ 阶可逆矩阵 T，使得 $T^{-1}\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}T$ 是对角矩阵.

证明: 令 $T = \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix}$, 则 $T^{-1} = \begin{bmatrix} T_1^{-1} & 0 \\ 0 & T_2^{-1} \end{bmatrix}$, 那么

$$T^{-1} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} T = \begin{bmatrix} T_1^{-1}AT_1 & 0 \\ 0 & T_2^{-1}A_2T_2 \end{bmatrix}$$

是对角矩阵。 □

练习 1.6.7.

1. 任取 $m\times n$ 矩阵 X，分块矩阵 $\begin{bmatrix}A & B\\ C & D\end{bmatrix}$，计算 $\begin{bmatrix}I_m & 0\\ X & I_n\end{bmatrix}\begin{bmatrix}A & B\\ C & D\end{bmatrix}$.

2. 由此判断 $\begin{bmatrix}1 & & & & c_1\\ & 1 & & & c_2\\ & & \ddots & & \vdots\\ & & & 1 & c_{n-1}\\ b_1 & b_2 & \cdots & b_{n-1} & a\end{bmatrix}$ 何时可逆，并在可逆时求其逆.

证明*:* 1. $\begin{bmatrix}I_m & 0\\ X & I_n\end{bmatrix}\begin{bmatrix}A & B\\ C & D\end{bmatrix}=\begin{bmatrix}A & B\\ XA+C & XB+D\end{bmatrix}$.

2. 令 $A=I_{n-1},C=\begin{bmatrix}c_1\\ \vdots\\ c_{n-1}\end{bmatrix},B=\begin{bmatrix}b_1 & \cdots & b_{n-1}\end{bmatrix},D=[a],X=-B$ 则

$$\begin{bmatrix}I_{n-1} & 0\\ X & 1\end{bmatrix}\begin{bmatrix}A & C\\ B & D\end{bmatrix}=\begin{bmatrix}A & C\\ XA+B & XC+D\end{bmatrix}.$$

因为 $XA+B=0$, 因此矩阵可逆，当且仅当 $XC+D=a-\sum_{i=1}^{n-1}b_ic_i\neq 0$, 且

$$\begin{bmatrix}A & C\\ B & D\end{bmatrix}^{-1}\begin{bmatrix}I_{n-1} & 0\\ -X & 1\end{bmatrix}=\begin{bmatrix}A & C\\ 0 & a-\sum_{i=1}^{n-1}b_ic_i\end{bmatrix}^{-1}=\begin{bmatrix}I_{n-1} & -\left(a-\sum_{i=1}^{n-1}b_ic_i\right)^{-1}C\\ 0 & \left(a-\sum_{i=1}^{n-1}b_ic_i\right)^{-1}\end{bmatrix}$$

所以

$$\begin{bmatrix}A & C\\ B & D\end{bmatrix}^{-1}=\begin{bmatrix}1+\frac{c_1b_1}{a-\sum_{i=1}^{n-1}b_ic_i} & \frac{c_1b_1}{a-\sum_{i=1}^{n-1}b_ic_i} & & & -\frac{c_1}{a-\sum_{i=1}^{n-1}b_ic_i}\\ \frac{c_2b_1}{a-\sum_{i=1}^{n-1}b_ic_i} & 1+\frac{c_2b_2}{a-\sum_{i=1}^{n-1}b_ic_i} & & & -\frac{c_1}{a-\sum_{i=1}^{n-1}b_ic_i}\\ & & \ddots & & \vdots\\ & & & 1+\frac{c_{n-1}b_{n-1}}{a-\sum_{i=1}^{n-1}b_ic_i} & -\frac{c_{n-1}}{a-\sum_{i=1}^{n-1}b_ic_i}\\ -\frac{b_1}{a-\sum_{i=1}^{n-1}b_ic_i} & -\frac{b_2}{a-\sum_{i=1}^{n-1}b_ic_i} & \cdots & -\frac{b_{n-1}}{a-\sum_{i=1}^{n-1}b_ic_i} & \left(a-\sum_{i=1}^{n-1}b_ic_i\right)^{-1}\end{bmatrix}$$

□

练习 1.6.8. 利用 *Sherman-Morrison* 公式判断下列矩阵何时可逆，并在可逆时求其逆.

$$\begin{bmatrix}a_1+1 & 1 & 1 & \cdots & 1\\ 1 & a_2+1 & 1 & \cdots & 1\\ 1 & 1 & a_3+1 & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & 1\\ 1 & 1 & \cdots & 1 & a_n+1\end{bmatrix}.$$

证明:

$$\begin{bmatrix} a_1+1 & 1 & 1 & \cdots & 1 \\ 1 & a_2+1 & 1 & \cdots & 1 \\ 1 & 1 & a_3+1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 \\ 1 & 1 & \cdots & 1 & a_n+1 \end{bmatrix} = \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix} + \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}$$

若 $\begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix}$ 不可逆，则存在 $a_i = 0$, 那么我们对原矩阵作初等行变换：用其他行减去第 i 行，就得到

$$\begin{bmatrix} a_1 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & \cdots & & 1 \\ \vdots & \vdots & \ddots & \ddots & \\ \cdots & & \cdots & & a_n \end{bmatrix}$$

这个矩阵第 i 行全是 1, 第 $j(\neq i)$ 对角线元素是 a_j, 其它元素都是 0. 因此原矩阵可逆当且仅当对任意 $j \neq i, a_j \neq 0$.

若 $\begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix}$ 可逆，即 $a_1 a_2 \cdots a_n \neq 0$, 则由 Sherman-Morrison 公式，原矩阵可逆当且仅当

$$1 + \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = 1 + a_1^{-1} + \cdots + a_n^{-1} \neq 0,$$

并且原矩阵的逆为

$$\begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix}^{-1} + \frac{1}{1 + a_1^{-1} + \cdots + a_n^{-1}} \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix}^{-1}$$

即

$$\begin{bmatrix} a_1^{-1} & & & \\ & a_2^{-1} & & \\ & & \ddots & \\ & & & a_n^{-1} \end{bmatrix}^{-1} + \frac{1}{1+a_1^{-1}+\cdots+a_n^{-1}} \begin{bmatrix} a_1^{-1} \\ \vdots \\ a_n^{-1} \end{bmatrix} \begin{bmatrix} a_1^{-1} & \cdots & a_n^{-1} \end{bmatrix}.$$

□

练习 1.6.9.

1. 求一个矩阵 A 使得 $\begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix}$ 为反对称矩阵.

2. 求一个 5 阶置换矩阵 P，满足 $P^5 \neq I_5, P^6 = I_5$.

3. 求一个对称矩阵 A，使得不存在 B，满足 $A = BB^{\mathrm{T}}$.

证明: 1. $\begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix}^T = \begin{bmatrix} 0 & A^{\mathrm{T}} \\ A^{\mathrm{T}} & 0 \end{bmatrix}$, 因此只需使得 $A = -A^{\mathrm{T}}$, 即 A 是反对称矩阵即可。例如 $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

2. 令 $P = \begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & & & \\ & & 0 & 1 & 0 \\ & & 0 & 0 & 1 \\ & & 1 & 0 & 0 \end{bmatrix}$, $P^5 = \begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & & & \\ & & 0 & 0 & 1 \\ & & 1 & 0 & 0 \\ & & 0 & 1 & 0 \end{bmatrix}$, 但是 $P^6 = I_5$.

3. 最简单的可以是 $A = [-1]$. 再复杂一些可以是 $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. 我们用反证法证明不存在 B，满足 $A = BB^{\mathrm{T}}$；若存在，则一方面 $\boldsymbol{e}_1^{\mathrm{T}} A \boldsymbol{e}_1 = -1 < 0$, 另一方面 $\boldsymbol{e}_1^{\mathrm{T}} A \boldsymbol{e}_1 = (B^{\mathrm{T}}\boldsymbol{e}_1)^T \cdot (B^{\mathrm{T}}\boldsymbol{e}_1) \geq 0$, 矛盾.

□

练习 1.6.10. 给定 $m \times n, n \times m$ 矩阵 A, B，求证：$I_m + AB$ 可逆当且仅当 $I_n + BA$ 可逆.

证明: 考虑分块矩阵 $\begin{bmatrix} I_m & -A \\ B & I_n \end{bmatrix}$, 我们对它做初等行变换, 把左下角消掉, 得到 $\begin{bmatrix} I_m & -A \\ 0 & I_n + BA \end{bmatrix}$; 我们对它作初等行变换，把右上角消掉，得到 $\begin{bmatrix} I_m + AB & 0 \\ B & I_n \end{bmatrix}$. 因此

$$I_m + AB \text{ 可逆} \Leftrightarrow \begin{bmatrix} I_m + AB & 0 \\ B & I_n \end{bmatrix} \text{ 可逆} \Leftrightarrow \begin{bmatrix} I_m & -A \\ B & I_n \end{bmatrix} \text{ 可逆}$$

同时

$$I_n + BA \text{ 可逆} \Leftrightarrow \begin{bmatrix} I_m & -A \\ 0 & I_n + BA \end{bmatrix} \text{ 可逆} \Leftrightarrow \begin{bmatrix} I_m & -A \\ B & I_n \end{bmatrix} \text{ 可逆}$$

□

这样的构造性证明看起来花里胡哨，而且有点“灵机一动”。事实上，我们有不那么花哨的证明。

证明*:* 由对称性，我们不妨只证明: 若 I_m+AB 可逆则 I_n+BA 可逆. 我们只需要证明齐次方程组 $(I_n+BA)\boldsymbol{x}=0$ 没有非零解。若 $\boldsymbol{x}$ 是 $(I_n+BA)\boldsymbol{x}=0$ 的解，我们考虑 $\boldsymbol{y}=A\boldsymbol{x}$, 则

$$(I_m+AB)\boldsymbol{y}=A\boldsymbol{x}+ABA\boldsymbol{x}=A(I_n+BA)\boldsymbol{x}=0$$

而已知 I_m+AB 可逆，所以 $\boldsymbol{y}=0$, 即 $A\boldsymbol{x}=0$. 于是 $\boldsymbol{x}=-BA\boldsymbol{x}=0$ □

练习 1.6.11. 设实分块方阵 $X=\begin{bmatrix}A & C\\ 0 & B\end{bmatrix}$，其中 A 是方阵. 如果 X 与 X^{T} 可交换，求证：$C=0$.

证明*:* $X^{\mathrm{T}}=\begin{bmatrix}A^{\mathrm{T}} & 0\\ C^{\mathrm{T}} & B^{\mathrm{T}}\end{bmatrix}$. X 和 X^{T} 可交换，所以 $XX^{\mathrm{T}}=X^{\mathrm{T}}X$, 即

$$\begin{bmatrix}AA^{\mathrm{T}}+CC^{\mathrm{T}} & CB^{\mathrm{T}}\\ BC^{\mathrm{T}} & BB^{\mathrm{T}}\end{bmatrix}=\begin{bmatrix}A^{\mathrm{T}}A & A^{\mathrm{T}}C\\ C^{\mathrm{T}}A & B^{\mathrm{T}}B\end{bmatrix}$$

因为 $\operatorname{trace}(XX^{\mathrm{T}})=\operatorname{trace}(X^{\mathrm{T}}X)$, 因此 $\operatorname{trace}(AA^{\mathrm{T}}+CC^{\mathrm{T}})+\operatorname{trace}(BB^{\mathrm{T}})=\operatorname{trace}(A^{\mathrm{T}}A)+\operatorname{trace}(B^{\mathrm{T}}B)$, 因此 $\operatorname{trace}(CC^{T})=0$. 记 $C=[c_{ij}]_{m\times n}$, 则 CC^{T} 的第 i 个对角元等于 $\sum_{j=1}^{n}c_{ij}^2$, 于是 $\operatorname{trace}(CC^{\mathrm{T}})=\sum_{i=1}^{m}\sum_{j=1}^{n}c_{ij}^2$, 所以 $c_{ij}=0,\forall i,j$. □

练习 1.6.12. 设分块对角矩阵 $J=\begin{bmatrix}J_1 & & \\ & J_2 & \\ & & J_3\end{bmatrix}$，其中 $J_1=\begin{bmatrix}2 & 1\\ & 2\end{bmatrix}, J_2=\begin{bmatrix}2 & 1 & \\ & 2 & 1\\ & & 2\end{bmatrix}, J_3=\begin{bmatrix}1 & 1\\ & 1\end{bmatrix}$. 试找出一个五次多项式 $f(x)$，使得 $f(J)=0$.

证明*:* 我们考虑到 $(J_1-2I_2)^2=0,(J_2-2I_3)^3=0,(J_3-I_2)^2=0$, 所以我们可以令 $f(x)=(x-2)^3(x-1)^2$. □

练习 1.6.13. 构造 $2n$ 阶实方阵 A，满足 $A^2=-I_{2n}$.

可以从 $n=1$ 的时候尝试。

证明*:* 我们可以令 $A=\begin{bmatrix}0 & I_n\\ -I_n & 0\end{bmatrix}$ □

练习 1.6.14. 设 A,B 为两个左相抵的行简化阶梯形矩阵.

1. 证明：$A\boldsymbol{x}=0$ 与 $B\boldsymbol{x}=0$ 同解.

2. 证明：如果 A 的最后一列不是主列，则存在 $\boldsymbol{x}$ 使得 $A\begin{bmatrix}\boldsymbol{x}\\ 1\end{bmatrix}=0$.

3. 证明：如果 A 最后一列是主列，则对任意 $\boldsymbol{x}$，$A\begin{bmatrix}\boldsymbol{x}\\ 1\end{bmatrix}\neq 0$.

4. 设 A,B 除了最后一列之外，其他列均相等，且 A 的最后一列不是主列，证明 $A=B$.

5. 设 A, B 除了最后一列之外，其他列均相等，且 A 的最后一列是主列，证明 $A = B$.
6. 用数学归纳法证明，左相抵的行简化阶梯形矩阵必然相等. 由此证明，一个矩阵 A 的行简化阶梯形矩阵唯一.

证明: A, B 为左相抵，则存在可逆矩阵 P, 使得 $B = PA$.

1. 若 $A\boldsymbol{x} = 0$，则 $PA\boldsymbol{x} = 0$, 即 $B\boldsymbol{x} = 0$。若 $B\boldsymbol{x} = 0$，则 $P^{-1}B\boldsymbol{x} = 0$, 即 $A\boldsymbol{x} = 0$. 所以 $A\boldsymbol{x} = 0$ 与 $B\boldsymbol{x} = 0$ 同解.
2. 我们把 A 写成分块矩阵 $\begin{bmatrix} A_1 & \boldsymbol{a} \end{bmatrix}$. A 的最后一列不是主列，根据定理 1.3.8，方程组 $A_1\boldsymbol{x} = -\boldsymbol{a}$ 有解，记为 $\boldsymbol{x}_0$; 此时 $A\begin{bmatrix} \boldsymbol{x}_0 \\ 1 \end{bmatrix} = A_1\boldsymbol{x}_0 + \boldsymbol{a} = 0$, 因此方程组 $A\begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} = 0$ 有解。
3. 如果 A 最后一列是主列，假设最后一个主列在第 i 行，那么第 i 行除了最后一个元素以外全等于零，于是 $A\begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} = 0$ 的第 i 个方程为 $1 = 0$, 从而该方程组无解，因此对任意 $\boldsymbol{x}$, $A\begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} \neq 0$.
4. 记 $A = \begin{bmatrix} A_1 & \boldsymbol{a} \end{bmatrix}, B = \begin{bmatrix} B_1 & \boldsymbol{b} \end{bmatrix}$. 由题意，$A_1 = B_1$. 由于 A 的最后一列不是主列，因此存在 $\boldsymbol{x}$，使得 $A\begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} = 0$; 那么由第 1 小题，我们知 $\begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix}$ 也是方程组 $B\boldsymbol{y} = 0$ 的解，因此 $B\begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} = 0$. 展开就有 $\boldsymbol{b} = -B_1\boldsymbol{x} = -A_1\boldsymbol{x} = \boldsymbol{a}$, 所以 $A = B$.
5. 记号同第 4 小题。我们只需证明 B 的最后一列也是主列。因为 $A_1 = B_1$ 本身也是行简化阶梯形矩阵，设非零行个数为 r. 当 A, B 的最后一行都是主列时，这个主列必须都等于 $\boldsymbol{e}_{r+1}$, 因此相等。由第 3 小题，对任意 $\boldsymbol{x}$，$A\begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} \neq 0$. 由于 P 可逆，对任意 $\boldsymbol{x}$，$PA\begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} \neq 0$, 即对任意 $\boldsymbol{x}$，$B\begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} \neq 0$. 由第 2 小题知，$B$ 的最后一列是主列。
6. 我们对矩阵的列数做归纳。若行简化阶梯形矩阵只有一列，那么要么它是零矩阵，那么等予以 $\begin{bmatrix} 1 \\ 0 \vdots \\ 0 \end{bmatrix}$. 假设命题对有 n 列的行简化阶梯型矩阵成立，现在假设 $A = \begin{bmatrix} A_1 & \boldsymbol{a} \end{bmatrix}, B = \begin{bmatrix} B_1 & \boldsymbol{b} \end{bmatrix}$ 为两个左相抵的有 $n+1$ 列的行简化阶梯形矩阵。那么 A_1, B_1 都是有 n 列的行简化阶梯形矩阵。由于 $B = PA = \begin{bmatrix} PA_1 & P\boldsymbol{a} \end{bmatrix}$, 我们有 $B_1 = PA_1$，即 A_1, B_1 左相抵，由归纳假设 $A_1 = B_1$. 进一步地，再由第 4、第 5 小题知 $A = B$. 由此完成了归纳。现在考虑 A 的两个行简化阶梯形矩阵 R_1, R_2, 那么他们都跟 A 左相抵，因此 R_1, R_2 也左相抵，从而 $R_1 = R_2$, 这就证明了 A 的行简化阶梯形矩阵唯一。

□

练习 1.6.15 (置换的不动点)**.** 设 P 是置换矩阵. 如果 P 对应的行变换保持第 i 行不变，则称 i 为 P 的不动点.

1. 证明 i 是 P 的不动点当且仅当 P 的第 i 个对角元为 1.

2. 证明 P 的不动点个数等于 $\operatorname{trace}(P)$.

3. 对任意置换矩阵 P_1, P_2，证明 P_1P_2 和 P_2P_1 的不动点个数相等.

证明: 1. 若 i 是 P 的不动点，则 $P\boldsymbol{e}_i = \boldsymbol{e}_i$. 而 P 的第 i 个对角元为 $\boldsymbol{e}_i^{\mathrm{T}} P \boldsymbol{e}_i = 1$。反之 P 的第 i 个对角元为 1，因为 P 是置换矩阵，它的第 i 行、第 i 列其他元素均为 0。对任意列向量 $\boldsymbol{x} = \begin{bmatrix} \vdots \\ x_i \\ \vdots \end{bmatrix}$, $P\boldsymbol{x}$ 的第 i 个分量是 P 的第 i 行和 $\boldsymbol{x}$ 的乘积，因此 $P\boldsymbol{x} = \begin{bmatrix} \vdots \\ x_i \\ \vdots \end{bmatrix}$, 从而对任意矩阵，左乘 P 都保持第 i 行不变，因此 i 是 P 的不动点。

2. 由于 P 的每个元素要么是 0, 要么是 1, 所以 P 的不动点个数等于对角元上 1 的个数；而其他对角元都是 0, 因此不动点个数等于 $\operatorname{trace}(P)$.

3. 对任意置换矩阵 P_1, P_2，P_1P_2 和 P_2P_1 也都是置换矩阵，因此它们的不动点个数等于它们各自的迹；而 $\operatorname{trace}(P_1P_2) = \operatorname{trace}(P_2P_1)$, 因此它们不动点个数相等。

□

练习 1.6.16. 考虑 $\mathbb{R}^n$ 上的变换 $f(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$，具有该形式的变换称为**仿射变换** 仿射变换通常不是线性变换.

1. 证明 $\begin{bmatrix} A & \boldsymbol{b} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} = \begin{bmatrix} f(\boldsymbol{x}) \\ 1 \end{bmatrix}$.

2. 对仿射变换 $f(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$, 记 $M_f := \begin{bmatrix} A & \boldsymbol{b} \\ 0 & 1 \end{bmatrix}$. 证明对任意仿射变换 $f(\boldsymbol{x}), g(\boldsymbol{x})$, $M_f M_g = M_{f\circ g}$.

3. 证明当仿射变换 f 可逆时，矩阵 M_f 可逆，且 $M_{f^{-1}} = (M_f)^{-1}$.

证明: 1. 由分块矩阵的乘法

$$\begin{bmatrix} A & \boldsymbol{b} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} = \begin{bmatrix} A\boldsymbol{x} + \boldsymbol{b} \\ 1 \end{bmatrix} = \begin{bmatrix} f(\boldsymbol{x}) \\ 1 \end{bmatrix}.$$

2. 设 $g(\boldsymbol{x}) = B\boldsymbol{x} + \boldsymbol{b}'$, 那么 $(f \circ g)(\boldsymbol{x}) = f(B\boldsymbol{x} + \boldsymbol{b}') = A(B\boldsymbol{x} + \boldsymbol{b}') + \boldsymbol{b} = AB\boldsymbol{x} + (A\boldsymbol{b}' + \boldsymbol{b})$, 因此 $M_{f\circ g} = \begin{bmatrix} AB & A\boldsymbol{b}' + \boldsymbol{b} \\ 0 & 1 \end{bmatrix}$, 而

$$M_f M_g = \begin{bmatrix} A & \boldsymbol{b} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} B & \boldsymbol{b}' \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} AB & A\boldsymbol{b}' + \boldsymbol{b} \\ 0 & 1 \end{bmatrix} = M_{f\circ g}.$$

□

练习 1.6.17. 图 1.4中有四个弹簧振子. 弹簧的初始长度分别为 l_1, l_2, l_3, l_4，振子在稳态的最终长度为 $l_1 + x_1, l_2 + x_2, l_3 + x_3, l_4 + x_4$.

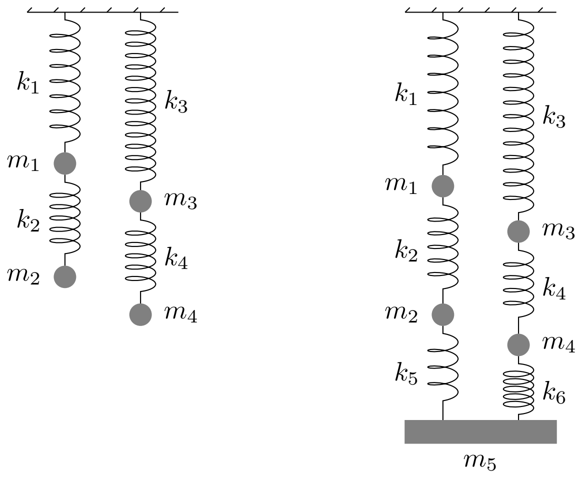


图 1.4

1. 找到矩阵 A 使得 $A\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=g\begin{bmatrix}m_1\\m_2\\m_3\\m_4\end{bmatrix}$. 此时 A 的分块结构有什么特点? 你能否从系统中看出原因?

2. 图 *1.4*中振子 m_2, m_4 下面，由劲度系数分别为 k_5, k_6 的弹簧挂住了一个质量为 m_5 的振子. 找到矩阵 A 使得 $A\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix}=g\begin{bmatrix}m_1\\m_2\\m_3\\m_4\\m_5\end{bmatrix}$. 此时 A 的分块结构是否还有之前的特点? 为什么?

证明: 1. 根据胡克定律, $k_1x_1=g(m_1+m_2), k_2x_2=gm_2, k_3x_3=g(m_3+m_4), k_4x_4=gm_4$, 因此

$$\begin{bmatrix}k_1 & -k_2 & 0 & 0\\0 & k_2 & 0 & 0\\0 & 0 & k_3 & -k_4\\0 & 0 & 0 & k_4\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=g\begin{bmatrix}m_1\\m_2\\m_3\\m_4\end{bmatrix}.$$

因此 $A=\begin{bmatrix}k_1 & -k_2 & 0 & 0\\0 & k_2 & 0 & 0\\0 & 0 & k_3 & -k_4\\0 & 0 & 0 & k_4\end{bmatrix}$, 它是分块上三角的。

2. 根据胡克定律，$k_5x_5+k_6x_6=gm_5, k_2x_2=gm_2+k_5x_5, k_1x_1=gm_1+k_2x_2, k_4x_4=gm_4+k_6x_6, k_3x_3=gm_3+k_4x_4$, 以及 $x_1+x_2+x_5=x_3+x_4+x_6$, 消去 x_6 并整理一

下有

$$\begin{cases} k_1x_1 & -k_2x_2 & & & & = gm_1 \\ & k_2x_2 & & & -k_5x_5 & = gm_2 \\ & & k_3x_3 & -k_4x_4 & & = gm_3 \\ -k_6x_1 & -k_6x_2 & +k_6x_3 & +(k_6+k_4)x_4 & -k_6x_5 & = gm_4 \\ k_6x_1 & +k_6x_2 & -k_3x_3 & -k_4x_4 & +(k_5+k_6)x_5 & = gm_5 \end{cases}$$

因此

$$A = \begin{bmatrix} k_1 & -k_2 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 & -k_5 \\ 0 & 0 & k_3 & -k_4 & 0 \\ -k_6 & -k_6 & k_6 & k_6+k_4 & -k_6 \\ k_6 & k_6 & -k_3 & -k_4 & k_5+k_6 \end{bmatrix}$$

A 不再是分块上三角矩阵，因为两侧的弹簧并联了。

□

练习 1.6.18. 考虑一个元素皆为复数的矩阵. 将每个元素分成实部和虚部，不难发现对任意复数矩阵，都存在实数矩阵 A, B，使得该复数矩阵为 $A + \mathrm{i}B$，其中 i 是虚数单位. 同理，任意复数向量，都存在实数向量 $\boldsymbol{v}, \boldsymbol{w}$，使得该复数向量为 $\boldsymbol{v} + \mathrm{i}\boldsymbol{w}$. 下面用 $\mathrm{Re}(A), \mathrm{Re}(\boldsymbol{v}), \mathrm{Im}(A), \mathrm{Im}(\boldsymbol{v})$ 来表达复矩阵 A 和复向量 $\boldsymbol{v}$ 的实部和虚部.

1. 用实数矩阵和向量 $A, B, \boldsymbol{v}, \boldsymbol{w}$ 来表达 $(A + \mathrm{i}B)(\boldsymbol{v} + \mathrm{i}\boldsymbol{w})$ 的实部和虚部;

2. 对任意实数矩阵 A, B，求矩阵 X，使得 $X \begin{bmatrix} \mathrm{Re}(\boldsymbol{v}+\mathrm{i}\boldsymbol{w}) \\ \mathrm{Im}(\boldsymbol{v}+\mathrm{i}\boldsymbol{w}) \end{bmatrix} = \begin{bmatrix} \mathrm{Re}\left((A+\mathrm{i}B)(\boldsymbol{v}+\mathrm{i}\boldsymbol{w})\right) \\ \mathrm{Im}\left((A+\mathrm{i}B)(\boldsymbol{v}+\mathrm{i}\boldsymbol{w})\right) \end{bmatrix}$ 对任意实数向量 $\boldsymbol{v}, \boldsymbol{w}$ 都成立.

3. 考虑映射 $f(a + \mathrm{i}b) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. 验证 $f\left((a+\mathrm{i}b)(c+\mathrm{i}d)\right) = f(a+\mathrm{i}b)f(c+\mathrm{i}d)$.

证明: 1. $(A + \mathrm{i}B)(\boldsymbol{v} + \mathrm{i}\boldsymbol{w}) = (A\boldsymbol{v} - B\boldsymbol{w}) + \mathrm{i}(A\boldsymbol{w} + B\boldsymbol{v})$, 所以

$$\mathrm{Re}((A+\mathrm{i}B)(\boldsymbol{v}+\mathrm{i}\boldsymbol{w})) = A\boldsymbol{v} - B\boldsymbol{w}, \mathrm{Im}((A+\mathrm{i}B)(\boldsymbol{v}+\mathrm{i}\boldsymbol{w})) = A\boldsymbol{w} + A\boldsymbol{v}$$

2. 由题意，我们需要求矩阵 X 满足 $X \begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{w} \end{bmatrix} = \begin{bmatrix} A\boldsymbol{v} - B\boldsymbol{w} \\ A\boldsymbol{a} + B\boldsymbol{v} \end{bmatrix} = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$. 由于它对任意 $\boldsymbol{v}, \boldsymbol{w}$ 都成立，因此 $X = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$.

3. $(a+\mathrm{i}b)(c+\mathrm{i}d) = (ac-bd)+\mathrm{i}(bc+ad)$, 因此 $f\left((a+\mathrm{i}b)(c+\mathrm{i}d)\right) = \begin{bmatrix} ac-bd & -(bc+ad) \\ bc+ad & ac-bd \end{bmatrix}$, 而

$$f(a+\mathrm{i}b)f(c+\mathrm{i}d) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} ac-bd & -ad-bc \\ bc+ad & -bd+ac \end{bmatrix}$$

□

练习 1.6.19 (错位分块对角矩阵)**.** 定义

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \triangle \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & a_{12} & 0 \\ 0 & b_{11} & 0 & b_{12} \\ a_{21} & 0 & a_{22} & 0 \\ 0 & b_{21} & 0 & b_{22} \end{bmatrix}.$$

1. 证明 $(A_1\triangle B_1)(A_2\triangle B_2) = (A_1A_2)\triangle(B_1B_2)$.

2. 证明 $A\triangle B$ 可逆当且仅当 A, B 都可逆；此时有 $(A\triangle B)^{-1} = A^{-1}\triangle B^{-1}$.

3. 求 X，使得对任意二阶方阵 A, B，都有 $X(A\triangle B)X^{-1} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$.

证明*:* 令 $P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ 则 $A\triangle B = P\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}P$

1.

$$\begin{aligned} (A_1\triangle B_1)(A_2\triangle B_2) &= P\begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix}PP\begin{bmatrix} A_2 & 0 \\ 0 & B_2 \end{bmatrix}P \\ &= P\begin{bmatrix} A_1A_2 & 0 \\ 0 & B_1B_2 \end{bmatrix}P \\ &= (A_1A_2)\triangle(B_1B_2). \end{aligned}$$

2. 证明 $A\triangle B$ 可逆，即 $P\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}P$ 可逆，当且仅当 A, B 都可逆；此时

$$(A\triangle B)^{-1} = P\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^{-1}P = P\begin{bmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{bmatrix}P = \left(A^{-1}\right)\triangle\left(B^{-1}\right).$$

3. 显然我们已经知道了，我们可以取 $X = P$.

□

1.7 LU 分解

练习 1.7.1. 求下列矩阵的 LU 分解，其中 L 为单位下三角矩阵：

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

2. $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{bmatrix}$.

3. $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 5 \end{bmatrix}$.

4. $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 6 \end{bmatrix}$.

5. $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}$.

6. $\begin{bmatrix} 1 & 5 & 5 & 5 \\ 1 & 2 & 6 & 6 \\ 1 & 2 & 3 & 7 \\ 1 & 2 & 3 & 4 \end{bmatrix}$.

注意观察，哪些 0 保留在了 L 中或者 U 中.

这类问题的做法是：对 $\begin{bmatrix} A & I_n \end{bmatrix}$ 做初等行变换，使得 A 变成上三角矩阵，那么整个矩阵就变成了 $\begin{bmatrix} U & L \end{bmatrix}$ 的形式，但是我们要注意，A 的 LU 分解并不是 $A = LU$, 而是 $A = L^{-1}U$. 下面我们只写答案。

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

2.
$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

3.
$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

4.
$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

5.
$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

6.
$$\begin{bmatrix} 1 & 5 & 5 & 5 \\ 1 & 2 & 6 & 6 \\ 1 & 2 & 3 & 7 \\ 1 & 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 5 & 5 & 5 \\ 0 & -3 & 1 & 1 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & -3 \end{bmatrix}.$$

□

练习 1.7.2. 利用行变换解下列线性方程组：

1.
$$\begin{bmatrix} 1 & 5 & 5 & 5 \\ 1 & 2 & 6 & 6 \\ 1 & 2 & 3 & 7 \\ 1 & 2 & 3 & 4 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 34 \\ 28 \\ 23 \\ 20 \end{bmatrix}.$$

2. $\begin{bmatrix}1&0&0&0\\1&1&0&0\\1&1&1&0\\1&1&1&1\end{bmatrix}\boldsymbol{y}=\begin{bmatrix}34\\28\\23\\20\end{bmatrix}$.

3. * $\begin{bmatrix}1&5&5&5\\0&-3&1&1\\0&0&-3&1\\0&0&0&-3\end{bmatrix}\boldsymbol{x}=\boldsymbol{y}$，这里的 $\boldsymbol{y}$ 是上一问的解.

我们稍微观察一下，就知道这一题是和上一题连贯的；而且如果记住计算过程的话，我们就知道

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

证明*:* 2. 第四行减第三行，第三行减第二行，第二行减第一行，很容易就得到

$$\boldsymbol{y}=\begin{bmatrix}34\\AA^{\mathrm{T}}+CC^{\mathrm{T}}-6\\-5\\-3\end{bmatrix}$$

3. 我们对增广矩阵作初等行变换

$$\left[\begin{array}{cccc|c}1&5&5&5&34\\0&-3&1&1&-6\\0&0&-3&1&-5\\0&0&0&-3&-3\end{array}\right]\longrightarrow\left[\begin{array}{cccc|c}1&5&5&0&29\\0&-3&1&0&-7\\0&0&-3&0&-6\\0&0&0&1&1\end{array}\right]\longrightarrow\left[\begin{array}{cccc|c}1&5&0&0&19\\0&-3&0&0&-9\\0&0&1&0&2\\0&0&0&1&1\end{array}\right]$$

$$\longrightarrow\left[\begin{array}{cccc|c}1&0&0&0&4\\0&1&0&0&3\\0&0&1&0&2\\0&0&0&1&1\end{array}\right]$$

因此方程组的解为 $\boldsymbol{x}=\begin{bmatrix}4\\3\\2\\1\end{bmatrix}$

1. 根据上一大题的第 6 小题，我们知道方程组的为 $\boldsymbol{x} = \begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$

□

练习 1.7.3. 设 n 阶方阵 $A = \begin{bmatrix} a_{11} & \boldsymbol{w}^{\mathrm{T}} \\ \boldsymbol{v} & B \end{bmatrix}$ 存在 LU 分解 $A = LU$，其中 $\boldsymbol{v} \neq 0$.

1. A 的左上角元素 a_{11} 是否非零?

2. 从上往下对 A 做行变换，将 $\boldsymbol{v}$ 变为 0 时，需要做多少次加法、乘法?

3. 再设 A 为对称矩阵. 此时 $\boldsymbol{v}$ 和 $\boldsymbol{w}$ 有什么关系? B 有什么性质?

4. 仍令 A 对称. 利用行变换将 $\boldsymbol{v}$ 变为 0 时，再利用对应的列变换将 $\boldsymbol{w}^{\mathrm{T}}$ 变为 0，求证此时 B 变为对称矩阵. 此时需要做几次加法、乘法?

由此看到，对称矩阵 LU 分解的计算量可以节省一半.

证明: 1. 因为 A 有 LU 分解等价于 A 的顺序子矩阵都可逆，而 a_{11} 是 A 的第一个顺序子矩阵，因此 $a_{11} \neq 0$.

2. 我们要计算的就是 $B - \frac{1}{a_{11}}\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}$ 的加法、乘法的次数，加法是 $(n-1)^2$ 次，乘法是 $n(n-1)$。

3. A 为对称矩阵，则 $\boldsymbol{v} = \boldsymbol{w}, B = B^T$.

4. 此时我们对 A 作的初等行变换和列变换用矩阵写出来就是

$$A \longmapsto \begin{bmatrix} 1 & 0 \\ -\frac{1}{a_{11}}\boldsymbol{v} & I_{n-1} \end{bmatrix} A \begin{bmatrix} 1 & -\frac{1}{a_11}\boldsymbol{w}^T \\ 0 & I_{n-1} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ 0 & B - \frac{1}{a_{11}}\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}} \end{bmatrix}.$$

A 对称，则 $\boldsymbol{v} = \boldsymbol{w}$，于是

$$\left(B - \frac{1}{a_{11}}\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}\right)^{\mathrm{T}} = B^{\mathrm{T}} - \frac{1}{a_{11}}\boldsymbol{w}\boldsymbol{v}^{\mathrm{T}} = B - \frac{1}{a_{11}}\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}$$

因此将 B 变成了对称矩阵. 根据对称性，我们不需要计算整个矩阵，只需要计算上三角或者下三角的部分，因此加法继续演了 $\frac{n(n-1)}{2}$, 乘法计算了 $\frac{n(n-1)}{2} + (n-1)$ 次，因此一共计算了 $n^2 - 1$ 次。那么整个 LU 分解用了 $\sum_{i=1}^{n}(i^2 - 1) = \frac{(n-1)n(n+1)}{3}$ 次。

□

练习 1.7.4. 设 n 阶方阵

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

利用初等变换证明，存在分解式 $T = LU$，其中 L 是下三角矩阵，U 是上三角矩阵. 据此求出 T^{-1}.

证明: 将 T 的第一行加到第二行，得到 T_1，把 T_1 的第二行加到第三行，得到 T_2, 把 T_2 的第三行行加到第四行，得到 T_3, 以此类推，我们得到

$$\begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 1 & 1 & & \\ \vdots & & \ddots & \ddots & \\ 1 & \cdots & & 1 & 1 \end{bmatrix} T = \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & 1 & \ddots & \\ & & & \ddots & -1 \\ & & & & 1 \end{bmatrix}.$$

所以 T 的 LU 分解为 $T = \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ 0 & -1 & 1 & & \\ & & \ddots & \ddots & \\ 0 & & & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & 1 & \ddots & \\ & & & \ddots & -1 \\ & & & & 1 \end{bmatrix}$,

$$T^{-1} = \begin{bmatrix} 1 & 1 & & & 1 \\ & 1 & 1 & & 1 \\ & & 1 & \ddots & \vdots \\ & & & \ddots & 1 \\ & & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 1 & 1 & & \\ \vdots & & \ddots & \ddots & \\ 1 & \cdots & & 1 & 1 \end{bmatrix}$$

□

第二章　子空间和维数

2.1　基 本 概 念

练习 2.1.1. 判断下列子集是否为 $\mathbb{R}^n$ 的子空间，其中的子空间是否可以写成由某个向量组线性生成的子空间；如果可以，写出一组基：

1. $\left\{\begin{bmatrix}a_1 & a_2 & \cdots & a_n\end{bmatrix}^{\mathrm{T}} \middle| \sum_{i=1}^{n} a_i = 0\right\}$.　*2.* $\left\{\begin{bmatrix}a_1 & a_2 & \cdots & a_n\end{bmatrix}^{\mathrm{T}} \middle| \sum_{i=1}^{n} a_i = 1\right\}$.

3. $\left\{\begin{bmatrix}a_1 & a_2 & \cdots & a_n\end{bmatrix}^{\mathrm{T}} \middle| \sum_{i=1}^{n} a_i^2 = 0\right\}$.　*4.* $\left\{\begin{bmatrix}a_1 & a_2 & \cdots & a_n\end{bmatrix}^{\mathrm{T}} \middle| a_1 = a_2 = 0\right\}$.

5. $\left\{\begin{bmatrix}a_1 & a_2 & \cdots & a_n\end{bmatrix}^{\mathrm{T}} \middle| a_1 \geq 0\right\}$.　*6.* $\left\{\begin{bmatrix}a_1 & a_2 & \cdots & a_n\end{bmatrix}^{\mathrm{T}} \middle| a_1 = a_2 = \cdots = a_n\right\}$.

7. $\left\{\begin{bmatrix}a_1 & a_2 & \cdots & a_n\end{bmatrix}^{\mathrm{T}} \middle| a_1^2 = a_2^2 = \cdots = a_n^2\right\}$.　*8.* $\left\{\begin{bmatrix}a_1 & a_2 & a_3\end{bmatrix}^{\mathrm{T}} \middle| a_1 a_2 a_3 = 0\right\}$.

9. $\left\{\begin{bmatrix}a_1 & a_2 & a_3\end{bmatrix}^{\mathrm{T}} \middle| a_1 = a_2\right\}$.　*10.* $\left\{\begin{bmatrix}a_1 & a_2 & a_3\end{bmatrix}^{\mathrm{T}} \middle| a_1 \leq a_2 \leq a_3\right\}$.

11. $\left\{\begin{bmatrix}a_1 & a_2 & a_3 & a_4\end{bmatrix}^{\mathrm{T}} \middle| \begin{bmatrix}a_1 & a_2 \\ a_3 & a_4\end{bmatrix}\text{是可逆矩阵}\right\}$.

13. $\left\{\begin{bmatrix}a_1 & a_2 & a_3 & a_4\end{bmatrix}^{\mathrm{T}} \middle| \begin{bmatrix}a_1 & a_2 \\ a_3 & a_4\end{bmatrix}\text{是不可逆矩阵}\right\}$.

15. $\left\{\begin{bmatrix}a_1 & a_2 & a_3 & a_4\end{bmatrix}^{\mathrm{T}} \middle| \begin{bmatrix}a_1 & a_2 \\ a_3 & a_4\end{bmatrix}\text{是对称矩阵}\right\}$.

17. $\left\{\begin{bmatrix}a_1 & a_2 & a_3 & a_4\end{bmatrix}^{\mathrm{T}} \middle| \begin{bmatrix}a_1 & a_2 \\ a_3 & a_4\end{bmatrix}\text{是反对称矩阵}\right\}$.

证明*:*　1.　是线性空间，它的一组基是 $\{\boldsymbol{e}_1 - \boldsymbol{e}_j \mid j = 2, 3, \cdots n\}$.

2.　不是线性空间，对加法和数乘都不封闭。

3.　是线性空间，因为它只包含零向量。

4.　是线性空间，它的一组基是 $\{\boldsymbol{e}_i : i = 3, 4, \cdots n\}$

5. 不是线性空间，对数乘不封闭。
6. 是线性空间，它的一组基是 $\left\{\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^{\mathrm{T}}\right\}$.
7. 不是线性空间，对加法不封闭。.
8. 不是线性空间，对加法不封闭。
9. 是线性空间，它的一组基是 $\{\boldsymbol{e}_1+\boldsymbol{e}_2, \boldsymbol{e}_3\}$.
10. 不是线性空间，对数乘不封闭。
11. 不是线性空间，它不包含零矩阵，对加法也不封闭。
12. 不是线性空间，对加法不封闭
13. 是线性空间，它的一组基是 $\left\{\begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 1 & 1 & 0\end{bmatrix}, \begin{bmatrix}0 & 0 & 0 & 1\end{bmatrix}\right\}$
14. 是线性空间，它的一组基是 $\left\{\begin{bmatrix}0 & 1 & -1 & 0\end{bmatrix}\right\}$.

□

练习 2.1.2. *判断满足下列性质的 $\mathbb{R}^3$ 子集 $\mathcal{M}$ 是否存在. 若存在，进一步判断哪些是子空间：*

1. $\mathcal{M}$ *含有* $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$*；对任意* $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{M}$*，都有* $\boldsymbol{v}+\boldsymbol{w} \in \mathcal{M}$*；存在* $\boldsymbol{v} \in \mathcal{M}$*，使得* $\dfrac{1}{2}\boldsymbol{v} \notin \mathcal{M}$.

2. $\mathcal{M}$ *含有所有形如* $\begin{bmatrix}\cos\theta \\ \sin\theta \\ 1\end{bmatrix}$ *的向量；对任意* $\boldsymbol{v} \in \mathcal{M}$*，都有* $k\boldsymbol{v} \in \mathcal{M}$*；存在* $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{M}$*，使得* $\boldsymbol{v}+\boldsymbol{w} \notin \mathcal{M}$.

3. $\mathcal{M}$ *含有* $\boldsymbol{e}_1, \boldsymbol{e}_2$ *但不含有* $\boldsymbol{e}_3$*；对任意* $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{M}, k \geq 0$*，都有* $\boldsymbol{v}+\boldsymbol{w}, k\boldsymbol{v} \in \mathcal{M}$.

证明*:* 1. 存在，例如 $\mathcal{M} := \{k_1\boldsymbol{e}_1 + k_2\boldsymbol{e}_2 + k_3\boldsymbol{e}_3 : k_1, k_2, k_3 \in \mathbb{Z}\}$, $\frac{1}{2}\boldsymbol{e}_1 \notin \mathcal{M}$. 它不是子空间。

2. 存在。例如 $\mathcal{M} = \left\{\begin{bmatrix}x_1 \\ x_2 \\ x_3\end{bmatrix} : x_3^2 = x_1^2 + x_2^2\right\}$, $\boldsymbol{v} = \begin{bmatrix}1 \\ 0 \\ 1\end{bmatrix}$, $\boldsymbol{w} = \begin{bmatrix}-1 \\ 0 \\ 1\end{bmatrix}$, 此时 $\boldsymbol{v}+\boldsymbol{w} = \begin{bmatrix}0 \\ 0 \\ 2\end{bmatrix} \notin \mathcal{M}$.

3. 存在，例如 $\mathcal{M} = \left\{\begin{bmatrix}x_1 \\ x_2 \\ 0\end{bmatrix} : x_1 \geq 0, x_2 \geq 0\right\}$.

□

练习 2.1.3. 证明命题 *2.1.9.*

证明*:* 1. • $\operatorname{Span}(S)$ 对加法封闭：取 $a_1\alpha_1 + \cdots + a_n\alpha_n, b_1\beta_n + \cdots + b_m\beta_m$, 其中 $a_i \in \mathbb{R}(1 \leq i \leq n), b_j \in \mathbb{R}(1 \leq j \leq m), \alpha_i \in S(1 \leq i \leq n), \beta_j \in S(1 \leq j \leq m)$, 则按照定义，

$$a_1\alpha_1 + \cdots + a_n\alpha_n + b_1\beta_n + \cdots + b_m\beta_m \in \operatorname{Span}(S).$$

• $\operatorname{Span}(S)$ 对数乘封闭：取 $a_1\alpha_1 + \cdots + a_n\alpha_n$, 其中 $a_i \in \mathbb{R}, \alpha_i \in S(1 \leq i \leq n)$, 取 $k \in \mathbb{R}$, 则

$$k(a_1\alpha_1 + \cdots + a_n\alpha_n) = ka_1\alpha_1 + \cdots + ka_n\alpha_n \in \operatorname{Span}(S).$$

因此 $\mathrm{Span}(S)$ 是线性子空间。

2. 设 V 是线性子空间，且 $S \subseteq V$, 我们需要证明 $\mathrm{Span}(S) \subset V$. 我们任取 $a_1\alpha_1 + \cdots + a_n\alpha_n \in \mathrm{Span}(S)$, 那么因为 $\alpha_i \in S \subset V$, 而 V 是子空间，所以 V 对加法和数乘都封闭，因此 $a_i\alpha_i \in V$ 且 $\sum_{i=1}^n a_i\alpha_i \in V$. 这就是说，$\mathrm{Span}(S)$ 中任意元素都属于 V, 即 $\mathrm{Span}(S) \subset V$.

□

练习 2.1.4. 证明，$\mathcal{R}(0_{m\times n}) = \{0_m\}, \mathcal{N}(0_{m\times n}) = \mathbb{R}^n$；如果 n 阶方阵 A 可逆，则 $\mathcal{R}(A) = \mathbb{R}^n$, $\mathcal{N}(A) = \{0\}$.

证明: 对任意 $\boldsymbol{x} \in \mathbb{R}^n$，$0_{m\times n}\boldsymbol{x} = 0_m$, 所以 $\mathcal{R}(0_{m\times n}) = \{0_m\}, \mathcal{N}0_{m\times n} = \mathbb{R}^n$.

若 A 可逆，对任意 $\boldsymbol{y} \in \mathbb{R}^n$, 我们有 $A(A^{-1}\boldsymbol{y}) = \boldsymbol{y}$, 因此 $\mathcal{R}(A) = \mathbb{R}^n$. 若 $A\boldsymbol{x} = 0$, 等式两边同时左乘 A^{-1}，就有 $\boldsymbol{x} = 0$, 因此 $\mathcal{N}(A) = \{0\}$. □

练习 2.1.5. 证明，对例 *2.1.3* 中的 $A = \begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3\end{bmatrix}$，这三个向量中的任意两个都可以作为 $\mathcal{R}(A)$ 的一组生成向量，但是，其中任意单个向量都不能生成 $\mathcal{R}(A)$.

证明: 我们以证明 $\mathcal{R}(A) = \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2)$ 为例。我们任取 $\boldsymbol{v} \in \mathcal{R}(A)$, 则 $\boldsymbol{v}$ 可以写成 $\boldsymbol{v} = k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2 + k_3\boldsymbol{a}_3$. 注意到 $\boldsymbol{a}_3 = \boldsymbol{a}_1 + \boldsymbol{a}_2$, 那么 $\boldsymbol{v} = (k_1 + k_3)\boldsymbol{a}_1 + (k_2 + k_3)\boldsymbol{a}_2 \in \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2)$. 这就证完了。剩下两种情形同理可得。

我们再以证明 $\mathcal{R}(A) \neq \mathrm{Span}(\boldsymbol{a}_1)$ 为例。这就更简单了：因为 $\boldsymbol{a}_2 \in \mathcal{R}(A)$ 但是 $\boldsymbol{a}_2 \notin \mathrm{Span}(\boldsymbol{a}_1)$. 其他情形同理可得。

□

练习 2.1.6. 判断下列 A, B 是否具有相同的列空间、零空间，证明或举出反例：

1. A 为任意矩阵，B 分别为 $2A, \begin{bmatrix}A & A\end{bmatrix}, \begin{bmatrix}A & AC\end{bmatrix}, \begin{bmatrix}A\\A\end{bmatrix}, \begin{bmatrix}A\\CA\end{bmatrix}, \begin{bmatrix}A & A\\0 & A\end{bmatrix}, PA, AQ$，其中 P, Q 可逆.

2. A 为 n 阶方阵，B 分别为 $A + I_n, A^2, A^{\mathrm{T}}$.

证明: 1.
- $2A$ 和 A 有相同的列空间和零空间，证明略。
- $\begin{bmatrix}A & A\end{bmatrix}$ 和 A 有相同的列空间，证明略。但零空间不同，反例 $A = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$, $\mathcal{N}(A) = \mathbb{R}^2, \mathcal{N}\left(\begin{bmatrix}A & A\end{bmatrix}\right) = \mathbb{R}^4$.
- $\begin{bmatrix}A & AC\end{bmatrix}$ 和 A 有相同的列空间，但是零空间不同。证明略，反例同上。
- $\begin{bmatrix}A\\A\end{bmatrix}$ 和 A 的列空间不同，但是有相同的零空间。证明略，反例:$A = I_2$, 则

$$\mathcal{R}(A) = \mathbb{R}^2, \mathcal{R}\left(\begin{bmatrix}A\\A\end{bmatrix}\right) = \left\{\begin{bmatrix}x_1\\x_2\\x_1\\x_2\end{bmatrix} : x_1, x_2 \in \mathbb{R}\right\}.$$

- $\begin{bmatrix} A \\ CA \end{bmatrix}$ 的列空间不同，但是有相同的零空间. 证明略，反例如 $A = C = I_2$.
- $\begin{bmatrix} A & A \\ 0 & A \end{bmatrix}$ 和 A 的列空间、零空间都不同。略。
- PA 和 A 有相同的零空间，但列空间不同。证明略，反例如 $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.
- AQ 和 A 有相同的列空间，但零空间不同。证明略，反例如：$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

。

2. $A + I_n, A^2, A^{\mathrm{T}}$ 和 A 的列空间、零空间都不一定相等。

□

练习 2.1.7. 设矩阵 A, B 具有相同的行数和列数，对下列判断证明或举出反例：

1. 如果 A, B 有相同的零空间，那么对任意向量 $\boldsymbol{b}$，$A\boldsymbol{x} = \boldsymbol{b}$ 与 $B\boldsymbol{x} = \boldsymbol{b}$ 一定同解.

2. 如果 A, B 有相同的列空间，那么对任意向量 $\boldsymbol{b}$，$A\boldsymbol{x} = \boldsymbol{b}$ 与 $B\boldsymbol{x} = \boldsymbol{b}$ 一定同解.

3. 如果 A, B 有相同的零空间和列空间，那么对任意向量 $\boldsymbol{b}$，$A\boldsymbol{x} = \boldsymbol{b}$ 与 $B\boldsymbol{x} = \boldsymbol{b}$ 一定同解.

证明*:* 1. 错误。例如：$A = I_2, B = 2I_2, \boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

2. 错误。例如：$A = I_2, B = 2I_2, \boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

3. 错误。例如：$A = I_2, B = 2I_2, \boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

□

练习 2.1.8. 设 $A\boldsymbol{x} = \boldsymbol{b}$ 有解，证明：

1. $\mathcal{R}(A) = \mathcal{R}\left(\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}\right)$.　　*2.* $\mathcal{N}(A^{\mathrm{T}}) = \mathcal{N}\left(\begin{bmatrix} A^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{bmatrix}\right)$.

证明*:* 1. 显然 $\mathcal{R}(A) \subseteq \mathcal{R}\left(\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}\right)$.我们还需要证明 $\mathcal{R}(A) \supseteq \mathcal{R}\left(\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}\right)$，只需要证明 $\boldsymbol{b} \in \mathcal{R}(A)$. 因为 $A\boldsymbol{x} = \boldsymbol{b}$ 有解，设解为 $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $A = \begin{bmatrix} \alpha_1 & \cdots & \alpha_n \end{bmatrix}$, 则 $\boldsymbol{b} = x_1\alpha_1 + \cdots + x_n\alpha_n$, 因此 $\boldsymbol{b} \in \mathrm{Span}(\alpha_1, \cdots, \alpha_n) = \mathcal{R}(A)$.

2. 若 $\boldsymbol{y} \in \mathcal{N}\left(\begin{bmatrix} A^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{bmatrix}\right)$, 则 $\begin{bmatrix} A^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{bmatrix} \boldsymbol{y} = 0$, 即 $\begin{bmatrix} A^{\mathrm{T}}\boldsymbol{y} \\ \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} \end{bmatrix} = 0$, 所以 $A^{\mathrm{T}}\boldsymbol{y} = 0$, 即 $\boldsymbol{y} \in \mathcal{N}(A^{\mathrm{T}})$, 这证明了 $\mathcal{N}\left(\begin{bmatrix} A^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{bmatrix}\right) \subseteq \mathcal{N}(A^{\mathrm{T}})$.

   若 $\boldsymbol{y} \in \mathcal{N}(A^{\mathrm{T}})$, 我们假设 $\boldsymbol{x}$ 是 $A\boldsymbol{x} = \boldsymbol{b}$ 的解，则

$$\begin{bmatrix} A^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{bmatrix} \boldsymbol{y} = \begin{bmatrix} A^{\mathrm{T}}\boldsymbol{y} \\ \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{y} \end{bmatrix} = 0$$

   即 $\boldsymbol{y} \in \mathcal{N}\left(\begin{bmatrix} A^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{bmatrix}\right)$, 这证明了 $\mathcal{N}(A^{\mathrm{T}}) \subseteq \mathcal{N}\left(\begin{bmatrix} A^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{bmatrix}\right)$.

□

练习 2.1.9. 把 $\mathbb{R}^4$ 中的向量 $\boldsymbol{b}$ 表示成 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4$ 的线性组合：

1. $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \boldsymbol{a}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \boldsymbol{a}_4 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}$.

2. $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \boldsymbol{a}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{a}_4 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$.

证明*:* 设 $\boldsymbol{b} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3 + x_4\boldsymbol{a}_4$, 那么就有

$$\begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 & \boldsymbol{a}_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \boldsymbol{b}$$

因此问题就转化成解线性方程组的问题。

1. 根据上面的分析，我们只需要解方程组

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}$$

   过程略，解得 $\boldsymbol{x} = \begin{bmatrix} \frac{5}{4} \\ \frac{1}{4} \\ \frac{-1}{4} \\ \frac{-1}{4} \end{bmatrix}$, 于是 $\boldsymbol{b} = \frac{5}{4}\boldsymbol{a}_1 + \frac{1}{4}\boldsymbol{a}_2 - \frac{1}{4}\boldsymbol{a}_3 - \frac{1}{4}\boldsymbol{a}_4$.

2. 根据上面的分析，我们只需要解方程组

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 3 & 0 & -1 \\ 1 & 1 & 0 & -1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

过程略，解得 $\boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$, 于是 $\boldsymbol{b} = \boldsymbol{a}_1 - \boldsymbol{a}_3$.

□

练习 2.1.10. 判断下列向量组是否线性相关：

1. $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \\ 4 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \boldsymbol{a}_3 = \begin{bmatrix} 1 \\ 3 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{a}_4 = \begin{bmatrix} 0 \\ -7 \\ 3 \\ 1 \end{bmatrix}.$

2. $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{a}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \boldsymbol{a}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \boldsymbol{a}_5 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}.$

根据定义，向量 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k$ 线性无关，当且仅当 $x_1\boldsymbol{a}_1 + \cdots + x_k\boldsymbol{a}_k = 0$ 没有非零解，因此我们把判断向量是否线性无关的问题转化成了齐次方程组有没有非零解的问题。

证明: 1. 令 $A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 & \boldsymbol{a}_4 \end{bmatrix}$, 考虑齐次方程组 $A\boldsymbol{x} = 0$. 我们对 A 作初等行变换，得到

$$\operatorname{rref}(A) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

因此方程组 $A\boldsymbol{x} = 0$ 没有非零解，从而向量组线性无关。

2. 令 $A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 & \boldsymbol{a}_4 & \boldsymbol{a}_5 \end{bmatrix}$, 考虑齐次方程组 $A\boldsymbol{x} = 0$. 把 A 的所有行都加到最后一行即可得到零行，因此 A 不是可逆矩阵，因此 $A\boldsymbol{x} = 0$you 非零解，从而向量组线性相关。

□

练习 2.1.11. 如果向量组 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ 中的任何两个都线性无关，该向量组是否一定线性无关？

证明: 不一定，例如 $\boldsymbol{a} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \boldsymbol{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ □

练习 2.1.12. 设 $\mathbb{R}^n$ 中向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 满足 $k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2 + k_3\boldsymbol{a}_3 = 0$，其中 $k_1k_2 \neq 0$. 求证：

$$\mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_3) = \mathrm{Span}(\boldsymbol{a}_2, \boldsymbol{a}_3).$$

证明: 我们先证明 $\mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_3) \subseteq \mathrm{Span}(\boldsymbol{a}_2, \boldsymbol{a}_3)$；只需要证明 $\boldsymbol{a}_1 \in \mathrm{Span}(\boldsymbol{a}_2, \boldsymbol{a}_3)$. 事实上，由 $k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2 + k_3\boldsymbol{a}_3 = 0$ 可得 $\boldsymbol{a}_1 = -k_1^{-1}k_2\boldsymbol{a}_2 - k_1^{-1}k_3\boldsymbol{a}_3 \in \mathrm{Span}(\boldsymbol{a}_2, \boldsymbol{a}_3)$.

同理可证 $\mathrm{Span}(\boldsymbol{a}_2, \boldsymbol{a}_3) \subseteq \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_3)$ □

练习 2.1.13. 设向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s$ 线性无关，证明向量组 $\boldsymbol{a}_1, \boldsymbol{a}_1 + \boldsymbol{a}_2, \ldots, \boldsymbol{a}_1 + \boldsymbol{a}_2 + \cdots + \boldsymbol{a}_s$ 线性无关.

证明: 假设 $x_1\boldsymbol{a}_1 + x_2(\boldsymbol{a}_1 + \boldsymbol{a}_2) + \cdots + x_s(\boldsymbol{a}_1 + \boldsymbol{a}_2 + \cdots + \boldsymbol{a}_s) = 0$, 那么

$$(x_1 + x_2 + \cdots + x_s)\boldsymbol{a}_1 + (x_2 + \cdots + x_s)\boldsymbol{a}_2 + \cdots + x_s\boldsymbol{a}_s = 0.$$

因为向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s$ 线性无关，所以

$$\begin{cases} x_1 + x_2 + \cdots + x_s & = 0 \\ \qquad\; x_2 + \cdots + x_s & = 0 \\ \qquad\qquad\quad \vdots & \\ \qquad\qquad\quad\; x_s & = 0 \end{cases}$$

逐步回代可得 $x_1 = x_2 = \cdots = x_s = 0$, 因此向量组 $\boldsymbol{a}_1, \boldsymbol{a}_1 + \boldsymbol{a}_2, \ldots, \boldsymbol{a}_1 + \boldsymbol{a}_2 + \cdots + \boldsymbol{a}_s$ 线性无关. □

练习 2.1.14. 设 $\mathbb{R}^n$ 中向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s$ 线性无关，A 是 n 阶可逆矩阵，求证：$A\boldsymbol{a}_1, A\boldsymbol{a}_2, \ldots, A\boldsymbol{a}_s$ 线性无关.

附注. 上题其实是本题的特殊情形. 另外，假设 **A** 是单射，是否已经足够?

证明: 假设 $x_1A\boldsymbol{a}_1 + x_2A\boldsymbol{a}_2 + \cdots + x_sA\boldsymbol{a}_s = 0$, 则

$$A(x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_s\boldsymbol{a}_s) = 0.$$

A 是单射，则 $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_s\boldsymbol{a}_s = 0$. 因为向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s$ 线性无关，所以 $x_1 = x_2 = \cdots = x_s = 0$. 因此 $A\boldsymbol{a}_1, A\boldsymbol{a}_2, \cdots, A\boldsymbol{a}_s$ 线性无关。 □

练习 2.1.15. 证明：一个线性无关向量组的任意部分组也线性无关；如果向量组有一个部分组线性相关，则该向量组也线性相关.

证明: 这个命题里面两个命题互为逆否命题，我们只需要证明其中一个即可。我们证明后一个。给定向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_k$，不妨设这个部分组为 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_s$. 若 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_s$ 线性相关，则存在不全为零的 $x_1, x_2, \cdots, x_s$ 使得 $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_s\boldsymbol{a}_s = 0$, 那么 $x_1, x_2, \cdots, x_s, 0, \cdots, 0$ 也不全为零，且

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_s\boldsymbol{a}_s + 0 \cdot x_{s+1} + \cdots + 0 \cdot \boldsymbol{a}_k = 0,$$

即：向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_k$ 线性相关。 □

练习 2.1.16. 证明：一个向量组线性相关当且仅当其中有一个向量可以被其他向量线性表示.

证明: 给定向量组 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_s$.

若 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_s$ 线性相关，则存在不全为零的 $x_1, \cdots, x_s$ 使得 $x_1\boldsymbol{a}_1 + \cdots + x_s\boldsymbol{a}_s = 0$. 设 $x_i \neq 0$, 那么

$$\boldsymbol{a}_i = -\frac{x_1}{x_i}\boldsymbol{a}_1 - \cdots - \frac{x_{i-1}}{x_i}\boldsymbol{a}_{i-1} - \frac{x_{i+1}}{x_i}\boldsymbol{a}_{i+1} - \cdots - \frac{x_s}{x_i}\boldsymbol{a}_s$$

即 $\boldsymbol{a}_i$ 可以被其他向量线性表示。

若向量组 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_s$ 中有一个向量可以被其他向量线性表示,设为 $\boldsymbol{a}_i$,那么存在 $x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_s$, 使得 $\boldsymbol{a}_i = x_1\boldsymbol{a}_1 + \cdots + x_{i-1}\boldsymbol{a}_{i-1} + x_{i+1}\boldsymbol{a}_{i+1} \cdots + x_s\boldsymbol{a}_s$, 那么 $x_1, \cdots, x_{i-1}, -1, x_{i+1}, \cdots, x_s$ 不全为零，而且

$$x_1\boldsymbol{a}_1 + \cdots + x_{i-1}\boldsymbol{a}_{i-1} + (-1)\boldsymbol{a}_i + x_{i+1}\boldsymbol{a}_{i+1} \cdots + x_s\boldsymbol{a}_s = 0.$$

因此向量组 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_s$ 线性相关。 □

练习 2.1.17. 给定 $\mathbb{R}^m$ 中向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$，从每个向量中去掉第 $i_1, i_2, \ldots, i_s$ 个分量，得到 $\mathbb{R}^{m-s}$ 中向量组 $\boldsymbol{a}'_1, \boldsymbol{a}'_2, \ldots, \boldsymbol{a}'_n$. 证明：

1. 如果 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ 线性相关，则 $\boldsymbol{a}'_1, \boldsymbol{a}'_2, \ldots, \boldsymbol{a}'_n$ 线性相关.

   附注. 有追求的读者可以尝试这种线性代数味道更浓的证明方法：可否找到一个矩阵 A，把每个 $\boldsymbol{a}_k$ 变成 $\boldsymbol{a}'_k, k = 1, 2, \ldots, n$?

2. 如果 $\boldsymbol{a}'_1, \boldsymbol{a}'_2, \ldots, \boldsymbol{a}'_n$ 线性无关，则 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ 线性无关.

证明: 1. 设 $x_1, \cdots, x_n$ 不全为零，使得 $x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n = 0$. 设 $\boldsymbol{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \cdots \\ a_{mi} \end{bmatrix}$, 则

$$x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n = \begin{bmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \cdots \\ x_1a_{m1} + x_2a_{m2} + \cdots + x_na_{mn} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$x_1\boldsymbol{a}'_1 + x_1\boldsymbol{a}'_2 + \cdots + x_n\boldsymbol{a}'_n$ 就是将矩阵 $\begin{bmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \cdots \\ x_1a_{m1} + x_2a_{m2} + \cdots + x_na_{mn} \end{bmatrix}$ 中第 $i_1, i_2, \ldots, i_s$ 个分量，所以 $x_1\boldsymbol{a}'_1 + x_1\boldsymbol{a}'_2 + \cdots + x_n\boldsymbol{a}'_n = 0$. 故 $\boldsymbol{a}'_1, \boldsymbol{a}'_2, \ldots, \boldsymbol{a}'_n$ 线性相关。
当然，我们也可以用附注的方法，这个 A 是将单位矩阵 I_m 的第 $i_1, i_2, \ldots, i_s$ 行去掉得到的矩阵。

2. 这是第 1 小题的逆否命题，不需重复证明。

□

练习 2.1.18. 证明，对任意 $\mathbb{R}^m$ 的非平凡子空间 $\mathcal{M}, \mathcal{N}$，都有 $\mathcal{M} \cup \mathcal{N} \neq \mathbb{R}^m$.

非平凡的意思是不等于零，也不等于全空间。这个问题里面，我们只需要他们不等于全空间即可。

证明: 我们用反证法。假设 $\mathcal{M} \cup \mathcal{N} = \mathbb{R}^m$。

首先，$\mathcal{M}, \mathcal{N}$ 互不包含，若不然，$\mathcal{M} \cup \mathcal{N}$ 就等于两者中较大的子空间，与它们的非平凡性矛盾。于是，存在 $\boldsymbol{a} \in \mathcal{M} - \mathcal{N}, \boldsymbol{b} \in \mathcal{N} - \mathcal{M}$. 我们考虑 $\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b} \in \mathbb{R}^m = \mathcal{M} \cup \mathcal{N}$, 那么

- 要么 $\boldsymbol{c} \in \mathcal{M}$，则 $\boldsymbol{b} = \boldsymbol{c} - \boldsymbol{a} \in \mathcal{M}$, 矛盾；
- 要么 $\boldsymbol{c} \in \mathcal{N}$，则 $\boldsymbol{a} = \boldsymbol{c} - \boldsymbol{b} \in \mathcal{N}$, 矛盾.

所以 $\mathcal{M} \cup \mathcal{N} \neq \mathbb{R}^m$ □

练习 2.1.19 (子空间的和)**.** 设 $\mathcal{M}, \mathcal{N}$ 是 $\mathbb{R}^m$ 的两个子空间，定义集合：

$$\mathcal{M} + \mathcal{N} := \{\boldsymbol{m} + \boldsymbol{n} : \boldsymbol{m} \in \mathcal{M}, \boldsymbol{n} \in \mathcal{N}\}.$$

证明：

1. 集合 $\mathcal{M} + \mathcal{N}$ 是 $\mathbb{R}^m$ 的子空间，称为子空间 $\mathcal{M}$ 与 $\mathcal{N}$ 的**和**.

2. 交集 $\mathcal{M} \cap \mathcal{N}$ 是 $\mathbb{R}^m$ 的子空间，称为子空间 $\mathcal{M}$ 与 $\mathcal{N}$ 的**交**.

3. 集合的交与并满足 $(S_1 \cup S_2) \cap S_3 = (S_1 \cap S_3) \cup (S_2 \cap S_3)$. 证明或举出反例：子空间的交与和满足 $(\mathcal{M} + \mathcal{N}) \cap \mathcal{W} = (\mathcal{M} \cap \mathcal{W}) + (\mathcal{N} \cap \mathcal{W})$.

4. 集合的交与并满足 $(S_1 \cap S_2) \cup S_3 = (S_1 \cup S_3) \cap (S_2 \cup S_3)$. 证明或举出反例：子空间的交与和满足 $(\mathcal{M} \cap \mathcal{N}) + \mathcal{W} = (\mathcal{M} + \mathcal{W}) \cap (\mathcal{N} + \mathcal{W})$.

证明:

1. 设 $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{M} + \mathcal{N}, k \in \mathbb{R}$, 设 $\boldsymbol{v} = \boldsymbol{m}_1 + \boldsymbol{n}_1, \boldsymbol{w} = \boldsymbol{m}_2 + \boldsymbol{n}_2$, 则

$$\boldsymbol{v} + \boldsymbol{w} = (\boldsymbol{m}_1 + \boldsymbol{n}_1) + (\boldsymbol{m}_2 + \boldsymbol{n}_2) = (\boldsymbol{m}_1 + \boldsymbol{m}_2) + (\boldsymbol{n}_1 + \boldsymbol{n}_2) \in \mathcal{M} + \mathcal{N}$$

且 $k\boldsymbol{v} = k(\boldsymbol{m}_1 + \boldsymbol{n}_1) = (k\boldsymbol{m}_1) + (k\boldsymbol{n}_1) \in \mathcal{M} + \mathcal{N}$. 所以 $\mathcal{M} + \mathcal{N}$ 是线性子空间。

2. 设 $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{M} \cap \mathcal{N}, k \in \mathbb{R}$, 则 $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{M}, \boldsymbol{v}, \boldsymbol{v} \in \mathcal{N}$, 因此 $k\boldsymbol{v}, \boldsymbol{v} + \boldsymbol{w} \in \mathcal{M}, k\boldsymbol{v}, \boldsymbol{v} + \boldsymbol{w} \in \mathcal{N}$, 因此 $k\boldsymbol{v} \in \mathcal{M} \cap \mathcal{N}, \boldsymbol{v} + \boldsymbol{w} \in \mathcal{M} \cap \mathcal{N}$. 故 $\mathcal{M} \cap \mathcal{N}$ 是 $\mathbb{R}^m$ 的子空间.

3. 不满足，例如我们考虑 $\mathbb{R}^2$, $\mathcal{M} = \mathrm{Span}(\boldsymbol{e}_1), \mathcal{N} = \mathrm{Span}(\boldsymbol{e}_2), \mathcal{W} = \mathrm{Span}(\boldsymbol{e}_1 + \boldsymbol{e}_2)$. 那么 $(\mathcal{M}+\mathcal{N})\cap\mathcal{W} = \mathcal{W}$, 而 $(\mathcal{M}\cap\mathcal{W}) = (\mathcal{N}\cap\mathcal{W}) = \{0\}$; 所以 $(\mathcal{M}+\mathcal{N})\cap\mathcal{W} \neq (\mathcal{M}\cap\mathcal{W})+(\mathcal{N}\cap\mathcal{W})$.

4. 不满足，例如我们考虑 $\mathbb{R}^2$, $\mathcal{M} = \mathrm{Span}(\boldsymbol{e}_1), \mathcal{N} = \mathrm{Span}(\boldsymbol{e}_2), \mathcal{W} = \mathrm{Span}(\boldsymbol{e}_1 + \boldsymbol{e}_2)$。那么 $\mathcal{M}\cap\mathcal{N} = \{0\}$, 所以 $\mathcal{M}\cap\mathcal{N}+\mathcal{W} = \mathcal{W}$. 而 $\mathcal{M}+\mathcal{W} = \mathcal{N}+\mathcal{W} = \mathbb{R}^2$, 因此 $(\mathcal{M}+\mathcal{W})\cap(\mathcal{N}+\mathcal{W}) = \mathbb{R}^2$.

□

练习 2.1.20. 设 $\mathbb{R}^m$ 中子空间 $\mathcal{M} = \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s), \mathcal{N} = \mathrm{Span}(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_t)$，证明：

$$\mathcal{M} + \mathcal{N} = \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s, \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_t).$$

证明:

- 由于 $\boldsymbol{a}_i, \boldsymbol{b}_j \in \mathcal{M}+\mathcal{N}$, 根据命题 2.1.9, $\mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s, \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_t) \subseteq \mathcal{M}+\mathcal{N}$.

- $\boldsymbol{a} \in \mathcal{M}, \boldsymbol{b} \in \mathcal{N}$, 设 $\boldsymbol{a} = x_1\boldsymbol{a}_1 + \cdots + x_s\boldsymbol{a}_s, \boldsymbol{b} = y_1\boldsymbol{b}_1 + \cdots + y_t\boldsymbol{b}_t$, 那么

$$\boldsymbol{a} + \boldsymbol{b} = x_1\boldsymbol{a}_1 + \cdots + x_s\boldsymbol{a}_s + y_1\boldsymbol{b}_1 + \cdots + y_t\boldsymbol{b}_t \in \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s, \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_t).$$

  因此 $\mathcal{M} + \mathcal{N} \subseteq \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s, \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_t)$

综合两方面即得 $\mathcal{M} + \mathcal{N} = \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s, \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_t)$. □

练习 2.1.21.

1. 给定 $m \times n, m \times s$ 矩阵 A, B，证明：$\mathcal{R}(A) + \mathcal{R}(B) = \mathcal{R}(C)$，其中 $C = \begin{bmatrix} A & B \end{bmatrix}$.
2. 给定 $m \times n, l \times n$ 矩阵 A, B，证明：$\mathcal{N}(A) \cap \mathcal{N}(B) = \mathcal{N}(D)$，其中 $D = \begin{bmatrix} A \\ B \end{bmatrix}$.

证明: 1. 设 $A = \begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{bmatrix}, B = \begin{bmatrix} \boldsymbol{b}_1 & \cdots & \boldsymbol{b}_s \end{bmatrix}$, 那么 $\mathcal{R}(A) = \mathrm{Span}(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n), \mathcal{R}(B) = \mathrm{Span}(\boldsymbol{b}_1, \cdots, \boldsymbol{b}_s), \mathcal{R}(C) = \mathrm{Span}(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n, \boldsymbol{b}_1, \cdots \boldsymbol{b}_s)$. 由上一题可得，

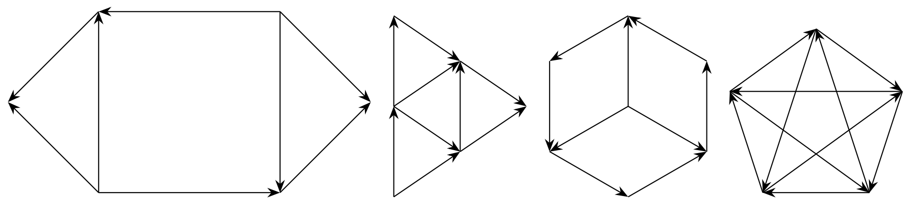
$$\mathcal{R}(A) + \mathcal{R}(B) = \mathrm{Span}(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n, \boldsymbol{b}_1, \cdots \boldsymbol{b}_s) = \mathcal{R}(C).$$

2. $\boldsymbol{x} \in \mathcal{N}(D) \Leftrightarrow D\boldsymbol{x} = 0 \Leftrightarrow \begin{bmatrix} A\boldsymbol{x} \\ B\boldsymbol{x} \end{bmatrix} = 0 \Leftrightarrow A\boldsymbol{x} = 0, B\boldsymbol{x} = 0 \Leftrightarrow \boldsymbol{x} \in \mathcal{N}(A) \cap \mathcal{N}(B)$.

□

练习 2.1.22 (Kirchhoff 电压定律). 对图 2.1中的电路网络，令 M_G 为对应的关联矩阵. 根据 *Kirchhoff* 电压定律，找出一个线性方程组，其解集恰为 $\mathcal{R}(M_G)$. 要求用尽量少的方程（但不需要写出证明）.

提示：对于最后一个图，它有 10 条边，5 个点，因此 M_G 是 10×5 的矩阵. 由于 $\mathcal{N}(M_G)$ 是 1 维的（所有点电势相等的情况），因此 $\mathcal{R}(M_G)$ 是几维的？至少需要几个等式才能确定？



(a) (b) (c) (d)

图 2.1: Kirchhoff 定律的练习

证明: 事实上，我们需要标记好点和边的的标号才能写出对应的矩阵，我们就不再画图了，可以根据我们写出的 M_G 反推点和边的标号。

(a) $M_G = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}$ 它有三个圈，因此需要三个线性方程，根据 Kirchhoff 电压定律，

$$\begin{cases} u_5 - u_6 - u_8 & = 0 \\ -u_1 + u_7 - u_4 + u_8 & = 0 \\ u_2 - u_3 - u_7 & = 0 \end{cases}$$

(b) $M_G = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix}$ 它有四个圈，因此需要四个线性方程，根据 Kirchhoff 电压定律，

$$\begin{cases} u_1 + u_6 - u_9 & = 0 \\ u_2 - u_3 + u_7 & = 0 \\ -u_4 + u_5 + u_8 & = 0 \\ -u_7 - u_8 + u_9 & = 0 \end{cases}$$

(c) $M_G = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{bmatrix}$ 它有三个圈，因此需要三个线性方程，根据 Kirchhoff 电压定律，

$$\begin{cases} -u_1 - u_2 - u_8 + u_7 & = 0 \\ u_8 - u_3 - u_4 - u_9 & = 0 \\ -u_7 + u_9 - u_5 - u_6 & = 0 \end{cases}$$

(d) $M_G = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 & 0 \end{bmatrix}$ 由维数公式，$\dim \mathcal{R}(M_G) = 4$, 因此至少 6 个方程。

因此我们需要找出“独立”的圈。

$$\begin{cases} u_1 + u_2 + u_{10} & = 0 \\ -u_{10} + u_3 - u_6 & = 0 \\ u_6 + u_4 + u_5 & = 0 \\ u_1 + u_8 + u_5 & = 0 \\ -u_7 + u_4 - u_8 & = 0 \\ u_3 + u_4 + u_9 & = 0 \end{cases}$$

□

练习 2.1.23 (电路网络拓展)**.** *设教材例 2.1.18 图 2.1.1 中的电路网络的关联矩阵是 M_G，定义图的 Laplace 矩阵 $L_G = M_G^{\mathrm{T}} M_G$.*

附注. 这类 Laplace 矩阵与图上的随机游走等问题密切相关.

1. 矩阵 L_G 的元素有什么意义?
2. 求子空间 $\mathcal{N}(L_G)$ 的一组生成向量.
3. 求子空间 $\mathcal{R}(L_G)$ 的一组生成向量，要求其中每个向量至少有两个分量为零.
4. 任取向量 $\boldsymbol{v} \in \mathcal{N}(L_G), \boldsymbol{w} \in \mathcal{R}(L_G)$，$\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w}$ 可能取哪些值?
5. 求线性方程组，其解集恰好是 $\mathcal{R}(L_G)$.

证明: 1. $(L_G)_{i,j} \neq 0$ 当且仅当 v_i, v_j 有边直接相连。

2. 显然 $\mathcal{N}(M_G) \subseteq \mathcal{N}(L_G)$. 对任意 $\boldsymbol{x} \in \mathcal{N}(L_G)$, 即 $L_G\boldsymbol{x} = 0$, 那么 $(M_G\boldsymbol{x})^{\mathrm{T}}(M_G\boldsymbol{x}) = 0$, 那么 $M_G\boldsymbol{x} = 0$, 即 $\boldsymbol{x} \in \mathcal{N}(M_G)$; 因此 $\mathcal{N}(L_G) \subseteq \mathcal{N}(M_G)$, 从而它们相等。所以 $\mathcal{N}(L_G) = \operatorname{Span}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}\right)$.

3. 我们可以把 L_G 计算出来，然后对 L_G 作初等行变换，找出主元列，这些主元列就是 $\mathcal{R}(L_G)$ 的一组生成向量。在学习了维数公式以后，我们还可以用下面的方法。
   由于 L_G 的列向量都是 M_G^{T} 的列向量的线性组合，因此 $\mathcal{R}(L_G) \subseteq \mathcal{R}(M_G^{\mathrm{T}})$，于是 $\operatorname{rank}(L_G) \leq \operatorname{rank}(M_G^{\mathrm{T}}) = \operatorname{rank}(M_G)$. 由上一小题知 $\mathcal{N}(L_G) = \mathcal{N}(M_G)$，根据教材第

四节中的维数公式可知 $\text{rank}(L_G) = \text{rank}(M_G^{\text{T}}) = 3$, 那么找 $\mathcal{R}(M_G^{\text{T}})$ 的一组生成向量即可。这样我们就省去了对 L_G 的计算。我们给出一组生成元仅供参考：

$$\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\-1\\0\\1\end{bmatrix}$$

4. 我们通过上述的生成元集发现，$\boldsymbol{v}^{\text{T}}\boldsymbol{w} = 0$.

5. 令 $A = \boldsymbol{v}^{\text{T}}$. 直接验证 $\mathcal{N}(A) = \text{Span}\left(\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\-1\\0\\1\end{bmatrix}\right) = \mathcal{R}(L_G)$

□

练习 2.1.24. *以下哪些向量组是矩阵* $\begin{bmatrix}1&1&1&1\\1&2&3&4\\4&3&2&1\end{bmatrix}$ *的零空间的基?*

1. $\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}$.

2. $\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\-3\\2\end{bmatrix}$.

3. $\begin{bmatrix}1\\-2\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\-2\\1\end{bmatrix}$.

4. $\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}, \begin{bmatrix}1\\-2\\1\\0\end{bmatrix}, \begin{bmatrix}-1\\2\\-1\\0\end{bmatrix}$.

5. $\begin{bmatrix}1\\-1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\-1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\-1\end{bmatrix}$.

证明*:* 我们对矩阵 $\begin{bmatrix}1&1&1&1\\1&2&3&4\\4&3&2&1\end{bmatrix}$ 作初等行变换转化为行阶梯形矩阵：

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

有两个主元和两个自由变元。我们可以验算，第 5 组向量都不属于矩阵零空间；第 1,2,3,4 组向量都属于矩阵零空间。但是我们注意到，第 1 组向量生成的空间不包含第 3 组向量中的元素，因此第 1 组向量不是基；第 4 组向量中有两个元素线性相关，因此也不是基。第 2,3 组向量中两个向量线性无关，是零空间的一组基。 □

2.2 基和维数

练习 2.2.1. 求下列向量组的极大线性无关部分组和秩：

1. $\begin{bmatrix}2\\3\\1\\1\end{bmatrix},\begin{bmatrix}4\\6\\2\\2\end{bmatrix},\begin{bmatrix}0\\1\\2\\1\end{bmatrix},\begin{bmatrix}2\\4\\3\\2\end{bmatrix}.$ *2.* $\begin{bmatrix}-1\\0\\1\\0\\0\end{bmatrix},\begin{bmatrix}1\\1\\1\\1\\0\end{bmatrix},\begin{bmatrix}0\\1\\2\\1\\0\end{bmatrix}.$

对于这样的问题，我们的方法如下。假设 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_s\in\mathbb{R}^m$, 要求这个向量组的极大线性无关组和秩，我们把这些向量排成一个矩阵，对这个矩阵做初等行变换，找出主元。主元的个数就是秩；主元所在的列就是一个极大线性无关组。

证明: 1. 我们把这组向量排成一个矩阵 $\begin{bmatrix}2&4&0&2\\3&6&1&4\\1&2&2&3\\1&2&1&2\end{bmatrix}$, 做初等行变换可得 $\begin{bmatrix}1&2&0&1\\0&0&1&1\\0&0&0&0\\0&0&0&0\end{bmatrix}$ 有两个主元，在第一列和第三列，因此秩为 2, $\begin{bmatrix}2\\3\\1\\1\end{bmatrix},\begin{bmatrix}0\\1\\2\\1\end{bmatrix}$ 是一个极大线性无关组.

2. 我们把这组向量排成一个矩阵 $\begin{bmatrix}-1&1&0\\0&1&1\\1&1&2\\0&1&1\\0&0&0\end{bmatrix}$, 做初等行变换可得 $\begin{bmatrix}1&-1&0\\0&1&1\\0&0&0\\0&0&0\\0&0&0\end{bmatrix}$ 有两个主元，在第一列和第二列，因此秩为 2，$\begin{bmatrix}-1\\0\\1\\0\\0\end{bmatrix},\begin{bmatrix}1\\1\\1\\1\\0\end{bmatrix}$ 是一个极大线性无关组。

□

练习 2.2.2. 在下列向量组中，除去哪个向量，得到的向量组和原向量组线性等价?

1. $\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}1\\2\\3\end{bmatrix}.$ *2.* $\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}1\\2\\0\end{bmatrix}.$

3. $\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}2\\0\\0\end{bmatrix}.$ *4.* $\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix}.$

证明: 我们去掉一个向量之后得到的向量组和原向量组等价，当且仅当去掉的这个向量可以被剩下的向量线性表出。

1. 去掉任意一个向量都符合题意。

2. $\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}1\\2\\0\end{bmatrix}$ 中去掉任意一个向量都可以。

3. $\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}2\\0\\0\end{bmatrix}$ 中去掉任意一个向量都可以。

4. 只有去掉 $\begin{bmatrix}0\\0\\0\end{bmatrix}$ 才可以。

□

练习 2.2.3. *用筛选法去掉方程组中所有多余的方程：* $\begin{bmatrix}1 & -1 & 0 & 0\\1 & 0 & -1 & 0\\1 & 0 & 0 & -1\\0 & 1 & -1 & 0\\0 & 1 & 0 & -1\\0 & 0 & 1 & -1\end{bmatrix}\boldsymbol{x} = \begin{bmatrix}-1\\0\\0\\1\\1\\0\end{bmatrix}.$

所谓的用筛选法去掉方程组中多余的方程，我们只需要对系数矩阵作初等行变换，找到主元所在的行即可。但是我们需要注意，如果在这个过程中有交换两行的操作，一定要记录下来，找到主元对应的最初的行即可。那么最保险的方法，就是把系数矩阵转置，对转置之后的矩阵作初等行变换，找到主元，那么主元所在的列就对应于原方程组的行。

证明: 我们把系数矩阵转置 $\begin{bmatrix}1 & 1 & 1 & 0 & 0 & 0\\-1 & 0 & 0 & 1 & 1 & 0\\0 & -1 & 0 & -1 & 0 & 1\\0 & 0 & -1 & 0 & -1 & -1\end{bmatrix}$, 对它做初等行变换可得

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

有三个主元，在第一、二、三列，因此原方程组的第一、二、三行以外的方程是多余的。 □

练习 2.2.4. *给定线性无关的向量组* $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$*，求* $\boldsymbol{a}_1 - \boldsymbol{a}_2, \boldsymbol{a}_2 - \boldsymbol{a}_3, \boldsymbol{a}_3 - \boldsymbol{a}_1$ *所有的极大线性无关部分组.*

证明: 因为 $(\boldsymbol{a}_1 - \boldsymbol{a}_2) + (\boldsymbol{a}_2 - \boldsymbol{a}_3) + (\boldsymbol{a}_3 - \boldsymbol{a}_1) = 0$, 因此它们线性相关，那么极大线性无关组至多有两个向量。

若 $x(\boldsymbol{a}_1 - \boldsymbol{a}_2) + y(\boldsymbol{a}_2 - \boldsymbol{a}_3) = 0$, 那么 $x\boldsymbol{a}_1 + (y - x)\boldsymbol{a}_2 - y\boldsymbol{a}_3 = 0$. 因为 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 线性无关，因此 $x = y - x = y = 0$，所以 $\boldsymbol{a}_1 - \boldsymbol{a}_2, \boldsymbol{a}_2 - \boldsymbol{a}_3$ 线性无关，从而是极大线性无关组。

同理可证 $\boldsymbol{a}_1 - \boldsymbol{a}_2, \boldsymbol{a}_3 - \boldsymbol{a}_1$ 和 $\boldsymbol{a}_2 - \boldsymbol{a}_3, \boldsymbol{a}_3 - \boldsymbol{a}_1$ □

练习 2.2.5. 证明，一个向量组的任意线性无关的部分组都可以扩充成它的一个极大线性无关部分组.

证明: 设 S 是一个向量组，它包含于 $\mathbb{R}^m$. T 是一个线性无关的部分组。若 $S \subset \operatorname{Span}(T)$, 则 T 已经是极大线性无关组了。若 $S \not\subset \operatorname{Span}(T)$, 则存在 $\boldsymbol{b}_1 \in S - \operatorname{Span}(T)$, 那么考虑 $T \cup \{\boldsymbol{b}_1\}$；我们断言这是一个线性无关组。事实上，设 $T = \{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_s\}$, 以及 $y\boldsymbol{b}_1 + x_1\boldsymbol{a}_1 + \cdots + x_s\boldsymbol{a}_s = 0$. 若 $y \neq 0$, 则 $\boldsymbol{b}_1 = -x^{-1}(x_1\boldsymbol{a}_1 + \cdots + x_s\boldsymbol{a}_s) \in \operatorname{Span}(T)$, 这和 $\boldsymbol{a}$ 的选取矛盾，因此 $x = 0$. 那么 $x_1\boldsymbol{a}_1 + \cdots + x_s\boldsymbol{a}_s = 0$ 而 T 是个线性无关组，所以 $x_1 = \cdots = x_s = 0$. 这就证明了 $T \cup \{\boldsymbol{b}_1\}$ 线性无关。

以此类推，我们可以找到 $\boldsymbol{b}_2, \cdots, \boldsymbol{b}_k, \cdots$，使得 $T \cup \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_k\}$ 都线性无关。事实上，这个步骤只能进行有限步，因为 $S \subset \mathbb{R}^m$, 它的一个极大无关组至多有 m 个元素。因此存在一个有限的 k, 使得 $S \subseteq \operatorname{Span}(T \cup \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_k\})$。那么我们就把 T 扩充成了 S 的一个极大线性无关组 $T \cup \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_k\}$ □

练习 2.2.6. 证明，如果向量组 S 可以被向量组 T 线性表示，则 $\operatorname{rank}(S) \leq \operatorname{rank}(T)$.

证明: 向量组 S 可以被向量组 T 线性表示，则 $S \subseteq \operatorname{Span}(T)$, 那么就有 $\operatorname{Span}(S) \subseteq \operatorname{Span}(T)$, 所以 $\dim \operatorname{Span}(S) \leq \dim \operatorname{Span}(T)$. 我们注意到 $\operatorname{rank}(S) = \dim \operatorname{Span}(S)$, 因此 $\operatorname{rank}(S) \leq \operatorname{rank}(T)$. □

练习 2.2.7. 证明向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 与向量组 $\boldsymbol{a}_1 + \boldsymbol{a}_2, \boldsymbol{a}_2 + \boldsymbol{a}_3, \boldsymbol{a}_3 + \boldsymbol{a}_1$ 线性等价.

附注. 有追求的读者可以尝试这种线性代数味道更浓的证明方法：可否找到一个可逆矩阵 A，把一个向量组变成另一个向量组？

证明: 我们只需证明向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 能被向量组 $\boldsymbol{a}_1 + \boldsymbol{a}_2, \boldsymbol{a}_2 + \boldsymbol{a}_3, \boldsymbol{a}_3 + \boldsymbol{a}_1$ 线性表出。设 $\boldsymbol{b}_1 = \boldsymbol{a}_1 + \boldsymbol{a}_2, \boldsymbol{b}_2 = \boldsymbol{a}_2 + \boldsymbol{a}_3, \boldsymbol{b}_3 = \boldsymbol{a}_3 + \boldsymbol{a}_1$, 那么

$$\begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \boldsymbol{b}_3 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

矩阵 $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ 可逆，它的逆为 $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$, 所以

$$\begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \boldsymbol{b}_3 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

即 $\boldsymbol{a}_1 = \frac{\boldsymbol{b}_1 - \boldsymbol{b}_2 + \boldsymbol{b}_3}{2}, \boldsymbol{a}_2 = \frac{\boldsymbol{b}_1 + \boldsymbol{b}_2 - \boldsymbol{b}_3}{2}, \boldsymbol{a}_3 = \frac{-\boldsymbol{b}_1 + \boldsymbol{b}_2 + \boldsymbol{b}_3}{2}$. 证毕。 □

练习 2.2.8. 证明，如果向量组和它的一个部分组的秩相同，则两个向量组线性等价.

证明: 设 S 是一个向量组，$T \subseteq S$ 是一个部分组，并且 $\operatorname{rank}(S) = \operatorname{rank}(T)$. 取 T_0 是 T 的一个极大线性无关组，那么 T_0 也是 S 中的线性无关组。根据 $\operatorname{rank}(S) = \operatorname{rank}(T)$, 则有 T_0 也是 S 的极大线性无关组，因此 S 中的任意向量否可以由 T_0 线性表出。从而 S 和 T 线性等价。 □

练习 2.2.9. 举例说明秩相等的向量组未必线性等价.

证明: 例如：我们考虑 $\mathbb{R}^2$, $S = \{\boldsymbol{e}_1\}, T = \{\boldsymbol{e}_2\}$. □

练习 2.2.10. 已知向量组的秩是 r，设 S 是一个包含 r 个向量的部分组. 证明：

1. 如果 S 线性无关，则 S 是原向量组的一个极大线性无关部分组.
2. 如果 S 与原向量组线性等价，则 S 是原向量组的一个极大线性无关部分组.

证明:

1. 若 S 不极大，则存在 $S' \supsetneq S$ 也线性无关，那么 S' 包含至少 $r+1$ 个元素，因此向量组的秩至少是 $r+1$, 矛盾。
2. 首先我们证明线性无关。记 $S = \{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r\}$, 假设 S 线性相关，则存在不全为零的 $x_1, \cdots, x_r$ 使得 $x_1\boldsymbol{a}_1 + \cdots + x_r\boldsymbol{a}_r = 0$. 我们不妨设 $x_i \neq 0$, 那么

$$\boldsymbol{a}_i = -x_i^{-1}(x_1\boldsymbol{a}_1 + \cdots + x_{i-1}\boldsymbol{a}_{i-1} + x_{i+1}\boldsymbol{a}_{i+1} + \cdots + x_r\boldsymbol{a}_r),$$

那么 $\mathrm{Span}(S) = \mathrm{Span}(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{i-1}, \boldsymbol{a}_{i+1}, \cdots, \boldsymbol{a}_r)$, 这就有 $\mathrm{rank}(S) \leq r-1$, 小于整个向量组的秩，因此不可能与原向量组线性等价，矛盾，所以 S 线性无关。
S 与原向量组线性等价，则任意 $\boldsymbol{a}$ 都可以被 S 线性表出，所以 $S \cup \{\boldsymbol{a}\}$ 线性相关，所以不存在真包含 S 的线性无关组。
综上，S 是极大的线性无关部分组。

□

练习 2.2.11. 证明命题 *2.2.17*.

我们需要明确一下子空间的基的定义：见定义 2.1.13，它满足：是生成元集且线性无关。

证明: 取 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ 是 $\mathcal{M}$ 的一组基。

1. 我们需要证明 $\mathcal{M}$ 中的元素都可以由 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r$ 表出，即 $\mathrm{Span}(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r) = \mathcal{M}$. 由于 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r$ 线性无关,根据定义,它们就是 $\mathrm{Span}(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r)$ 的一组基。若 $\mathrm{Span}(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r) \subsetneq \mathcal{M}$, 则我们可以将 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r$ 扩张成为 $\mathcal{M}$ 的一组基 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r, \boldsymbol{a}_{r+1}, \cdots, \boldsymbol{a}_s$, 那么 $\dim \mathcal{M} = s \geq r+1$. 矛盾。
2. 我们需要证明 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r$ 线性无关, 那么 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r$ 的一个极大线性无关组 $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_s$ 至多包含 $s \leq r-1$ 个元素，并且 $\mathrm{Span}(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r) = \mathrm{Span}(\boldsymbol{b}_1, \cdots, \boldsymbol{b}_s)$, 因此 $\dim \mathcal{M} = s \leq r-1$. 矛盾。

□

练习 2.2.12. 求下列子空间的基和维数：

1. 空间 $\mathbb{R}^3$ 中平面 $x-y=0$ 与平面 $x+y-2z=0$ 的交集.
2. 空间 $\mathbb{R}^3$ 中与向量 $\begin{bmatrix}1 & -1 & 0\end{bmatrix}, \begin{bmatrix}1 & 1 & -2\end{bmatrix}$ 都垂直的向量组成的子空间.
3. 齐次线性方程组 $\begin{bmatrix}1 & 1 & 1\\ 0 & 0 & 1\\ 0 & 0 & 1\end{bmatrix}\boldsymbol{x} = 0$ 的解集.

证明: 1. 平面 $x-y=0$ 与平面 $x+y-2z=0$ 的交集. 即方程组 $\begin{bmatrix}1 & -1 & 0\\ 1 & 1 & -2\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$ 的解集，该解集为 $\left\{k\begin{bmatrix}1\\1\\1\end{bmatrix}:k\in\mathbb{R}\right\}$。

2. 向量 $\begin{bmatrix}x\\y\\z\end{bmatrix}$ 与向量 $\begin{bmatrix}1 & -1 & 0\end{bmatrix}$ 垂直就是 $\begin{bmatrix}1 & -1 & 0\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=0$, 即 $x-y=0$. 同理，我们有 $x+y-2z=0$. 因此所求子空间为 $\left\{k\begin{bmatrix}1\\1\\1\end{bmatrix}:k\in\mathbb{R}\right\}$。

3. 过程略，该方程组解集为 $\left\{k\begin{bmatrix}1\\-1\\0\end{bmatrix}:k\in\mathbb{R}\right\}$。

□

练习 2.2.13. 一个三阶方阵，如果每行、每列以及两个对角线上的元素之和都相等，则称为幻方矩阵. 判断 $\left\{\boldsymbol{a}=\begin{bmatrix}a_1 & a_2 & \cdots & a_9\end{bmatrix}^{\mathrm{T}}\in\mathbb{R}^9 \,\middle|\, \begin{bmatrix}a_1 & a_2 & a_3\\ a_4 & a_5 & a_6\\ a_7 & a_8 & a_9\end{bmatrix}\text{是幻方矩阵}\right\}$ 是否为 $\mathbb{R}^9$ 的子空间. 如果是，求它的一组基.

提示. 一个常见的幻方矩阵是 $\begin{bmatrix}2 & 9 & 4\\ 7 & 5 & 3\\ 6 & 1 & 8\end{bmatrix}$. 你能否根据它找到许多幻方矩阵?

证明: 一个矩阵是幻方矩阵，当且仅当它满足

$$\begin{aligned} a_1+a_2+a_3 &= a_4+a_5+a_6 &= a_7+a_8+a_9\\ =a_1+a_4+a_7 &= a_2+a_5+a_8 &= a_3+a_6+a_9\\ =a_1+a_5+a_9 &= a_3+a_5+a_7 & \end{aligned}$$

当且仅当 $a_1,\cdots,a_9$ 是齐次方程组

$$\begin{bmatrix}1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1\\ -1 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 1\\ 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0\\ 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1\\ -1 & 0 & 1 & 0 & -1 & 1 & 0 & 0 & 0\\ 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\ \vdots\\x_9\end{bmatrix}=\begin{bmatrix}0\\0\\ \vdots\\0\end{bmatrix}$$

的解。记系数矩阵为 A, 那么所求的空间即 $\mathcal{N}(A)$, 因此是 $\mathbb{R}^9$ 的子空间。要求它的一组基，这是

后面章节的任务。我们给出一组基仅供参考，解题方法请学习了相关内容后自行解决

$$\begin{bmatrix}2/3 & 2/3 & -1/3 & -2/3 & 1/3 & 4/3 & 1 & 0 & 0\end{bmatrix}^{\mathrm{T}}$$
$$\begin{bmatrix}2/3 & -1/3 & 2/3 & 1/3 & 1/3 & 1/3 & 0 & 1 & 0\end{bmatrix}^{\mathrm{T}}$$
$$\begin{bmatrix}-1/3 & 2/3 & 2/3 & 4/3 & 1/3 & -2/3 & 0 & 0 & 1\end{bmatrix}^{\mathrm{T}}$$

□

练习 2.2.14. 任取非零常数 $k_1, k_2, \ldots, k_n$ 满足 $\frac{1}{k_1}+\frac{1}{k_2}+\cdots+\frac{1}{k_n}+1 \neq 0$，求如下向量组的秩：

$$\boldsymbol{a}_1=\begin{bmatrix}1+k_1\\1\\\vdots\\1\end{bmatrix}, \boldsymbol{a}_2=\begin{bmatrix}1\\1+k_2\\\vdots\\1\end{bmatrix}, \ldots, \boldsymbol{a}_n=\begin{bmatrix}1\\1\\\vdots\\1+k_n\end{bmatrix}.$$

证明: 我们考虑矩阵

$$A=\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n\end{bmatrix}=\begin{bmatrix}k_1 & & & \\ & k_2 & & \\ & & \ddots & \\ & & & k_n\end{bmatrix}+\begin{bmatrix}1 & 1 & \cdots & 1\\1 & 1 & \cdots & 1\\\vdots & & & \vdots\\1 & 1 & \cdots & 1\end{bmatrix}$$

我们记 $\boldsymbol{v}=\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}$, 那么 $A=\operatorname{diag}(k_1,\cdots,k_n)+\boldsymbol{v}\boldsymbol{v}^T$. 根据 Sherman-Morrison 公式，A 可逆当且仅当 $1+\boldsymbol{v}^{\mathrm{T}}\operatorname{diag}(k_1,\cdots,k_n)^{-1}\boldsymbol{v}\neq 0$. 由题目条件，

$$1+\boldsymbol{v}^{\mathrm{T}}\operatorname{diag}(k_1,\cdots,k_n)^{-1}\boldsymbol{v}=\frac{1}{k_1}+\frac{1}{k_2}+\cdots+\frac{1}{k_n}+1.$$

所以矩阵 A 可逆，那么 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 线性无关，所以该向量组的秩为 n. □

练习 2.2.15. 设 $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_r$ 是子空间 $\mathcal{M}$ 的一组基.令 $\begin{bmatrix}\boldsymbol{b}_1 & \boldsymbol{b}_2 & \ldots & \boldsymbol{b}_r\end{bmatrix}=\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \ldots & \boldsymbol{a}_r\end{bmatrix}P$，其中 P 是 r 阶方阵. 证明，$\boldsymbol{b}_1,\boldsymbol{b}_2,\ldots,\boldsymbol{b}_r$ 是 $\mathcal{M}$ 的一组基当且仅当矩阵 P 可逆.

证明: 因为 $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_r$ 是子空间 $\mathcal{M}$ 的一组基，那么 $\dim\mathcal{M}=r$.

- 假设 $\boldsymbol{b}_1,\boldsymbol{b}_2,\ldots,\boldsymbol{b}_r$ 是 $\mathcal{M}$ 的一组基,那么它们线性无关,即方程组 $\begin{bmatrix}\boldsymbol{b}_1 & \boldsymbol{b}_2 & \ldots & \boldsymbol{b}_r\end{bmatrix}\boldsymbol{x}=0$ 只有零解，即 $\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \ldots & \boldsymbol{a}_r\end{bmatrix}P\boldsymbol{x}=0$ 只有零解；因为 $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_r$ 线性无关，因此 $P\boldsymbol{x}=0$ 只有零解，从而 P 可逆。
- 若 P 可逆，则 $\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \ldots & \boldsymbol{a}_r\end{bmatrix}=\begin{bmatrix}\boldsymbol{b}_1 & \boldsymbol{b}_2 & \ldots & \boldsymbol{b}_r\end{bmatrix}P^{-1}$, 也就是说，任意 $\boldsymbol{a}_i(1\leq i\leq r)$ 都可以被 $\boldsymbol{b}_1,\boldsymbol{b}_2,\ldots,\boldsymbol{b}_r$ 线性表出，所以 $\mathcal{M}=\operatorname{Span}(\boldsymbol{b}_1,\cdots,\boldsymbol{b}_r)$. 由命题 2.2.17 知 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_r$ 是 $\mathcal{M}$ 的一组基。

□

练习 2.2.16 (Steinitz 替换定理)**.** *设* $S\colon \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_r$ *线性无关*,*可被* $T\colon \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_t$ *线性表示*,*求证*:

1. $r \leq t$.

2. *可以选择* T *中的* r *个向量换成* S,*得到的新的向量组与* T *线性等价*.

证明: 由题意,对任意 $1 \leq i \leq r$, $\boldsymbol{a}_i = c_{1i}\boldsymbol{b}_1 + c_{2i}\boldsymbol{b}_2 + \cdots + c_{ti}\boldsymbol{b}_t$, 写成矩阵形式:

$$\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_r\end{bmatrix} = \begin{bmatrix}\boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_t\end{bmatrix} C.$$

其中 $C = [c_{ij}] \in M_{t\times r}$.

1. 我们用反证法,假设 $r > t$, 那么齐次方程组 $C\boldsymbol{x} = 0$ 一定有非零解,我们取一个非零解 $\boldsymbol{x}_0$, 那么

$$\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_r\end{bmatrix} \boldsymbol{x}_0 = \begin{bmatrix}\boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_t\end{bmatrix} C\boldsymbol{x}_0 = 0$$

   这说明 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r$ 线性相关,矛盾。

2. 我们对 r 作归纳。先考虑 $r = 1$; 因为 $\boldsymbol{a}_1$ 能被 T 线性表出,那么 $x_1, \cdots, x_t$, 使得 $\boldsymbol{a}_1 = x_1\boldsymbol{b}_1 + \cdots + x_t\boldsymbol{b}_t$. 那么 $x_1, \cdots, x_t$ 中存在非零数,否则 $\boldsymbol{a}_1 = 0$, 那么 S 线性相关,矛盾。不妨设 $x_1 \neq 0$, 那我们用 $\boldsymbol{a}_1$ 来替换 $\boldsymbol{b}_1$, 我们证明 T 和 $T_1 : \boldsymbol{a}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_t$ 线性等价。显然 T_1 能被 T 线性表出,那么我们只需证明 T 能被 T_1 线性表出;而这只需要证明 $\boldsymbol{b}_1$ 能被 T_1 线性表出。事实上,我们有

$$\boldsymbol{b}_1 = -x_1^{-1}(-\boldsymbol{a}_1 + x_2\boldsymbol{b}_2 \cdots + x_t\boldsymbol{b}_t).$$

   这就证明了 T 和 T_1 线性等价。

   下面我们假设对 $1 \leq r < t$, 该命题成立,我们来证明 $r+1$ 的情形。那么规矩归纳假设,我们可以在 T 中找到 r 个向量,不妨设为 $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_r$,把他们替换成 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r$,得到的向量组 $T' : \boldsymbol{a}_1, \cdots, \boldsymbol{a}_r, \boldsymbol{b}_{r+1}, \cdots, \boldsymbol{b}_t$ 和 T 线性等价。现在考虑 $\boldsymbol{a}_{r+1}$, 因为 $\boldsymbol{a}_{r+1}$ 能被 T 线性表出,那么它也能被 T' 线性表出,因此存在 $y_1, \cdots, y_t$ 使得

$$\boldsymbol{a}_{r+1} = y_1\boldsymbol{a}_1 + \cdots + y_r\boldsymbol{a}_r + y_{r+1}\boldsymbol{b}_{r+1} + \cdots + y_t\boldsymbol{b}_t;$$

   我们断言 $y_{r+1}, \cdots, y_t$ 中存在非零数,若不然则 $y_1\boldsymbol{a}_1 + \cdots + y_r\boldsymbol{a}_r - \boldsymbol{a}_{r+1} = 0$, 这说明 S 线性相关矛盾。那么我们不妨设 $y_{r+1} \neq 0$, 同理我们可以证明 T' 和 T'' : $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{r+1}, \boldsymbol{b}_{r+2}, \cdots, \boldsymbol{b}_t$ 线性等价,于是 T, T'' 也线性等价。这就完成了归纳。

□

练习 2.2.17 (平行的平面)**.** *考虑方程* $x - 3y - z = 12$ *和* $x - 3y - z = 0$ *的解集. 这两个解集的交集是什么? 如果* $\boldsymbol{x}_1, \boldsymbol{x}_2$ *分别是第一个和第二个方程的解,那么* $\boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{x}_1 + 2\boldsymbol{x}_2, 2\boldsymbol{x}_1 + \boldsymbol{x}_2$ *分别是哪个方程的解?*

附注. 从几何上看,这两个解集是 $\mathbb{R}^3$ 中两个平行的平面,其中一个经过原点,因此是二维子空间;另一个不经过原点,因此不是子空间.

证明: 两个方程解集的交集即方程组 $\begin{bmatrix}1 & -3 & -1\\1 & -3 & -1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}12\\0\end{bmatrix}$ 的解集。显然这个方程组无解，因此这两个方程的解集的交集为空集。

令 $\boldsymbol{v}=\begin{bmatrix}1\\-3\\-1\end{bmatrix}$, $\boldsymbol{v}^{\mathrm{T}}\begin{bmatrix}x\\y\\z\end{bmatrix}=x-3y-z$ 那么

- $\boldsymbol{v}^{\mathrm{T}}(\boldsymbol{x}_1+\boldsymbol{x}_2)=\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}_1+\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}_2=12$, 因此 $\boldsymbol{x}_1+\boldsymbol{x}_2$ 是方程 $x-3y-z=12$ 的解；
- $\boldsymbol{v}^{\mathrm{T}}(\boldsymbol{x}_1+2\boldsymbol{x}_2)=\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}_1+2\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}_2=12$, 因此 $\boldsymbol{x}_1+2\boldsymbol{x}_2$ 是方程 $x-3y-z=12$ 的解；
- $\boldsymbol{v}^{\mathrm{T}}(2\boldsymbol{x}_1+\boldsymbol{x}_2)=2\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}_1+\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}_2=23$, 因此 $2\boldsymbol{x}_1+\boldsymbol{x}_2$ 是方程 $x-3y-z=24$ 的解。

□

练习 2.2.18. 设 $\mathcal{M},\mathcal{N}$ 是 $\mathbb{R}^m$ 的两个子空间，求证维数公式：

$$\dim(\mathcal{M}+\mathcal{N})=\dim\mathcal{M}+\dim\mathcal{N}-\dim(\mathcal{M}\cap\mathcal{N}).$$

提示. $\mathcal{M}\cap\mathcal{N}$ 的一组基，根据基扩充定理，分别扩充成 $\mathcal{M}$ 和 $\mathcal{N}$ 的一组基，证明这两组基的并集是 $\mathcal{M}+\mathcal{N}$ 的一组基.

证明: 取 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r$ 是 $\mathcal{M}\cap\mathcal{N}$ 的一组基，根据基扩充定理，它可以扩张

- $\mathcal{M}$ 的一组基 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r,\boldsymbol{b}_1,\cdots,\boldsymbol{b}_s$；和
- $\mathcal{N}$ 的一组基 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r,\boldsymbol{c}_1,\cdots,\boldsymbol{c}_t$.

我们证明：$\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r,\boldsymbol{b}_1,\cdots,\boldsymbol{b}_s,\boldsymbol{c}_1,\cdots,\boldsymbol{c}_t$ 是 $\mathcal{M}+\mathcal{N}$ 的一组基.

- $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r,\boldsymbol{b}_1,\cdots,\boldsymbol{b}_s,\boldsymbol{c}_1,\cdots,\boldsymbol{c}_t$ 线性无关。设

  $$x_1\boldsymbol{a}_1+\cdots+x_r\boldsymbol{a}_r+y_1\boldsymbol{b}_1+\cdots+y_s\boldsymbol{b}_s+z_1\boldsymbol{c}_1+\cdots+z_t\boldsymbol{c}_t=0.$$

  那么 $y_1\boldsymbol{b}_1+\cdots+y_s\boldsymbol{b}_s+z_1\boldsymbol{c}_1=-(x_1\boldsymbol{a}_1+\cdots+x_r\boldsymbol{a}_r+\cdots+z_t\boldsymbol{c}_t)$. 等式左边属于 $\mathcal{M}$, 等式右边属于 $\mathcal{N}$, 因此 $y_1\boldsymbol{b}_1+\cdots+y_s\boldsymbol{b}_s=-(x_1\boldsymbol{a}_1+\cdots+x_r\boldsymbol{a}_r+z_1\boldsymbol{c}_1+\cdots+z_t\boldsymbol{c}_t)\in\mathcal{M}\cap\mathcal{N}$. $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r$ 是 $\mathcal{M}\cap\mathcal{N}$ 的一组基，因此

  $$y_1\boldsymbol{b}_1+\cdots+y_s\boldsymbol{b}_s=-(x_1\boldsymbol{a}_1+\cdots+x_r\boldsymbol{a}_r+z_1\boldsymbol{c}_1+\cdots+z_t\boldsymbol{c}_t)=w_1\boldsymbol{a}_1+\cdots+w_r\boldsymbol{a}_r.$$

  因此

  $$\begin{cases} w_1\boldsymbol{a}_1+\cdots+w_r\boldsymbol{a}_r-(y_1\boldsymbol{b}_1+\cdots+y_s\boldsymbol{b}_s) & =0\\ (w_1+x_1)\boldsymbol{a}_1+\cdots+(w_r+x_r)\boldsymbol{a}_r+z_1\boldsymbol{c}_1+\cdots+z_t\boldsymbol{c}_t & =0\end{cases}$$

  由于 $\mathcal{M},\mathcal{N}$ 的基的选取，我们知道

  $$w_1=\cdots=w_r=y_1=\cdots=y_s=w_1+x_1=\cdots=w_r+x_r=z_1=\cdots=z_t=0.$$

  这就得到了线性无关性。

- $\mathcal{M}+\mathcal{N}=\mathrm{Span}(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r,\boldsymbol{b}_1,\cdots,\boldsymbol{b}_s,\boldsymbol{c}_1,\cdots,\boldsymbol{c}_t)$. 任取 $\boldsymbol{v}\in\mathcal{M}+\mathcal{N}$, 那么存在 $\boldsymbol{v}_1\in\mathcal{M},\boldsymbol{v}_2\in\mathcal{N}$ 使得 $\boldsymbol{v}=\boldsymbol{v}_1+\boldsymbol{v}_2$. 那么

  $$\boldsymbol{v}_1=x_1\boldsymbol{a}_1+\cdots x_r\boldsymbol{a}_r+y_1\boldsymbol{b}_1+\cdots+y_s\boldsymbol{b}_s,\boldsymbol{v}_2=z_1\boldsymbol{a}_1+\cdots+z_r\boldsymbol{a}_r+w_1\boldsymbol{c}_1+\cdots+w_t\boldsymbol{c}_t,$$

  于是 $\boldsymbol{v}=(x_1+z_1)\boldsymbol{a}_1+\cdots+(x_r+z_r)\boldsymbol{a}_r+y_1\boldsymbol{b}_1+\cdots+y_s\boldsymbol{b}_s+w_1\boldsymbol{c}_1+\cdots+w_t\boldsymbol{c}_t$.

这就证明了 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r,\boldsymbol{b}_1,\cdots,\boldsymbol{b}_s,\boldsymbol{c}_1,\cdots,\boldsymbol{c}_t$ 是 $\mathcal{M}+\mathcal{N}$ 的一组基. 所以

$$\dim(\mathcal{M}+\mathcal{N})=r+s+t=(r+s)+(r+t)-r=\dim\mathcal{M}+\dim\mathcal{N}-\dim(\mathcal{M}\cap\mathcal{N})$$

□

2.3 矩阵的秩

练习 2.3.1. 给定矩阵 $A=\boldsymbol{u}_1\boldsymbol{v}_1^{\mathrm{T}}+\boldsymbol{u}_2\boldsymbol{v}_2^{\mathrm{T}}$.

1. 写出 A 的行空间、列空间.

2. 如果 $\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{v}_1,\boldsymbol{v}_2$ 都不为零, 求 $\operatorname{rank}(A)$ 的所有可能值, 并讨论这四个向量与 $\operatorname{rank}(A)$ 之间的关系.

证明: 1. A 的列都是 $\boldsymbol{u}_1,\boldsymbol{u}_2$ 的线性组合,因此 $\mathcal{R}(A)$ 有四种可能:$\{0\},\operatorname{Span}(\boldsymbol{u}_1),\operatorname{Span}(\boldsymbol{u}_2),\operatorname{Span}(\boldsymbol{u}_1,\boldsymbol{u}_2)$. 同理 $\mathcal{R}(A^{\mathrm{T}})$ 有四种可能：$\{0\},\operatorname{Span}(\boldsymbol{v}_1),\operatorname{Span}(\boldsymbol{v}_2),\operatorname{Span}(\boldsymbol{v}_1,\boldsymbol{v}_2)$.

2. 由上一题，我们知道 $\operatorname{rank}(A)$ 可能的值为 $0,1,2$.

   - 当且仅当存在 $\lambda\neq 0$ 使得 $\boldsymbol{u}_1=\lambda\boldsymbol{u}_2,\boldsymbol{v}_1=-\lambda^{-1}\boldsymbol{v}_2$, 则 $\operatorname{rank}(A)=0$;
   - 当且仅当 $A\neq 0$ 且 $\boldsymbol{u}_1,\boldsymbol{u}_2$ 或者 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 线性相关时，$\operatorname{rank}(A)=1$;
   - 当且仅当 $\boldsymbol{u}_1,\boldsymbol{u}_2$ 线性无关且 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 线性无关时，$\operatorname{rank}(A)=2$.

□

练习 2.3.2. 求下列矩阵列空间的一组基:

1. $\begin{bmatrix}1&-1&2&1&0\\2&-2&4&-2&0\\3&0&6&-1&1\\0&3&0&0&1\end{bmatrix}$. *2.* $\begin{bmatrix}1&0&0&1&4\\0&1&0&2&5\\0&0&1&3&6\\1&2&3&14&32\\4&5&6&32&77\end{bmatrix}$. *3.* $\begin{bmatrix}1&0&1&0&0\\1&1&0&0&0\\0&1&1&0&0\\0&0&1&1&0\\0&1&0&1&1\end{bmatrix}$.

证明: 这样的问题有一个固定的程序，那就是对矩阵作初等行变换，找出主元，主元所在列的**原矩阵**的列向量构成矩阵列空间的一组基。以下我们省略过程，直接写结果。

1. 对矩阵做初等行变换可得 $\begin{bmatrix}1&-1&2&1&0\\0&3&0&-4&1\\0&0&0&4&0\\0&0&0&0&0\end{bmatrix}$ 主元在第 1,2,4 列，因此原矩阵的第 1,2,4 个列向量是矩阵列空间的一组基，即

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

2. 对矩阵做初等行变换可得 $\begin{bmatrix}1&0&0&1&4\\0&1&0&2&5\\0&0&1&3&6\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix}$ 主元在第 1,2,3 列，因此原矩阵的第 1,2,3 个列向量是矩阵列空间的一组基，即

$$\begin{bmatrix}1\\0\\0\\1\\4\end{bmatrix},\begin{bmatrix}0\\1\\0\\2\\5\end{bmatrix},\begin{bmatrix}0\\0\\1\\3\\6\end{bmatrix}$$

3. 对矩阵做初等行变换可得 $\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}$ 主元在第 1,2,3,4，5 列，因此原矩阵的第 1,2,3,4,5 个列向量是矩阵列空间的一组基。当然，此时 $\mathcal{R}(A)=\mathbb{R}^5$, 我们也可以直接选取 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_5$ 作为 $\mathcal{R}(A)$ 的一组基。

□

练习 2.3.3. 求满足下列条件的 3×4 矩阵 A 的行简化阶梯形矩阵和秩：

1. $A\boldsymbol{x}=0$ 解集的一组基为 $\begin{bmatrix}-3\\1\\0\\0\end{bmatrix},\begin{bmatrix}-2\\0\\6\\1\end{bmatrix}$.

2. A 的 (i,j) 元为 4.

3. A 的 (i,j) 元为 $a_{ij}=i+j+1$.

4. A 的 (i,j) 元为 $a_{ij}=(-1)^{i+j}$.

证明*:* 1. 若 x_1 不是主元，则 $\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$ 是方程组的解，根据我们的解的情形可知 x_1 是主元。

若 x_2 也是主元，那么 A 的行简化阶梯形矩阵第二行形如 $x_2+ax_3+bx_4=0$，但是 $\begin{bmatrix}-3\\1\\0\\0\end{bmatrix}$ 不是解；若 x_4 是主元，那么 A 的行简化阶梯形矩阵的第二行形如 $x_4=0$，那

么 $\begin{bmatrix}-2\\0\\6\\1\end{bmatrix}$ 不是解。因此 x_1, x_3 是主元，那么 A 的行简化梯形矩阵形如

$$\begin{bmatrix}1 & a & 0 & b\\0 & 0 & 1 & c\\0 & 0 & 0 & 0\end{bmatrix}$$

把两个解代进去反解可得 $a=3, b=2, c=-6$, 因此 A 的行简化阶梯型矩阵为

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

A 的秩为 2.

2. 这个提示容易的，A 的行简化阶梯形矩阵为 $\begin{bmatrix}1 & 1 & 1 & 1\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix}$ 它的秩为 1.

3. 矩阵 $A=\begin{bmatrix}3 & 4 & 5 & 6\\4 & 5 & 6 & 7\\5 & 6 & 7 & 8\end{bmatrix}$, 那么 A 的行简化阶梯型矩阵为

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

A 的秩为 2.

4. 矩阵 A 为 $\begin{bmatrix}1 & -1 & 1 & -1\\-1 & 1 & -1 & 1\\1 & -1 & 1 & -1\end{bmatrix}$, 它的行简化阶梯形矩阵为

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

它的秩为 1.

□

练习 2.3.4. 设矩阵 A 的行简化阶梯形矩阵为 R，求下列 B 的行简化阶梯形矩阵.

1. $B=\begin{bmatrix}A & A\end{bmatrix}$.

2. $B=\begin{bmatrix}A & AC\end{bmatrix}$.

3. $B=\begin{bmatrix}A\\A\end{bmatrix}$.

4. $B=\begin{bmatrix}A\\CA\end{bmatrix}$.

5. $B=\begin{bmatrix}A & A\\0 & A\end{bmatrix}$.

6. $B=PA$，其中 P 是可逆矩阵.

7. A 有 LU 分解 $A=LU$，这里 L 为单位下三角阵，令 $B=U$.

证明: 1. B 的行简化阶梯型矩阵为 $\begin{bmatrix} R & R \end{bmatrix}$

2. B 的行简化阶梯形矩阵为 $\begin{bmatrix} R & RC \end{bmatrix}$

3. B 的行简化阶梯形矩阵为 $\begin{bmatrix} R \\ 0 \end{bmatrix}$

4. 注意到 $\begin{bmatrix} I_n & 0 \\ -C & I_n \end{bmatrix} B = \begin{bmatrix} A \\ 0 \end{bmatrix}$, 所以 B 的行简化阶梯形矩阵为 $\begin{bmatrix} R \\ 0 \end{bmatrix}$

5. 注意到 $\begin{bmatrix} I_n & -I_n \\ 0 & I_n \end{bmatrix} B = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$, 那么 B 的行简化阶梯型矩阵为 $\begin{bmatrix} R_1 & 0 \\ 0 & R_1 \\ 0 & 0 \end{bmatrix}$ 其中 R_1 是 R 的非零行。

6. 因为 P 是可逆矩阵，因此 B 和 A 的行简化阶梯形矩阵相同，也是 R.

7. 因为 L 是单位下三角矩阵，因此可逆，那么 $B=L^{-1}A$, 所以 B 的行简化阶梯形矩阵是 R.

□

练习 2.3.5. 求矩阵 A 中 "$*$" 处的元素，满足相应的条件：

1. $A=\begin{bmatrix} 1 & 2 & 4 \\ 2 & * & * \\ 4 & * & * \end{bmatrix}$，且 A 的秩为 1.

2. $A=\begin{bmatrix} * & 9 & * \\ 1 & * & * \\ 2 & 6 & -3 \end{bmatrix}$，且 A 的秩为 1.

3. $A=\begin{bmatrix} a & b \\ c & * \end{bmatrix}$，$a \neq 0$，且 A 的秩为 1.

4. $A=\begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 7 \\ 4 & 5 & * \end{bmatrix}$，且 A 的秩为 2.

证明: 1. A 的秩为 1 且第一行非零,所以 A 的所有行是第一行的倍数,所以 $A=\begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \\ 4 & 8 & 16 \end{bmatrix}$

2. A 的所有行是第三行的倍数，所以 $A=\begin{bmatrix} 3 & 9 & -\frac{9}{2} \\ 1 & 3 & -\frac{3}{2} \\ 2 & 6 & -3 \end{bmatrix}$

3. A 的所有行是第一行的倍数，所以 $A=\begin{bmatrix} a & b \\ c & \frac{bc}{a} \end{bmatrix}$

4. A 的前两列线性无关，它的秩为 2，因此第三列是第二列的线性组合，所以我们可以假设

$$\begin{bmatrix} 3 \\ 7 \\ * \end{bmatrix} = x \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} + y \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$

解出来 $x=y=1, *=9$，所以 $A=\begin{bmatrix}1&2&3\\3&4&7\\4&5&9\end{bmatrix}$

$\square$

练习 2.3.6. 构造满足下列条件的 4×8 矩阵 R，要求其中为 1 的元素尽量多：

1. R 为阶梯形矩阵，其中主列为第二、四、五列.
2. R 为阶梯形矩阵，其中主列为第一、三、六、八列.
3. R 为阶梯形矩阵，其中主列为第四、六列.
4. R 为行简化阶梯形矩阵，其中自由列为第二、四、五、六列.
5. R 为行简化阶梯形矩阵，其中自由列为第一、三、六、七、八列.

证明: 阶梯形矩阵，主元的下方都是 0; 行简化阶梯形矩阵主元所在列除了主元全是 0.

1. $R=\begin{bmatrix}0&1&1&1&1&1&1&1\\0&0&0&1&1&1&1&1\\0&0&0&0&1&1&1&1\\0&0&0&0&0&0&0&0\end{bmatrix}$
2. $R=\begin{bmatrix}1&1&1&1&1&1&1&1\\0&0&1&1&1&1&1&1\\0&0&0&0&0&1&1&1\\0&0&0&0&0&0&0&1\end{bmatrix}$
3. $R=\begin{bmatrix}0&0&0&1&1&1&1&1\\0&0&0&0&0&1&1&1\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{bmatrix}$
4. 主列是一三七八列，因此 $R=\begin{bmatrix}1&1&0&1&1&1&0&0\\0&0&1&1&1&1&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&1\end{bmatrix}$
5. 主列是二四五列 $R=\begin{bmatrix}0&1&1&0&0&1&1&1\\0&0&0&1&0&1&1&1\\0&0&0&0&1&1&1&1\\0&0&0&0&0&0&0&0\end{bmatrix}$

$\square$

练习 2.3.7. 说明满足下列条件的 4×7 矩阵 A 是否存在. 如果存在，举例说明，要求其中为 0 的元素尽量多.

1. A 为阶梯形矩阵，其中主列为第二、四、五列，A^{T} 的主列为第一、三列.
2. A 为阶梯形矩阵，其中主列为第二、四、五列，A^{T} 的主列为第二、三、四列.

证明: 1. 不存在，因为 A, A^{T} 有相同的秩，它们的主列个数相等。

2. 不存在，因为 A 有三个主列，且它是阶梯形矩阵，因此它的最后一行是零行，前三行非零，因此 A 形如
$$A=\begin{bmatrix}0&1&*&*&*&*&*\\0&0&0&1&*&*&*\\0&0&0&0&1&*&*\\0&0&0&0&0&0&0\end{bmatrix}$$
那么 A^{T} 的第一列是主列。

□

练习 2.3.8. 证明，$\operatorname{rank}(kA)=\operatorname{rank}(A)\ (k\neq 0),\quad \operatorname{rank}(A+B)\leq\operatorname{rank}(A)+\operatorname{rank}(B)$.

证明: 因为 $k\neq 0$, 我们对 kA 作倍乘变换可以得到 A, 因此 kA 和 A 有相同的行简化阶梯形矩阵。

注意到 $\operatorname{rank}(A+B)=\dim\mathcal{R}(A+B)$; 因为 $A+B$ 的列向量是 A 的列向量和 B 的列向量的线性组合，因此 $\mathcal{R}(A+B)\subseteq\mathcal{R}(A)+\mathcal{R}(B)$, 所以
$$\dim\mathcal{R}(A+B)\leq\dim(\mathcal{R}(A)+\mathcal{R}(B))\leq\dim\mathcal{R}(A)+\dim\mathcal{R}(B)=\operatorname{rank}(A)+\operatorname{rank}(B).$$

□

练习 2.3.9. 证明：$\max\{\operatorname{rank}(A),\operatorname{rank}(B)\}\leq\operatorname{rank}\left(\begin{bmatrix}A&B\end{bmatrix}\right)\leq\operatorname{rank}(A)+\operatorname{rank}(B)$.

证明: A 的列向量是 $\begin{bmatrix}A&B\end{bmatrix}$ 的列向量的一部分，因此 $\mathcal{R}(A)\subseteq\mathcal{R}(\begin{bmatrix}A&B\end{bmatrix})$, 所以 $\operatorname{rank}(A)\leq\operatorname{rank}(\begin{bmatrix}A&B\end{bmatrix})$;同理 $\operatorname{rank}(B)\leq\operatorname{rank}(\begin{bmatrix}A&B\end{bmatrix})$, 所以 $\max\{\operatorname{rank}(A),\operatorname{rank}(B)\}\leq\operatorname{rank}\left(\begin{bmatrix}A&B\end{bmatrix}\right)$.

另一方面
$$\operatorname{rank}\left(\begin{bmatrix}A&B\end{bmatrix}\right)=\dim(\mathcal{R}(A)+\mathcal{R}(B))\leq\dim(\mathcal{R}(A))+\dim(\mathcal{R}(B))=\operatorname{rank}(A)+\operatorname{rank}(B).$$

□

练习 2.3.10.

1. 对分块对角矩阵 $C=\begin{bmatrix}A&0\\0&B\end{bmatrix}$，证明：$\operatorname{rank}(C)=\operatorname{rank}(A)+\operatorname{rank}(B)$.
2. 对分块上三角矩阵 $C=\begin{bmatrix}A&X\\0&B\end{bmatrix}$，证明：$\operatorname{rank}(C)\geq\operatorname{rank}(A)+\operatorname{rank}(B)$. 由此证明，当 A,B 可逆时，C 也可逆.

证明: 设 A 有 n 行，B 有 m 行。

1. 我们对前 n 行作初等行变换，将 A 变成行简化阶梯形矩阵 R_A，保持后 m 行不变；然后对后 m 行作初等行变换，将 B 变成行简化阶梯形矩阵 R_B，保持前 n 行不变；然

后将 R_A 的零行和 R_B 的非零行交换，就得到了一个行简化阶梯形矩阵，这个行简化阶梯形矩阵的非零行的个数等于 R_A, R_B 非零行的个数的和。

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \longrightarrow \begin{bmatrix} R_A & 0 \\ 0 & B \end{bmatrix} \longrightarrow \begin{bmatrix} R_A & 0 \\ 0 & R_B \end{bmatrix} \longrightarrow \begin{bmatrix} R'_A & 0 \\ 0 & R'_B \\ 0 & 0 \end{bmatrix}$$

其中 R'_A, R'_B 是 R_A, R_B 的非零行。因此

$$\operatorname{rank}(C) = \operatorname{rank}(A) + \operatorname{rank}(B).$$

2. 类似于上面的过程，我们能够把矩阵通过初等行变换转化成

$$\begin{bmatrix} R'_A & C_1 \\ 0 & R'_B \\ 0 & C_2 \end{bmatrix}$$

其中 $\begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$ 是对 A 作初等行变换将 C 转化成的矩阵。由于 $\begin{bmatrix} R'_A & C_1 \\ 0 & R'_B \end{bmatrix}$ 是阶梯形矩阵，所以 C 的行简化阶梯形矩阵非零行个数大于等于 R_A, R_B 非零行个数的和，这就说明

$$\operatorname{rank}(C) \geq \operatorname{rank}(A) + \operatorname{rank}(B).$$

若 A, B 均可逆，那么 $\operatorname{rank}(A) = n, \operatorname{rank}(B) = m$, 因此 $\operatorname{rank}(C) \geq n + m$, 从而只能有 $\operatorname{rank}(C) = m + n$, 即 C 可逆。

□

练习 2.3.11. 设 A, B, C 分别为 $m \times n, n \times k, k \times s$ 矩阵，证明：

$$\operatorname{rank}(AB) + \operatorname{rank}(BC) \leq \operatorname{rank}(ABC) + \operatorname{rank}(B).$$

提示 构造合适的分块上三角矩阵.

证明: 考虑分块矩阵 $\begin{bmatrix} AB & 0 \\ B & BC \end{bmatrix}$, 我们作分块行列变换

$$\begin{bmatrix} AB & 0 \\ B & BC \end{bmatrix} \xrightarrow{\text{行}} \begin{bmatrix} 0 & -ABC \\ B & BC \end{bmatrix} \xrightarrow{\text{列}} \begin{bmatrix} 0 & -ABC \\ B & 0 \end{bmatrix}$$

因此

$$\operatorname{rank}(AB) + \operatorname{rank}(BC) \leq \operatorname{rank}\left(\begin{bmatrix} AB & 0 \\ B & BC \end{bmatrix}\right) = \operatorname{rank}\left(\begin{bmatrix} 0 & -ABC \\ B & 0 \end{bmatrix}\right) = \operatorname{rank}(ABC) + \operatorname{rank}(B).$$

□

练习 2.3.12. 设 $m \times n$ 矩阵 A 的秩为 1，证明，存在非零向量 $\boldsymbol{a} \in \mathbb{R}^m, \boldsymbol{b} \in \mathbb{R}^n$，使得 $A = \boldsymbol{a}\boldsymbol{b}^{\mathrm{T}}$.

证明: 我们任取 A 的一个非零列 $\boldsymbol{a}$；它的秩为 1，则它任意一列都是 $\boldsymbol{a}$ 的倍数，因此

$$A = \begin{bmatrix} b_1\boldsymbol{a} & b_2\boldsymbol{a} & \cdots & b_n\boldsymbol{a} \end{bmatrix} = \boldsymbol{a}\boldsymbol{b}^{\mathrm{T}}$$

其中 $\boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$. □

练习 2.3.13. *试分析矩阵 A 满足什么条件时，$AB = AC$ 可以推出 $B = C$.*

附注. 类比如下问题：对集合之间的映射 $f\colon X \to Y, g\colon Z \to X, h\colon Z \to X$，当一个映射 f 满足什么条件时，$f \circ g = f \circ h$ 可以推出 $g = h$?

证明: 设 $A \in M_{m\times n}(\mathbb{R})$. 我们取 C 是零向量，则题目条件是说 $A\boldsymbol{x} = 0$ 没有非零解，因此 A 的列向量线性无关。下面我们证明 A 的列向量线性无关也是充分条件。

A 的列向量线性无关则 A^{T} 的行向量线性无关，那就是说 A^{T} 的行简化阶梯形矩阵有 n 个非零行, 因此 $m \geq n$, 且对任意 $\boldsymbol{b} \in \mathbb{R}^n$, $A\boldsymbol{x} = \boldsymbol{b}$ 都有解，特别地，我们取 $\boldsymbol{x}_i (1 \leq i \leq n)$ 是方程组 $A^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{e}_i$ 的一个解, $X = \begin{bmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_n \end{bmatrix}$, 那么 $A^{\mathrm{T}}X = I_n$, 因此 $X^T A = I_n$. 那么 $AB = AC$ 推出 $X^{\mathrm{T}}AB = X^{\mathrm{T}}AC$，即 $B = C$. □

练习 2.3.14. *设 $m \times n$ 矩阵 A 列满秩，求证：存在行满秩的 $n \times m$ 矩阵 B，使得 $BA = I_n$.*

附注. 类比如下问题：设集合 X, Y 都只有有限个元素，映射 $f\colon X \to Y$ 为单射，是否存在映射 $g\colon Y \to X$，满足 $g \circ f$ 是 X 上的恒同变换?

证明: 考虑矩阵转置，则原命题等价于：$n \times m$ 矩阵 A 行满秩，则存在列满秩的 $m \times n$ 矩阵 B，使得 $AB = I_n$. 这一点在上一小题的解答中已经给出了证明。我们就不再重复了。 □

练习 2.3.15. *证明命题 2.3.10：方阵 A 可逆当且仅当它满秩。*

证明: 假设 A 可逆，则它的行简化阶梯形矩阵为 I_n, 因此 $\mathrm{rank}(A) = n$, 即它满秩。

假设 A 满秩，则它列满秩，根据上一题，存在 B 使得 $AB = I_n$, 因此它可逆。 □

练习 2.3.16. *设 A, B 是 n 阶方阵，利用不等式 $\mathrm{rank}(AB) \leq \min\{\mathrm{rank}(A), \mathrm{rank}(B)\}$，证明：*

1. *如果 $AB = I_n$，则 A, B 都可逆，且 $BA = I_n$.*
2. *如果 AB 可逆，则 A, B 都可逆.*

证明:

1. $n = \mathrm{rank}(I_n) = \mathrm{rank}(AB) \leq \min\{\mathrm{rank}(A), \mathrm{rank}(B)\}$, 而 $\mathrm{rank}(A), \mathrm{rank}(B) \leq n$, 所以 $\mathrm{rank}(A) = \mathrm{rank}(B) = n$, 因此 $A = B^{-1}$, 以及 $BA = I_n$.
2. 如果 AB 可逆, 则 $n = \mathrm{rank}(AB) \leq \min\{\mathrm{rank}(A), \mathrm{rank}(B)\}$, 而 $\mathrm{rank}(A), \mathrm{rank}(B) \leq n$, 所以 $\mathrm{rank}(A) = \mathrm{rank}(B) = n$, 因此 A, B 均可逆。

□

练习 2.3.17. *证明 2.3.15：给定两个 $m \times n$ 矩阵 A, B, 那么二者相抵，当且仅当存在 m 阶可逆矩阵 P 和 n 阶可逆矩阵 Q 使得 $PAQ = B$.*

证明: 我们需要注意到，作初等行变换等价于在矩阵左边乘以初等矩阵，作初等列变换等价于在右边乘以初等矩阵。因此，矩阵 A,B 相抵，当且仅当在矩阵 A 左边乘以一些初等矩阵，右边乘以一些初等矩阵可以得到 B. 而任何可逆矩阵都等于初等矩阵的乘积，因此矩阵 A,B 相抵，当且仅当存在 m 阶可逆矩阵 P 和 n 阶可逆矩阵 Q 使得 $PAQ=B$ □

练习 2.3.18. 证明，当 A 行满秩时，仅用初等列变换就可以把它化为相抵标准形；当 A 列满秩时，仅用初等行变换就可以把它化为相抵标准形.

证明: 设 $A\in M_{m\times n}(\mathbb{R})$ 这等价于说，如果 A 行满秩，则存在可逆矩阵 $P\in M_{n\times n}$ 使得 $AP=\begin{bmatrix}I_m & 0\end{bmatrix}$. 命题的第二部分可以考虑矩阵 A 的转置，从第一部分推出，证明我们略过。

因为 A 行满秩，因此对任意 $\boldsymbol{b}$, $A\boldsymbol{x}=\boldsymbol{b}$ 都有解，特别地，我们取 $\boldsymbol{a}_i\in\mathbb{R}^n(1\le i\le m)$ 是 $A\boldsymbol{x}=\boldsymbol{e}_i$ 的解。我们断言 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m$ 线性无关：假设 $a_1\boldsymbol{x}_1+\cdots+a_m\boldsymbol{a}_m=0$, 那么 $A(a\boldsymbol{x}_1+\cdots+a_m\boldsymbol{a}_m)=0$, 即 $a_1\boldsymbol{e}_1+\cdots+a_m\boldsymbol{e}_m=0$, 所以 $a_1=\cdots=a_m=0$. 那么我们把 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m$ 扩张成 $\mathbb{R}^n$ 的一组基 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$，此时 $P_1=\begin{bmatrix}\boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n\end{bmatrix}$ 是可逆矩阵，而且 $AP_1=\begin{bmatrix}I_m & A'\end{bmatrix}$. 令 $P_2=\begin{bmatrix}I_m & -A'\\ 0 & I_{n-m}\end{bmatrix}$, 那么 $AP_1P_2=\begin{bmatrix}I_m & 0\end{bmatrix}$. 我们取 $P=P_1P_2$ 即可。 □

练习 2.3.19. 对二阶方阵 A，如果存在 $n>2$，使得 $A^n=0$，求证：$A^2=0$. **提示**. 根据 A 的秩分类讨论

证明: A 是二阶方阵，则 $\operatorname{rank}(A)=0,1,2$.

- 若 $\operatorname{rank}(A)=2$, 则 A 可逆，那么 $A=A^n(A^{-1})^{n-1}=0$, 矛盾；所以 $\operatorname{rank}(A)\neq 2$.
- 若 $\operatorname{rank}(A)=1$, 存在 $\boldsymbol{v},\boldsymbol{w}\in\mathbb{R}^2$ 使得 $A=\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}$, 那么 $A^n=(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{v})^{n-1}A=0$，所以 $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{v}=0$, 于是 $A^2=(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{v})A=0$，命题成立。
- 若 $\operatorname{rank}(A)=0$，则 $A=0$, 那么显然有 $A^2=0$.

□

练习 2.3.20. 多项式 $f(x)$ 满足 $f(0)=0$，求证：对任意方阵 A，都有 $\operatorname{rank}(f(A))\le\operatorname{rank}(A)$.

证明: 设 $A\in M_{m\times m}(\mathbb{R}), f=a_nx^n+\cdots+a_1x+a_0$. 从 $f(0)=0$ 得知 $a_0=0$. 于是 $f(A)=a_nA^n+\cdots+a_1A=AB$, 其中 $B=a_nA^{n-1}+\cdots+a_1I_m$. 那么

$$\operatorname{rank}(f(A))=\operatorname{rank}(AB)\le\operatorname{rank}(A)$$

□

练习 2.3.21. 考虑反对称矩阵的秩.

1. 证明反对称矩阵的秩不能是 1.
2. 对反对称矩阵 A，去掉首行首列得到矩阵 B. 求证：B 也是反对称矩阵，且 $\operatorname{rank}(B)$ 等于 $\operatorname{rank}(A)$ 或 $\operatorname{rank}(A)-2$.

   提示. 注意 $A=\begin{bmatrix}0 & -\boldsymbol{v}^{\mathrm{T}}\\ \boldsymbol{v} & B\end{bmatrix}$，然后根据是否有 $\boldsymbol{v}\in\mathcal{R}(B)$ 进行分类讨论.
3. 证明反对称矩阵的秩必然是偶数. 由此证明，奇数阶反对称矩阵一定不可逆.

证明: 设 $A \in M_{m\times m}(\mathbb{R})$ 是反对称矩阵。

1. 若 $\operatorname{rank}(A)=1$, 那么存在列向量 $\boldsymbol{v},\boldsymbol{w}$ 使得 $A=\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}$. 那么首先 $\mathcal{R}(A)=\operatorname{Span}(\boldsymbol{v})$, 其次 $A^{\mathrm{T}}=\boldsymbol{w}\boldsymbol{v}^{\mathrm{T}}=-A$, 于是 $\mathcal{R}(A^{\mathrm{T}})=\operatorname{Span}(\boldsymbol{w})=\mathcal{R}(A)$, 因此 $\boldsymbol{w}=c\boldsymbol{v}$, 即 $A=c\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$, 但是一方面 $A^{\mathrm{T}}=c\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$, 另一方面 $A^{\mathrm{T}}=-A=-c\boldsymbol{v}\boldsymbol{v}^{T}$, 所以 $c=0$. 矛盾。
2. 按题意，$A=\begin{bmatrix}0 & -\boldsymbol{v}^{\mathrm{T}}\\ \boldsymbol{v} & B\end{bmatrix}$. 注意到 $\operatorname{rank}(B)\le\operatorname{rank}\left(\begin{bmatrix}\boldsymbol{v} & B\end{bmatrix}\right)\le\operatorname{rank}(B)+1$, 并且，根据高斯消元法，前一个等号成立当且仅当方程组 $B\boldsymbol{x}=\boldsymbol{v}$ 有解当且仅当 $\boldsymbol{v}\in\mathcal{R}(B)$。
   - $\boldsymbol{v}=B\boldsymbol{x}\in\mathcal{R}(B)$, 那么 $A=\begin{bmatrix}1 & -\boldsymbol{x}^{\mathrm{T}}\\ 0 & I_{m-1}\end{bmatrix}\begin{bmatrix}0 & 0\\ 0 & B\end{bmatrix}\begin{bmatrix}1 & 0\\ \boldsymbol{x} & I_{m-1}\end{bmatrix}$, 所以 $\operatorname{rank}(A)=\operatorname{rank}(B)$.
   - 若 $\boldsymbol{v}\notin\mathcal{R}(B)$, 注意到 $\operatorname{rank}\left(\begin{bmatrix}\boldsymbol{v} & B\end{bmatrix}\right)\le\operatorname{rank}(A)\le\operatorname{rank}\left(\begin{bmatrix}\boldsymbol{v} & B\end{bmatrix}\right)+1$. 我们要证明后一个等号成立, 或者等价地，前一个等号不成立。因为 $\operatorname{rank}(A)=\operatorname{rank}\left(\begin{bmatrix}-\boldsymbol{v}^{\mathrm{T}} & 0\\ B & \boldsymbol{v}\end{bmatrix}\right)$, $\operatorname{rank}\left(\begin{bmatrix}\boldsymbol{v} & B\end{bmatrix}\right)=\operatorname{rank}\left(\begin{bmatrix}\boldsymbol{v}^{\mathrm{T}}\\ -B\end{bmatrix}\right)$, 所以 $\operatorname{rank}\left(\begin{bmatrix}\boldsymbol{v} & B\end{bmatrix}\right)=\operatorname{rank}(A)$ 当且仅当方程组 $\begin{bmatrix}\boldsymbol{v}^{\mathrm{T}}\\ -B\end{bmatrix}\boldsymbol{x}=\begin{bmatrix}0\\ \boldsymbol{v}\end{bmatrix}$ 有解，当且仅当 $\begin{bmatrix}\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}\\ -B\boldsymbol{x}\end{bmatrix}=\begin{bmatrix}0\\ \boldsymbol{v}\end{bmatrix}$, 这说明 $\boldsymbol{v}=B(-\boldsymbol{x})\in\mathcal{R}(B)$, 矛盾。所以我们有
   $$\operatorname{rank}(A)=\operatorname{rank}\left(\begin{bmatrix}\boldsymbol{v} & B\end{bmatrix}\right)+1=\operatorname{rank}(B)+2.$$
3. 设 A 去掉首行首列得到的矩阵为 A_1, A_1 去掉首行首列得到的矩阵为 A_2，以此类推，我们有 $A_1,A_2,\cdots,A_{m-1}=\begin{bmatrix}0\end{bmatrix}$, 根据上一小题，$\operatorname{rank}(A),\operatorname{rank}(A_1),\cdots,\operatorname{rank}(A_m)$ 有相同的奇偶性；因为 $\operatorname{rank}(A_m)=0$，所以 $\operatorname{rank}(A)$ 是偶数。
   若 m 是奇数，那么 $\operatorname{rank}(A)\neq m$, 所以 A 不可逆。

□

练习 2.3.22. 设 A 是 n 阶可逆实反对称矩阵，$\boldsymbol{b}$ 是 n 维实列向量，求证：$\operatorname{rank}(A+\boldsymbol{b}\boldsymbol{b}^{\mathrm{T}})=n$.

证明: 令 $B=A+\boldsymbol{b}\boldsymbol{b}^{\mathrm{T}}$，我们证明齐次方程组 $B\boldsymbol{x}=0$ 没有非零解。假设 $B\boldsymbol{x}=0$, 那么 $A\boldsymbol{x}=-\boldsymbol{b}\boldsymbol{b}^{T}\boldsymbol{x}$。令 $c=\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}\in\mathbb{R}$, $-\boldsymbol{x}^{\mathrm{T}}\boldsymbol{b}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$, 即 $-c^2=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$. 又有

$$(\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x})^{\mathrm{T}}=\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{x}=-\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x},$$

所以 $-c^2=c^2$, 因此 $c=0$. 于是 $A\boldsymbol{x}=-c\boldsymbol{b}=0$. A 可逆，所以 $\boldsymbol{x}=0$. 这就证明了 $B\boldsymbol{x}=0$ 只有零解。 □

练习 2.3.23.

1. 求向量 $\boldsymbol{u},\boldsymbol{v}$，使得 $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}=\begin{bmatrix}3 & 6 & 6\\ 1 & 2 & 2\\ 4 & 8 & 8\end{bmatrix}$.

2. 设 A 是秩为 $r>0$ 的 $m\times n$ 的矩阵，令 C 为 A 的主列按顺序组成的矩阵，则 C 有几行几列？令 R 为 A 的行简化阶梯形矩阵的非零行按顺序组成的矩阵，则 R 有几行几列？求证 $A=CR$.

3. 求证：任意秩为 $r>0$ 的 $m\times n$ 矩阵 A 可以分解成列满秩矩阵和行满秩矩阵的乘积，即分别存在 $m\times r, r\times n$ 矩阵 C,R，且 $\operatorname{rank}(C)=\operatorname{rank}(R)=r$，使得 $A=CR$.
4. 证明，任意线性映射 f 都存在分解 $f=g\circ h$，其中 h 是线性满射，g 是线性单射.

证明:

1. 秩 1 矩阵我们讨论过很多次了，具体过程略过。$\boldsymbol{u}=c\begin{bmatrix}3\\1\\4\end{bmatrix}, \boldsymbol{v}=c^{-1}\begin{bmatrix}1\\2\\2\end{bmatrix}$, 其中 $c\in\mathbb{R}^{\times}$.
2. $\operatorname{rank}(A)=r$，所以 A 有 r 个主列，因此 C 有 m 行 r 列。A 的行简化阶梯形矩阵有 r 行，因此 R 有 r 行 n 列。
3. 我们对 A 的列数作归纳。当 A 只有一列时，若 $A=0$, 则结论是显然的；若 $A\neq 0$，那么 $C=A, R=\begin{bmatrix}1\end{bmatrix}$, 结论也是明显的。现在假设命题对任意 $m\times n$ 的矩阵成立，并且 $A\in M_{m\times(n+1)}(\mathbb{R})$. 我们把 A 分块：$A=\begin{bmatrix}A' & \boldsymbol{a}\end{bmatrix}$. 那么由归纳假设，$A'=C'R'$，$C',R'$ 的含义如题意。若 A 的最后一列是主列，那么 $C=\begin{bmatrix}C' & \boldsymbol{a}\end{bmatrix}, R=\begin{bmatrix}R' & 0\\0 & 1\end{bmatrix}$，那么 $CR=\begin{bmatrix}C' & \boldsymbol{a}\end{bmatrix}\begin{bmatrix}R' & 0\\0 & 1\end{bmatrix}=\begin{bmatrix}C'R' & \boldsymbol{a}\end{bmatrix}=A$; 若 A 的最后一列不是主列，那么 $C=C'$，且存在 $\boldsymbol{x}\in\mathbb{R}^r$, 使得 $\boldsymbol{a}=C\boldsymbol{x}$；那么 $A=C\begin{bmatrix}R' & \boldsymbol{x}\end{bmatrix}$. 我们还需要证明 $R=\begin{bmatrix}R' & \boldsymbol{x}\end{bmatrix}$ 是 A 的行简化阶梯形矩阵的非零行按顺序组成的矩阵；或者等价地要证明，我们在 R 的下方添加上 $m-r$ 个零行得到的矩阵 $\tilde{R}$ 是 A 的行简化阶梯形矩阵。事实上，我们把 C 的列向量扩张成 $\mathbb{R}^m$ 的一组基，把新得到的向量排在 A 的右边，得到了矩阵 $P=\begin{bmatrix}C & D\end{bmatrix}$, 那么 P 是可逆矩阵，且 $P\tilde{R}=\begin{bmatrix}C & D\end{bmatrix}\begin{bmatrix}R\\0\end{bmatrix}=CR=A$, 即有 $\tilde{R}=P^{-1}A$. 在 A 的左边乘以可逆矩阵，相当于对 A 作了一系列的初等行变换；又注意到 $\tilde{R}$ 已经是行简化阶梯形矩阵，由行简化阶梯形矩阵的唯一性知，$\tilde{R}$ 就是 A 对应的行简化阶梯形矩阵。至此归纳完成，证毕。
4. 这是上一小题的推论。
5. 假设线性映射 $f:\mathbb{R}^n\to\mathbb{R}^m$ 对应了矩阵 A, 那么我们按照上一小题，把 A 分解为 $A=CR$, 其中 C 列满秩，R 行满秩。令 $g:\mathbb{R}^r\to\mathbb{R}^m, h:\mathbb{R}^n\to\mathbb{R}^r$ 分别对应了矩阵 C,R，那么就有 $f=g\circ h$.

□

练习 2.3.24. 证明，任意秩为 $r>0$ 的矩阵 A 可以分解成 r 个秩为 1 的矩阵的和.

证明: 我们把 A 按照上一题作满秩分解 $A=CR$, 其中 C 列满秩，R 行满秩。我们把它们写成：

$$C=\begin{bmatrix}\boldsymbol{a}_1 & \cdots & \boldsymbol{a}_r\end{bmatrix}, R=\begin{bmatrix}\boldsymbol{b}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{b}_r^{\mathrm{T}}\end{bmatrix}$$

那么 $A=CR=\boldsymbol{a}_1\boldsymbol{b}_1^{\mathrm{T}}+\cdots+\boldsymbol{a}_r\boldsymbol{b}_r^{\mathrm{T}}$. 由 C,R 的满秩性可知 $\boldsymbol{a}_i,\boldsymbol{b}_j(1\le i,j\le r)$ 都不是零矩阵，因此 $\boldsymbol{a}_i\boldsymbol{b}_i^{\mathrm{T}}$ 都是秩 1 矩阵。 □

2.4 线性方程组的解集

练习 2.4.1. 求下列矩阵零空间的一组基：

1. $\begin{bmatrix}1&1&1&1&1\\3&2&1&1&-3\\0&1&2&2&6\\5&4&3&3&-1\end{bmatrix}$. *2.* $\begin{bmatrix}1&-2&1&1&-1\\2&1&-1&-1&-1\\1&7&-5&5&5\\3&-1&-2&1&-1\end{bmatrix}$. *3.* $\begin{bmatrix}2&1&-1&-1&1\\1&-1&1&1&-2\\3&3&-3&-3&4\\4&5&-5&-5&7\end{bmatrix}$.

求矩阵零空间的一组基步骤为：

1. 对矩阵做初等行变换，转化为行（简化）阶梯形矩阵，找出主元和自由变元。
2. 逐个地令一个自由变元为 1，其它自由变元为 0，求出方程组相应的解。

这样我们得到的解就是零空间的一组基。

证明: 以下我们仅对第一小题作详细的展示。其他小题仅写出答案。

1. 原矩阵的行简化阶梯形矩阵为 $\begin{bmatrix}1&0&-1&-1&-5\\0&1&2&2&6\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix}$ 因此主元是 x_1,x_2, 自由变元是 x_3,x_4,x_5.
   - 令 $x_3=1,x_4=x_5=0$, 解得 $x_1=1,x_2=-2$, 得到一个解 $\begin{bmatrix}1&-2&1&0&0\end{bmatrix}^{\mathrm{T}}$.
   - 令 $x_4=1,x_3=x_5=0$, 解得 $x_1=1,x_2=-2$, 得到一个解 $\begin{bmatrix}1&-2&0&1&0\end{bmatrix}^{\mathrm{T}}$.
   - 令 $x_5=1,x_3=x_4=0$, 解得 $x_1=5,x_2=-6$, 得到一个解 $\begin{bmatrix}5&-6&0&0&1\end{bmatrix}^{\mathrm{T}}$

   因此 $\begin{bmatrix}1\\-2\\1\\0\\0\end{bmatrix},\begin{bmatrix}1\\-2\\0\\1\\0\end{bmatrix},\begin{bmatrix}5\\-6\\0\\0\\1\end{bmatrix}$ 是原矩阵零空间的一组基。

2. 原矩阵的行简化阶梯形矩阵为 $\begin{bmatrix}1&0&0&0&-\frac{4}{7}\\0&1&0&0&\frac{2}{7}\\0&0&1&0&-\frac{2}{7}\\0&0&0&1&\frac{3}{7}\end{bmatrix}$ 因此 $\begin{bmatrix}\frac{4}{7}\\-\frac{2}{7}\\\frac{2}{7}\\-\frac{3}{7}\\1\end{bmatrix}$ 是原矩阵零空间的一组基。

3. 原矩阵的行简化阶梯形矩阵为 $\begin{bmatrix}1&0&0&0&0\\0&1&-1&-1&\frac{5}{3}\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix}$ 因此 $\begin{bmatrix}0\\1\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\\1\\0\end{bmatrix},\begin{bmatrix}\frac{1}{3}\\-\frac{5}{3}\\0\\0\\1\end{bmatrix}$ 是原矩阵零空间的一组基。

□

练习 2.4.2. 求如下线性方程组的全部解：

$$\begin{cases} 2x_1 - 2x_2 + x_3 - x_4 + x_5 = 1, \\ x_1 + 2x_2 - x_3 + x_4 - 2x_5 = 1, \\ 4x_1 - 10x_2 + 5x_3 - 5x_4 + 7x_5 = 1. \end{cases}$$

求方程组的全部解也有固定的程序：先求一个特解 $\boldsymbol{x}_0$，然后求系数矩阵 A 零空间 $\mathcal{N}(A)$ 的一组基，那么方程组的全部解就是 $\{\boldsymbol{x}_0 + \boldsymbol{x} : \boldsymbol{x} \in \mathcal{N}(A)\}$. 我们上一题已经总结了了求 $\mathcal{N}(A)$ 的基的方法；求特解的方法是对增广矩阵作初等行变换，然后令所有的自由变元全为零。

证明: 方程组的增广矩阵为

$$\left[\begin{array}{ccccc|c} 2 & -2 & 1 & -1 & 1 & 1 \\ 1 & 2 & -1 & 1 & -2 & 1 \\ 4 & -10 & 5 & -5 & 7 & 1 \end{array}\right]$$

它的行简化阶梯形矩阵为

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & -\frac{1}{3} & \frac{2}{3} \\ 0 & 1 & -\frac{1}{2} & \frac{1}{2} & -\frac{5}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

主元是 x_1, x_2, 我们令 $x_3 = x_4 = x_5 = 0$, 得到 $x_1 = \frac{2}{3}, x_2 = \frac{1}{6}$. 我们考虑齐次方程组，然后逐个令一个自由变元为 1, 其它自由变元是 0, 得到 $\mathcal{N}(A)$ 的一组基

$$\begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -\frac{1}{2} \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{3} \\ \frac{5}{6} \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

因此原方程组的全部解为

$$\left\{ \begin{bmatrix} \frac{2}{3} \\ \frac{1}{6} \\ 0 \\ 0 \\ 0 \end{bmatrix} + a_1 \begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \\ 0 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ -\frac{1}{2} \\ 0 \\ 1 \\ 0 \end{bmatrix} + a_3 \begin{bmatrix} \frac{1}{3} \\ \frac{5}{6} \\ 0 \\ 0 \\ 1 \end{bmatrix} : a_1, a_2, a_3 \in \mathbb{R} \right\}$$

□

练习 2.4.3. 求下列矩阵零空间的一组基：

1. $\begin{bmatrix} I_n & I_n \end{bmatrix}$. *2.* $\begin{bmatrix} I_n & I_n \\ 0 & 0 \end{bmatrix}$. *3.* $\begin{bmatrix} I_n & I_n & I_n \end{bmatrix}$.

证明: 1. 矩阵已经是行简化阶梯形矩阵了，$x_1, \cdots, x_n$ 是主元，$x_{n+1}, \cdots, x_{2n}$ 是自由变元。矩阵零空间的一组基为 $\{\boldsymbol{e}_1 - \boldsymbol{e}_{n+1}, \boldsymbol{e}_2 - \boldsymbol{e}_{n+2}, \cdots, \boldsymbol{e}_n - \boldsymbol{e}_{2n}\}$.

2. 矩阵已经是行简化阶梯形矩阵了，跟上一小题相比，只多了一些零行，相当于多了一些 $0=0$ 的方程，因此方程组的解没有变化。$x_1,\cdots,x_n$ 是主元，$x_{n+1},\cdots,x_{2n}$ 是自由变元。矩阵零空间的一组基为 $\{\boldsymbol{e}_1-\boldsymbol{e}_{n+1},\boldsymbol{e}_2-\boldsymbol{e}_{n+2},\cdots,\boldsymbol{e}_n-\boldsymbol{e}_{2n}\}$.
3. 矩阵已经是行简化阶梯形矩阵了，$x_1,\cdots,x_n$ 是主元，$x_{n+1},\cdots,x_{3n}$ 是自由变元。矩阵零空间的一组基为 $\{\boldsymbol{e}_1-\boldsymbol{e}_{n+1},\boldsymbol{e}_2-\boldsymbol{e}_{n+2},\cdots,\boldsymbol{e}_n-\boldsymbol{e}_{2n},\boldsymbol{e}_1-\boldsymbol{e}_{2n+1},\boldsymbol{e}_2-\boldsymbol{e}_{2n+2},\cdots,\boldsymbol{e}_n-\boldsymbol{e}_{3n}\}$.

□

练习 2.4.4. 给定 $A=\begin{bmatrix}1&&\\6&1&\\9&8&1\end{bmatrix}\begin{bmatrix}1&2&3&4\\&1&2&3\\&&1&2\end{bmatrix}$，求 A 的零空间、列空间、行空间、左零空间的一组基.

证明: 我们已经对 A 作了 LU 分解。我们记 $L=\begin{bmatrix}1&&\\6&1&\\9&8&1\end{bmatrix}$,$U=\begin{bmatrix}1&2&3&4\\&1&2&3\\&&1&2\end{bmatrix}$,

- A 的零空间 $\mathcal{N}(A)=\mathcal{N}(U)=\operatorname{Span}\left(\begin{bmatrix}0\\1\\-2\\1\end{bmatrix}\right)$
- A 有 3 个主元，因此 $\operatorname{rank}(A)=3$，而 $\mathcal{R}(A)\subseteq\mathbb{R}^3$, 所以 $\mathcal{R}(A)=\mathbb{R}^3$。基的选取就很自由了，例如 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$
- 作初等行变换不改变矩阵的行空间，所以 A 的行空间是 $\operatorname{Span}\left(\begin{bmatrix}1\\2\\3\\4\end{bmatrix},\begin{bmatrix}0\\1\\2\\3\end{bmatrix},\begin{bmatrix}0\\0\\1\\2\end{bmatrix}\right)$
- $\operatorname{rank}(A^{\mathrm{T}})=\operatorname{rank}(A)=3$, 而 $A^{\mathrm{T}}\boldsymbol{x}=0$ 只有三个变元，因此 A 的左零空间是 0.

□

练习 2.4.5. 线性方程组 $A\boldsymbol{x}=\begin{bmatrix}1\\3\end{bmatrix}$ 的全部解是 $\boldsymbol{x}=\begin{bmatrix}1\\0\end{bmatrix}+k\begin{bmatrix}0\\1\end{bmatrix}$,$k\in\mathbb{R}$，求 A.

证明: 因为 $\begin{bmatrix}0\\1\end{bmatrix}$ 是 $A\boldsymbol{x}=0$ 的一组基，因此 A 形如 $\begin{bmatrix}a&0\\b&0\end{bmatrix}$, 再利用 $A\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}1\\3\end{bmatrix}$ 知 $a=1,b=3$. □

练习 2.4.6. 求常数 a,b,c，使得方程 $\begin{bmatrix}1&-3&-1\end{bmatrix}\boldsymbol{x}=12$ 的所有解都具有如下形式：

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}a\\0\\0\end{bmatrix}+\begin{bmatrix}b\\1\\0\end{bmatrix}x_2+\begin{bmatrix}c\\0\\1\end{bmatrix}x_3.$$

证明: 由题意，$\begin{bmatrix} b \\ 1 \\ 0 \end{bmatrix}$ 和 $\begin{bmatrix} c \\ 0 \\ 1 \end{bmatrix}$ 是方程组 $\begin{bmatrix} 1 & -3 & -1 \end{bmatrix} \boldsymbol{x} = 0$ 的解的一组基，所以

$$b-3=0, c-1=0,\ \text{即} b=3, c=1.$$

同时，$\begin{bmatrix} 1 & -3 & -1 \end{bmatrix} \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix} = 12$, 所以 $a=12$. □

练习 2.4.7. 设 3×4 矩阵 A 的零空间的一组基是 $\begin{bmatrix} 2 \\ 3 \\ 1 \\ 0 \end{bmatrix}$.

1. 求 $\operatorname{rank}(A)$ 和 $\operatorname{rref}(A)$.
2. 线性方程组 $A\boldsymbol{x} = \boldsymbol{b}$ 对哪些 $\boldsymbol{b}$ 有解?

证明:

1. 按题意，A 只有一个自由变元，有三个主元。若 x_1 是自由变元，则零空间的一组基形如 $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$；若 x_2 是自由变元，则零空间的一组基形如 $\begin{bmatrix} * \\ 1 \\ 0 \\ 0 \end{bmatrix}$；若 x_4 是自由变元，则零空间的一组基形如 $\begin{bmatrix} * \\ \\ * \\ 1 \end{bmatrix}$；这三种向量和给定的向量都不线性相关，因此矛盾，从而 x_3 是自由变元，x_1, x_2, x_4 是主元，因此 $\operatorname{rref}(A)$ 形如

$$\begin{bmatrix} 1 & 0 & a & 0 \\ 0 & 1 & b & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

而 $\operatorname{rref}(A) \begin{bmatrix} 2 \\ 3 \\ 1 \\ 0 \end{bmatrix} = 0$, 于是 $a=-2, b=-3$. 这就得到了

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

2. 因为 $\mathcal{R}(A) \subseteq \mathbb{R}^3$ 且 $\dim \mathcal{R}(A) = \operatorname{rank}(A) = 3$, 所以 $\mathcal{R}(A) = \mathbb{R}^3$. 那么，对任意 $\boldsymbol{b} \in \mathbb{R}^3$, $A\boldsymbol{x} = \boldsymbol{b}$ 都有解。

□

练习 2.4.8. 设 $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 1 & 0 & 0 & 0 \end{bmatrix}$, 对任意 $\boldsymbol{v} \in \mathbb{R}^4$, 令 $S_{\boldsymbol{v}} := \{\boldsymbol{v} + \boldsymbol{x}_0 \mid \boldsymbol{x}_0 \in \mathcal{N}(A)\}$, 即由 $\mathcal{N}(A)$ 沿着 $\boldsymbol{v}$ 平移后得到的子集. 令 $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 6 \end{bmatrix}$, 求 $\boldsymbol{b}_i\,(i = 1, 2, 3)$ 使得 $S_{\boldsymbol{v}_i}$ 是 $A\boldsymbol{x} = \boldsymbol{b}_i$ 的解集，并分析 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ 之间的关系.

证明: 对任意 $\boldsymbol{w} \in S_{\boldsymbol{v}}$, $A\boldsymbol{w} = A(\boldsymbol{v} + \boldsymbol{x}_0) = A\boldsymbol{v}$. 因此 $\boldsymbol{b}_i = A\boldsymbol{v}_i$, 即得

$$\boldsymbol{b}_1 = \begin{bmatrix} 5 \\ 17 \\ 1 \end{bmatrix}, \boldsymbol{b}_2 = \begin{bmatrix} 11 \\ 23 \\ 0 \end{bmatrix} \boldsymbol{b}_1 = \begin{bmatrix} 38 \\ 86 \\ 1 \end{bmatrix}$$

另一方面，$\operatorname{rank}(A) = 3$，那么 A 有三个主元，一个自由变元，从而 $A\boldsymbol{x} = \boldsymbol{b}$ 的解都是某个 $S_{\boldsymbol{v}}$。故以上所求 $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$ 就是答案。

观察一下就知道 $\boldsymbol{v}_3 = \boldsymbol{v}_1 + 3\boldsymbol{v}_2$. □

练习 2.4.9. 把国际象棋的棋盘以及棋子的初始位置分别抽象成如下矩阵 B 和 C:

$$B = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} r & n & b & q & k & b & n & r \\ p & p & p & p & p & p & p & p \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p & p & p & p & p & p & p & p \\ r & n & b & q & k & b & n & r \end{bmatrix}.$$

分别求其列空间、行空间、零空间和左零空间的一组基，其中 r, n, b, q, k, p 是两两不等的非零实数.

证明: B 是实对称矩阵，其列空间和行空间相同，零空间和左零空间相同。对 B 作初等行变换

有

$$\mathrm{rref}(B)=\begin{bmatrix}1&0&1&0&1&0&1&0\\0&1&0&1&0&1&0&1\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{bmatrix}$$

因此，

$$\mathcal{R}(B)=\mathrm{Span}\left(\begin{bmatrix}1\\0\\1\\0\\1\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\\1\\0\\1\\0\\1\end{bmatrix}\right),\mathcal{N}(B)=\mathrm{Span}\left(\boldsymbol{e}_1-\boldsymbol{e}_3,\boldsymbol{e}_1-\boldsymbol{e}_5,\boldsymbol{e}_1-\boldsymbol{e}_7,\boldsymbol{e}_2-\boldsymbol{e}_4,\boldsymbol{e}_2-\boldsymbol{e}_6,\boldsymbol{e}_2-\boldsymbol{e}_8\right)$$

我们通过观察知 $\mathrm{rank}(C)=2$，其第一二行线性无关，第一二列线性无关。因此第一二行是行空间的一组基，第一二列是列空间的一组基。

$$\mathrm{rref}(C)=\begin{bmatrix}1&0&\frac{n-b}{n-r}&\frac{n-q}{n-r}&\frac{n-k}{n-r}&\frac{n-b}{n-r}&0&1\\0&1&\frac{b-r}{n-r}&\frac{q-r}{n-r}&\frac{k-r}{n-r}&\frac{b-r}{n-r}&1&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{bmatrix}$$

$\mathcal{N}(C)$ 的一组基是

$$\begin{bmatrix}-\frac{n-b}{n-r}\\-\frac{b-r}{n-r}\\1\\0\\0\\0\\0\\0\end{bmatrix},\begin{bmatrix}-\frac{n-q}{n-r}\\-\frac{q-r}{n-r}\\0\\1\\0\\0\\0\\0\end{bmatrix},\begin{bmatrix}-\frac{n-k}{n-r}\\-\frac{k-r}{n-r}\\0\\0\\1\\0\\0\\0\end{bmatrix},\begin{bmatrix}-\frac{n-b}{n-r}\\-\frac{b-r}{n-r}\\0\\0\\0\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\-1\\0\\0\\0\\0\\1\\0\end{bmatrix},\begin{bmatrix}-1\\0\\0\\0\\0\\0\\0\\1\end{bmatrix}$$

$$\text{rref}(C^{\mathrm{T}}) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\mathcal{N}(C^{\mathrm{T}})$ 的一组基是

$$\boldsymbol{e}_1 - \boldsymbol{e}_8, \boldsymbol{e}_2 - \boldsymbol{e}_7, \boldsymbol{e}_3, \boldsymbol{e}_4, \boldsymbol{e}_5, \boldsymbol{e}_6$$

□

练习 2.4.10. 设 A, B, C, D 为二阶方阵，如果分块矩阵 $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ 满足每一行、每一列以及四个二阶子方阵中的四个元素都是 $1, 2, 3, 4$，则称 M 为四阶数独矩阵. 写出一个四阶数独矩阵，并分别求其列空间、行空间、零空间和左零空间的一组基.

证明: 这样的数独矩阵有很多，我们给出一例进攻参考：$M = \begin{bmatrix} 1 & 3 & 2 & 4 \\ 2 & 4 & 1 & 3 \\ 3 & 1 & 4 & 2 \\ 4 & 2 & 3 & 1 \end{bmatrix}$. M 是对称矩阵，所以它的行空间和列空间相同，左零空间和右零空间相同。

$$\text{rref}(M) = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

主元列是第 1,2,3 列，因此

$$\mathcal{R}(M) = \text{Span}\left(\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 4 \\ 3 \end{bmatrix} \right), \mathcal{N}(M) = \text{Span}\left(\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \right)$$

□

练习 2.4.11. 在平面直角坐标系下给定点 $A(a_1, a_2), B(b_1, b_2), C(c_1, c_2)$，证明，$A, B, C$ 三点不共线当且仅当矩阵 $\begin{bmatrix} a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \\ c_1 & c_2 & 1 \end{bmatrix}$ 可逆.

证明: 平面上直线方程的一般形式为 $\lambda X+\mu Y+\nu=0$, 其中 λ,μ,ν 不同时为零。那么 A,B,C 三点共线，当且仅当方程组

$$\begin{cases} a_1\lambda+a_2\mu+\nu &=0\\ b_1\lambda+b_2\mu+\nu &=0\\ c_1\lambda+c_2\mu+\nu &=0\end{cases}$$

有非零解，当且仅当矩阵 $\begin{bmatrix} a_1 & a_2 & 1\\ b_1 & b_2 & 1\\ c_1 & c_2 & 1\end{bmatrix}$ 不可逆。 □

练习 2.4.12. 设 $\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_t$ 是线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 的解，其中 $\boldsymbol{b}\neq 0$，证明，$c_0\boldsymbol{x}_0+c_1\boldsymbol{x}_1+\cdots+c_t\boldsymbol{x}_t$ 也是解当且仅当 $c_0+c_1+\cdots+c_t=1$.

证明:

$$\begin{aligned} c_0\boldsymbol{x}_0+c_1\boldsymbol{x}_1+\cdots+c_t\boldsymbol{x}_t\text{ 也是解} &\Leftrightarrow A(c_0\boldsymbol{x}_0+c_1\boldsymbol{x}_1+\cdots+c_t\boldsymbol{x}_t)=\boldsymbol{b}\\ &\Leftrightarrow c_0A\boldsymbol{x}_0+c_1A\boldsymbol{x}_1+\cdots+c_tA\boldsymbol{x}_t=\boldsymbol{b}\\ &\Leftrightarrow c_0\boldsymbol{b}+c_1\boldsymbol{b}+\cdots+c_t\boldsymbol{b}=\boldsymbol{b}\\ &\Leftrightarrow c_0+c_1+\cdots+c_t=1\end{aligned}$$

□

练习 2.4.13. 设 $\boldsymbol{x}_0$ 是线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 的一个解，其中 $\boldsymbol{b}\neq 0$，而 $\boldsymbol{k}_1,\boldsymbol{k}_2,\ldots,\boldsymbol{k}_t$ 是 $\mathcal{N}(A)$ 的一组基. 令 $\boldsymbol{x}_i=\boldsymbol{x}_0+\boldsymbol{k}_i, i=1,2,\ldots,t$，证明，线性方程组的任意解都可唯一地表示成 $c_0\boldsymbol{x}_0+c_1\boldsymbol{x}_1+\cdots+c_t\boldsymbol{x}_t$，其中 $c_0+c_1+\cdots+c_t=1$.

证明: 设 $\boldsymbol{x}$ 是方程组的一个解，那么 $\boldsymbol{x}-\boldsymbol{x}_0\in\mathcal{N}(A)$，所以存在 $a_1,\cdots,a_t$ 使得

$$\boldsymbol{x}-\boldsymbol{x}_0=a_1\boldsymbol{k}_1+\cdots+a_t\boldsymbol{k}_t.$$

我们令 $c_0=1-(a_1+\cdots+a_t), c_1=a_1,\cdots,c_t=a_t$, 就有 $c_0+c_1+\cdots+c_t=1$ 且 $\boldsymbol{x}=c_0\boldsymbol{x}_0+c_1\boldsymbol{x}_1+\cdots+c_t\boldsymbol{x}_t$.

现在来证明唯一性。我们先断言：$\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_t$ 是线性无关的，那么若

$$c_0\boldsymbol{x}_0+c_1\boldsymbol{x}_1+\cdots+c_t\boldsymbol{x}_t=c_0'\boldsymbol{x}_0+c_1'\boldsymbol{x}_1+\cdots+c_t'\boldsymbol{x}_t,$$

即 $(c_0-c_0')\boldsymbol{x}_0+(c_1-c_1')\boldsymbol{x}_1+\cdots+(c_t-c_t')\boldsymbol{x}_t=0$, 线性无关性推出 $c_0-c_0'=c_1-c_1'=\cdots=c_t-c_t'=0$, 所以 $c_i=c_i',\forall i$, 这就是表示的唯一性。现在我们来证明线性无关性。假设 $a_0\boldsymbol{x}_0+a_1\boldsymbol{x}_1+\cdots+a_t\boldsymbol{x}_t=0$, 那么

$$(a_0+a_1+\cdots+a_t)\boldsymbol{x}_0+a_1\boldsymbol{k}_1+\cdots+a_t\boldsymbol{k}_t=0;$$

等号的左右两边同时左乘 A 即得 $(a_0+a_1+\cdots+a_t)\boldsymbol{b}=0$, 所以 $a_0+a_1+\cdots+a_t=0$; 从而有 $a_1\boldsymbol{k}_1+\cdots+a_t\boldsymbol{k}_t=0$, 利用 $\boldsymbol{k}_1,\boldsymbol{k}_2,\ldots,\boldsymbol{k}_t$ 的线性无关性知 $a_1=a_2=\cdots=a_t=0$, 进一步的有 $a_0=0$. 这就证明了 $\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_t$ 是线性无关的。 □

练习 2.4.14. 对任意 $\mathbb{R}^n$ 中线性无关的向量组 $\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_t$，证明，存在满足如下条件的非齐次线性方程组：

1. $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_t$ 都是此方程组的解；
2. 该方程组的任意解都能被 $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_t$ 线性表示.

证明: 令 $\boldsymbol{x}_i' = \boldsymbol{x}_i - \boldsymbol{x}_0$，那么 $\boldsymbol{x}_1' \cdots, \boldsymbol{x}_t'$ 线性无关；令 $X = \begin{bmatrix} \boldsymbol{x}_1' & \cdots & \boldsymbol{x}_t' \end{bmatrix}^{\mathrm{T}}$，则 $\operatorname{rank}(X) = t$，所以 $\dim \mathcal{N}(X) = n - t \geq 1$. 取 $\mathcal{N}(X)$ 的一组基 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{n-t} \in \mathbb{R}^n$, 令 $A = \begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_{n-t} \end{bmatrix}^{\mathrm{T}}$.

- $A\boldsymbol{x}_i' = 0$. 这是因为 $A\boldsymbol{x}_i = \begin{bmatrix} \boldsymbol{a}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{a}_{n-t}^{\mathrm{T}} \end{bmatrix} \boldsymbol{x}_i' = \begin{bmatrix} \boldsymbol{a}_1^{\mathrm{T}} \boldsymbol{x}_i' \\ \vdots \\ \boldsymbol{a}_{n-t}^{\mathrm{T}} \boldsymbol{x}_i' \end{bmatrix}$, 而由 $X\boldsymbol{a}_j = \begin{bmatrix} \boldsymbol{x}_1'^{\mathrm{T}} \boldsymbol{a}_j \\ \vdots \\ \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{a}_j \end{bmatrix} = 0$ 知 $\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{a}_j = 0, \forall 1 \leq i \leq t, 1 \leq j \leq n - t$, 从而 $A\boldsymbol{x}_i' = 0$.
- $\boldsymbol{b} := A\boldsymbol{x}_0 \neq 0$. 若不然, 则 $\boldsymbol{a}_j^{\mathrm{T}} \boldsymbol{x}_0 = 0, \forall 1 \leq j \leq n-t$, 从而 $\boldsymbol{a}_j^{\mathrm{T}} \boldsymbol{x}_i = \boldsymbol{a}_j^{\mathrm{T}} \boldsymbol{a}_i' + \boldsymbol{a}_j^{\mathrm{T}} \boldsymbol{x}_0 = 0$, 也就是说, $\boldsymbol{a}_j$ 都是方程组 $X'\boldsymbol{x} = 0$ 的解, 其中 $X' = \begin{bmatrix} \boldsymbol{x}_0 & \cdots & \boldsymbol{x}_t \end{bmatrix}^{\mathrm{T}}$. 而 $\operatorname{rank}(X') = t+1$, 于是
$$\dim \mathcal{N}(X) = \dim \operatorname{Span}(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{n-t}) \leq \dim \mathcal{N}(X') = n - (t+1)$$
矛盾。

现在我们考虑非齐次方程组 $A\boldsymbol{x} = \boldsymbol{b}$. 由上面的构造过程就知道这个齐次方程组的全部解为
$$\{\boldsymbol{x}_0 + a_1\boldsymbol{x}_1' + \cdots + a_t\boldsymbol{x}_t' : a_1, \cdots, a_t \in \mathbb{R}\},$$
从而都能被 $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_t$ 线性表示。 □

练习 2.4.15. 给定线性方程组 $A\boldsymbol{x} = \boldsymbol{b}$，和分块矩阵 $B = \begin{bmatrix} A & \boldsymbol{b} \\ \boldsymbol{b}^{\mathrm{T}} & 0 \end{bmatrix}$. 证明，如果 $\operatorname{rank}(A) = \operatorname{rank}(B)$，则方程组有解.

证明: 记 $B' = \begin{bmatrix} A \\ \boldsymbol{b}^{\mathrm{T}} \end{bmatrix}, \boldsymbol{b}' = \begin{bmatrix} \boldsymbol{b} \\ 0 \end{bmatrix}$, 那么 $B = \begin{bmatrix} B' & \boldsymbol{b}' \end{bmatrix}$. 注意到我们有不等式 $\operatorname{rank}(A) \leq \operatorname{rank}(B') \leq \operatorname{rank}(B)$. 如果 $\operatorname{rank}(A) = \operatorname{rank}(B)$, 那么这个不等式的两个不等号必须都取等号。

注意到 $\operatorname{rank}(B') = \operatorname{rank}(B)$ 当且仅当 $\boldsymbol{b}' \in \mathcal{R}(B')$, 即存在 $\boldsymbol{x}_0$ 使得 $\boldsymbol{b}' = B'\boldsymbol{x}_0$, 所以 $A\boldsymbol{x}_0 = \boldsymbol{b}$ 且 $\boldsymbol{b}^{\mathrm{T}} \boldsymbol{x}_0 = 0$; 前者说明 $A\boldsymbol{x} = \boldsymbol{b}$ 有解。 □

练习 2.4.16 (Fredholm 二择一定理)**.** 线性方程组 $A\boldsymbol{x} = \boldsymbol{b}$ 有解当且仅当 $\begin{bmatrix} A^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{bmatrix} \boldsymbol{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 无解.

附注. 前一个方程组中 $\boldsymbol{x}$ 为未知向量，后一个方程组中 $\boldsymbol{y}$ 为未知向量.

证明: 根据高斯消元法，$\begin{bmatrix} A^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{bmatrix} \boldsymbol{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 无解当且仅当 $\operatorname{rank}\left(\begin{bmatrix} A^{\mathrm{T}} & 0 \\ \boldsymbol{b}^{\mathrm{T}} & 1 \end{bmatrix}\right) = \operatorname{rank}\left(\begin{bmatrix} A^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{bmatrix}\right) + 1$，当且仅当 $\operatorname{rank}\left(\begin{bmatrix} A & \boldsymbol{b} \\ 0 & 1 \end{bmatrix}\right) = \operatorname{rank}\left(\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}\right) + 1$. 而 $\operatorname{rank}\left(\begin{bmatrix} A & \boldsymbol{b} \\ 0 & 1 \end{bmatrix}\right) = \operatorname{rank}(A) + 1$，所以 $\begin{bmatrix} A^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{bmatrix} \boldsymbol{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 无解当且仅当 $\operatorname{rank}(A) = \operatorname{rank}\left(\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}\right)$, 当且仅当（也是根据高斯消元法）$A\boldsymbol{x} = \boldsymbol{b}$ 有解。 □

练习 2.4.17. 如果 10 阶方阵 A 满足 $A^2=0$，证明 $\operatorname{rank}(A)\le 5$. 是否存在 $A^2=0, \operatorname{rank}(A)=5$ 的 10 阶方阵 A?

提示. 子空间 $\mathcal{N}(A)$ 和 $\mathcal{R}(A)$ 有何关系?

证明*:* $A^2=0$ 推出 $\mathcal{R}(A)\subseteq\mathcal{N}(A)$, 那么

$$2\operatorname{rank}(A)\le\operatorname{rank}(A)+\dim\mathcal{N}(A)=10$$

所以 $\operatorname{rank}(A)\le 5$.

存在 $A^2=0, \operatorname{rank}(A)=5$ 的 10 阶方阵 A，例如

$$\operatorname{diag}\left(\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix}\right)$$

□

练习 2.4.18. 设 A,B 分别为 $m\times n, n\times k$ 矩阵，证明，$\operatorname{rank}(AB)\ge\operatorname{rank}(A)+\operatorname{rank}(B)-n$.

证明*:*

(法一) 利用题 2.3.11，我们有 $\operatorname{rank}(AI_n)+\operatorname{rank}(I_nB)\le\operatorname{rank}(AI_nB)+\operatorname{rank}(I_n)$.

(法二) 利用维数公式，这等价于证明

$$k-\dim\mathcal{N}(AB)\ge(n-\dim\mathcal{N}(A))+(k-\dim\mathcal{N}(B))-n$$

即 $\dim\mathcal{N}(A)+\dim\mathcal{N}(B)\ge\dim\mathcal{N}(AB)$. 我们注意到 $\mathcal{N}(B)\subseteq\mathcal{N}(AB)$, 那么取 $\mathcal{N}(B)$ 的一组基 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r$, 扩张成为 $\mathcal{N}(AB)$ 的一组基 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r,\boldsymbol{a}_{r+1},\cdots,\boldsymbol{a}_{r+t}$; 那我们只需要证明 $\dim\mathcal{N}(A)\ge t$. 考虑 $\boldsymbol{b}_1=B\boldsymbol{a}_{r+1},\cdots,\boldsymbol{b}_t=B\boldsymbol{a}_{r+t}$, 因为 $A\boldsymbol{b}_i=AB\boldsymbol{a}_{r+i}=0$, 所以 $\boldsymbol{b}_i\in\mathcal{N}(A)$. 那么我们只需要证明 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_t$ 是线性无关的, 就可以得到 $\dim\mathcal{N}(A)\ge t$. 假设 $y_1\boldsymbol{b}_1+\cdots+y_t\boldsymbol{b}_t=0$, 那么 $B(y_1\boldsymbol{a}_{r+1}+\cdots+y_t\boldsymbol{a}_{r+t})=0$, 于是 $y_1\boldsymbol{a}_{r+1}+\cdots+y_t\boldsymbol{a}_{r+t}\in\mathcal{N}(B)$. 因此存在 $x_1,\cdots,x_r$, 使得

$$x_1\boldsymbol{a}_1+\cdots+x_r\boldsymbol{a}_r=y_1\boldsymbol{a}_{r+1}+\cdots+y_t\boldsymbol{a}_{r+t}.$$

根据 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{r+t}$ 的线性无关性就有 $x_1=\cdots=x_r=y_1=\cdots=y_t=0$. 证毕。

□

练习 2.4.19. 对 n 阶方阵 A，求证：

1. $A^2=A$ 当且仅当 $\operatorname{rank}(A)+\operatorname{rank}(I_n-A)=n$.
提示. 考虑 A 和 $I-A$ 的相关的子空间有何关系?

2. $A^2=I_n$ 当且仅当 $\operatorname{rank}(I_n+A)+\operatorname{rank}(I_n-A)=n$.
提示. 考虑 $I+A$ 和 $I-A$ 的相关的子空间有何关系?

证明*:* 1. 考虑 $V=\mathcal{R}(A)$ 和 $W=\mathcal{R}(I_n-A)$. 对任意 $\boldsymbol{x}\in\mathbb{R}^n$, $\boldsymbol{x}=A\boldsymbol{x}+(I_n-A)\boldsymbol{x}$, 这说明 $\mathbb{R}^n=V+W$. 根据维数公式，就有 $n=\operatorname{rank}(A)+\operatorname{rank}(I_n-A)-\dim(V\cap W)$.

- 若 $A^2=A$, 设 $\boldsymbol{x}\in V\cap W$, 则存在 $\boldsymbol{v},\boldsymbol{w}\in\mathbb{R}^n$ 使得 $\boldsymbol{x}=A\boldsymbol{v}=(I_n-A)\boldsymbol{w}$, 于是

$$(I_n-A)\boldsymbol{x}=(I_n-A)A\boldsymbol{v}=0,\quad A\boldsymbol{x}=A(I_n-A)\boldsymbol{w}=0,$$

所以 $\boldsymbol{x}=(I_n-A)\boldsymbol{x}+A\boldsymbol{x}=0$. 这就说明 $V\cap W=\{0\}$.

- 若 $\operatorname{rank}(A)+\operatorname{rank}(I_n-A)=n$，则 $\dim\mathcal{N}(A)+\dim\mathcal{N}(I_n-A)=n$. 设 $\boldsymbol{x}\in\mathcal{N}(A)\cap\mathcal{N}(I_n-A)$, 则 $A\boldsymbol{x}=(I_n-A)\boldsymbol{x}=0$, 所以 $\boldsymbol{x}=A\boldsymbol{x}+(I_n-A)\boldsymbol{x}=0$, 因此 $\mathcal{N}(A)\cap\mathcal{N}(I_n-A)=\{0\}$. 根据维数公式，有

$$\dim(\mathcal{N}(A)+\mathcal{N}(I_n-A))=\dim\mathcal{N}(A)+\dim\mathcal{N}(I_n-A)=n,$$

所以 $\mathcal{N}(A)+\mathcal{N}(I_n-A)=\mathbb{R}^n$. 从而对任意 $\boldsymbol{x}\in\mathbb{R}^n$, 我们把它写成 $\boldsymbol{x}=\boldsymbol{x}_1+\boldsymbol{x}_2$，其中 $\boldsymbol{x}_1\in\mathcal{N}(A),\boldsymbol{x}_2=\mathcal{N}(I_n-A)$. 那么

$$(A^2-A)\boldsymbol{x}=(A^2-A)\boldsymbol{x}_1+(A^2-A)\boldsymbol{x}_2=(I_n-A)A\boldsymbol{x}_1+A(I_n-A)\boldsymbol{x}_2=0.$$

这就证明了 $A^2-A=0$.

2. 考虑 $V=\mathcal{R}(I_n+A)$ 和 $W=\mathcal{R}(I_n-A)$. 对任意 $\boldsymbol{x}\in\mathbb{R}^n$,

$$\boldsymbol{x}=\frac{(I_n+A)\boldsymbol{x}}{2}+\frac{(I_n-A)\boldsymbol{x}}{2},$$

这说明 $\mathbb{R}^n=V+W$. 根据维数公式，就有 $n=\operatorname{rank}(A)+\operatorname{rank}(I_n-A)-\dim(V\cap W)$.

- 若 $A^2=I_n$, 设 $\boldsymbol{x}\in V\cap W$, 则存在 $\boldsymbol{v},\boldsymbol{w}\in\mathbb{R}^n$ 使得 $\boldsymbol{x}=(I_n+A)\boldsymbol{v}=(I_n-A)\boldsymbol{w}$, 于是

$$(I_n-A)\boldsymbol{x}=(I_n-A)(I_n+A)\boldsymbol{v}=0,\quad (I_n+A)\boldsymbol{x}=(I_n+A)(I_n-A)\boldsymbol{w}=0$$

, 所以 $\boldsymbol{x}=\frac{(I_n+A)\boldsymbol{x}}{2}+\frac{(I_n-A)\boldsymbol{x}}{2}=0$. 这就说明 $V\cap W=\{0\}$.

- 若 $\operatorname{rank}(I_n+A)+\operatorname{rank}(I_n-A)=n$，则 $\dim\mathcal{N}(I_n+A)+\dim\mathcal{N}(I_n-A)=n$. 设 $\boldsymbol{x}\in\mathcal{N}(I_n+A)\cap\mathcal{N}(I_n-A)$, 则 $(I_n+A)\boldsymbol{x}=(I_n-A)\boldsymbol{x}=0$, 所以 $\boldsymbol{x}=\frac{(I_n+A)\boldsymbol{x}}{2}+\frac{(I_n-A)\boldsymbol{x}}{2}=0$, 那么因此 $\mathcal{N}(I_n+A)\cap\mathcal{N}(I_n-A)=\{0\}$. 根据维数公式，有

$$\dim(\mathcal{N}(I_n+A)+\mathcal{N}(I_n-A))=\dim\mathcal{N}(I_n+A)+\dim\mathcal{N}(I_n-A)=n,$$

所以 $\mathcal{N}(A)+\mathcal{N}(I_n-A)=\mathbb{R}^n$. 从而对任意 $\boldsymbol{x}\in\mathbb{R}^n$, 我们把它写成 $\boldsymbol{x}=\boldsymbol{x}_1+\boldsymbol{x}_2$，其中 $\boldsymbol{x}_1\in\mathcal{N}(I_n+A),\boldsymbol{x}_2=\mathcal{N}(I_n-A)$. 那么

$$(I_n-A^2)\boldsymbol{x}=(I_n-A)(I_n+A)\boldsymbol{x}_1+(I_n+A)(I_n-A)\boldsymbol{x}_2=0.$$

这就证明了 $I_n-A^2=0$.

$\square$

练习 2.4.20. 证明或否定：如果对任意 $\boldsymbol{b}$，线性方程组 $A_1\boldsymbol{x}=\boldsymbol{b}$ 和 $A_2\boldsymbol{x}=\boldsymbol{b}$ 总有相同的解集，则 $A_1=A_2$.

证明: 这个命题我们在第一章就证明过了。在此，我们再证明一遍。对任意 i, 我们取 $\boldsymbol{b}$ 是 A_1 的第 i 个列向量, 那么 $\boldsymbol{e}_i$ 就是 $A_1\boldsymbol{x}=\boldsymbol{b}$ 的解, 那么它也是 $A_2\boldsymbol{x}=\boldsymbol{b}$ 的解, 也就是说 $A_2\boldsymbol{e}_i=A_1\boldsymbol{e}_i$, 即 A_2 和 A_1 的第 i 个列向量相同。由 i 的任意性知，A_2 每一个列向量都等于 A_1 相应的列向量，证毕。 □

练习 2.4.21. 证明，$\mathbb{R}^n$ 的任意子空间一定是某个矩阵的零空间.

证明: 设 $V\subseteq\mathbb{R}^n$ 是一个子空间，我们证明的是勋在矩阵 A 使得 $V=\mathcal{N}(A)$.

若 $V=\{0\}$, 那么我们取 $A=I_n$ 即可；若 $V=\mathbb{R}^n$, 我们取 $A=0$ 即可。现在假设 V 是真子空间，我们取 V 的一组基 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r(0<r<n)$, 并记 $B=\begin{bmatrix}\boldsymbol{v}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{v}_r^{\mathrm{T}}\end{bmatrix}$, 那么 $\dim\mathcal{N}(B)=n-r$, 我们取 $\mathcal{N}(B)$ 的一组基 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{n-r}$, 记 $A=\begin{bmatrix}\boldsymbol{x}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{x}_{n-r}^{\mathrm{T}}\end{bmatrix}$, 下面我们来证明 $V=\mathcal{N}(A)$. 我们需要注意到 $\operatorname{rank}(A)=\dim\mathcal{N}(B)=n-\operatorname{rank}(B)=n-r$.

- 按照我们的构造，$\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{x}_j=0(\forall i,j)$, 即 $\boldsymbol{x}_j^{\mathrm{T}}\boldsymbol{v}_i=0$, 这说明 $\boldsymbol{v}_i\in\mathcal{N}(A)$, 所以 $V\subseteq\mathcal{N}(A)$.
- 若 $\boldsymbol{v}\notin V$, 记 $\tilde{B}=\begin{bmatrix}\boldsymbol{v}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{v}_r^{\mathrm{T}}\\ \boldsymbol{v}^{\mathrm{T}}\end{bmatrix}$, 由 $\boldsymbol{v}$ 的选取, $\operatorname{rank}(\tilde{B})=\operatorname{rank}(B)+1$, 且 $\mathcal{N}(\tilde{B})\subseteq\mathcal{N}(B)$; 由维数公式, $\dim\mathcal{N}(\tilde{B})=\dim\mathcal{N}(B)-1$. 因此存在 $\boldsymbol{x}_0\in(\mathcal{N}(B)-\mathcal{N}(\tilde{B}))$. 这说明 $\boldsymbol{v}^T\boldsymbol{x}_0\neq 0$, 因此 $\boldsymbol{v}\notin\mathcal{N}(A)$。这就证明了 $\mathcal{N}(A)\subseteq V$.

□

练习 2.4.22. 给定 $l\times n$ 矩阵 A 和 $m\times n$ 矩阵 B，证明，$\mathcal{N}(A)\subseteq\mathcal{N}(B)$ 当且仅当存在 $m\times l$ 矩阵 C，使得 $B=CA$.

证明:

- 假设存在 $m\times l$ 矩阵 C，使得 $B=CA$。若 $A\boldsymbol{x}=0$，则 $B\boldsymbol{x}=CA\boldsymbol{x}=0$, 因此 $\mathcal{N}(A)\subseteq\mathcal{N}(B)$.
- 假设 $\mathcal{N}(A)\subseteq\mathcal{N}(B)$. 我们断言：$\boldsymbol{b}\in\mathcal{R}(B^{\mathrm{T}})$ 当且仅当对任意 $\boldsymbol{x}\in\mathcal{N}(B),\boldsymbol{b}\cdot\boldsymbol{x}=0$. 设 $B^{\mathrm{T}}=\begin{bmatrix}\boldsymbol{b}_1 & \cdots & \boldsymbol{b}_m\end{bmatrix}$，那么

$$B\boldsymbol{x}=\begin{bmatrix}\boldsymbol{b}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{b}_m^{\mathrm{T}}\end{bmatrix}\boldsymbol{x}=\begin{bmatrix}\boldsymbol{b}_1\cdot\boldsymbol{x}\\ \vdots\\ \boldsymbol{b}_m\cdot\boldsymbol{x}\end{bmatrix}$$

因此 $\boldsymbol{x}\in\mathcal{N}(B)$ 当且仅当 $\boldsymbol{b}_i\cdot\boldsymbol{x}=0,\forall 1\le i\le m$.

设 $\boldsymbol{b}\in\mathcal{R}(B^{\mathrm{T}})$, 则 $\boldsymbol{b}=c_1\boldsymbol{b}_1+\cdots+c_m\boldsymbol{b}_m$, 那么对任意 $\boldsymbol{x}\in\mathcal{N}(B)$, $\boldsymbol{b}\cdot\boldsymbol{x}=c_1\boldsymbol{b}_1\cdot\boldsymbol{x}+\cdots+c_m\boldsymbol{b}_m\cdot\boldsymbol{x}=0$. 反之，假设 $\boldsymbol{b}\notin\mathcal{R}(B^{\mathrm{T}})$, 令 $\tilde{B}^{\mathrm{T}}=\begin{bmatrix}\boldsymbol{b}_1 & \cdots & \boldsymbol{b}_m & \boldsymbol{b}\end{bmatrix}$, 那么 $\operatorname{rank}(\tilde{B})=\operatorname{rank}(B)+1$, 且 $\mathcal{N}(\tilde{B})\subseteq\mathcal{N}(B)$. 由维数公式,

$$\dim\mathcal{N}(\tilde{B})=n-\operatorname{rank}(\tilde{B})=n-\operatorname{rank}(B)-1=\dim\mathcal{N}(B)-1,$$

那么就存在 $\boldsymbol{x} \in \left(\mathcal{N}(B) - \mathcal{N}(\tilde{B})\right)$, 这说明 $\boldsymbol{x} \in \mathcal{N}(B)$ 且 $\boldsymbol{b} \cdot \boldsymbol{x} \neq 0$. 至此就证明了断言。
同理我们对 A 有 $\boldsymbol{a} \in \mathcal{R}(A^{\mathrm{T}})$ 当且仅当对任意 $\boldsymbol{x} \in \mathcal{N}(A), \boldsymbol{a} \cdot \boldsymbol{x} = 0$.
$\boldsymbol{b} \in \mathcal{R}(B^{\mathrm{T}})$, 则 $\boldsymbol{b} \cdot \boldsymbol{x} = 0, \forall \boldsymbol{x} \in \mathcal{N}(B)$, 则 $\boldsymbol{b} \cdot \boldsymbol{x} = 0, \forall \boldsymbol{x} \in \mathcal{N}(A)$, 则 $\boldsymbol{b} \in \mathcal{R}(A^{\mathrm{T}})$. 因此 B^{T} 的任意列向量都是 A^{T} 的列向量的线性组合，即存在 C^{T}，使得 $B^{\mathrm{T}} = A^{\mathrm{T}}C^{\mathrm{T}}$, 所以 $B = CA$.

□

练习 2.4.23. 给定 $m \times n$ 矩阵 A, B，证明，$\mathcal{N}(A) = \mathcal{N}(B)$ 当且仅当存在 m 阶可逆矩阵 T，使得 $B = TA$.

证明: 由上题的断言，$\mathcal{N}(A) = \mathcal{N}(B)$ 当且仅当 $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B^{\mathrm{T}})$.

- 若存在 m 阶可逆矩阵 T 使得 $B = TA$，容易的我们有 $\mathcal{N}(A) = \mathcal{N}(B)$.
- 现在假设 $\mathcal{N}(A) = \mathcal{N}(B)$；由上题，存在方阵 $C, D \in M_{m\times m}$, 使得 $B = CA, A = DB$, 那么由 $B = CDB$。若 B 都是行满秩的, 那么存在矩阵 P 使得 $BP = I_m$, 因此 $CD = I_m$, 即 C, D 可逆。我们取 $T = C$ 即可。
  若 B 不是行满秩的,那么 A 也不行满秩,假设可逆矩阵 P_1, P_2 使得 $P_1B = \mathrm{rref}(B), P_2A = \mathrm{rref}(A)$. 设 R_A, R_B 分别是 $\mathrm{rref}(A), \mathrm{rref}(B)$ 的非零行构成的矩阵。那么 R_A, R_B 行满秩，且 $\mathcal{N}(R_A) = \mathcal{N}(A) = \mathcal{N}(B) = \mathcal{N}(R_B)$, 于是存在可逆矩阵 $C' \in M_{r\times r}(\mathbb{R})$ 使得 $R_B = C'R_A$, 那么 $\mathrm{rref}(B) = \begin{bmatrix} C' & 0 \\ 0 & I_{m-r} \end{bmatrix} \mathrm{rref}(A)$. 即 $P_1B = \begin{bmatrix} C' & 0 \\ 0 & I_{m-r} \end{bmatrix} P_2A$, 因此我们令 $T = P_1^{-1} \begin{bmatrix} C' & 0 \\ 0 & I_{m-r} \end{bmatrix} P_2$ 即可。

□

练习 2.4.24. 给定 n 阶方阵 A.

1. 对任意 k，证明 $\mathcal{R}(A^k) \supseteq \mathcal{R}(A^{k+1})$；
2. 假设 $\mathcal{R}(A^k) = \mathcal{R}(A^{k+1})$，求证 $\mathcal{R}(A^{k+1}) = \mathcal{R}(A^{k+2})$；
3. 求证：存在 $k \leq n$，满足 $\mathrm{rank}(A^k) = \mathrm{rank}(A^{k+1}) = \cdots$. 由此证明，如果存在 p 使得 $A^p = 0$，则 $A^n = 0$.

证明:

1. 因为 $A^{k+1} = A^kA$, 因此 A^{k+1} 的列向量都是 A^k 的列向量的线性组合，因此 $\mathcal{R}(A^{k+1}) \subseteq \mathcal{R}(A^k)$.
2. 由列空间的包含关系，$\mathcal{R}(A^k) = \mathcal{R}(A^{k+1}) \Leftrightarrow \mathrm{rank}(A^k) = \mathrm{rank}(A^{k+1}) \Leftrightarrow \dim \mathcal{N}(A^k) = \dim \mathcal{N}(A^{k+1})$. 我们也有 $\mathcal{N}(A^k) \subseteq \mathcal{N}(A^{k+1})$, 因此 $\mathcal{R}(A^k) = \mathcal{R}(A^{k+1}) \Leftrightarrow \mathcal{N}(A^k) = \mathcal{N}(A^{k+1})$；这对任意 k 都成立。从而要证明 $\mathcal{R}(A^{k+1}) = \mathcal{R}(A^{k+2})$, 只需要证明 $\mathcal{N}(A^{k+1}) = \mathcal{N}(A^{k+2})$. $\mathcal{N}(A^{k+1}) \subseteq \mathcal{N}(A^{k+2})$ 已知；对任意 $\boldsymbol{x} \in \mathcal{N}(A^{k+2})$, 则 $A^{k+2}\boldsymbol{x} = 0$, 则 $A\boldsymbol{x} \in \mathcal{N}(A^{k+1})$, 则 $A\boldsymbol{x} \in \mathcal{N}(A^k)$, 则 $A^{k+1}\boldsymbol{x} = 0$, 即 $\boldsymbol{x} \in \mathcal{N}(A^{k+1})$；因此 $\mathcal{N}(A^{k+2}) \subseteq \mathcal{N}(A^{k+1})$.
3. 我们有不等式的链：$\mathrm{rank}(A) \geq \mathrm{rank}(A^2) \geq \cdots \geq \mathrm{rank}(A^k) \geq \mathrm{rank}(A^{k+1}) \geq \cdots$；而且上一小题已经证明了，若某一个 $\geq$ 取等号，那么从它往后的所有 $\geq$ 都取等号；所以这些 $\geq$ 中，最多只有开始的 $\mathrm{rank}(A)$ 个可能取严格的 $>$。考虑到 $\mathrm{rank}(A) \leq n$，若 $\mathrm{rank}(A) = n$，那么 A 可逆，且所有的 A^k 都可逆，因此 $\mathrm{rank}(A^k) = n$, 对任

意的 $k \geq 1$; 若 $\mathrm{rank}(A) < n$，那么至多有 n 个严格的 $>$，因此存在 $k \leq n$ 使得 $\mathrm{rank}(A^k) = \mathrm{rank}(A^{k+1}) = \cdots$。因此，若存在 p 使得 $A^p = 0$, 那么存在 $k \leq n$ 使得 $\mathrm{rank}(A^k) = \mathrm{rank}(A^p) = 0$, 即 $A^k = 0$, 那么 $A^n = A^k A^{n-k} = 0$.

□

练习 2.4.25 (Kirchhoff 电流定律)**.** 对练习 2.1.22中的电路网络，令 M_G 为对应的关联矩阵. 根据 *Kirchhoff* 电流定律，求 M_G 左零空间的一组基.

证明: 我们要求的就是 $M_G^T\boldsymbol{x} = 0$ 的解。在已知 M_G 的情况下，这仅仅是高斯消元法的实践操作而已，我们忽略具体过程，给出一组基仅供参考。

(a) x_6, x_7, x_8 是自由变元，M_G^{T} 零空间的一组基是

$$\begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

(b) x_6, x_7, x_8, x_9 是自由变元，M_G^{T} 零空间的一组基是

$$\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

(c) x_8, x_9 是自由变元，M_G^{T} 零空间的一组基是

$$\begin{bmatrix} 0 \\ 0 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

- $x_5, x_6, x_7, x_8, x_9, x_{10}$ 是自由变元，M_G^{T} 零空间的一组基是

$$\begin{bmatrix}1\\1\\1\\1\\1\\0\\0\\0\\0\\0\end{bmatrix},\begin{bmatrix}-1\\-1\\-1\\0\\0\\1\\0\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\1\\0\\0\\0\\1\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\-1\\-1\\-1\\0\\0\\0\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\\1\\0\\0\\0\\0\\1\\0\end{bmatrix}\begin{bmatrix}1\\1\\0\\0\\0\\0\\0\\0\\0\\1\end{bmatrix}$$

□

练习 2.4.26. *在例 2.4.7 中，有结论* $\operatorname{rank}(M_G^{\mathrm{T}}R^{-1}M_G) = \operatorname{rank}(M_G)$*，其中* R^{-1} *为对角元素都大于零的对角矩阵. 根据下列思路证明该结论：*

1. *若* $\boldsymbol{y}^{\mathrm{T}}R^{-1}\boldsymbol{y} = 0$*，则* $\boldsymbol{y} = 0$.
2. $M_G\boldsymbol{x} = 0$*，当且仅当* $M_G^{\mathrm{T}}R^{-1}M_G\boldsymbol{x} = 0$.
3. $\operatorname{rank}(M_G^{\mathrm{T}}R^{-1}M_G) = \operatorname{rank}(M_G)$.

证明: 1. 令 $R = \operatorname{diag}(r_1, r_2, \cdots, r_n)$, $r_i > 0$. 设 $\boldsymbol{y} = \begin{bmatrix}y_1 & \cdots & y_n\end{bmatrix}^{\mathrm{T}}$, 则 $\boldsymbol{y}^{\mathrm{T}}R^{-1}\boldsymbol{y} = \sum_{i=1}^n r_i^{-1}y_i^2$. 若 $\boldsymbol{y}^{\mathrm{T}}R^{-1}\boldsymbol{y} = 0$, 即 $\sum_{i=1}^n r_i^{-1}y_i^2 = 0$, 则 $y_1 = \cdots = y_n = 0$.

2. 当 $M_G\boldsymbol{x} = 0$ 时,显然有 $M_G^{\mathrm{T}}R^{-1}M_G\boldsymbol{x} = 0$. 反之,若 $M_G^{\mathrm{T}}R^{-1}M_G\boldsymbol{x} = 0$, 则 $(M_G\boldsymbol{x})^{\mathrm{T}}R^{-1}(M_G\boldsymbol{x}) = \boldsymbol{x}^T M_G^{\mathrm{T}}R^{-1}M_G\boldsymbol{x} = 0$. 由上一小题知 $M_G\boldsymbol{x} = 0$.

3. 上一小题实际上证明了 $\mathcal{N}(M_G) = \mathcal{N}(M_G^{\mathrm{T}}R^{-1}M_G)$, 由维数公式就可以得到 $\operatorname{rank}(M_G^{\mathrm{T}}R^{-1}M_G) = \operatorname{rank}(M_G)$.

□

第三章　内积和正交性

3.1　基 本 概 念

练习 3.1.1. 证明命题 *3.1.3:* 向量内积满足如下性质：

1. 对称：$\boldsymbol{a}^T\boldsymbol{b} = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}$;
2. 双线性：$\boldsymbol{a}^{\mathrm{T}}(k_1\boldsymbol{b}_1 + k_2\boldsymbol{b}_2) = k_1\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}_1 + k_2\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}_2, (k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2)^{\mathrm{T}}\boldsymbol{b} = k_1\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{b} + k_2\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{b}$.
3. 正定：$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} \geq 0$, 且 $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = 0$ 当且仅当 $\boldsymbol{a} = 0$.

证明*:*　1. 设 $\boldsymbol{a} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, 那么

$$\begin{aligned} \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} &= \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} &= x_1y_1 + \cdots + x_ny_n \\ \boldsymbol{b}^{\mathrm{T}}\boldsymbol{a} &= \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} &= y_1x_1 + \cdots + y_nx_n \end{aligned}$$

从而 $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{b}^T\boldsymbol{a}$.

2. 我们只证明前半个命题，后半个命题同理。我们设 $\boldsymbol{a} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \boldsymbol{b}_1 = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, , \boldsymbol{b}_2 = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$,

那么

$$\begin{aligned} \boldsymbol{a}^{\mathrm{T}}(k_1\boldsymbol{b}_1 + k_2\boldsymbol{b}_2) &= \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} k_1y_1 + k_2z_1 \\ \vdots \\ k_1y_n + k_2z_n \end{bmatrix} \\ &= x_1(k_1y_1 + k_2z_1) + \cdots + x_n(k_1y_n + k_2z_n) \\ &= k_1(x_1y_1 + \cdots + x_ny_n) + k_2(x_1z_1 + \cdots + x_nz_n) \\ &= k_1\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}_1 + k_2\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}_2 \end{aligned}$$

3. 设 $\boldsymbol{a} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, 那么 $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = x_1^2 + \cdots + x_n^2 \geq 0$, 且 $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = 0$ 当且仅当 $x_1^2 + \cdots + x_n^2 = 0$,

当且仅当 $x_1 = \cdots = x_n = 0$, 当且仅当 $\boldsymbol{a} = 0$.

$\square$

练习 3.1.2. 在 $\mathbb{R}^4$ 中求向量 $\boldsymbol{a}, \boldsymbol{b}$ 的夹角：

1. $\boldsymbol{a} = \begin{bmatrix} 2 \\ 1 \\ 3 \\ 2 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 1 \\ 2 \\ -2 \\ 1 \end{bmatrix}$. *2.* $\boldsymbol{a} = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 3 \\ 1 \\ 5 \\ 1 \end{bmatrix}$. *3.* $\boldsymbol{a} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 0 \end{bmatrix}$.

证明*:* 设 $\boldsymbol{a}, \boldsymbol{b}$ 的夹角为 θ

1. $\boldsymbol{a} \cdot \boldsymbol{b} = 2 \cdot 1 + 1 \cdot 2 + 3 \cdot (-2) + 2 \cdot 1 = 0$, 所以 $\theta = \frac{\pi}{2}$.
2. $\cos\theta = \dfrac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|\boldsymbol{a}\| \|\boldsymbol{b}\|} = \dfrac{3+2+10+3}{\sqrt{1+4+4+9} \cdot \sqrt{9+1+25+1}} = \dfrac{1}{\sqrt{2}}$, 因此 $\theta = \dfrac{\pi}{4}$.
3. $\cos\theta = \dfrac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|\boldsymbol{a}\| \|\boldsymbol{b}\|} = \dfrac{3+1+(-1)+0}{\sqrt{1+1+1+4} \cdot \sqrt{9+1+1+0}} = \dfrac{3}{\sqrt{77}}$, 因此 $\theta = \arccos\dfrac{3}{\sqrt{77}}$.

$\square$

练习 3.1.3. 求证：

1. 在 $\mathbb{R}^n$ 中的非零向量 $\boldsymbol{a}, \boldsymbol{b}$ 夹角为 0，当且仅当存在 $k > 0$，使得 $\boldsymbol{a} = k\boldsymbol{b}$.

2. 在 $\mathbb{R}^n$ 中的两向量 $\boldsymbol{a}, \boldsymbol{b}$ 正交，当且仅当对任意实数 t，有 $\|\boldsymbol{a} + t\boldsymbol{b}\| \geq \|\boldsymbol{a}\|$.

3. 在 $\mathbb{R}^n$ 中的非零向量 $\boldsymbol{a}, \boldsymbol{b}$ 正交，当且仅当 $\|\boldsymbol{a} + \boldsymbol{b}\| = \|\boldsymbol{a} - \boldsymbol{b}\|$.

证明*:* 设 $\boldsymbol{a} = \begin{bmatrix} x_1 \\ \cdots \\ x_n \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$

1. $\boldsymbol{a}, \boldsymbol{b}$ 夹角为 0 当且仅当 $\boldsymbol{a} \cdot \boldsymbol{b} = \|\boldsymbol{a}\| \|\boldsymbol{b}\|$, 当且仅当

$$(x_1 y_1 + \cdots + x_n y_n) = (x_1^2 + \cdots + x_n^2)^{1/2} (y_1^2 + \cdots + y_n^2)^{1/2}.$$

   根据 Cauchy-Schwartz 不等式取等号的条件，当且仅当存在 k 使得 $a_i = kb_i, \forall 1 \leq i \leq n$. 由于 $\boldsymbol{a} \cdot \boldsymbol{b} > 0$ 知 $k > 0$.

2.
   - 若 $\boldsymbol{a}, \boldsymbol{b}$ 正交，则任意实数 t，有 $\|\boldsymbol{a} + t\boldsymbol{b}\| = (\|\boldsymbol{a}\|^2 + \|t\boldsymbol{b}\|)^{1/2} \geq \|\boldsymbol{a}\|$.
   - 若 $\boldsymbol{a}, \boldsymbol{b}$ 不正交，则 $\boldsymbol{a} \cdot \boldsymbol{b} \neq 0$, 那么不等式 $2t\boldsymbol{a} \cdot \boldsymbol{b} + t^2 \boldsymbol{b} \cdot \boldsymbol{b} < 0$ 有解，我们任取一个解 t_0, 那么

$$\|\boldsymbol{a} + t_0 \boldsymbol{b}\| = (\|\boldsymbol{a}\|^2 + 2t_0 \boldsymbol{a} \cdot \boldsymbol{b} + t_0^2 \boldsymbol{b} \cdot \boldsymbol{b})^{1/2} < \|\boldsymbol{a}\|.$$

3. $\|\boldsymbol{a} + \boldsymbol{b}\| = \|\boldsymbol{a} - \boldsymbol{b}\| \Leftrightarrow (\boldsymbol{a} + \boldsymbol{b}) \cdot (\boldsymbol{a} + \boldsymbol{b}) = (\boldsymbol{a} - \boldsymbol{b}) \cdot (\boldsymbol{a} - \boldsymbol{b}) \Leftrightarrow \boldsymbol{a} \cdot \boldsymbol{b} = 0$.

$\square$

练习 3.1.4. 证明推论 *3.1.5*：$\|\boldsymbol{a} + \boldsymbol{b}\| \leq \|\boldsymbol{a}\| + \|\boldsymbol{b}\|$.

证明: 设 $\boldsymbol{a}=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}, \boldsymbol{b}=\begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}$, 则

$$\begin{aligned}\|\boldsymbol{a}+\boldsymbol{b}\| \leq \|\boldsymbol{a}\|+\|\boldsymbol{b}\| \quad &\Leftrightarrow \quad \sum_{i=1}^{n}(x_i+y_i)^2 \leq \left(\left(\sum_{i=1}^{n} x_i^2\right)^{1/2}+\left(\sum_{i=1}^{n} x_i^2\right)^{1/2}\right)^2 \\ &\Leftrightarrow \quad \sum_{i=1}^{n} x_i y_i \leq \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}\left(\sum_{i=1}^{n} x_i^2\right)^{1/2}\end{aligned}$$

由 Cauchy-Schwartz 不等式知命题成立。 □

练习 3.1.5 (Cauchy-Schwarz 不等式的其他证明)**.**

1. 先证明 $\boldsymbol{a}, \boldsymbol{b}$ 都是单位向量的情形：$\left|\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}\right| \leq 1$，且等号成立当且仅当 $\boldsymbol{a}=\pm\boldsymbol{b}$. 再由单位向量的情形推广到一般的情形.
2. 根据内积的正定性，对任意实数 t，都有 $(\boldsymbol{a}+t\boldsymbol{b})^{\mathrm{T}}(\boldsymbol{a}+t\boldsymbol{b})=\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}+2t\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}+t^2\boldsymbol{b}^{\mathrm{T}}\boldsymbol{b} \geq 0$. 利用判别式证明结论.

证明: 设 $\boldsymbol{a}=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}, \boldsymbol{b}=\begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}$,

1. 若 $\boldsymbol{a}, \boldsymbol{b}$ 都是单位向量，则 $\sum_{i=1}^{n} x_i^2=\sum_{i=1}^{n} y_i^2=1$.

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}=\sum_{i=1}^{n} x_i y_i \leq \sum_{i=1}^{n} \frac{x_i^2+y_i^2}{2}=1=\|\boldsymbol{a}\|\,\|\boldsymbol{b}\|\,.$$

   对于非零向量 $\boldsymbol{a}, \boldsymbol{b}$, 设 $\tilde{\boldsymbol{a}}=\dfrac{\boldsymbol{a}}{\|\boldsymbol{a}\|}, \tilde{\boldsymbol{b}}=\dfrac{\boldsymbol{b}}{\|\boldsymbol{b}\|}$，则 $\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}$ 都是单位向量，因此 $\tilde{\boldsymbol{a}}^{\mathrm{T}}\tilde{\boldsymbol{b}} \leq 1$, 即 $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} \leq \|\boldsymbol{a}\|\,\|\boldsymbol{b}\|$.
2. 这是一个关于 t 的二次多项式，由于它大于等于零恒成立，所以
   - 要么 $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{b}=\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}=0, \boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} \geq 0$，即 $\boldsymbol{b}=0$;
   - 要么 $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{b}>0, (2\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b})^2-4\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{b} \leq 0$, 即 $\boldsymbol{b} \neq 0$ 且 $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} \leq \|\boldsymbol{a}\|\,\|\boldsymbol{b}\|$.

□

练习 3.1.6. 给定 $\boldsymbol{a} \in \mathbb{R}^3$.计算 $\boldsymbol{a}$ 与坐标向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 的夹角的余弦**,**并计算这三个余弦值的平方和.

证明: 设 $\boldsymbol{a}=\begin{bmatrix}x\\ y\\ z\end{bmatrix}$, $\boldsymbol{a}$ 与 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 的夹角分别为 $\theta_1, \theta_2, \theta_3$. 则

$$\cos\theta_1=\frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{e}_1}{\|\boldsymbol{a}\|}=\frac{x}{\sqrt{x^2+y^2+z^2}}, \cos\theta_2=\frac{y}{\sqrt{x^2+y^2+z^2}}, \cos\theta_3=\frac{z}{\sqrt{x^2+y^2+z^2}}.$$

那么 $\cos^2\theta_1+\cos^2\theta_2+\cos^2\theta_3=1$. □

练习 3.1.7. 设 $\|\boldsymbol{a}\| = 3, \|\boldsymbol{b}\| = 4$，确定 $\|\boldsymbol{a}-\boldsymbol{b}\|$ 的取值范围.

证明: $\|\boldsymbol{a}-\boldsymbol{b}\|^2 = \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 - 2\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}$. 根据 Cauchy-Schwartz 不等式，

$$-\|\boldsymbol{a}\|\,\|\boldsymbol{b}\| \leq \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} \leq \|\boldsymbol{a}\|\,\|\boldsymbol{b}\|$$

因此 $9+16-24 \leq \|\boldsymbol{a}-\boldsymbol{b}\| \leq 9+16+24$, 即 $1 \leq \|\boldsymbol{a}-\boldsymbol{b}\| \leq 7$. □

练习 3.1.8. 设 $\boldsymbol{a} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} z \\ x \\ y \end{bmatrix}$，且 $x+y+z=0$. 确定 $\boldsymbol{a}, \boldsymbol{b}$ 夹角的取值范围.

证明: 设 $\boldsymbol{a}, \boldsymbol{b}$ 的夹角为 θ. 那么

$$\cos\theta = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|\,\|\boldsymbol{b}\|} = \frac{xz+yx+zy}{x^2+y^2+z^2}$$

$x+y+z=0$ 则 $x^2+y^2+z^2+2xy+2yz+2xz=0$, 所以 $\cos\theta = -\dfrac{1}{2}$. 则 $\theta = \dfrac{2\pi}{3}$ □

练习 3.1.9.

1. 找到 $\mathbb{R}^4$ 中的四个两两正交的向量，且每个向量的每个分量只能是 ± 1.
2. $\mathbb{R}^n$ 中最多有多少个两两正交的向量?

证明: 1. 有很多可能，我们仅给出一个选择：

$$\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \boldsymbol{a}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \boldsymbol{a}_4 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix}$$

2. $\mathbb{R}^n$ 中最多有 n 个两两正交的向量。首先，n 个坐标向量两两正交。其次，我们断言两两正交的向量构成的向量组线性无关，而 $\mathbb{R}^n$ 中线性无关的向量组至多有 n 个向量，所以 $\mathbb{R}^n$ 中两两正交的向量至多有 n 个。现在我们来证明断言：假设 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r$ 两两正交，且 $x_1\boldsymbol{a}_1 + \cdots + x_r\boldsymbol{a}_r = 0$, 对任意 i, $\boldsymbol{a}_i^{\mathrm{T}}(x_1\boldsymbol{a}_1 + \cdots + x_r\boldsymbol{a}_r) = 0$, 即 $x_i = 0, \forall i$.

□

练习 3.1.10.

1. 找到 $\mathbb{R}^2$ 中的三个向量，使它们之间两两内积为负.
2. 找到 $\mathbb{R}^3$ 中的四个向量，使它们之间两两内积为负.
3. $\mathbb{R}^n$ 中最多有多少个向量，使它们之间两两内积为负?

证明: 1. 例如 $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix}, \boldsymbol{a}_3 = \begin{bmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{bmatrix}$

2. 例如 $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 0 \\ \frac{1}{100} \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \\ \frac{1}{100} \end{bmatrix}, \boldsymbol{a}_3 = \begin{bmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \\ \frac{1}{100} \end{bmatrix}, \boldsymbol{a}_4 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$

3. 受前两问的启发我们可以猜测，最多有 $n+1$ 个向量它们之间两两内积为负。接下来我们证明这一点。

- 首先，不能有 $n+2$ 个向量它们之间两两内积为负。我们用反证法：假设 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{n+1},\boldsymbol{a}_{n+2}$ 两两内积为负。由于 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{n+1}$ 线性相关，就存在不全为零的 $x_1,\cdots,x_{n+1}$ 使得 $x_1\boldsymbol{a}_1+\cdots+x_{n+1}\boldsymbol{a}_{n+1}=0$; 若 $x_1,\cdots,x_{n+1}$ 中既有大于零的项，又有小于零的项，那么不失一般性，我们不妨假设 $x_1,\cdots,x_r>0, x_{r+t},\cdots,x_{n+1}<0$, 那么

$$x_1\boldsymbol{a}_1+\cdots+x_r\boldsymbol{a}_r=-(x_{r+t}\boldsymbol{a}_{r+t}+\cdots+x_{n+1}\boldsymbol{a}_{n+1}).$$

  等式两边同时跟等式左边做内积，就有

$$\begin{aligned}0 &\le (x_1\boldsymbol{a}_1+\cdots+x_r\boldsymbol{a}_r)\cdot(x_1\boldsymbol{a}_1+\cdots+x_r\boldsymbol{a}_r)\\ &=-(x_1\boldsymbol{a}_1+\cdots+x_r\boldsymbol{a}_r)\cdot(x_{r+t}\boldsymbol{a}_{r+t}+\cdots+x_{n+1}\boldsymbol{a}_{n+1}).\end{aligned}$$

  但是上式最右边 $=\sum_{i=1}^{r}\sum_{j=r+t}^{n+1}(-x_ix_j)\boldsymbol{a}_i\cdot\boldsymbol{a}_j<0$, 矛盾。所以所有系数同时大于等于零或者同时小于等于零，我们不妨设 $x_i\ge 0, \forall 1\le i\le n+1$. 那么

$$0=(x_1\boldsymbol{a}_1+\cdots+x_{n+1}\boldsymbol{a}_{n+1})\cdot\boldsymbol{a}_{n+2}=\sum_{i=1}^{n+1}x_i\boldsymbol{a}_i\cdot\boldsymbol{a}_{n+2}<0,$$

  同样矛盾。

- 我们证明存在 $n+1$ 个向量它们之间两两内积为负。我们对 n 用归纳法。$n=1,2$ 时，已经给出了具体构造。假设 $\mathbb{R}^n$ 中存在 $n+1$ 个向量它们之间两两内积为负，设为 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{n+1}$, 我们取

$$\boldsymbol{b}_i=\begin{bmatrix}\boldsymbol{a}_i\\ \epsilon\end{bmatrix}, 1\le i\le n+1, \boldsymbol{b}_{n+2}=\begin{bmatrix}0\\ -1\end{bmatrix}$$

  其中 $\epsilon>0$ 待定. 对 $1\le i\le n+1$, $\boldsymbol{b}_i\cdot\boldsymbol{b}_{n+2}=-\epsilon<0$. 对 $1\le i,j\le n+1$, $\boldsymbol{b}_i\cdot\boldsymbol{b}_j=\boldsymbol{a}_i\cdot\boldsymbol{a}_j+\epsilon^2$, 因此我们只需要取 ϵ 满足

$$0<\epsilon<\min\{-\boldsymbol{a}_i\cdot\boldsymbol{a}_j:1\le i,j\le n+1, i\ne j\}$$

  即可。这就完成了归纳。

□

练习 3.1.11. 在 $\mathbb{R}^4$ 中求一单位向量与下列向量正交：

$$\boldsymbol{a}_1=\begin{bmatrix}1\\1\\-1\\1\end{bmatrix},\quad \boldsymbol{a}_2=\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix},\quad \boldsymbol{a}_3=\begin{bmatrix}2\\1\\1\\3\end{bmatrix}.$$

证明: 设 $\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}$ 与 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 正交，当且仅当 $\boldsymbol{a}_i\cdot\boldsymbol{x}=0, i=1,2,3$, 即

$$\begin{bmatrix}1&1&-1&1\\1&-1&-1&1\\2&1&1&3\end{bmatrix}\boldsymbol{x}=0$$

解方程组即得零空间的一组基是 $\boldsymbol{x}=\begin{bmatrix}-\frac{4}{3}\\0\\-\frac{1}{3}\\1\end{bmatrix}$,然后我们单位化,可得所求向量为 $\pm\begin{bmatrix}-\frac{4}{\sqrt{26}}\\0\\-\frac{1}{\sqrt{26}}\\3/\sqrt{26}\end{bmatrix}$ □

练习 3.1.12. 设 $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_n$ 是 $\mathbb{R}^n$ 的一组标准正交基，证明，

1. 如果 $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}_i=0\,(i=1,2,\ldots,n)$，则 $\boldsymbol{b}=0$.
2. 如果 $\boldsymbol{b}_1^{\mathrm{T}}\boldsymbol{a}_i=\boldsymbol{b}_2^{\mathrm{T}}\boldsymbol{a}_i\,(i=1,2,\ldots,n)$，则 $\boldsymbol{b}_1=\boldsymbol{b}_2$.

证明:

1. 我们可以把 $\boldsymbol{b}$ 写成 $\boldsymbol{b}=x_1\boldsymbol{a}_1+\cdots+x_n\boldsymbol{a}_n$, 由于 $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}_i=\sum_{j=1}^n x_j\boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{a}_i=x_i$, 故所有的 $x_1,\cdots,x_n$ 都是零。因此 $\boldsymbol{b}=0$.
2. 考虑 $\boldsymbol{b}=\boldsymbol{b}_1-\boldsymbol{b}_2$, 则 $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}_i=0(i=1,2,\cdots,n)$, 由上一小题，$\boldsymbol{b}=0$, 所以 $\boldsymbol{b}_1=\boldsymbol{b}_2$.

□

练习 3.1.13. 设 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 是 $\mathbb{R}^3$ 的一组标准正交基，证明下列向量组也是一组标准正交基：

$$\boldsymbol{b}_1=\frac{1}{3}(2\boldsymbol{a}_1+2\boldsymbol{a}_2-\boldsymbol{a}_3),\quad \boldsymbol{b}_2=\frac{1}{3}(2\boldsymbol{a}_1-\boldsymbol{a}_2+2\boldsymbol{a}_3),\quad \boldsymbol{b}_3=\frac{1}{3}(\boldsymbol{a}_1-2\boldsymbol{a}_2-2\boldsymbol{a}_3).$$

证明: 这等价于证明 $B=\begin{bmatrix}\boldsymbol{b}_1&\boldsymbol{b}_2&\boldsymbol{b}_3\end{bmatrix}$ 是正交矩阵，即 $B^{\mathrm{T}}B=I_3$. 事实上

$$B=\frac{1}{3}\begin{bmatrix}\boldsymbol{a}_1&\boldsymbol{a}_2&\boldsymbol{a}_3\end{bmatrix}\begin{bmatrix}2&2&1\\2&-1&-2\\-1&2&-2\end{bmatrix}$$

所以

$$\begin{aligned}B^{\mathrm{T}}B&=\tfrac{1}{9}\begin{bmatrix}2&2&1\\2&-1&-2\\-1&2&-2\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{a}_1^{\mathrm{T}}\\\boldsymbol{a}_2^{\mathrm{T}}\\\boldsymbol{a}_3^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{a}_1&\boldsymbol{a}_2&\boldsymbol{a}_3\end{bmatrix}\begin{bmatrix}2&2&1\\2&-1&-2\\-1&2&-2\end{bmatrix}\\&=\tfrac{1}{9}\begin{bmatrix}2&2&-1\\2&-1&2\\1&-2&-2\end{bmatrix}\begin{bmatrix}2&2&1\\2&-1&-2\\-1&2&-2\end{bmatrix}\\&=I_3\end{aligned}$$

□

练习 3.1.14. 设 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4, \boldsymbol{a}_5$ 是 $\mathbb{R}^5$ 的一组标准正交基，令

$$\boldsymbol{b}_1 = \boldsymbol{a}_1 + \boldsymbol{a}_5, \quad \boldsymbol{b}_2 = \boldsymbol{a}_1 - \boldsymbol{a}_2 + \boldsymbol{a}_4, \quad \boldsymbol{b}_3 = 2\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3.$$

求 $\mathrm{Span}(\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3)$ 的一组标准正交基.

证明*:*

$$\begin{aligned}
\boldsymbol{c}_1 &= \boldsymbol{b}_1 \\
\boldsymbol{c}_2 &= \boldsymbol{b}_2 - \frac{\boldsymbol{b}_2 \cdot \boldsymbol{c}_1}{\boldsymbol{c}_1 \cdot \boldsymbol{c}_1}\boldsymbol{c}_1 && = \tfrac{1}{2}\boldsymbol{a}_1 - \boldsymbol{a}_2 + \boldsymbol{a}_4 - \tfrac{1}{2}\boldsymbol{a}_5 \\
\boldsymbol{c}_3 &= \boldsymbol{b}_3 - \frac{\boldsymbol{b}_3 \cdot \boldsymbol{c}_1}{\boldsymbol{c}_1 \cdot \boldsymbol{c}_1}\boldsymbol{c}_1 - \frac{\boldsymbol{b}_3 \cdot \boldsymbol{c}_2}{\boldsymbol{c}_2 \cdot \boldsymbol{c}_2}\boldsymbol{c}_2 && = \boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3 - \boldsymbol{a}_5
\end{aligned}$$

然后我们做单位化即可得到 $\mathrm{Span}(\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3)$ 的一组标准正交基：

$$\frac{\boldsymbol{a}_1 + \boldsymbol{a}_5}{\sqrt{2}}, \quad \frac{\boldsymbol{a}_1 - 2\boldsymbol{a}_2 + 2\boldsymbol{a}_4 - \boldsymbol{a}_5}{\sqrt{10}}, \quad \frac{\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3 - \boldsymbol{a}_5}{2}$$

□

练习 3.1.15. 求齐次线性方程组

$$\begin{cases} 2x_1 + x_2 - x_3 + x_4 - 3x_5 = 0, \\ x_1 + x_2 - x_3 \qquad\quad + x_5 = 0 \end{cases}$$

解空间的一组标准正交基.

这类题目的固定步骤是：先用高斯消元法，求出零空间的一组基，然后用 Gram-Schmidt 正交化，将这组基正交化。

证明*:* 系数矩阵为 $\begin{bmatrix} 2 & 1 & -1 & 1 & -3 \\ 1 & 1 & -1 & 0 & 1 \end{bmatrix}$, 它的行简化阶梯型矩阵是 $\begin{bmatrix} 1 & 0 & 0 & 1 & -4 \\ 0 & 1 & -1 & -1 & 5 \end{bmatrix}$, 零空间的一组基是

$$\boldsymbol{a}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \boldsymbol{a}_3 = \begin{bmatrix} 4 \\ -5 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

那么

$$\boldsymbol{b}_1 = \boldsymbol{a}_1, \boldsymbol{b}_2 = \boldsymbol{a}_2 - \frac{\boldsymbol{b}_1 \cdot \boldsymbol{a}_2}{\boldsymbol{b}_1 \cdot \boldsymbol{b}_1}\boldsymbol{b}_1 = \begin{bmatrix} -1 \\ \frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \boldsymbol{b}_3 = \boldsymbol{a}_3 - \frac{\boldsymbol{b}_1 \cdot \boldsymbol{a}_3}{\boldsymbol{b}_1 \cdot \boldsymbol{b}_1}\boldsymbol{b}_1 - \frac{\boldsymbol{b}_2 \cdot \boldsymbol{a}_3}{\boldsymbol{b}_2 \cdot \boldsymbol{b}_2}\boldsymbol{b}_2 = \begin{bmatrix} \frac{7}{5} \\ -\frac{6}{5} \\ \frac{6}{5} \\ \frac{13}{5} \\ 1 \end{bmatrix}$$

将 $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$ 单位化，就可以得到解空间的一组标准正交基：

$$\begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2/\sqrt{10} \\ 1/\sqrt{10} \\ -1/\sqrt{10} \\ 2/\sqrt{10} \\ 0 \end{bmatrix}, \begin{bmatrix} 7/\sqrt{315} \\ -6/\sqrt{315} \\ 6/\sqrt{315} \\ 13/\sqrt{315} \\ 5/\sqrt{315} \end{bmatrix}.$$

$\square$

练习 3.1.16. 利用 *Gram-Schmidt* 正交化方法求由下列向量线性生成的子空间的标准正交基：

1. $\boldsymbol{a}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}$, $\boldsymbol{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix}$, $\boldsymbol{a}_3 = \begin{bmatrix} -2 \\ 1 \\ 1 \\ 2 \end{bmatrix}$. *2.* $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}$, $\boldsymbol{a}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$, $\boldsymbol{a}_3 = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 1 \end{bmatrix}$.

证明*:* 1. $\boldsymbol{b}_1 = \boldsymbol{a}_1$, $\boldsymbol{b}_2 = \boldsymbol{a}_2 - \dfrac{\boldsymbol{b}_1 \cdot \boldsymbol{a}_2}{\boldsymbol{b}_1 \cdot \boldsymbol{b}_1}\boldsymbol{b}_1 = \begin{bmatrix} -1 \\ 1/2 \\ 2 \\ 3/2 \end{bmatrix}$, $\boldsymbol{b}_3 = \boldsymbol{a}_3 - \dfrac{\boldsymbol{b}_1 \cdot \boldsymbol{a}_3}{\boldsymbol{b}_1 \cdot \boldsymbol{b}_1}\boldsymbol{b}_1 - \dfrac{\boldsymbol{b}_2 \cdot \boldsymbol{a}_3}{\boldsymbol{b}_2 \cdot \boldsymbol{b}_2}\boldsymbol{b}_2 = \begin{bmatrix} -2/3 \\ 2/3 \\ -1 \\ 2/3 \end{bmatrix}$. 将 $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$ 单位化，就可以得到解空间的一组标准正交基：

$$\begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ 0 \\ 1/\sqrt{6} \end{bmatrix}, \begin{bmatrix} -2/\sqrt{30} \\ 1/\sqrt{30} \\ 4/\sqrt{30} \\ 3/\sqrt{30} \end{bmatrix}, \begin{bmatrix} -2/\sqrt{21} \\ 2/\sqrt{21} \\ -3/\sqrt{21} \\ 2/\sqrt{21} \end{bmatrix}.$$

2. $\boldsymbol{b}_1 = \boldsymbol{a}_1$, $\boldsymbol{b}_2 = \boldsymbol{a}_2 - \dfrac{\boldsymbol{b}_1 \cdot \boldsymbol{a}_2}{\boldsymbol{b}_1 \cdot \boldsymbol{b}_1}\boldsymbol{b}_1 = \begin{bmatrix} 4/3 \\ -1/3 \\ 2/3 \\ 1 \end{bmatrix}$, $\boldsymbol{b}_3 = \boldsymbol{a}_3 - \dfrac{\boldsymbol{b}_1 \cdot \boldsymbol{a}_3}{\boldsymbol{b}_1 \cdot \boldsymbol{b}_1}\boldsymbol{b}_1 - \dfrac{\boldsymbol{b}_2 \cdot \boldsymbol{a}_3}{\boldsymbol{b}_2 \cdot \boldsymbol{b}_2}\boldsymbol{b}_2 = \begin{bmatrix} -6/5 \\ 13/10 \\ 7/5 \\ 11/10 \end{bmatrix}$.

将 $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$ 单位化，就可以得到解空间的一组标准正交基：

$$\begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \\ 0 \end{bmatrix}, \begin{bmatrix} 4/\sqrt{30} \\ -1/\sqrt{30} \\ 2/\sqrt{30} \\ 3/\sqrt{30} \end{bmatrix}, \begin{bmatrix} -12/\sqrt{630} \\ 13/\sqrt{630} \\ 14/\sqrt{630} \\ 11/\sqrt{630} \end{bmatrix}.$$

$\square$

练习 3.1.17. 证明命题 *3.1.13*: 设 $\mathcal{M}, \mathcal{N}$ 是 $\mathbb{R}^n$ 的两个子空间，如果 $\mathcal{M} \subseteq \mathcal{N}$, 则 $\mathcal{M}$ 的任意一组标准正交基都可以扩充成 $\mathcal{N}$ 的一组标准正交基。

证明*:* 设 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r$ 是 $\mathcal{M}$ 的一组标准正交基,我们首先把它扩充为 $\mathcal{N}$ 的一组基 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r, \boldsymbol{a}_{r+1}, \cdots, \boldsymbol{a}_t$. 然后我们对这组基作 Gram-Schmidt 正交化，就得到 $\mathcal{N}$ 的一组标准正交基 $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_t$。我们只需要说明，对 $1 \le i \le r$，我们都有 $\boldsymbol{b}_i = \boldsymbol{a}_i$。首先 $\boldsymbol{b}_1 = \boldsymbol{a}_1$, 假设对 $1 \le i \le r-1$ 都有 $\boldsymbol{b}_i = \boldsymbol{a}_i$，那么 $\boldsymbol{b}_{i+1}$ 是向量

$$\boldsymbol{a}_{i+1} - \sum_{j=1}^{i} \frac{\boldsymbol{a}_{i+1} \cdot \boldsymbol{b}_j}{\boldsymbol{b}_j \cdot \boldsymbol{b}_j}\boldsymbol{b}_j$$

的单位化；根据归纳假设，对 $1 \le j \le i, \boldsymbol{b}_j = \boldsymbol{a}_j$, 因此 $\boldsymbol{a}_{i+1} \cdot \boldsymbol{b}_j = 0$, 所以 $\boldsymbol{b}_{i+1} = \boldsymbol{a}_{i+1}$. □

练习 3.1.18 (勾股定理的高维推广)**.**

1. 向量 $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ 围出一个三角形，证明其面积的平方为 $\frac{1}{4}\left(\|\boldsymbol{a}\|^2\|\boldsymbol{b}\|^2 - (\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b})^2\right)$.
2. 两两垂直的向量 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^n$ 围出一个四面体，证明其斜面上三角形面积的平方等于其余三个直角三角形面积的平方和.

证明: 1. 设 $\boldsymbol{a}, \boldsymbol{b}$ 夹角为 θ, 那么所求三角形面积为 $S = \frac{1}{2}\|\boldsymbol{a}\|\|\boldsymbol{b}\|\sin\theta$. 而 $\cos\theta = \dfrac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|}$, 因此

$$S^2 = \frac{1}{4}\|\boldsymbol{a}\|^2\|\boldsymbol{b}\|^2\left(1 - \cos^2\theta\right) = \frac{1}{4}\left(\|\boldsymbol{a}\|^2\|\boldsymbol{b}\|^2 - (\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b})^2\right)$$

2. 记三个直角三角形为 $S_{\boldsymbol{a},\boldsymbol{b}}, S_{\boldsymbol{a},\boldsymbol{c}}, S_{\boldsymbol{b},\boldsymbol{c}}$, 斜面上的三角形为 S, 那么 S 是由 $\boldsymbol{c} - \boldsymbol{a}, \boldsymbol{b} - \boldsymbol{a}$ 围出的三角形。因此

$$S_{\boldsymbol{a},\boldsymbol{b}}^2 = \frac{\|\boldsymbol{a}\|^2\|\boldsymbol{b}\|^2}{4}, \quad S_{\boldsymbol{a},\boldsymbol{c}}^2 = \frac{\|\boldsymbol{a}\|^2\|\boldsymbol{c}\|^2}{4}, \quad S_{\boldsymbol{b},\boldsymbol{c}}^2 = \frac{\|\boldsymbol{b}\|^2\|\boldsymbol{c}\|^2}{4},$$

而 $S^2 = \dfrac{\|\boldsymbol{b}-\boldsymbol{a}\|^2\|\boldsymbol{c}-\boldsymbol{a}\|^2 - ((\boldsymbol{b}-\boldsymbol{a})^{\mathrm{T}}(\boldsymbol{c}-\boldsymbol{a}))^2}{4}$, $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ 两两垂直，那么

$$S^2 = \frac{(\|\boldsymbol{b}\|^2 + \|\boldsymbol{a}\|^2)(\|\boldsymbol{c}\|^2 + \|\boldsymbol{a}\|^2) - (\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a})^2}{4} = \frac{\|\boldsymbol{a}\|^2\|\boldsymbol{b}\|^2 + \|\boldsymbol{b}\|^2\|\boldsymbol{c}\|^2 + \|\boldsymbol{a}\|^2\|\boldsymbol{c}\|^2}{4}$$

所以我们有 $S^2 = S_{\boldsymbol{a},\boldsymbol{b}}^2 + S_{\boldsymbol{a},\boldsymbol{c}}^2 + S_{\boldsymbol{b},\boldsymbol{c}}^2$.

□

练习 3.1.19. 取定非零向量 $\boldsymbol{a} \in \mathbb{R}^n$，考虑 $\mathbb{R}^n$ 上的一个变换，它将每个向量 $\boldsymbol{b}$ 映射到其向直线 $\mathrm{Span}\{\boldsymbol{a}\}$ 正交投影后平行于 $\boldsymbol{a}$ 的部分.

1. 证明这是一个线性变换，其表示矩阵为 $A = \dfrac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}$.
2. 证明 $A^2 = A, A^{\mathrm{T}} = A$.

证明: 1. 设 $f : \mathbb{R}^n \to \mathbb{R}^n$ 是题中的映射，那么对任意 $\boldsymbol{b}$, $f(\boldsymbol{b}) = \dfrac{\boldsymbol{b}\cdot\boldsymbol{a}}{\boldsymbol{a}\cdot\boldsymbol{a}}\boldsymbol{a}$, 我们用矩阵乘法来表示点积，就有

$$f(\boldsymbol{b}) = \frac{\boldsymbol{a}(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b})}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}} = \frac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}\boldsymbol{b}$$

因此结论是明显的。

2. $A^2 = \dfrac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}\dfrac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}} = \dfrac{(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a})(\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}})}{(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a})^2} = A$, $A^{\mathrm{T}} = \dfrac{1}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}(\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}})^{\mathrm{T}} = \dfrac{1}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}} = A$.

□

练习 3.1.20 (内积决定转置)**.** 求证：设 $m \times n$ 矩阵 A 和 $n \times m$ 矩阵 B，如果对任意 $\boldsymbol{v} \in \mathbb{R}^m, \boldsymbol{w} \in \mathbb{R}^n$，都有 $(B\boldsymbol{v})^{\mathrm{T}}\boldsymbol{w} = \boldsymbol{v}^{\mathrm{T}}(A\boldsymbol{w})$，则 $B = A^{\mathrm{T}}$.

证明: 按题意，对任意 $\boldsymbol{v} \in \mathbb{R}^m, \boldsymbol{w} \in \mathbb{R}^n$，都有 $\boldsymbol{v}^{\mathrm{T}}B^{\mathrm{T}}\boldsymbol{w} = \boldsymbol{v}^{\mathrm{T}}(A\boldsymbol{w})$. 我们取 $\boldsymbol{v} = \boldsymbol{e}_i \in \mathbb{R}^m, \boldsymbol{w} = \boldsymbol{f}_j \in \mathbb{R}^n$ 分别是第 i 个和第 j 个单位向量，那么 $\boldsymbol{v}^{\mathrm{T}}B^{\mathrm{T}}\boldsymbol{w}$ 是 B^{T} 第 j 行第 i 列的元素，$\boldsymbol{v}^{\mathrm{T}}A\boldsymbol{w}$ 是 A 第 j 行第 i 列的元素，因此它们相等。由 i, j 的任意性知 $B = A^{\mathrm{T}}$. □

练习 3.1.21 (Riesz 表示定理)**.** 设 $f:\mathbb{R}^n\to\mathbb{R}$ 是线性映射，证明，存在向量 $\boldsymbol{b}$，使得对任意 $\boldsymbol{a}\in\mathbb{R}^n$，都有 $f(\boldsymbol{a})=\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}$.

证明*:* 对 $i=1,2,\cdots,n$, 令 $b_i=f(\boldsymbol{e}_i)$, $\boldsymbol{b}=\begin{bmatrix}b_1\\ \vdots\\ b_n\end{bmatrix}$, 那么对任意 $\boldsymbol{a}=\begin{bmatrix}a_1\\ \vdots\\ a_n\end{bmatrix}$, $\boldsymbol{a}=a_1\boldsymbol{e}_1+\cdots+a_n\boldsymbol{e}_n$.

那么

$$f(\boldsymbol{a})=\sum_{i=1}^n a_i f(\boldsymbol{e}_i)=\sum_{i=1}^n a_i b_i=\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}.$$

□

练习 3.1.22.

1. （平行四边形法则）证明 $\mathbb{R}^n$ 中任意平行四边形的两条对角线长度的平方和，等于其四条边长的平方和.

2. （极化公式）证明 $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w}=\dfrac{1}{4}\left(\|\boldsymbol{v}+\boldsymbol{w}\|^2-\|\boldsymbol{v}-\boldsymbol{w}\|^2\right)$.

附注. 这意味着，长度决定夹角.

设 $\|\boldsymbol{a}\|_4=\left(\sum_{i=1}^n a_i^4\right)^{\frac{1}{4}}$，定义关于 $\boldsymbol{v},\boldsymbol{w}$ 的二元函数 $\dfrac{1}{4}\left(\|\boldsymbol{v}+\boldsymbol{w}\|_4^2-\|\boldsymbol{v}-\boldsymbol{w}\|_4^2\right)$. 这个二元函数是否满足内积定义中的对称、双线性、正定三条性质?

设 $\|\boldsymbol{a}\|_\infty=\max\limits_{1\le i\le n}|a_i|$，定义关于 $\boldsymbol{v},\boldsymbol{w}$ 的二元函数 $\dfrac{1}{4}\left(\|\boldsymbol{v}+\boldsymbol{w}\|_\infty^2-\|\boldsymbol{v}-\boldsymbol{w}\|_\infty^2\right)$. 这个二元函数是否满足内积定义中的对称、双线性、正定三条性质? 此时，三角不等式和 Cauchy-Schwarz 不等式是否仍然成立?

证明*:* 1. 假设平行四边形一对相邻的边对应于向量 $\boldsymbol{v},\boldsymbol{w}$, 那么两条对角线分别为 $\boldsymbol{v}+\boldsymbol{w},\boldsymbol{v}-\boldsymbol{w}$, 那么要证明的就是 $2(\|\boldsymbol{v}\|^2+\|\boldsymbol{w}\|^2)=\|\boldsymbol{w}-\boldsymbol{v}\|^2+\|\boldsymbol{w}+\boldsymbol{v}\|^2$. 这是因为

$$\begin{aligned}\|\boldsymbol{w}-\boldsymbol{v}\|^2+\|\boldsymbol{w}+\boldsymbol{v}\|^2 &=(\boldsymbol{w}-\boldsymbol{v})\cdot(\boldsymbol{w}-\boldsymbol{v})+(\boldsymbol{w}+\boldsymbol{v})(\boldsymbol{w}+\boldsymbol{v})\\ &=(\|\boldsymbol{w}\|^2-2\boldsymbol{w}\cdot\boldsymbol{v}+\|\boldsymbol{v}\|^2)+(\|\boldsymbol{w}\|^2+2\boldsymbol{w}\cdot\boldsymbol{v}+\|\boldsymbol{v}\|^2)\ .\\ &=2(\|\boldsymbol{w}\|^2+\|\boldsymbol{v}\|^2)\end{aligned}$$

2.

$$\begin{aligned}\|\boldsymbol{v}+\boldsymbol{w}\|^2-\|\boldsymbol{v}-\boldsymbol{w}\|^2 &=(\boldsymbol{v}+\boldsymbol{w})\cdot(\boldsymbol{v}+\boldsymbol{w})+(\boldsymbol{v}-\boldsymbol{w})(\boldsymbol{v}-\boldsymbol{w})\\ &=(\|\boldsymbol{v}\|^2+2\boldsymbol{v}\cdot\boldsymbol{w}+\|\boldsymbol{v}\|^2)-(\|\boldsymbol{v}\|^2-2\boldsymbol{v}\cdot\boldsymbol{w}+\|\boldsymbol{w}\|^2)\ .\\ &=4\boldsymbol{v}\cdot\boldsymbol{w}\end{aligned}$$

我们记二元函数 $\dfrac{1}{4}\left(\|\boldsymbol{v}+\boldsymbol{w}\|_4^2-\|\boldsymbol{v}-\boldsymbol{w}\|_4^2\right)$ 为 $H_4(\boldsymbol{v},\boldsymbol{w})$. 那么我们显然有

$$H_4(\boldsymbol{v},\boldsymbol{w})=H_4(\boldsymbol{w},\boldsymbol{v});\quad H_4(\boldsymbol{v},\boldsymbol{v})=\|\boldsymbol{v}\|\ge 0, H_4(\boldsymbol{v},\boldsymbol{v})=0\text{当且仅当}\boldsymbol{v}=0.$$

但是不满足双线性性：例如我们考虑 $\mathbb{R}^2$, $\boldsymbol{v}=\begin{bmatrix}1\\1\end{bmatrix}$, $\boldsymbol{w}=\begin{bmatrix}1\\0\end{bmatrix}$, 那么 $H_4(\boldsymbol{v},\boldsymbol{w})=\dfrac{\sqrt{17}-1}{4}$, 而 $H_4(2\boldsymbol{v},\boldsymbol{w})=\dfrac{\sqrt{97}-\sqrt{17}}{4}\neq 2H_4(\boldsymbol{v},\boldsymbol{w})$.

我们记二元函数 $\frac{1}{4}\left(\|\boldsymbol{v}+\boldsymbol{w}\|_\infty^2-\|\boldsymbol{v}-\boldsymbol{w}\|_\infty^2\right)$ 为 $H_\infty(\boldsymbol{v},\boldsymbol{w})$. 那么我们显然有

$$H_\infty(\boldsymbol{v},\boldsymbol{w})=H_\infty(\boldsymbol{w},\boldsymbol{v});\quad H_\infty(\boldsymbol{v},\boldsymbol{v})=\|\boldsymbol{a}\|_\infty\geq 0, H_\infty(\boldsymbol{v},\boldsymbol{v})=0\text{当且仅当}\boldsymbol{v}=0.$$

但是不满足双线性性：例如我们考虑 $\mathbb{R}^2$, $\boldsymbol{v}=\begin{bmatrix}1\\1\end{bmatrix}$, $\boldsymbol{w}=\begin{bmatrix}1\\0\end{bmatrix}$, 那么 $H_\infty(\boldsymbol{v},\boldsymbol{w})=\frac{3}{4}$, 而 $H_\infty(2\boldsymbol{v},\boldsymbol{w})=\frac{5}{4}\neq 2H_\infty(\boldsymbol{v},\boldsymbol{w})$.

□

3.2 正交矩阵和 QR 分解

练习 3.2.1. 求一个四阶正交矩阵，其中前两个列向量分别为：$\frac{1}{3}\begin{bmatrix}1\\-2\\0\\2\end{bmatrix}$, $\frac{1}{\sqrt{6}}\begin{bmatrix}-2\\0\\1\\1\end{bmatrix}$.

我们先求出和前两个向量正交的子空间一组基，然后对其做 Gram-Schmidt 正交化。

证明: 设前两个列向量为 $\boldsymbol{a}_1,\boldsymbol{a}_2$, 我们先要找出齐次方程组

$$\begin{cases}\boldsymbol{a}_1\cdot\boldsymbol{x} &=0\\ \boldsymbol{a}_2\cdot\boldsymbol{x} &=0\end{cases}\qquad \text{即}\begin{bmatrix}1&-2&0&2\\-2&0&1&1\end{bmatrix}\boldsymbol{x}=0$$

的零空间的一组基。解方程得：$\boldsymbol{x}_1=\begin{bmatrix}1/2\\1/4\\1\\0\end{bmatrix}$, $\boldsymbol{x}_2=\begin{bmatrix}1/2\\5/4\\0\\1\end{bmatrix}$, 对这两个向量做 Gram-Schmidt 正交化可得：

$$\begin{bmatrix}2/\sqrt{21}\\1/\sqrt{21}\\4/\sqrt{21}\\0\end{bmatrix},\begin{bmatrix}2/\sqrt{126}\\8/\sqrt{126}\\-3/\sqrt{126}\\7/\sqrt{126}\end{bmatrix}$$

□

练习 3.2.2. 写出元素都是 0 或 1 的所有三阶正交矩阵.

证明: 由于正交矩阵的每一个行向量和列向量都是单位向量，印着符合题意的正交矩阵每一行每一列都只有唯一的 1. 因此这样的矩阵有 6 个：

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix},\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix},\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix},\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix},\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix},\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}$$

□

阅读 3.2.3 (Hadamard 矩阵). 给定 n 阶矩阵 A，如果 A 的元素都是 1 或 -1，且 $A^{\mathrm{T}}A = nI_n$，则称 A 是一个 n 阶 *Hadamard* 矩阵.

显然 *Hadamard* 是所有元素绝对值相同的正交矩阵的倍数. 可以证明 *Hadamard* 矩阵的阶只能是 $1, 2$ 或 $4k, k = 1, 2, \ldots$.

然而是否存在 $4k$ 阶 *Hadamard* 矩阵，还是一个悬而未决的问题，称为 *Hadamard* 猜想。

Hadamard 矩阵在信号处理中有应用.

练习 3.2.4.

1. 列举所有的一，二阶 *Hadamard* 矩阵.
2. 说明不存在三阶 *Hadamard* 矩阵.
3. 找出一个 4 阶 *Hadamard* 矩阵.
4. 证明如果 A 是 *Hadamard* 矩阵，则 $\begin{bmatrix} A & A \\ A & -A \end{bmatrix}$ 也是 *Hadamard* 矩阵. 以此说明存在 2^n 阶 *Hadamard* 矩阵.

   附注. 这与阅读 3.1.23 中的 Haar 小波基相关.

证明: 1. 一阶的有两个:$[1], [-1]$. 二阶的有

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}.$$

2. 由于 $A^{\mathrm{T}}A$ 的所有元素都是三个 $1, -1$ 的求和，因此是奇数，从而 $A^{\mathrm{T}}A$ 的任意分量都不是 0, 这就不可能是 $3I_n$.
3. 如果硬要寻找可能还是比较麻烦，但是第四小题就告诉了我们一种方法，所以我们可以写出来一个 hadamard 矩阵：

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

4. 若 A 是 Hamamard 矩阵，那么 $A^{\mathrm{T}}A = nI_n$, 那么

$$\begin{bmatrix} A & A \\ A & -A \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} A & A \\ A & -A \end{bmatrix} = \begin{bmatrix} A^{\mathrm{T}} & A^{\mathrm{T}} \\ A^{\mathrm{T}} & -A^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} A & A \\ A & -A \end{bmatrix} = \begin{bmatrix} 2A^{\mathrm{T}}A & 0 \\ 0 & 2A^{\mathrm{T}}A \end{bmatrix} = \begin{bmatrix} 2nI_n & \\ & 2nI_n \end{bmatrix} = 2nI_{2n}$$

由第 2 小题，存在 $1, 2$ 阶的 Hamamard 的矩阵，通过题中的方法，可以将阶数不断地翻倍，因此存在 2^n 阶 Hamamard 矩阵。

□

练习 3.2.5. 证明，分块上三角矩阵 $\begin{bmatrix} c & \boldsymbol{a}^{\mathrm{T}} \\ 0 & Q \end{bmatrix}$ 是正交矩阵时，必有 $c = \pm 1, \boldsymbol{a} = 0$，$Q$ 是正交矩阵.

证明: 由题意：$\begin{bmatrix} c & \boldsymbol{a}^{\mathrm{T}} \\ 0 & Q \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} c & \boldsymbol{a}^{\mathrm{T}} \\ 0 & Q \end{bmatrix} = I_{n+1}$,

$$\text{左边} = \begin{bmatrix} c & 0 \\ \boldsymbol{a} & Q^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} c & \boldsymbol{a}^{\mathrm{T}} \\ 0 & Q \end{bmatrix} = \begin{bmatrix} c^2 & c\boldsymbol{a}^{\mathrm{T}} \\ c\boldsymbol{a} & \boldsymbol{a}\boldsymbol{a}^{\mathrm{T}} + Q^{\mathrm{T}}Q \end{bmatrix}$$

那么就得到了：$c^2 = 1, c\boldsymbol{a} = 0, \boldsymbol{a}\boldsymbol{a}^{\mathrm{T}} + Q^{\mathrm{T}}Q = I_n$, 因此 $c = \pm 1, \boldsymbol{a} = 0, Q$ 是正交矩阵。 □

练习 3.2.6. 证明，上三角矩阵是正交矩阵时，必是对角矩阵，且对角元素是 ± 1.

法一. 我们对方阵的阶数 n 作归纳。$n = 1$ 结论是显然的；假设命题对 n 成立，那么对于 $n+1$，我们对该上三角矩阵作分块：

$$U = \begin{bmatrix} c & \boldsymbol{a}^{\mathrm{T}} \\ 0 & Q \end{bmatrix}$$

其中 Q 是 n 阶上三角矩阵。那么由上一题知：$c = \pm 1, \boldsymbol{a} = 0, Q$ 是正交矩阵；再由归纳假设，Q 是对角矩阵，对角元素是 ± 1. 结合起来就完成了归纳。 □

证明: 设 U 是对角线为 $a_1, \cdots, a_n$ 的上三角矩阵。若 U 是正交阵，则 $U^{-1} = U^{\mathrm{T}}$。U^{-1} 是对角线为 $a_1^{-1}, \cdots, a_n^{-1}$ 的上三角矩阵，U^{T} 是对角线为 $a_1, \cdots, a_n$ 的下三角矩阵，因此只能是对角阵，且 $a_i^{-1} = a_i, i = 1, 2, \cdots, n$, 因此 $a_i = \pm 1$. □

练习 3.2.7. 对标准基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$，显然 $\displaystyle\sum_{i=1}^{n} \boldsymbol{e}_i \boldsymbol{e}_i^{\mathrm{T}} = I_n$. 对任意标准正交基 $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n$，求证 $\displaystyle\sum_{i=1}^{n} \boldsymbol{q}_i \boldsymbol{q}_i^{\mathrm{T}} = I_n$.

证明: 令 $A = \displaystyle\sum_{i=1}^{n} \boldsymbol{q}_i \boldsymbol{q}_i^{\mathrm{T}}, Q = \begin{bmatrix} \boldsymbol{q}_1 & \cdots & \boldsymbol{q}_n \end{bmatrix}$, 那么 Q 是可逆矩阵，且 $AQ = \begin{bmatrix} A\boldsymbol{q}_1 & \cdots & A\boldsymbol{q}_n \end{bmatrix}$ 注意到，对 $j = 1, 2, \cdots, n$,

$$A\boldsymbol{q}_j = \left(\sum_{i=1}^{n} \boldsymbol{q}_i \boldsymbol{q}_i^{\mathrm{T}} \right) \boldsymbol{q}_j = \sum_{i=1}^{n} \boldsymbol{q}_i (\boldsymbol{q}_i^{\mathrm{T}} \boldsymbol{q}_j) = \boldsymbol{q}_j,$$

这说明 $AQ = Q$, 因此 $A = I_n$. □

练习 3.2.8. 设 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ 和 $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n$ 是 $\mathbb{R}^n$ 的两组标准正交基，证明存在正交矩阵 Q，使得 $Q\boldsymbol{a}_i = \boldsymbol{b}_i, i = 1, 2, \ldots, n$.

证明: 我们按题意，把 Q 的要求用矩阵的方式写出来：

$$\begin{bmatrix} \boldsymbol{b}_1 & \cdots & \boldsymbol{b}_n \end{bmatrix} = \begin{bmatrix} Q\boldsymbol{a}_1 & \cdots & Q\boldsymbol{a}_n \end{bmatrix} = Q \begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{bmatrix},$$

这说明 $Q = BA^{-1}$, 其中 $A = \begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{bmatrix}, B = \begin{bmatrix} \boldsymbol{b}_1 & \cdots & \boldsymbol{b}_n \end{bmatrix}$.

现在，我们只需要说明 $Q = BA^{-1}$ 是正交矩阵即可，事实上这是因为 A, B 都是正交矩阵，所以 A^{-1} 也是正交矩阵，BA^{-1} 也是正交矩阵。 □

练习 3.2.9. 回顾命题 *3.2.4*，设 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ 是 $\mathbb{R}^n$ 的一组基，考虑下列放松的条件.

1. A 在一组基上保距，即如果对 $i=1,2,\dots,n$，$\|A\boldsymbol{a}_i\|=\|\boldsymbol{a}_i\|$，那么 A 是否一定是正交矩阵？

2. A 在一组基上保内积，即如果对 $i,j=1,2,\dots,n$，$A\boldsymbol{a}_i$ 与 $A\boldsymbol{a}_j$ 的内积等于 $\boldsymbol{a}_i$ 与 $\boldsymbol{a}_j$ 的内积，那么 A 是否一定是正交矩阵？

证明*:* 1. 不一定，例如我们考虑 $\mathbb{R}^2$, $\boldsymbol{a}_1=\begin{bmatrix}1\\0\end{bmatrix},\boldsymbol{a}_2=\begin{bmatrix}0\\1\end{bmatrix},A=\begin{bmatrix}1&3/5\\0&4/5\end{bmatrix}$

2. 不一定，例如我们考虑 $\mathbb{R}^2$, $\boldsymbol{a}_1=\begin{bmatrix}1\\0\end{bmatrix},\boldsymbol{a}_2=\begin{bmatrix}0\\1\end{bmatrix},A=\begin{bmatrix}2&0\\0&1/2\end{bmatrix}$

□

练习 3.2.10. 设 $H_{\boldsymbol{v}}:=I_n-2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$ 是 $\mathbb{R}^n$ 上反射变换的表示矩阵，Q 是 n 阶正交矩阵，证明 $Q^{-1}H_{\boldsymbol{v}}Q$ 也是某个反射变换的表示矩阵.

证明*:* Q 是正交矩阵，则 $Q^{-1}=Q^{\mathrm{T}}$. 那么

$$Q^{-1}H_vQ=Q^{-1}(I_n-2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}})Q=I_n-2(Q^{\mathrm{T}}\boldsymbol{v})(Q^{\mathrm{T}}\boldsymbol{v})^{\mathrm{T}},$$

要说明 $Q^{-1}H_vQ$ 只需要 $Q^{\mathrm{T}}\boldsymbol{v}$ 是单位向量即可；这由正交矩阵的保距性可知。 □

练习 3.2.11. 证明命题 *3.2.6.* 给定 $\mathbb{R}^n$ 中的向量 $\boldsymbol{x},\boldsymbol{y}$, 满足 $\|\boldsymbol{x}\|=\|\boldsymbol{y}\|$, 则存在反射 $H_{\boldsymbol{v}}$, 其中 $\boldsymbol{v}=\dfrac{\boldsymbol{x}-\boldsymbol{y}}{\|\boldsymbol{x}-\boldsymbol{y}\|}$, 使得 $H_{\boldsymbol{v}}(\boldsymbol{x})=\boldsymbol{y}$.

证明*:* 因为 $\boldsymbol{v}$ 是单位向量，$H_{\boldsymbol{v}}$ 良定义。我们只需要验证 $H_{\boldsymbol{v}}(\boldsymbol{x})=\boldsymbol{y}$.

$$H_{\boldsymbol{v}}(\boldsymbol{x})=\boldsymbol{y}\Leftrightarrow\boldsymbol{x}-2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{y}\Leftrightarrow\boldsymbol{x}-\boldsymbol{y}=2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}$$

注意到 $\boldsymbol{x}-\boldsymbol{y}=\|\boldsymbol{x}-\boldsymbol{y}\|\,\boldsymbol{v}$, 我们有

$$H_{\boldsymbol{v}}(\boldsymbol{x})=\boldsymbol{y}\Leftrightarrow\|\boldsymbol{x}-\boldsymbol{y}\|=2\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}\Leftrightarrow\|\boldsymbol{x}-\boldsymbol{y}\|^2=2(\boldsymbol{x}^{\mathrm{T}}-\boldsymbol{y}^{\mathrm{T}})\boldsymbol{x}$$

我们利用 $\|x\|=\|y\|$, 就有

$$\|\boldsymbol{x}-\boldsymbol{y}\|=\boldsymbol{x}^T\boldsymbol{x}-2\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}+\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}=2\boldsymbol{x}^T\boldsymbol{x}-2\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}=2(\boldsymbol{x}^{\mathrm{T}}-\boldsymbol{y}^{\mathrm{T}})\boldsymbol{x}.$$

□

练习 3.2.12. 计算 QR 分解：

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

对矩阵作 QR 分解其实就是把 Gram-Schmidt 正交化的过程用矩阵写下来。当然，对于特殊的矩阵，如果我们能瞪眼看出结果，那么也没有必要去计算.

证明*:* 我们记 $A=\begin{bmatrix}\boldsymbol{a}_1&\boldsymbol{a}_2&\boldsymbol{a}_3\end{bmatrix}$

1. 简单观察就知道 A 的三列两两正交，因此可以直接写出结论

$$A = QR,\ 其中Q = \frac{1}{3}A, , R = 3I_3.$$

2. 不需要观察和计算就知道 $A = QR$, 其中 $Q = I_3, R = A$.

3. $\tilde{\boldsymbol{q}}_1 = \boldsymbol{a}_1, \tilde{\boldsymbol{q}}_2 = \boldsymbol{a}_2 - \frac{\boldsymbol{a}_2 \cdot \tilde{\boldsymbol{q}}_1}{\tilde{\boldsymbol{q}}_1 \cdot \tilde{\boldsymbol{q}}_1}\tilde{\boldsymbol{q}}_1 = \begin{bmatrix} -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}, \tilde{\boldsymbol{q}}_3 = \boldsymbol{a}_3 - \frac{\boldsymbol{a}_3 \cdot \tilde{\boldsymbol{q}}_1}{\tilde{\boldsymbol{q}}_1 \cdot \tilde{\boldsymbol{q}}_1}\tilde{\boldsymbol{q}}_1 - \frac{\boldsymbol{a}_3 \tilde{\boldsymbol{q}}_2}{\tilde{\boldsymbol{q}}_2 \cdot \tilde{\boldsymbol{q}}_2}\tilde{\boldsymbol{q}}_2 = \begin{bmatrix} 0 \\ -1/2 \\ 1/2 \end{bmatrix}$

   因此

$$\begin{aligned} Q &= \begin{bmatrix} \frac{\tilde{\boldsymbol{q}}_1}{\|\tilde{\boldsymbol{q}}_1\|} & \frac{\tilde{\boldsymbol{q}}_2}{\|\tilde{\boldsymbol{q}}_2\|} & \frac{\tilde{\boldsymbol{q}}_3}{\|\tilde{\boldsymbol{q}}_3\|} \end{bmatrix} & = \begin{bmatrix} 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix} \\ R &= \operatorname{diag}(\|\tilde{\boldsymbol{q}}\|_i) \begin{bmatrix} 1 & \frac{\boldsymbol{a}_2 \cdot \tilde{\boldsymbol{q}}_1}{\tilde{\boldsymbol{q}}_1 \cdot \tilde{\boldsymbol{q}}_1} & \frac{\boldsymbol{a}_3 \cdot \tilde{\boldsymbol{q}}_1}{\tilde{\boldsymbol{q}}_1 \cdot \tilde{\boldsymbol{q}}_1} \\ & 1 & \frac{\boldsymbol{a}_3 \cdot \tilde{\boldsymbol{q}}_2}{\tilde{\boldsymbol{q}}_2 \cdot \tilde{\boldsymbol{q}}_2} \\ & & 1 \end{bmatrix} & = \begin{bmatrix} \sqrt{3} & 2/\sqrt{3} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix} \end{aligned}$$

□

练习 3.2.13. *设向量组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ 线性无关，首先令 $\boldsymbol{q}_1$ 为与 $\boldsymbol{v}_1$ 平行的单位向量，然后令 $\boldsymbol{q}_2$ 为二维子空间 $\mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ 中垂直于直线 $\mathrm{Span}\{\boldsymbol{v}_1\}$ 的单位向量，再令 $\boldsymbol{q}_3$ 为三维子空间 $\mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ 中垂直于平面 $\mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ 的单位向量，以此类推. 这样得到的 $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_k$ 与 Gram-Schmidt 正交化得到的结果是否一致？如果有区别的话，区别在哪里？从 QR 分解的角度如何解释？*

证明*:* 这样得到的 $bmq_1, \cdots, \boldsymbol{q}_k$ 和 Gram-Schmidt 正交化得到的结果未必一致，这是因为每一次计算 $\boldsymbol{q}_i$ 时，我们都有两个选择，这两个选择相差 -1 倍。这也是唯一的区别。从 QR 分解的角度看，我们把 Q 的某些列变成原来的 -1 倍，得到的新矩阵依然是正交矩阵，并且它的前 i 列张成的子空间也没有改变。 □

练习 3.2.14 (QR 分解的其他理解)**.** *考虑 $A = \begin{bmatrix} 1 & 0 \\ -2 & 5 \\ 0 & 1 \end{bmatrix}$ 的列向量围出的平行四边形.*

1. *通过列变换，把第一列的若干倍加到第二列，使得平行四边形变成长方形. 这对应着 A 右乘哪个矩阵？*

2. *通过旋转和反射，把平行四边形的第一条边变到 x_1 轴正半轴上，第二条边变到 x_1x_2 平面中 $x_2 > 0$ 的那一半. 这对应着 A 左乘哪个正交矩阵？*

证明*:* • 对应着 A 右乘一个倍加矩阵 $\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$, 其中 a 满足 $\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ -2a+5 \\ 1 \end{bmatrix} = 0$, 解得 $a = 2$.

- 当我们做完这一步的操作，那么第一步的操作就等价于将第二条边变到 x_2 轴正半轴。因为 $A\begin{bmatrix}1&2\\0&1\end{bmatrix}=\begin{bmatrix}1&2\\-2&1\\0&1\end{bmatrix}$, 所以 A 左乘的正交矩阵 P 满足

$$P\begin{bmatrix}1/\sqrt{5}&2/\sqrt{6}\\-2/\sqrt{5}&1/\sqrt{6}\\0&1/\sqrt{6}\end{bmatrix}=\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix},\text{ 即 }\begin{bmatrix}1/\sqrt{5}&2/\sqrt{6}\\-2/\sqrt{5}&1/\sqrt{6}\\0&1/\sqrt{6}\end{bmatrix}=P^{\mathrm{T}}\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}$$

那么 P^{T} 的前两列是 $\begin{bmatrix}1/\sqrt{5}\\-2/\sqrt{5}\\0\end{bmatrix},\begin{bmatrix}2/\sqrt{6}\\1/\sqrt{6}\\1/\sqrt{6}\end{bmatrix}$. 再利用 P^{T} 是正交矩阵知，它第三列是和前两列正交的单位向量，解得第三列为 $\pm\begin{bmatrix}2/\sqrt{30}\\1/\sqrt{30}\\-5/\sqrt{30}\end{bmatrix}$，所以

$$P=\begin{bmatrix}1/\sqrt{5}&-2/\sqrt{5}&0\\2/\sqrt{6}&1/\sqrt{6}&1/\sqrt{6}\\\pm2/\sqrt{30}&\pm1/\sqrt{30}&\mp5/\sqrt{30}\end{bmatrix}$$

□

练习 3.2.15. 证明若矩阵列满秩，则其简化 QR 分解唯一. **提示**. 可以利用练习 *3.2.6*.

证明: 假设 A 列满秩，$A=Q_1R_1=Q_2R_2$ 是它的两个 QR 分解，则 Q_1,Q_2 的列向量是两两正交的单位向量，R_1,R_2 是对角线为正的上三角矩阵，那么 $Q_2^{\mathrm{T}}Q_1=Q_2^{\mathrm{T}}Q_2R_2R_1^{-1}=R_2R_1^{-1}$. 这里 R_1^{-1} 也是对角线为正的上三角矩阵。

$$(Q_2^{\mathrm{T}}Q_1)^{\mathrm{T}}(Q_2^{\mathrm{T}}Q_1)=Q_1^{\mathrm{T}}Q_2Q_2^{\mathrm{T}}Q_1=I_n$$

因此 $Q_2^{\mathrm{T}}Q_1$ 是正交矩阵，从而 $R_2R_1^{-1}$ 也是正交矩阵。利用练习 3.2.6可知 $R_2R_1^{-1}=I_n$，即 $R_2=R_1$, 进而 $Q_1=Q_2$. □

练习 3.2.16. 证明任意 n 阶正交矩阵可以表示成不多于 n 个反射的乘积.

证明: 我们对维数 n 做归纳。$n=1$ 时，结论是显然的。假设命题对任意 n 阶正交矩阵成立，现在我们考虑 $n+1$.

- 若对任意 $\boldsymbol{v}\in\mathbb{R}^{n+1}$, $A\boldsymbol{v}=\boldsymbol{v}$，那么命题也是显然的：此时 $A=I_{n+1}$.
- 若存在 $\boldsymbol{v}\in\mathbb{R}^{n+1}$，使得 $A\boldsymbol{v}\neq\boldsymbol{v}$，不妨设 $\|\boldsymbol{v}\|=1$. 令 $\boldsymbol{v}'=\dfrac{\boldsymbol{v}-A\boldsymbol{v}}{\|A\boldsymbol{v}-\boldsymbol{v}\|}$，由题 3.2.10 知 $H_{\boldsymbol{v}}\boldsymbol{v}=A\boldsymbol{v}$。我们把 $\boldsymbol{v}$ 扩张成 $\mathbb{R}^{n+1}$ 的一组标准正交基 $\boldsymbol{v},\boldsymbol{v}_2,\cdots,\boldsymbol{v}_{n+1}$，令 $P=\begin{bmatrix}\boldsymbol{v}&\boldsymbol{v}_2&\cdots&\boldsymbol{v}_{n+1}\end{bmatrix}$。因为 $(H_{\boldsymbol{v}}A)(\boldsymbol{v})=\boldsymbol{v}$, 那么

$$H_{\boldsymbol{v}}AP=P\begin{bmatrix}1&\boldsymbol{a}^{\mathrm{T}}\\0&B\end{bmatrix},\text{ 即 }\begin{bmatrix}1&\boldsymbol{a}^{\mathrm{T}}\\0&B\end{bmatrix}=P^{\mathrm{T}}H_{\boldsymbol{v}}AP$$

所以 $\begin{bmatrix}1 & \boldsymbol{a}^{\mathrm{T}}\\ 0 & B\end{bmatrix}$ 是正交矩阵，根据练习 3.2.5知 $\boldsymbol{a}=0, B$ 是正交矩阵。那么根据归纳假设，存在 $k\geq n$, 以及单位向量 $\boldsymbol{w}_1,\cdots,\boldsymbol{w}_k\in\in\mathbb{R}^n$ 使得 $B=H_{\boldsymbol{w}_1}\cdots H_{\boldsymbol{w}_k}$. 令 $\boldsymbol{w}_i'=\begin{bmatrix}0\\ \boldsymbol{w}_i\end{bmatrix}, i=1,\cdots,k$ 那么 $\begin{bmatrix}1 & 0\\ 0 & H_{\boldsymbol{w}_i}\end{bmatrix}=H_{\boldsymbol{w}_i'}$.

$$P^{\mathrm{T}}H_{\boldsymbol{v}}AP=H_{\boldsymbol{w}_1'}\cdots H_{\boldsymbol{w}_k'},\ \text{即}A=H_{\boldsymbol{v}}(PH_{\boldsymbol{w}_1'}P^{\mathrm{T}})\cdots(PH_{\boldsymbol{w}_k'}P^{\mathrm{T}}).$$

根据练习 3.2.10知 $PH_{\boldsymbol{w}_i'}P^{\mathrm{T}}$ 也都是反射。这就完成了归纳。

□

练习 3.2.17. *设 $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_s$ 和 $\boldsymbol{b}_1,\boldsymbol{b}_2,\ldots,\boldsymbol{b}_s$ 是 $\mathbb{R}^n$ 的两个向量组，证明，存在正交矩阵 Q，使得 $Q\boldsymbol{a}_i=\boldsymbol{b}_i, i=1,2,\ldots,s$，当且仅当 $\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{a}_j=\boldsymbol{b}_i^{\mathrm{T}}\boldsymbol{b}_j, i,j=1,2,\ldots,s$.*

提示. *化简成两个向量组均为线性无关向量组的情形，再用练习 3.2.15.*

证明:

- 若存在正交矩阵 Q，使得 $Q\boldsymbol{a}_i=\boldsymbol{b}_i, i=1,2,\ldots,s$，则 $\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{a}_j=\boldsymbol{b}_i^{\mathrm{T}}Q^{\mathrm{T}}Q\boldsymbol{b}_j=\boldsymbol{b}_i^{\mathrm{T}}\boldsymbol{b}_j, i,j=1,2,\ldots,s$.
- 反之，假设 $\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{a}_j=\boldsymbol{b}_i^{\mathrm{T}}\boldsymbol{b}_j, i,j=1,2,\ldots,s$.我们先证明：$\boldsymbol{a}_{i_1},\cdots,\boldsymbol{a}_{i_r}$ 线性无关，当且仅当 $\boldsymbol{b}_{i_1},\cdots,\boldsymbol{b}_{i_r}$ 线性无关。令 $A_r=\begin{bmatrix}\boldsymbol{a}_{i_1} & \cdots & \boldsymbol{a}_{i_r}\end{bmatrix}, B_r=\begin{bmatrix}\boldsymbol{b}_{i_1} & \cdots & \boldsymbol{b}_{i_r}\end{bmatrix}$. 则 $\boldsymbol{a}_{i_1},\cdots,\boldsymbol{a}_{i_r}$ 线性无关当且仅当 A_r 列满秩，当且仅当 $A_r^{\mathrm{T}}A_r$ 可逆；注意到 $A_r^{\mathrm{T}}A_r=B^{\mathrm{T}}B_r$, 所以 $\boldsymbol{a}_{i_1},\cdots,\boldsymbol{a}_{i_r}$ 线性无关，当且仅当 $\boldsymbol{b}_{i_1},\cdots,\boldsymbol{b}_{i_r}$ 线性无关。我们取 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_s$ 的一个极大线性无关组，不妨设为 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r$，那么当且仅当 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_r$ 也是 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_s$ 的极大线性无关组。我们再来证明：若 $\boldsymbol{a}_j=x_1\boldsymbol{a}_1+\cdots+x_r\boldsymbol{a}_r$, 则 $\boldsymbol{b}_j=x_1\boldsymbol{b}_1+\cdots+x_r\boldsymbol{b}_r$. 由于我们选取了极大线性无关组，因此存在唯一的线性表出：$\boldsymbol{b}_j=y_1\boldsymbol{b}_1+\cdots+y_r\boldsymbol{b}_r$，即 $B_r\begin{bmatrix}y_1\\ \vdots\\ y_r\end{bmatrix}=\boldsymbol{b}_j$, 那么

$$\begin{bmatrix}y_1\\ \vdots\\ y_r\end{bmatrix}=(B_r^{\mathrm{T}}B_r)^{-1}B_r^{\mathrm{T}}\boldsymbol{b}_j=(A_r^{\mathrm{T}}A)^{-1}A_r\boldsymbol{a}_j=\begin{bmatrix}x_1\\ \vdots\\ x_r\end{bmatrix}$$

  这说明若正交矩阵 Q 满足 $Q\boldsymbol{a}_i=\boldsymbol{b}_i, i=1,\cdots,r$, 则 $Q\boldsymbol{a}_i=\boldsymbol{b}_i, i=1,\cdots s$. 所以我们不妨设 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_s$ 和 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_s$ 都线性无关。即 $A=\begin{bmatrix}\boldsymbol{a}_1 & \cdots & \boldsymbol{a}_s\end{bmatrix}, B=\begin{bmatrix}\boldsymbol{b}_1 & \cdots & \boldsymbol{b}_s\end{bmatrix}$, 我们作 QR 分解：$A=Q_1R_1, B=Q_2R_2$, 那么由题意，$A^{\mathrm{T}}A=B^{\mathrm{T}}B$，即 $R_1^{\mathrm{T}}R_1=R_2^{\mathrm{T}}R_2$, 所以 $(R_2^{\mathrm{T}})^{-1}R_1^{\mathrm{T}}=R_2R_1^{-1}=(R_1R_2^{-1})^{\mathrm{T}}$; 这说明 $R_2R_1^{-1}$ 是正交矩阵，且是对角线为正的上三角矩阵，所以只能是单位阵，即 $R_2=R_1$. 这个时候，我们取 $Q=Q_2Q_1^T$ 即可。

□

练习 3.2.18 (保角变换). *设可逆矩阵 A 对应的线性变换保持向量之间的夹角不变.*

1. *对 A 进行 QR 分解，证明 R 也保持向量之间的夹角不变.*
2. *证明 R 为对角矩阵.*

3. 证明 $R = kI_n$，这里 k 为常数. 由此得到，A 必是某个正交矩阵的倍数.

证明: 证明: 1. 因为 A 保持夹角,所以对任意非零向量 $\boldsymbol{a}, \boldsymbol{b}$, 我们有 $\frac{(A\boldsymbol{a}) \cdot (A\boldsymbol{b})}{\|A\boldsymbol{a}\| \|A\boldsymbol{b}\|} = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|\boldsymbol{a}\| \|\boldsymbol{b}\|}$, 那么

$$\frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|\boldsymbol{a}\| \|\boldsymbol{b}\|} = \frac{\boldsymbol{a}^{\mathrm{T}} R^{\mathrm{T}} Q^{\mathrm{T}} QR\boldsymbol{b}}{\|QR\boldsymbol{a}\| \|QR\boldsymbol{b}\|} = \frac{\boldsymbol{a}^{\mathrm{T}} R^{\mathrm{T}} R\boldsymbol{b}}{\|R\boldsymbol{a}\| \|R\boldsymbol{b}\|} = \frac{(R\boldsymbol{a}) \cdot (R\boldsymbol{b})}{\|R\boldsymbol{a}\| \|R\boldsymbol{b}\|}$$

2. 考虑 R 的列向量 $\boldsymbol{a}_1 = R\boldsymbol{e}_1, \cdots, \boldsymbol{a}_n = R\boldsymbol{e}_n$，由保角性知 $\boldsymbol{a}_i \cdot \boldsymbol{a}_j = 0 (i \neq j)$. 取 $i = 1$ 知 R 第一行除了第一列以外全是零；再取 $i = 2$ 知知 R 第二行除了第二列以外全是零，以此类推；即可得 R 是对角矩阵。

3. 设 $R = \mathrm{diag}(k_1, \cdots, k_n), k_i > 0, i = 1, 2. \cdots, n$. 当 $i \neq j$ 时，$(\boldsymbol{e}_i + \boldsymbol{e}_j) \cdot (\boldsymbol{e}_i - \boldsymbol{e}_j) = 0$, 故 $(R\boldsymbol{e}_i + R\boldsymbol{e}_j) \cdot (R\boldsymbol{e}_i - R\boldsymbol{e}_j) = 0$, 这就得到 $k_i^2 = k_j^2$，所以 $k_i = k_j$.

□

□

练习 3.2.19. 设 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ 是 $\mathbb{R}^n$ 中的 m 个向量，定义矩阵

$$G(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m) := \begin{bmatrix} \boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_1 & \boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_2 & \cdots & \boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_m \\ \boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{a}_1 & \boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{a}_2 & \cdots & \boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_m \\ \vdots & \vdots & & \vdots \\ \boldsymbol{a}_m^{\mathrm{T}}\boldsymbol{a}_1 & \boldsymbol{a}_m^{\mathrm{T}}\boldsymbol{a}_2 & \cdots & \boldsymbol{a}_m^{\mathrm{T}}\boldsymbol{a}_m \end{bmatrix},$$

称为 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ 的 *Gram* 矩阵. 证明：

1. $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ 是正交单位向量组当且仅当 $G(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m) = I_m$.
2. *Gram* 矩阵 $G = G(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$ 是 m 阶对称矩阵,且对任意 $\boldsymbol{x} \in \mathbb{R}^m$,都有 $\boldsymbol{x}^{\mathrm{T}} G\boldsymbol{x} \geq 0$.
3. $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ 线性无关当且仅当 $G = G(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$ 可逆，也等价于对任意非零 $\boldsymbol{x} \in \mathbb{R}^m$，都有 $\boldsymbol{x}^{\mathrm{T}} G\boldsymbol{x} > 0$.

证明: 1. $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ 是正交单位向量组当且仅当

$$\boldsymbol{a}_i^{\mathrm{T}} \boldsymbol{a}_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

当且仅当 $G(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m) = I_m$.

2. 令 $A = \begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_m \end{bmatrix}$, 则 $G(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m) = A^{\mathrm{T}} A$, 故 G 是对称矩阵，且对任意 $\boldsymbol{x} \in \mathbb{R}^m$，都有

$$\boldsymbol{x}^{\mathrm{T}} G\boldsymbol{x} = (A\boldsymbol{x})^{\mathrm{T}} (A\boldsymbol{x}) \geq 0.$$

3. $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ 线性无关当且仅当 $A\boldsymbol{x} = 0$ 只有零解;当且仅当对任意非零 $\boldsymbol{x}$, $(A\boldsymbol{x})^{\mathrm{T}} (A\boldsymbol{x}) > 0$，此时 $G\boldsymbol{x} = 0$ 只有零解，所以 G 可逆。反之若 G 可逆，则 $A^{\mathrm{T}} A\boldsymbol{x} = G\boldsymbol{x} = 0$ 只有零解, 则 $A\boldsymbol{x} = 0$ 只有零解。

□

3.3 子空间和投影

练习 3.3.1. 设 $\mathcal{M}$ 是如下齐次线性方程组的解空间：

$$\begin{cases} 2x_1+x_2+3x_3-\ x_4=0, \\ 3x_1+\ 2x_2 \qquad -2x_4=0, \\ 3x_1+x_2+9x_3-\ x_4=0, \end{cases}$$

分别求 $\mathcal{M}$ 和 $\mathcal{M}^{\perp}$ 的一组标准正交基.

证明: 系数矩阵 $A=\begin{bmatrix} 2 & 1 & 3 & -1 \\ 3 & 2 & 0 & -2 \\ 3 & 1 & 9 & -1 \end{bmatrix}$，则 $\mathrm{rref}(A)=\begin{bmatrix} 1 & 0 & 6 & 0 \\ 0 & 1 & -9 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ 那么 $\mathcal{M}$ 的一组基为 $\begin{bmatrix} -6 \\ 9 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathcal{M}^{\perp}$ 的一组基为 $\begin{bmatrix} 1 \\ 0 \\ 6 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -9 \\ -1 \end{bmatrix}$, 直接做 Gram-Schmidt 正交化即可：为了计算简单，我们其实稍微做了点变化；

- $\mathcal{M}$ 的一组标准正交基为 $\begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} -12/\sqrt{310} \\ 9/\sqrt{155} \\ 2/\sqrt{310} \\ -9/\sqrt{310} \end{bmatrix}$；
- $\mathcal{M}^{\perp}$ 的一组标准正交基为 $\begin{bmatrix} 2/\sqrt{15} \\ 1/\sqrt{15} \\ 3/\sqrt{15} \\ -1/\sqrt{15} \end{bmatrix}, \begin{bmatrix} -5/\sqrt{93} \\ -4/\sqrt{93} \\ 6/\sqrt{93} \\ 4/\sqrt{93} \end{bmatrix}$,

□

练习 3.3.2. 设 $\mathbb{R}^4$ 中的向量 $\boldsymbol{a}_1=\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{a}_2=\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$ 生成子空间 $\mathcal{M}=\mathrm{Span}(\boldsymbol{a}_1,\boldsymbol{a}_2)$，求 $\mathcal{M}^{\perp}$ 的一组标准正交基.

证明: 我们需要先解方程组：$\begin{cases} \boldsymbol{a}_1\cdot\boldsymbol{x}=0 \\ \boldsymbol{a}_2\cdot\boldsymbol{x}=0 \end{cases}$ 也就是求出矩阵 $\begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$ 零空间的一组基，然后做 Gram-Schmidt 正交化。我们省略过程，按照标准程序求出来零空间的一组基为 $\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$.

这两个向量已经正交了，直接单位化即可，因此 $\mathcal{M}^{\perp}$ 的一组标准正交基为

$$\begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$$

□

练习 3.3.3.

1. 设 $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 2 \end{bmatrix}, B = \begin{bmatrix} 5 & 4 \\ 6 & 3 \\ 5 & 1 \end{bmatrix}$，求两个矩阵列空间的交集中的一个非零向量；由此判断两个列空间是否正交.

2. 求标准正交基 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$，使得 $\boldsymbol{v}_1 \in \mathcal{R}(A) \cap \mathcal{R}(B)$，$\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathcal{R}(A)$，且 $\boldsymbol{v}_1, \boldsymbol{v}_3 \in \mathcal{R}(B)$.

3. 求一组向量 $\boldsymbol{x}, \boldsymbol{y}$，使得 $A\boldsymbol{x} = B\boldsymbol{y}$，并计算 $\begin{bmatrix} A & -B \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix}$.

4. 求 $\begin{bmatrix} A & -B \end{bmatrix}$ 零空间的一组基，并求所有的 $\boldsymbol{x}, \boldsymbol{y}$，满足 $A\boldsymbol{x} = B\boldsymbol{y}$.

证明: 1. 因为 $\mathcal{R}(A) = \mathcal{N}(A^{\mathrm{T}})^{\perp}, \mathcal{R}(B) = \mathcal{N}(B^{\mathrm{T}})^{\perp}$, 那么 $\mathcal{R}(A) \cap \mathcal{R}(B)$ 就是和 $\mathcal{N}(A^{\mathrm{T}}), \mathcal{N}(B^{\mathrm{T}})$ 都正交的向量。直接解方程组可得 $\mathcal{N}(A^{\mathrm{T}}) = \mathrm{Span}\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right), \mathcal{N}(B^{\mathrm{T}}) = \mathrm{Span}\left(\begin{bmatrix} 1 \\ -\frac{5}{3} \\ 1 \end{bmatrix}\right)$，因此 $\mathcal{R}(A) \cap \mathcal{R}(B)$ 就是矩阵 $\begin{bmatrix} 1 & 0 & -1 \\ 1 & -\frac{5}{3} & 1 \end{bmatrix}$ 的零空间，即 $\mathrm{Span}\left(\begin{bmatrix} 1 \\ \frac{6}{5} \\ 1 \end{bmatrix}\right)$。这两个列空间不正交，因为正交的子空间的交集 $\{0\}$.

2. 我们可以先取 $\boldsymbol{w}_1 = \begin{bmatrix} 1 \\ \frac{6}{5} \\ 1 \end{bmatrix}, \boldsymbol{w}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \boldsymbol{w}_3 = \begin{bmatrix} 4 \\ 3 \\ 1 \end{bmatrix}$, 然后再做 Gram-Schmidt 正交化即可，那么

$$\boldsymbol{v}_1 = \begin{bmatrix} 5/\sqrt{86} \\ 6/\sqrt{86} \\ 5/\sqrt{86} \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 3/\sqrt{43} \\ -5/\sqrt{43} \\ 3/\sqrt{43} \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$$

3. 因为 $A\boldsymbol{x} \in \mathcal{R}(A), B\boldsymbol{y} \in \mathcal{R}(B)$, 所以 $A\boldsymbol{x} = B\boldsymbol{y} \in \mathcal{R}(A) \cap \mathcal{R}(B)$, 所以我们不妨令 $A\boldsymbol{x} = B\boldsymbol{y} = \begin{bmatrix} 5 \\ 6 \\ 5 \end{bmatrix}$, 那么解得 $\boldsymbol{x} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \boldsymbol{y} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, 并且

$$\begin{bmatrix} A & -B \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} = A\boldsymbol{x} - B\boldsymbol{y} = 0.$$

4. 由于 $\mathcal{R}(A)\cap\mathcal{R}(B)=\operatorname{Span}\left(\begin{bmatrix}1\\ \frac{6}{5}\\ 1\end{bmatrix}\right)$，因此 $\begin{bmatrix}A & -B\end{bmatrix}$ 的一组基为 $\begin{bmatrix}3\\1\\1\\0\end{bmatrix}$，所有满足 $A\boldsymbol{x}=B\boldsymbol{y}$ 的 $\boldsymbol{x},\boldsymbol{y}$ 即为 $\left(k\begin{bmatrix}3\\1\end{bmatrix},k\begin{bmatrix}1\\0\end{bmatrix}\right),k\in\mathbb{R}$.

□

练习 3.3.4.

1. $\mathbb{R}^5$ 中的两个三维子空间是否可能正交?

2. 设 $A=\begin{bmatrix}0&1&&&&\\&0&1&&&\\&&0&0&&\\&&&0&1&\\&&&&0&1\\&&&&&0\end{bmatrix}$，则 $\mathcal{R}(A)$ 和 $\mathcal{N}(A)$ 是否正交? 是否互为正交补?

证明: 1. 不可能，因为 $\dim(V_1\cap V_2)=\dim(V_1)+\dim(V_2)-\dim(V_1+V_2)\geq 1$, 从而 $V_1\cap V_2\neq\{0\}$.

2. 直接计算：$\mathcal{R}(A)=\operatorname{Span}(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_4,\boldsymbol{e}_5),\mathcal{N}(A)=\operatorname{Span}(\boldsymbol{e}_1,\boldsymbol{e}_4)$, 它们不正交。

□

练习 3.3.5. 设 6 阶方阵 A 满足 $A^3=0$，它的秩最大为多少? 举例说明. 在 A,A^2 的行空间、列空间、零空间和左零空间之中，哪些互相正交?

证明: 根据不等式 $\operatorname{rank}(XY)\geq\operatorname{rank}(X)+\operatorname{rank}(Y)-n$ 可得:

$$0=\operatorname{rank}(A^3)\geq\operatorname{rank}(A^2)+\operatorname{rank}(A)-6\geq\operatorname{rank}(A)+\operatorname{rank}(A)-6+\operatorname{rank}(A)-6,$$

所以 $\operatorname{rank}(A)\leq 4$. 事实上，取 $A=\begin{bmatrix}0&1&&&&\\&0&1&&&\\&&0&0&&\\&&&0&1&\\&&&&0&1\\&&&&&0\end{bmatrix}$，就有 $A^3=0,\operatorname{rank}(A)=4$.

- $\mathcal{R}(A)=\operatorname{Span}(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_4,\boldsymbol{e}_5),\mathcal{N}(A)=\operatorname{Span}(\boldsymbol{e}_1,\boldsymbol{e}_4),\mathcal{R}(A^{\mathrm{T}})=\operatorname{Span}(\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_5,\boldsymbol{e}_6),\mathcal{N}(A^T)=\operatorname{Span}(\boldsymbol{e}_3,\boldsymbol{e}_6)$
- $\mathcal{R}(A^2)=\operatorname{Span}(\boldsymbol{e}_1,\boldsymbol{e}_4),\mathcal{N}(A)=\operatorname{Span}(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_4,\boldsymbol{e}_5),\mathcal{R}((A^2)^{\mathrm{T}})=\operatorname{Span}(\boldsymbol{e}_3,\boldsymbol{e}_6),\mathcal{N}((A^2)^T)=\operatorname{Span}(\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_5,\boldsymbol{e}_6)$

哪些子空间互相正交请自行判断。

□

练习 3.3.6. 设 $\mathbb{R}^n$ 的子空间 $\mathcal{M}=\operatorname{Span}(\boldsymbol{a}_1,\boldsymbol{a}_2,\dots,\boldsymbol{a}_s)$，证明，$\mathcal{M}^\perp=\left\{\boldsymbol{b}\in\mathbb{R}^n:\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}_i=0,i=1,2,\dots,s\right\}$.

证明: 我们先明确一下定义:

$$\mathcal{M}^{\perp} := \left\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x}^{\mathrm{T}}\boldsymbol{a} = 0, \forall \boldsymbol{a} \in \mathcal{M}\right\}.$$

我们记 $\mathcal{N} = \left\{\boldsymbol{b} \in \mathbb{R}^n : \boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}_i = 0, i = 1, 2, \ldots, s\right\}$.

- 任取 $\boldsymbol{x} \in \mathcal{M}^{\perp}$, 由于 $\boldsymbol{a}_i \in \mathcal{M}$, 那么 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{a}_i = 0$, 这说明 $\boldsymbol{x} \in \mathcal{N}$; 即 $\mathcal{M}^{\perp} \subseteq \mathcal{N}$.
- 任取 $\boldsymbol{x} \in \mathcal{N}$ 以及 $\boldsymbol{a} \in \mathcal{M}$. 因为 $\mathcal{M} = \mathrm{Span}(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_s)$, 那么 $\boldsymbol{a} = y_1\boldsymbol{a}_1 + \cdots + y_s\boldsymbol{a}_s$, 所以

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{a} = \sum_{i=1}^{s} y_i \boldsymbol{x}^{\mathrm{T}}\boldsymbol{a}_i = 0$$

  这说明 $\boldsymbol{x} \in \mathcal{M}^{\perp}$, 即 $\mathcal{N} \subseteq \mathcal{M}^{\perp}$.

□

练习 3.3.7. 对向量组 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$, 定义 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}^{\perp}$ 为与这些向量都正交的向量所构成的子集.

1. 证明 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}^{\perp}$ 是一个子空间.
2. 构造矩阵 A, 使得 $\mathcal{N}(A) = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}^{\perp}$.
3. 证明 $(\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}^{\perp})^{\perp} = \mathrm{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k)$.

证明:
- 记 $W = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}^{\perp}$. 对任意 $\boldsymbol{w} \in W, a \in \mathbb{R}$, $(a\boldsymbol{w})^{\mathrm{T}}\boldsymbol{v}_i = a(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}_i) = 0$, 所以 $a\boldsymbol{w} \in W$; 对任意 $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$, $(\boldsymbol{w}_1+\boldsymbol{w}_2)^{\mathrm{T}}\boldsymbol{v}_i = \boldsymbol{w}_1^{\mathrm{T}}\boldsymbol{v}_i+\boldsymbol{w}_2^{\mathrm{T}}\boldsymbol{v}_i = 0$, 所以 $\boldsymbol{w}_1+\boldsymbol{w}_2 \in W$. 这说明 W 是子空间。
- 令 $A = \begin{bmatrix} \boldsymbol{v}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{v}_k^{\mathrm{T}} \end{bmatrix}$. 那么

$$\boldsymbol{x} \in \mathcal{N}(A) \Leftrightarrow A\boldsymbol{x} = 0 \Leftrightarrow \begin{bmatrix} \boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{x} \\ \vdots \\ \boldsymbol{v}_k^{\mathrm{T}}\boldsymbol{x} \end{bmatrix} = 0 \Leftrightarrow \boldsymbol{x} \in \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}^{\perp}$$

- $(\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}^{\perp})^{\perp} = (\mathcal{N}(A))^{\perp} = \mathcal{R}(A^{\mathrm{T}}) = \mathrm{Span}\,(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k)$.

□

练习 3.3.8. 集合运算有 *De Morgan* 定律: 对给定集合的两个子集 X, Y, $X \cap Y$ 的补集等于 X 的补集与 Y 的补集的并集; $X \cup Y$ 的补集等于 X 的补集与 Y 的补集的交集. 子空间是否也有类似的法则呢? 设 $\mathbb{R}^m$ 的两个子空间 $\mathcal{M}, \mathcal{N}$, 不妨设存在矩阵 A, B, 使得 $\mathcal{M} = \mathcal{R}(A), \mathcal{N} = \mathcal{R}(B)$.

1. $\mathcal{M} + \mathcal{N}$ 是哪个矩阵的列空间? 因此, $(\mathcal{M} + \mathcal{N})^{\perp}$ 是该矩阵的什么空间?
2. $\mathcal{M}^{\perp}, \mathcal{N}^{\perp}, \mathcal{M}^{\perp} \cap \mathcal{N}^{\perp}$ 分别是哪个矩阵的零空间?
3. 证明 *De Morgan* 定律的子空间版本: $(\mathcal{M} + \mathcal{N})^{\perp} = \mathcal{M}^{\perp} \cap \mathcal{N}^{\perp}, (\mathcal{M} \cap \mathcal{N})^{\perp} = \mathcal{M}^{\perp} + \mathcal{N}^{\perp}$.

证明:
1. $\mathcal{M} + \mathcal{N}$ 是矩阵 $\begin{bmatrix} A & B \end{bmatrix}$ 的列空间, $(\mathcal{M} + \mathcal{N})^{\perp}$ 是该矩阵的左零空间。
2. $\mathcal{M}^{\perp}, \mathcal{N}^{\perp}, \mathcal{M}^{\perp} \cap \mathcal{N}^{\perp}$ 分别是矩阵 $A, B, \begin{bmatrix} A & B \end{bmatrix}$ 的左零空间。

3.
- $\boldsymbol{x} \in (\mathcal{M}+\mathcal{N})^{\perp} \Leftrightarrow \forall \boldsymbol{v} \in \mathcal{M}, \boldsymbol{w} \in \mathcal{N}, \boldsymbol{x}\cdot\boldsymbol{v}=0, \boldsymbol{x}\cdot\boldsymbol{w}=0 \Leftrightarrow \boldsymbol{x} \in \mathcal{M}^{\perp}\cap\mathcal{N}^{\perp}$
- 将前一个结论中的 $\mathcal{M},\mathcal{N}$ 分别换成 $\mathcal{M}^{\perp},\mathcal{N}^{\perp}$，我们有

$$(\mathcal{M}^{\perp}+\mathcal{N}^{\perp})^{\perp} = (\mathcal{M}^{\perp})^{\perp}\cap(\mathcal{N}^{\perp})^{\perp} = \mathcal{M}\cap\mathcal{N}$$

再取正交即有 $(\mathcal{M}\cap\mathcal{N})^{\perp} = \mathcal{M}^{\perp}+\mathcal{N}^{\perp}$.

□

练习 3.3.9. 设 $\boldsymbol{a}_1 = \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \mathcal{M} = \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2)$ 是 $\mathbb{R}^3$ 的子空间，求 $\boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 在 $\mathcal{M}$ 上的正交投影.

证明: 瞪眼一看，$\boldsymbol{a}_1, \boldsymbol{a}_2$ 线性无关，那么我们就可以套公式了：令 $A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 \end{bmatrix}$, 那么求 $\boldsymbol{b}$ 在 $\mathcal{M}$ 上的正交投影为

$$A(A^{\mathrm{T}}A)^{-1}A^{T}\boldsymbol{b} = \begin{bmatrix} 8/9 \\ 11/9 \\ 7/9 \end{bmatrix}$$

□

练习 3.3.10. 设 $\mathcal{L}$ 是 $\mathbb{R}^3$ 中的直线：

$$\mathcal{L}: \begin{cases} x_1 + x_2 + x_3 = 0, \\ 2x_1 - x_2 - 2x_3 = 0. \end{cases}$$

求 $\boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ 在直线 $\mathcal{L}$ 上的正交投影.

证明: 我们解方程组，得到 $\mathcal{L}$ 的一组基 $\begin{bmatrix} \frac{1}{3} \\ -\frac{4}{3} \\ 1 \end{bmatrix}$, 然后代公式就有 $\boldsymbol{b}$ 在直线 $\mathcal{L}$ 上的正交投影是

$$\begin{bmatrix} \frac{1}{3} \\ -\frac{4}{3} \\ 1 \end{bmatrix} \left(\begin{bmatrix} \frac{1}{3} & -\frac{4}{3} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ -\frac{4}{3} \\ 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} \frac{1}{3} & -\frac{4}{3} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{13} \\ -\frac{8}{13} \\ \frac{6}{13} \end{bmatrix}$$

□

练习 3.3.11. 设 $\mathcal{M}$ 是 $\mathbb{R}^3$ 中由方程 $x_1 - x_2 + x_3 = 0$ 决定的平面，求 $\boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ 在平面 $\mathcal{M}$ 上的正交投影，并求 $\boldsymbol{b}$ 到平面 $\mathcal{M}$ 的距离.

证明: 我们可以算出 $\mathcal{M}$ 的一组基，然后代公式，求正交投影。然而，我们知道 $\mathcal{M}^{\perp}$ 的一组基是 $\begin{bmatrix}1\\-1\\1\end{bmatrix}$, 因此我们可以先求出到 $\mathcal{M}^{\perp}$ 的正交投影来。

我们直接代公式就有 $\boldsymbol{b}$ 在 $\mathcal{M}^{\perp}$ 的正交投影是

$$\begin{bmatrix}1\\-1\\1\end{bmatrix}\left(\begin{bmatrix}1&-1&1\end{bmatrix}\begin{bmatrix}1\\-1\\1\end{bmatrix}\right)^{-1}\begin{bmatrix}1&-1&1\end{bmatrix}\begin{bmatrix}1\\1\\2\end{bmatrix}=\begin{bmatrix}\frac{2}{3}\\-\frac{2}{3}\\\frac{2}{3}\end{bmatrix}$$

$\boldsymbol{b}$ 到平面 $\mathcal{M}$ 的距离是 $\left\|\begin{bmatrix}\frac{2}{3}\\-\frac{2}{3}\\\frac{2}{3}\end{bmatrix}\right\|=\frac{2\sqrt{3}}{3}$, $\boldsymbol{b}$ 在平面 $\mathcal{M}$ 上的正交投影是

$$\boldsymbol{b}-\begin{bmatrix}\frac{2}{3}\\-\frac{2}{3}\\\frac{2}{3}\end{bmatrix}=\begin{bmatrix}\frac{1}{3}\\\frac{5}{3}\\\frac{4}{3}\end{bmatrix}$$

□

练习 3.3.12. 将下列问题中的 $\boldsymbol{x}$ 分解成 $\mathcal{N}(A)$ 与 $\mathcal{R}(A^{\mathrm{T}})$ 中向量的和:

1. $\boldsymbol{x}=\begin{bmatrix}2\\0\end{bmatrix}, A=\begin{bmatrix}1&-1\\0&0\\0&0\end{bmatrix}$. *2.* $\boldsymbol{x}=\begin{bmatrix}1\\0\end{bmatrix}, A=\begin{bmatrix}1&1\\1&1\end{bmatrix}$. *3.* $\boldsymbol{x}=\begin{bmatrix}5\\5\\9\end{bmatrix}, A=\begin{bmatrix}1&2&3\\2&2&4\\2&3&5\end{bmatrix}$.

由于 $\mathcal{N}(A)$ 和 $\mathcal{R}(A^{\mathrm{T}})$ 互为正交补空间，因此我们实际上只需要算一个正交投影即可。这两个空间哪个维数少，我们算哪个；哪个算起来简单，我们算哪个。

证明:

1. $\mathcal{R}(A^{\mathrm{T}})=\mathrm{Span}\left(\begin{bmatrix}1\\-1\end{bmatrix}\right)$,

$$\begin{bmatrix}1\\-1\end{bmatrix}\left(\begin{bmatrix}1&-1\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix}\right)^{-1}\begin{bmatrix}1&-1\end{bmatrix}\begin{bmatrix}2\\0\end{bmatrix}=\begin{bmatrix}1\\-1\end{bmatrix},$$

因此 $\boldsymbol{x}=\begin{bmatrix}1\\1\end{bmatrix}+\begin{bmatrix}1\\-1\end{bmatrix}$.

2. $\mathcal{R}(A^{\mathrm{T}})=\mathrm{Span}\left(\begin{bmatrix}1\\1\end{bmatrix}\right)$,

$$\begin{bmatrix}1\\1\end{bmatrix}\left(\begin{bmatrix}1&1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}\right)^{-1}\begin{bmatrix}1&1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}1/2\\1/2\end{bmatrix},$$

因此 $\boldsymbol{x}=\begin{bmatrix}1/2\\-1/2\end{bmatrix}+\begin{bmatrix}1/2\\1/2\end{bmatrix}$.

3. $\mathcal{N}(A)=\{0\},\mathcal{R}(A^{\mathrm{T}})=\mathbb{R}^3$，因此 $\boldsymbol{x}=0+\boldsymbol{x}$.

□

练习 3.3.13. 证明命题 *3.3.12.* 给定 $\mathbb{R}^n$ 的子空间 $\mathcal{M}$ 和向量 $\boldsymbol{a}$, 而 $\boldsymbol{a}_1=P_{\mathcal{M}}(\boldsymbol{a})$ 为 $\boldsymbol{a}$ 在 $\mathcal{M}$ 上的正交投影，则 $\|\boldsymbol{a}-\boldsymbol{a}_1\|=\min_{\boldsymbol{x}\in\mathcal{M}}\|\boldsymbol{a}-\boldsymbol{x}\|$

证明: 由题意 $\boldsymbol{a}-\boldsymbol{a}_1\in\mathcal{M}^{\perp}$. 对任意 $\boldsymbol{x}\in\mathcal{M}$,

$$\|\boldsymbol{a}-\boldsymbol{x}\|=\|(\boldsymbol{a}-\boldsymbol{a}_1)+(\boldsymbol{a}_1-\boldsymbol{x})\|=\sqrt{\|\boldsymbol{a}-\boldsymbol{a}_1\|^2+\|\boldsymbol{a}_1-\boldsymbol{x}\|^2}\geq\|\boldsymbol{a}-\boldsymbol{a}_1\|$$

取等号当且仅当 $\boldsymbol{x}=\boldsymbol{a}_1$. □

练习 3.3.14. 说明满足下列条件的矩阵是否存在，如果存在，举例说明：

1. $\begin{bmatrix}1\\2\\-3\end{bmatrix},\begin{bmatrix}2\\-3\\5\end{bmatrix}\in\mathcal{R}(A),\begin{bmatrix}1\\1\\1\end{bmatrix}\in\mathcal{N}(A)$.

2. $\begin{bmatrix}1\\2\\-3\end{bmatrix},\begin{bmatrix}2\\-3\\5\end{bmatrix}\in\mathcal{R}(A^{\mathrm{T}}),\begin{bmatrix}1\\1\\1\end{bmatrix}\in\mathcal{N}(A)$.

3. $A\boldsymbol{x}=\begin{bmatrix}1\\1\\1\end{bmatrix}$ 有解，且 A 的左零空间包含 $\begin{bmatrix}1\\0\\0\end{bmatrix}$.

4. A 不是零矩阵，且 A 的每一行的转置垂直于 A 的每一列.

5. A 非零，且 $\mathcal{R}(A)=\mathcal{N}(A)$.

6. A 的所有列向量的和是零向量，且所有行向量的和是分量均为 1 的向量.

7. A,B 均为非零的正交投影矩阵，且 $A+B$ 仍是正交投影矩阵.

8. A,B 均为正交投影矩阵，但 $A+B$ 并不是正交投影矩阵.

9. A,B,C 均为非零的三阶正交投影矩阵，且 $A+B+C=I_3$.

证明: 1. 存在，例如 $A=\begin{bmatrix}1&2&-3\\2&-3&1\\-3&5&-2\end{bmatrix}$

2. 不存在，因为 $\mathcal{R}(A^{\mathrm{T}})=\mathcal{N}(A)^{\perp}$, 但是题中的向量 $\begin{bmatrix}1\\2\\-3\end{bmatrix},\begin{bmatrix}2\\-3\\5\end{bmatrix}$ 和 $\begin{bmatrix}1\\1\\1\end{bmatrix}$ 并不正交.

3. 不存在。由题意，$\begin{bmatrix}1\\1\\1\end{bmatrix}\in\mathcal{R}(A),\begin{bmatrix}1\\0\\0\end{bmatrix}\in\mathcal{N}(A^{\mathrm{T}})$. 但是这两个向量也并不正交。

4. 不存在。由题意 $A^{\mathrm{T}}A=0$, 我们计算 $A^{\mathrm{T}}A$ 的迹就知道 A 是零矩阵。

5. 存在，例如 $A=\begin{bmatrix}0&1\\0&0\end{bmatrix}$

6. 不存在。令 $\boldsymbol{x}=\begin{bmatrix}1\\\vdots\\1\end{bmatrix}$, 考虑 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$, 那么由题意：$\boldsymbol{x}^{\mathrm{T}}(A\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}0=0$, 且 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=(\boldsymbol{x}^{\mathrm{T}}A)\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=n$, 矛盾。
7. 存在，例如 $A=\begin{bmatrix}1&0\\0&0\end{bmatrix},B=\begin{bmatrix}0&0\\0&1\end{bmatrix}$
8. 存在，例如 $A=B=I_2$.
9. 存在，例如 $A=\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix},B=\begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix},C=\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}$

□

练习 3.3.15. 下列说法中，哪些正确？

1. A 的所有行的转置与 A^{-1} 的所有行的转置对应正交.
2. A 的所有行的转置与 A^{-1} 的所有列对应正交.
3. A 的所有列与 A^{-1} 的所有行的转置对应正交.
4. A 的所有列与 A^{-1} 的所有列对应正交.
5. 如果向量 $\boldsymbol{v}$ 与 $\boldsymbol{w}$ 正交，则 $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}=0$ 与 $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}=0$ 的解集互相正交.
6. 如果 A 是正交投影矩阵，则 A 的第 k 列的长度的平方等于 A 的第 k 个对角元素.
7. 如果 A,B 是正交投影矩阵，则 $AB=BA$ 当且仅当 AB 也是正交投影矩阵.
8. 如果 A,B 是正交投影矩阵，则 $A+B$ 是正交投影矩阵当且仅当 $\mathcal{R}(A),\mathcal{R}(B)$ 互相正交.

证明: 6，7，8 是正确的，其它是错误的。 □

练习 3.3.16. 如果矩阵 A 的列向量线性无关，那么向 $\mathcal{R}(A)$ 的正交投影矩阵为 $A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$. 试分析以下化简中可能出现的问题：$A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}=AA^{-1}(A^{\mathrm{T}})^{-1}A^{\mathrm{T}}=I_nI_n=I_n$.

1. 证明 $(A^{\mathrm{T}}A)^{-1}=A^{-1}(A^{\mathrm{T}})^{-1}$ 并不一定总成立.
2. 当 A 满足什么条件时，上式一定成立？试分析此时正交投影矩阵等于 I_n 的原因.

证明:
1. A 仅仅是列满秩，它如果不是方阵，无法谈论 A^{-1}.
2. 若 A^{-1} 存在，即 A 是可逆方阵，那么上式是成立的，直接做矩阵计算可知正交投影矩阵等于 I_n. 其原因是，A 可逆时，$\mathcal{R}(A)=\mathbb{R}^n$, 所以正交投影其实是不需要投影，就是恒同映射。

□

练习 3.3.17. 设 $A=\begin{bmatrix}3&6&6\\4&8&8\end{bmatrix}$. 求向 A 的列空间的正交投影矩阵 P_1 和向 A 的行空间的正交投影矩阵 P_2，并计算 P_1AP_2.

证明: $\boldsymbol{v}=\begin{bmatrix}3\\4\end{bmatrix}$ 是 $\mathcal{R}(A)$ 的一组基，那么 $P_1=\begin{bmatrix}3\\4\end{bmatrix}\left(\begin{bmatrix}3&4\end{bmatrix}\begin{bmatrix}3\\4\end{bmatrix}\right)^{-1}\begin{bmatrix}3&4\end{bmatrix}=\frac{1}{25}\begin{bmatrix}9&12\\12&25\end{bmatrix}$

$$\boldsymbol{w}\begin{bmatrix}3\\6\\6\end{bmatrix} \text{ 是 } \mathcal{R}(A^T) \text{ 的一组基，那么 } P_2=\begin{bmatrix}3\\6\\6\end{bmatrix}\left(\begin{bmatrix}3&6&6\end{bmatrix}\begin{bmatrix}3\\6\\6\end{bmatrix}\right)^{-1}\begin{bmatrix}3&6&6\end{bmatrix}=\tfrac{1}{9}\begin{bmatrix}1&2&2\\2&4&4\\2&4&4\end{bmatrix}$$

$$P_1AP_2=(\boldsymbol{v}(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v})^{-1}\boldsymbol{v}^{\mathrm{T}})(\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}})(\boldsymbol{w}(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w})^{-1}\boldsymbol{w})=\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}=A.$$

□

练习 3.3.18. 给定 $A=\begin{bmatrix}1&0\\2&1\\0&1\end{bmatrix}$. 设 P_1 是关于 A 的第一列的正交投影矩阵，P_2 是关于 A 的正交投影矩阵，计算 P_2P_1.

证明:

$$P_1=\begin{bmatrix}1\\2\\0\end{bmatrix}\left(\begin{bmatrix}1&2&0\end{bmatrix}\begin{bmatrix}1\\2\\0\end{bmatrix}\right)^{-1}\begin{bmatrix}1&2&0\end{bmatrix}=\frac{1}{5}\begin{bmatrix}1&2&0\\2&4&0\\0&0&0\end{bmatrix}$$

$$P_2=\begin{bmatrix}1&0\\2&1\\0&1\end{bmatrix}\left(\begin{bmatrix}1&2&0\\0&1&1\end{bmatrix}\begin{bmatrix}1&0\\2&1\\0&1\end{bmatrix}\right)^{-1}\begin{bmatrix}1&2&0\\0&1&1\end{bmatrix}=\frac{1}{6}\begin{bmatrix}2&2&-2\\2&5&1\\-2&1&5\end{bmatrix}$$

P_2P_1 对应的线性映射是先向第一列做正交投影，然后再向 A 的列空间作正交投影；那么第二步实际上是无效步骤，所以 $P_2P_1=P_1$.

□

练习 3.3.19. 一个 n 阶方阵 P 是关于 A 的正交投影矩阵，当且仅当对任意向量 $\boldsymbol{x}\in\mathbb{R}^n$，都有 $P\boldsymbol{x}\in\mathcal{R}(A), \boldsymbol{x}-P\boldsymbol{x}\in\mathcal{N}(A^{\mathrm{T}})$.

证明: 这个没什么好证明的，仅仅是 $\mathcal{N}(A^{\mathrm{T}})=\mathcal{R}(A)^{\perp}$ 的直接推论而已。 □

练习 3.3.20. 当 A 分别为对称矩阵、反对称矩阵、正交矩阵或上三角矩阵时，判断 $\mathcal{R}(A)$ 和 $\mathcal{N}(A)$ 是否互为正交补. 证明或给出反例.

证明: 因为 $\mathcal{R}(A)^{\perp}=\mathcal{N}(A^{\mathrm{T}})$; 那么 $\mathcal{R}(A)^{\perp}=\mathcal{N}(A)$ 当且仅当 $\mathcal{N}(A^{\mathrm{T}})=\mathcal{N}(A)$. 因此当矩阵 A 是对称矩阵或者反对称矩阵时，$\mathcal{R}(A)$ 和 $\mathcal{N}(A)$ 互为正交补。

若 A 是正交阵，则 A 可逆，此时 $\mathcal{N}(A^{\mathrm{T}})=\mathcal{N}(A)=\{0\}$, 所以 $\mathcal{R}(A)$ 和 $\mathcal{N}(A)$ 互为正交补。

A 是上三角矩阵时，我们有反例：$A=\begin{bmatrix}0&1\\0&0\end{bmatrix}, \mathcal{N}(A)=\mathrm{Span}(\boldsymbol{e}_1), \mathcal{N}(A^{\mathrm{T}})=\mathrm{Span}(\boldsymbol{e}_2)$. □

练习 3.3.21. 给定 $\mathbb{R}^m$ 中的子空间 $\mathcal{M}_1,\mathcal{M}_2$，$\mathbb{R}^n$ 中的子空间 $\mathcal{N}_1,\mathcal{N}_2$.

1. 是否一定存在矩阵 A，使得 $\mathcal{R}(A)=\mathcal{M}_1, \mathcal{N}(A^{\mathrm{T}})=\mathcal{M}_2, \mathcal{R}(A^{\mathrm{T}})=\mathcal{N}_1, \mathcal{N}(A)=\mathcal{N}_2$?
2. 如果不一定存在，那么当四个子空间满足什么条件时，这样的矩阵才一定存在?

证明: 不一定存在。这些子空间需要满足矩阵四个基本子空间的关系时，才存在：

- $\dim\mathcal{M}_1=\dim\mathcal{N}_1$;

- $\dim \mathcal{M}_1 + \dim \mathcal{N}_2 = n$;
- $\mathcal{M}_1 \perp \mathcal{M}_2$;
- $\mathcal{N}_1 \perp \mathcal{N}_2$.

此时，为了说明存在性，我们只需要构造矩阵 A，使得 $\mathcal{R}(A) = \mathcal{M}_1, \mathcal{R}(A^{\mathrm{T}}) = \mathcal{N}_1$。设 $r = \dim \mathcal{M}_1 = \dim \mathcal{N}_1$, 我们取 $\mathcal{M}_1$ 的一组标准正交基 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_r$, $\mathcal{N}_1$ 的一组标准正交基 $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_r$, 令 $A = \boldsymbol{v}_1\boldsymbol{w}_1^T + \cdots + \boldsymbol{v}_r\boldsymbol{w}_r^{\mathrm{T}}$, 那么显然 $\mathcal{R}(A) \subseteq \mathrm{Span}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_r) = \mathcal{M}_1$; 另一方面，$A\boldsymbol{w}_i = \sum_{j=1}^{r} \boldsymbol{v}_j\boldsymbol{w}_j^{\mathrm{T}}\boldsymbol{w}_i = \boldsymbol{v}_i$, 因此 $\boldsymbol{v}_i \in \mathcal{R}(A)$, 所以 $\mathcal{M} \subseteq \mathcal{R}(A)$. 总而言之，$\mathcal{R}(A) = \mathcal{M}_1$; 同理可证 $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{N}_1$. □

练习 3.3.22. 给定向量 $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$，不难确知其平均值 $\boldsymbol{x}_m = \dfrac{1}{m}\displaystyle\sum_{i=1}^{m} \boldsymbol{b}_i$. 再给出向量 $\boldsymbol{b}_{m+1}$，这 $m+1$ 个向量的平均值就是 $\boldsymbol{x}_{m+1} = \dfrac{1}{m+1}\displaystyle\sum_{i=1}^{m+1} \boldsymbol{b}_i$. 则 $\boldsymbol{x}_{m+1}$ 可以用 $\boldsymbol{b}_{m+1}, \boldsymbol{x}_m, m$ 来表示：$\boldsymbol{x}_{m+1} = \boldsymbol{x}_m + \dfrac{1}{m+1}(\boldsymbol{b}_{m+1} - \boldsymbol{x}_m)$.

若给出向量组成的矩阵 $B_n = \begin{bmatrix} \boldsymbol{b}_{m+1} & \cdots & \boldsymbol{b}_{m+n} \end{bmatrix}$,如何用 $B_n, \boldsymbol{x}_m, m, n$ 来表示 $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_{m+n}$ 这 $m+n$ 个向量的平均值 $\boldsymbol{x}_{m+n}$?

证明:

$$(m+n)\boldsymbol{x}_{m+n} = \sum_{i=1}^{m+n} \boldsymbol{b}_i = \sum_{i=1}^{m} \boldsymbol{b}_i + \sum_{i=m+1}^{m+n} \boldsymbol{b}_i = mx_m + B_n \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

所以

$$\boldsymbol{x}_{m+n} = x_m + \frac{1}{m+n}\left(B_n \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} - n\boldsymbol{x}_m \right)$$

□

练习 3.3.23. 一个方阵如果仅仅满足 $P^2 = P$，则称之为**斜投影矩阵**，其对应的线性变换称为**斜投影**. 给定一个 n 阶斜投影矩阵 P.

1. 证明 $I_n - P$ 也是 n 阶斜投影矩阵.
2. 证明 $\mathcal{R}(P) = \mathcal{N}(I_n - P), \mathcal{R}(I_n - P) = \mathcal{N}(P)$.
3. 对任意向量 $\boldsymbol{v} \in \mathbb{R}^n$，是否一定存在分解 $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$，满足 $\boldsymbol{v}_1 \in \mathcal{R}(P), \boldsymbol{v}_2 \in \mathcal{R}(I_n - P)$? 分解如果存在，是否唯一？
4. 构造一个二阶斜投影矩阵，但不是正交投影矩阵.

证明: 1. $(I_n - P)^2 = I_n - 2P + P^2 = I_n - P$, 所以 $I_n - P$ 是斜投影矩阵。

2. 对任意 $\boldsymbol{v} \in \mathcal{R}(P)$, 则存在 $\boldsymbol{x} \in \mathbb{R}^n$ 使得 $\boldsymbol{v} = P\boldsymbol{x}$, 那么 $(I_n - P)\boldsymbol{v} = (I_n - P)P\boldsymbol{x} = 0$, 这就是说 $\mathcal{R}(P) \subseteq \mathcal{N}(I_n - P)$. 对任意 $\boldsymbol{v} \in \mathcal{N}(I_n - P)$，则 $(I_n - P)\boldsymbol{v} = 0$, 那么 $\boldsymbol{v} = P\boldsymbol{v} \in \mathcal{R}(A)$, 这就是说 $\mathcal{N}(I_n - P) \subseteq \mathcal{R}(P)$. 所以 $\mathcal{R}(P) = \mathcal{N}(I_n - P)$.
   因为 $I_n - P$ 也是斜投影矩阵，在上一个等式中，我们可以用 $I_n - P$ 来替换 P，这就得到了 $\mathcal{R}(I_n - P) = \mathcal{N}(P)$.

3. 对任意向量 $\boldsymbol{v}\in\mathbb{R}^n$, $\boldsymbol{v}=P\boldsymbol{v}+(I_n-P)\boldsymbol{v}$, 那么 $P\boldsymbol{v}\in\mathcal{R}(P),(I_n-P)\boldsymbol{v}\in\mathcal{R}(I_n-P)$, 所以分解存在。若存在 $\boldsymbol{v}\in\mathbb{R}^n$ 有两个这样的分解：$\boldsymbol{v}=\boldsymbol{v}_1+\boldsymbol{v}_2=\boldsymbol{w}_1+\boldsymbol{w}_2$，那么 $\boldsymbol{x}:=\boldsymbol{v}_1-\boldsymbol{w}_1=\boldsymbol{w}_2-\boldsymbol{v}_2\in\mathcal{R}(P)\cap\mathcal{R}(I_n-P)=\mathcal{N}(I_n-P)\cap\mathcal{N}(P)$. 那么 $(I_n-P)\boldsymbol{x}=P\boldsymbol{x}=0$，于是 $\boldsymbol{x}=(I_n-P)\boldsymbol{x}+P\boldsymbol{x}=0$, 所以 $\boldsymbol{v}_1=\boldsymbol{w}_1,\boldsymbol{v}_2=\boldsymbol{w}_2$，即分解唯一。

4. 我们任取一组非正交基，例如 $\boldsymbol{v}_1:=\begin{bmatrix}1\\0\end{bmatrix},\boldsymbol{v}_2:=\begin{bmatrix}1\\1\end{bmatrix}$；对任意 $\boldsymbol{v}\in\mathbb{R}^n$, $\boldsymbol{v}$ 可以唯一第写成 $x_1\boldsymbol{v}_1+x_2\boldsymbol{v}_2$ 的形式。令 P 满足 $P\boldsymbol{v}=x_1\boldsymbol{v}_1$. 那么 $P=\begin{bmatrix}1&-1\\0&0\end{bmatrix}$, 直接验证 $P^2=P$.

□

练习 3.3.24. 设平面上的四个点 $\begin{bmatrix}x_i\\y_i\end{bmatrix}$ 分别是 $\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}1\\8\end{bmatrix},\begin{bmatrix}3\\8\end{bmatrix},\begin{bmatrix}4\\20\end{bmatrix}$，利用最小二乘法求下列直线或曲线：

1. 求平行于 x 轴的直线 $y=b$ 使得 $\displaystyle\sum_{i=1}^4|y_i-b|^2$ 最小.

2. 求经过原点的直线 $y=kx$ 使得 $\displaystyle\sum_{i=1}^4|y_i-kx_i|^2$ 最小.

3. 求直线 $y=kx+b$ 使得 $\displaystyle\sum_{i=1}^4|y_i-(kx_i+b)|^2$ 最小.

4. 求抛物线 $y=ax^2+bx+c$ 使得 $\displaystyle\sum_{i=1}^4\left|y_i-(ax_i^2+bx_i+c)\right|^2$ 最小.

证明*:* 我们按照固定的套路来做。

1. 考虑线性方程组 $Ab=\boldsymbol{y}$, 其中 $A=\begin{bmatrix}1\\1\\1\\1\end{bmatrix},\boldsymbol{y}=\begin{bmatrix}0\\8\\8\\20\end{bmatrix}$. 这个方程组的最小二乘解为

$$b_0=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{y}=9$$

那么要求的最小值就是

$$\min\left(\sum_{i=1}^4|y_i-b|^2\right)=\|\boldsymbol{y}-Ab_0\|=\left\|\begin{bmatrix}-9\\-1\\-1\\11\end{bmatrix}\right\|=\sqrt{204}$$

2. 考虑线性方程组 $Ak=\boldsymbol{y}$, 其中 $A=\begin{bmatrix}0\\1\\3\\4\end{bmatrix},\boldsymbol{y}=\begin{bmatrix}0\\8\\8\\20\end{bmatrix}$. 这个方程组的最小二乘解为

$$k_0=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{y}=\frac{56}{13}$$

那么要求的最小值就是

$$\min\left(\sum_{i=1}^{4}|y_i-kx_i|^2\right)=\|\boldsymbol{y}-Ak_0\|=\left\|\begin{bmatrix}-\frac{56}{13}\\ \frac{48}{13}\\ -\frac{64}{13}\\ \frac{36}{13}\end{bmatrix}\right\|=\frac{4\sqrt{677}}{13}$$

3. 考虑线性方程组 $A\boldsymbol{x}=\boldsymbol{y}$, 其中 $A=\begin{bmatrix}0&1\\1&1\\3&1\\4&1\end{bmatrix}$, $\boldsymbol{x}=\begin{bmatrix}k\\b\end{bmatrix}$, $\boldsymbol{y}=\begin{bmatrix}0\\8\\8\\20\end{bmatrix}$. 这个方程组的最小二乘解为

$$\boldsymbol{x}_0=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{y}=\frac{1}{6}\begin{bmatrix}-2\\49\end{bmatrix}$$

那么要求的最小值就是

$$\min\left(\sum_{i=1}^{4}|y_i-kx_i|^2\right)=\|\boldsymbol{y}-A\boldsymbol{x}_0\|=\left\|\begin{bmatrix}-\frac{49}{6}\\ \frac{1}{6}\\ \frac{5}{6}\\ \frac{79}{6}\end{bmatrix}\right\|=\frac{\sqrt{2167}}{3}$$

4. 考虑线性方程组 $A\boldsymbol{x}=\boldsymbol{y}$, 其中 $A=\begin{bmatrix}0&0&1\\1&1&1\\9&3&1\\16&4&1\end{bmatrix}$, $\boldsymbol{x}=\begin{bmatrix}a\\b\\c\end{bmatrix}$, $\boldsymbol{y}=\begin{bmatrix}0\\8\\8\\20\end{bmatrix}$. 这个方程组的最小二乘解为

$$\boldsymbol{x}_0=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{y}=\begin{bmatrix}\frac{2}{3}\\ \frac{4}{3}\\ 2\end{bmatrix}$$

那么要求的最小值就是

$$\min\left(\sum_{i=1}^{4}|y_i-kx_i|^2\right)=\|\boldsymbol{y}-A\boldsymbol{x}_0\|=\left\|\begin{bmatrix}-2\\4\\-4\\2\end{bmatrix}\right\|=2\sqrt{10}$$

$\square$

第四章　行列式

4.2　行列式函数

练习 4.2.1. 计算下列行列式：

1. $\begin{vmatrix} 1 & 2 & 1 & 4 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 3 \end{vmatrix}$.　*2.* $\begin{vmatrix} 1 & -1 & 2 \\ 3 & 2 & 1 \\ 0 & 1 & 4 \end{vmatrix}$.　*3.* $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$.

4. $\begin{vmatrix} x & y & x+y \\ y & x+y & x \\ x+y & x & y \end{vmatrix}$.　*5.* $\begin{vmatrix} 1+x & 1 & 1 & 1 \\ 1 & 1-x & 1 & 1 \\ 1 & 1 & 1+y & 1 \\ 1 & 1 & 1 & 1-y \end{vmatrix}$.

证明: 1. 这是上三角矩阵，行列式就等于对角线元素的成绩，因此行列式是 -6.

2. 我们稍微做点初等行变换：

$$\begin{vmatrix} 1 & -1 & 2 \\ 3 & 2 & 1 \\ 0 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 2 \\ 0 & 5 & -5 \\ 0 & 1 & 4 \end{vmatrix} = 5 \times 4 - (-5) \times 1 = 25$$

3. 我们还是要作初等行变换

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 10 & 10 & 10 & 10 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix} = 10 \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & -2 & -1 \\ 0 & -3 & -2 & -1 \end{vmatrix}$$
$$= 10 \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 4 & -4 \end{vmatrix} = 160$$

4. 我们作初等行列变换

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

5. 我们作初等行列变换

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

□

练习 4.2.2. 设 A 是三阶方阵，$\det(A)=5$，求下列矩阵 B 的行列式：

1. $B = 2A, -A, A^2, A^{-1}$.

2. $B = \begin{bmatrix} \boldsymbol{a}_1^{\mathrm{T}} - \boldsymbol{a}_3^{\mathrm{T}} \\ \boldsymbol{a}_2^{\mathrm{T}} - \boldsymbol{a}_1^{\mathrm{T}} \\ \boldsymbol{a}_3^{\mathrm{T}} - \boldsymbol{a}_2^{\mathrm{T}} \end{bmatrix}$，$\begin{bmatrix} \boldsymbol{a}_1^{\mathrm{T}} + \boldsymbol{a}_3^{\mathrm{T}} \\ \boldsymbol{a}_2^{\mathrm{T}} + \boldsymbol{a}_1^{\mathrm{T}} \\ \boldsymbol{a}_3^{\mathrm{T}} + \boldsymbol{a}_2^{\mathrm{T}} \end{bmatrix}$，其中 $A = \begin{bmatrix} \boldsymbol{a}_1^{\mathrm{T}} \\ \boldsymbol{a}_2^{\mathrm{T}} \\ \boldsymbol{a}_3^{\mathrm{T}} \end{bmatrix}$.

证明*:* 1.
- $\det(B) = 2^3\det(A) = 40$.
- $\det(B) = (-1)^3\det(A) = -5$.
- $\det(B) = \det(A)^2 = 25$.
- $\det(B) = \det(A)^{-1} = \frac{1}{5}$.

2.
- $B = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} A$, 所以 $\det(B) = \det\left(\begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}\right)\det(A) = 0$
- $B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} A$, 所以 $\det(B) = \det\left(\begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}\right)\det(A) = 10$

□

练习 4.2.3. 设 $A_n = \begin{bmatrix} -1 & 1 \\ -6 & 4 \end{bmatrix} + nI_2$.

1. 求 A_0, A_1, A_2, A_3 的行列式.

2. 求 $\begin{bmatrix}-1 & 1\\ -6 & 4\end{bmatrix}+xI_2$ 的行列式，并将其写成 $(x+a)(x+b)$ 的形式.

3. 分别求 $A_0^2, A_0^2+I_2, A_0^2+3A_0+2I_2, A_0^3-2A_0^2+3A_0-4I_2$ 的行列式，并分析它们与 a,b 的关系.

证明: 1. $\det(A_0)=\begin{vmatrix}-1 & 1\\ -6 & 4\end{vmatrix}=2$; $\det(A_1)=\begin{vmatrix}0 & 1\\ -6 & 5\end{vmatrix}=6$; $\det(A_2)=\begin{vmatrix}1 & 1\\ -6 & 6\end{vmatrix}=12$; $\det(A_3)=\begin{vmatrix}2 & 1\\ -6 & 7\end{vmatrix}=20$.

2. $\det\left(\begin{bmatrix}-1 & 1\\ -6 & 4\end{bmatrix}+xI_2\right)=\begin{vmatrix}-1+x & 1\\ -6 & 4+x\end{vmatrix}=(4+x)(-1+x)+6=x^2+3x+2=(x+1)(x+2)$.

3.
   - $\det(A_0^2)=\det(A_0)^2=4$.
   - $A_0^2+I_2=\begin{bmatrix}-4 & 3\\ -18 & 11\end{bmatrix}, \det(A_0^2+I_2)=10$.
   - $\det(A_0^2+3A_0^2+2I_2)=\det(A_0+2I_2)\det(A_0+I_2)=72$
   - $A_0^3-2A_0+3A_0-4I_2=4A_0-6I_2=\begin{bmatrix}-10 & 4\\ -24 & 10\end{bmatrix}$, 所以 $\det(A_0^3-2A_0+3A_0-4I_2=4A_0-6I_2=-4$

   设 $g(X)$ 是一个多项式，那么 $\det(g(A_0))=g(a)g(b)$.

□

练习 4.2.4. 计算 $\det(A)$:

1. $A=\Big[i+j\Big]_{n\times n}$.

2. $A=\Big[ij\Big]_{n\times n}$.

证明: 1. $n=1$ 时 $A=[2]$, 则 $\det(A)=2$; $n=2$ 时，$A=\begin{bmatrix}2 & 3\\ 3 & 4\end{bmatrix}$, 则 $\det(A)=-1$. $n\geq 3$ 时，若 $A=\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 & \cdots\end{bmatrix}$，则 $\boldsymbol{a}_1-2\boldsymbol{a}_2+\boldsymbol{a}_3=0$, 因此 A 不可逆，则 $\det(A)=0$;

2. $n=1$ 时，$A=[1]$，则 $\det(A)=1$; $n\geq 2$ 时，A 的前两列成比例，因此 $\det(A)=0$

□

练习 4.2.5. 计算 $\begin{vmatrix}1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1\\ -1 & & & 1\end{vmatrix}$.

证明: 我们把第 1 至 $(n-1)$ 列都加到第 n 列, 此时第 n 列是 0, 所以 $\begin{vmatrix}1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1\\ -1 & & & 1\end{vmatrix}=0$ □

练习 4.2.6. 计算 $\begin{vmatrix} 1+x_1y_1 & 1+x_1y_2 & \cdots & 1+x_1y_n \\ 1+x_2y_1 & 1+x_2y_2 & \cdots & 1+x_2y_n \\ \vdots & \vdots & & \vdots \\ 1+x_ny_1 & 1+x_ny_2 & \cdots & 1+x_ny_n \end{vmatrix}$.

证明:
- $n=1$ 时，$A=[1+x_1y_1]$, 那么 $\det(A)=1+x_1y_1$.
- $n=2$ 时，$A=\begin{bmatrix} 1+x_1y_1 & 1+x_1y_2 \\ 1+x_2y_1 & 1+x_2y_2 \end{bmatrix}$,

$$\det(A)=(1+x_1y_1)(1+x_2y_2)-(1+x_1y_2)(1+x_2y_2)=(x_1-x_2)(y_1-y_2).$$

- $n\geq 3$ 时，$A=\begin{bmatrix}1\\ \vdots \\ 1\end{bmatrix}\begin{bmatrix}1 & \cdots & 1\end{bmatrix}+\begin{bmatrix}x_1\\ \vdots \\ x_n\end{bmatrix}\begin{bmatrix}y_1 & \cdots & y_n\end{bmatrix}$, 所以 $\operatorname{rank}(A)\leq 2$, A 不满秩，因此 $\det(A)=0$

□

练习 4.2.7. 计算 $\begin{vmatrix} a_{11} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & \cdots & a_{2,n-1} & 0 \\ \vdots & \iddots & \iddots & \vdots \\ a_{n1} & 0 & \cdots & 0 \end{vmatrix}$.

证明: 交换第 1 列和第 n 列，第 2 列和第 $n-1$ 列，...，交换所以第 i 列和第 $n+1-i$ 列，我们得到上三角矩阵，对角线元素为 $a_{1,n},a_{2,n-1},\cdots,a_{n,1}$. 若 n 是偶数，我们需要交换 $\frac{n}{2}$; 若 n 是奇数，我们需要交换 $\frac{n-1}{2}$; 总而言之，需要交换 $\left[\frac{n}{2}\right]$(高斯取整函数)，所以

$$\begin{vmatrix} a_{11} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & \cdots & a_{2,n-1} & 0 \\ \vdots & \iddots & \iddots & \vdots \\ a_{n1} & 0 & \cdots & 0 \end{vmatrix}=(-1)^{\left[\frac{n}{2}\right]}a_{1,n}a_{2,n-1}\cdots a_{n,1}.$$

□

附注. 我们也可以这么看：最后一列逐次和前一列交换，直至交换至第一列，需要 $n-1$ 次；再将新矩阵的最后一列逐次和前一列交换，直至交换至第二列，需要 $n-2$ 次，以此类推，一共需要 $(n-1)+(n-2)+\cdots+1=\dfrac{n(n-1)}{2}$ 次，所以行列式等于

$$(-1)^{\frac{n(n-1)}{2}}a_{1,n}a_{2,n-1}\cdots a_{n,1}.$$

这个值与前面的结果是吻合的。

练习 4.2.8.

1. 令 A_n 是从右上到左下对角线上的元素全为 1，其余元素全为 0 的 n 阶方阵. 求 A_2,A_3,A_4,A_5 的行列式，分析其规律，推断出 A_n 的行列式.

2. 令 $A_2=\begin{bmatrix}1&1\\1&2\end{bmatrix}, A_3=\begin{bmatrix}1&1&1\\1&2&2\\1&2&3\end{bmatrix}$，以此类推 A_n. 求 A_2,A_3,A_4 的行列式，分析其规律，推断出 A_n 的行列式. 利用 LU 分解.

3. 设 A 具有 QR 分解 $A=Q\begin{bmatrix}1&2&3\\0&4&5\\0&0&6\end{bmatrix}$，求 $\det(A)$ 的所有可能值.

4. 定义 Hilbert 矩阵 [Hilbert matrix] $H_n=\left[\dfrac{1}{i+j-1}\right]_{n\times n}$. 计算 $\det(H_2),\det(H_3)$. *Hilbert* 矩阵是一种常见的难于计算的矩阵，常用来测试算法.

证明*:* 证明*:*

1. 由上题直接得到 $\det(A_n)=(-1)^{\frac{n(n-1)}{2}}$

2. $\det(A_2)=1,\det(A_3)=1,\det(A_4)=1$. 我们对 A_n 作 LU 分解，有

$$A_n=\begin{bmatrix}1&0&\cdots&0\\1&1&\ddots&0\\\vdots&&\ddots&\vdots\\1&1&\cdots&1\end{bmatrix}\begin{bmatrix}1&1&\cdots&1\\0&1&\ddots&1\\\vdots&&\ddots&\vdots\\0&\cdots&0&1\end{bmatrix}$$

所以 $\det(A_n)=1$.

3. $\det(A)=24\det(Q)$; Q 是正交矩阵则 $Q^{\mathrm T}Q=I_3$, 那么 $\det(Q)^2=1$, 所 $\det(Q)=\pm1$, suoy$\det(A)$ 的所有可能值就是 ±24.

4.
   - $H_2=\begin{bmatrix}1&\frac12\\\frac12&\frac13\end{bmatrix}=\dfrac13-\dfrac12\cdot\dfrac12=\dfrac1{12}$
   - $H_2=\begin{bmatrix}1&\frac12&\frac13\\\frac12&\frac13&\frac14\\\frac13&\frac14&\frac15\end{bmatrix}=\dfrac1{15}+\dfrac1{24}+\dfrac1{24}-\dfrac1{27}-\dfrac1{20}-\dfrac1{16}=\dfrac1{2160}$

□

□

练习 4.2.9. 证明或举出反例.

1. $AB-BA$ 的行列式必然是零.

2. A 的行列式等于其行简化阶梯形矩阵的行列式.

3. A 为 n 阶反对称矩阵，当 n 为奇数时，$\det(A)=0$.

4. A 为 n 阶反对称矩阵，当 n 为偶数时，$\det(A)=0$.

5. 如果 $|\det(A)|>1$，那么当 n 趋于无穷时，A^n 中必然有元素的绝对值趋于无穷.

6. 如果 $|\det(A)|<1$，那么当 n 趋于无穷时，A^n 中的所有元素都趋于 0.

证明*:*

1. 错误。反例：$A=\begin{bmatrix}0&1\\1&0\end{bmatrix}, B=\begin{bmatrix}1&2\\4&8\end{bmatrix}$, 则 $AB-BA=\begin{bmatrix}2&7\\-7&-2\end{bmatrix}=45$.

2. 错误。反例:$A=\begin{bmatrix}0&1\\1&0\end{bmatrix}$, $\mathrm{rref}(A)=I_2,\det(A)=-1,\det(I_2)=1$.

3. 正确。因为 $\det(A)=\det(A^{\mathrm{T}})=\det(-A)=(-1)^n\det(A)$; n 是奇数，则 $\det(A)=-\det(A)$, 所以 $\det(A)=0$.

4. 错误。反例：$A=\begin{bmatrix}0&1\\-1&0\end{bmatrix}$, $\det(A)=1$.

5. 正确。若 A 是 m 阶方阵方阵，A^n 行列式的展开式中有 $m!$ 项，若 A^n 中元素有界 M，那么 $|\det(A^n)|\le m\,M^m$ 也是有界的。但是 $|\det(A)^n|=|\det(A)|^n\to\infty$, 矛盾。

6. 错误。反例：$A=\begin{bmatrix}1&0\\0&\frac{1}{2}\end{bmatrix}$, $A^n=\begin{bmatrix}1&0\\0&\frac{1}{2^n}\end{bmatrix}$

□

练习 4.2.10. 以下均为某些学生出现过的错误. 请找到错误的原因.

1. 对二阶可逆矩阵 $A=\begin{bmatrix}a&b\\c&d\end{bmatrix}$，计算其逆矩阵的行列式：

$$\det(A^{-1})=\det\left(\frac{1}{ad-bc}\begin{bmatrix}d&-b\\-c&a\end{bmatrix}\right)=\frac{ad-bc}{ad-bc}=1.$$

这个结论很奇怪. 请问错在哪里?

2. 对分块矩阵 $M=\begin{bmatrix}A&B\\C&D\end{bmatrix}$，计算其行列式：$\det\left(\begin{bmatrix}A&B\\C&D\end{bmatrix}\right)=AD-BC$，得到的是矩阵，而不是数. 如果 A 可逆，正确的公式是什么?

3. 计算正交投影矩阵 P 的行列式 $\det(P)=\det(A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}})=\dfrac{\det(A)\det(A^{\mathrm{T}})}{\det(A^{\mathrm{T}}A)}=1$，然而正交投影矩阵常常不可逆. 错在哪里?

4. 如果 $AB=-BA$, 那么 $\det(A)\det(B)=-\det(B)\det(A)$, 由此得到 $2\det(A)\det(B)=0$，所以 A,B 必有一个矩阵不可逆. 这是否正确? 如果不是，请指出错误并举出反例.

5. 对矩阵 $\begin{bmatrix}a&b\\c&d\end{bmatrix}$，同时做两个行变换得到 $\begin{bmatrix}a+sc&b+sd\\c+ta&d+tb\end{bmatrix}$，行列式是否一定保持不变? s,t 满足什么条件时，行列式一定保持不变?

附注. 这不是初等行变换.

证明: 1. $\det\left(\dfrac{1}{ad-bc}\begin{bmatrix}d&-b\\-c&a\end{bmatrix}\right)=\dfrac{1}{(ad-bc)^2}\det\left(\begin{bmatrix}d&-b\\-c&a\end{bmatrix}\right)=\dfrac{1}{ad-bc}$

2. $\det\left(\begin{bmatrix}A&B\\C&D\end{bmatrix}\right)=AD-BC$ 这个等式是错误的，和行列式的定义无关。若 A 可逆，我们可以对 $\begin{bmatrix}A&B\\C&D\end{bmatrix}$ 做分块初等行变换：

$$\begin{bmatrix}A&B\\C&D\end{bmatrix}\longrightarrow\begin{bmatrix}A&B\\&D-CA^{-1}B\end{bmatrix}$$

那么 $\det\left(\begin{bmatrix}A&B\\C&D\end{bmatrix}\right)=\det(A)\det(D-CA^{-1}B)$.

3. A 通常不是方阵，不能求行列式。

4. 若 A, B 是 n 阶方阵,那么 $\det(-BA) = (-1)^n \det(B)\det(A)$. 反例: $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, 都是可逆矩阵，并且 $AB = -BA = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

5. 我们直接计算 $\det\left(\begin{bmatrix} a+sc & b+sd \\ c+ta & d+tb \end{bmatrix}\right) = (1-st)(ad-bc)$. 那么很明显了，当 $st = 0$ 时行列式不变。

□

练习 4.2.11. 设 n 阶方阵 A 的对角元素全为 0，其他元素全为 1，令 $A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix}$.

1. 求向量 $\boldsymbol{u}$，使得 $\boldsymbol{a}_i\,(i = 1, 2, \ldots, n)$ 可以写成 $\boldsymbol{e}_i$ 和 $\boldsymbol{u}$ 的线性组合.

2. 根据行列式满足列多线性，求 $\det(A)$.

证明*:* 1. 观察就有 $\boldsymbol{a}_i + \boldsymbol{e}_i = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$, 那么我们令 $\boldsymbol{u} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$

2. $A = \begin{bmatrix} \boldsymbol{u}-\boldsymbol{e}_1 & \boldsymbol{u}-\boldsymbol{e}_2 & \cdots & \boldsymbol{u}-\boldsymbol{e}_n \end{bmatrix}$. 那么按列展开，$\det(A)$ 是 2^n 个新矩阵的行列式的和，但是在这 2^n 个新矩阵中，若 $\boldsymbol{u}$ 作为列向量如果出现两次，该矩阵行列式为零，因此

$$\begin{aligned}\det(A) &= \det\left(\begin{bmatrix} -\boldsymbol{e}_1 & -\boldsymbol{e}_2 & \cdots & -\boldsymbol{e}_n \end{bmatrix}\right) + \sum_{i=1}^{n} \det\left(\begin{bmatrix} \cdots & -\boldsymbol{e}_{i-1} & \boldsymbol{u} & -\boldsymbol{e}_{i+1} & \cdots \end{bmatrix}\right) \\ &= (-1)^n + n(-1)^{n-1} = (n-1)(-1)^{n-1}\end{aligned}$$

□

练习 4.2.12. 证明命题 *4.2.5.*

证明*:* 1. 这是行列式定义中的第 2 条。

2. 将行列式定义中的第 1 条中 $\boldsymbol{a}_i'$ 令为 0 可得。

3. 行列式定义中的第 1 条中, 令 $k = 1, \boldsymbol{a}_i' = \boldsymbol{a}_j$, 然后结合命题 4.2.3d 的第 1 条可得。

□

练习 4.2.13. 证明命题 *4.2.11.*

证明*:* 我们只需证明：满足命题 4.2.11 要求的函数 δ 也满足定义 4.2.1 中要求的列多线性、列反对称性和单位化三个性质。

- 由于单位阵是参数为 1 的倍乘矩阵，我们就得到了单位化 $\delta(I_n) = 1$.
- 令 $B = P_{ij}$, 我们就得到了列反对称性。
- 令 $B = E_{ii,k}(k \neq 0)$, 我们得到 $\delta\left(\begin{bmatrix} \cdots & k\boldsymbol{a}_i & \cdots \end{bmatrix}\right) = k\delta\left(\begin{bmatrix} \cdots & \boldsymbol{a}_i & \cdots \end{bmatrix}\right)$; 结合 δ 在不可逆矩阵上取值为零可知这一条对 $k = 0$ 也成立。再取 B 是倍加矩阵可知对 A 作

列倍加初等行变换，不改变 δ 的取值。结合起来就有 δ 在矩阵

$$\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_{i-1} & \boldsymbol{x} & \cdots & \boldsymbol{e}_n\end{bmatrix} = \begin{bmatrix}1 & 0 & \cdots & x_1 & 0 & \cdots & 0\\ 0 & 1 & \cdot & x_2 & 0 & \cdots & 0\\ & & \ddots & \vdots & & & \\ & & & x_i & & & \\ & & & \vdots & \ddots & & \\ 0 & \cdots & & x_n & 0 & \cdots & 1\end{bmatrix}$$

上的取值为 x_i. 现在我们来证明 $\delta\left(\begin{bmatrix}\cdots & k\boldsymbol{a}_i + k'\boldsymbol{a}_i' & \cdots\end{bmatrix}\right) = k\delta\left(\begin{bmatrix}\cdots & \boldsymbol{a}_i & \cdots\end{bmatrix}\right) + k'\delta\left(\begin{bmatrix}\cdots & \boldsymbol{a}_i' & \cdots\end{bmatrix}\right)$. 根据已有的结论，我们只需证明

$$\delta\left(\begin{bmatrix}\cdots & \boldsymbol{a}_i + \boldsymbol{a}_i' & \cdots\end{bmatrix}\right) = \delta\left(\begin{bmatrix}\cdots & \boldsymbol{a}_i & \cdots\end{bmatrix}\right) + \delta\left(\begin{bmatrix}\cdots & \boldsymbol{a}_i' & \cdots\end{bmatrix}\right) \tag{$*$}$$

设 $V = \mathrm{Span}(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{i-1}, \boldsymbol{a}_{i+1}, \cdots, \boldsymbol{a}_n)$; 若 $\dim V \le n-2$ 或者 $\boldsymbol{a}, \boldsymbol{a}'$ 都属于 V（此时 $\boldsymbol{a} + \boldsymbol{a}'$ 也属于 V），那么上式中三个矩阵都不可逆，因此左右两边为零；等号成立。下面我们不妨假设 $\dim V = n-1, \boldsymbol{a} \notin V$，那么 $\begin{bmatrix}\cdots & \boldsymbol{a}_i & \cdots\end{bmatrix}$ 是可逆矩阵，设它的逆矩阵为 P；由于可逆矩阵都是初等矩阵的乘积，因此 $\delta(P) \ne 0$. 注意到

$$\begin{aligned} P\begin{bmatrix}\cdots & \boldsymbol{a}_i' & \cdots\end{bmatrix} &= \begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_{i-1} & \boldsymbol{x} & \cdots & \boldsymbol{e}_n\end{bmatrix}\\ P\begin{bmatrix}\cdots & \boldsymbol{a}_i + \boldsymbol{a}_i' & \cdots\end{bmatrix} &= \begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_{i-1} & \boldsymbol{e}_i + \boldsymbol{x} & \cdots & \boldsymbol{e}_n\end{bmatrix}\end{aligned}$$

那么我们把 $(*)$ 式等号两边同时乘以 $\delta(P)$ 之后，等式左右两边都等于 $1 + x_i$; 因为 $\delta(P) \ne 0$, 因此 $(*)$ 式等号成立。

□

练习 4.2.14. *证明命题 4.2.11 中第 2 条冗余，即由第 1 条和第 3 条至第 5 条可以推出该条.*

证明*:* 由于可逆矩阵都是初等矩阵的乘积，第 3 条至第 5 条说 δ 在初等矩阵上取值非零，根据第 1 条，δ 在可逆矩阵上取值非零。由于 A 不可逆，对 A 作初等列变换可以将 A 的最后一列变成 0，设此时得到的矩阵为 B; 那么 $\delta(A) = 0$ 当且仅当 $\delta(B) = 0$.

因为 B 最后一列是 0, 则 $BE_{nn,2} = B$; 于是

$$\delta(B) = \delta(BE_{nn,2}) = \delta(B)\delta(E_{nn,2}) = 2\delta(B).$$

所以 $\delta(B) = 0$. □

练习 4.2.15. *用行列式证明奇数阶反对称矩阵不可逆.*

证明*:* 设 n 是奇数，且 A 是反对称矩阵，那么

$$\det(A) = \det(A^{\mathrm{T}}) = \det(-A) = (-1)^n \det(A) = -\det(A)$$

所以 $\det(A) = 0$, 即 A 不可逆。 □

练习 4.2.16. *证明任意可逆矩阵 A 都可以只用倍加变换化为 $\mathrm{diag}(1, 1, \ldots, 1, \det(A))$.*

证明: 我们对方阵的阶数做归纳。$n=1$ 时是平凡的; 以下假设命题对 n 阶方阵成立，A 是 $n+1$ 阶方阵。

由于 A 可逆，那么它的第一列有非零元素，因此我们取一个非零元素，并且把该非零元素所在行的适当的倍数加到第一行，可以使得 A 左上角元素是 1: 若 a_{i1} 是非零元，我们把第 i 行的 $\dfrac{1-a_{11}}{a_{i1}}$ 倍加到第一行即可。这个过程是一个倍加变换。我们再通过倍加变换，把矩阵变成

$$\begin{bmatrix} 1 & \boldsymbol{v}^{\mathrm{T}} \\ 0 & B \end{bmatrix}$$

由于倍加变换不改变行列式，就有 $\det(B)=\det(A)\neq 0$, 所以 B 可逆。根据归纳假设，只通过倍加变换可以把 B 化为 $\operatorname{diag}(1,1,\ldots,1,\det(A))$; 这些倍加变换可以看成是对 A 的后 n 行作倍加变换，因此我们可以通过倍加变换，把 A 化为

$$\begin{bmatrix} 1 & \boldsymbol{v}^{\mathrm{T}} \\ 0 & \operatorname{diag}(1,1,\ldots,1,\det(A)) \end{bmatrix}$$

那么显然，我们可以只通过倍加变换把 $\begin{bmatrix} 1 & \boldsymbol{v}^{\mathrm{T}} \\ 0 & \operatorname{diag}(1,1,\ldots,1,\det(A)) \end{bmatrix}$ 化为

$$\begin{bmatrix} 1 & 0 \\ 0 & \operatorname{diag}(1,1,\ldots,1,\det(A)) \end{bmatrix} = \operatorname{diag}(1,1,1,\ldots,1,\det(A)).$$

□

练习 4.2.17. 给定 $A=\begin{bmatrix} a_{ij} \end{bmatrix}_{n\times n}$，而 $B=\begin{bmatrix} a_{ij}c^{i-j} \end{bmatrix}_{n\times n}$，其中 $c\neq 0$，证明，$\det(A)=\det(B)$.

证明: 令 $C=\operatorname{diag}(c,c^2,\ldots,c^n)$, 则 $C^{-1}=\operatorname{diag}(c^{-1},c^{-2},\ldots,c^{-n})$, 且 $B=CAC^{-1}$, 所以 $\det(B)=\det(A)$. □

练习 4.2.18. 给定 $n-1$ 个互不相同的数 $a_1,a_2,\ldots,a_{n-1}$，令

$$P(x)=\begin{vmatrix} 1 & x & \cdots & x^{n-1} \\ 1 & a_1 & \cdots & a_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & a_{n-1} & \cdots & a_{n-1}^{n-1} \end{vmatrix}.$$

证明 $P(x)$ 是一个关于 x 的 $n-1$ 次多项式，并求 $P(x)$ 的 $n-1$ 个根.

证明: 这是一个范德蒙行列式，那么

$$p(x)=\prod_{i=1}^{n-1}(a_i-x)\prod_{1\leq j<i\leq n-1}(a_i-a_j)$$

显见这是关于 x 的 $n-1$ 次多项式，它的 $n-1$ 个根是 $a_1,\cdots,a_{n-1}$. □

练习 4.2.19. 设 $f_i(x)$ 是 i 次多项式，$i=0,1,2,\ldots,n-1$，其首项系数是 a_i. 又设 $b_0,b_1,\ldots,b_{n-1}$ 是 n 个数，计算如下的 n 阶行列式：

$$\begin{vmatrix} f_0(b_0) & f_0(b_1) & \cdots & f_0(b_{n-1}) \\ f_1(b_0) & f_1(b_1) & \cdots & f_1(b_{n-1}) \\ \vdots & \vdots & & \vdots \\ f_{n-1}(b_0) & f_{n-1}(b_1) & \cdots & f_{n-1}(b_{n-1}) \end{vmatrix}.$$

证明: 由题意，$a_i \neq 0, 0 \leq i \leq n-1$. 我们可以作初等行倍加变换，把第 i 行中次数小于 i 的项消去，因此

$$\begin{vmatrix} a_0 & a_0 & \cdots & a_0 \\ a_1 b_0 & a_1 b_0 & \cdots & a_1 b_0 \\ \vdots & \vdots & & \vdots \\ a_{n-1} b_0^{n-1} & a_{n-1} b_1^{n-1} & \cdots & a_{n-1} b_{n-1}^{n-1} \end{vmatrix} = \prod_{i=0}^{n-1} a_i \prod_{0 \leq j < i \leq n-1} (b_i - b_j)$$

□

练习 4.2.20.

1. 对分块对角矩阵 $X = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$，其中 A, B 是方阵，证明，$\det(X) = \det(A)\det(B)$.

2. 对分块上三角矩阵 $X = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$，其中 A, B 是方阵，证明，$\det(X) = \det(A)\det(B)$.

证明: 1. 首先我们考虑函数 $\delta : M_{n\times n}(\mathbb{R}) \to \mathbb{R}, A \mapsto \det\left(\begin{bmatrix} A & 0 \\ 0 & I_n \end{bmatrix}\right)$, 那么 δ 满足列线性性、列反对称性和单位化，因此 $\delta(A) = \det(A)$.
若 A 不可逆，那么 X 的前 n 列线性相关，所以 $\det(X) = 0 = \det(A)\det(B)$。若 A 可逆，那么我们考虑函数：$\delta' : M_{m\times m}(\mathbb{R}) \to \mathbb{R} : B \mapsto \det(A)^{-1}\det\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\right)$, 那么 δ' 满足列线性性、列反对称性和单位化，因此 $\delta'(B) = \det(B)$，所以 $\det\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\right) = \det(A)\det(B)$.

2. 首先我们考虑函数 $\delta : M_{n\times n}(\mathbb{R}) \to \mathbb{R}, A \mapsto \det\left(\begin{bmatrix} A & C \\ 0 & I_n \end{bmatrix}\right)$, 那么 δ 满足列线性性、列反对称性和单位化，因此 $\delta(A) = \det(A)$.
若 A 不可逆，那么 X 的前 n 列线性相关，所以 $\det(X) = 0 = \det(A)\det(B)$。若 A 可逆，那么我们考虑函数：$\delta' : M_{m\times m}(\mathbb{R}) \to \mathbb{R} : B \mapsto \det(A)^{-1}\det\left(\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}\right)$, 那么 δ' 满足行线性性、行反对称性和单位化，因此 $\delta'(B) = \det(B)$，所以 $\det\left(\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}\right) = \det(A)\det(B)$.

□

练习 4.2.21. 设 A 可逆，D 是方阵，证明，$\det\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right) = \det(A)\det(D - CA^{-1}B)$.

证明: A 是可逆矩阵，则

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

由上一题知 $\det\left(\begin{bmatrix} I_n & 0 \\ -CA^{-1} & I_m \end{bmatrix}\right) = 1, \det\left(\begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix}\right) = \det(A)\det(D - CA^{-1}B)$.

所以

$$\det\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right) = \det(A)\det(D - CA^{-1}B).$$

□

练习 4.2.22. 设 A, B 是 n 阶方阵，证明，$\det\left(\begin{bmatrix} A & B \\ B & A \end{bmatrix}\right) = \det(A+B)\det(A-B)$.

证明: 作点分块初等行列变换即可：

$$\det\left(\begin{bmatrix} A & B \\ B & A \end{bmatrix}\right) = \det\left(\begin{bmatrix} A+B & A+B \\ B & A \end{bmatrix}\right) = \det(A+B)\det\left(\begin{bmatrix} I_n & I_n \\ B & A \end{bmatrix}\right) = \det(A+B)\det(A-B);$$

上式最后一个等号可以用上一题的结论；也可以再做分块初等行变换。 □

练习 4.2.23. 设 A, B 分别是 $m \times n, n \times m$ 矩阵，证明，$\det(I_m + AB) = \det(I_n + BA)$. 由此推出，$I_m + AB$ 可逆当且仅当 $I_n + BA$ 可逆.

提示. 构造分块矩阵.

证明: 考虑分块矩阵

$$\begin{bmatrix} I_m & A \\ -B & I_n \end{bmatrix}$$

- 作分块倍加行变换，消去左下角得：$\begin{bmatrix} I_m & A \\ 0 & I_n + BA \end{bmatrix}$
- 作分块倍加行变换，消去右上角得：$\begin{bmatrix} I_m + AB & 0 \\ -B & I_n \end{bmatrix}$

那么

$$\det\left(\begin{bmatrix} I_m & A \\ 0 & I_n + BA \end{bmatrix}\right) = \det\left(\begin{bmatrix} I_m & A \\ -B & I_n \end{bmatrix}\right) = \det\left(\begin{bmatrix} I_m + AB & 0 \\ -B & I_n \end{bmatrix}\right)$$

所以 $\det(I_m + AB) = \det(I_n + BA)$。 □

练习 4.2.24. 计算 $\begin{vmatrix} 1+a_1^2 & a_1a_2 & \cdots & a_1a_n \\ a_2a_1 & 1+a_2^2 & \cdots & a_2a_n \\ \vdots & \vdots & & \vdots \\ a_na_1 & a_na_2 & \cdots & 1+a_n^2 \end{vmatrix}$.

证明: 令 $\boldsymbol{a}=\begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix}$, 那么所求的行列式就是 $\det(I_n+\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}})$, 由上一题,

$$\det(I_n+\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}})=\det(I_1+\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a})=1+a_1^2+\cdots+a_n^2$$

□

练习 4.2.25. 设 A 是三阶矩阵，已知 $\det(A-I_3)=\det(A-2I_3)=\det(A-3I_3)=0$.

1. 证明存在非零向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$，满足 $A\boldsymbol{v}_i=i\boldsymbol{v}_i, i=1,2,3$.
2. 设 $k_1\boldsymbol{v}_1+k_2\boldsymbol{v}_2+k_3\boldsymbol{v}_3=0$，证明 $k_1\boldsymbol{v}_1+2k_2\boldsymbol{v}_2+3k_3\boldsymbol{v}_3=0, k_1\boldsymbol{v}_1+4k_2\boldsymbol{v}_2+9k_3\boldsymbol{v}_3=0$.
3. 证明存在可逆 *Vandermonde* 矩阵 V，使得 $\begin{bmatrix}k_1\boldsymbol{v}_1 & k_2\boldsymbol{v}_2 & k_3\boldsymbol{v}_3\end{bmatrix}V=0$.
4. 证明 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 构成 $\mathbb{R}^3$ 的一组基，因此矩阵 $B=\begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3\end{bmatrix}$ 可逆.
5. 证明存在对角矩阵 D，使得 $AB=BD$，并计算 $\det(A)$.

证明: 1. 对 $i=1,2,3$, 因为 $\det(A-iI_3)=0$, $A-iI_3$ 不可逆，那么 $\mathcal{N}(A-iI_3)\neq\{0\}$, 所以存在非零向量 $\boldsymbol{v}_i\in\mathcal{N}(A-iI_3)$, 即 $A\boldsymbol{v}_i=i\boldsymbol{v}_i$.

2. 若 $k_1\boldsymbol{v}_1+k_2\boldsymbol{v}_2+k_3\boldsymbol{v}_3=0$，则 $A(k_1\boldsymbol{v}_1+k_2\boldsymbol{v}_2+k_3\boldsymbol{v}_3)=0$, 那么

$$k_1A\boldsymbol{v}_1+k_2A\boldsymbol{v}_2+k_3A\boldsymbol{v}_3=k_1\boldsymbol{v}_1+2k_2\boldsymbol{v}_2+3k_3\boldsymbol{v}_3=0.$$

再左乘一次 A, 就有 $k_1\boldsymbol{v}_1+4k_2\boldsymbol{v}_2+9k_3\boldsymbol{v}_3=0$.

3. 取 $V=\begin{bmatrix}1&1&1\\1&2&4\\1&3&9\end{bmatrix}$ 即可。

4. V 可逆，上式右乘 V^{-1}, 则 $\begin{bmatrix}k_1\boldsymbol{v}_1 & k_2\boldsymbol{v}_2 & k_3\boldsymbol{v}_3\end{bmatrix}=0$, 所以 $k_1\boldsymbol{v}_1=k_2\boldsymbol{v}_2=k_3\boldsymbol{v}_3=0$, 即 $k_1=k_2=k_3=0$. 至此，我们的证明说明 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 线性无关，所以 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 构成 $\mathbb{R}^3$ 的一组基，且矩阵 $B=\begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3\end{bmatrix}$ 可逆.

5. $AB=\begin{bmatrix}A\boldsymbol{v}_1 & A\boldsymbol{v}_2 & A\boldsymbol{v}_3\end{bmatrix}=\begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3\end{bmatrix}\operatorname{diag}(1,2,3)$, 所以取 $D=\operatorname{diag}(1,2,3)$. 因为 B 可逆，$\det(B)\neq 0$, 所以 $\det(A)=\det(D)=6$.

□

练习 4.2.26. 对函数 $f(t)=\det(I_n+tA)$ 在 $t=0$ 处求导. 设 $A=\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_n\end{bmatrix}$，则 I_n+tA 的第 i 列是 $\boldsymbol{e}_i+t\boldsymbol{a}_i$.

1. 当 $n=1,2,3$ 时，用 A 的元素表示 $f'(0)$；分析其规律，求 $f'(0)$ 的一般表达式.
2. 利用 $\det(I_n+AB)=\det(I_m+BA)$，证明 $\operatorname{trace}(AB)=\operatorname{trace}(BA)$.

证明: 1.
- $n=1$ 时，$A=[a]$, $f(t)=1+at$, $f'(0)=a$;
- $n=2$ 时，$A=\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}$,

$$f(t)=(1+a_{11}t)(1+a_{22}t)-t^2a_{12}a_{21}=1+(a_{11}+a_{22})t+(a_{11}a_{22}-a_{12}a_{21})t^2,$$

所以 $f'(0)=a_{11}+a_{22}$.

- $n=3$ 时，$A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, 以下的计算繁琐，但是不难，我们省略过程

$$\begin{aligned} f(t)= & \ 1+(a_{11}+a_{22}+a_{33})t \\ & +((a_{11}a_{22}+a_{11}a_{33}+a_{22}a_{33})-(a_{12}a_{21}+a_{13}a_{31}+a_{23}a_{32}))t^2 \\ & +\det(A)t^3. \end{aligned}$$

所以 $f'(0)=a_{11}+a_{22}+a_{33}$.

进一步地，对于 n 阶方阵 $A=[a_{ij}]$, 我们可以猜测，$f'(0)=a_{11}+\cdots+a_{nn}=\operatorname{trace}(A)$. 事实上，$f(t)$ 是个 n 次多项式，$f'(0)$ 等于该多项式一次项的系数。而 $f(t)$ 的一次项只出现在对角线元素的乘积中（这个需要了解行列式的完全展开式，除了对角线的乘积，其他项贡献的 t 的次数至少是 $n-1$），所以 $f'(0)$ 就是 $(1+a_{11}t)(1+a_{22}t)\cdots(1+a_{nn}t)$ 中一次项的系数，于是 $f'(0)=a_{11}+a_{22}+\cdots+a_{nn}$.

- 考虑 $f(t)=\det(I_n+tAB)$, 那么由练习 4.2.23,

$$f(t)=\det(I_n+(tA)B)=\det(I_m+B(tA))=\det(I_m+tBA)$$

那么就有 $f'(0)=\operatorname{trace}(AB)=\operatorname{trace}(BA)$.

□

练习 4.2.27. 设 A 是 $m\times n$ 矩阵，任取其中的 k 行和 k 列交叉点上的元素构成 k 阶方阵，它的行列式称为 A 的一个 k 阶子式. 定义 $\operatorname{rank}_{\det}(A)=\max\{k \mid A \text{ 有非零的 } k \text{ 阶子式}\}$. 证明，$\operatorname{rank}_{\det}(A)=\operatorname{rank}(A)$.

证明: 方便起见，记 $r=\operatorname{rank}(A), d=\operatorname{rank}_{\det}(A), A=\begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix}$.

$(r\le d)$. 我们取 A 的列向量的一个极大线性无关组 $\boldsymbol{a}_{i_1},\cdots,\boldsymbol{a}_{i_r}(i_1<i_2<\cdots<i_r)$，那么 $B=\begin{bmatrix} \boldsymbol{a}_{i_1} & \cdots & \boldsymbol{a}_{i_r} \end{bmatrix}$ 的秩和 A 的秩相等：$\operatorname{rank}(B)=\operatorname{rank}(A)$. 那么 B 的行向量组的极大线性无关组也是有 r 个向量，我们取这样的一个极大线性无关组，不妨设为第 $j_1<\cdots<j_r$ 个行向量。这样得到的子矩阵 C 它的行秩等于 B 的行秩，所以 $\operatorname{rank}(C)=\operatorname{rank}(B)=\operatorname{rank}(A)=r$，所以 C 是可逆矩阵，从而 $\det(C)\neq 0$。另一方面，C 是 A 的第 $j_1,\cdots,j_r$ 行第 $i_1,\cdots,i_r$ 列的交叉点上的元素构成的子矩阵；按照定义，就有 $d\ge\operatorname{rank}(C)=r$.

$(d\le r)$. 设 A 的第 $j_1,\cdots,j_d$ 行第 $i_1,\cdots,i_d$ 列的交叉点上的元素组成的矩阵 C 行列式非零，那么 C 可逆，所以 $\operatorname{rank}(C)=d$; 设 B 是 A 的设 A 的第 $j_1,\cdots,j_d$ 行组成的矩阵，那么我们有 $r=\operatorname{rank}(A)\ge\operatorname{rank}(B)\ge\operatorname{rank}(C)=d$.

□

练习 4.2.28 (Hadamard 不等式).

1. 利用 QR 分解证明，对任意 n 阶矩阵 $T=\begin{bmatrix} \boldsymbol{t}_1 & \boldsymbol{t}_2 & \cdots & \boldsymbol{t}_n \end{bmatrix}$, $|\det(T)|\le\|\boldsymbol{t}_1\|\,\|\boldsymbol{t}_2\|\ldots\|\boldsymbol{t}_n\|$.

2. 说明阅读 3.2.3中的 *Hadamard* 矩阵使得等号成立.

证明: 1. 若 T 不可逆，则 $\det(T)=0$, 所以不等式是显然的。若 T 可逆，设 $T=QR$ 是他的 QR 分解，那么 $|\det(Q)|=|\det(R)|$, 等于 R 的对角线元素的乘积的绝对值。设 $R=\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n\end{bmatrix}$, 显然 $\|\boldsymbol{v}_i\|\geq|x_i|$, 其中 x_i 是 R 的第 i 个对角元，也是 $\boldsymbol{v}_i$ 的第 i 个分量。另一方面，$\boldsymbol{t}_i=Q\boldsymbol{v}_i$, 因此 $\|\boldsymbol{t}_i\|=\|\boldsymbol{v}_i\|\geq|x_i|$. 所以

$$|\det(T)|=|x_1\cdots x_n|\leq\|\boldsymbol{v}_1\|\ldots\|\boldsymbol{v}_n\|=\|\boldsymbol{t}_1\|\|\boldsymbol{t}_2\|\ldots\|\boldsymbol{t}_n\|$$

2. 设 $A=\begin{bmatrix}\boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n\end{bmatrix}$ 是 Hadamard 矩阵, 于是 $A^{\mathrm{T}}A=nI_n$，所以 $\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{a}_i=n,\forall 1\leq i\leq n$, 那么 $\|\boldsymbol{a}_i\|=n^{\frac{1}{2}}$. 另一方面 $\det(A)^2=\det(A^{\mathrm{T}}A)=\det(nI_n)$，所以

$$|\det(A)|=n^{\frac{n}{2}}=\|\boldsymbol{a}_1\|\|\boldsymbol{a}_2\|\ldots\|\boldsymbol{a}_n\|.$$

□

练习 4.2.29. 设 n 阶对称矩阵 A 有 LDL^{T} 分解 $A=LDL^{\mathrm{T}}$，其中 $D=\operatorname{diag}(d_i)$，并记 A 的第 i 个顺序主子阵为 A_i. 证明 $d_i=\dfrac{\det(A_i)}{\det(A_{i-1})}$.

矩阵 A 的第 i 个顺序主子阵的行列式称为其第 i 个**顺序主子式**.

证明: 对 $1\leq i\leq n$, 我们把 L 分块：$L=\begin{bmatrix}L_{1i} & 0\\ L_{2i} & L_{3i}\end{bmatrix}$, 其中 L_{1i},L_{3i} 是阶数分别为 $i,n-i$、对角线为 1 的下三角矩阵，D 也做相应的分块:$D=\operatorname{diag}(D_{1i},D_{2i})$, 那么

$$A=\begin{bmatrix}L_{1i} & 0\\ L_{2i} & L_{3i}\end{bmatrix}\begin{bmatrix}D_{1i} & 0\\ 0 & D_{2i}\end{bmatrix}\begin{bmatrix}L_{1i}^{\mathrm{T}} & L_{2i}^{\mathrm{T}}\\ 0 & L_{3i}^{\mathrm{T}}\end{bmatrix}$$

直接计算，就有 $A_i=L_{1i}D_{1i}L_{1i}^{\mathrm{T}}$, 所以 $\det(A_i)=\det(D_{1i})=d_1\cdots d_i$. 命题结论就很明显了。□

练习 4.2.30 (行列式在多元微积分中的应用). 一个多元函数 $f(x_1,x_2,\ldots,x_n)$，把 x_i 之外的变量都看做常数，对 x_i 的导数称为 f 对 x_i 的偏导数，记作 $\dfrac{\partial f}{\partial x_i}$. 例如，若 $f(x,y)=x^2y$，则 $\dfrac{\partial f}{\partial x}=2xy,\dfrac{\partial f}{\partial y}=x^2$.

平面 $\mathbb{R}^2$ 有直角坐标系 (x,y),和极坐标系 (r,θ),其中 $r\geq0$ 是该点到原点的距离,$\theta\in[0,2\pi)$ 是从 x 轴正方向开始，逆时针旋转，到达该点和原点连线所需要的角度. 两种坐标之间的关系是 $x=r\cos\theta,y=r\sin\theta$.

分别计算 $J_1=\begin{bmatrix}\dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial\theta}\\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial\theta}\end{bmatrix}$ 和 $J_2=\begin{bmatrix}\dfrac{\partial r}{\partial x} & \dfrac{\partial r}{\partial y}\\ \dfrac{\partial\theta}{\partial x} & \dfrac{\partial\theta}{\partial y}\end{bmatrix}$ 的行列式，将结果都写成 r,θ 的函数. 这两个矩阵 J_1,J_2 有什么关系?

证明: $\dfrac{\partial x}{\partial r}=\cos\theta,\dfrac{\partial x}{\partial\theta}=-r\sin\theta,\dfrac{\partial y}{\partial r}=\sin\theta,\dfrac{\partial y}{\partial\theta}=r\cos\theta$, 所以

$$J_1=\begin{bmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{bmatrix}$$

$r=(x^2+y^2)^{1/2}, \tan\theta = y/x$, 那么

$$\frac{\partial r}{\partial x} = x(x^2+y^2)^{-1/2} = \cos\theta, \frac{\partial r}{\partial y} = y(x^2+y^2)^{-1/2} = \sin\theta,$$

且

$$\frac{1}{\cos^2\theta}\frac{\partial\theta}{\partial x} = -\frac{y}{x^2}, \frac{1}{\cos^2\theta}\frac{\partial\theta}{\partial y} = \frac{1}{x}, \quad \text{即} \quad \frac{\partial\theta}{\partial x} = -\frac{\sin\theta}{r}, \frac{\partial\theta}{\partial y} = \frac{\cos\theta}{r}$$

所以

$$J_2 = \begin{bmatrix} \cos\theta & \sin\theta \\ -\dfrac{\sin\theta}{r} & \dfrac{\cos\theta}{r} \end{bmatrix}$$

直接计算就有 $J_1J_2 = I_2$.

J_1, J_2 的关系也可以由下述的链式法则得到：

$$\begin{cases} \cos\theta\dfrac{\partial r}{\partial x} - r\sin\theta\dfrac{\partial\theta}{\partial x} & = 1 \\ \cos\theta\dfrac{\partial r}{\partial y} - r\sin\theta\dfrac{\partial\theta}{\partial y} & = 0 \\ \sin\theta\dfrac{\partial r}{\partial x} + r\cos\theta\dfrac{\partial\theta}{\partial x} & = 0 \\ \sin\theta\dfrac{\partial r}{\partial y} + r\cos\theta\dfrac{\partial\theta}{\partial y} & = 1 \end{cases}$$

□

练习 4.2.31. 设 $f(a,b,c,d) = \ln(ad-bc)$.

1. 接上题，求偏导数 $\dfrac{\partial f}{\partial a}, \dfrac{\partial f}{\partial b}, \dfrac{\partial f}{\partial c}, \dfrac{\partial f}{\partial d}$.
2. 证明 $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} \dfrac{\partial f}{\partial a} & \dfrac{\partial f}{\partial b} \\ \dfrac{\partial f}{\partial c} & \dfrac{\partial f}{\partial d} \end{bmatrix}^{\mathrm{T}}$.
3. 三阶矩阵时是否有类似的结论?

证明: 1. $\dfrac{f}{\partial a} = \dfrac{d}{ad-bc}, \dfrac{f}{\partial b} = -\dfrac{c}{ad-bc}, \dfrac{f}{\partial c} = -\dfrac{b}{ad-bc}, \dfrac{f}{\partial d} = \dfrac{a}{ad-bc}$.

2. 直接计算矩阵乘法

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} \dfrac{\partial f}{\partial a} & \dfrac{\partial f}{\partial b} \\ \dfrac{\partial f}{\partial c} & \dfrac{\partial f}{\partial d} \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} \dfrac{d}{ad-bc} & -\dfrac{b}{ad-bc} \\ -\dfrac{c}{ad-bc} & \dfrac{a}{ad-bc} \end{bmatrix} = I_2$$

3. 一般地，n 阶矩阵也有这样的公式，我们这里不作证明，只叙述一下。设 $f(x_{ij}) = \ln(\det(A))$, 其中 $A=[x_{ij}]$, 令 $B = \left[\dfrac{\partial f}{x_{ji}}\right] = \left[\dfrac{\partial f}{x_{ij}}\right]^{\mathrm{T}}$ 则 $AB=BA=I_n$. 这可以由行列式的完全展开式和 Laplace 展开来说明。

□

阅读 4.2.32. 在 *Vandermonde* 矩阵中

4.3 行列式的展开式

练习 4.3.1. 计算：

1. $\begin{vmatrix} 1 & 1 & 4 \\ 1 & 2 & 2 \\ 1 & 2 & 5 \end{vmatrix}$.　　*2.* $\begin{vmatrix} 1 & 1 & 10 \\ 1 & 2 & 2 \\ 1 & 2 & 5 \end{vmatrix}$.

证明*:* 1. $\begin{vmatrix} 1 & 1 & 4 \\ 1 & 2 & 2 \\ 1 & 2 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 4 \\ 0 & 1 & -2 \\ 0 & 0 & 3 \end{vmatrix} = 3$

2. $\begin{vmatrix} 1 & 1 & 10 \\ 1 & 2 & 2 \\ 1 & 2 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 10 \\ 0 & 1 & -8 \\ 0 & 0 & 3 \end{vmatrix} = 3.$

□

练习 4.3.2. 利用按行（列）展开求下列行列式；按哪一行（列）展开，使得计算最简单？

1. $\begin{vmatrix} 1 & 0 & 0 & 2 \\ 0 & 3 & 4 & 5 \\ 5 & 4 & 0 & 3 \\ 2 & 0 & 0 & 1 \end{vmatrix}$.　　*2.* $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 1 & 0 & 0 & 4 \end{vmatrix}$.

证明*:* 1. 按照第三列展开，有

$$\text{原行列式} = (-1)^{2+3}4\begin{vmatrix} 1 & 0 & 2 \\ 5 & 4 & 3 \\ 2 & 0 & 1 \end{vmatrix} = -4\cdot(-1)^{2+2}4\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = 48$$

2. 按照第二行展开，有

$$\text{原行列式} = (-1)^{1+2}\begin{vmatrix} 1 & 1 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{vmatrix} + (-1)^{2+2}2\begin{vmatrix} 1 & 1 & 1 \\ 1 & 3 & 0 \\ 1 & 0 & 4 \end{vmatrix} = -12+10 = 2$$

□

练习 4.3.3. 计算 $\begin{vmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \\ b_1 & b_2 & b_3 & b_4 & b_5 \\ c_1 & c_2 & 0 & 0 & 0 \\ d_1 & d_2 & 0 & 0 & 0 \\ e_1 & e_2 & 0 & 0 & 0 \end{vmatrix}$.

证明*:* 后三行只有两个非零列向量，因此后三行构成则子矩阵秩小于等于 2，那么整个矩阵的秩小于等于 4，即不满秩，所以行列式为 0. □

练习 4.3.4. 求 $\begin{vmatrix} 2x & x & 1 & 2 \\ 1 & x & 1 & -1 \\ 3 & 2 & x & 1 \\ 1 & 1 & 1 & x \end{vmatrix}$ 中 x^4, x^3 的系数.

证明: 行列式的完全展开式中，每一项都是在每一行每一列取唯一的元素相乘。那么

- x^4 项必须每一行都取含有 x 的项相乘，这样的项只有唯一的一个，那就是对角元相乘，因此 x^4 系数是 2.
- x^3 项必须有三行取含有 x 的项,另一行取数字项,也只有一项:分别是矩阵 $(2,1),(1,2).(3,3),(4,4)$ 位置的元素；考虑符号，那么 x^3 项的系数是 (-1)

□

练习 4.3.5. 计算 $\begin{vmatrix} \lambda & & & & a_n \\ -1 & \lambda & & & a_{n-1} \\ & \ddots & \ddots & & \vdots \\ & & -1 & \lambda & a_2 \\ & & & -1 & \lambda+a_1 \end{vmatrix}$.

证明: 我们对比较小的 n 进行计算，找规律，然后用数学归纳法证明规律。这是一个比较常见的方法。$n=1$ 时，行列式等于 $\lambda+a_1$；$n=2$ 时，行列式等于 $\lambda^2+a_1\lambda+a_2$；$n=3$ 时，行列式等于 $\lambda^3+a_1\lambda^2+a_2\lambda+a_3$；那么我们猜测一般情形行列式的为 $\lambda^n+a_1\lambda^{n-1}+a_2\lambda^{n-2}+\cdots+a_n$.

假设上述论断对 n 成立，那么对 $n+1$ 阶方阵，我们按第一行展开，有

$$\text{原行列式} = \lambda \begin{vmatrix} \lambda & & & & a_n \\ -1 & \lambda & & & a_{n-1} \\ & \ddots & \ddots & & \vdots \\ & & -1 & \lambda & a_2 \\ & & & -1 & \lambda+a_1 \end{vmatrix} + (-1)^{1+(1+n)}a_{n+2} \begin{vmatrix} -1 & \lambda & & \\ & \ddots & \ddots & \\ & & -1 & \lambda \\ & & & -1 \end{vmatrix}$$

前一项由归纳假设代入，后一项的矩阵是上三角矩阵，因此

$$\begin{aligned} \text{原行列式} &= \lambda(\lambda^n+a_1\lambda^{n-1}+a_2\lambda^{n-2}+\cdots+a_n)+(-1)^{n+2}a_{n+1}(-1)^n \\ &= \lambda^{n+1}+a_1\lambda^n+a_2\lambda^{n-1}+\cdots+a_n\lambda+a_{n+1}. \end{aligned}$$

□

练习 4.3.6. 计算 $\begin{vmatrix} \lambda & 1 & & & & \\ n & \lambda & 2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & 2 & \lambda & n \\ & & & & 1 & \lambda \end{vmatrix}$.

证明*:* 我们把原行列式的值记作 $f_n(\lambda)$. 对 $1\le i\le n-1$, 我们把第 $i+1,i+2,\cdots,n$ 行都加到第 i 行，不改变行列式的值，所以 $f_n(\lambda)$ 的等于

$$\begin{vmatrix}\lambda+n & \lambda+n & \lambda+n & \cdots & \cdots & \lambda+n\\ n & \lambda+n-1 & \lambda+n & \cdots & \cdots & \lambda+n\\ & \ddots & \ddots & \ddots & & \vdots\\ & & \ddots & \ddots & \ddots & \vdots\\ & & & 2 & \lambda+1 & \lambda+n\\ & & & & 1 & \lambda\end{vmatrix}.$$

对 $2\le j\le n$, 我们把第 j 列减去第 $j-1$ 列，不改变行列式的值，所以 $f_n(\lambda)$ 等于

$$\begin{vmatrix}\lambda+n & 0 & & & & \\ n & \lambda-1 & 1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & 2 & \lambda-1 & n-1\\ & & & & 1 & \lambda-1\end{vmatrix}.$$

按照第一行展开，就有

$$f_n(\lambda)=(\lambda+n)\begin{vmatrix}\lambda-1 & 1 & & & \\ \ddots & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & 2 & \lambda-1 & n-1\\ & & & 1 & \lambda-1\end{vmatrix}=(\lambda+n)f_{n-1}(\lambda-1),$$

我们很容易就能计算 $f_1(\lambda),f_2(\lambda)$; 那么递推就可以算出来 $f_n(\lambda)$; 剩下来的交给读者完成。 $\square$

练习 4.3.7. *回顾??中的对称、上三角和下三角 Pascal 矩阵. 在四阶的情形，这三种矩阵分别为* $S_4=\begin{bmatrix}1&1&1&1\\1&2&3&4\\1&3&6&10\\1&4&10&20\end{bmatrix}$, $U_4=\begin{bmatrix}1&1&1&1\\0&1&2&3\\0&0&1&3\\0&0&0&1\end{bmatrix}$, $L_4=\begin{bmatrix}1&0&0&0\\1&1&0&0\\1&2&1&0\\1&3&3&1\end{bmatrix}$. *另外，存在 LU 分解* $S_n=L_nU_n$.

1. *求* $\det(L_n),\det(U_n),\det(S_n)$.
2. *求* S_n *右下角元素的代数余子式.*
3. *将* S_n *右下角的元素减 1 得到矩阵* A_n*，求* $\det(A_n)$.

证明*:*

1. L_n 是对角线为 1 的上三角矩阵，所以 $\det(L_n)=1$；U_n 是对角线为 1 的上三角矩阵，所以 $\det(U_n)=1$. $\det(S_n)=\det(L_n)\det(U_n)=1$.
2. S_n 右下角元素的代数余子式是 $(-1)^{n+n}\det(S_{n-1})=1$.

3. 设 S_n' 是前 $n-1$ 行和 S_n 相同，最后一行是 $\begin{bmatrix}0 & \cdots & 0 & 1\end{bmatrix}$, 那么根据行列式的行线性性，所求行列式即 $\det(S_n)-\det(S_n')$. 对 S_n' 按照最后一行展开，就有 $\det(S_n')=\det(S_{n-1})$, 因此所求行列式为 $\det(S_n)-\det(S_{n-1})=0$。

□

练习 4.3.8. 给定 $A_n=\begin{bmatrix}2 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2\end{bmatrix}, B_n=\begin{bmatrix}1 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2\end{bmatrix}$.

1. 利用展开式得到 $\det(B_n)$ 关于 n 的递推关系，并计算 $\det(B_n)$.

2. 利用 $\det(A_n)$ 与 $\det(B_n)$ 的关系计算 $\det(A_n)$.

证明*:* 1. $B_1=1, B_2=1$; 当 $n\geq 3$ 时，我们对最后一列用 Laplace 展开：

$$\det(B_n)=(-1)\cdot(-1)^{2n-1}\begin{vmatrix}1 & -1 & & & \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ \vdots & & -1 & 2 & -1 \\ 0 & & \cdots & 0 & -1\end{vmatrix}+2(-1)^{2n}\begin{vmatrix}1 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2\end{vmatrix}$$

所以我们得到了递推式 $\det(B_n)=-\det(B_{n-2})+2\det(B_{n-1})$. 这说明 $\det(B_n)$ 构成等差数列，所以 $\det(B_n)=1$.

我们也可以把 B_n 的第一行加到第二行，然后对第一列作 Laplace 展开，这样直接得到 $\det(B_n)=\det(B_{n-1})$. 我们上面之所以用了比较复杂的计算是因为那是所谓的“三对角矩阵”求行列式的一般方法。

2. 根据 A_n 的第一行和 B_n 第一行的关系，我们利用行列式的行线性性，就有

$$\det(A_n)=\det(B_n)+\begin{vmatrix}1 & 0 \\ * & A_{n-1}\end{vmatrix}=1+\det(A_{n-1});$$

而 $\det(A_1)=2$, 所以 $\det(A_n)=n+1$.

□

练习 4.3.9 (行列式中的 Fibonacci 数列). 如果一个矩阵比上（下）三角矩阵仅仅多一排非零对角元，则称之为上（下）*Hessenberg* 矩阵. 例如，$H_4=\begin{bmatrix}2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2\end{bmatrix}$ 就是上 *Hessenberg* 矩阵. 上 *Hessenberg* 矩阵在数值分析中很有用.

1. 令 H_n 为 n 阶上 *Hessenberg* 矩阵，其对角元素都是 2，其他非零元素都是 1. 证明 $\det(H_{n+2})=\det(H_{n+1})+\det(H_n)$，即这些行列式组成了 *Fibonacci* 数列.

2. 令 S_n 是对角元素为 3，与对角元相邻的元素为 1 的 n 阶三对角矩阵（见练习 1.3.12），例如，$S_4 = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix}$. 它既是上 *Hessenberg* 矩阵，也是下 *Hessenberg* 矩阵. 求 S_n 的递归公式，并分析与 *Fibonacci* 数列的关系.

3. 设 n 阶三对角矩阵的行列式的完全展开式中，最多有 t_n 项非零，求 t_n 的递归公式.

证明: 1. 我们首先利用行列式的行线性性知

$$\det(H_{n+2}) = \begin{vmatrix} 1 & 1 & 1 & \cdots \\ 1 & 2 & 1 & \cdots \\ 0 & 1 & 2 & \cdots \\ & \cdots & & \cdots \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 & \cdots \\ 1 & 2 & 1 & \cdots \\ 0 & 1 & 2 & \cdots \\ & \cdots & & \cdots \end{vmatrix}$$

后一项按第一行展开就得到他等于 $\det(H_{n+1})$; 前一行我们做一个行倍加：将第二行减去第一行，就有：

$$\begin{vmatrix} 1 & 1 & 1 & \cdots \\ 1 & 2 & 1 & \cdots \\ 0 & 1 & 2 & \cdots \\ & \cdots & & \cdots \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 1 & 2 & \cdots \\ & \cdots & & \cdots \end{vmatrix} = \det(H_n)$$

最后一个等号是对中间的行列式先按第一列展开，然后对得到的新行列式按照第一行展开。那么就有 $\det(H_{n+2}) = \det(H_{n+1}) + \det(H_n)$. 又 $\det(H_1) = 2, \det(H_2) = 3$, 所以这些行列式组成了 Fibanacci 数列。

2. 我们在上一题中已经演示了计算三对角矩阵的方法。首先，$S_1 = 3, S_2 = 8, S_3 = 21$. 另一方面，我们把 S_n 按照第一行展开就有

$$S_{n+2} = 3S_{n+1} - \begin{bmatrix} 1 & * \\ 0 & S_n \end{bmatrix} = 3S_{n+1} - S_n.$$

我们记 Fibonacci 数列为如下归纳定义的数列：$F_0 = F_1 = 1, F_{n+2} = F_{n+1} + F_n$. 比较前几项，我们可以猜测 $S_n = F_{2n+2}$. 可以通过计算通项公式进行比较，也可以用归纳法；我们用归纳法来证明，因为这更符合我们"找规律——验证规律"这样的思维过程。$n = 1, 2, 3$ 是已知的，我们假设对任意 $k \leq n$ 都有 $S_k = F_{2k+2}$; 那么

$$S_{n+1} = 3S_n - S_{n-1} = 3F_{2n+2} - F_{2n} = 2F_{2n+2} + F_{2n+1} = F_{2n+2} + F_{2n+3} = F_{2n+4}$$

这就完成了归纳。

3. $t_1 = 1\ t_2 = 2, t_3 = 3$, 同样是按第一行展开，我们有

$$t_{n+2} = t_{n+1} + t_n.$$

□

练习 4.3.10. 求下列推广的 *Vandermonde* 行列式：

$$\begin{vmatrix} 1 & x_1 & \cdots & x_1^{n-2} & x_1^n \\ 1 & x_2 & \cdots & x_2^{n-2} & x_2^n \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-2} & x_n^n \end{vmatrix}.$$

证明: 我们把这个行列式扩展一下：

$$f(y) := \begin{vmatrix} 1 & x_1 & \cdots & x_1^{n-2} & x_1^{n-1} & x_1^n \\ 1 & x_2 & \cdots & x_2^{n-2} & x_2^{n-1} & x_2^n \\ \vdots & \vdots & & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^{n-2} & x_n^{n-1} & x_n^n \\ 1 & y & \cdots & y^{n-2} & y^{n-1} & y^n \end{vmatrix}.$$

我们对新行列最后一行作 Laplace 展开，那么原行列式是 $f(y)$ 中 y^{n-1} 项的系数的 $(-1)^{(n+1)+n}$ 倍。而 $f(y)$ 本身也是个 Vandermonde 行列式，所以

$$f(y) = \prod_{1\le j<i\le n}(x_i - x_j)\prod_{\ell=1}^{n}(y - x_\ell),$$

那么 $f(y)$ 中 y^{n-1} 项的系数就是 $-\left(\sum_{\ell=1}^{n} x_\ell\right)\prod_{1\le j<i\le n}(x_i - x_j)$, 所以原行列式等于

$$\left(\sum_{\ell=1}^{n} x_\ell\right)\prod_{1\le j<i\le n}(x_i - x_j)$$

□

练习 4.3.11. 对 n 阶方阵 A，证明，$\det(\lambda I_n - A)$ 是 λ 的首项系数为 1 的 n 次多项式.

证明: 我们可以用 Laplace 展开和数学归纳法来证明。这里我们用行列式的完全展开式来证。根据行列式的完全展开式，行列式是将矩阵每一行每一列都去一个元素相乘然后配以合适的 ±1 倍之后再相加，那么 $\det(\lambda I_n - A)$ 的完全展开式中显然每一项都是 λ 的多项式，从而 $\det(\lambda I_n - A)$ 也是 λ 的多项式。当我们观察最高次的时候，会发现这些展开式中除了只有全部取对角线元素以外的项，其 λ 的次数最多是 $n-2$，而全部取对角线元素的项是 $(\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn})$，它最高次是 λ^n. 这就完成了我们的证明。 □

附注. 这个证明过程也说明 $\det(\lambda I_n - A)$ 中 λ^{n-1} 项的系数是矩阵 A 的迹的 -1 倍。

练习 4.3.12.

1. 设 A 是正交矩阵，且 $\det(A) < 0$，证明 $I_n + A$ 不可逆. 由此可得，存在非零向量 $\boldsymbol{x}$，使得 $A\boldsymbol{x} = -\boldsymbol{x}$.
   提示. 考虑 $A^{\mathrm{T}}(A+I)$ 与 $A+I$ 的行列式之间的关系.
2. 设 A 是奇数阶正交矩阵，且 $\det(A) > 0$，证明 $I_n - A$ 不可逆. 由此可得，存在非零向量 $\boldsymbol{x}$，使得 $A\boldsymbol{x} = \boldsymbol{x}$. 偶数阶的情形，结论是否成立?

证明: A 是正交阵，则 $A^TA = I_n$, 所以 $\det(A)^2 = 1$，即 $\det(A) = \pm 1$.

1. $\det(A) < 0$ 则 $\det(A) = -1$. $\det(I_n + A) = \det(AA^{\mathrm{T}} + A) = \det(A)\det(A^T + I_n) = -\det(I_n + A)$, 所以 $\det(I_n + A) = 0$. 于是 $I_n + A$ 不可逆，因此存在非零向量 $\boldsymbol{x}$ 使得 $(I_n + A)\boldsymbol{x} = 0$, 即 $A\boldsymbol{x} = -\boldsymbol{x}$.
2. $\det(A) > 0$ 则 $\det(A) = 1$. $\det(I_n - A) = \det(AA^{\mathrm{T}} - A) = \det(A)\det(A^{\mathrm{T}} - I_n) = (-1)^n \det(I_n - A)$, 所以 $I_n - A$ 不可逆，那么存在非零向量 $\boldsymbol{x}$ 使得 $(I_n - A)\boldsymbol{x} = 0$, 即 $A\boldsymbol{x} = \boldsymbol{x}$.

   n 偶数时结论不成立，例如 $A = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$

□

练习 4.3.13. 设 n 阶方阵 $A = [a_{ij}]$ 对角占优（见**??**），且对角元素全是正数，证明 $\det(A) > 0$.

证明: 我们对阶数 n 作归纳。$n = 1$ 时是平凡的；假设命题对任意复合题意的 n 阶方阵成立，现在考虑对角元都是正数的对角占优矩阵 $A = \begin{bmatrix} a_{ij} \end{bmatrix}_{1 \le i,j \le n+1}$. 由于 $a_{11} > 0$, 我们作初等列倍加变换，把第一行除 a_{11} 外圈化为 0, 设此时得到的矩阵形如: $\begin{bmatrix} a_{11} & 0 \\ \boldsymbol{a} & B \end{bmatrix}$, 并设 $B = \begin{bmatrix} b_{ij} \end{bmatrix}_{2 \le i,j \le n+1}$ 那么 $b_{ij} = a_{ij} - \frac{a_{1j}}{a_{11}} a_{i1}$. 下面我们来证明 B 是对角元都是正数的对角占优矩阵。首先

$$b_{ii} = a_{ii} - \frac{a_{1i}}{a_{11}} a_{i1} = \frac{a_{ii}a_{11} - a_{1i}a_{i1}}{a_{11}} > 0.$$

其次

$$\sum_{\substack{2 \le j \le n+1 \\ j \ne i}} |b_{ij}| = \sum_{\substack{2 \le j \le n+1 \\ j \ne i}} \left| \left(a_{ij} - \frac{a_{1j}}{a_{11}} a_{i1} \right) \right| \le \sum_{\substack{2 \le j \le n+1 \\ j \ne i}} \left(|a_{ij}| - \left| \frac{a_{1j}}{a_{11}} a_{i1} \right| \right)$$

我们需要注意到

$$\begin{aligned} \sum_{\substack{2 \le j \le n+1 \\ j \ne i}} |a_{ij}| & < a_{ii} - |a_{i1}| \\ \sum_{\substack{2 \le j \le n+1 \\ j \ne i}} \left| \frac{a_{1j}}{a_{11}} a_{i1} \right| & = \left| \frac{a_{i1}}{a_{11}} \right| \sum_{\substack{2 \le j \le n+1 \\ j \ne i}} |a_{1j}| < \left| \frac{a_{i1}}{a_{11}} \right| (a_{11} - |a_{1i}|) \end{aligned}$$

所以

$$\sum_{\substack{2 \le j \le n+1 \\ j \ne i}} |b_{ij}| < b_{11}$$

根据归纳假设，$\det(B) > 0$, 那么 $\det(A) = a_{11}\det(B) > 0$. □

证明: 我们还有更精巧的证明。令 $f(t) = \det(tI_n + A)$, 我们把 $f(t)$ 视作 $[0, +\infty)$ 上的函数，因为 f 是关于 t 的次数为 n 首项系数为 1 的多项式，所以 $f(t)$ 是连续的，且当 $t \to +\infty$ 时，$f(t) \to +\infty$.

另一方面，对任意的 $t \ge 0$, $tI_n + A$ 都是对角线元素都是正数的对角占优矩阵，所以 $\det(tI_n + A) \ne 0, \forall t \ge 0$. 那么 $f(t)$ 在 $[0, +\infty)$ 上没有零点。结合 $t \to +\infty$ 时，$f(t) \to +\infty$ 可知 $f(t) > 0, \forall t \ge 0$; 特别地，$\det(A) = f(0) > 0$. □

练习 4.3.14. 证明若 A 不可逆，则其补矩阵的秩是 0 或 1.

证明: A 不可逆，则 $\operatorname{rank}(A) \le n-1$，且 $AA^* = \det(A)I_n = 0$，所以 $\mathcal{R}(A^*) \subseteq \mathcal{N}(A)$，即 $\operatorname{rank}(A^*) \le n - \operatorname{rank}(A)$

- 若 $\operatorname{rank}(A) = n-1$，则 $\operatorname{rank}(A^*) \le 1$. 另一方面，根据练习 4.2.27，$A$ 存在 $n-1$ 阶子方阵 B，使得 $\det(B) \ne 0$；而 $\det(B)$ 是 A^* 的某个元素，所以 $A^* \ne 0$；故 $\operatorname{rank}(A^*) \ge 1$.
- 若 $\operatorname{rank}(A) \le n-2$，那么根据练习 4.2.27，$A$ 的任意 $n-1$ 阶子方阵都不可逆，所以 A^* 是 0 矩阵，故 $\operatorname{rank}(A^*) = 0$.

$\square$

练习 4.3.15. 给定所有元素全为整数的可逆矩阵 A，证明，A^{-1} 的所有元素全为整数，当且仅当 $|\det(A)| = 1$.

证明: 因为 A 和 A^{-1} 的元素都是整数，所以 $\det(A), \det(A^{-1})$ 都是整数。又因为 $\det(A)\det(A^{-1}) = 1$，所以 $\det(A) = \pm 1$. $\square$

练习 4.3.16. 利用 *Cramer* 法则把未知数 x_1, x_2 表示成 t 的函数：

$$\begin{cases} e^t x_1 + \ \ e^{-2t} x_2 = 3\sin t, \\ e^t x_1 - 2e^{-2t} x_2 = t\cos t. \end{cases}$$

证明: 根据 Cramer 法则，

$$x_1 = \frac{\det\left(\begin{bmatrix} 3\sin t & e^{-2t} \\ t\cos t & 2e^{-2t} \end{bmatrix}\right)}{\det\left(\begin{bmatrix} e^t & e^{-2t} \\ e^t & 2e^{-2t} \end{bmatrix}\right)} = (6\sin t - t\cos t)\, e^{-t}$$

$$x_2 = \frac{\det\left(\begin{bmatrix} e^t & 3\sin t \\ e^t & t\cos t \end{bmatrix}\right)}{\det\left(\begin{bmatrix} e^t & e^{-2t} \\ e^t & 2e^{-2t} \end{bmatrix}\right)} = (t\cos t - 3\sin t)\, e^{2t}$$

$\square$

阅读 4.3.17 (Cramer 法则的另一证明).

练习 4.3.18. 设 n 阶方阵 A, B 满足 $AB = BA$，证明：

1. $\left|\det\left(\begin{bmatrix} A & B \\ -B & A \end{bmatrix}\right)\right| = \left|\det(A^2 + B^2)\right|$.

   提示. 将 $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} a^2+b^2 & 0 \\ 0 & a^2+b^2 \end{bmatrix}$ 推广到分块矩阵.

2. $\det\left(\begin{bmatrix} A & B \\ -B & A \end{bmatrix}\right) = \det(A^2 + B^2)$.

   提示. 利用完全展开式比较特定项如 $(a_{11}a_{22}\dots a_{nn})^2$ 的系数.

附注. 此题也有一些不利用完全展开的方法，读者可以自行思考. 法一：先考虑 A 可逆的情形，再对不可逆的 A 构造可逆序列 $A_k \to A$. 法二：利用复数.

证明: 1. 注意到对任意 n 阶方阵 A, B, 我们有

$$\begin{bmatrix} A & B \\ -B & A \end{bmatrix}\begin{bmatrix} A & -B \\ B & A \end{bmatrix} = \begin{bmatrix} A^2+B^2 & -AB+BA \\ -BA+AB & A^2+B^2 \end{bmatrix}$$

$AB=BA$ 则有 $\begin{bmatrix} A & B \\ -B & A \end{bmatrix}\begin{bmatrix} A & -B \\ B & A \end{bmatrix} = \begin{bmatrix} A^2+B^2 & 0 \\ 0 & A^2+B^2 \end{bmatrix}$, 取行列式就有

$$\det(A^2+B^2)^2 = \det\left(\begin{bmatrix} A & B \\ -B & A \end{bmatrix}\right)\det\left(\begin{bmatrix} A & -B \\ B & A \end{bmatrix}\right)$$

再注意到 $\begin{bmatrix} A & -B \\ B & A \end{bmatrix} = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}\begin{bmatrix} A & B \\ -B & A \end{bmatrix}\begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}$, 所以

$$\det\left(\begin{bmatrix} A & B \\ -B & A \end{bmatrix}\right) = \det\left(\begin{bmatrix} A & -B \\ B & A \end{bmatrix}\right),$$

那么就有 $\left|\det\left(\begin{bmatrix} A & B \\ -B & A \end{bmatrix}\right)\right| = \left|\det(A^2+B^2)\right|$。

2. 若 A 可逆，那么根据练习 4.2.21知

$$\det\left(\begin{bmatrix} A & B \\ -B & A \end{bmatrix}\right) = \det(A)\det(A-(-B)A^{-1}B) = \det(A^2+B^2).$$

若 A 不可逆，考虑 $A_t = A+tI_n$。因为 $f(t)=\det(A_t)$ 是关于 t 的首项系数为 1 的 n 次多项式，当 $t\to+\infty, f(t)\to+\infty$, 因此存在 $t_0>0$, 使得对任意 $t>t_0$ 时，A_t 可逆。令

$$F(t) = \det\left(\begin{bmatrix} A_t & B \\ -B & A_t \end{bmatrix}\right), G(t) = \det(A_t^2+B^2),$$

那么 $F(t), G(t)$ 都是关于 t 的首项系数为 1 的 $2n$ 次多项式。根据 A 可逆的情形，我们知道当 $t>t_0$ 时，$F(t)=G(t)$, 那么 $F(t)$ 和 $G(t)$ 作为多项式也是相等的。特别地，我们有 $F(0)=G(0)$.

我们也可以考虑 $f(t)$ 在 $t=0$ 附近的行为。因为 $f(t)$ 是多项式，所以 $f(t)=0$ 只有有限多个解，那设非零的绝对值最小的解为 t_0, 那么 $f(t)$ 在开区间 $U=(-|t_0|,|t_0|)$ 上只有唯一解 $t=0$. 同样是根据 A 可逆的情形知当 $t\in U, t\neq 0$ 时，$F(t)=G(t)$. 根据 $F(t), G(t)$ 的连续性知 $F(0)=G(0)$.

□

练习 4.3.19. 利用完全展开式求行列式.

1. 设 A 为 4 阶矩阵，所有元素均为 1，在其行列式的完全展开式中，多少项为 1? 多少项为 -1? 由此计算 A 的行列式.
2. 把 A 的 (i,j) 元乘以 $\frac{i}{j}$ 得到 B，在 A 和 B 行列式的完全展开式中，每一项如何变化? 行列式如何变化?

3. 设 $A=\begin{bmatrix} a & 0 & b & 0 \\ 0 & c & 0 & d \\ e & 0 & f & 0 \\ 0 & g & 0 & h \end{bmatrix}$，在行列式的完全展开式中，有多少项非零？这个完全展开式是否有因式分解？（跟分块对角矩阵的情形进行类比）

证明: 1. 行列式的完全展开式每一项都是在 A 的每一行每一列取唯一一个元素相乘，然后再乘以相应的逆序数。那么按题意，取 1 的项数即逆序数是偶数的项数；取 -1 的项数即逆序数是奇数的项数，因此都是 $\frac{4!}{2}=12$. 所以行列式为零。

2. A 和 B 的行列式的完全展开式每一项都相等，因为和 A 相比，B 相同位置的项相乘分子分母均多出来一个 $1\times 2\times 3\times 4$，所以总乘积不变。

3. 一共有 4 项非零：我们先看第一三行：两种选取：af, be, 第二四行两种选取：ch, dg, 所以一共 4 项。要将这个完全展开式因式分解，我们先交换第二三行，再交换第二三列，得到的矩阵 $\begin{bmatrix} a & b & 0 & 0 \\ e & f & 0 & 0 \\ 0 & 0 & c & d \\ 0 & 0 & g & h \end{bmatrix}$, 所以完全展开式可以因式分解为 $(af-be)(dh-dg)$.

□

练习 4.3.20. 由 $\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{vmatrix}=0$，证明，$1,2,\ldots,n$ 的所有排列中，奇、偶排列各占一半.

证明: 因为该矩阵所有元素都是 1, 所以行列式完全展开式中每一项都是 ±1，取 1 当且仅当相应的项对应于一个偶排列，取 -1 当且仅当相应的项对应于一个奇排列。

行列式为 0 说明取值 $1,-1$ 的项一样多；从而奇、偶排列各占一半。 □

练习 4.3.21. 在空间 $\mathbb{R}^3$ 中，证明由向量 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 围出的平行四边形面积是 $\sqrt{\det(A^{\mathrm{T}}A)}$，其中 $A=\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2\end{bmatrix}$.

提示. 进行 QR 分解 $A=QR$，那么 A 和 R 对应的平行四边形有何关系？$\sqrt{\det(A^{\mathrm{T}}A)}$ 和 $\sqrt{\det(R^{\mathrm{T}}R)}$ 有何关系？

证明: 根据练习 3.1.18, 由 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 围出平行四边形面积 $S=\left(\|\boldsymbol{a}_1\|^2\|\boldsymbol{a}_2\|^2-(\boldsymbol{a}_1\cdot\boldsymbol{a}_2)^2\right)^{1/2}$.

另一方面

$$\det(A^{\mathrm{T}}A)=\det\left(\begin{bmatrix}\boldsymbol{a}_1^{\mathrm{T}} \\ \boldsymbol{a}_2^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2\end{bmatrix}\right)=\det\left(\begin{bmatrix}\|\boldsymbol{a}_1\|^2 & \boldsymbol{a}_1\cdot\boldsymbol{a}_2 \\ \boldsymbol{a}_2\cdot\boldsymbol{a}_2 & \|\boldsymbol{a}_2\|^2\end{bmatrix}\right)=S^2$$

□

练习 4.3.22 (行列式游戏). 甲和乙两个人构造一个 n 阶方阵，轮流填写矩阵中的元素，直到填满为止. 如果矩阵的行列式非零，则甲胜；否则，乙胜.

1. 如果乙先开始，且 $n=2$，则谁有必胜策略？

2. 如果乙先开始，且 $n=3$，则谁有必胜策略？

3. 如果甲先开始，且 $n=3$，则谁有必胜策略？

4. 如果甲先开始，且 n 为偶数，则谁有必胜策略？

证明*:* 1. 乙有必胜策略，策略如下：乙左上角填 0, 那么右上和左下位置，至少有一个还是乙填，再填 0,z 这样行列式为 0.

2.

□

第五章　特征值和特征向量

5.1　引　　子

练习 5.1.1. 证明??。

5.2　基 本 概 念

练习 5.2.1. 求下列复矩阵的全部特征值和特征向量：

1. $\begin{bmatrix} 3 & 4 \\ 5 & 2 \end{bmatrix}$.　*2.* $\begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}$.　*3.* $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.　*4.* $\begin{bmatrix} 0 & 2 & 1 \\ -2 & 0 & 3 \\ -1 & -3 & 0 \end{bmatrix}$.

5. $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$.

证明*:* 我们就第一小题写一下过程，以下只写答案。其中 $\boldsymbol{i} = \sqrt{-1}$.

1. $\det(\lambda I_2 - A) = \det\left(\begin{bmatrix} \lambda - 3 & -4 \\ -5 & \lambda - 2 \end{bmatrix}\right) = (\lambda - 3)(\lambda - 2) - 20 = \lambda^2 - 5\lambda - 14$. 所以它有两个特征值：$\lambda_1 = 7, \lambda_2 = -2$.
    - $\mathcal{N}(\lambda_1 I_2 - A) = \operatorname{Span}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$, 即属于 $\lambda_1 = 7$ 的特征向量为 $k\begin{bmatrix} 1 \\ 1 \end{bmatrix}, k \neq 0$.
    - $\mathcal{N}(\lambda_1 I_2 - A) = \operatorname{Span}\left(\begin{bmatrix} -4 \\ 5 \end{bmatrix}\right)$, 即属于 $\lambda_2 = -2$ 的特征向量为 $k\begin{bmatrix} -4 \\ 5 \end{bmatrix}, k \neq 0$.
2. 特征值是 $\pm a\boldsymbol{i}$.
    - 属于 $a\boldsymbol{i}$ 的特征向量为 $k\begin{bmatrix} 1 \\ \boldsymbol{i} \end{bmatrix}, k \neq 0$.
    - 属于 $-a\boldsymbol{i}$ 的特征向量为 $k\begin{bmatrix} 1 \\ -\boldsymbol{i} \end{bmatrix}, k \neq 0$.
3. 特征值是 ± 1.
    - 属于 1 的特征向量为 $\begin{bmatrix} a \\ b \\ a \end{bmatrix}$, a, b 不全为零。

- 属于 -1 的特征向量为 $k\begin{bmatrix}1\\0\\-1\end{bmatrix}, k\neq 0.$

4. 特征值是 $0, \pm\sqrt{14}\boldsymbol{i}$.

   - 属于 0 的特征向量为 $k\begin{bmatrix}3/2\\-1/2\\1\end{bmatrix}, k\neq 0.$
   - 属于 $\sqrt{14}\boldsymbol{i}$ 的特征向量为 $k\begin{bmatrix}-(6+\sqrt{14}\boldsymbol{i})/10\\-(-2+3\sqrt{14}\boldsymbol{i})/10\\1\end{bmatrix}, k\neq 0.$
   - 属于 $-\sqrt{14}\boldsymbol{i}$ 的特征向量为 $k\begin{bmatrix}-(6-\sqrt{14}\boldsymbol{i})/10\\-(-2-3\sqrt{14}\boldsymbol{i})/10\\1\end{bmatrix}, k\neq 0.$

5. 特征值是 ± 2

   - 属于 2 的特征向量为 $\begin{bmatrix}x_1+x_2+x_3\\x_1\\x_2\\x_3\end{bmatrix}, x_1, x_2, x_3$ 不全为零。
   - 属于 -2 的特征向量为 $k\begin{bmatrix}1\\-1\\-1\\-1\end{bmatrix}, k\neq 0.$

□

练习 5.2.2. 构造符合要求的矩阵 A：

1. A 的特征多项式为 $\lambda^2-9\lambda+20$，构造三个不同的 A.

2. $A=\begin{bmatrix}0&1\\ *&*\end{bmatrix}$，且 A 的特征值为 $4, 7$.

3. $A=\begin{bmatrix}0&1&0\\0&0&1\\ *&*&*\end{bmatrix}$，且 A 的特征值为 $1, 2, 3$.

证明*:* 1. 按题意，A 阶数为 2, 迹为 9, 行列式为 20. 选择很多，我们随便给几个

$$\begin{bmatrix}4&0\\0&5\end{bmatrix},\quad \begin{bmatrix}2&2\\-3&7\end{bmatrix},\quad \begin{bmatrix}3&-2\\1&6\end{bmatrix}$$

2. 特征值是 $4, 7$，那么迹是 11，行列式是 28，所以 $A=\begin{bmatrix}0&1\\-28&11\end{bmatrix}$. 我们也可以假设 $A=\begin{bmatrix}0&1\\x&y\end{bmatrix}$，那么 A 的特征多项式是 $\det\left(\begin{bmatrix}\lambda&-1\\-x&\lambda-y\end{bmatrix}\right)=\lambda^2-y\lambda-x$; 由题意，$4, 7$ 是 $\lambda^2-y\lambda-x=0$ 的两个根。即可求出 $x=-28, y=11$.

3. 设 $A=\begin{bmatrix}0&1&0\\0&0&1\\x&y&z\end{bmatrix}$，$A$ 的特征多项式是

$$\det\left(\begin{bmatrix}\lambda&-1&0\\0&\lambda&-1\\-x&-y&\lambda-z\end{bmatrix}\right)=\lambda^3-z\lambda^2-y\lambda-x.$$

根据题意有 $\lambda^3-z\lambda^2-y\lambda-x=(\lambda-1)(\lambda-2)(\lambda-3)$, 所以

$$x=6,\quad ,y=-11,\quad ,z=6.$$

□

练习 5.2.3. 设 $A=\begin{bmatrix}2&-2&0\\-2&x&-2\\-2&-2&0\end{bmatrix}$,$B=\begin{bmatrix}2&&\\&2&\\&&y\end{bmatrix}$，已知 A,B 特征多项式相同，求 x,y.

证明: (法一) $\det(\lambda I_3-A)=\det\left(\begin{bmatrix}\lambda-2&2&0\\2&\lambda-x&2\\2&2&\lambda\end{bmatrix}\right)=\lambda^3-(2+x)\lambda^2+(2x-8)\lambda+16$,

$\det(\lambda I_3-B)=\det\left(\begin{bmatrix}\lambda-2&&\\&\lambda-2&\\&&\lambda-y\end{bmatrix}\right)=\lambda^3-(4+y)\lambda^2+(4+4y)\lambda-4y$, 比较系数就有 $y=-4,x=-2$.

(法二) A,B 的特征多项式相同，就他们有相同的迹和行列式。A 的行列式通过多第三列进行 Laplace 展开，B 是对角矩阵，即得 $4y=-16$，所以 $y=-4$；再考虑迹，就有 $2+x=2+2+y$，故 $x=-2$.

□

练习 5.2.4. 设 $A=\begin{bmatrix}2&1&1\\1&2&1\\1&1&a\end{bmatrix}$ 可逆，且 $\boldsymbol{x}=\begin{bmatrix}1\\b\\1\end{bmatrix}$ 是 A 的特征向量，求 a,b 的值.

证明: 按定义，存在 λ 使得 $A\boldsymbol{x}=\lambda\boldsymbol{x}$, 即

$$\begin{cases}2+b+1&=\lambda\\1+2b+1&=\lambda b\\1+b+a&=\lambda\end{cases}$$

解方程组得

$$\begin{cases}a&=2\\b&=-2\\\lambda&=1\end{cases}\quad\text{或}\quad\begin{cases}a&=2\\b&=1\\\lambda&=4\end{cases}$$

□

练习 5.2.5. 设 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

1. 利用一元二次方程求根公式，写出 A 的两个特征值 λ_1, λ_2 的表达式.

2. 构造一个非对角的 A，满足 $\lambda_1 = \lambda_2$.

3. 设 λ 为 A 的特征值，证明 $A \begin{bmatrix} b \\ \lambda - a \end{bmatrix} = \lambda \begin{bmatrix} b \\ \lambda - a \end{bmatrix}, A \begin{bmatrix} \lambda - d \\ c \end{bmatrix} = \lambda \begin{bmatrix} \lambda - d \\ c \end{bmatrix}$.

附注. 如果这两个向量不是零向量，那么它们就是特征向量.

4. 若上述两个向量中有且仅有一个是零向量，求 A 的特征值和特征向量.

5. 若上述两个向量都是零向量，求 A 的特征值和特征向量.

证明: 1. A 的特征多项式是 $\det(\lambda I_2 - A) = \lambda^2 - (a+d)\lambda + (ad - bc)$, 所以

$$\lambda_{1,2} = \frac{(a+d) \pm \sqrt{(a+d)^2 - 4(ad-bc)}}{2} = \frac{(a+d) \pm \sqrt{(a-d)^2 + 4bc}}{2}$$

2. $\lambda_1 = \lambda_2 \Leftrightarrow (a-d)^2 + 4bc = 0$. 例如我们可以令 $a = d \neq 0, b \neq 0, c = 0$.

3.
$$A \begin{bmatrix} b \\ \lambda - a \end{bmatrix} = \begin{bmatrix} ab + b(\lambda - a) \\ bc + d(\lambda - a) \end{bmatrix} = \begin{bmatrix} b\lambda \\ \lambda^2 - a\lambda \end{bmatrix} = \lambda \begin{bmatrix} b \\ \lambda - a \end{bmatrix}$$
$$A \begin{bmatrix} \lambda - d \\ c \end{bmatrix} = \begin{bmatrix} a(\lambda - d) + bc \\ c(\lambda - d) + dc \end{bmatrix} = \begin{bmatrix} \lambda^2 - d\lambda \\ c\lambda \end{bmatrix} = \lambda \begin{bmatrix} \lambda - d \\ c \end{bmatrix}$$

4. 若 $\begin{bmatrix} b \\ \lambda - a \end{bmatrix} = 0$, 则 $\begin{bmatrix} \lambda - d \\ c \end{bmatrix} \neq 0$, 那么 $b = 0, \lambda = a$.

   - $a \neq d$, 那么 a, d 是 A 的两个特征值,属于特征值 a 的特征向量是 $k \begin{bmatrix} a - d \\ c \end{bmatrix}, k \neq 0$; 属于 d 的特征向量是 $\begin{bmatrix} 0 \\ k \end{bmatrix}, k \neq 0$.
   - $a = d$, 那么 $c \neq 0$, 此时 A 有唯一的特征值 a, 属于特征值 a 的特征向量是 $\begin{bmatrix} 0 \\ k \end{bmatrix}$

5. 此时 $\lambda - a = b = \lambda - d = c = 0$, 因此 A 是数量矩阵 aI_2, $\mathbb{R}^2$ 的任何非零向量都是 A 的特征向量。

□

练习 5.2.6. 给定向量 $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$，计算 $A = \boldsymbol{a}\boldsymbol{b}^{\mathrm{T}}$ 的所有特征对.

证明: $A = 0$ 当且仅当 $\boldsymbol{a}$ 或者 $\boldsymbol{b}$ 等于零。此时特征对为 $(0, \boldsymbol{x}), \forall \boldsymbol{x} \neq 0$. 以下假设 $\boldsymbol{a} \neq 0, \boldsymbol{b} \neq 0$.

设 $(\lambda, \boldsymbol{x})$ 是 A 的一个特征对，即 $A\boldsymbol{x} = \lambda \boldsymbol{x}$, 那么

$$\boldsymbol{a}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} = \lambda \boldsymbol{x}.$$

- 若 $\lambda = 0$, $(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x})\boldsymbol{a} = 0$, 那么这等价于 $\boldsymbol{x} \in \boldsymbol{b}^{\perp}$.

- 若 $\lambda \neq 0$, 那么 $\boldsymbol{x} = (\lambda^{-1}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x})\boldsymbol{a}$, 于是

$$\boldsymbol{a}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} = (\lambda^{-1}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x})\boldsymbol{a}(\boldsymbol{b}^{\mathrm{T}})\boldsymbol{a} \neq 0,$$

  此时需要 $\boldsymbol{b}\boldsymbol{a} \neq 0$, 并且进一步地有 $\lambda = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}$.

综合起来就有：

- 若 $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a} = 0$, A 的特征对为 $\Big\{(0, \boldsymbol{x}) : \boldsymbol{x} \in \boldsymbol{b}^{\perp} - \{0\}\Big\}$;
- 若 $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a} \neq 0$, A 的特征对为 $\Big\{(0, \boldsymbol{x}) : \boldsymbol{x} \in \boldsymbol{b}^{\perp} - \{0\}\Big\} \cup \Big\{(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}, k\boldsymbol{a}) : k \neq 0\Big\}$.

□

练习 5.2.7. 已知 A 的特征值为 $1, 2, 3$，求下列矩阵的特征值：

1. $2A, A+I_3, A^2, \overline{A}, A^{\mathrm{T}}, A^{-1}$（$A$ 为何可逆？）. *2.* $\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}, \begin{bmatrix} A & A \\ 0 & A \end{bmatrix}$.

证明*:*

1.
   - $2A$ 的特征值是 $2, 4, 6$;
   - $A + I_3$ 的特征值是 $2, 3, 4$;
   - A^2 的特征值是 $1, 4, 9$;
   - $\overline{A}$ 的特征值是 $1, 2, 3$;
   - A^{T} 的特征值是 $1, 2, 3$;
   - 因为 A 的行列式是它所有特征值（算重数）的乘积，所以 A 可逆，A^{-1} 的特征值是 $1, 2^{-1}, 3^{-1}$
2. 这两个矩阵的特征多项式都是 A 的特征多项式的平方，因此特征值也是 $1, 2, 3$; 但是每个特征值的重数都是 A 的重数的两倍。

□

练习 5.2.8. 证明，如果 $(\lambda, \boldsymbol{x})$ 是 A 的特征对，则 $(f(\lambda), \boldsymbol{x})$ 是 $f(A)$ 的特征对，其中 $f(x)$ 是任意多项式.

证明*:* 因为 $A\boldsymbol{x} = \lambda\boldsymbol{x}$, 那么 $A^n\boldsymbol{x} = \lambda A^{n-1}\boldsymbol{x} = \cdots = \lambda^n\boldsymbol{x}$. 若 $f(x) = \sum_{i=0}^{n} a_i x^i$, 那么

$$f(A)\boldsymbol{x} = \sum_{i=0}^{n} a_i A^i \boldsymbol{x} = \sum_{i=0}^{n} a_i \lambda^i \boldsymbol{x} = f(\lambda)\boldsymbol{x},$$

所以 $(f(\lambda), \boldsymbol{x})$ 是 $f(A)$ 的特征对。 □

练习 5.2.9. 设 A 是可逆矩阵，证明，A 的特征值都不为 0；若 λ_0 是 A 的一个特征值，则 $\dfrac{1}{\lambda_0}$ 是 A^{-1} 的一个特征值.

证明*:* $f(\lambda) = \det(\lambda I_n - A)$, A 的所有特征值是 $f(\lambda)$ 的根, 那么

$$\lambda_1 \cdots \lambda_n = \det(A).$$

A 可逆当且仅当 $\det(A) \neq 0$, 所以 A 的所有特征值都非零。

另一方面，若 $\lambda_0 \neq 0$，那么

$$\det(\lambda_0 I_n - A) = \det(A)\det(\lambda_0 A^{-1} - I_n) = (-1)^n \lambda_0^n \det(\lambda_0^{-1} I_n - A^{-1}),$$

所以 λ_0 是 A 的特征值当且仅当 λ_0^{-1} 是 A^{-1} 的特征值。 □

练习 5.2.10. 设 $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3 \in \mathbb{R}^3$ 为一组标准正交基，分别求 $\boldsymbol{q}_1\boldsymbol{q}_1^{\mathrm{T}}, \boldsymbol{q}_1\boldsymbol{q}_1^{\mathrm{T}}+\boldsymbol{q}_2\boldsymbol{q}_2^{\mathrm{T}}, \boldsymbol{q}_1\boldsymbol{q}_1^{\mathrm{T}}+\boldsymbol{q}_2\boldsymbol{q}_2^{\mathrm{T}}+\boldsymbol{q}_3\boldsymbol{q}_3^{\mathrm{T}}$ 所有特征值和特征向量.

证明:

- 记 $A = \boldsymbol{q}_1\boldsymbol{q}_1^{\mathrm{T}}$, 那么

$$A\boldsymbol{q}_1 = \boldsymbol{q}_1, \quad A\boldsymbol{q}_2 = A\boldsymbol{q}_3 = 0$$

  所以 A 的特征对为 $\{(1, k\boldsymbol{q}_1) : k \neq 0\}$ 和 $\{(0, k_2\boldsymbol{q}_2 + k_3\boldsymbol{q}_3) : k_2, k_3\text{不全为零}\}$

- 记 $A = \boldsymbol{q}_1\boldsymbol{q}_1^{\mathrm{T}} + \boldsymbol{q}_2\boldsymbol{q}_2^{\mathrm{T}}$, 那么

$$A\boldsymbol{q}_1 = \boldsymbol{q}_1, \quad A\boldsymbol{q}_2 = \boldsymbol{q}_2, \quad A\boldsymbol{q}_3 = 0$$

  所以 A 的特征对为 $\{(1, k_1\boldsymbol{q}_1 + k_2\boldsymbol{q}_2) : k_1, k_2\text{不全为零}\}$ 和 $\{(0, k_3\boldsymbol{q}_3) : k_3 \neq 0\}$

- 事实上这个矩阵是单位阵，所以 1 是它唯一的特征值，所有非零向量都是属于特征值 1 的特征向量。

□

练习 5.2.11. 设 n 阶实方阵 A 满足 $A^{\mathrm{T}}\boldsymbol{v} = \lambda\boldsymbol{v}$，其中 $\boldsymbol{v} \in \mathbb{R}^n, \boldsymbol{v} \neq 0$.

1. 设 $A\boldsymbol{w} = \mu\boldsymbol{w}$，且 $\boldsymbol{w} \in \mathbb{R}^n, \lambda \neq \mu$，证明 $\boldsymbol{v}, \boldsymbol{w}$ 正交。
2. 证明，实对称矩阵的属于不同特征值的实特征向量正交。

证明:

1. 考虑 $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w}$, 那么

$$\lambda\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w} = (\lambda\boldsymbol{v})^{\mathrm{T}}\boldsymbol{w} = (A^{\mathrm{T}}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{w} = \boldsymbol{v}^{\mathrm{T}}A\boldsymbol{w} = \mu\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w}$$

   因为 $\lambda \neq \mu$，所以 $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w} = 0$。

2. 设 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_r$ 分别是实对称矩阵 A 的属于不同的特征值 $\lambda_1, \cdots, \lambda_r$ 的是特征向量，那么由第一题，$\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{v}_j = 0$, 故 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_r$ 两两互相正交。

□

练习 5.2.12. 证明：

1. 若存在正整数 k，使得 $A^k = 0$，则 A 的特征值只能是 0.
2. 若 $A^2 = I_n$，则 A 的特征值只能是 1 或 -1.
3. 若 $A^2 = A$，则 A 的特征值只能是 1 或 0.

证明: 设 $(\lambda, \boldsymbol{x})$ 是 A 的一个特征对，即 $A\boldsymbol{x} = \lambda\boldsymbol{x}$, 在练习 5.2.8那么我们已经证明过，对任意 k, $A^k\boldsymbol{x} = \lambda^k\boldsymbol{x}$.

1. 根据题意 $A^k\boldsymbol{x} = 0$, 故 $\lambda^k\boldsymbol{x} = 0$, 所以 $\lambda = 0$.
2. 根据题意 $A^2\boldsymbol{x} = \boldsymbol{x}$, 故 $\lambda^2 = 1$, 所以 $\lambda = \pm 1$.

3. 根据题意 $A^2\boldsymbol{x} = A\boldsymbol{x}$, 故 $\lambda^2 = \lambda$, 所以 $\lambda = 0$ 或 1.

□

练习 5.2.13. 对方阵 A，若多项式 $f(x)$ 满足 $f(A) = 0$，则称 $f(x)$ 是 A 的**化零多项式**.

给定 A 的特征值 λ，证明，若 $f(x)$ 是 A 的化零多项式，则 $f(\lambda) = 0$.

证明: 这是练习 5.2.8的直接推论：若 $(\lambda, \boldsymbol{x})$ 是 A 的一个特征对，那么

$$f(\lambda)\boldsymbol{x} = f(A)\boldsymbol{x} = 0$$

所以 $f(\lambda) = 0$ □

练习 5.2.14. 设方阵 A, B 可交换，λ_0 是 A 的一个特征值，V_{λ_0} 是 A 的特征值为 λ_0 的特征子空间。证明，对任意 $\boldsymbol{x} \in V_{\lambda_0}$，都有 $B\boldsymbol{x} \in V_{\lambda_0}$.

当 A, B 不可交换时，结论是否成立?

证明: 对任意 $\boldsymbol{x} \in V_{\lambda_0} = \mathcal{N}(\lambda_0 I_n - A)$，那么

$$A(B\boldsymbol{x}) = BA\boldsymbol{x} = B(\lambda_0\boldsymbol{x}) = \lambda_0 B\boldsymbol{x},$$

即 $(\lambda_0 I_n - A)B\boldsymbol{x} = 0$, 所以 $B\boldsymbol{x} \in V_{\lambda_0}$.

A, B 不交换时，结论不成立，反例如：

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \lambda_0 = 0, \boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

□

练习 5.2.15. 证明，A 和 $T^{-1}AT$ 具有相同的特征多项式。

证明:

$$\det(\lambda I_n - T^{-1}AT) = \det(T^{-1}(\lambda I_n - A)T) = \det(T^{-1})\det(\lambda I_n - A)\det(T) = \det(\lambda I_n - A)$$

这就说明 A 和 $T^{-1}AT$ 具有相同的特征多项式。 □

练习 5.2.16. 设 λ_1, λ_2 是 A 的两个不同特征值，$\boldsymbol{x}_1, \boldsymbol{x}_2$ 是分别属于 λ_1, λ_2 的特征向量. 证明，$\boldsymbol{x}_1 + \boldsymbol{x}_2$ 不是 A 的特征向量.

证明: 首先我们说明 $\boldsymbol{x}_1, \boldsymbol{x}_2$ 线性无关：若 $a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 = 0$, 那么

$$a_1\lambda_1\boldsymbol{x}_1 + a_2\lambda_2\boldsymbol{a}_2 = A(a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2) = 0$$

于是 $a_1(\lambda_2 - \lambda_1)\boldsymbol{x}_1 = 0$, 所以 $a_1 = 0$; 进一步地，$a_2 = 0$.

下面我们用反证法来证明 $\boldsymbol{x}_1 + \boldsymbol{x}_2$ 不是 A 的特征向量。若不然，存在 λ 使得 $A(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \lambda(\boldsymbol{x}_1 + \boldsymbol{x}_2)$, 于是

$$\lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2 = \lambda(\boldsymbol{x}_1 + \boldsymbol{x}_2),$$

由线性无关性有 $\lambda_1 = \lambda = \lambda_2$, 矛盾。 □

练习 5.2.17. 证明或举出反例：

1. 如果 A, B 具有相同的特征值、代数重数和特征向量，则 $A = B$.
2. 如果 A, B 有相同的特征值和代数重数，则 $A - B$ 所有特征值之和为零.
3. $A + B$ 的特征值之和等于 A 的特征值之和与 B 的特征值之和的和.
4. $A + B$ 的特征值之积等于 A 的特征值之积与 B 的特征值之积的积.
5. AB 的特征值之积等于 A 的特征值之积与 B 的特征值之积的积.
6. AB 和 BA 具有相同的特征值和代数重数.
7. 如果 A 的特征值全为零，则 A 是零矩阵.
8. 将 A 的第 i 行加到第 j 行上，再将第 i 列从第 j 列中减去，得到的矩阵 B 和 A 有相同的特征值. 若正确，则对应的特征向量有何联系？
9. 将 A 的第 i 行加到第 j 行上，再将第 j 列从第 i 列中减去，得到的矩阵 B 和 A 有相同的特征值. 若正确，则对应的特征向量有何联系？
10. 将 A 的第 i, j 行交换，再将第 i, j 列交换，得到的矩阵 B 和 A 有相同的特征值. 若正确，则对应的特征向量有何联系？
11. 对角矩阵的特征向量一定是标准坐标向量.
12. 正交矩阵的特征值都是绝对值等于 1 的复数.
13. 所有 n 阶置换矩阵都有一个共同的特征向量.

证明:

1. 错误，例如 $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$. 在有相同的特征值、代数重数的基础上，需要相同的特征对。
2. 正确，因为特征值之和（计算重数）等于矩阵的迹，因此 $A - B$ 的特征值之和等于 $\operatorname{trace}(A - B) = \operatorname{trace}(A) - \operatorname{trace}(B) = 0$.
3. 正确，证明方法同上，考虑迹即可。
4. 错误；我们考虑行列式，特征值之积（计算重数）等于矩阵的行列式，但是一般来说 $\det(A + B) \neq \det(A)\det(B)$. 反例如 $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
5. 正确，我们考虑行列式即可。
6. 若 A, B 都是方阵，那结论是正确的，若不是方阵，则结论是错误的。若 $A \in M_{m\times n}(\mathbb{R}), B \in M_{n\times m}(\mathbb{R})$, 则 $\lambda^n \det(\lambda I_m - AB) = \lambda^m \det(\lambda I_n - BA)$.
7. 错误。例如 $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.
8. 错误。例如 $A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}, i = 1, j = 2$, 则 $B = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$, $\det(\lambda I_2 - A) = \lambda^2 - 3\lambda + 2$, $\det(\lambda I - 2 - B) = \lambda^2 - 4\lambda + 2$.
9. 正确，因为 $B = E_{ij}(1)AE_{ij}(1)^{-1}$。$\boldsymbol{v}$ 是 A 的特征向量当且仅当 $E_{ij}(1)\boldsymbol{v}$ 是 B 的特征向量。
10. 正确，因为 $B = P_{ij}AP_{ij}$。$\boldsymbol{v}$ 是 A 的特征向量当且仅当 $P_{ij}\boldsymbol{v}$ 是 B 的特征向量。

11. 错误，例如: $\begin{bmatrix}1\\1\end{bmatrix}$ 是 I_2 的的特征向量。

12. 正确。设 A 是正交矩阵，$(\lambda, \boldsymbol{v})$ 是特征对，那么

$$\overline{\lambda}\lambda\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} = \overline{\lambda\boldsymbol{v}}^{\mathrm{T}}(\lambda\boldsymbol{v}) = \overline{\boldsymbol{v}}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{v} = \overline{\boldsymbol{v}}^{\mathrm{T}}\boldsymbol{v},$$

所以 $\overline{\lambda}\lambda = 1$, 即 λ 绝对值等于 1。

13. 正确，$\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}$ 是所以置换矩阵共同的特征向量。

□

练习 5.2.18. 设 A, B 分别是 $m \times n, n \times m$ 矩阵，证明

$$\lambda^n \det(\lambda I_m - AB) = \lambda^m \det(\lambda I_n - BA).$$

特别地，当 $m = n$ 时，$\det(\lambda I_n - AB) = \det(\lambda I_n - BA)$.

证明: 考虑分块矩阵 $C = \begin{bmatrix}0 & A\\B & 0\end{bmatrix}$, C 的特征多项式是矩阵 $\begin{bmatrix}\lambda I_m & A\\B & \lambda I_n\end{bmatrix}$ 的行列式。我们分别作分块初等行倍加变换把 C 转化为上三角矩阵和下三角矩阵，则得到了：

$$\begin{bmatrix}\lambda I_m & A\\0 & \lambda I_n - \frac{1}{\lambda}BA\end{bmatrix} \quad 和 \quad \begin{bmatrix}\lambda I_m - \frac{1}{\lambda}AB & 0\\B & \lambda I_n\end{bmatrix}$$

这两个矩阵的行列式相等，所以

$$\lambda^m \det\left(\lambda I_n - \frac{1}{\lambda}BA\right) = \lambda^n \det\left(\lambda I_m - \frac{1}{\lambda}AB\right)$$

整理一下就有

$$\lambda^{2m} \det(\lambda^2 I_n - BA) = \lambda^{2n} \det(\lambda^2 I_m - AB).$$

用 λ 来替换 λ^2 即可得

$$\lambda^n \det(\lambda I_m - AB) = \lambda^m \det(\lambda I_n - BA).$$

□

练习 5.2.19. 如果复矩阵 A, B 可交换，证明 A, B 至少有一个公共的特征向量.

证明: 设 λ_0 是 A 的一个特征值,考虑 V_{λ_0}, 它的维数设为 $k \geq 1$. 我们取 V_{λ_0} 的一组基 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$. 由练习 5.2.14，$B\boldsymbol{v}_i \in V_{\lambda_0}$，存在 c_ij 使得 $B\boldsymbol{v}_i = \sum_{j=1}^{k} c_{ij}\boldsymbol{v}_j$, 写成矩阵形式就有

$$B\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k\end{bmatrix} = \begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k\end{bmatrix}\begin{bmatrix}c_{ji}\end{bmatrix}$$

令矩阵 $C = \begin{bmatrix}c_{ji}\end{bmatrix}$, 取 C 的一个特征对 $(\mu, \boldsymbol{y})$, 其中 $0 \neq \boldsymbol{y} \in \mathbb{R}^k$, 那么

$$B\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k\end{bmatrix}\boldsymbol{y} = \begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k\end{bmatrix}\begin{bmatrix}c_{ji}\end{bmatrix}\boldsymbol{y} = \begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k\end{bmatrix}\mu\boldsymbol{y}$$

即：

$$B(y_1\boldsymbol{v}_1+\cdots+y_k\boldsymbol{v}_k)=\mu(y_1\boldsymbol{v}_1+\cdots+y_k\boldsymbol{v}_k).$$

根据特征对的要求和 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k$ 的线性无关性知 $\boldsymbol{w}=y_1\boldsymbol{v}_1+\cdots+y_k\boldsymbol{v}_k\neq 0$, 这就是说 $(\mu,\boldsymbol{w})$ 是 B 的一个特征对。另一方面，$\boldsymbol{w}\in V_{\lambda_0}$; 所以 $\boldsymbol{w}$ 是 A,B 共同的特征向量。 □

练习 5.2.20. 设 A 是四阶数独矩阵（见练习 *2.4.10*），证明其绝对值最大的特征值为 10，且属于该特征值的特征向量的所有分量都相等.

证明: 设 $(\lambda,\boldsymbol{x})$ 是 A 的一个特征对，其中 $0\neq\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}$. 设 x_i 使得 $|x_i|=\max(|x_1|,|x_2|,|x_3|,|x_4|)>0$. 因为 $A\boldsymbol{x}=\lambda\boldsymbol{x}$，我们考虑第 i 分量，就有

$$a_{i1}x_1+a_{i2}x_2+a_{i3}x_3+a_{i4}x_2=\lambda x_i$$

我们取绝对值就有

$$|\lambda|\,|x_i|\leq\sum_{j=1}^{s}|a_{ij}|\,|x_i|\leq\sum_{j=1}^{4}|a_{ij}|\,|x_i|$$

所以 $|\lambda|\leq\sum_{j=1}^{4}|a_{ij}|=10$. 因为 $a_{ij}\in\{1,2,3,4\}$, 那么取等号仅当 x_i 正负号相同且绝对值相同，即 x_i 均相等。

反之，若 x_i 都相等，那么由于 A 是数独矩阵，每一行元素相加等于 10, 所以就有 $A\boldsymbol{x}=10\boldsymbol{x}$. □

练习 5.2.21. 设方阵 A 的每个元素都是整数，证明：$\dfrac{1}{2}$ 一定不是 A 的特征值.

证明: 因为方阵 A 的每个元素都是整数，那么根据行列式的完全展开式，A 的特征多项式是首项系数为 1 的，系数都是整数的 n 次多项式。我们证明下面这个一般的结论：首项系数为 1 的整系数多项式的有理数根都是整数。

设 $x^n+a_1x^{n-1}+\cdots+a_n=0(a_i\in\mathbb{Z},1\leq i\leq n)$, 有理数 $\frac{r}{s}$ 是它的根，其中 $r,s\in\mathbb{Z}$ 且 r,s 互素。要证明 $\frac{r}{s}$ 是整数，即要证明 $s=\pm 1$.

$$\frac{r^n}{s^n}+a_1\frac{r^{n-1}}{s^{n-1}}+\cdots+a_n=0,$$

那么我们有 $r^n=-s(a_1r^{n-1}+a_2r^{n-2}s+\cdots+a_ns^{n-1})$, 等式右边是 s 的倍数，左边和 s 互素，那就只能用 $s=\pm 1$. 证毕。 □

练习 5.2.22. 回顾**??**中的矩阵 $A=\begin{bmatrix}0.6&0.1\\0.4&0.9\end{bmatrix}$，以及稳定状态对应的矩阵 $A^{\infty}=\begin{bmatrix}0.2&0.2\\0.8&0.8\end{bmatrix}$.

1. 如果 A^n 和 A^{∞} 中对应的元素相差不超过 0.01，那么 n 至少是多少？
2. 交换 A 的两行，特征值是否不变？

证明: 1. $\det(\lambda I_2 - A) = \lambda^2 - 1.5\lambda + 0.5 = (\lambda-1)(\lambda-0.5)$, 所以 A 的两个特征值是 1 和 0.5. 我们再计算

$$\mathcal{N}(I_2 - A) = \operatorname{Span}\left(\begin{bmatrix}1\\4\end{bmatrix}\right), \mathcal{N}(0.5I_2 - A) = \operatorname{Span}\left(\begin{bmatrix}1\\-1\end{bmatrix}\right)$$

那么我们有 $A\begin{bmatrix}1 & 1\\4 & -1\end{bmatrix} = \begin{bmatrix}1 & 1\\4 & -1\end{bmatrix}\begin{bmatrix}1 & 0\\0 & \frac{1}{2}\end{bmatrix}$。令 $P = \begin{bmatrix}1 & 1\\4 & -1\end{bmatrix}$，则 $A = P\begin{bmatrix}1 & 0\\0 & \frac{1}{2}\end{bmatrix}P^{-1}$, 从而

$$A^n = P\begin{bmatrix}1 & 0\\0 & \frac{1}{2^n}\end{bmatrix}P^{-1} = \begin{bmatrix}\frac{1}{5}+\frac{4}{5\cdot 2^n} & \frac{1}{5}-\frac{1}{5\cdot 2^n}\\ \frac{4}{5}-\frac{4}{5\cdot 2^n} & \frac{4}{5}+\frac{1}{5\cdot 2^n}\end{bmatrix}$$

为了使误差不超过 0.01, 则 $\frac{4}{5\cdot 2^n} \leq 0.01$, 则 $n \geq 7$.

2. 交换两行特征值要变化。我们直接计算可以知道交换两行后，特征值变为 1 和 -0.5.

□

练习 5.2.23. 给定 m 阶方阵 A_1，n 阶上三角矩阵 A_2 和 $m\times n$ 矩阵 B. 证明如果 A_1 和 A_2 没有相同的特征值，关于 $m\times n$ 矩阵 X 的矩阵方程 $A_1X - XA_2 = B$ 有唯一解.

矩阵方程 $A_1X - XA_2 = B$ 称为 *Sylvester* **方程**，在控制论中有不少应用。

证明: 我们先证矩阵方程 $A_1X - XA_2 = B$ 有唯一解，而这等价于 $A_1X = XA_2$ 只有 X 是零矩阵这一个解。我们记 $A_2 = \begin{bmatrix}a_{ij}\end{bmatrix}$，因为 A_2 上三角，所以当 $i > j$ 时，$a_{ij} = 0$，且 A_2 的特征值是它的对角元 $a_{11}, \cdots, a_{nn}$，根据题意，它们都不是 A_1 的特征值。设 $X = \begin{bmatrix}\boldsymbol{x}_1 & \cdots & \boldsymbol{a}_n\end{bmatrix}$ 是 $A_1X = XA_2$ 的解，那么就有

$$A\boldsymbol{x}_1 = a_{11}\boldsymbol{x}_1, \quad A\boldsymbol{x}_2 = a_{12}\boldsymbol{x}_1 + a_{22}\boldsymbol{x}_2, \quad \cdots, \quad A\boldsymbol{x}_n = a_{1n}\boldsymbol{x}_1 + \cdots + a_{nn}\boldsymbol{x}_n.$$

那么 $(A - a_{11}I_m)\boldsymbol{x}_1 = 0$，$a_{11}$ 不是 A 的特征值，所以 $A - a_{11}I_m$ 可逆，所以 $\boldsymbol{x}_1 = 0$. 于是 $A\boldsymbol{x}_2 = a_{12}\boldsymbol{x}_1 + a_{22}\boldsymbol{x}_2 = a_{22}\boldsymbol{x}_2$，$a_{22}$ 不是 A 的特征值，所以 $A - a_{22}I_m$ 可逆，所以 $\boldsymbol{x}_2 = 0$. 以此类推就证得 $\boldsymbol{x}_1 = \boldsymbol{x}_2 = \cdots = \boldsymbol{x}_n = 0$, 所以方程 $A_1X = XA_2$ 只有零矩阵这一个解。

我们把 $M_{m\times n}(\mathbb{R})$ 通过如下的方式等同于 $\mathbb{R}^{mn}$：

$$\varphi : M_{m\times n}(\mathbb{R}) \to \mathbb{R}^{mn}, X = \begin{bmatrix}\boldsymbol{x}_1 & \cdots & \boldsymbol{x}_n\end{bmatrix} \mapsto \begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\\ \vdots\\ \boldsymbol{x}_n\end{bmatrix}$$

这是一个双射。给定 A_1, A_2, 我们定义 $f : \mathbb{R}^{mn} \to \mathbb{R}^{mn}$ 是如下的复合映射：

$$\mathbb{R}^{mn} \xrightarrow{\varphi^{-1}} M_{m\times n}(\mathbb{R}) \xrightarrow{X\mapsto A_1X - XA_2} M_{m\times n}(\mathbb{R}) \xrightarrow{\varphi} \mathbb{R}^{m\times n}$$

我们可以验证 f 是线性映射（请读者自行验证）。方程 $A_1X = XA_2$ 只有零矩阵这一个解等价于 $M_{m\times n}(\mathbb{R}) \xrightarrow{X\mapsto A_1X - XA_2} M_{m\times n}(\mathbb{R})$ 是单射，从而复合映射 f 也是单射；那么 f 也是满射，从而映射 $M_{m\times n}(\mathbb{R}) \xrightarrow{X\mapsto A_1X - XA_2} M_{m\times n}(\mathbb{R})$ 是满射，意即：对任意 $B \in M_{m\times n}(\mathbb{R})$, 都存在 $X \in M_{m\times n}(\mathbb{R})$ 使得 $A_1X - XA_2 = B$.

□

5.3 对角化和谱分解

练习 5.3.1. 设三阶方阵 A 的特征值及对应特征向量分别是 1,1,3 和 $\begin{bmatrix}2\\1\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\1\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}$，求 A.

证明: 按照题意，我们有

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

那么 $A = \begin{bmatrix}2 & -1 & 0\\1 & 0 & 0\\0 & 1 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 3\end{bmatrix}\begin{bmatrix}2 & -1 & 0\\1 & 0 & 0\\0 & 1 & 1\end{bmatrix}^{-1} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\2 & -4 & 3\end{bmatrix}$ □

练习 5.3.2. 判断下列方阵是否可对角化：

1. $\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix}$.
2. $\begin{bmatrix}0 & 1 & 1\\0 & 0 & 1\\0 & 0 & 0\end{bmatrix}$.

证明: 这样的问题分两步走：计算特征多项式和特征值，计算每个特征值的代数重数和几何重数。

1. $\det(\lambda I_3 - A) = \lambda(\lambda^2 - 15\lambda - 18)$, 故 A 有 3 个不同的特征值，那么 A 可以相似对角化。
2. A 只有唯一的特征值 0, 其代数重数数 3，几何重数是 1，所以不可以相似对角化。

□

练习 5.3.3. 计算 A^n：

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

证明:

1. $A = \begin{bmatrix}0\\0\\1\end{bmatrix}\begin{bmatrix}4 & 0 & 1\end{bmatrix}$, 那么 $A^n = \left(\begin{bmatrix}4 & 0 & 1\end{bmatrix}\begin{bmatrix}0\\0\\1\end{bmatrix}\right)^{n-1} A = A.$
2. $\det(\lambda I_3 - A) = (\lambda - 2)^2(\lambda - 6)$, 特征值是 $2, 6$; $\mathcal{N}(2I_3 - A)$ 的一组基是 $\begin{bmatrix}1\\-1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}$, $\mathcal{N}(6I_3-$

A) 的一组基是 $\begin{bmatrix}1\\-2\\3\end{bmatrix}$, 那么 $A=\begin{bmatrix}1&1&1\\-1&0&-2\\0&1&3\end{bmatrix}\begin{bmatrix}2&&\\&2&\\&&6\end{bmatrix}\begin{bmatrix}1&1&1\\-1&0&-2\\0&1&3\end{bmatrix}^{-1}$, 所以

$$\begin{aligned}A^n&=\begin{bmatrix}1&1&1\\-1&0&-2\\0&1&3\end{bmatrix}\begin{bmatrix}2^n&&\\&2^n&\\&&6^n\end{bmatrix}\begin{bmatrix}1&1&1\\-1&0&-2\\0&1&3\end{bmatrix}^{-1}\\&=\frac{1}{4}\begin{bmatrix}5\cdot 2^n-6^n&2^n-6^n&6^n-2^n\\2\cdot 6^n-2\cdot 2^n&2\cdot 6^n+2\cdot 2^n&2\cdot 2^n-2\cdot 6^n\\3\cdot 2^n-3\cdot 6^n&3\cdot 2^n-3\cdot 6^n&2^n+3\cdot 6^n\end{bmatrix}\end{aligned}$$

□

练习 5.3.4. 设 $A=\begin{bmatrix}3&2&-2\\-k&-1&k\\4&2&-3\end{bmatrix}$. 当 k 取何值时，A 可对角化？当 A 可对角化时，写出其谱分解.

证明: A 的特征多项式是 $\det(\lambda I_3-A)=(\lambda+1)^2(\lambda-1)$, 它的特征值是 $1,-1$, 代数重数分别是 $1,2$. $\mathcal{N}(I_3-A)$ 的一组基是 $\begin{bmatrix}1\\0\\1\end{bmatrix}$, A 可以相似对角化，当且仅当 -1 的几何重数是 2, 即 $\operatorname{rank}(A+I_3)=1$. 而 $A+I_3=\begin{bmatrix}4&2&-2\\-k&0&k\\4&2&-2\end{bmatrix}$, 那么 $\operatorname{rank}(A)=1$ 当且仅当 $k=0$. 此时 $\mathcal{N}(A+I_3)$ 的一组基是 $\begin{bmatrix}1\\-2\\0\end{bmatrix}$, $\begin{bmatrix}1\\0\\2\end{bmatrix}$. 那么 A 的谱分解是

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

□

练习 5.3.5. 设 A 是 4 阶方阵，其对角元都是 4，非对角元都是 -1. 令 H 为 4 阶 *Hadamard* 矩阵，即 $H=\begin{bmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\\1&-1&-1&1\end{bmatrix}$（见阅读 *3.2.3*）. 计算 AH，由此得出 A 的谱分解，并求 $\det(A)$, A^{-1}.

证明: $A = \begin{bmatrix} 4 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & 4 \end{bmatrix}$, 我们直接计算矩阵乘法

$$AH = \begin{bmatrix} 1 & 5 & 5 & 5 \\ 1 & -5 & 5 & -5 \\ 1 & 5 & -5 & -5 \\ 1 & -5 & -5 & 5 \end{bmatrix} = H \begin{bmatrix} 1 & & & \\ & 5 & & \\ & & 5 & \\ & & & 5 \end{bmatrix}$$

所以 A 的谱分解为 $A = H\mathrm{diag}(1,5,5,5)H^{-1}, \det(A) = 125, A^{-1} = H\mathrm{diag}(1,5^{-1},5^{-1},5^{-1})H^{-1}$.

□

练习 5.3.6. 利用矩阵，求下列数列的通项公式和极限.

1. （*Lucas* 数）$L_1 = 1, L_2 = 3, L_{n+2} = L_{n+1} + L_n \, (n = 1, 2, \dots)$.
2. $a_0 = 0, a_1 = 1, a_{n+2} = \dfrac{1}{2}a_{n+1} + \dfrac{1}{2}a_n \, (n = 0, 1, 2, \dots)$.
3. $a_0 = a_1 = 0, a_2 = 1, a_{n+3} = \dfrac{1}{2}a_{n+2} + \dfrac{1}{4}a_{n+1} + \dfrac{1}{4}a_n \, (n = 0, 1, 2, \dots)$.
4. $a_0 = 1, a_{n+1} = a_n + 2 \times 3^n \, (n = 0, 1, 2, \dots)$. 考虑向量 $\begin{bmatrix} a_n \\ 3^n \end{bmatrix}$.
5. $a_0 = 0, b_0 = 1, a_{n+1} = a_n + b_n, b_{n+1} = 3a_n + 3b_n \ (n = 0, 1, 2, \dots)$. 考虑向量 $\begin{bmatrix} a_n \\ b_n \end{bmatrix}$.

证明: 对于递推式 $x_{n+k} = a_1 x_{n+k-1} + \cdots + a_{k-1}x_{n+1} + a_k x_n$, 我们考虑向量 $\boldsymbol{x}_n = \begin{bmatrix} x_n & \cdots & x_{n+k-1} \end{bmatrix}^{\mathrm{T}}$, 根据递推式就有

$$\boldsymbol{x}_{n+1} = A\boldsymbol{x}_n, \text{其中} A = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ a_k & a_{k-1} & \cdots & a_1 \end{bmatrix}$$

要计算 x_n，我们只需要计算 A^n, 那么就需要对 A 进行谱分解。A 的特征多项式恰好是 $\lambda^k - (a_1\lambda_{k-1} + \cdots + a_{k-1}\lambda + a_k)$。在 A 可以相似对角化的情况下（例如方程 $\lambda^k = (a_1\lambda_{k-1} + \cdots + a_{k-1}\lambda + a_k$ 有 k 个互不相同的根），走完整个流程我们可以验证：存在 $b_1, \cdots, b_k$ 使得 $x_n = b_1\lambda^n + \cdots + b_k\lambda^k$. 那么，我们可以代数初始值，求出 $b_1, \cdots, b_k$ 即可。

1. $\lambda^2 = \lambda + 1$ 有两个根 $\frac{1 \pm \sqrt{5}}{2}$, 那么 $L_n = b_1\left(\frac{1+\sqrt{5}}{2}\right)^n + b_2\left(\frac{1-\sqrt{5}}{2}\right)^n$; 考虑 $n = 1, 2$，可以求出 $b_1 = 1, b_2 = 1$. 所以
$$L_n = \left(\frac{1+\sqrt{5}}{2}\right)^n + \left(\frac{1-\sqrt{5}}{2}\right)^n$$
2. $\lambda^2 = \frac{1}{2}\lambda + \frac{1}{2}$ 有两个根 $1, -\frac{1}{2}$, 那么 $a_n = b_1 + b_2(-\frac{1}{2})^n$; 代入 $n = 0, 1$ 得 $b_1 = \frac{2}{3}, b_2 = -\frac{2}{3}$，所以
$$a_n = \frac{2}{3} - \frac{2}{3}(-2)^{-n}$$

3. $\lambda^3=\frac{1}{2}\lambda^2+\frac{1}{4}\lambda+\frac{1}{4}$ 有一个实根 1，两个虚根 $\frac{-1\pm\sqrt{-3}}{4}$，所以 $a_n=b_1+b_2(\frac{-1+\sqrt{-3}}{4})^n+b_3(\frac{-1-\sqrt{-3}}{4})^n$, 代入 $n=0,1,2$ 解得 $b_1=\frac{4}{7},b_2=\frac{-10-2\sqrt{-3}}{7\sqrt{-3}},b_3=\frac{10-2\sqrt{-3}}{7\sqrt{-3}}$, 所以

$$a_n=\frac{4}{7}+\frac{-10-2\sqrt{-3}}{7\sqrt{-3}}\left(\frac{-1+\sqrt{-3}}{2}\right)^n+\frac{10-2\sqrt{-3}}{7\sqrt{-3}}\left(\frac{-1-\sqrt{-3}}{2}\right)^n$$

4. 令 $\boldsymbol{a}_n=\begin{bmatrix}a_n\\3^n\end{bmatrix}$, 则

$$\boldsymbol{a}_{n+1}=\begin{bmatrix}a_n+2\times3^n\\3^{n+1}\end{bmatrix}=\begin{bmatrix}1&2\\0&3\end{bmatrix}\begin{bmatrix}a_n\\3^n\end{bmatrix}$$

记 $A=\begin{bmatrix}1&2\\0&3\end{bmatrix}$，则 $A=\begin{bmatrix}1&1\\0&1\end{bmatrix}\begin{bmatrix}1&0\\0&3\end{bmatrix}\begin{bmatrix}1&1\\0&1\end{bmatrix}^{-1}$,

$$A^n=\begin{bmatrix}1&1\\0&1\end{bmatrix}\begin{bmatrix}1&0\\0&3^n\end{bmatrix}\begin{bmatrix}1&1\\0&1\end{bmatrix}^{-1}=\begin{bmatrix}1&3^n-1\\0&3^n\end{bmatrix}$$

那么 $\boldsymbol{a}_{n+1}=A^n\boldsymbol{a}_1=\begin{bmatrix}3^{n+1}\\3^{n+1}\end{bmatrix}$, 所以 $a_n=3^n$.

5. 令 $\boldsymbol{a}_n=\begin{bmatrix}a_n\\b_n\end{bmatrix}$, 那么

$$\boldsymbol{a}_{n+1}=\begin{bmatrix}1&1\\3&3\end{bmatrix}\boldsymbol{a}_n$$

令 $A=\begin{bmatrix}1&1\\3&3\end{bmatrix}=\begin{bmatrix}1\\3\end{bmatrix}\begin{bmatrix}1&1\end{bmatrix}$, 那么 $n\geq2$ 时

$$\boldsymbol{a}_n=A^n\boldsymbol{a}_0=\left(\begin{bmatrix}1&1\end{bmatrix}\begin{bmatrix}1\\3\end{bmatrix}\right)^{n-1}A\boldsymbol{a}_0=4^{n-1}\begin{bmatrix}1\\3\end{bmatrix}$$

所以 $n\geq2$ 时，$a_n=4^{n-1},b_n=3\cdot4^{n-1}$.

□

练习 5.3.7. 利用谱分解说明如下事实：

1. $\begin{bmatrix}3&2\\1&4\end{bmatrix}^{1024}$ 的每个元素都大于 10^{700}.

2. $\begin{bmatrix}3&2\\-5&-3\end{bmatrix}^{1024}=I_2$.

3. $\begin{bmatrix}5&7\\-3&-4\end{bmatrix}^{1024}=-\begin{bmatrix}5&7\\-3&-4\end{bmatrix}$.

4. $\begin{bmatrix}5&6.9\\-3&-4\end{bmatrix}^{1024}$ 的每个元素都小于 10^{-70}.

证明*:* 1. $\begin{bmatrix}3&2\\1&4\end{bmatrix}=\begin{bmatrix}1&2\\1&-1\end{bmatrix}\begin{bmatrix}5&\\&2\end{bmatrix}\begin{bmatrix}1&2\\1&-1\end{bmatrix}^{-1}$, 那么

$$\begin{bmatrix}3&2\\1&4\end{bmatrix}^{1024}=\begin{bmatrix}1&2\\1&-1\end{bmatrix}\begin{bmatrix}5^{1024}&\\&2^{1024}\end{bmatrix}\begin{bmatrix}1&2\\1&-1\end{bmatrix}^{-1}=\frac{1}{3}\begin{bmatrix}5^{1024}+2^{1024}&2\cdot5^{2014}-2^{1025}\\5^{1024}-2^{1024}&2\cdot5^{1024}+2^{1024}\end{bmatrix}$$

最小的是 $5^{1024}-2^{1024}>5^{1020}$, 我们只需证明 $5^{320}>2^{700}\Leftrightarrow 5^{16}>2^{35}$. 而 $5^{16}=625^4>512^4=2^{36}$.

2. 该矩阵有两个不同的特征值 $\pm\sqrt{-1}$, 所以存在复矩阵 X，使得

$$\begin{bmatrix}3 & 2\\ -5 & -3\end{bmatrix}=X\begin{bmatrix}\sqrt{-1} & 0\\ 0 & -\sqrt{-1}\end{bmatrix}X^{-1}$$

所以该矩阵的 8 次方等于 I_2。于是 $\begin{bmatrix}3 & 2\\ -5 & -3\end{bmatrix}^{1024}=\left(\begin{bmatrix}3 & 2\\ -5 & -3\end{bmatrix}^{8}\right)^{128}=I_2$.

3. 矩阵的特征多项式是 $\lambda^2-\lambda+1$, 设它的两个特征值是 λ_1,λ_2, 则 $\lambda_1\neq\lambda_2,\lambda_1^3=\lambda_2^3=-1$. 所以存在复矩阵 X，使得

$$\begin{bmatrix}5 & 7\\ -3 & -4\end{bmatrix}=X\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}X^{-1}$$

所以该矩阵的 3 次方等于 $-I_2$. 于是

$$\begin{bmatrix}5 & 7\\ -3 & -4\end{bmatrix}^{1024}=\left(\begin{bmatrix}5 & 7\\ -3 & -4\end{bmatrix}^{3}\right)^{341}\begin{bmatrix}5 & 7\\ -3 & -4\end{bmatrix}=-\begin{bmatrix}5 & 7\\ -3 & -4\end{bmatrix}.$$

4. 矩阵的特征多项式是 $\lambda^2-\lambda+0.7$, 它有两个不同的复特征值 λ_1,λ_2, 则 $|\lambda_1|^2=|\lambda_2|^2=0.7$. 我们不需要算得很精确,只需要估算即可。将它谱分解之后,我们会发现 $\begin{bmatrix}5 & 6.9\\ -3 & -4\end{bmatrix}^{1024}$ 的每个分量的模长都小于 $100(0.7)^{512}$, 我们只需证明 $\left(\dfrac{7}{10}\right)^{512}<10^{-72}$, 或者等价地 $7^{512}<10^{440}$, 或者等价地 $7^{64}<10^{55}$. 而 $7^{64}=49^{32}<50^{32}$, 我们只需证 $50^{32}<10^{55}$, 或者等价的 $5^{32}<10^{23}$, 或者等价的 $5^9<2^{23}$. 又 $5^9=125^3<128^3=2^{21}<2^{23}$. 证毕

□

练习 5.3.8. 设有谱分解 $A=X\Lambda X^{-1}$，求下列矩阵的谱分解：

1. A^{-1}. *2.* A^{T}. *3.* $\begin{bmatrix}A & 0\\ 0 & 2A\end{bmatrix}$. *4.* $\begin{bmatrix}0 & A\\ A & 0\end{bmatrix}$.

证明*:* 1. $A^{-1}=X\Lambda^{-1}X^{-1}$.

2. $A^{\mathrm{T}}=(X^{\mathrm{T}})^{-1}\lambda^{\mathrm{T}}X^{\mathrm{T}}$.

3. $\begin{bmatrix}A & 0\\ 0 & 2A\end{bmatrix}=\begin{bmatrix}X & 0\\ 0 & X\end{bmatrix}\begin{bmatrix}\Lambda & 0\\ 0 & 2\Lambda\end{bmatrix}\begin{bmatrix}X^{-1} & 0\\ 0 & X^{-1}\end{bmatrix}$

4. $\begin{bmatrix}0 & A\\ A & 0\end{bmatrix}=\tilde{X}\tilde{\Lambda}\tilde{X}^{-1}$, 其中 $\tilde{X}=\begin{bmatrix}I_n & I_n\\ I_n & -I_n\end{bmatrix}\begin{bmatrix}X & \\ & X\end{bmatrix},\tilde{\Lambda}=\begin{bmatrix}\Lambda & \\ & -\Lambda\end{bmatrix}$

□

练习 5.3.9. 利用特征值计算下列 n 阶矩阵的行列式：

1. $\begin{vmatrix} a & b & \cdots & b \\ b & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & a \end{vmatrix}$. *2.* $\begin{vmatrix} 1+a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & 1+a_2b_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1}b_n \\ a_nb_1 & \cdots & a_nb_{n-1} & 1+a_nb_n \end{vmatrix}$.

证明*:* 记矩阵为 A

1. $\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ 是 A 的特征向量，特征值是 $a+(n-1)b$; 那么其余特征值的代数重数、几何重数至多是 $n-1$; 注意到 $(a-b)I_n - A$ 所有元素都是 $-b$，所以秩是 1, 那么 $\dim \mathcal{N}((a-b)I_n - A) = n-1$，即 $a-b$ 是特征值，几何重数是 $n-1$，那么这就是所有的特征值了。特征值的乘积是行列式，所以 $\det(A) = (a-b)^{n-1}(a+(n-1)b)$.
2. 记矩阵 $B = \begin{bmatrix} -a_ib_j \end{bmatrix}$, $f(\lambda) = \det(\lambda I_n - B)$, 那么 $\det(A) = f(1)$.

   现在我们考虑 $B = -\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}$, B 是秩 1 矩阵, 所以 $\dim \mathcal{N}(B) = n-1$, 所以 0 是 B 的几何重数为 $n-1$ 的特征值。$\operatorname{trace}(B) = -(a_1b_1 + \cdots + a_nb_n)$, 因为 $\operatorname{trace}(B)$ 是 B 的所有特征值的和，所以 B 还有一个特征值 $-(a_1b_1+\cdots+a_nb_n)$; 那么 $f(\lambda) = \lambda^{n-1}(\lambda + (a_1b_1 + \cdots + a_nb_n))$。从而 $\det(A) = f(1) = 1 + (a_1b_1 + \cdots + a_nb_n)$.

□

练习 5.3.10. 证明：

1. n 阶方阵 $J_n(\lambda_0) = \begin{bmatrix} \lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{bmatrix}$ 只有一个特征值 λ_0，其代数重数是 n，几何重数是 1.

2. 分块对角矩阵 $A = \begin{bmatrix} J_{n_1}(\lambda_0) & 0 \\ 0 & J_{n_2}(\lambda_0) \end{bmatrix}$ 只有一个特征值 λ_0，其代数重数是 n_1+n_2，几何重数是 2.

证明*:* 我们主要用到的结论是上三角矩阵的行列式等于对角线的乘积以及维数公式。

1. $\lambda I_n - J_n(\lambda_0) = J_n(\lambda - \lambda_0)$, 所以 $\det(\lambda I_n - J_n(\lambda_0)) = (\lambda - \lambda_0)^n$. 所以代数重数是 n. $\lambda_0 I_n - J_n(\lambda_0) = J_n(0)$, $\operatorname{rank}(J_n(0)) = n-1$, 所以 $\dim \mathcal{N}(\lambda_0 I_n - J_n(\lambda_0)) = 1$, 所以几何重数是 1.
2. $\lambda I_n - A = \begin{bmatrix} J_{n_1}(\lambda-\lambda_0) & \\ & J_{n_2}(\lambda-\lambda_0) \end{bmatrix}$, 那么 $\det(\lambda I_n - J_n(\lambda_0)) = (\lambda - \lambda_0)^{n_1+n_2}$. 所以代数重数是 n_1+n_2. $\lambda_0 I_n - J_n(\lambda_0) = \begin{bmatrix} J_{n_1}(0) & \\ & J_{n_2}(0) \end{bmatrix}$, $\operatorname{rank}(J_n(0)) = (n_1-1)+(n_2-1)$, 所以 $\dim \mathcal{N}(\lambda_0 I_n - J_n(\lambda_0)) = 2$, 所以几何重数是 2.

$\square$

练习 5.3.11. 试分析秩为 1 的方阵何时可对角化.

证明: 设 $A = \boldsymbol{a}\boldsymbol{b}^{\mathrm{T}}$ 是秩为 1 的矩阵，那么 $\dim \mathcal{N}(A) = n-1$, 即 0 是 A 的几何重数位 $n-1$ 的特征值，那么 0 的代数重数等于 $n-1$ 或 n。另一方面，$\mathrm{trace}(A) = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}$, 由于矩阵的迹是它的特征值的和，所以 A 还有一个特征值是 $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}$.

- 若 $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a} = 0$，那么 0 的代数重数等于 n，不等于其几何重数，所以 A 不可以对角化。
- 若 $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a} \neq 0$，由于 0 的几何重数是 $n-1$，那么 $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a} \neq 0$ 的几何重数只能是 1 了，它的代数重数也是 1, 所以矩阵可以对角化。

$\square$

练习 5.3.12. 设 A 是实二阶方阵，且 $\det(A) < 0$，证明 A 在 $\mathbb{R}$ 上可对角化.

证明: 二阶方阵 A 的特征多项式是 $\lambda^2 - \mathrm{trace}(A)\lambda + \det(A)$. 当 $\det(A) < 0$ 时，它有两个不同的特征值，所以可以对角化。 $\square$

练习 5.3.13. 证明 $A = \begin{bmatrix} I_r & 0 \\ B & -I_{n-r} \end{bmatrix}$ 可对角化.

证明: $\det(\lambda I_n - A) = (\lambda-1)^r(\lambda+1)^{n-r}$, 所以 A 有两个特征值 $1, -1$, 代数重数分别为 $r, n-r$.

- $I_n - A = \begin{bmatrix} 0 & 0 \\ -B & 2I_{n-r} \end{bmatrix}$, 那么 $\mathrm{rank}(I_n - A) = n-r, \dim \mathcal{N}(I_n - A) = r$, 即特征值 1 的几何重数等于 r.
- $-I_n - A = \begin{bmatrix} -2I_r & 0 \\ -B & 0 \end{bmatrix}$, 那么 $\mathrm{rank}(I_n - A) = r, \dim \mathcal{N}(I_n - A) = n-r$, 即特征值 -1 的几何重数等于 $n-r$.

由此我们看出 A 的特征值的代数重数均等于其几何重数，所以 A 可以对角化。 $\square$

练习 5.3.14. 证明：

1. 若 $A^2 = A$，则 A 可对角化.
2. 若 $A^2 = 0$，且 $A \neq 0$，则 A 不可对角化.
3. 若 $A^2 + A + I_n = 0$，则 A 在 $\mathbb{R}$ 上不可对角化.

证明:

1. 令 $V = \mathcal{N}(A), W = \mathcal{N}(I_n - A)$. 我们断言：$V + W = \mathbb{R}^n, V \cap W = \{0\}$. 若 $\boldsymbol{x} \in V \cap W$, 那么 $A\boldsymbol{x} = (I_n - A)\boldsymbol{x} = 0$, 于是 $\boldsymbol{x} = (I_n - A)\boldsymbol{x} + A\boldsymbol{x} = 0$, 这就是说 $V \cap W = \{0\}$. 对任意 $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x} = A\boldsymbol{x} + (I_n - A)\boldsymbol{x}$, 那么 $(I_n - A)A\boldsymbol{x} = A(I_n - A)\boldsymbol{x} = (A - A^2)\boldsymbol{x} = 0$, 所以 $A\boldsymbol{x} \in W, (I_n - A)\boldsymbol{x} \in V$. 这样我们就有 $\dim V + \dim W = n$. 注意到 $\dim V$ 是 A 的特征值 0 的几何重数，$\dim W$ 是 A 的特征值 1 的几何重数。所以 A 可以对角化。
2. 设 $(\lambda, \boldsymbol{x})$ 是 A 的一个特征对，即 $A\boldsymbol{x} = \lambda\boldsymbol{x}$, 那么 $0 = A^2\boldsymbol{x} = \lambda^2\boldsymbol{x}$, 所以 $\lambda = 0$，所以 0 是 A 唯一我特征值，那么它的代数重数就等于 n. 另一方面 $A \neq 0$，则 $\dim \mathcal{N}(A) = n - \mathrm{rank}(A) \leq n-1$，所以几何重数不等于代数重数，因此不可以对角化。

3. 设 $(\lambda, \boldsymbol{x})$ 是 A 的一个特征对，即 $A\boldsymbol{x} = \lambda \boldsymbol{x}$，则 $0 = (A^2 + A + I_n)\boldsymbol{x} = (\lambda^2 + \lambda + 1)\boldsymbol{x}$，于是 $\lambda^2 + \lambda + 1 = 0$. 这说明 A 没有实数特征值，那就在 $\mathbb{R}$ 上不可对角化。

□

练习 5.3.15. 证明或举出反例：

1. 如果 A 所有的特征向量是 $k \begin{bmatrix} 1 \\ 4 \end{bmatrix}, k \neq 0$，则 A 一定有不单的特征值.

2. 如果 A 所有的特征向量是 $k \begin{bmatrix} 1 \\ 4 \end{bmatrix}, k \neq 0$，则 A 一定不可对角化.

3. 如果 A 是上三角矩阵但不是对角矩阵，则 A 不可对角化.

4. 如果 A 是上三角矩阵但不是对角矩阵，则存在对角矩阵 D，使得 $A - D$ 不可对角化.

5. 如果 A 可以被对角矩阵对角化，则 A 也是对角矩阵.

证明: 1. 正确。按题意，A 是 2 阶方阵；如果 A 的特征值都是单特征值，那么它应该有两个不同的特征值，于是有两个线性无关的特征向量，与题意矛盾。

2. 正确。按照上一小题，A 只有一个代数重数为 2 的特征值，但是几何重数是 1, 所以 A 不可以相似对角化。

3. 错误。例如 $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$, 他有两个不同的特征值，那么一定可以对角化。

4. 正确。我们取 D 的对角元就是 A 的对角元，那么 $A - D$ 是对角线元素全是 0 的非零上三角矩阵，因此它只有一个代数重数为 n 的特征值 0; 但是 $A - D \neq 0$, 所以 $\dim \mathcal{N}(A - D) < n$, 即几何重数小于代数重数。所以 $A - D$ 不可以对角化。

5. 正确、按题意，A 是三个对角矩阵的乘积，所以是对角矩阵。

□

练习 5.3.16. 证明定理 *5.3.9* 第 *2* 条。

证明: 证明略。这个证明与第 1 条不仅仅是类似，而是完全照抄即可，只需要把复数域 $\mathbb{C}$ 替换成实数域 $\mathbb{R}$ 即可. □

练习 5.3.17. 证明，对反射矩阵 $H_{\boldsymbol{v}} = I_n - 2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$，其中 $\|\boldsymbol{v}\| = 1$，存在正交矩阵 Q，使得 $Q^{-1} H_{\boldsymbol{v}} Q = \mathrm{diag}(-1, 1, 1, \ldots, 1)$.

证明: 我们知道反射矩阵 $H_{\boldsymbol{v}}$ 满足 $H_{\boldsymbol{v}}\boldsymbol{v} = -\boldsymbol{v}, H_{\boldsymbol{w}}\boldsymbol{w} = \boldsymbol{w}, \forall \boldsymbol{w} \in \{\boldsymbol{v}\}^{\perp}$. $\dim\{\boldsymbol{v}\}^{\perp} = n - 1$, 我们取 $\{\boldsymbol{v}\}^{\perp}$ 的一组标准正交基 $\boldsymbol{w}_2, \cdots, \boldsymbol{w}_n$，那么 $\boldsymbol{v}, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_n$ 就是 $\mathbb{R}^n$ 的一组标准正交基，于是矩阵 $Q = \begin{bmatrix} \boldsymbol{v} & \boldsymbol{w}_2 & \cdots & \boldsymbol{w}_n \end{bmatrix}$ 是正交矩阵，且

$$H_{\boldsymbol{v}} Q = \begin{bmatrix} H_{\boldsymbol{v}}\boldsymbol{v} & H_{\boldsymbol{v}}\boldsymbol{w}_2 & \cdots & H_{\boldsymbol{v}}\boldsymbol{w}_n \end{bmatrix} = Q\mathrm{diag}(-1, 1, 1, \ldots, 1).$$

□

练习 5.3.18 (Householder 矩阵的推广). 设 $\mathcal{M}$ 是 $\mathbb{R}^n$ 的子空间，对任意 $\boldsymbol{v} \in \mathbb{R}^n$，存在唯一的分解 $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$，其中 $\boldsymbol{v}_1 \in \mathcal{M}, \boldsymbol{v}_2 \in \mathcal{M}^{\perp}$. 定义 $\mathbb{R}^n$ 上的变换 $\mathbf{H}$，使得 $\mathbf{H}(\boldsymbol{v}_1 + \boldsymbol{v}_2) = \boldsymbol{v}_1 - \boldsymbol{v}_2$.

1. 证明 $\mathbf{H}$ 是线性变换.

2. 设 H 为该线性变换的表示矩阵，证明 $H^2 = I_n$.

3. 求 H 的特征值、代数重数及特征子空间. H 能否被对角化?

4. 证明，存在矩阵 A，使得 $H = I_n - 2AA^{\mathrm{T}}$.

线性变换 $\mathbf{H}$ 保持 $\mathcal{M}$ 不变，把 $\mathcal{M}^{\perp}$ 中的向量映射到其负向量，因此是关于子空间 $\mathcal{M}$ 的反射变换.

证明: 1. 论证的关键是正交分解的唯一性。我们称满足题中条件的分解为关于 $\mathcal{M}$ 的正交分解。若 $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$ 是 $\boldsymbol{v}$ 关于 $\mathcal{M}$ 的正交分解，那么 $a\boldsymbol{v} = a\boldsymbol{v}_1 + a\boldsymbol{v}_2$ 是 $a\boldsymbol{v}$ 关于 $\mathcal{M}$ 的正交分解. 那么 $\mathbf{H}(a\boldsymbol{v}) = a\boldsymbol{v}_1 - a\boldsymbol{v}_2 = a(\boldsymbol{v}_1 - \boldsymbol{v}_2) = a\mathbf{H}(\boldsymbol{v})$. 若 $\boldsymbol{w} = \boldsymbol{w}_1 + \boldsymbol{w}_2$ 是 $\boldsymbol{w}$ 关于 $\mathcal{M}$ 的正交分解, 那么 $\boldsymbol{v} + \boldsymbol{w} = (\boldsymbol{v}_1 + \boldsymbol{w}_1) + (\boldsymbol{v}_2 + \boldsymbol{w}_2)$ 是 $\boldsymbol{v} + \boldsymbol{w}$ 关于 $\mathcal{M}$ 的正交分解, 所以

$$\mathbf{H}(\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{v}_1 + \boldsymbol{w}_1) - (\boldsymbol{v}_2 + \boldsymbol{w}_2) = (\boldsymbol{v}_1 - \boldsymbol{v}_2) + (\boldsymbol{w}_1 - \boldsymbol{w}_2) = \mathbf{H}(\boldsymbol{v}) + \mathbf{H}(\boldsymbol{w}).$$

因此 $\mathbf{H}$ 是线性映射。

2. 因为 $\mathbf{H}(\boldsymbol{v})$ 关于 $\mathcal{M}$ 的正交分解是 $\mathbf{H}(\boldsymbol{v}) = \boldsymbol{v}_1 + (-\boldsymbol{v}_2)$, 所以 $\mathbf{H}^2(\boldsymbol{v}) = \boldsymbol{v}_1 - (-\boldsymbol{v}_2) = \boldsymbol{v}$, 所以 $\mathbf{H}^2 = \mathrm{id}$, 那么 $H^2 = I_n$.

3. 对任意 $\boldsymbol{v} \in \mathcal{M}, H\boldsymbol{v} = \boldsymbol{v}$, 这说明 1 是 H 的特征值，其几何重数至少等于 $\dim \mathcal{M}$. 对任意 $\boldsymbol{v} \in \mathcal{M}^{\perp}, H(\boldsymbol{v}) = -\boldsymbol{v}$, 这说明 -1 是 H 的特征值，其几何重数至少等于 $\dim \mathcal{M}^{\perp}$. 因为 $\dim \mathcal{M} + \dim \mathcal{M}^{\perp} = n$, 这说明 A 可以对角化。

4. 我们取 $\mathcal{M}$ 的一组标准正交基 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_r$, 令 $A = \begin{bmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_r \end{bmatrix}$. 下面我们验证 $H = I_n - 2AA^{\mathrm{T}}$. 对任意 $\boldsymbol{v} \in \mathcal{M}$, 因为 $\boldsymbol{v}^i\mathrm{T}\boldsymbol{v} = 0 (1 \le i \le r)$, 那么 $A^{\mathrm{T}}\boldsymbol{v} = 0$. 所以 $(I_n - 2AA^{\mathrm{T}})\boldsymbol{v} = \boldsymbol{v}$. 对任意 $\boldsymbol{v}_j (1 \le j \le r)$, 因为我们取的标准正交基，所以 $A^{\mathrm{T}}\boldsymbol{v}_j = \boldsymbol{e}_j$, 那么 $AA^{\mathrm{T}}\boldsymbol{v}_j = A\boldsymbol{e}_j = \boldsymbol{v}_j$; 于是 $(I_n - 2AA^{\mathrm{T}})(\boldsymbol{v}_j) = \boldsymbol{v}_j - 2\boldsymbol{v}_j = -\boldsymbol{v}_j$. 因为 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_r$ 是 $\mathcal{M}^{\perp}$ 的一组基, 所以对任意 $\boldsymbol{v} \in \mathcal{M}^{\perp}$, 就有 $(I_n - 2AA^{\mathrm{T}})(\boldsymbol{v}) = -\boldsymbol{v}$. 所以 $H = I_n - 2AA^{\mathrm{T}}$. □

练习 5.3.19. 设 A 是 n 阶实方阵，满足 $A^2 = I_n$，证明 A 在 $\mathbb{R}$ 上可对角化，并判断 A 是否为关于某个子空间的反射变换.

证明: 令 $V = \mathcal{N}(I_n - A), W = \mathcal{N}(I_n + A)$. 我们断言: $V + W = \mathbb{R}^n, V \cap W = \{0\}$. 若 $\boldsymbol{x} \in V \cap W$, 那么 $(I_n + A)\boldsymbol{x} = (I_n - A)\boldsymbol{x} = 0$, 于是 $\boldsymbol{x} = \dfrac{(I_n - A)\boldsymbol{x} + (I_n + A)\boldsymbol{x}}{2} = 0$, 这就是说 $V \cap W = \{0\}$. 对任意 $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x} = \dfrac{(I_n - A)\boldsymbol{x} + (I_n + A)\boldsymbol{x}}{2}$, 那么 $(I_n - A)(I_n + A)\boldsymbol{x} = (I_n + A)(I_n - A)\boldsymbol{x} = (I_n - A^2)\boldsymbol{x} = 0$, 所以 $(I_n - A)\boldsymbol{x} \in W, (I_n + A)\boldsymbol{x} \in V$. 这样我们就有 $\dim V + \dim W = n$. 注意到 $\dim V$ 是 A 的特征值 1 的几何重数，$\dim W$ 是 A 的特征值 -1 的几何重数。所以 A 可以对角化。

A 未必是关于某个子空间的反射变换。例如 $A = \begin{bmatrix} 1 & -2 \\ 0 & -1 \end{bmatrix}$; 因为根据上一题第 4 小题知，关于某个子空间的反射变换对应的矩阵是对称矩阵，而 $\begin{bmatrix} 1 & -2 \\ 0 & -1 \end{bmatrix}$ 并不对称。

事实上，若 A 对称，则 A 是关于 $V=\mathcal{N}(I_n-A)$ 的投影矩阵。要证明这一点，只需证明 $V\perp W$. 对任意 $\boldsymbol{v}\in V,\boldsymbol{w}\in W$

$$\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w}=(A\boldsymbol{v})^{\mathrm{T}}\boldsymbol{w}=\boldsymbol{v}^{\mathrm{T}}(A\boldsymbol{w})=-\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w}$$

所以 $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w}=0$. □

练习 5.3.20. 设矩阵 A 的特征多项式是 $p_A(x)$.

1. 设 A 为对角矩阵，证明 $p_A(A)=0$.
2. 设 A 为可对角化的矩阵，证明 $p_A(A)=0$.
3. 设 $A=\begin{bmatrix}0&1\\-1&2\end{bmatrix}$，它是否可对角化？是否满足 $p_A(A)=0$?

证明: 1. 我们首先注意到：若干个对角矩阵的乘积也是对角矩阵，并且乘积的对角元是每个对角矩阵相同位置对角元的乘积。设 $A=\operatorname{diag}(a_1,\cdots,a_n)$, 那么 $p_A(x)=(x-a_1)\cdots(x-a_n)$, 于是 $p_A(A)=(A-a_1I_n)\cdots(A-a_nI_n)$. 对任意 $1\le i\le n$, $A-a_iI_n$ 也是对角矩阵，且它的第 i 个对角元是 0. 所以 $p_A(A)=0$

2. 我们注意到：若 X 是可逆矩阵，且 $B=XAX^{-1}$, 那么 $p_A(x)=p_B(x)=p(x)$, 且 $p(B)=Xp(A)X^{-1}$. 若 A 可对角化，设可逆矩阵 X 使得 $XAX^{-1}=B$ 是对角矩阵，那么 $p_B(B)=0$, 所以 $p_A(X)=X^{-1}p_B(B)X=0$.

3. $\det(\lambda I_2-A)=\lambda^2-2\lambda+1=(\lambda-1)^2$, 所以它有唯一的特征值 1; $I_2-A=\begin{bmatrix}1&-1\\1&-1\end{bmatrix}$, 那么 $\dim\mathcal{N}(I_2-A)=1$, 代数重数大于几何重数，所以不能对角化。而

$$p_A(A)=(I_2-A)^2=\begin{bmatrix}1&-1\\1&-1\end{bmatrix}^2=0.$$

□

练习 5.3.21. 求下列矩阵的特征值，并判断是否可以对角化：

1. $\begin{bmatrix}10&1&\\&10&1\\&&10\end{bmatrix},\begin{bmatrix}10&1&\\&10.001&1\\&&10.002\end{bmatrix}$. 2. $\begin{bmatrix}0&1&&\\&0&1&\\&&0&1\\&&&0\end{bmatrix},\begin{bmatrix}0&1&&\\&0&1&\\&&0&1\\0.001&&&0\end{bmatrix}$.

附注. 由此可见，矩阵变化不大时，特征值和可对角化性质可能变化很大.

证明: 1. 第一个矩阵有唯一的特征值 10, 代数重数是 3, 几何重数是 1, 所以不可以对角化。

第二个矩阵有三个不同的特征值 $10,10.001,10.002$, 所以可以对角化

2. 第一个矩阵有唯一的特征值 0, 代数重数是 4, 几何重数是 1, 所以不可以对角化。

第二个矩阵的特征多项式是 $\lambda^4+0.001$, 在复数域上有 4 个不同的根，所有矩阵有 4 个不同的特征值，因此在复数域上可以对角化。

□

5.4 相　　似

练习 5.4.1. 设 A, B 是 n 阶方阵，且满足 $AB = BA$.

1. 证明，若 A 有 n 个不同的特征值，则 B 可对角化.
2. 若 A 有代数重数大于 1 的特征值，B 是否一定可对角化?

证明: 1. 设 $(\lambda_1, \boldsymbol{v}_1), \cdots, (\lambda_n, \boldsymbol{v}_n)$ 是 A 的特征对，因为 $\lambda_1, \cdots, \lambda_n$ 互不相等，那么 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 线性无关，且 $\dim \mathcal{N}(\lambda_i I_n - A) = 1$。对 $1 \le i \le n$, $A(B\boldsymbol{v}_i) = B(A\boldsymbol{v}_i) = \lambda_i(B\boldsymbol{v}_i)$, 所以存在 μ_i 使得 $B\boldsymbol{v}_i = \mu_i \boldsymbol{v}_i$, 所以 B 有 n 个线性无关的特征向量，所以 B 可以对角化。

2. 不一定，例如 $A = I_n, B = J_n(0)$.

□

练习 5.4.2. 对下列矩阵 A, B，求 X 使得 $A = XBX^{-1}$.

1. $A = MN, B = NM$，其中 M, N 是方阵，且 M 可逆.
2. $A = \begin{bmatrix} MN & 0 \\ N & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ N & NM \end{bmatrix}$，其中 M, N 不必是方阵.
3. $A = \begin{bmatrix} M & -N \\ N & M \end{bmatrix}, B = \begin{bmatrix} M + iN & 0 \\ 0 & M - iN \end{bmatrix}$，其中 M, N 是方阵.

证明: 1. 取 $X = M$.

2. 我们把 X 分块，然后求解 $AX = XB$ 即可。若 $M \in M_{m\times n}(\mathbb{R})$, 我们可以取 $X = \begin{bmatrix} I_n & M \\ 0 & I_n \end{bmatrix}$

3. 我们先把 M, N 看成是数，然后这个问题要做的就是对 A 进行对角化；然后我们类比，猜出 X 即可。我们省略其中的计算，把关键步骤写下来。当 M, N 都是数时，进行对角化的过程，我们有

$$A \begin{bmatrix} 1 & 1 \\ -\boldsymbol{i} & \boldsymbol{i} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -\boldsymbol{i} & \boldsymbol{i} \end{bmatrix} B$$

那么我们猜测的过程就是把数字换成矩阵:1 换成 I_n, $\boldsymbol{i}$ 换成 $\boldsymbol{i}I_n$ 即可。所以我们令 $X = \begin{bmatrix} I_n & I_n \\ -\boldsymbol{i}I_n & \boldsymbol{i}I_n \end{bmatrix}$，作初等行变换可以验证 X 是可逆的，然后再验证 $AX = XB$ 即可；这仅仅是矩阵乘法的运算而已，我们忽略。

□

练习 5.4.3. 给定 n 阶方阵 A, B 满足 $AB = BA$.

1. 证明，若 A 有 n 个不同的特征值，则存在次数不超过 $n-1$ 的多项式 $f(x)$，使得 $B = f(A)$.
2. 证明，若 $A = J_n(\lambda)$，则存在次数不超过 $n-1$ 的多项式 $f(x)$，使得 $B = f(A)$.
3. 举例说明，存在 A, B 满足 $AB = BA$，但不存在多项式 $f(x)$，使得 $B = f(A)$.

证明: 1. 设 $(\lambda_1, \boldsymbol{v}_1), \cdots, (\lambda_n, \boldsymbol{v}_n)$ 是 A 的特征对，在练习 5.4.1中，我们已经证明了存在 μ_i 使得 $B\boldsymbol{v}_i = \mu_i \boldsymbol{v}_i$。我们令 $X = \begin{bmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{bmatrix}$, 则 $\tilde{A} = X^{-1}AX = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$, $\tilde{B} =$

$X^{-1}BX = \operatorname{diag}(\mu_1, \cdots, \mu_n)$。另一方面，$f(A) = B$ 当且仅当 $f(\tilde{A}) = \tilde{B}$；而 $f(\tilde{A}) = \operatorname{diag}(f(\lambda_1), \cdots, f(\lambda_n))$，所以问题转化为证明：存在次数不超过 $n-1$ 的多项式 $f(x)$ 使得 $f(\lambda_i) = \mu_i (1 \le i \le n)$. 设 $f(x) = a_{n-1}x^{n-1} + \cdots + a_0$, 那么我们就是要解方程组

$$\begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & & & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}$$

系数矩阵是 Vandermonde 矩阵，$\lambda_1, \cdots, \lambda_n$ 互不相同则系数矩阵可逆，因此该方程组有解。

2. 设 $B = \begin{bmatrix} b_{ij} \end{bmatrix}$. 因为 $A = J_n(0) + \lambda I_n$, 那么 $AB = BA$ 当且仅当 $J_n(0)B = BJ_n(0)$。我们直接观察：
   - $BJ_n(0)$ 第一列是 0，对 $2 \le i \le n$, 它的第 i 列是 B 的第 $i-1$ 列；
   - $J_n(0)B$ 第 n 行是 0，对 $1 \le i \le n-1$, 它的第 i 行是 B 的第 $i+1$ 行。

   我们比较 $BJ_n(0)$ 和 $J_n(0)B$ 的每一个分量，就得到 $b_{21} = \cdots = b_{n1} = b_{n2} = \cdots = b_{n,n-1} = 0$, 且对任意 $1 \le i, j \le n$ 都有 $b_{ij} = b_{i+1,j+1}$; 那么 B 形如

$$\begin{bmatrix} b_1 & b_2 & \cdots & & b_{n-1} & b_n \\ 0 & b_1 & b_2 & \ddots & b_{n-2} & b_{n-1} \\ 0 & 0 & b_1 & & & \\ \vdots & & \ddots & \ddots & \ddots & \\ 0 & & \cdots & 0 & b_1 & b_2 \\ 0 & & \cdots & 0 & 0 & b_1 \end{bmatrix}$$

   那么 $B = b_1 I_n + b_2 J_n(0) + \cdots + b_{n-1}J_n(0)^{n-2} + b_n J_n(0)^{n-1}$, 那么我们取

$$f(x) = b_n(x-\lambda)^{n-1} + b_{n-1}(x - \lambda^{n-2}) + \cdots + b_2(x-\lambda) + b_1$$

   即可。

3. 例如 $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$.

□

练习 5.4.4. 证明，任意迹为 0 的方阵相似于一个对角元素全为 0 的方阵.

利用数学归纳法. 先取非零向量 $\boldsymbol{v}$ 且 $\boldsymbol{v}$ 不是 A 的特征向量，再把 $\boldsymbol{v}, A\boldsymbol{v}$ 扩充成一组基做相似变换.

证明: 我们对阶数 n 归纳。$n = 1$ 时，命题是显然的。假设命题对 $n-1$ 成立，我们考虑 n 阶方阵 A.

- 若任意非零向量都是 A 的特征向量，那么 $A = cI_n$. $\operatorname{trace}(A) = nc$, 所以 $c = 0$, 即 A 是零矩阵，所以命题成立。

- 若存在非零向量 $\boldsymbol{v}$ 不是 A 的特征向量，那么 $A\boldsymbol{v} \neq 0$ 且 $\boldsymbol{v}, A\boldsymbol{v}$ 线性无关。那么我们把 $\boldsymbol{v}, A\boldsymbol{v}$ 扩充成 $\mathbb{R}^n$ 的一组基：$\boldsymbol{v}_1 = \boldsymbol{v}, \boldsymbol{v}_2 = A\boldsymbol{v}, \boldsymbol{v}_3, \cdots, \boldsymbol{v}_n$, 令 $X = \begin{bmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{bmatrix}$, 那么

$$AX = \begin{bmatrix} A\boldsymbol{v}_1 & A\boldsymbol{v}_2 & \cdots & A\boldsymbol{v}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}_2 & A\boldsymbol{v}_2 & \cdots & A\boldsymbol{v}_n \end{bmatrix} = X \begin{bmatrix} 0 & \alpha^{\mathrm{T}} \\ \beta & B \end{bmatrix}$$

其中，$\beta = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$. 也就是说 A 和 $\begin{bmatrix} 0 & \alpha^{\mathrm{T}} \\ \beta & B \end{bmatrix}$ 相似，那么 $\mathrm{trace}(A) = \mathrm{trace}(B)$. 根据归纳假设，$B$ 相似于对角元素全是 0 的方阵 C，因此存在 $n-1$ 阶可逆矩阵 Y 使得 $B = CYC^{-1}$, 那么

$$A = X \begin{bmatrix} 0 & \alpha^{\mathrm{T}} \\ \beta & YCY^{-1} \end{bmatrix} X^{-1} = X \begin{bmatrix} 1 & 0 \\ 0 & Y \end{bmatrix} \begin{bmatrix} 0 & \alpha^{\mathrm{T}} Y \\ Y^{-1}\beta & C \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & Y \end{bmatrix}^{-1} X^{-1}$$

这说明 A 相似于 $\begin{bmatrix} 0 & \alpha^{\mathrm{T}} Y \\ Y^{-1}\beta & C \end{bmatrix}$; 而 $\begin{bmatrix} 0 & \alpha^{\mathrm{T}} Y \\ Y^{-1}\beta & C \end{bmatrix}$ 对角元都是 0, 这就完成了归纳。

□

练习 5.4.5. *利用 Jordan 标准形证明，A 与 A^{T} 相似.*

证明: 根据 Jordan 标准型理论，存在可逆矩阵 X 和约当块 $J_{n_i}(\lambda_1), \cdots, J_{n_r}(\lambda_r)$, 使得

$$A = X \mathrm{diag}\left(J_{n_i}(\lambda_1), \cdots, J_{n_r}(\lambda_r)\right) X^{-1}$$

此时 $A^{\mathrm{T}} = (X^{\mathrm{T}})^{-1} \mathrm{diag}\left(J_{n_i}(\lambda_1)^{\mathrm{T}}, \cdots, J_{n_r}(\lambda_r)^{\mathrm{T}}\right) X^{\mathrm{T}}$, 所以我们只需要证明：Jordan 块 $J_n(\lambda)$ 与 $J_n(\lambda)^{\mathrm{T}}$ 相似，进一步地只需要证明 $J_n(0)$ 和 $J_n(0)^{\mathrm{T}}$ 相似。这个我们直接观察：令 P 是反对角线上元素都是 1 其他元素都是 0 的 n 阶方阵，那么 $P^2 = I_n$ 且 $J_n(0)^{\mathrm{T}} = PJ_n(0)P$. □

练习 5.4.6 (Jordan 链). *对任意 n 阶方阵 A，一组向量 $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_s$，如果存在 $\lambda \in \mathbb{C}$，使得*

$\boldsymbol{x}_1 \neq 0$, $(A - \lambda I)\boldsymbol{x}_1 = 0$, $(A - \lambda I)\boldsymbol{x}_2 = \boldsymbol{x}_1$, $\ldots$, $(A - \lambda I)\boldsymbol{x}_s = \boldsymbol{x}_{s-1}$, 而 $(A - \lambda I)\boldsymbol{y} = \boldsymbol{x}_s$ 无解,

就称其为 A 的一个关于 λ 长度为 s 的 Jordan **链**.

显然，$\boldsymbol{x}_1 \in \mathcal{N}(A - \lambda I)$ 是特征向量，$\boldsymbol{x}_2 \in \mathcal{N}\left((A - \lambda I)^2\right), \ldots, \boldsymbol{x}_s \in \mathcal{N}((A - \lambda I)^s)$，但 $\boldsymbol{x}_s \notin \mathcal{R}(A - \lambda I)$.

求证：

1. *$\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_s$ 线性无关.*
2. *令 $X_1 = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_s \end{bmatrix}$，则 $AX_1 = X_1 J_s(\lambda)$.*
3. *若 A 只有一个特征值 λ，且其几何重数为 1，则 A 相似于 $J_n(\lambda)$，且有一个关于 λ 的长度为 n 的 Jordan 链 $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$.*

证明: 1. 设 $a_1\boldsymbol{x}_1 + \cdots + a_s\boldsymbol{x}_s = 0$; 那么 $(A - \lambda I)^{s-1}(a_1\boldsymbol{x}_1 + \cdots + a_s\boldsymbol{x}_s) = 0$, 那么 $a_s(A - \lambda I)^{s-1}(\boldsymbol{x}) = 0$. 而 $(A - \lambda I)^{s-1}(\boldsymbol{x}_s) = \boldsymbol{x}_1$, 所以 $a_s = 0$. 然后我们再左乘 $(A - \lambda I)^{s-2}$ 可得 $a_{s-1} = 0$, 同理类推可知 $a_1 = \cdots = a_s = 0$.

2. 由 $(A-\lambda)\boldsymbol{x}_{i+1}=\boldsymbol{x}_i$，即 $A\boldsymbol{x}_{i+1}=\boldsymbol{x}_i+\lambda\boldsymbol{x}_{i+1}$, 所以

$$AX_1=\begin{bmatrix}A\boldsymbol{x}_1 & A\boldsymbol{x}_2 & \cdots & A\boldsymbol{x}_s\end{bmatrix}=\begin{bmatrix}\lambda\boldsymbol{x}_1 & \boldsymbol{x}_1+\lambda\boldsymbol{x}_2 & \cdots & \boldsymbol{a}_{s-1}+\lambda\boldsymbol{x}_s\end{bmatrix}=X_1J_s(\lambda).$$

3. 令 $B=A-\lambda I_n$. 我们考虑

$$0\subseteq\mathcal{N}(B)\subseteq\mathcal{N}(B^2)\subseteq\cdots\subseteq\mathcal{N}(B^k)\subseteq\cdots \qquad *$$

首先我们证明：对 $k\geq 1$, 若 $\mathcal{N}(B^k)=\mathcal{N}(B^{k+1})$, 那么 $\mathcal{N}(B^{k+1})=\mathcal{N}(B^{k+2})$. 我们只需证明 $\mathcal{N}(B^{k+2})\subseteq\mathcal{N}(B^{k+1})$. 对任意 $\boldsymbol{v}\in\mathcal{N}(B^{k+2})$, $B^{k+1}(B\boldsymbol{v})=B^{k+2}\boldsymbol{v}=0$, 所以 $B\boldsymbol{v}\in\mathcal{N}(B^{k+1})=\mathcal{N}(B^k)$, 故 $B^{k+1}\boldsymbol{v}=B^k(B\boldsymbol{v})=0$, 即 $\boldsymbol{v}\in\mathcal{N}(B^{k+1})$.
这说明在 $*$ 式中，一旦某个 $\subseteq$ 取等号，那么后面的 $\subseteq$ 全是等号。我们假设 s 满足

$$0\subsetneq\mathcal{N}(B)\subsetneq\cdots\subsetneq\mathcal{N}(B^s)=\mathcal{N}(B^{s+1})=\cdots$$

由维数公式，我们有

$$\mathbb{R}^n\supsetneq\mathcal{R}(B)\supsetneq\cdots\supsetneq\mathcal{R}(B^s)=\mathcal{R}(B^{s+1})=\cdots$$

其次我们证明 $\mathcal{N}(B^s)\cap\mathcal{R}(B^s)=\{0\}$. 对任意 $\boldsymbol{v}\in\mathcal{N}(B^s)\cap\mathcal{R}(B^s)$, $\boldsymbol{v}\in\mathcal{R}(B^s)$ 说明 $\boldsymbol{v}=B^s\boldsymbol{w}$, $\boldsymbol{v}\in\mathcal{N}(B^s)$ 说明 $B^{2s}\boldsymbol{w}=B^s\boldsymbol{v}=0$, 于是 $\boldsymbol{w}\in\mathcal{N}(B^{2s})=\mathcal{N}(B^s)$, 所以 $\boldsymbol{v}=B^s\boldsymbol{w}=0$.
若 $\dim\mathcal{R}(B^s)=m\neq 0$，我们取 $\mathcal{R}(B^s)$ 的一组基 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_m$, 由于 $\mathcal{R}(B^s)=\mathcal{R}(B^{s+1})$，存在矩阵 $C\in M_{m\times m}(\mathbb{C})$ 使得

$$B\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_m\end{bmatrix}=\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_m\end{bmatrix}C$$

我们取 C 的一个特征对 $(\mu,\boldsymbol{y})$, 那么我们就有

$$B\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_m\end{bmatrix}\boldsymbol{y}=\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_m\end{bmatrix}C\boldsymbol{y},$$

即 $B(y_1\boldsymbol{v}_1+\cdots+y_m\boldsymbol{v}_m)=\mu(y_1\boldsymbol{v}_1+\cdots+y_m\boldsymbol{v}_m)$, 也就是说 $(\mu,y_1\boldsymbol{v}_1+\cdots+y_m\boldsymbol{v}_m)$ 也是 B 的一个特征对；这和 B 只有几何重数为 1 的唯一一个特征值矛盾。所以 $\mathcal{R}(B^s)=0$, 进而由维数公式：$\dim\mathcal{N}(B^s)=n-\dim\mathcal{R}\,B^s=n$, 所以 $\mathcal{N}(B^s)=\mathbb{R}^n$.
取 $\boldsymbol{x}_s\in\mathcal{N}(B^s)-\mathcal{N}(B^{s-1})$, 对 $1\leq i\leq s-1$, 令 $\boldsymbol{x}_i=B^{s-i}\boldsymbol{x}_s\in\mathcal{N}(B^i)$, 那么对任意 $1\leq i\leq s$, 都有 $\boldsymbol{x}_i=B^{j-i}\boldsymbol{x}_j$. 我们断言：$\boldsymbol{x}_i\notin\mathcal{N}(B^{i-1})$, 若不然则有 $B^{i-1}\boldsymbol{x}_i=0$, 即 $B^{s-1}\boldsymbol{x}_s=0$, 也即 $\boldsymbol{x}_s\in\mathcal{N}(B^{s-1})$; 这和我们的选取矛盾。下面我们来证明：$\boldsymbol{x}_1,\cdots,\boldsymbol{x}_s$ 构成 $\mathcal{N}(B^s)$ 的一组基。

- 若 $a_1\boldsymbol{x}_1+\cdots+a_s\boldsymbol{x}_s=0$; 两边左乘 B^{s-1} 即得 $a_s=0$; 再两边左乘 B^{s-2} 即得 $a_{s-1}=0$; 以此类推即可得 $a_s=a_{s-1}=\cdots=a_1=0$.
- 对任意 $\boldsymbol{x}\in\mathcal{N}(B^s)$, 那么 $B^{s-1}\boldsymbol{x}\in\mathcal{N}(B)$；因为 $\dim\mathcal{N}(B)=1$，那么存在 a_1 使得

  $$B^{s-1}\boldsymbol{x}=a_1\boldsymbol{x}_1=a_1B^{s-1}\boldsymbol{x}_s;$$

  所以 $\boldsymbol{x}-a_1\boldsymbol{x}_s\in\mathcal{N}(B^{s-1})$, 那么 $B^{s-2}(\boldsymbol{x}-a_1\boldsymbol{x}_s)\in\mathcal{N}(B)$, 于是存在 a_2 使得

  $$B^{s-2}(\boldsymbol{x}-a_1\boldsymbol{x}_s)=a_2\boldsymbol{x}_1=a_2B^{s-2}\boldsymbol{x}_{s-1};$$

所以 $\boldsymbol{x}-a_1\boldsymbol{x}_s-a_2\boldsymbol{x}_{s-1}\in\mathcal{N}(B^{s-2})$. 以此类推,我们可以证明存在 $a_3,a_4,\cdots,a_{s-1}$, 使得 $\boldsymbol{x}-a_1\boldsymbol{x}_s-a_2\boldsymbol{x}_{s-1}-\cdots-a_{s-1}\boldsymbol{x}_2\in\mathcal{N}(B)$, 那么存在 a_s 使得 $\boldsymbol{x}-a_1\boldsymbol{x}_s-a_2\boldsymbol{x}_{s-1}-\cdots-a_{s-1}\boldsymbol{x}_2=a_s\boldsymbol{x}_1$. 这就说明了 $\mathcal{N}(B^s)=\mathrm{Span}(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_s)$.

综合起来就有 $s=\dim\mathcal{N}(B^s)=\dim\mathbb{R}^n=n$.

□

练习 5.4.7. 给定 m 阶方阵 A_1，n 阶方阵 A_2 和 $m\times n$ 矩阵 B. 证明：

1. 如果 A_1 和 A_2 没有相同的特征值，则关于 $m\times n$ 矩阵 X 的 *Sylvester* 方程 $A_1X-XA_2=B$ 有唯一解.

2. 如果 A_1 和 A_2 没有相同的特征值，则存在唯一的矩阵 X 满足

$$\begin{bmatrix} I_m & X \\ 0 & I_n \end{bmatrix}\begin{bmatrix} A_1 & B \\ 0 & A_2 \end{bmatrix}\begin{bmatrix} I_m & X \\ 0 & I_n \end{bmatrix}^{-1}=\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}.$$

3. 对 n 阶方阵 A，存在可逆矩阵 X，使得

$$X^{-1}AX=\begin{bmatrix} \lambda_1I+N_1 & & & \\ & \lambda_2I+N_2 & & \\ & & \ddots & \\ & & & \lambda_sI+N_s \end{bmatrix},$$

其中 $N_1,N_2,\ldots,N_s$ 是严格上三角矩阵.

4. 如果 $m>n$ 且 $A_1=J_m(0),A_2=J_n(0),B=\boldsymbol{e}_1\boldsymbol{b}^{\mathrm{T}}$，即 B 除第一行外元素全为 0，则关于 $m\times n$ 矩阵 X 的 *Sylvester* 方程 $A_1X-XA_2=B$ 有解，于是存在矩阵 X 满足

$$\begin{bmatrix} I_m & X \\ 0 & I_n \end{bmatrix}\begin{bmatrix} A_1 & B \\ 0 & A_2 \end{bmatrix}\begin{bmatrix} I_m & X \\ 0 & I_n \end{bmatrix}^{-1}=\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}.$$

5. 对 n 阶方阵 $A=\begin{bmatrix} J_{k_1}(0) & \boldsymbol{e}_1\boldsymbol{a}_2^{\mathrm{T}} & \cdots & \boldsymbol{e}_1\boldsymbol{a}_r^{\mathrm{T}} \\ & J_{k_2}(0) & & \\ & & \ddots & \\ & & & J_{k_r}(0) \end{bmatrix}$，其中 $k_1>k_2\geq\cdots\geq k_r$（即 A 是一个 *Jordan* 标准形与除第一行外元素全为 0 的矩阵的和），存在可逆矩阵 X，使得

$$X^{-1}AX=\begin{bmatrix} J_{k_1}(0) & & & \\ & J_{k_2}(0) & & \\ & & \ddots & \\ & & & J_{k_r}(0) \end{bmatrix}.$$

证明*:* 1. 由命题 5.4.6，存在 n 阶可逆矩阵 P 使得 $\tilde{A}_2=P^{-1}A_2P$ 是上三角矩阵, 那么

$$\begin{aligned} A_1X-XA_2=B \text{ 有唯一解} &\Leftrightarrow A_1XP-XP\tilde{A}_2=BP \text{ 有唯一解} \\ &\Leftrightarrow A_1Y-Y\tilde{A}_2=BP \text{ 有唯一解} \end{aligned}$$

因为 $\tilde{A}_2$ 和 A_2 有相同的特征值，那么 $\tilde{A}_2$ 和 A_1 没有相同的特征值；根据练习 5.2.23，$A_1Y-Y\tilde{A}_2=BP$ 有唯一解。

2. 等价于证明：存在矩阵 X 使得：$\begin{bmatrix} I_m & X \\ 0 & I_n \end{bmatrix}\begin{bmatrix} A_1 & B \\ 0 & A_2 \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}\begin{bmatrix} I_m & X \\ 0 & I_n \end{bmatrix}$，即

$$\begin{bmatrix} A_1 & B + XA_2 \\ 0 & A_2 \end{bmatrix} = \begin{bmatrix} A_1 & A_1X \\ 0 & A_2 \end{bmatrix}$$

所以等价于证明：存在矩阵 X 使得 $B + XA_2 = A_1X$. 因为 A_1, A_2 没有相同的特征值，那么由上一小题，知存在唯一的 X 满足题意。

3. 根据命题 5.4.6，存在可逆矩阵 X_1, 使得

$$X_1^{-1}AX_1 = \begin{bmatrix} \lambda_1 I + N_1 & A_{12} & \cdots & A_{1s} \\ & \lambda_2 I + N_2 & A_{23} & \vdots \\ & & \ddots & \\ & & & \lambda_s I + N_s \end{bmatrix},$$

其中 $\lambda_1, \cdots, \lambda_s$ 互不相同，$N_1, \cdots, N_s$ 都是严格上三角矩阵。下一步就是通过相似，逐步把右上角的 A_{ij} 消去。我们重新分块：

$$A_1 = \begin{bmatrix} \lambda_1 I + N_1 & A_{12} & \cdots & A_{1,s-1} \\ & \lambda_2 I + N_2 & A_{23} & \vdots \\ & & \ddots & \\ & & & \lambda_{s-1} I + N_{s-1} \end{bmatrix}, A_2 = \lambda_s I + N_s$$

那么 A_1, A_2 没有相同的特征值，所以存在 Y_s 满足

$$\begin{bmatrix} I_{n-n_s} & Y_s \\ 0 & I_{n_s} \end{bmatrix}\begin{bmatrix} A_1 & \star \\ 0 & A_2 \end{bmatrix}\begin{bmatrix} I_{n-n_s} & Y_s \\ 0 & I_{n_s} \end{bmatrix}^{-1} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

那么我们再对 A_1 重新分块，用同样的操作，把 $\lambda_{s-1}I + N_{s-1}$ 上方的矩阵全消去，直至消去所有的 A_{ij}。这样就证明了命题。

4. 记 $\boldsymbol{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$；我们还是逐列构造 $X = \begin{bmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_n \end{bmatrix}$. 由于 A_2 第一列是 0, 那么该方程的第一列就是 $A\boldsymbol{x}_1 = b_1\boldsymbol{e}_1$, 因此 $\boldsymbol{x}_1 = \begin{bmatrix} x_1 \\ b_1 \\ 0 \\ \vdots \end{bmatrix}$, 其中 x_1 是自由的. 假设我们求出了 X 的第 i 列 $\boldsymbol{x}_i$，事实上，我们逐个归纳，可以假设 $\boldsymbol{x}_i = \begin{bmatrix} x_i \\ x_{i-1} + b_i \\ \vdots \\ x_1 + b_2 \\ b_1 \\ 0 \\ \vdots \end{bmatrix}$, 其中 x_i 是自由的，

那么考虑 $A_1X - XA_2 = B$ 的第 $i+1$ 列，则得到方程 $A\boldsymbol{x}_{i+1} - \boldsymbol{x}_i = b_{i+1}\boldsymbol{e}_1$. 这个方程组有解，解形如 $\boldsymbol{x}_{i+1} = \begin{bmatrix} x_{i+1} \\ x_i + b_{i+1} \\ \vdots \\ x_1 + b_2 \\ b_1 \\ 0 \\ \vdots \end{bmatrix}$. 因为 $m > n$，这样的步骤可以一直进行下去，知道解出 X 的所有列。

5. 证明方法和第 3 小题类似；留给读者自己完成。

$\square$

第六章　实对称矩阵

6.1　实对称矩阵的谱分解

练习 6.1.1. 求下列实对称矩阵的谱分解：

1. $\begin{bmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{bmatrix}$.　*2.* $\begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$.　*3.* $\begin{bmatrix} 0 & 0 & 4 & 1 \\ 0 & 0 & 1 & 4 \\ 4 & 1 & 0 & 0 \\ 1 & 4 & 0 & 0 \end{bmatrix}$.　*4.* $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$.

步骤和一般的矩阵相似对角化类似，仅仅多了两步步：对同一特征值的特征向量多 Gram-Schmidt 正交化，再对所有特征向量作单位化。

证明*:* 记原矩阵为 A.

1. $\det(\lambda I_3 - A) = (\lambda-1)(\lambda-4)(\lambda+2)$, 所以 A 有三个不同的特征值 $1, 4, -2$, 对应的单位特征向量分别是 $\begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}, \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$, 那么 A 的谱分解为
$$A = \begin{bmatrix} -2/3 & 2/3 & 1/3 \\ -1/3 & -2/3 & 2/3 \\ 2/3 & 1/3 & 2/3 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 4 & \\ & & -2 \end{bmatrix} \begin{bmatrix} -2/3 & 2/3 & 1/3 \\ -1/3 & -2/3 & 2/3 \\ 2/3 & 1/3 & 2/3 \end{bmatrix}^{\mathrm{T}}$$

2. $\det(\lambda I_3 - A) = (\lambda-1)^2(\lambda-10)$, 所以 A 有两个不同的特征值 $10, 1$. 特征值 10 的一个单位特征向量是：$\begin{bmatrix} -1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$。用固定步骤解齐次方程组，可以得到 $\mathcal{N}(I_3 - A)$ 的一组基是 $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$, 作 Gram-Schmidt 正交化得到：$\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2/5 \\ 4/5 \\ 1 \end{bmatrix}$；再单位化就得到: $\begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix}, \begin{bmatrix} 2/\sqrt{45} \\ 4/\sqrt{45} \\ 5/\sqrt{45} \end{bmatrix}$。那么 A 的谱分解为
$$A = \begin{bmatrix} -1/3 & -2/\sqrt{5} & 2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{5} \\ 2/3 & 0 & 5/\sqrt{5} \end{bmatrix} \begin{bmatrix} 1 & & \\ & 4 & \\ & & -2 \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} & -2/\sqrt{5} & 2/\sqrt{45} \\ -2/\sqrt{5} & 1/\sqrt{5} & 4/\sqrt{5} \\ 2/\sqrt{5} & 0 & 5/\sqrt{5} \end{bmatrix}^{\mathrm{T}}$$

3. $\det(\lambda I_4 - A) = (\lambda^2-25)(\lambda^2-9)$, 所以 A 有三个不同的特征值 $5, -5, 3\ -3$, 对应的单位特征向量分别是 $\begin{bmatrix}1/2\\1/2\\1/2\\1/2\end{bmatrix}, \begin{bmatrix}-1/2\\-1/2\\1/2\\1/2\end{bmatrix}, \begin{bmatrix}-1/2\\1/2\\-1/2\\1/2\end{bmatrix} \begin{bmatrix}1/2\\-1/2\\-1/2\\1/2\end{bmatrix}$, 那么 A 的谱分解为

$$A = \begin{bmatrix}1/2 & -1/2 & -1/2 & 1/2\\1/2 & -1/2 & 1/2 & -1/2\\1/2 & 1/2 & -1/2 & -1/2\\1/2 & 1/2 & 1/2 & 1/2\end{bmatrix}\begin{bmatrix}5 & & & \\ & -5 & & \\ & & 3 & \\ & & & -3\end{bmatrix}\begin{bmatrix}1/2 & -1/2 & -1/2 & 1/2\\1/2 & -1/2 & 1/2 & -1/2\\1/2 & 1/2 & -1/2 & -1/2\\1/2 & 1/2 & 1/2 & 1/2\end{bmatrix}^{\mathrm{T}}$$

4. $\det(\lambda I_4 - A) = \lambda^3(\lambda-4)$, 所以 A 有两个不同的特征值 $4, 0$, 对应与特征值 4 的单位特征向量是 $\begin{bmatrix}1/2\\1/2\\1/2\\1/2\end{bmatrix}$. 用固定步骤解齐次方程组 $A\boldsymbol{x}=0$，可以得到 $\mathcal{N}(A)$ 的一组基是 $\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\1\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\0\\1\end{bmatrix}$, 作 Gram-Schmidt 正交化得到：$\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}-1/2\\-1/2\\1\\0\end{bmatrix}, \begin{bmatrix}-1/3\\-1/3\\-1/3\\1\end{bmatrix}$. 再单位化就得到: $\begin{bmatrix}-1/\sqrt{2}\\1/\sqrt{2}\\0\\0\end{bmatrix}, \begin{bmatrix}-1/\sqrt{6}\\-1/\sqrt{6}\\2/\sqrt{6}\\0\end{bmatrix}, \begin{bmatrix}-\sqrt{3}/6\\-\sqrt{3}/6\\-\sqrt{3}/6\\\sqrt{3}/2\end{bmatrix}$。那么 A 的谱分解为

$$A = \begin{bmatrix}1/2 & -1/\sqrt{2} & -1/\sqrt{6} & -\sqrt{3}/6\\1/2 & 1/\sqrt{2} & -1/\sqrt{6} & -\sqrt{3}/6\\1/2 & 0 & 2/\sqrt{6} & -\sqrt{3}/6\\1/2 & 0 & 0 & \sqrt{3}/2\end{bmatrix}\begin{bmatrix}4 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0\end{bmatrix}\begin{bmatrix}1/2 & -1/\sqrt{2} & -1/\sqrt{6} & -\sqrt{3}/6\\1/2 & 1/\sqrt{2} & -1/\sqrt{6} & -\sqrt{3}/6\\1/2 & 0 & 2/\sqrt{6} & -\sqrt{3}/6\\1/2 & 0 & 0 & \sqrt{3}/2\end{bmatrix}^{\mathrm{T}}$$

□

练习 6.1.2. 设三阶实对称矩阵 A 的特征值是 $0, 3, 3$，$\begin{bmatrix}1\\1\\1\end{bmatrix} \in \mathcal{N}(A)$, $\begin{bmatrix}-1\\1\\0\end{bmatrix} \in \mathcal{N}(3I-A)$，求 A.

证明: 由于 A 实对称，我们可以取第三个特征向量和已知的两个特征向量正交，这个特征向量属于特征值 3. 设这个特征向量为 $\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$, 那么我们可以得到方程组

$$\begin{cases} x_1+x_2+x_3 &= 0\\ -x_1+x_2 &= 0\end{cases}$$

那么我们可以取 $x_1=1,x_2=1,x_3=-2$. 把三个向量排成一个矩阵，得 $P=\begin{bmatrix}1&-1&1\\1&1&1\\1&0&-2\end{bmatrix}$ 那么

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

□

练习 6.1.3. 设实对称矩阵 $A=\begin{bmatrix}a&1&1&-1\\1&a&-1&1\\1&-1&a&1\\-1&1&1&a\end{bmatrix}$，有一个单特征值 -3，求 a 的值和 A 的谱分解.

证明: 我们直接计算特征多项式 $\det(\lambda I_4-A)=(\lambda-(a+1))^3(\lambda-(a-3))$; 所以 A 的单特征值是 $a-3$. 根据题意：$a-3=-3$, 所以 $a=0$. □

练习 6.1.4. 设三阶实对称矩阵 A 的各行元素之和均为 3，$\boldsymbol{a}_1=\begin{bmatrix}-1\\2\\-1\end{bmatrix},\boldsymbol{a}_2=\begin{bmatrix}0\\-1\\1\end{bmatrix}\in\mathcal{N}(A)$，求 A 及其谱分解.

证明: 设 $\boldsymbol{a}_3=\begin{bmatrix}1\\1\\1\end{bmatrix}$, 那么由题意 $A\boldsymbol{a}_3=3\boldsymbol{a}_3$. 令 $P=\begin{bmatrix}-1&0&1\\2&-1&1\\-1&1&1\end{bmatrix}$ 那么

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

□

练习 6.1.5. 求下列实对称矩阵 A 的谱分解：

1. A 满足 $A^3=0$.
2. $A=a_1\boldsymbol{x}_1\boldsymbol{x}_1^{\mathrm{T}}+a_2\boldsymbol{x}_2\boldsymbol{x}_2^{\mathrm{T}}$，其中 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 是 $\mathbb{R}^2$ 的一组标准正交基，a_1,a_2 为实数.
3. $A=\begin{bmatrix}0&M\\M&0\end{bmatrix}$，其中 M 是 n 阶对称矩阵，有谱分解 $M=Q\Lambda Q^{\mathrm{T}}$.

证明: 1. 因为 A 只有 0 特征值，所以 A 是零矩阵。

2. 令 $P=\begin{bmatrix}\boldsymbol{x}_1&\boldsymbol{x}_2\end{bmatrix}$, 那么 P 是正交矩阵，且 $A=P\begin{bmatrix}a_1&\\&a_2\end{bmatrix}P^{\mathrm{T}}$.

3. 设 $Q=\begin{bmatrix}\boldsymbol{v}_1&\cdots&\boldsymbol{v}_n\end{bmatrix},\Lambda=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$, 那么 $M\boldsymbol{v}_i=\lambda_i$. 令 $\boldsymbol{w}_i\in\mathbb{R}^{2n}$ 如下：

$$\boldsymbol{w}_i=\begin{bmatrix}\boldsymbol{v}_i\\\boldsymbol{v}_i\end{bmatrix},1\le i\le n;\quad \boldsymbol{w}_i=\begin{bmatrix}\boldsymbol{v}_{i-n}\\-\boldsymbol{v}_{i-n}\end{bmatrix},n+1\le i\le 2n.$$

令 $\tilde{Q}=\begin{bmatrix}\boldsymbol{w}_1 & \cdots & \boldsymbol{w}_{2n}\end{bmatrix}=\begin{bmatrix}Q & Q\\ Q & -Q\end{bmatrix}=\begin{bmatrix}I_n & I_n\\ I_n & -I_n\end{bmatrix}\begin{bmatrix}Q & \\ & Q\end{bmatrix}$, 于是 $\tilde{Q}$ 可逆，且

$$\begin{aligned}A\tilde{Q} &= \begin{bmatrix}\lambda_1\boldsymbol{w}_1 & \cdots & \lambda_n\boldsymbol{w}_n & -\lambda_1\boldsymbol{w}_{n+1} & \cdots & -\lambda_n\boldsymbol{w}_{2n}\end{bmatrix}\\ &= \tilde{Q}\mathrm{diag}(\lambda_1,\cdots,\lambda_n,-\lambda_1,\cdots,\lambda_n)\end{aligned}$$

所以 A 的谱分解为 $A=\tilde{Q}\begin{bmatrix}\Lambda & \\ & -\Lambda\end{bmatrix}\tilde{Q}^{-1}$.

□

练习 6.1.6. 设矩阵 $A=\begin{bmatrix}2 & b\\ 1 & 0\end{bmatrix}$. 计算满足下列条件的 b 的取值范围：

1. A 不可逆. *2.* A 可以正交对角化. *3.* A 不可对角化.

证明*:*
1. A 不可逆当且仅当 $\det(A)=0$, 所以 $b=0$.
2. A 可以正交对角化，当且仅当 A 实对称，所以 $b=1$.
3. A 是 2 阶方阵，那么 A 不可对角化当且仅当 A 只有唯一的特征值，且其几何重数为 1; 那么因为 $\mathrm{trace}(A)=2$, 所以 A 的特征值是 1, 且 $\dim\mathcal{N}(I_2-A)=1$. 所以 $b=-1$.

□

练习 6.1.7. 计算下列矩阵及其特征值. 哪些是对称矩阵? 哪些矩阵的特征值是 ±1? 由此看出正交相似的特殊性.

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

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

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

证明*:*
1. $\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}=\begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix}$, 是对称矩阵，特征多项式 $\det(\lambda I_2-A)=\lambda^2-\lambda-1$. 特征值不是 ±1.
2. 该矩阵相似于 $\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$, 特征值是 ±1. $\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1 & 0\\ -1 & 1\end{bmatrix}=\begin{bmatrix}1 & 0\\ 2 & -1\end{bmatrix}$, 不是对称矩阵。
3. 该矩阵相似于 $\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$, 特征值是 ±1. $\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1 & 0\\ -1 & 1\end{bmatrix}=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}$, 是对称矩阵。

□

练习 6.1.8. 证明命题 *6.1.6*：实方阵的正交相似关系是等价关系。

证明*:*
- $A=I_nAI_n^{\mathrm{T}}$, 所以 A 正交相似于自身。
- 若 A 正交相似于 B, 即存在正交矩阵 P 使得 $A=PBP^{\mathrm{T}}$, 那么 $B=P^{\mathrm{T}}A(P^{\mathrm{T}})^{\mathrm{T}}$, 因为 P^{T} 也是正交矩阵，所以 B 正交相似于 A.

- 若 A 正交相似于 B, B 正交相似于 C, 即存在正交矩阵 P,Q 使得 $A=PBP^{\mathrm{T}},B=QCQ^{\mathrm{T}}$, 那么 $A=(PQ)C(PQ)^{\mathrm{T}}$, 所以 A 正交相似于 C.

这就证明了正交相似关系是等价关系 □

练习 6.1.9. 证明，两个实对称矩阵正交相似，当且仅当它们具有相同的特征多项式.

证明:

- 相似的矩阵有相同的特征多项式，所以两个正交相似的实对称矩阵也有相同的特征多项式。
- 假设 A,B 是两个实对称矩阵，且有相同的特征多项式，那么他们有相同的特征值 $\lambda_1,\cdots,\lambda_n$; 于是存在正交矩阵 P,Q 使得 $A=P\Lambda P^{\mathrm{T}},B=Q\Lambda Q^{\mathrm{T}}$, 其中 $\Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$. 那么 $A=(PQ^{-1})B(PQ^{-1})^{-1}$. 正交矩阵的乘积还是正交矩阵，所以 A,B 正交相似。

□

练习 6.1.10. 设实对称矩阵 A 满足 $A^5=I_n$，证明 $A=I_n$.

证明: 设 $(\lambda,\boldsymbol{v})$ 是 A 的特征对，那么 $\lambda\in\mathbb{R}$. $A\boldsymbol{x}=\boldsymbol{x}$, 那么 $\boldsymbol{x}=A^5\boldsymbol{x}=\lambda^5\boldsymbol{x}$, 从而 $\lambda^5=1$, 所以 $\lambda=1$. 这就是说 A 的特征值全是 1, 那么 A 正交相似于单位阵，因此 $A=I_n$. □

练习 6.1.11. 设 $A_1,A_2,\ldots,A_m$ 是 m 个两两可交换的实对称矩阵，证明它们可以同时正交对角化.

证明: 我们对阶数 n 做归纳。$n=1$ 时，命题是平凡的。我们假设命题对任意阶数小于等于 n 的方阵都成立，现在考虑 $n+1$ 阶方阵。

若这些 $A_1,\cdots,A_m$ 中任意一个方阵都只有唯一的特征值，那么每个方阵都是数量矩阵，即已经都是对角矩阵，结论是显然的。假设存在某个方阵的特征值不唯一，不妨假设该方阵为 A_1, 取它的一个特征值 λ_1, 那么存在正交矩阵 P, 使得 $A=P\begin{bmatrix}\lambda_1 I_r & \\ & \Lambda\end{bmatrix}P^{\mathrm{T}}$, 其中 Λ 是对角矩阵，且对角元都和 λ_1 不相同。$A_1,\cdots,A_m$ 两两可交换，当且仅当 $\begin{bmatrix}\lambda_1 I_r & \\ & \Lambda\end{bmatrix}, B_2=P^{\mathrm{T}}A_2P,\cdots,B_m=P^{\mathrm{T}}A_mP$ 两两可交换。我们把这些 B_j 分块：$B_i=\begin{bmatrix}C_i & D_i\\ X_i & Y_i\end{bmatrix}$, 那么 $\begin{bmatrix}\lambda_1 I_r & \\ & \Lambda\end{bmatrix}B_i=B_i\begin{bmatrix}\lambda_1 I_r & \\ & \Lambda\end{bmatrix}$ 推出

$$\lambda_1 D_i=D_i\Lambda,\lambda_1 X_i=\Lambda X_i,\lambda_1 Y_i=Y_i\Lambda.$$

我们对 D_i, 考虑它的列，那么 λD_i 和 ΛD_i 都是对 D_i 作列倍乘变换。由于 Λ 是对角矩阵的对角元都和 λ 不相同，所以 $D_i=0$; 同理考虑 X_i 的行，那么 $X_i=0$.

那么 $A_1,\cdots,A_m$ 两两可交换，当且仅当 $\lambda_1 I_r,C_2,\cdots,C_m$ 两两可交换且 $\Lambda,Y_1,\cdots,Y_m$ 两两可交换。根据归纳假设，存在正交矩阵 Q_1,Q_2, 使得 $Q_1^{\mathrm{T}}(\lambda_1 I_r)Q_1,Q_1^{\mathrm{T}}C_2Q_1,\cdots,Q_1^{\mathrm{T}}C_mQ_1$, $Q_2^{\mathrm{T}}\Lambda Q_2,Q_2^{\mathrm{T}}Y_1Q_2,\cdots,Q_2^{\mathrm{T}}Y_mQ_2$ 同时是对角矩阵。那么令 $Q=P\begin{bmatrix}Q_1 & \\ & Q_2\end{bmatrix}$, 则 Q 是正交矩阵，且 $Q^{\mathrm{T}}A_1Q,\cdots,Q^{\mathrm{T}}A_mQ$ 都是对角矩阵。这就完成了归纳。

□

练习 6.1.12. 当 n 充分大时，矩阵 $A = \begin{bmatrix} 1 & 10^{-n} \\ 0 & 1+10^{-n} \end{bmatrix}$ 非常接近对称矩阵. 计算其谱分解，并求两个线性无关的特征向量的夹角.

证明: A 有两个不同的特征值 $1, 1+10^{-n}$, $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 是属于特征值 1 的特征向量，$\boldsymbol{v}_2 = \begin{bmatrix} 1 & 1 \end{bmatrix}$ 是属于特征值 $1+10^{-n}$ 的特征向量，因为我们有谱分解：

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & 1+10^{-n} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1}.$$

$\boldsymbol{v}_1, \boldsymbol{v}_2$ 的夹角记为 θ, 则 $\cos\theta = \dfrac{\boldsymbol{v}_1 \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_1\| \, \|\boldsymbol{v}_2\|} = \dfrac{\sqrt{2}}{2}$, 所以 $\theta = \dfrac{\pi}{4}$. $\square$

练习 6.1.13. 设 λ_1 是实对称矩阵 A 的最大的特征值. 证明 A 的左上角元素 $a_{11} \le \lambda_1$.

证明: 我们需要注意到，最大特征值 λ_1 满足：

$$\lambda_1 = \sup_{\|\boldsymbol{x}\|=1} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}.$$

因为 $a_{11} = \boldsymbol{e}_1^{\mathrm{T}} A \boldsymbol{e}_1$, 所以 $a_{11} \le \lambda_1$. $\square$

练习 6.1.14. 将如下函数表示成对称矩阵的 *Rayleigh* 商，并通过 *Rayleigh* 商求表达式的最大值和最小值.

1. $\dfrac{3x^2+2xy+3y^2}{x^2+y^2}$.　　*2.* $\dfrac{(x+4y)^2}{x^2+y^2}$.

证明:

1. 令 $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, 那么 $\dfrac{3x^2+2xy+3y^2}{x^2+y^2} = \dfrac{\boldsymbol{v}^{\mathrm{T}} A \boldsymbol{v}}{\boldsymbol{v}^{\mathrm{T}} \boldsymbol{v}}$, 那么表达式的最大最小值分别是 A 的最大和最小特征值，分别是 $4, 2$.
2. $\dfrac{(x+4y)^2}{x^2+y^2} = \dfrac{x^2+8xy+16y^2}{x^2+y^2}$. 令 $A = \begin{bmatrix} 1 & 4 \\ 4 & 16 \end{bmatrix}$, 那么 $\dfrac{(x+4y)^2}{x^2+y^2} = \dfrac{\boldsymbol{v}^{\mathrm{T}} A \boldsymbol{v}}{\boldsymbol{v}^{\mathrm{T}} \boldsymbol{v}}$, 那么表达式的最大最小值分别是 A 的最大和最小特征值，分别是 $17, 0$.

$\square$

练习 6.1.15 (非对称矩阵的 Rayleigh 商).

1. 如下对实矩阵的特征值都是实数的证明，哪里有问题？
设 $A\boldsymbol{x} = \lambda \boldsymbol{x}, \boldsymbol{x} \ne 0$，于是 $\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} = \lambda \boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}, \lambda = \dfrac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}$. 由于分子、分母都是实数，特征值 λ 也是实数.

2. 设 $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$，计算 $\max\limits_{\boldsymbol{x} \ne 0} \dfrac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}$.

3. 设 $A = \begin{bmatrix} 2 & 10 \\ -2 & -2 \end{bmatrix}$，计算 $\max\limits_{\boldsymbol{x} \ne 0} \dfrac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}$.

4. 对任意实矩阵 A 与非零实向量 $\boldsymbol{x}$, 证明 $\dfrac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}} = \dfrac{\boldsymbol{x}^{\mathrm{T}} A^{\mathrm{T}} \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}} = \dfrac{\boldsymbol{x}^{\mathrm{T}} B \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}$, 其中 $B = \dfrac{A + A^{\mathrm{T}}}{2}$ 是对称矩阵.

证明: 1. 我们并不知道 $\dfrac{\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}$ 分子分母都是实数。事实上，如果对于实矩阵我们能证明有实特征向量，那么该向量属于的特征值是实数。

2. $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=\begin{bmatrix}x & y\end{bmatrix}\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}y & -x\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=0$, 所以所求最大值是 0.

3. 若 $\boldsymbol{x}=\begin{bmatrix}x\\ y\end{bmatrix}$, 则 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=2x^2+8xy-2y^2$. 令 $B=\begin{bmatrix}2 & 4\\ 4 & -2\end{bmatrix}$, 则 $\max_{\boldsymbol{x}\neq 0}\dfrac{\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}$ 的最大最小值分别是 B 的最大和最小特征值，分别是 $\sqrt{20}$ 和 $-\sqrt{20}$

4. $(\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x})^{\mathrm{T}}=\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}(\boldsymbol{x}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{x}\in\mathbb{R}$, 那么要证的等式是明显的。

□

练习 6.1.16. 设实对称矩阵 $S=\begin{bmatrix}0 & A\\ A^{\mathrm{T}} & 0\end{bmatrix}$.

1. 证明，$S\boldsymbol{x}=\lambda\boldsymbol{x}$，当且仅当 $\boldsymbol{x}=\begin{bmatrix}\boldsymbol{y}\\ \boldsymbol{z}\end{bmatrix}$，满足 $A\boldsymbol{z}=\lambda\boldsymbol{y}, A^{\mathrm{T}}\boldsymbol{y}=\lambda\boldsymbol{z}$.

2. 证明，如果 λ 是 S 的特征值，则 $-\lambda$ 也是 S 的特征值.

3. 证明，如果 $\lambda\neq 0$ 是 S 的特征值，则 λ^2 是 $A^{\mathrm{T}}A$ 的特征值，也是 AA^{T} 的特征值.

4. 证明，AA^{T} 和 $A^{\mathrm{T}}A$ 的非零特征值相同，且有相同的重数.

5. 分别取 $A=I_2$ 或 $A=\begin{bmatrix}1\\ 1\end{bmatrix}$，求对应 S 的谱分解.

证明: 1. $S\boldsymbol{x}=\lambda\boldsymbol{x}\Leftrightarrow\begin{bmatrix}0 & A\\ A^{\mathrm{T}} & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{y}\\ \boldsymbol{z}\end{bmatrix}=\lambda\begin{bmatrix}\boldsymbol{y}\\ \boldsymbol{z}\end{bmatrix}\Leftrightarrow\begin{bmatrix}A\boldsymbol{z}\\ A^{\mathrm{T}}\boldsymbol{y}\end{bmatrix}=\lambda\begin{bmatrix}\boldsymbol{y}\\ \boldsymbol{z}\end{bmatrix}$, 所以 $S\boldsymbol{x}=\lambda\boldsymbol{x}$，当且仅当 $A\boldsymbol{z}=\lambda\boldsymbol{y}, A^{\mathrm{T}}\boldsymbol{y}=\lambda\boldsymbol{z}$.

2. 设 $\boldsymbol{x}=\begin{bmatrix}\boldsymbol{y}\\ \boldsymbol{z}\end{bmatrix}$ 是属于 λ 的特征向量；考虑 $\boldsymbol{x}'=\begin{bmatrix}\boldsymbol{y}\\ -\boldsymbol{z}\end{bmatrix}$, 则

$$S\boldsymbol{x}'=\begin{bmatrix}-A\boldsymbol{z}\\ A^{\mathrm{T}}\boldsymbol{y}\end{bmatrix}=\begin{bmatrix}-\lambda\boldsymbol{y}\\ \lambda\boldsymbol{z}\end{bmatrix}=-\lambda\boldsymbol{x}',$$

所以 $-\lambda$ 也是 S 的特征值。

3. 设 $\boldsymbol{x}=\begin{bmatrix}\boldsymbol{y}\\ \boldsymbol{z}\end{bmatrix}$ 是属于 λ 的特征向量，那么 $\boldsymbol{y}$ 或者 $\boldsymbol{z}$ 非零；若 $\boldsymbol{y}\neq 0$, 则 $A\boldsymbol{z}=\lambda\boldsymbol{y}\neq 0$, 所以 $\boldsymbol{z}\neq 0$; 若 $\boldsymbol{z}\neq 0$, 则 $A^{\mathrm{T}}\boldsymbol{y}=\lambda\boldsymbol{z}\neq 0$, 所以 $\boldsymbol{y}\neq 0$. 那么

$$A^{\mathrm{T}}A\boldsymbol{z}=\lambda A^{\mathrm{T}}\boldsymbol{y}=\lambda^2\boldsymbol{z};\quad AA^{\mathrm{T}}\boldsymbol{y}=A(\lambda z)=\lambda^2\boldsymbol{y}$$

所以 λ^2 是 $A^{\mathrm{T}}A$ 和 AA^{T} 的特征值。

4. 设 $A\in M_{m\times n}(\mathbb{R})$, 则我们有 $\lambda^n\det(\lambda I_m-AA^{\mathrm{T}})=\lambda^m\det(\lambda I_n-A^{\mathrm{T}}A)$. 而 AA^{T} 的非零特征值是 $\lambda^n\det(\lambda I_m-AA^{\mathrm{T}})=0$ 的非零解，重数是也是该方程解的重数，这就证明了我们的命题。

5. $A^{\mathrm{T}}A=\begin{bmatrix}2\end{bmatrix}$，只有一个特征值 2, 所以 S 有两个非零特征值 $\sqrt{2},-\sqrt{2}$, 还有一个特征值是 0. 根据前题的经验，我们取 $A^{\mathrm{T}}A$ 的一个特征向量 $\boldsymbol{z}=\begin{bmatrix}\sqrt{2}\end{bmatrix}$, 令 $\boldsymbol{y}=\frac{1}{\sqrt{2}}A\boldsymbol{z}=\begin{bmatrix}1\\ 1\end{bmatrix}$,

那么 $\boldsymbol{x}_1=\begin{bmatrix}1\\1\\\sqrt{2}\end{bmatrix},\boldsymbol{x}_2=\begin{bmatrix}1\\1\\-\sqrt{2}\end{bmatrix}$ 分别是 S 的属于特征值 $\sqrt{2},-\sqrt{2}$ 的特征值。要求属于特征值 0 的特征向量，我们可以直接解方程组 $S\boldsymbol{x}=0$, 也可以求和 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 都正交的向量，可以求得 $\boldsymbol{x}_2=\begin{bmatrix}1\\-1\\0\end{bmatrix}$ 是属于特征值 0 的特征向量。所以 S 的谱分解为

$$S=\begin{bmatrix}1&1&1\\1&1&-1\\\sqrt{2}&-\sqrt{2}&0\end{bmatrix}\begin{bmatrix}\sqrt{2}&&\\&-\sqrt{2}&\\&&0\end{bmatrix}\begin{bmatrix}1&1&1\\1&1&-1\\\sqrt{2}&-\sqrt{2}&0\end{bmatrix}^{-1}$$

□

练习 6.1.17. 构造一个实方阵 A，满足 $AA^{\mathrm{T}}=A^{\mathrm{T}}A$ 但 $A\neq A^{\mathrm{T}}$，并验证 A 和 A^{T} 具有相同的特征值和特征向量. 注意，这里相同的特征向量不意味着对应的特征值相同.

附注. 事实上，对实方阵 A，如果 $AA^{\mathrm{T}}=A^{\mathrm{T}}A$，则 A 和 A^{T} 具有相同的特征值和特征向量.

证明: 例如我们可以取一个正交矩阵：$A=\begin{bmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{bmatrix},A^{\mathrm{T}}=\begin{bmatrix}\cos\theta&\sin\theta\\-\sin\theta&\cos\theta\end{bmatrix}$ A,A^{T} 有相同的特征多项式：$\lambda^2-2\cos\theta\lambda+1$, 有两个不同的特征值：$\cos\theta\pm\boldsymbol{i}\sin\theta$.

- A 的一个属于特征值 $\cos\theta+\boldsymbol{i}\sin\theta$ 的特征向量是：$\begin{bmatrix}1\\-\boldsymbol{i}\end{bmatrix}$, 一个属于特征值 $\cos\theta-\boldsymbol{i}\sin\theta$ 的特征向量是：$\begin{bmatrix}1\\\boldsymbol{i}\end{bmatrix}$
- A^{T} 的一个属于特征值 $\cos\theta+\boldsymbol{i}\sin\theta$ 的特征向量是：$\begin{bmatrix}1\\\boldsymbol{i}\end{bmatrix}$, 一个属于特征值 $\cos\theta-\boldsymbol{i}\sin\theta$ 的特征向量是：$\begin{bmatrix}1\\-\boldsymbol{i}\end{bmatrix}$

□

练习 6.1.18. 考虑下列复对称矩阵.

1. 设复对称矩阵 $A=\begin{bmatrix}\boldsymbol{i}&1\\1&-\boldsymbol{i}\end{bmatrix}$，计算 A^2，并判断 A 是否可对角化.
2. 设复对称矩阵 $F_4=\begin{bmatrix}1&1&1&1\\1&\boldsymbol{i}&-1&-\boldsymbol{i}\\1&-1&1&-1\\1&-\boldsymbol{i}&-1&\boldsymbol{i}\end{bmatrix}$，计算 F_4^2 和 $F_4\overline{F_4}^{\mathrm{T}}$；根据 F_4^2 的特征值、F_4 的迹和行列式，求 F_4 的所有特征值. 是否存在实正交矩阵 Q，使得 $Q^{\mathrm{T}}F_4Q$ 是对角阵?

   这里 F_4 是一个 4 阶**离散** *Fourier* **变换**的表示矩阵，这类矩阵在信号处理和数值积分中有重要应用.

证明: 1. $A^2=\begin{bmatrix}0&0\\0&0\end{bmatrix}$. A 不可对角化：由 $A^2=0$ 可知 A 只有零特征值；那么若 A 可对角化，则 A 也是零矩阵，矛盾。

2. $F_4^2=\begin{bmatrix}4&0&0&0\\0&0&0&4\\0&0&4&0\\0&4&0&0\end{bmatrix}, F_4\overline{F_4}^{\mathrm{T}}=\begin{bmatrix}4&0&0&0\\0&4&0&0\\0&0&4&0\\0&0&0&4\end{bmatrix}$. $\mathrm{trace}(F_4)=2+2\boldsymbol{i}$. 我们通过对 F_4 作初等行变换可以求得 $\det(F_4)=2^4\boldsymbol{i}$. F_4^2 的特征值为 4(三重)、-4（一重），因此 F_4 有三个特征值取 ±2, 一个特征值取 $\pm2\boldsymbol{i}$. 根据迹，三个实特征值的和是 2, 纯虚特征值等于 $2\boldsymbol{i}$, 所以 F_4 的特征值为 $2,2,-2,2\boldsymbol{i}$.

假设存在实正交矩阵使得 $Q^{\mathrm{T}}F_4Q$ 是对角矩阵,那么我们可以假设 $Q^{\mathrm{T}}F_4Q=\mathrm{diag}(2,2,-2,2\boldsymbol{i})$, 即 $F_4=Q\mathrm{diag}(2,2,-2,2\boldsymbol{i})Q^{\mathrm{T}}$, 那么

$$F_4-\overline{F}_4=Q\mathrm{diag}(0,0,0,4\boldsymbol{i})Q^{\mathrm{T}}=\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&4\boldsymbol{i}\end{bmatrix}$$

矛盾。

□

练习 6.1.19 (二阶差分矩阵与二阶导数). 函数 $f(x)$ 的二阶导数 $f''(x)\approx\dfrac{f(x+\delta)+f(x-\delta)-2f(x)}{\delta^2}$. 如果将一个区间等分，将函数 $f(x)$ 在等分点处的取值列成一个向量 $\boldsymbol{f}$，则 $\boldsymbol{f}$ 就是离散化的函数，计算中可以代替 $f(x)$. 而 $f(x)$ 的二阶导数就可以用 $\dfrac{1}{h^2}D\boldsymbol{f}$ 来表示，其中 h 是每个分出的小区间的长度，而 $D=\begin{bmatrix}*&1&&&&\\1&-2&1&&&\\&1&-2&\ddots&&\\&&\ddots&\ddots&1&\\&&&1&-2&1\\&&&&1&*\end{bmatrix}$，其中 “$*$” 处元素与函数在区间端点处满足的条件有关.

1. 求 $A_3=\begin{bmatrix}2&-1&0\\-1&2&-1\\0&-1&2\end{bmatrix}$ 的谱分解，并将特征向量写成 $\begin{bmatrix}\sin d\\\sin 2d\\\sin 3d\end{bmatrix}$ 的形式.

   附注. 这是离散化的正弦函数.

2. 求 $B_4=\begin{bmatrix}1&-1&0&0\\-1&2&-1&0\\0&-1&2&-1\\0&0&-1&1\end{bmatrix}$ 的谱分解，并将特征向量写成 $\begin{bmatrix}\cos d\\\cos 3d\\\cos 5d\\\cos 7d\end{bmatrix}$ 的形式.

   附注. 这是离散化的余弦函数.

注意，$\dfrac{\mathrm{d}^2}{\mathrm{d}x^2}\sin kx=-k^2\sin kx,\dfrac{\mathrm{d}^2}{\mathrm{d}x^2}\cos kx=-k^2\cos kx$. 对离散化的情形，也有类似结论. 这一点将在第 *8* 章中进一步地讨论. 在 *Fourier* 分析中，将函数写成正弦函数和余弦函数的和有极大的好处；而在计算上，将向量写成离散正弦或离散余弦的线性组合，也有类似的好处. 例如，图片的 *JPEG* 压缩就常用 B_8 谱分解产生的正交基来进行处理.

证明: 1. $\det(\lambda I_3-A)=(\lambda-2)(\lambda^2-4\lambda+2)$, 有三个不同的特征值 $2-\sqrt{2},2,2+\sqrt{2}$

- $\begin{bmatrix}\sqrt{2}/2\\1\\\sqrt{2}/2\end{bmatrix}=\begin{bmatrix}\sin(\pi/4)\\\sin(\pi/2)\\\sin(3\pi/4)\end{bmatrix}$ 是属于特征值 $2-\sqrt{2}$ 的特征向量。
- $\begin{bmatrix}1\\0\\-1\end{bmatrix}=\begin{bmatrix}\sin(\pi/2)\\\sin(\pi)\\\sin(3\pi/2)\end{bmatrix}$ 是属于特征值 2 的特征向量。
- $\begin{bmatrix}\sqrt{2}/2\\-1\\\sqrt{2}/2\end{bmatrix}=\begin{bmatrix}\sin(3\pi/4)\\\sin(3\pi/2)\\\sin(9\pi/4)\end{bmatrix}$ 是属于特征值 $2+\sqrt{2}$ 的特征向量。

令 $Q=\begin{bmatrix}\sin(\pi/4)&\sin(\pi/2)&\sin(3\pi/4)\\\sin(2\pi/4)&\sin(\pi)&\sin(3\pi/2)\\\sin(3\pi/4)&\sin(3\pi/2)&\sin(9\pi/4)\end{bmatrix}$, 则 A 的谱分解为

$$A=Q\begin{bmatrix}2-\sqrt{2}&&\\&2&\\&&2+\sqrt{2}\end{bmatrix}Q^{-1}$$

2. $\det(\lambda I_3-B)=\lambda(\lambda-2)(\lambda^2-4\lambda+2)$, 有三个不同的特征值 $0,2-\sqrt{2},2,2+\sqrt{2}$

- $\begin{bmatrix}\cos(0)\\\cos(0)\\\cos(0)\\\cos(0)\end{bmatrix}$ 是属于特征值 0 的特征向量。
- $\begin{bmatrix}\cos(\pi/8)\\\cos(3\pi/8)\\\cos(5\pi/8)\\\cos(7\pi/8)\end{bmatrix}$ 是属于特征值 $2-\sqrt{2}$ 的特征向量。
- $\begin{bmatrix}\cos(\pi/4)\\\cos(3\pi/4)\\\cos(5\pi/4)\\\cos(7\pi/4)\end{bmatrix}$ 是属于特征值 2 的特征向量。
- $\begin{bmatrix}\cos(3\pi/8)\\\cos(9\pi/8)\\\cos(15\pi/8)\\\cos(21\pi/8)\end{bmatrix}$ 是属于特征值 $2+\sqrt{2}$ 的特征向量。

令 $Q=\begin{bmatrix}\cos(0) & \cos(\pi/8) & \cos(\pi/4) & \cos(3\pi/8)\\ \cos(0) & \cos(3\pi/8) & \cos(3\pi/4) & \cos(9\pi/8)\\ \cos(0) & \cos(5\pi/8) & \cos(5\pi/4) & \cos(15\pi/8)\\ \cos(0) & \cos(7\pi/8) & \cos(7\pi/4) & \cos(21\pi/8)\end{bmatrix}$, 则 A 的谱分解为

$$A=Q\begin{bmatrix}0 & & & \\ & 2-\sqrt{2} & & \\ & & 2 & \\ & & & 2+\sqrt{2}\end{bmatrix}Q^{-1}$$

□

练习 6.1.20. 考虑正方形铁丝上的热扩散问题，设四个顶点的温度组成向量 $\boldsymbol{t}$. 在热量扩散时，如果一个点的温度低于周围点温度的平均值，则该点温度升高；反之则该点温度降低. 假设每经过一个时间单位，一个点的温度变化与周围点平均温度和该点温度之差成正比，比例系数为 k.

1. 写出矩阵 A，使得 $A\boldsymbol{t}$ 表示经过一个单位时间之后四个点的温度.
2. 令 H 为 4 阶 Hadamard 矩阵（见练习 5.3.5），计算 AH.
3. 求 A 的谱分解.
4. 求 $\lim\limits_{t\to\infty}A^t$，它是否与 k 有关？经过足够长的时间后，最终的温度分布是什么？

证明: 1. 根据题意，$\Delta\boldsymbol{t}=\begin{bmatrix}k\left(\frac{t_2+t_4}{2}-t_1\right)\\ k\left(\frac{t_1+t_3}{2}-t_2\right)\\ k\left(\frac{t_2+t_4}{2}-t_3\right)\\ k\left(\frac{t_1+t_3}{2}-t_4\right)\end{bmatrix}$, 那么

$$A\boldsymbol{t}=\boldsymbol{t}+\Delta\boldsymbol{t}=\begin{bmatrix}(1-k)t_1+\frac{k}{2}t_2+\frac{k}{2}t_4\\ \frac{k}{2}t_1+(1-k)t_2+\frac{k}{2}t_3\\ \frac{k}{2}t_2+(1-k)t_3+\frac{k}{2}t_4\\ \frac{k}{2}t_1+\frac{k}{2}t_3+(1-k)t_4\end{bmatrix}=\begin{bmatrix}1-k & \frac{k}{2} & 0 & \frac{k}{2}\\ \frac{k}{2} & 1-k & \frac{k}{2} & 0\\ 0 & \frac{k}{2} & 1-k & \frac{k}{2}\\ \frac{k}{2} & 0 & \frac{k}{2} & 1-k\end{bmatrix}\boldsymbol{t}$$

那么 $A=\begin{bmatrix}1-k & \frac{k}{2} & 0 & \frac{k}{2}\\ \frac{k}{2} & 1-k & \frac{k}{2} & 0\\ 0 & \frac{k}{2} & 1-k & \frac{k}{2}\\ \frac{k}{2} & 0 & \frac{k}{2} & 1-k\end{bmatrix}$

2. 我们直接计算

$$\begin{aligned}AH &= \begin{bmatrix}1-k & \frac{k}{2} & 0 & \frac{k}{2}\\ \frac{k}{2} & 1-k & \frac{k}{2} & 0\\ 0 & \frac{k}{2} & 1-k & \frac{k}{2}\\ \frac{k}{2} & 0 & \frac{k}{2} & 1-k\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & -1 & 1 & -1\\ 1 & 1 & -1 & -1\\ 1 & -1 & -1 & 1\end{bmatrix}\\ &= \begin{bmatrix}1 & 1-2k & 1-k & 1-k\\ 1 & -(1-2k) & 1-k & -(1-k)\\ 1 & 1-2k & -(1-k) & -(1-k)\\ 1 & -(1-2k) & -(1-k) & 1-k\end{bmatrix}\end{aligned}$$

3. 由上一题，$AH = H\begin{bmatrix}1 & & & \\ & 1-2k & & \\ & & 1-k & \\ & & & 1-k\end{bmatrix}$, 所以 A 的谱分解为

$$A = H\begin{bmatrix}1 & & & \\ & 1-2k & & \\ & & 1-k & \\ & & & 1-k\end{bmatrix}H^{-1}$$

4. 在不和外界发生热量交换，且遵循热力学第二定律的条件下，$k \in (0,1)$, 那么

$$\lim_{t\to\infty} A^t = H\lim_{t\to\infty}\begin{bmatrix}1 & & & \\ & (1-2k)^t & & \\ & & (1-k)^t & \\ & & & (1-k)^t\end{bmatrix}H^{-1} = \frac{1}{4}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{bmatrix}$$

□

练习 6.1.21. 采集 n 个人的三项数据，例如身高、体重、收入，组成向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3 \in \mathbb{R}^n$. 令 x_1, x_2, x_3 为三组数据的平均值，$\mathbf{1} = \begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}$. 定义向量 $\boldsymbol{a}_i, \boldsymbol{a}_j$ 的协方差为 $\operatorname{cov}(\boldsymbol{a}_i, \boldsymbol{a}_j) = \frac{1}{n}(\boldsymbol{a}_i - x_i\mathbf{1})^{\mathrm{T}}(\boldsymbol{a}_j - x_j\mathbf{1})$. 称 $C = \begin{bmatrix}\operatorname{cov}(\boldsymbol{a}_i, \boldsymbol{a}_j)\end{bmatrix}_{3\times 3}$ 为数据的协方差矩阵.

1. 令 $A = \begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3\end{bmatrix}$, U 为所有元素都是 1 的 $n\times 3$ 矩阵，求常数 k, 使得 $kUU^{\mathrm{T}}A = \begin{bmatrix}x_1\mathbf{1} & x_2\mathbf{1} & x_3\mathbf{1}\end{bmatrix}$.

2. 用 A, U 来表示 C，并证明 C 是对称矩阵.

3. 任取 $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^3$，证明向量 $A\boldsymbol{y}_1, A\boldsymbol{y}_2$ 的协方差是 $\boldsymbol{y}_1^{\mathrm{T}}C\boldsymbol{y}_2$.

4. 设 C 有谱分解 $C = Q\Lambda Q^{\mathrm{T}}$，考虑 $\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3\end{bmatrix}Q$ 的三个列向量对应的三组数据，证明这三组数据两两协方差为零.

协方差为零的两组数据称为不相关的数据，利用谱分解从原本相关的数据中得到不相关的数据的线性组合，是统计学中很重要的一个方法.

证明: 根据题意，$UU^{\mathrm{T}} = 3\cdot\mathbf{1}\mathbf{1}^{\mathrm{T}}, nx_1 = \mathbf{1}^{\mathrm{T}}\boldsymbol{a}_1, nx_2 = \mathbf{1}^{\mathrm{T}}\boldsymbol{a}_2, nx_3 = \mathbf{1}^{\mathrm{T}}\boldsymbol{a}_3$.

1. $UU^{\mathrm{T}}A = 3\cdot\mathbf{1}\mathbf{1}^{\mathrm{T}}A = 3\cdot\mathbf{1}\begin{bmatrix}\mathbf{1}^{\mathrm{T}}\boldsymbol{a}_1 & \mathbf{1}^{\mathrm{T}}\boldsymbol{a}_2 & \mathbf{1}^{\mathrm{T}}\boldsymbol{a}_3\end{bmatrix} = 3n\begin{bmatrix}x_1\mathbf{1} & x_2\mathbf{1} & x_3\mathbf{1}\end{bmatrix}$, 所以 $k = \frac{1}{3n}$.
2. 根据题意，令 $B = A - \begin{bmatrix}x_1\mathbf{1} & x_2\mathbf{1} & x_3\mathbf{1}\end{bmatrix}$, 那么 $C = \frac{1}{n}B^{\mathrm{T}}B$, 所以 C 是对称矩阵。$B = A - \frac{1}{3n}UU^{\mathrm{T}}A$, 所以

$$C = \frac{1}{n}A^{\mathrm{T}}\left(I_n - \frac{1}{3n}UU^{\mathrm{T}}\right)^2 A = \frac{1}{n}A^{\mathrm{T}}\left(I_n - \frac{1}{3n}UU^{\mathrm{T}}\right)A$$

3. 将 $nx_i = \boldsymbol{a}_i^{\mathrm{T}}\mathbf{1} = \mathbf{1}^{\mathrm{T}}\boldsymbol{a}_i$ 代入 $\operatorname{cov}(\boldsymbol{a}_i, \boldsymbol{a}_j)$ 可知

$$\operatorname{cov}(\boldsymbol{a}_i, \boldsymbol{a}_j) = \frac{1}{n}\boldsymbol{a}^T\left(I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^{\mathrm{T}}\right)\boldsymbol{a}_j,$$

所以

$$\operatorname{cov}(A\boldsymbol{y}_i, A\boldsymbol{y}_j) = \frac{1}{n}\boldsymbol{y}^{\mathrm{T}}A^T\left(I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^{\mathrm{T}}\right)A\boldsymbol{y}_j = \boldsymbol{y}_i^{\mathrm{T}}C\boldsymbol{y}_j$$

4. 设 $Q = \begin{bmatrix}\boldsymbol{y}_1 & \boldsymbol{y}_2 & \boldsymbol{y}_3\end{bmatrix}$, 则 $\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3$ 是两两正交的单位向量。根据题意，我们需要证明的是 $\operatorname{cov}(A\boldsymbol{y}_i, A\boldsymbol{y}_j) = 0, \forall 1 \le i, j \le 3, i \le j$. 由上一小题，

$$\operatorname{cov}(A\boldsymbol{y}_i, A\boldsymbol{y}_j) = \boldsymbol{y}_i^{\mathrm{T}}C\boldsymbol{y}_j = \boldsymbol{y}_i^{\mathrm{T}}Q\Lambda Q^{\mathrm{T}}\boldsymbol{y}_j = \boldsymbol{e}_i^{\mathrm{T}}\Lambda\boldsymbol{e}_j,$$

$\boldsymbol{e}_i^{\mathrm{T}}\Lambda\boldsymbol{e}_j$ 是 Λ 的第 (i,j) 分量，因为 Λ 是对角矩阵，$i \ne j$ 时，Λ 的第 (i,j) 分量等于 0.

□

6.2 正定矩阵

练习 6.2.1. 判断下列矩阵是否正定：

1. $\begin{bmatrix}5 & 6\\ 6 & 7\end{bmatrix}$. *2.* $\begin{bmatrix}-1 & -2\\ -2 & -5\end{bmatrix}$. *3.* $\begin{bmatrix}1 & 10\\ 10 & 100\end{bmatrix}$. *4.* $\begin{bmatrix}1 & 10\\ 10 & 101\end{bmatrix}$.
5. $\begin{bmatrix}1 & 1 & 1\\ 1 & 0 & 2\\ 1 & 2 & 5\end{bmatrix}$. *6.* $\begin{bmatrix}1 & 1 & 1\\ 1 & -1 & 2\\ 1 & 2 & 5\end{bmatrix}$.

证明:
1. 不正定，因为行列式小于 0;
2. 不正定，因为 1 阶顺序主子式小于 0;
3. 不正定，因为行列式等于 0.
4. 正定，因为所欲偶顺序主子式都大于 0;
5. 不正定，因为 2 阶顺序主子式小于 0;
6. 不正定，因为 2 阶顺序主子式小于 0.

□

练习 6.2.2. 考虑实矩阵 $S = \begin{bmatrix} 1 & b \\ b & 1 \end{bmatrix}$. 求使得下列条件成立的 b 的取值范围：

1. S 不正定但是半正定. *2.* S 不定. *3.* S 半负定.

证明*:* 因为 S 的 1 阶顺序主子式为 1, 那么

1. S 不正定但是半正定，则行列式为 0, 所以 $b = \pm 1$.
2. S 不定则行列式小于 0, 那么 $b > 1$ 或者 $b < -1$.
3. 不存在使得 S 半负定的 b.

□

练习 6.2.3. 下列实矩阵中未知元素满足什么条件时，矩阵正定？半正定？

1. $\begin{bmatrix} 1 & b \\ b & 9 \end{bmatrix}$. *2.* $\begin{bmatrix} 2 & 4 \\ 4 & c \end{bmatrix}$. *3.* $\begin{bmatrix} c & b \\ b & c \end{bmatrix}$. *4.* $\begin{bmatrix} c & 1 & 1 \\ 1 & c & 1 \\ 1 & 1 & c \end{bmatrix}$.

5. $\begin{bmatrix} 1 & 2 & 3 \\ 2 & d & 4 \\ 3 & 4 & 5 \end{bmatrix}$. *6.* $\begin{bmatrix} s & -4 & -4 \\ -4 & s & -4 \\ -4 & -4 & s \end{bmatrix}$. *7.* $\begin{bmatrix} t & 3 & 0 \\ 3 & t & 4 \\ 0 & 4 & t \end{bmatrix}$. *8.* $\begin{bmatrix} a & a & a \\ a & a+c & a-c \\ a & a-c & a+c \end{bmatrix}$.

我们可以对矩阵作 LDL^{T}，或者计算矩阵的特征值来判断。

证明*:*

1. $\begin{bmatrix} 1 & b \\ b & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & 9-b^2 \end{bmatrix} \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$, 因此 $-3 < b < 3$ 时正定，$-3 \le b \le b$ 时半正定。
2. $\begin{bmatrix} 2 & 4 \\ 4 & c \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & \\ & c-8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, 因此 $c > 8$ 时正定，$c \ge 8$ 时半正定。
3. 若 $c = 0$, 则行列式为 $-b^2$, 所以只有 $b = 0$ 时半正定。若 $c \ne 0$, 则

$$\begin{bmatrix} c & b \\ b & c \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{b}{c} & 1 \end{bmatrix} \begin{bmatrix} c & 0 \\ 0 & c - \frac{b^2}{c} \end{bmatrix} \begin{bmatrix} 1 & \frac{b}{c} \\ 0 & 1 \end{bmatrix}$$

正定当且仅当 $c > 0, c - \frac{b^2}{c} > 0$, 即 $c > |b|$. 半正定当且仅当 $c \ge 0, c - \frac{b^2}{c} \ge 0$, 即 $c > 0$ 且 $c \ge |b|$.

4. $\det\left(\lambda I_3 - \begin{bmatrix} c & 1 & 1 \\ 1 & c & 1 \\ 1 & 1 & c \end{bmatrix}\right) = (\lambda - c + 1)^2(\lambda - c - 2)$ 因此矩阵有两个不等的特征值 $c - 1, c + 2$. 所以正等当且仅当 $c > 1$，半正定当且仅当 $c \ge 1$.
5. 半正定则顺序主子式都大于等于 0, 那么

$$d - 4 \ge 0, -4d + 8 \ge 0,$$

无解，因此矩阵不可能半正定，因此也不可能正定。

6. $\det\left(\lambda I_3-\begin{bmatrix}s&-4&-4\\-4&s&-4\\-4&-4&s\end{bmatrix}\right)=(\lambda-s-4)^2(\lambda-s+8)$ 因此矩阵有两个不等的特征值 $s+4,s-8$. 所以正定当且仅当 $s>8$，半正定当且仅当 $s\geq 8$.

7. $\det\left(\lambda I_3-\begin{bmatrix}t&3&4\\3&t&4\\0&4&t\end{bmatrix}\right)=(\lambda-t)(\lambda-t+5)(\lambda-t-5)$ 因此矩阵有两个不等的特征值 $t,t+5,t-5$. 所以正等当且仅当 $t>5$，半正定当且仅当 $t\geq 5$.

8. $$\begin{bmatrix}a&a&a\\a&a+c&a-c\\a&a-c&a+c\end{bmatrix}=\begin{bmatrix}1&0&0\\1&1&0\\0&-1&1\end{bmatrix}\begin{bmatrix}a&&\\&c&\\&&0\end{bmatrix}\begin{bmatrix}1&0&0\\1&1&0\\0&-1&1\end{bmatrix}^{\mathrm{T}}$$

所以矩阵不正定；当且仅当 $a,c\geq 0$ 时半正定。

□

练习 6.2.4. 设 A 对称正定，B 是实矩阵.

1. 证明，对任意整数，A^k 也正定.
2. 若存在正整数 r，使得 $A^rB=BA^r$，证明 $AB=BA$.

证明: 1. A 对称正定,则 A 的特征值全大于 0, 因此存在正交矩阵 P 使得 $A=P\mathrm{diag}(\lambda_1,\cdots,\lambda_n)P^{\mathrm{T}}$, 其中 $\lambda_1,\cdots,\lambda_n>0$, 那么 $A^k=P\mathrm{diag}(\lambda_1^k,\cdots,\lambda_n^k)P^T$, 即 A 正交相似于对角线都是正数的对角矩阵，所以 A^k 正定。

2. 设 $\lambda_1,\cdots,\lambda_k$ 是 A 的不同的特征值。A 正定，则 $\lambda_1,\cdots,\lambda_k>0$, 那么对任意正整数 r, $\lambda_1^r,\cdots,\lambda_k^r$ 也互不相同。那么存在正交矩阵 Q，使得 $A=Q\mathrm{diag}(\lambda_1I_{n_1},\cdots,\lambda_kI_{n_k})Q^{\mathrm{T}}$, 那么 $A^r=Q\mathrm{diag}(\lambda_1^rI_{n_1},\cdots,\lambda_k^rI_{n_k})Q^T$. 那么

$$\mathrm{diag}(\lambda_1^rI_{n_1},\cdots,\lambda_k^rI_{n_k})Q^{\mathrm{T}}BQ=Q^{\mathrm{T}}BQ\mathrm{diag}(\lambda_1^rI_{n_1},\cdots,\lambda_k^rI_{n_k}). \qquad *$$

我们对 $Q^{\mathrm{T}}BQ$ 作分块：$Q^{\mathrm{T}}BQ=\Big[B_{ij}\Big]_{1\leq i,j\leq k}$, 其中 $B_{ij}\in M_{n_i\times n_j}(\mathbb{R})$. 那么 $(*)$ 式推出 $\lambda_i^rB_{ij}=\lambda_j^rB_{ij}$, 所以当 $i\neq j$ 时 $B_{ij}=0$, 因此 $Q^{\mathrm{T}}BQ$ 是分块对角矩阵，那么

$$\mathrm{diag}(\lambda_1I_{n_1},\cdots,\lambda_rI_{n_r})Q^{\mathrm{T}}BQ=Q^{\mathrm{T}}BQ\mathrm{diag}(\lambda_1I_{n_1},\cdots,\lambda_rI_{n_r}),$$

从而 $AB=BA$.

□

练习 6.2.5. 对 n 阶实对称矩阵 A，证明，当实数 t 充分大时，tI_n+A 正定.

证明: 因为 tI_n+A 的所有主子式都是关于 t 的首项系数为 1 的多项式，那么当 t 充分大时，这些多项式都大于 0，即当 t 充分大时，所有的主子式都大于 0。所以 tI_n+A 正定。 □

练习 6.2.6. 对 n 阶实对称矩阵 A，证明，存在正实数 c，使得对任意 $\boldsymbol{x}\in\mathbb{R}^n$，都有 $\left|\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}\right|\leq c\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$.

证明: 设 $\lambda_1, \cdots, \lambda_n$ 是 A 的特征值，满足 $|\lambda_1| \geq \cdots \geq |\lambda_n|$, 且 Q 是正交矩阵使得 $A = Q^{\mathrm{T}}\Lambda Q$, 其中 $\Lambda = \operatorname{diag}(\lambda_1, \cdots, \lambda_r)$；那么对 $\boldsymbol{x} \neq 0$, 我们有

$$\max_{\boldsymbol{x}\neq 0}\left|\frac{\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}\right| = \max_{\|\boldsymbol{x}\|=1}\left|\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}\right| \leq \max_{\|\boldsymbol{y}\|=1}\left|\boldsymbol{y}^{\mathrm{T}}\Lambda\boldsymbol{y}\right| \leq \max_{\|\boldsymbol{y}\|=1}|\lambda_1|\,(y_1^2 + \cdots + y_n^2) = |\lambda_1|\,.$$

所以我们取 $c = |\lambda_1|$. $\square$

练习 6.2.7 (Hadamard 不等式). 给定 n 阶对称正定矩阵 A，求证：

1. 对任意 $\boldsymbol{y}$，$\det\left(\begin{bmatrix} A & \boldsymbol{y} \\ \boldsymbol{y}^{\mathrm{T}} & 0 \end{bmatrix}\right) \leq 0$；
2. 记 $A = \begin{bmatrix} a_{ij} \end{bmatrix}$，则 $\det(A) \leq a_{nn}\det(A_{n-1})$，其中 A_{n-1} 是 A 的 $n-1$ 阶顺序主子阵；
3. $\det(A) \leq a_{11}a_{22}\cdots a_{nn}$.

利用上述结论证明：如果实矩阵 $T = \begin{bmatrix} t_{ij} \end{bmatrix}$ 可逆，那么 $\det(T)^2 \leq \prod_{i=1}^{n}(t_{1i}^2 + t_{2i}^2 + \cdots + t_{ni}^2)$.

附注. 练习 4.2.28用不同方法证明了相同结论.

证明: 1. 由于 A 可逆，我们作初等行倍加：$\begin{bmatrix} A & \boldsymbol{y} \\ \boldsymbol{y}^{\mathrm{T}} & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} A & \boldsymbol{y} \\ 0 & -\boldsymbol{y}^{\mathrm{T}}A^{-1}\boldsymbol{y} \end{bmatrix}$, 所以

$$\det\left(\begin{bmatrix} A & \boldsymbol{y} \\ \boldsymbol{y}^{\mathrm{T}} & 0 \end{bmatrix}\right) = -(\boldsymbol{y}^{\mathrm{T}}A^{-1}\boldsymbol{y})\det(A).$$

由于 A 对称，则 A^{-1} 也对称；A 正定，则其特征值都大于零，那么 A^{-1} 的特征值都是 A 特征值的倒数，也大于零，所以 A^{-1} 也正定，因此

$$\det\left(\begin{bmatrix} A & \boldsymbol{y} \\ \boldsymbol{y}^{\mathrm{T}} & 0 \end{bmatrix}\right) = -(\boldsymbol{y}^{\mathrm{T}}A^{-1}\boldsymbol{y})\det(A) \leq 0.$$

2. 利用行列式的行线性性，

$$\begin{aligned}\det(A) &= \det\left(\begin{bmatrix} A_{n-1} & \boldsymbol{y} \\ \boldsymbol{y}^{\mathrm{T}} & 0 \end{bmatrix}\right) + \det\left(\begin{bmatrix} A_{n-1} & \boldsymbol{y} \\ 0 & a_{nn} \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} A_{n-1} & \boldsymbol{y} \\ \boldsymbol{y}^{\mathrm{T}} & 0 \end{bmatrix}\right) + a_{nn}\det(A_{n-1}) \\ &\leq a_{nn}\det(A_{n-1})\end{aligned}.$$

3. 对 $2 \geq i \geq n$，上一小题的证明可以推出 $\det(A_i) \leq a_{ii}\det(A_{i-1})$, 所以

$$\det(A) \leq a_{nn}\det(A_{n-1}) \leq a_{nn}a_{n-1,n-1}\det(A_{n-2}) \leq \cdots \leq a_{11}a_{22}\cdots a_{nn}.$$

如果实矩阵 T 可逆，那么考虑 $\tilde{T} = T^{\mathrm{T}}T$, 那么 $\tilde{T}$ 是对称正定矩阵，且它的第 i 个对角元是 $t_{1i}^2 + t_{2i}^2 + \cdots + t_{ni}^2$；利用第 3 小题的结论就有

$$\det(T)^2 \leq \prod_{i=1}^{n}(t_{1i}^2 + t_{2i}^2 + \cdots + t_{ni}^2).$$

$\square$

练习 6.2.8. 证明 $A=\left[\dfrac{1}{i+j}\right]_{n\times n}$ 正定.

法一：考虑 $\int_0^1 t(x_1+x_2t+\cdots+x_nt^{n-1})^2\mathrm{d}t$. 法二：证明 $\det(A)=\prod\limits_{\substack{1\le i\le n\\1\le j\le n}}\dfrac{1}{i+j}\prod\limits_{1\le i<j\le n}(i-j)^2$.

证明: 我们记 $A_n=\left[\dfrac{1}{i+j}\right]_{n\times n}$, 那么 A 的 k 阶顺序子矩阵恰好是 A_k，所以我们只需证明：对任意 n, $\det(A_n)>0$. 我们对矩阵作初等列倍加变换：先将所有列都减去最后一列：

$$\begin{bmatrix}\frac12 & \frac13 & \cdots & \frac1n & \frac1{n+1}\\ \frac13 & \frac14 & \cdots & \frac1{n+1} & \frac1{n+2}\\ \vdots & & & & \vdots\\ \frac1n & \frac1{n+1} & \cdots & \frac1{2n-2} & \frac1{2n-1}\\ \frac1{n+1} & \frac1{n+2} & \cdots & \frac1{2n-1} & \frac1{2n}\end{bmatrix}\to\begin{bmatrix}\frac{n-1}{2(n+1)} & \frac{n-2}{3(n+1)} & \cdots & \frac{1}{n(n+1)} & \frac1{n+1}\\ \frac{n-1}{3(n+2)} & \frac{n-2}{4(n+2)} & \cdots & \frac{1}{(n+1)(n+2)} & \frac1{n+2}\\ \vdots & & & & \vdots\\ \frac{n-1}{n(2n-1)} & \frac{n-2}{(n+1)(2n-1)} & \cdots & \frac{1}{(2n-2)(2n-1)} & \frac1{2n-1}\\ \frac{n-1}{(n+1)(2n)} & \frac{n-2}{(n+2)(2n)} & \cdots & \frac{1}{(2n-1)(2n)} & \frac1{2n}\end{bmatrix}$$

那么

$$\det(A_n)=\prod_{\ell=1}^{n-1}(n-\ell)\prod_{\ell=1}^{n}\frac{1}{n+\ell}\begin{vmatrix}\frac12 & \frac13 & \cdots & \frac1n & 1\\ \frac13 & \frac14 & \cdots & \frac1{n+1} & 1\\ \vdots & & & & \vdots\\ \frac1n & \frac1{n+1} & \cdots & \frac1{2n-2} & 1\\ \frac1{n+1} & \frac1{n+2} & \cdots & \frac1{2n-1} & 1\end{vmatrix}$$

我们再对 $\begin{vmatrix}\frac12 & \frac13 & \cdots & \frac1n & 1\\ \frac13 & \frac14 & \cdots & \frac1{n+1} & 1\\ \vdots & & & & \vdots\\ \frac1n & \frac1{n+1} & \cdots & \frac1{2n-2} & 1\\ \frac1{n+1} & \frac1{n+2} & \cdots & \frac1{2n-1} & 1\end{vmatrix}$ 作行倍加变换，每一行都减去最后一行：

$$\begin{vmatrix}\frac12 & \frac13 & \cdots & \frac1n & 1\\ \frac13 & \frac14 & \cdots & \frac1{n+1} & 1\\ \vdots & & & & \vdots\\ \frac1n & \frac1{n+1} & \cdots & \frac1{2n-2} & 1\\ \frac1{n+1} & \frac1{n+2} & \cdots & \frac1{2n-1} & 1\end{vmatrix}\to\begin{vmatrix}\frac{n-1}{2(n+1)} & \frac{n-1}{3(n+2)} & \cdots & \frac{n-1}{n(2n-1)} & 0\\ \frac{n-2}{3(n+1)} & \frac{n-2}{4(n+2)} & \cdots & \frac{n-2}{(n+1)(2n-1)} & 0\\ \vdots & & & & \vdots\\ \frac{1}{n(n+1)} & \frac{1}{(n+1)(n+2)} & \cdots & \frac{1}{(2n-2)(2n-1)} & 0\\ \frac1{n+1} & \frac1{n+2} & \cdots & \frac1{2n-1} & 1\end{vmatrix}$$

那么

$$\begin{vmatrix}\frac12 & \frac13 & \cdots & \frac1n & 1\\ \frac13 & \frac14 & \cdots & \frac1{n+1} & 1\\ \vdots & & & & \vdots\\ \frac1n & \frac1{n+1} & \cdots & \frac1{2n-2} & 1\\ \frac1{n+1} & \frac1{n+2} & \cdots & \frac1{2n-1} & 1\end{vmatrix}=\prod_{\ell=1}^{n-1}(n-\ell)\prod_{\ell=1}^{n-1}\frac{1}{n+\ell}\det(A_{n-1})$$

所以 $\det(A_n)=\left(\prod_{\ell=1}^{n-1}(n-\ell)\right)^2\prod_{\ell=1}^{n}\frac{1}{n+\ell}\prod_{\ell=1}^{n-1}\frac{1}{n+\ell}A_{n-1}$. 而 $\det(A_1)=\frac12,\det(A_2)=\frac1{72}$, 递归地知道 $\det(A_n)>0,\forall n$. □

练习 6.2.9. 证明 *Hilbert* 矩阵 $H_n = \left[\dfrac{1}{i+j-1}\right]_{n\times n}$ 正定.

证明: 跟上题一样，对 H_n 作相同步骤的初等列变换和初等行变换，可得关于 H_n 的递推公式。具体步骤我们不再重复。 □

练习 6.2.10. 设 A, B 是实对称矩阵，如果 $A-B$ 正定，则记作 $A \succ B$. 求证：

1. $A \succ B, B \succ C$ 可以推出 $A \succ C$.
2. $A \succ B$ 和 $B \succ A$ 不可能同时成立.
3. 对任意实对称矩阵 A，都存在实数 k_1, k_2 使得 $k_1 I_n \succ A \succ k_2 I_n$.

证明: $A-B$ 正定，当且仅当对任意 $\boldsymbol{x}\neq 0$，有 $\boldsymbol{x}^{\mathrm{T}}(A-B)\boldsymbol{x}>0$, 即 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}>\boldsymbol{x}^{\mathrm{T}}B\boldsymbol{x}>0$.

1. $A \succ B, B \succ C$ 当且仅当对任意 $\boldsymbol{x}\neq 0, \boldsymbol{x}^{T}A\boldsymbol{x}>\boldsymbol{x}^{\mathrm{T}}B\boldsymbol{x}>\boldsymbol{x}^{\mathrm{T}}C\boldsymbol{x}$, 则 $A \succ C$.
2. $A \succ B$ 且 $B \succ A$ 当且仅当对任意 $\boldsymbol{x}\neq 0, \boldsymbol{x}^{T}A\boldsymbol{x}>\boldsymbol{x}^{\mathrm{T}}B\boldsymbol{x}>\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$, 这显然不可能成立。
3. 对任意实对称矩阵，令 λ_1 是 A 最大的特征值，λ_2 是 A 最小的特征值。我们取 $k_1 > \lambda_1, k_2 < \lambda_2$, 那么 $k_1 I_n - A$ 所有特征值都大于 0，所以 $k_1 I_n - A$ 是正定矩阵，即 $k_1 I_n \succ A$. 同理：$A \succ k_2 I_n$.

□

练习 6.2.11. 举例说明，实对称矩阵 A 的所有顺序主子式都非负，但 A 并不半正定.

证明: $A = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$, 顺序主子式都是 0, 但是 $\boldsymbol{e}_2^{\mathrm{T}}A\boldsymbol{e}_2 = -1$. □

练习 6.2.12. 证明命题 *6.2.4.* 对实对称矩阵 A，以下叙述等价：

1. A 半正定；
2. A 的特征值都是非负数；
3. 存在矩阵 T，使得 $A = TT^{\mathrm{T}}$;
4. A 存在 LDL^{T} 分解，且 D 的对角线元素都是非负数。

证明$(1 \Rightarrow 2)$. 设 A 的谱分解位 $A = P\Lambda P^{\mathrm{T}}$, 其中 $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_n), P = \begin{bmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{bmatrix}$, 那么对 $i = 1, 2, \cdots, n$, $\lambda_i = \boldsymbol{v}_i^{\mathrm{T}}A\boldsymbol{v}_i \geq 0$.

$(2 \Rightarrow 3)$. 设 A 的谱分解如上, 那么我们取 $T = P\mathrm{diag}(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n})$ 即可。

$(3 \Rightarrow 1)$. $\forall \boldsymbol{x}\neq 0, \boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} = \left\|T^{\mathrm{T}}\boldsymbol{x}\right\|^2 \geq 0$，即 A 半正定。

$(4 \Rightarrow 1)$. $\forall \boldsymbol{x}$, 设 $L^{\mathrm{T}}\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $D = \mathrm{diag}(d_1, \cdots, d_n)$, 那么 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} = d_1x_1^2 + \cdots + d_nx_n^2 \geq 0$，即 A 半正定。

$(1 \Rightarrow 4)$. 我们对 A 的阶数 n 用归纳法. $n=1$ 命题是平凡的。假设命题结论对所有 n 阶实对称矩阵成立，那么对 $n+1$ 阶实对称矩阵 A, 我们对其进行分块：$A = \begin{bmatrix} B & \boldsymbol{a} \\ \boldsymbol{a}^{\mathrm{T}} & a_{n+1,n+1} \end{bmatrix}$。

对任意 $\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{x}^{\mathrm{T}}B\boldsymbol{x}=\begin{bmatrix}\boldsymbol{x}\\0\end{bmatrix}^{\mathrm{T}}A\begin{bmatrix}\boldsymbol{x}\\0\end{bmatrix}\geq 0$，所以 B 也半正定。若 B 可逆，则

$$A=\begin{bmatrix}I_n & 0\\ \boldsymbol{a}B^{-1} & 1\end{bmatrix}\begin{bmatrix}B & 0\\ 0 & a_{n+1,n+1}-\boldsymbol{a}^{\mathrm{T}}B^{-1}\boldsymbol{a}\end{bmatrix}\begin{bmatrix}I_n & 0\\ \boldsymbol{a}B^{-1} & 1\end{bmatrix}^{\mathrm{T}}$$

由归纳假设，B 存在 LDL^{T} 分解，所以

$$A=\left(\begin{bmatrix}I_n & 0\\ \boldsymbol{a}B^{-1} & 1\end{bmatrix}\begin{bmatrix}L & 0\\ 0 & 1\end{bmatrix}\right)\begin{bmatrix}D & 0\\ 0 & a_{n+1,n+1}-\boldsymbol{a}^{\mathrm{T}}B^{-1}\boldsymbol{a}\end{bmatrix}\left(\begin{bmatrix}I_n & 0\\ \boldsymbol{a}B^{-1} & 1\end{bmatrix}\begin{bmatrix}L & 0\\ 0 & 1\end{bmatrix}\right)^{\mathrm{T}}$$

是符合题意的 A 的分解。若 B 不可逆，我们断言 $\boldsymbol{a}\in\mathcal{R}(B)$, 若不然：$\boldsymbol{a}\notin\mathcal{R}(B)$, 则由练习 2.4.16 知方程组 $\begin{bmatrix}B^{\mathrm{T}}\\ \boldsymbol{a}^{\mathrm{T}}\end{bmatrix}\boldsymbol{x}=\begin{bmatrix}0\\1\end{bmatrix}$ 有解 $\boldsymbol{x}_0$, 对 $t\in\mathbb{R}$, 考虑 $f(t):=\begin{bmatrix}t\boldsymbol{x}_0\\1\end{bmatrix}^{\mathrm{T}}A\begin{bmatrix}t\boldsymbol{x}_0\\1\end{bmatrix}$. 因为 A 半正定，所以 $f(t)\geq 0,\forall t\in\mathbb{R}$. 另一方面，我们按照分块矩阵展开可得 $f(t)=\boldsymbol{x}_0^{\mathrm{T}}B\boldsymbol{x}_0+2t\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}_0+a_{n+1,n+1}=2t+a_{n+1,n+1}$, 显然 $f(t)$ 不能恒非负，矛盾。至此，我们证明了断言。于是存在 $\boldsymbol{b}\in\mathbb{R}^n$ 使得 $\boldsymbol{a}=B\boldsymbol{b}$。那么

$$A=\begin{bmatrix}I_n & 0\\ \boldsymbol{b}^{\mathrm{T}} & 1\end{bmatrix}\begin{bmatrix}B & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}I_n & 0\\ \boldsymbol{b}^{\mathrm{T}} & 1\end{bmatrix}^{\mathrm{T}}$$

由归纳假设，B 存在 LDL^{T} 分解，所以

$$A=\left(\begin{bmatrix}I_n & 0\\ \boldsymbol{b} & 1\end{bmatrix}\begin{bmatrix}L & 0\\ 0 & 1\end{bmatrix}\right)\begin{bmatrix}D & 0\\ 0 & a_{n+1,n+1}\end{bmatrix}\left(\begin{bmatrix}I_n & 0\\ \boldsymbol{b} & 1\end{bmatrix}\begin{bmatrix}L & 0\\ 0 & 1\end{bmatrix}\right)^{\mathrm{T}}$$

是符合题意的 A 的分解。

□

练习 6.2.13. 证明，实对称矩阵半正定，当且仅当它的所有主子式都非负.

法一：利用数学归纳法. 法二：考虑矩阵 A 的微小变化得到的正定矩阵 $A+\varepsilon I$.

证明: 若实对称矩阵 A 半正定，考虑 $A_\epsilon=A+\epsilon I_n(\epsilon\geq 0)$, 那么对任意 $\epsilon>0$, $\boldsymbol{x}^{\mathrm{T}}A_\epsilon\boldsymbol{x}=\boldsymbol{x}^TA\boldsymbol{x}+\epsilon\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}\geq 0+\epsilon\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}>0$, 所以 A_ϵ 是正定矩阵，那么 A_ϵ 的所有主子式都大于 0. 设 B 是 A 的一个 k-阶主子阵，那么 $B+\epsilon I_k$ 是 A_ϵ 的一个 k 阶主子阵，所以 $\det(B_\epsilon)>0$. 那么

$$\det(B)=\lim_{\epsilon\to 0}\det(B_\epsilon)\geq 0.$$

反之，当实对称矩阵 A 的所有主子式都非负时，我们通过对阶数 n 归纳法来证明 A 半正定。$n=1$ 是平凡的；假设命题结论对所有 n 阶实对称矩阵成立，那么对 $n+1$ 阶实对称矩阵 A, 我们对其进行分块：$A=\begin{bmatrix}B & \boldsymbol{x}\\ \boldsymbol{x}^{\mathrm{T}} & a_{n+1,n+1}\end{bmatrix}$。因为 $a_{n+1,n+1}$ 是 A 的一个 1 阶主子式，那么 $a_{n+1,n+1}\geq 0$

- 若 $a_{n+1,n+1}=0$, 则 $\boldsymbol{x}=0$. 若不然，存在 $x_i\neq 0$, 那么 $\begin{bmatrix}a_{ii} & x_i\\ x_i & a_{n+1,n+1}\end{bmatrix}$ 是 A 的一个 2 阶主子阵，但是它的行列式等于 $-x_i^2<0$, 矛盾。因为 B 的主子式都是 A 的主子式，所以都非负，根据归纳假设，B 半正定。所以此时 $A=\begin{bmatrix}B & 0\\ 0 & 0\end{bmatrix}$ 是半正定的。

- 若 $a_{n+1,n+1}>0$, 则 A 半正定，当且仅当

$$\begin{bmatrix} B-\boldsymbol{x}a_{n+1,n+1}^{-1}\boldsymbol{x}^{\mathrm{T}} & 0 \\ 0 & a_{n+1,n+1} \end{bmatrix} = \begin{bmatrix} I_n & -a_{n+1,n+1}^{-1}\boldsymbol{x} \\ 0 & 1 \end{bmatrix} A \begin{bmatrix} I_n & -a_{n+1,n+1}^{-1}\boldsymbol{x} \\ 0 & 1 \end{bmatrix}^{\mathrm{T}}$$

半正定。根据归纳假设，我们需要验证 $B-\boldsymbol{x}a_{n+1,n+1}^{-1}\boldsymbol{x}^{\mathrm{T}}$ 的主子式都非负。事实上，$B-\boldsymbol{x}a_{n+1,n+1}^{-1}\boldsymbol{x}^{\mathrm{T}}$ 的第 $i_1,\cdots,i_r$ 行、第 $i_1,\cdots,i_r$ 列交叉点构成的主子阵的行列式乘以 $a_{n+1,n+1}$ 等于 A 的第 $i_1,\cdots,i_r\ n+1$ 行、第 $i_1,\cdots,i_r,n+1$ 列交叉点构成的主子阵的行列式；所以 $B-\boldsymbol{x}a_{n+1,n+1}^{-1}\boldsymbol{x}^{\mathrm{T}}$ 的主子式都非负。

□

练习 6.2.14. 设 A 是实对称矩阵，证明：

1. *A 半正定，当且仅当存在实对称矩阵 B，使得 $A=B^2$.*
2. *若 A 半正定，则存在唯一的半正定实对称矩阵 B，使得 $A=B^2$.*

证明*:* A 是实对称矩阵，则存在正交矩阵 Q 和对角矩阵 $\Lambda=\mathrm{diag}(\lambda_1 I_{n_1},\cdots,\lambda_r I_{n_r})$ 使得 $A=Q\Lambda Q^{\mathrm{T}}$, 其中 $\lambda_1>\cdots>\lambda_r$ 是 A 互不相同的特征值.

1. A 半正定，则 $\lambda_i\geq 0(1\leq i\leq r)$. 我们取 $B=Q\mathrm{diag}(\sqrt{\lambda_1}I_{n_1},\cdots,\sqrt{\lambda_r}I_{n_r})Q^{\mathrm{T}}$, 那么 B 也是半正定的实对称矩阵，且 $A=B^2$. 反之，若存在实对称矩阵 B 使得 $A=B^2$，那么 A 的特征值都是 B 的特征值的平方，所以都是非负的，所以 A 是半正定的。
2. 存在性在上一小题的证明中我们已经证明了，我们只需要证明唯一性。设 B 的特征值是 $\mu_1,\cdots,\mu_n(\mu_1\geq\cdots\mu_n\geq 0)$. 因为 $A=B^2$, 那么 A 的特征值就是 $\mu_1^2,\cdots,\mu_n^2$. 根据命题 5.4.6, 我们可以取到正交矩阵 P 使得 $B=P\mathrm{diag}(\mu_1,\cdots,\mu_n)P^{\mathrm{T}}$ 且 $\mathrm{diag}(\mu_1^2,\cdots,\mu_n^2)=\Lambda$. 此时，

$$\mathrm{diag}(\mu_1,\cdots,\mu_n)=\mathrm{diag}(\sqrt{\lambda_1}I_{n_1},\cdots,\sqrt{\lambda_r}I_{n_r}).$$

并且 $A=B^2$ 即 $(Q^{\mathrm{T}}P)\Lambda=\Lambda(Q^{\mathrm{T}}P)$. 我们将 $Q^{\mathrm{T}}P$ 按照 Λ 进行相应的分块 $Q^{\mathrm{T}}P=\begin{bmatrix}X_{ij}\end{bmatrix}_{1\leq i,j\leq r}$, 那么 $\begin{bmatrix}\lambda_j X_{ij}\end{bmatrix}=\begin{bmatrix}\lambda_i X_{ij}\end{bmatrix}$，所以 $i\neq j$ 时，$X_{ij}=0$. 因为 $Q^{\mathrm{T}}P$ 也是正交矩阵，那么 $\begin{bmatrix}X_{ij}\end{bmatrix}\begin{bmatrix}X_{ij}\end{bmatrix}^{\mathrm{T}}=I_n$, 所以 $X_{ii}X_{ii}^{\mathrm{T}}=I_{n_i}$, 那么 $P=Q\mathrm{diag}(X_{11},\cdots,X_{nn})$, 那么

$$B=Q\mathrm{diag}(X_{11},\cdots,X_{nn})\mathrm{diag}(\sqrt{\lambda_1}I_{n_1},\cdots,\sqrt{\lambda_r}I_{n_r})\mathrm{diag}(X_{11},\cdots,X_{nn})^{\mathrm{T}}Q^{\mathrm{T}},$$

即 B 是唯一的，它就等于 $Q\mathrm{diag}(\sqrt{\lambda_1}I_{n_1},\cdots,\sqrt{\lambda_r}I_{n_r})Q^{\mathrm{T}}$

□

练习 6.2.15. 证明，若实对称矩阵对角占优，且对角元素全为正数，则该矩阵正定.

证明*:* 对角元素为正数的对角占优矩阵的顺序主子阵也都是对角元素为正数的对角占优矩阵，根据练习 4.3.13, 对角元素为正数的对角占优矩阵的顺序主子式都是正的，所以该矩阵正定。 □

练习 6.2.16. 证明或举出反例：

1. *如果 A 对称正定，则 A^{-1} 正定.*
2. *如果 A,B 对称正定，则 $A+B$ 正定.*

3. 如果 A, B 对称半正定，则 $A+B$ 半正定.

4. 如果 A, B 对称不定，则 $A+B$ 不定.

5. 如果 A 列满秩，B 对称正定，则 $A^{\mathrm{T}}BA$ 正定.

6. 如果 $S=A^{\mathrm{T}}A$ 且 A 有简化 QR 分解 $A=QR$，则 $S=R^{\mathrm{T}}R$ 是 S 的 *Cholesky* 分解.

7. 如果 A, B 对称正定，则 AB 的特征值都是正数.

证明:

1. 正确。A 对称正定，则它的特征值都是正的，那么 A^{-1} 的特征值也是正的，所以 A^{-1} 也是对称正定矩阵。
2. 正确。对任意 $\boldsymbol{x}\neq 0$, $\boldsymbol{x}^{\mathrm{T}}(A+B)\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}+\boldsymbol{x}^{\mathrm{T}}B\boldsymbol{x}>0$. 所以 $A+B$ 正定。
3. 正确。对任意 $\boldsymbol{x}\neq 0$, $\boldsymbol{x}^{\mathrm{T}}(A+B)\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}+\boldsymbol{x}^{\mathrm{T}}B\boldsymbol{x}\geq 0$. 所以 $A+B$ 半正定。
4. 错误。反例：$A=\begin{bmatrix}3 & 1\\ 1 & 0\end{bmatrix}, B=\begin{bmatrix}0 & 1\\ 1 & 3\end{bmatrix}$
5. 正确。对任意 $\boldsymbol{x}\neq 0$, 则 $A\boldsymbol{x}\neq 0$, 那么 $\boldsymbol{x}^{\mathrm{T}}(A^{\mathrm{T}}BA)\boldsymbol{x}=(A\boldsymbol{x})^{\mathrm{T}}B(A\boldsymbol{x})>0$.
6. 正确。这符合 Cholesky 分解的定义。
7. 正确。A, B 对称正定，则存在可逆矩阵 A_1, B_1 使得 $A=A_1^{\mathrm{T}}A_1, B=B_1^{\mathrm{T}}B_1$, 那么

$$AB=A_1^{\mathrm{T}}A_1B_1^{\mathrm{T}}B_1=A_1^{\mathrm{T}}\left((B_1A_1^{\mathrm{T}})^{\mathrm{T}}(B_1A_1^{\mathrm{T}})\right)(A_1^{\mathrm{T}})^{-1}$$

所以 AB 相似于对称正定矩阵 $A_1B_1^{\mathrm{T}}B_1A_1^{\mathrm{T}}$, 那么特征值都是正数。

□

练习 6.2.17. *证明命题 6.2.7. 方阵的合同关系是等价关系。*

证明:

1. 反身性：对任意 $A\in M_{n\times n}(\mathbb{R}), A=I_n^{\mathrm{T}}AI_n$; 这说明 A 合同于自身。
2. 对称性: 若 A 合同于 B, 则存在可逆矩阵 P 使得 $A=P^{\mathrm{T}}BP$, 那么 $B=(P^{-1})^{\mathrm{T}}AP^{-1}$, 即 B 合同于 A.
3. 传递性：若 A 合同于 B, B 合同于 C，则存在可逆矩阵 P, Q 使得 $A=P^{T}BP, B=Q^{\mathrm{T}}CQ$, 那么 $A=(PQ)^{\mathrm{T}}C(PQ)$, 即 A 合同于 C.

所以合同关系是等价关系。 □

练习 6.2.18. *考虑实对称矩阵* $S=\begin{bmatrix}A & B\\ B^{\mathrm{T}} & C\end{bmatrix}$. *证明 S 正定当且仅当 A 及其 Schur 补都正定.*

证明:

- 若 S 正定，则其顺序主子式都是正的，所以 A 正定。那么

$$S=\begin{bmatrix}I_n & 0\\ B^{\mathrm{T}}A^{-1} & I_m\end{bmatrix}\begin{bmatrix}A & \\ & C-B^{\mathrm{T}}A^{-1}B\end{bmatrix}\begin{bmatrix}I_n & 0\\ B^{\mathrm{T}}A^{-1} & I_m\end{bmatrix}^{\mathrm{T}}$$

因此 $\begin{bmatrix}A & \\ & C-B^{\mathrm{T}}A^{-1}B\end{bmatrix}$ 正定。注意到 $C-B^{\mathrm{T}}A^{-1}B$ 是对称矩阵，且它的特征值都是 S 的特征值，所以 $C-B^{\mathrm{T}}A^{-1}B$ 的特征值都是正的，所以是正定矩阵。

- 若 A 及其 Schur 补都正定，那么它们的特征值都是正的，从而 $\begin{bmatrix} A & \\ & C-B^{\mathrm{T}}A^{-1}B \end{bmatrix}$ 的特征值也是正的，即它是正定矩阵。那么同样因为

  $$S=\begin{bmatrix} I_n & 0 \\ B^{\mathrm{T}}A^{-1} & I_n \end{bmatrix}\begin{bmatrix} A & \\ & C-B^{\mathrm{T}}A^{-1}B \end{bmatrix}\begin{bmatrix} I_n & 0 \\ B^{\mathrm{T}}A^{-1} & I_n \end{bmatrix}^{\mathrm{T}}$$

  则 S 是正定矩阵。

□

练习 6.2.19. 设 A, B 是 n 阶实对称矩阵，A 正定. 证明，存在可逆矩阵 T，使得 $T^{\mathrm{T}}AT$ 和 $T^{\mathrm{T}}BT$ 同时是对角矩阵.

证明: A 正定，则存在可逆矩阵 P 使得 $A=P^{\mathrm{T}}P$. 我们考虑 $B'=(P^{-1})^{\mathrm{T}}BP^{-1}$, 那么 B' 也是实对称矩阵，那么存在正交矩阵 Q 和对角矩阵 Λ 使得 $B=Q^{\mathrm{T}}\Lambda Q$. 令 $T=P^{-1}Q^{\mathrm{T}}$, 那么

$$T^{\mathrm{T}}AT=Q(P^{\mathrm{T}})^{-1}AP^{-1}Q^{-1}=I_n, T^{\mathrm{T}}BT=QB'Q^{\mathrm{T}}=\Lambda,$$

都是对角矩阵。 □

练习 6.2.20. 设 A, B 是 n 阶实对称矩阵，A, B 半正定. 证明，存在可逆矩阵 T，使得 $T^{\mathrm{T}}AT$ 和 $T^{\mathrm{T}}BT$ 同时是对角矩阵.

提示. 按 $\mathcal{N}(A)\cap\mathcal{N}(B)$ 是否为平凡子空间分类讨论，转化为练习 6.2.19.

证明: 首先，我们断言：实对称半正定矩阵的对角元都大于等于 0, 且若某对角元是 0, 那么该对角元所在的行和列所有元素都是 0。这一点可由练习 4.2.13 得到。

我们考虑 $A+B$, 它是实对称半正定矩阵，因此存在可逆矩阵 P 使得 $P^{\mathrm{T}}(A+B)P=\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$。而 $P^{\mathrm{T}}AP, P^{\mathrm{T}}BP$ 也都是实对称半正定矩阵，所以他们的第 $r+1$ 至 n 个对角元都大于等于 0, 但是相加是 0, 所以 $P^{\mathrm{T}}AP, P^{\mathrm{T}}BP$ 的第 $r+1$ 至 n 个对角元都等于 0，于是存在 r 阶对称矩阵 C 使得 $P^{\mathrm{T}}AP=\begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix}$. 我们把 C 正交对角化，即取 r 阶正交矩阵 Q 和对角矩阵 Λ_r 使得 $C=Q\Lambda_rQ^{\mathrm{T}}$, 那么

$$\left(P\begin{bmatrix} Q & 0 \\ 0 & I_{n-r} \end{bmatrix}\right)^{\mathrm{T}}A\left(P\begin{bmatrix} Q & 0 \\ 0 & I_{n-r} \end{bmatrix}\right)=\begin{bmatrix} \Lambda_r & 0 \\ 0 & 0 \end{bmatrix},$$

那么此时

$$\left(P\begin{bmatrix} Q & 0 \\ 0 & I_{n-r} \end{bmatrix}\right)^{\mathrm{T}}B\left(P\begin{bmatrix} Q & 0 \\ 0 & I_{n-r} \end{bmatrix}\right)=\begin{bmatrix} I_r-\Lambda_r & 0 \\ 0 & 0 \end{bmatrix}$$

也是对角矩阵。 □

6.3 奇异值分解

练习 6.3.1. 求下列矩阵的奇异值分解：

1. $\begin{bmatrix}3 & 4 & 0\end{bmatrix}$. *2.* $\begin{bmatrix}1 & 1 & 0\\0 & 1 & 1\end{bmatrix}$. *3.* $\begin{bmatrix}2 & 2\\-1 & 1\end{bmatrix}$. *4.* $\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}$.

5. $\begin{bmatrix}1 & 1\\3 & 3\end{bmatrix}$. *6.* $\begin{bmatrix}0 & 1 & 0 & 0\\0 & 0 & 0 & 2\\3 & 0 & 0 & 0\\0 & 0 & 4 & 0\end{bmatrix}$. *7.* $\begin{bmatrix}1 & 0\\2 & 1\end{bmatrix}$.

8. $\begin{bmatrix}A & 0\\0 & 0\end{bmatrix}$，其中 A 有奇异值分解 $A=U\Sigma V^{\mathrm{T}}$.

奇异值分解可以对 AA^{T} 对正交相似对角化，也可以对 $A^{\mathrm{T}}A$ 作正交相似对角化。哪个计算量小，就用哪个。

证明: 1. $AA^{\mathrm{T}}=\begin{bmatrix}25\end{bmatrix}$, 所以非零奇异值是 5, 左奇异向量是 $\boldsymbol{v}_1=\begin{bmatrix}1\end{bmatrix}$, 右奇异向量是 $\boldsymbol{w}_1=\frac{1}{5}A^{\mathrm{T}}\boldsymbol{v}_1=\begin{bmatrix}3/5\\4/5\\0\end{bmatrix}$. 我们把 $\boldsymbol{w}_1$ 扩张成 $\mathbb{R}^3$ 的一组标准正交基 $\begin{bmatrix}-4/5\\3/5\\0\end{bmatrix}$, $\begin{bmatrix}0\\0\\1\end{bmatrix}$, 则 A 的奇异值分解为：

$$A=\begin{bmatrix}1\end{bmatrix}\begin{bmatrix}5 & 0 & 0\end{bmatrix}\begin{bmatrix}3/5 & -4/5 & 0\\4/5 & 3/5 & 0\\0 & 0 & 1\end{bmatrix}^{\mathrm{T}}$$

2. $AA^{\mathrm{T}}=\begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}$. AA^{T} 的特征值是 $3,1$，对应的单位特征向量是 $\boldsymbol{v}_1=\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$, $\boldsymbol{v}_2=\begin{bmatrix}1/\sqrt{2}\\-1/\sqrt{2}\end{bmatrix}$. $\boldsymbol{w}_1=\frac{1}{\sqrt{3}}A^{\mathrm{T}}\boldsymbol{v}_2=\begin{bmatrix}1/\sqrt{6}\\2/\sqrt{6}\\1/\sqrt{6}\end{bmatrix}$, $\boldsymbol{w}_2=A^{\mathrm{T}}\boldsymbol{v}_1=\begin{bmatrix}1/\sqrt{2}\\0\\-1/\sqrt{2}\end{bmatrix}$, 把 $\boldsymbol{w}_1,\boldsymbol{w}_2$ 扩张成 $\mathbb{R}^3$ 的一组标准正交基：$\boldsymbol{w}_1,\boldsymbol{w}_2,\boldsymbol{w}_3=\begin{bmatrix}1/\sqrt{3}\\-1\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$. 那么 A 的奇异值分解为

$$A=\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}\begin{bmatrix}\sqrt{3} & 0 & 0\\0 & 1 & 0\end{bmatrix}\begin{bmatrix}1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3}\\2/\sqrt{6} & 0 & -1/\sqrt{3}\\1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3}\end{bmatrix}^{\mathrm{T}}$$

3. $AA^{T}=\begin{bmatrix}8 & 0\\0 & 2\end{bmatrix}$, AA^{T} 的特征值是 $8,2$，对应的单位特征向量是 $\boldsymbol{v}_1=\begin{bmatrix}1\\0\end{bmatrix}$, $\boldsymbol{v}_2=\begin{bmatrix}0\\1\end{bmatrix}$.

$\boldsymbol{w}_1 = \frac{1}{\sqrt{8}} A^{\mathrm{T}} \boldsymbol{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \boldsymbol{w}_2 = \frac{1}{\sqrt{2}} A^{\mathrm{T}} \boldsymbol{w}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$, 那么 A 的奇异值分解为

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{8} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^{\mathrm{T}}$$

4. 直接观察可得 A 的奇异值分解是 $A = UI_2V$，其中 $U = I_2, V = A$.

5. $AA^{\mathrm{T}} = \begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix}$, AA^{T} 的特征值是 $20, 0$，对应的单位特征向量是 $\boldsymbol{v}_1 = \begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} -3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$. $\boldsymbol{w}_1 = \frac{1}{\sqrt{20}} A^{\mathrm{T}} \boldsymbol{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, A^{\mathrm{T}} \boldsymbol{v}_2 = 0$, 我们把 $\boldsymbol{w}_1$ 扩张成 $\mathbb{R}^2$ 的一组基 $\boldsymbol{w}_1, \boldsymbol{w}_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$. 那么 A 的奇异值分解为

$$A = \begin{bmatrix} 1/\sqrt{10} & -3/\sqrt{10} \\ 3/\sqrt{10} & 1/\sqrt{10} \end{bmatrix} \begin{bmatrix} \sqrt{20} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}^{\mathrm{T}}$$

6. 我们考虑 A 的行列置换有：$A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$, 因为置换矩阵都是正交阵，所以这就是 A 的奇异值分解。

7. $AA^{\mathrm{T}} = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$, AA^{T} 的特征值是 $3 + 2\sqrt{2}, 3 - 2\sqrt{2}$，对应的单位特征向量是 $\boldsymbol{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{4+2\sqrt{2}}} \\ \frac{1+\sqrt{2}}{\sqrt{4+2\sqrt{2}}} \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{4-2\sqrt{2}}} \\ \frac{1-\sqrt{2}}{\sqrt{4-2\sqrt{2}}} \end{bmatrix}$.

$$\boldsymbol{w}_1 = \frac{1}{\sqrt{3+2\sqrt{2}}} A^{\mathrm{T}} \boldsymbol{v}_1 = \begin{bmatrix} \frac{1+\sqrt{2}}{\sqrt{4+2\sqrt{2}}} \\ \frac{1}{\sqrt{4+2\sqrt{2}}} \end{bmatrix}, \boldsymbol{w}_2 = \frac{1}{\sqrt{3-2\sqrt{2}}} A^{\mathrm{T}} \boldsymbol{v}_1 = \begin{bmatrix} \frac{1-\sqrt{2}}{\sqrt{4-2\sqrt{2}}} \\ \frac{1}{\sqrt{4-2\sqrt{2}}} \end{bmatrix}$$

那么 A 的奇异值分解为

$$A = \begin{bmatrix} \frac{1+\sqrt{2}}{\sqrt{4+2\sqrt{2}}} & \frac{1-\sqrt{2}}{\sqrt{4-2\sqrt{2}}} \\ \frac{1}{\sqrt{4+2\sqrt{2}}} & \frac{1}{\sqrt{4-2\sqrt{2}}} \end{bmatrix} \begin{bmatrix} 1+\sqrt{2} & 0 \\ 0 & \sqrt{2}-1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{4+2\sqrt{2}}} & \frac{1}{\sqrt{4-2\sqrt{2}}} \\ \frac{1+\sqrt{2}}{\sqrt{4+2\sqrt{2}}} & \frac{1-\sqrt{2}}{\sqrt{4-2\sqrt{2}}} \end{bmatrix}^{\mathrm{T}}$$

8. $\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & 1 \end{bmatrix}^{\mathrm{T}}$. 这就是 $\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}$ 的奇异值分解。

$\square$

练习 6.3.2. 矩阵 A 的 QR 分解 $A = QR$，且 R 有奇异值分解 $R = U\Sigma V^{\mathrm{T}}$，求 A 的奇异值分解.

证明: $A=QR=(QU)\Sigma V^{\mathrm{T}}$, 这就是 A 的奇异值分解。 $\square$

练习 6.3.3. 设 A 的奇异值分解为 $A=U\Sigma V^{\mathrm{T}}$，求矩阵 $\begin{bmatrix}0 & A^{\mathrm{T}}\\ A & 0\end{bmatrix}$ 的谱分解.

证明: 设 $A\in M_{m\times n}(\mathbb{R}), U=\begin{bmatrix}\boldsymbol{u}_1 & \cdots & \boldsymbol{u}_m\end{bmatrix}\in M_{m\times m}(\mathbb{R}), V=\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n\end{bmatrix}\in M_{n\times n}(\mathbb{R}), \sigma_1\geq\cdots\geq\sigma_r>0$ 是 A 的所有非零奇异值。

令 $S=\begin{bmatrix}0 & A^{\mathrm{T}}\\ A & 0\end{bmatrix}$ 那么对 $1\leq i\leq r$

$$S\begin{bmatrix}\boldsymbol{v}_i\\ \boldsymbol{u}_i\end{bmatrix}=\begin{bmatrix}A^{\mathrm{T}}\boldsymbol{u}_i\\ A\boldsymbol{v}_i\end{bmatrix}=\begin{bmatrix}\sigma_i\boldsymbol{v}_i\\ \sigma_i\boldsymbol{u}_i\end{bmatrix}=\sigma_i\begin{bmatrix}\boldsymbol{v}_i\\ \boldsymbol{u}_i\end{bmatrix};S\begin{bmatrix}\boldsymbol{v}_i\\ -\boldsymbol{u}_i\end{bmatrix}=\begin{bmatrix}-A^{\mathrm{T}}\boldsymbol{u}_i\\ A\boldsymbol{v}_i\end{bmatrix}=\begin{bmatrix}-\sigma_i\boldsymbol{v}_i\\ \sigma_i\boldsymbol{u}_i\end{bmatrix}=-\sigma_i\begin{bmatrix}\boldsymbol{v}_i\\ -\boldsymbol{u}_i\end{bmatrix}.$$

对 $r+1\leq i\leq n$

$$S\begin{bmatrix}0\\ \boldsymbol{u}_i\end{bmatrix}=\begin{bmatrix}A^{\mathrm{T}}\boldsymbol{u}_i\\ 0\end{bmatrix}=0;S\begin{bmatrix}\boldsymbol{v}_i\\ 0\end{bmatrix}=\begin{bmatrix}0\\ A\boldsymbol{v}_i\end{bmatrix}=0$$

且:

$$\left\{\begin{bmatrix}\boldsymbol{v}_i\\ \boldsymbol{u}_i\end{bmatrix}:1\leq i\leq r\right\}\cup\left\{\begin{bmatrix}\boldsymbol{v}_i\\ -\boldsymbol{u}_i\end{bmatrix}:1\leq i\leq r\right\}\cup\left\{\begin{bmatrix}\boldsymbol{v}_i\\ 0\end{bmatrix}:r+1\leq i\leq n\right\}\cup\left\{\begin{bmatrix}0\\ \boldsymbol{u}_i\end{bmatrix}:r+1\leq i\leq m\right\}$$

我们把 U,V 分成两块，前 r 列是一块，剩下的列是一块：$U=\begin{bmatrix}U_1 & U_2\end{bmatrix}, V=\begin{bmatrix}V_1 & V_2\end{bmatrix}$ 令 $P=\begin{bmatrix}V_1 & V_1 & V_2 & 0\\ U_1 & -U_1 & 0 & U_2\end{bmatrix}$, 则 A 的谱分解为

$$A=P\mathrm{diag}(\sigma_1,\cdots,\sigma_r,-\sigma_1,\cdots,-\sigma_r,0,\cdots,0)P^{-1}.$$

$\square$

练习 6.3.4. 设矩阵 $A=\begin{bmatrix}1 & 0\\ -1 & 1\end{bmatrix}$，考虑单位圆 $C=\left\{\boldsymbol{v}\in\mathbb{R}^2 \mid \|\boldsymbol{v}\|=1\right\}$ 及其在 A 对应的线性变换 $\mathbf{A}$ 下的像 $A(C)=\left\{A\boldsymbol{v}\in\mathbb{R}^2:\|\boldsymbol{v}\|=1\right\}$.

1. 设 $\boldsymbol{w}\in A(C)$，证明 $\boldsymbol{w}^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}\boldsymbol{w}=1$.
2. 求 A 的奇异值分解 $A=U\Sigma V^{\mathrm{T}}$.
3. 注意 V,U 为二阶正交矩阵，对应的线性变换是旋转或反射，而 Σ 是对角矩阵，对应伸缩变换. 从几何上看，曲线 $V^{\mathrm{T}}(C),\Sigma V^{\mathrm{T}}(C),U\Sigma V^{\mathrm{T}}(C)$ 分别是什么形状?

证明:

1. $\boldsymbol{w}\in A(C)$, 则存在 $\boldsymbol{v}\in\mathbb{R}^2,\|\boldsymbol{v}\|=1$ 使得 $\boldsymbol{w}=A\boldsymbol{v}$, 那么

$$\boldsymbol{w}^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}\boldsymbol{w}=\boldsymbol{v}^{\mathrm{T}}A^{\mathrm{T}}(A^{\mathrm{T}})^{-1}A^{-1}A\boldsymbol{v}=\|\boldsymbol{v}\|^2=1.$$

2. $AA^{\mathrm{T}}=\begin{bmatrix}1 & -1\\ -1 & 2\end{bmatrix}$, 它的两个特征值是 $\frac{3+\sqrt{5}}{2},\frac{3-\sqrt{5}}{2}$，对应的单位特征向量是 $\boldsymbol{v}_1=$

$$\begin{bmatrix}\dfrac{1-\sqrt{5}}{\sqrt{10-2\sqrt{5}}}\\ \dfrac{2}{\sqrt{10-2\sqrt{5}}}\end{bmatrix}, \boldsymbol{v}_2=\begin{bmatrix}\dfrac{1+\sqrt{5}}{\sqrt{10+2\sqrt{5}}}\\ \dfrac{2}{\sqrt{10+2\sqrt{5}}}\end{bmatrix}.$$

$$\boldsymbol{w}_1=\frac{2}{1+\sqrt{5}}A^{\mathrm{T}}\boldsymbol{v}_1=\begin{bmatrix}\dfrac{-2}{\sqrt{10-2\sqrt{5}}}\\ \dfrac{\sqrt{5}-1}{\sqrt{10-2\sqrt{5}}}\end{bmatrix}, \boldsymbol{w}_2=\frac{2}{\sqrt{5}-1}A^{\mathrm{T}}\boldsymbol{v}_2=\begin{bmatrix}\dfrac{2}{\sqrt{10+2\sqrt{5}}}\\ \dfrac{\sqrt{5}+1}{\sqrt{10+2\sqrt{5}}}\end{bmatrix}$$

那么 A 的奇异值分解为

$$A=\begin{bmatrix}\dfrac{1-\sqrt{5}}{\sqrt{10-2\sqrt{5}}} & \dfrac{1+\sqrt{5}}{\sqrt{10+2\sqrt{5}}}\\ \dfrac{2}{\sqrt{10-2\sqrt{5}}} & \dfrac{2}{\sqrt{10+2\sqrt{5}}}\end{bmatrix}\begin{bmatrix}\dfrac{\sqrt{5}+1}{2} & 0\\ 0 & \dfrac{\sqrt{5}-1}{2}\end{bmatrix}\begin{bmatrix}\dfrac{-2}{\sqrt{10-2\sqrt{5}}} & \dfrac{2}{\sqrt{10+2\sqrt{5}}}\\ \dfrac{\sqrt{5}-1}{\sqrt{10-2\sqrt{5}}} & \dfrac{\sqrt{5}+1}{\sqrt{10+2\sqrt{5}}}\end{bmatrix}^{\mathrm{T}}$$

3. $V^{\mathrm{T}}(C)$ 是单位圆，$\Sigma V^{\mathrm{T}}(C)$ 是椭圆，$U\Sigma V^{\mathrm{T}}(C)$ 也是椭圆。

□

练习 6.3.5. 设矩阵 A 的奇异值分解是 $A=U\Sigma V^{\mathrm{T}}$.

1. 证明 $AA^{\mathrm{T}}=U(\Sigma\Sigma^{\mathrm{T}})U^{\mathrm{T}}, A^{\mathrm{T}}A=V(\Sigma^{\mathrm{T}}\Sigma)V^{\mathrm{T}}$ 分别是这两个对称矩阵的谱分解，并得到 AA^{T} 和 $A^{\mathrm{T}}A$ 的非零特征值相同.

2. 对任意 A 的奇异值 $\sigma\neq 0$，设 $\boldsymbol{v}$ 和 $\boldsymbol{w}$ 分别是 $A^{\mathrm{T}}A$ 和 AA^{T} 的属于 σ^2 的特征向量，证明 $A\boldsymbol{v}$ 和 $A^{\mathrm{T}}\boldsymbol{w}$ 分别是 AA^{T} 和 $A^{\mathrm{T}}A$ 的属于 σ^2 的特征向量.

证明*:* 1. 只需要说明 $\Sigma\Sigma^{\mathrm{T}},\Sigma^{\mathrm{T}}\Sigma$ 都是对角矩阵。设 $A\in M_{m\times n}(\mathbb{R}), r=\operatorname{rank}(A)$, 则 $\Sigma=\begin{bmatrix}\Lambda_r & 0_{r\times(n-r)}\\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)}\end{bmatrix}\in M_{m\times n}(\mathbb{R})$, 其中 $\Lambda_r=\operatorname{diag}(\sigma_1,\cdots,\sigma_r)$, 那么直接计算：

$$\Sigma\Sigma^{\mathrm{T}}=\begin{bmatrix}\Lambda_r^2 & 0\\ 0 & 0_{(m-r)\times(m-r)}\end{bmatrix}, \Sigma^{\mathrm{T}}\Sigma=\begin{bmatrix}\Lambda_r^2 & 0\\ 0 & 0_{(n-r)\times(n-r)}\end{bmatrix}$$

□

练习 6.3.6 (极分解). 对 n 阶方阵 A，存在正交矩阵 Q 和对称半正定矩阵 S，使得 $A=QS$.

分解式 $A=QS$ 称为 A 的**极分解**. 容易看到，$A=S_1Q_1$，即方阵分解为对称半正定矩阵和正交矩阵的乘积，也存在.

证明*:* 设 A 的奇异值分解为 $A=U\Sigma V^{\mathrm{T}}$, 其中 U,V 为正交矩阵。A 是方阵，则 Σ 是对角线非负。那么 $A=(UV^{\mathrm{T}})(V\Sigma V^{\mathrm{T}})$，令 $Q=UV^{\mathrm{T}},S=V\Sigma V^{\mathrm{T}}$.

$A=(U\Sigma U^{\mathrm{T}})(UV^{\mathrm{T}})$, 令 $S_1=U\Sigma U^{\mathrm{T}},Q_1=UV^{\mathrm{T}}$ 即符合题意。 □

练习 6.3.7. 证明矩阵的广义逆唯一.

提示. 耐心计算.

我们首先需要理解一下问题。广义逆的定义方式如下：设 A 有奇异值分解：$A = U\Sigma V^{\mathrm{T}}$；注意，奇异值分解中只有 $\Sigma = \begin{bmatrix}\Sigma_r & 0\\ 0 & 0\end{bmatrix}$ 是唯一的，其中 $\Sigma_r = \mathrm{diag}(\sigma_1, \cdots, \sigma_r), \sigma_1 \geq \cdots \geq \sigma_r > 0$, U, V 都不唯一。然后我们把 U, V 也分块：$U = \begin{bmatrix}U_r & *\end{bmatrix}, V = \begin{bmatrix}V_r & *\end{bmatrix}$, 其中 U_r, V_r 分别是 U, V 的前 r 列，那么 A 的广义逆定义为

$$A^+ = V_r\Sigma_r^{-1}U_r^{\mathrm{T}}$$

表面看起来 A^+ 取决于 U, V 特别是 U_r, V_r，那么本题的目的就是要证明广义逆和它们无关。

证明: 设 $A = \tilde{U}\Sigma\tilde{V}^{\mathrm{T}}$ 也是 A 的奇异值分解，符号类比于 $\tilde{U}_r, \tilde{V}_r$ 分别是 $\tilde{U}, \tilde{V}$ 的前 r 列.

因为 $U_r, \tilde{U}_r$ 的列向量分别都是 $\mathcal{R}A$ 的一组标准正交基，那么存在可逆矩阵 $X \in M_{r\times r}$ 使得 $\tilde{U}_r = U_rX$; 又因为 $U_r^{\mathrm{T}}U_r = \tilde{U}_r\tilde{U}_r = I_r$, 那么 $X^{\mathrm{T}}X = I_r$, 即 X 是正交阵。同理，存在正交矩阵 Y 使得 $\tilde{V}_r = V_rY$

同时，我们对 A 的两个奇异值分解用分块矩阵乘法：

$$A = \begin{bmatrix}U_r & *\end{bmatrix}\begin{bmatrix}\Sigma_r & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}V_r & 0\end{bmatrix}^T = U_r\Sigma_rV_r^{\mathrm{T}},$$

同理 $A = \tilde{U}_r\Sigma_r\tilde{V}_r^{\mathrm{T}}$, 这就有 $U_r\Sigma_rV_r^{\mathrm{T}} = \tilde{U}_r\Sigma_r\tilde{V}_r^{\mathrm{T}}$, 即

$$U_r\Sigma_rV_r^{\mathrm{T}} = U_rX\Sigma_rY^{\mathrm{T}}V_r^{\mathrm{T}}.$$

再利用 $U_r^{\mathrm{T}}U_r = V_r^{\mathrm{T}}V_r = I_r$ 知 $\Sigma_r = X\Sigma_rY^{\mathrm{T}}$, 所以 $Y\Sigma_r^{-1}X^{\mathrm{T}} = \Sigma_r^{-1}$；进一步地就有

$$V_r\Sigma_r^{-1}U_r^{\mathrm{T}} = V_rY\Sigma_r^{-1}X^{\mathrm{T}}U_r^{\mathrm{T}} = \tilde{V}_r\Sigma_r^{-1}\tilde{U}_r^{\mathrm{T}}.$$

□

练习 6.3.8. *证明命题 6.3.7. 矩阵的谱范数满足：*

1. $\|A\| \geq 0$, 且 $\|A\| = 0$ 当且仅当 $A = 0$;
2. $\|kA\| = |k|\,\|A\|$;
3. $\|A + B\| \leq \|A\| + \|B\|$;
4. $\|AB\| \leq \|A\|\,\|B\|$;
5. 如果 U, V 正交，则 $\left\|UAV^{\mathrm{T}}\right\| = \|A\|$.

证明:

1. $\|A\| \geq 0$ 是明显的。$\|A\| = 0$ 当且仅当对任意 $\boldsymbol{x}$, $\|A\boldsymbol{x}\| = 0$, 当且仅当对任意 $\boldsymbol{x}$, $A\boldsymbol{x} = 0$, 当且仅当 $A = 0$.
2. $$\|kA\| = \max_{\boldsymbol{x}\neq 0}\frac{\|kA\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = |k|\max_{\boldsymbol{x}\neq 0}\frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = |k|\,\|A\|\,.$$
3. $\|A + B\| = \max_{\boldsymbol{x}\neq 0}\dfrac{\|(A+B)\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \max_{\boldsymbol{x}\neq 0}\left(\dfrac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|} + \dfrac{\|B\boldsymbol{x}\|}{\|\boldsymbol{x}\|}\right) \leq \|A\| + \|B\|$.
4. 对任意 $\boldsymbol{x} \neq 0$, 若 $B\boldsymbol{x} = 0$, 则 $AB\boldsymbol{x} = 0$; 若 $B\boldsymbol{x} \neq 0$, 则 $\|AB\boldsymbol{x}\| \leq \|A\|\,\|B\boldsymbol{x}\| \leq \|A\|\,\|B\|\,\|\boldsymbol{x}\|$. 所以 $\|AB\| \leq \|A\|\,\|B\|$.

5. $\left\|UAV^{\mathrm{T}}\right\| = \max_{\boldsymbol{x}\neq 0}\dfrac{\left\|UAV^{\mathrm{T}}\boldsymbol{x}\right\|}{\|\boldsymbol{x}\|} = \max_{\boldsymbol{x}\neq 0}\dfrac{\|A\boldsymbol{x}\|}{\|V\boldsymbol{x}\|} = \max_{\boldsymbol{x}\neq 0}\dfrac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = \|A\|.$

□

练习 6.3.9. 证明矩阵任意特征值的绝对值不大于其最大的奇异值.

证明: 设 λ 是 A 的绝对值最大的特征值，且 $\boldsymbol{x}\neq 0$ 是属于 λ 的特征向量，那么

$$|\lambda| = \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \max_{\boldsymbol{x}\neq 0}\frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|},$$

而 A 的最大奇异值即 $\max_{\boldsymbol{x}\neq 0}\dfrac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|}$. □

练习 6.3.10. 证明或者举出反例.

1. n 阶方阵 A 为正交矩阵当且仅当它的 n 个奇异值都是 1.
2. n 阶方阵的 n 个奇异值的乘积等于所有特征值的乘积.
3. 设 n 阶方阵 A 和 $A+I_n$ 的奇异值分解分别为 $A=U\Sigma V^{\mathrm{T}}, A+I_n=U(\Sigma+I_n)V^{\mathrm{T}}$. 证明 A 是对称矩阵.
4. 如果 n 阶方阵 A 的 n 个奇异值就是它的 n 个特征值，则 A 是对称矩阵.

证明:

1. 正确。若 A 是正交矩阵，那么 $A=AI_nI_n$ 就是 A 的奇异值分解，所以它的 n 个奇异值都是 1. 另一方面，若 A 的 n 个奇异值都是 1，那么 A 的奇异值分解为 $A=UI_nV^{\mathrm{T}}$，所以 A 是正交矩阵。
2. 错误。例如 $A=\begin{bmatrix}-1\end{bmatrix}$, A 的奇异值是 1，特征值是 -1.
3. 正确。$A+I_n=U(\Sigma+I_n)V^{\mathrm{T}}=U\Sigma V^{\mathrm{T}}+UV^{\mathrm{T}}$, 所以 $UV^{\mathrm{T}}=I_n$, 那么 $U=V, A=U\Sigma U^{\mathrm{T}}$. Σ 此时是对角矩阵，所以对称，那么 A 对称。
4. 设 A 的奇异值分解为 $A=U\Sigma V^{\mathrm{T}}$, 则 $AA^{\mathrm{T}}=U\Sigma\Sigma^{\mathrm{T}}U^{\mathrm{T}}$, 所以

$$\operatorname{trace}(AA^{\mathrm{T}}) = \operatorname{trace}(\Sigma\Sigma^{\mathrm{T}}) = \sum_{i=1}^{n}\sigma_i^2 = \sum_{i=1}^{n}\lambda_i^2 = \operatorname{trace}(A^2)$$

最后一个等号可由命题 5.4.6 得到。设 $A=\begin{bmatrix}a_{ij}\end{bmatrix}$, 那么

$$\operatorname{trace}(AA^{\mathrm{T}}) = \sum_{i=1}^{n}(AA^{\mathrm{T}})_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2; \quad \operatorname{trace}(A^2) = \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}a_{ji};$$

注意到

$$\operatorname{trace}(A^2) \leq \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij}^2+a_{ji}^2\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2$$

取等号当且仅当对任意 $i,j,a_{ij}=a_{ji}$; 所以 $A=A^{\mathrm{T}}$ 是对称矩阵。

□

练习 6.3.11. 证明命题 *6.3.15.* 矩阵的 *Frobenius* 范数满足:

1. $\|A\|_F \geq 0$, 且 $\|A\|_F = 0$ 当且仅当 $A=0$;
2. $\|kA\|_F = |k|\,\|A\|_F$;

3. $\|A+B\|_F \leq \|A\|_F + \|B\|_F$;

4. $\|AB\|_F \leq \|A\| \|B\|_F$, $\|AB\|_F \leq \|A\|_F \|B\|$;

5. 如果 U, V 正交，则 $\left\|UAV^{\mathrm{T}}\right\|_F = \|A\|_F$.

证明: 1. 根据定义，若 $A = \left[a_{ij}\right]$, $\|A\|_F = \left(\sum_{i,j} |a_{ij}|^2\right)^{1/2}$，显然 $\|A\|_F \geq 0$；$\|A\|_F = 0$ 当且仅当 $|a_{ij}| = 0, \forall i, j$, 即 $A = 0$.

2. $\|kA\|_F = \left(\sum_{i,j} |ka_{ij}|^2\right)^{1/2} = |k| \|A\|_F$.

3. $\|A+B\|_F^2 = \left(\sum_{i,j} |a_{ij}+b_{ij}|^2\right) = \sum_{ij} a_{ij}^2 + 2\sum_{i,j} a_{ij}b_{ij} + \sum_{i,j} b_{ij}^2$, 而

$$\left(\|A\|_F + \|B\|_F\right)^2 = \sum_{i,j} a_{ij}^2 + 2\left(\sum_{i,j} a_{ij}^2\right)^{1/2}\left(\sum_{i,j} b_{ij}^2\right)^{1/2} + \sum_{i,j} b_{ij}^2$$

所以 $\|A+B\|_F \leq \|A\|_F + \|B\|_F$ 当且仅当 $\sum_{i,j} a_{ij}b_{ij} \leq \left(\sum_{i,j} a_{ij}^2\right)^{1/2}\left(\sum_{i,j} b_{ij}^2\right)^{1/2}$，而这恰好就是 Cauchy-Schwartz 不等式。

4. 我们需要注意到，对列向量 $\boldsymbol{v}$ 我们 $\|\boldsymbol{v}\|_F = \|\boldsymbol{v}\|$. 我们先证明 $B = \boldsymbol{v}$ 是个列向量的情形。此时

$$\|A\boldsymbol{v}\|_F = \|A\boldsymbol{v}\| \leq \|A\| \|\boldsymbol{v}\| = \|A\| \|\boldsymbol{v}\|_F.$$

若 $B = \begin{bmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k \end{bmatrix}$, 则 $\|B\|_F^2 = \sum_{i=1}^k \|\boldsymbol{v}_i\|_F^2$（这是因为等式两边都是 B 中所有元素的平方和）。那么

$$\|AB\|_F^2 = \sum_{i=1}^k \|A\boldsymbol{v}_i\|^2 \leq \sum_{i=1}^k \|A\|^2 \|\boldsymbol{v}_i\|^2 = \|A\|^2 \sum_{i=1}^k \|\boldsymbol{v}_i\|^2 = \|A\|^2 \|B\|_F^2.$$

开方即可。

另一方面，$\|A\|_F = \sqrt{\operatorname{trace}(A^{\mathrm{T}}A)} = \sqrt{\operatorname{trace}(AA^{\mathrm{T}})} = \left\|A^{\mathrm{T}}\right\|_F$; 又 A 和 A^{T} 有相同的奇异值，这就说明 $\|A\| = \left\|A^{\mathrm{T}}\right\|$；所以

$$\|AB\|_F = \left\|B^{\mathrm{T}}A^{\mathrm{T}}\right\|_F \leq \left\|B^T\right\| \left\|A^T\right\|_F = \|B\| \|A\|_F.$$

5. 若 U, V 正交，则

$$\begin{aligned} \left\|UAV^{\mathrm{T}}\right\|_F &= \sqrt{\operatorname{trace}(VA^{\mathrm{T}}U^{\mathrm{T}}UAV^{\mathrm{T}})} \\ &= \sqrt{\operatorname{trace}(VA^{\mathrm{T}}AV^{\mathrm{T}})} \\ &= \sqrt{\operatorname{trace}(A^{\mathrm{T}}AV^{\mathrm{T}}V)} \\ &= \sqrt{\operatorname{trace}(A^{\mathrm{T}}A)} = \|A\|_F \end{aligned}.$$

□

练习 6.3.12. 证明命题 *6.3.16*：对任意矩阵 A，其 *Frobenius* 范数平方 $\|A\|_F^2$ 等于 A 的所有奇异值的平方和，因此 $\|A\|_F \geq \|A\|$.

证明*:* 设 A 的奇异值分解为 $A = U\Sigma V^{\mathrm{T}}$, 由上一题第 5 小题：

$$\|A\|_F^2 = \|\Sigma\| = \mathrm{trace}(\Sigma\Sigma^{\mathrm{T}}) = \sigma_1^2 + \cdots + \sigma_r^2.$$

而 $\|A\| = \sigma_1$, 所以 $\|A\|_F \geq \|A\|$. □

练习 6.3.13 (樊畿迹定理). 对任意 n 阶对称矩阵 $A \in M_{n\times n}(\mathbb{R})$，假设特征值为 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$，对应特征向量为 $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$，则 $\max\limits_{\substack{n\times m \text{ 矩阵 } Q: \\ Q^{\mathrm{T}}Q=I}} \mathrm{trace}(Q^{\mathrm{T}}AQ) = \sum\limits_{i=1}^{m} \lambda_i$，且 $Q = \begin{bmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_m\end{bmatrix}$ 时取得最大值.

证明*:* 令 $P^{\mathrm{T}} = \begin{bmatrix}\boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n\end{bmatrix}, \Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$, 那么 $A = P^{\mathrm{T}}\Lambda P$. 另一方面，对 $Q \in M_{n\times m}(\mathbb{R})$, $Q^{\mathrm{T}}Q = I_m$ 当且仅当 $(PQ)^{\mathrm{T}}(PQ) = I_m$. 所以

$$\max_{\substack{n\times m \text{ 矩阵 } Q: \\ Q^{\mathrm{T}}Q=I}} \mathrm{trace}(Q^{\mathrm{T}}AQ) = \max_{\substack{n\times m \text{ 矩阵 } Q: \\ Q^{\mathrm{T}}Q=I}} \mathrm{trace}\left((PQ)^{\mathrm{T}}\Lambda(PQ)\right) = \max_{\substack{n\times m \text{ 矩阵 } Q: \\ Q^{\mathrm{T}}Q=I}} \mathrm{trace}(Q^{\mathrm{T}}\Lambda Q)$$

对于 $Q \in M_{n\times m}(\mathbb{R}), Q^{\mathrm{T}}Q = I_m$, 我们记 $Q = \begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_m\end{bmatrix}$, 那么

$$\mathrm{trace}(Q^{\mathrm{T}}\Lambda Q) = \sum_{i=1}^{m} \boldsymbol{e}_i^{\mathrm{T}}Q^{\mathrm{T}}\Lambda Q\boldsymbol{e}_i = \sum_{i=1}^{m} \boldsymbol{v}_i^{\mathrm{T}}\Lambda\boldsymbol{v}_i$$

令 $\boldsymbol{v}_i = \begin{bmatrix}x_{1i} \\ \vdots \\ x_{ni}\end{bmatrix}$, 则 $\sum_{j=1}^{n} x_{ji}^2 = 1$. 另一方面，由于 Q 的列向量是两两正交的单位向量，所以我们可以把 Q 的列向量扩张成 $\mathbb{R}^n$ 的一个标准正交基，这样我们往 Q 增加了一些列向量补成一个正交矩阵，这说明 Q 的行向量的模长都小于等于 1，即对任意 $1 \leq j \leq n$, 都有 $\sum_{i=1}^{m} x_{ji}^2 \leq 1$.

对任意 $1 \leq i \leq m$,

$$\boldsymbol{v}_i^{\mathrm{T}}\Lambda\boldsymbol{v}_i = \sum_{j=1}^{n} \lambda_j x_{ji}^2 = \lambda_m + \sum_{j=1}^{n}(\lambda_j - \lambda_m)x_{ji}^2 \leq \lambda_m + \sum_{j=1}^{m}(\lambda_j - \lambda_m)x_{ji}^2$$

所以

$$\sum_{i=1}^{m} \lambda_i - \sum_{i=1}^{m} \boldsymbol{v}_i^{\mathrm{T}}\Lambda\boldsymbol{v}_i \geq \sum_{i=1}^{m}(\lambda_i - \lambda_m) - \sum_{i=1}^{m}\sum_{j=1}^{m}(\lambda_j - \lambda_m)x_{ji}^2$$

我们将上式最右端第一项改一下求和指标，并交换第二项的求和顺序：

$$\sum_{i=1}^{m}(\lambda_i - \lambda_m) - \sum_{i=1}^{m}\sum_{j=1}^{m}(\lambda_j - \lambda_m)x_{ji}^2 = \sum_{j=1}^{m}(\lambda_j - \lambda_m)\left(1 - \sum_{i=1}^{m} x_{ji}^2\right) \geq 0$$

所以

$$\sum_{i=1}^{m} \lambda_i \geq \sum_{i=1}^{m} \boldsymbol{v}_i^{\mathrm{T}}\Lambda\boldsymbol{v}_i.$$

显然，当 $Q = \begin{bmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_m\end{bmatrix}$ 时可以取等号 (注意：不是充要条件)。 □

练习 6.3.14 (低秩逼近与数据拟合). 考虑平面上的点 $\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1\\8\end{bmatrix}, \begin{bmatrix}3\\8\end{bmatrix}, \begin{bmatrix}4\\20\end{bmatrix}$. 现在想找到一条直线，使得点到直线距离的平方和最小. 令 $A = \begin{bmatrix}0 & 1 & 3 & 4\\0 & 8 & 8 & 20\end{bmatrix}$.

1. 找到函数 $f(a,b,c)$，使得 $f(a,b,c)$ 就是每个点到直线 $ax+by+c=0$ 的距离的平方和.
2. 假设 a,b 已知，用导数证明此时最好的 c 是 $-\begin{bmatrix}a & b\end{bmatrix}\begin{bmatrix}x_0\\y_0\end{bmatrix}$，这里 x_0, y_0 代表四个点的 x 坐标平均值和 y 坐标平均值.

   附注. 这意味着欲求直线应为 $a(x-x_0)+b(y-y_0)+c=0$.
3. 对 A 进行"中心化"，即对每一行所有元素减去该行的平均值，得到 B. 此时对 B 的列对应的四个点来说，最佳直线为何应该经过原点？A, B 对应的最佳直线为何一定平行？
4. 计算 B 的最佳秩 1 逼近.
5. 计算欲求直线.
6. 思考：如果考虑的不是平面上的点，而是 n 维空间中的点，问题如何处理？

证明:

1. 点 (x_0, y_0) 到直线 $ax+by+c=0$ 的距离为 $\dfrac{|ax_0+by_0+c|}{\sqrt{a^2+b^2}}$, 所以
$$f(a,b,c) = \frac{1}{a^2+b^2}\left(c^2+(a+8b+c)^2+(3a+8b+c)^2+(4a+20b+c)^2\right)$$
2. a,b 已知，那么要求 c 的话，考虑 $\dfrac{\partial f}{\partial c} = \dfrac{2}{a^2+b^2}(4c+(a+8b+3a+8b+4a+20b))$, 利用导数考虑函数的极值点和单调性可知 $c = -\begin{bmatrix}a & b\end{bmatrix}\begin{bmatrix}x_0\\y_0\end{bmatrix}$.
3. 从 A 变到 B 是对给定的四个点同时平移了一个固定的向量，所以最佳直线自然也平移同一个固定的向量。对于 B 而言，此时四个点的 x, y 的坐标的平均值都是 0，利用上一小题知最佳直线经过原点。
4. $B = \begin{bmatrix}-2 & -1 & 1 & 2\\-9 & -1 & -1 & 11\end{bmatrix}$, B 的奇异值分解数字过于离谱，省略。由奇异值分解可得 B 的最佳秩 1 逼近。
5. 所求直线是由 B 的最大奇异值对应的左奇异向量张成的直线。

□

练习 6.3.15. 考虑子空间 $\mathcal{M}, \mathcal{N}$，其对应的正交投影矩阵为 P, Q. 我们想要研究矩阵

$$H = P(P+Q)^+Q + Q(P+Q)^+P.$$

1. 计算 $(P+Q)(P+Q)^+$ 的列空间和零空间，该矩阵是否为一个正交投影矩阵？
2. 计算 $(P+Q)^+(P+Q)$ 的列空间和零空间，该矩阵是否为一个正交投影矩阵？和前一矩阵有何关联？
3. 证明 $Q(P+Q)^+(P+Q) = Q$，$(P+Q)(P+Q)^+Q = Q$.
4. 证明 $H = 2P(P+Q)^+Q = 2Q(P+Q)^+P$.

5. 假设 T 是 $\mathcal{M}\cap\mathcal{N}$ 上的正交投影矩阵，证明 $HP=HQ=HT=H$.

6. 证明 $HT=T$.

注意 $PT=QT=T$，再利用第 1 条，可得 $H=HT=T$，由此即得 $\mathcal{M}\cap\mathcal{N}$ 的正交投影矩阵的表达式.

证明: 令 $A=P+Q, V=\mathcal{M}\cap\mathcal{N}$.

1. AA^+ 是 $\mathcal{R}(A)$ 上的正交投影的表示矩阵，它的列空间是 $\mathcal{R}(A)=\mathcal{R}(P+Q), \mathcal{N}(A)=\mathcal{N}(P+Q)$. 令 $\boldsymbol{x}\in\mathcal{N}(P+Q)$, 则 $P\boldsymbol{x}=-Q\boldsymbol{x}, P\boldsymbol{x}\in\mathcal{M}, Q\boldsymbol{x}\in\mathcal{N}$, 所以 $P\boldsymbol{x}=-Q\boldsymbol{x}\in V$, 那么两边左乘 T 就有 $T\boldsymbol{x}=-T\boldsymbol{x}$, 所以 $T\boldsymbol{x}=0$. 因为 $P\boldsymbol{x}\in V$, $P\boldsymbol{x}=TP\boldsymbol{x}=T\boldsymbol{x}=0$, 所以 $P\boldsymbol{x}=Q\boldsymbol{x}=0$, 即 $\boldsymbol{x}\in\mathcal{N}(P)\cap\mathcal{N}(Q)=\mathcal{M}^\perp\cap\mathcal{N}^\perp=(\mathcal{M}+\mathcal{N})^\perp$, 这说明 $\mathcal{N}(P+Q)\subseteq(\mathcal{M}+\mathcal{N})^\perp$. 另一个方向的包含关系是显然的，所以 $\mathcal{N}(P+Q)=(\mathcal{M}+\mathcal{N})^\perp$. 进一步地，$\mathcal{R}(P+Q)=\mathcal{M}+\mathcal{N}$.
2. A^+A 是 $\mathcal{R}(A^{\mathrm{T}})$ 上的正交投影的表示矩阵；注意到 $A^{\mathrm{T}}=(P+Q)^{\mathrm{T}}=P^{\mathrm{T}}+Q^{\mathrm{T}}=P+Q=A$, 因此 $A^+A=AA^+$，所以它们的行空间列空间都相等。
3. 因为 $Q, A^+A=AA^+$ 分别是子空间 $\mathcal{N}, \mathcal{M}+\mathcal{N}$ 对应的正交投影矩阵，因此 $Q(P+Q)^+(P+Q)=Q$，$(P+Q)(P+Q)^+Q=Q$。同理我们还有 $P(P+Q)^+(P+Q)=P$，$(P+Q)(P+Q)^+P=P$.
4. 一方面，$Q=Q(P+Q)^+(P+Q)=Q(P+Q)^+P+Q(P+Q)^+Q$, 另一方面，$Q=(P+Q)(P+Q)^+Q=P(P+Q)^+Q+Q(P+Q)^+Q$, 那么我们有
$$P(P+Q)^+Q=Q(P+Q)^+P.$$
剩下的是明显的。
5.
   - $HP=2Q(P+Q)^+PP=2Q(P+Q)^+P=H$,
   - $HQ=2P(P+Q)^+QQ=2P(P+Q)^+Q=H$,
   - 令 $B=(I-P)+(I-Q)=2I-P-Q$, 那么
   $$T=P_{\mathcal{M}\cap\mathcal{N}}=I-P_{(\mathcal{M}\cap\mathcal{N})^\perp}=I-P_{\mathcal{M}^\perp+\mathcal{N}^\perp}=I-BB^+$$
   那么
   $$HT=H-HB=H-(2H-HP-HQ)B^+=H.$$
6. 对任意 $\boldsymbol{x}\in(\mathcal{M}\cap\mathcal{N})^\perp=\mathcal{N}(T)$, $HT\boldsymbol{x}=0=T\boldsymbol{x}$; 对任意 $\boldsymbol{x}\in\mathcal{M}\cap\mathcal{N}$, $P\boldsymbol{x}=Q\boldsymbol{x}=(P+Q)(P+Q)^+=T\boldsymbol{x}=\boldsymbol{x}$
$$\begin{aligned}HT\boldsymbol{x}&=H\boldsymbol{x}=P(P+Q)^+Q\boldsymbol{x}+Q(P+Q)^+P\boldsymbol{x}\\&=P(P+Q)^+\boldsymbol{x}+Q(P+Q)^+\boldsymbol{x}\\&=(P+Q)(P+Q)^+\boldsymbol{x}=\boldsymbol{x}\end{aligned}$$
由正交投影的含义知 $HT=T$

□

第七章 线性空间和线性映射

7.1 线 性 空 间

练习 7.1.1. 在所有正实数构成的集合 $\mathbb{R}^+$ 上，定义加法和数乘运算：

$$a \oplus b := ab, \quad k \odot a := a^k, \quad \forall a, b \in \mathbb{R}^+, k \in \mathbb{R}.$$

判断 $\mathbb{R}^+$ 对这两个运算是否构成 $\mathbb{R}$ 上的线性空间.

证明: 我们需要验证这两个运算是否符合线性空间的八条运算法则。

1. 加法结合律，即要验证 $(a \oplus b) \oplus c = a \oplus (b \oplus c)$. 事实上，

$$(a \oplus b) \oplus c = (ab) \oplus c = abc, a \oplus (b \oplus c) = a \oplus (bc) = abc.$$

2. 加法交换律，即要验证 $a \oplus b = b \oplus a$. 事实上

$$a \oplus b = ab = ba = b \oplus a.$$

3. 零元素，即存在元素 $\mathbf{0}$, 对任意 $a \in \mathbb{R}^+$, $\mathbf{0} \oplus a = a$. 事实上，我们取 $\mathbf{0} = 1$.
4. 负元素。需要注意的是，我们的零元素 $\mathbf{0}$ 它不是 0, 而是 1, 所以对任意 $a \in \mathbb{R}^+$, 我们需要证明存在 b 使得 $a \oplus b = \mathbf{0} = 1$. 那么显然，我们取 $b = a^{-1}$.
5. 单位数。单位数是系数域中的元素，不是线性空间中的元素。因为 $1 \odot a = a^1 = a$, 所以 1 是单位数。
6. 数乘结合律，即要验证对任意 $a \in \mathbb{R}^+, k, l \in \mathbb{R}$, 有 $(kl) \odot a = k \odot (l \odot a)$. 事实上:

$$k \odot (l \odot a) = k \odot (a^l) = (a^l)^k = a^{kl} = (kl) \odot a.$$

7. 数乘对数的分配律: 即要验证对任意 $a \in \mathbb{R}^+, k, l \in \mathbb{R}$, 都有 $(k+l) \odot a = (k \odot a) \oplus (l \odot a)$. 事实上，

$$(k + l) \odot a = a^{k+1} = a^k a^l = (a^k) \oplus (a^l) = (k \cdot a) \oplus (l \odot a).$$

8. 数乘对向量的分配律：即要验证对任意 $a, b \in \mathbb{R}^+, k \in \mathbb{R}$, 都有 $k \odot (a \oplus b) = (k \odot a) \oplus (k \odot b)$. 事实上，

$$k \odot (a \oplus b) = (a \oplus b)^k = (ab)^k = a^k b^k = (a^k) \oplus (b^k) = (k \cdot a) \oplus (k \odot b).$$

□

练习 7.1.2. 令 $\omega = \dfrac{-1+\sqrt{3}\boldsymbol{i}}{2}$，$\mathbb{Q}[\omega] = \{a + b\omega \mid a, b \in \mathbb{Q}\}$.

1. 证明 $\mathbb{Q}[\omega]$ 关于数的加法和数乘构成 $\mathbb{Q}$ 上的一个线性空间.

2. 证明子集 $\mathbb{Q}$ 和 $\mathcal{M} = \{b\omega \mid b \in \mathbb{Q}\}$ 都是 $\mathbb{Q}[\omega]$ 的子空间. 并求二者的交与和.

3. 判断 $\mathbb{Q}[\omega]$ 是否为数域.

证明*:* 1. 要验证 $\mathbb{Q}[\omega]$ 是线性空间，也就是要验证它满足线性空间的八条运算法则。按照固定步骤取验证即可，这是容易的，我们略过。

2. 我们只需要验证 $\mathbb{Q}$ 和 $\mathcal{M}$ 对以 $\mathbb{Q}$ 为系数的线性组合封闭，这也是容易的，略去。考虑 $z \in \mathbb{Q} \cap \mathcal{M}$, 那么存在 $b \in \mathbb{Q}$ 使得 $z = b\omega = -\frac{b}{2} + \frac{b\sqrt{3}}{2}\boldsymbol{i}$, 那么 $z \in \mathbb{Q}$ 就推出 $b = 0$，所以 $\mathbb{Q} \cap \mathcal{M} = \{0\}$. 按照定义，$\mathbb{Q} + \mathcal{M} = \mathbb{Q}[\omega]$.

3. 我们这里需要说明一下，考虑 $\mathbb{Q}[\omega]$ 是否是 $\mathbb{Q}$ 上的线性空间和是否是数域这两者之间的异同。其中一个差异就是前者我们考虑的乘法是 $\mathbb{Q}$ 和 $\mathbb{Q}[\omega]$ 的乘法，而后者我们考虑的是 $\mathbb{Q}[\omega]$ 自身元素的乘法。

   首先，我们需要说明 $\mathbb{Q}[\omega]$ 对加法和乘法封闭；对加法封闭是显然的，对乘法封闭我们需要注意到：$\omega^2 + \omega + 1 = 0$, 那么 $(a + b\omega)(c + d\omega) = (ac - bd) + (ad + dc - bd)\omega$. 我们还需要说明：若 $a + b\omega \neq 0$（这等价于说 a, b 不全为零），则存在 $c + d\omega$ 使得 $(a+b\omega)(c+d\omega) = 1$. 事实上，考虑 $\bar{\omega} = \dfrac{-1-\sqrt{3}\boldsymbol{i}}{2} = 1 - \omega \in \mathbb{Q}[\omega]$, 那么 $a + b\bar{\omega} \in \mathbb{Q}[\omega]$, 且 $(a + b\omega)(a + b\bar{\omega}) = a^2 - ab + b^2$。因为 a, b 不全为零，所以 $a^2 - ab + b^2 \neq 0$, 那么我们令 $c + d\omega = \dfrac{a + b\bar{\omega}}{a^2 - ab + b^2}$ 即可。

□

练习 7.1.3. 把复数域 $\mathbb{C}$ 看作有理数域 $\mathbb{Q}$ 上的线性空间，子集 $\mathbb{R}$ 是否为子空间?

证明*:* 是子空间。我们只需要说明，子集 $\mathbb{R}$ 对以 $\mathbb{Q}$ 为系数的线性组合是封闭的，即：$\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}, a, b \in \mathbb{Q}$, 都有 $a\boldsymbol{x} + b\boldsymbol{y} \in \mathbb{R}$；这一点是明显的。 □

练习 7.1.4. 设 $\mathbb{Q}[\boldsymbol{i}] = \{a + b\boldsymbol{i} \mid a, b \in \mathbb{Q}\}$.

1. 证明 $\mathbb{Q}[\boldsymbol{i}]$ 关于数的加法和数乘构成 $\mathbb{Q}$ 上的一个线性空间.

2. 证明 $\mathbb{Q}[\boldsymbol{i}]$ 是数域.

3. 把复数域 $\mathbb{C}$ 看作数域 $\mathbb{Q}[\boldsymbol{i}]$ 上的线性空间，子集 $\mathbb{R}$ 是否为子空间?

证明*:* 1. 要验证 $\mathbb{Q}[\boldsymbol{i}]$ 是线性空间，也就是要验证它满足线性空间的八条运算法则。按照固定步骤取验证即可，这是容易的，我们略过。

2. 和 $\mathbb{Q}[\omega]$ 的情形类似，我们要验证：$\mathbb{Q}[\boldsymbol{i}]$ 对乘法封闭，并且非零元都有乘法逆元。首先，因为 $\boldsymbol{i}^2 = -1$, 则 $(a + b\boldsymbol{i})(c + d\boldsymbol{i}) = (ac - bd) + (ad + bc)\boldsymbol{i} \in \mathbb{Q}[\boldsymbol{i}]$，所以 $\mathbb{Q}[\boldsymbol{i}]$ 对乘法封闭. 其次，$a + b\boldsymbol{i} \neq 0$，则 a, b 不全为零，并且 $(a + b\boldsymbol{i}) \cdot \dfrac{a - b\boldsymbol{i}}{a^2 + b^2} = 1$，所以 $\mathbb{Q}[\boldsymbol{i}]$ 中非零元都有乘法逆元。

3. 由于子集 $\mathbb{R}$ 对 $\mathbb{Q}[\boldsymbol{i}]$ 上的数乘不封闭，因此不是线性子空间。

□

练习 7.1.5. 设 $\mathcal{V}$ 是以 0 为极限的实数序列全体：$\mathcal{V}=\left\{\{a_n\}\,\middle|\,\lim\limits_{n\to\infty}a_n=0\right\}$. 定义加法和数乘分别为

$$\{a_n\}+\{b_n\}=\{a_n+b_n\};\quad k\{a_n\}=\{ka_n\},k\in\mathbb{R}.$$

证明 $\mathcal{V}$ 是 $\mathbb{R}$ 上的线性空间。

证明: 首先，$\mathcal{V}$ 对加法和数乘都是封闭的。那么验证 $\mathcal{V}$ 满足线性空间的八条运算法则即可。

1. 加法结合律，即要验证 $(\{a_n\}+\{b_n\})+\{c_n\}=\{a_n\}+(\{b_n\}+\{c_n\})$. 事实上，上式左右两边都等于 $\{a_n+b_n+c_n\}$.
2. 加法交换律，即要验证 $a\oplus b=b\oplus a$. 事实上，上式左右两边都等于 $\{a_n+b_n\}$.
3. 零元素，即存在元素 $\mathbf{0}$, 对任意 $a\in\mathbb{R}^+$, $\mathbf{0}\oplus a=a$. 事实上，我们取 $\mathbf{0}$ 是所有项都等于 0 的序列。
4. 负元素。$-\{a_n\}=\{-a_n\}$.
5. 单位数。因为 $1\{a_n\}=\{a_n\}$, 所以 1 是单位数。
6. 数乘结合律，即要验证对任意 $\{a_n\}\in\mathcal{V},k,l\in\mathbb{R}$, 有 $(kl)\{a_n\}=k(l\{a_n\})$. 事实上，上式左右两边都等于 $\{kla_n\}$.
7. 数乘对数的分配律：即要验证对任意 $\{a_n\}\in\mathcal{V},k,l\in\mathbb{R}$, 都有 $(k+l)\{a_n\}=(k\{a_n\})+(l\{a_n\})$. 事实上，上式左右两边都等于 $\{ka_n+la_n\}$.
8. 数乘对向量的分配律：即要验证对任意 $\{a_n\},\{b_n\}\in\mathcal{V},k\in\mathbb{R}$, 都有 $k(\{a_n\}+\{b_n\})=(k\{a_n\})+(k\{b_n\})$. 事实上，上式左右两边都等于 $\{ka_n+kb_n\}$.

□

练习 7.1.6. 考虑 $\mathbb{R}^3$ 中的平面 $S=\left\{\begin{bmatrix}x\\y\\z\end{bmatrix}\,\middle|\,x+y-z=1\right\}$. 对于平面上的任意两点 $\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix},\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}$ 和实数 k，定义加法和数乘分别为

$$\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix}\oplus\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}=\begin{bmatrix}x_1+x_2-1\\y_1+y_2\\z_1+z_2\end{bmatrix};\quad k\otimes\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}kx+1-k\\ky\\kz\end{bmatrix},k\in\mathbb{R}.$$

1. 证明 S 是 $\mathbb{R}$ 上线性空间.
2. 求单射 $f\colon S\to\mathbb{R}^3$，使得 $f(\boldsymbol{v}\oplus\boldsymbol{w})=f(\boldsymbol{v})+f(\boldsymbol{w})$，且 $f(k\otimes\boldsymbol{v})=kf(\boldsymbol{v})$.

可以看到，尽管这个平面不经过原点，但不妨在平面上任取一点“装作”原点，则任意取法都会产生一个线性空间，且不同取法对应不同的线性空间.

证明: 1. 我们首先要说明，S 对加法和数乘封闭。对平面上任意两点 $\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix},\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}$ 和实数 k, 我们有

$$\begin{aligned}
&(x_1+x_2-1)+(y_1+y_2)-(z_1+z_2) &&=(x_1+y_1-z_1)+(x_2+y_2-z_2)-1 &&=1\\
&(kx+1-k)+ky-kz &&=k(x+y-z)+(1-k) &&=1
\end{aligned}$$

这就说明了 S 对加法和数乘封闭。下面，我们需要验证它满足线性空间的八条运算法则。

- 加法结合律, 即要验证 $\left(\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix}\oplus\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}\right)\oplus\begin{bmatrix}x_3\\y_3\\z_3\end{bmatrix}=\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix}\oplus\left(\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}\oplus\begin{bmatrix}x_3\\y_3\\z_3\end{bmatrix}\right)$.

  事实上,

  $$\left(\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix}\oplus\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}\right)\oplus\begin{bmatrix}x_3\\y_3\\z_3\end{bmatrix}=\begin{bmatrix}x_1+x_2-1\\y_1+y_2\\z_1+z_2\end{bmatrix}+\begin{bmatrix}x_3\\y_3\\z_3\end{bmatrix}=\begin{bmatrix}(x_1+x_2-1)+x_3-1\\(y_1+y_2)+y_3\\(z_1+z_2)+z_3\end{bmatrix}$$

  类似地计算等式另外一边，就可以得到它们相等。

- 加法交换律，即要验证 $\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix}\oplus\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}=\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}\oplus\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix}$. 事实上，上式左右两边都等于 $\begin{bmatrix}x_1+x_2-1\\y_1+y_2\\z_1+z_2\end{bmatrix}$.

- 零元素，即存在元素 $\mathbf{0}$, 对任意 $\boldsymbol{v}=\begin{bmatrix}x\\y\\z\end{bmatrix}\in S$, $\mathbf{0}\oplus\boldsymbol{v}=\boldsymbol{v}$. 事实上，$\mathbf{0}=\begin{bmatrix}1\\0\\0\end{bmatrix}$.

- 负元素。$-\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}2-x\\-y\\-z\end{bmatrix}$.

- 单位数。因为 $1\otimes\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}x+1-1\\1y\\1z\end{bmatrix}=\begin{bmatrix}x\\y\\z\end{bmatrix}$, 所以 1 是单位数。

- 数乘结合律, 即要验证对任意 $\boldsymbol{v}=\begin{bmatrix}x\\y\\z\end{bmatrix}\in S, k,\ell\in\mathbb{R}$, 有 $(k\ell)\otimes\boldsymbol{v}=k\otimes(\ell\otimes\boldsymbol{v})$.

  $$k\otimes(\ell\otimes\boldsymbol{v})=k\otimes\begin{bmatrix}\ell x+1-\ell\\\ell y\\\ell z\end{bmatrix}=\begin{bmatrix}k(\ell x+1-\ell)+1-k\\k(\ell y)\\k(\ell z)\end{bmatrix}=\begin{bmatrix}k\ell x+1-k\ell\\k\ell y\\k\ell z\end{bmatrix}=(k\ell)\otimes\begin{bmatrix}x\\y\\z\end{bmatrix}$$

- 数乘对数的分配律: 即要验证对任意 $\boldsymbol{v}=\begin{bmatrix}x\\y\\z\end{bmatrix}\in S, k,\ell\in\mathbb{R}$, 都有 $(k+\ell)\otimes\boldsymbol{v}=$

$(k\otimes \boldsymbol{v})\oplus(\ell\otimes \boldsymbol{v})$. 事实上，

$$\begin{aligned}(k\otimes \boldsymbol{v})\oplus(\ell\otimes \boldsymbol{v}) &= \begin{bmatrix} kx+1-k \\ ky \\ kz \end{bmatrix}\oplus\begin{bmatrix} \ell x+1-\ell \\ \ell y \\ \ell z \end{bmatrix} \\ &= \begin{bmatrix} (kx+1-k)+(\ell x+1-\ell)-1 \\ ky+\ell y \\ kz+\ell z \end{bmatrix} \\ &= \begin{bmatrix} (k+\ell)x+1-(k+\ell) \\ (k+\ell)y \\ (k+\ell)z \end{bmatrix} \\ &= (k+\ell)\otimes \boldsymbol{v}\end{aligned}$$

- 数乘对向量的分配律：即要验证对任意 $\boldsymbol{v}=\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}, \boldsymbol{w}=\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}, k\in\mathbb{R}$, 都有 $k\otimes(\boldsymbol{v}+\boldsymbol{w})=(k\otimes \boldsymbol{v})\oplus(k\otimes \boldsymbol{w})$.

  $$(k\otimes \boldsymbol{v})\oplus(k\otimes \boldsymbol{w})=\begin{bmatrix} kx_1+1-k \\ ky_1 \\ kz_1 \end{bmatrix}\oplus\begin{bmatrix} kx_2+1-k \\ ky_2 \\ kz_2 \end{bmatrix}=\begin{bmatrix} k(x_1+x_2)+1-2k \\ k(y_1+y_2) \\ k(z_1+z_2) \end{bmatrix}$$

  而

  $$k\otimes(\boldsymbol{v}+\boldsymbol{w})=k\otimes\begin{bmatrix} x_1+x_2-1 \\ y_1+y_2 \\ z_1+z_2 \end{bmatrix}=\begin{bmatrix} k(x_1+x_2-1)+1-k \\ k(y_1+y_2) \\ k(z_1+z_2) \end{bmatrix}$$

  因此两者相等。

□

练习 7.1.7. 对 n 阶方阵 A，令 $P(A)=\{f(A)\mid f(x)\in\mathbb{F}[x]\}$. 证明 $P(A)$ 关于矩阵的加法和数乘构成 $\mathbb{F}$ 上的线性空间.

证明: 这实际上是要证明 $P(A)$ 是 $M_{n\times n}(\mathbb{R})$ 的子空间；等价于证明：$\forall X,Y\in P(A), a,b\in\mathbb{R}, aX+bY\in P(A)$.

$X,Y\in P(A)$, 则存在 $f(x),g(x)\in\mathbb{F}[x]$ 使得 $X=f(A), Y=g(A)$, 那么令 $h(x)=af(x)+bg(x)$, 则 $aX+bY=h(A)\in P(A)$. □

练习 7.1.8. 证明加法交换律在线性空间的定义中冗余，即由其他七条运算法则可以推出加法交换律.

证明: 对任意 $\boldsymbol{a},\boldsymbol{b}\in\mathcal{V}$, 根据第 7 条和第 8 条，我们有

$$(\boldsymbol{a}+\boldsymbol{b})+(\boldsymbol{a}+\boldsymbol{b})=2(\boldsymbol{a}+\boldsymbol{b})=2\boldsymbol{a}+2\boldsymbol{b}$$

根据第 4 条，上式左边加 $-\boldsymbol{a}$，右边加 $-\boldsymbol{b}$，再根据第 1 条和第 8 条，就有

$$\boldsymbol{b}+\boldsymbol{a}=-\boldsymbol{a}+(\boldsymbol{a}+\boldsymbol{b})+(\boldsymbol{a}+\boldsymbol{b})+(-\boldsymbol{b})=(-\boldsymbol{a})+2\boldsymbol{a}+2\boldsymbol{b}+(-\boldsymbol{b})=\boldsymbol{a}+\boldsymbol{b}.$$

□

练习 7.1.9. 对 n 阶方阵 A，令 $\mathrm{Com}(A)=\{n \text{ 阶方阵 } B \mid AB=BA\}$.

1. 证明，$\mathrm{Com}(A)$ 是 $\mathbb{F}^{n\times n}$ 的子空间.
2. 证明，对任意 $B,C\in\mathrm{Com}(A)$，都有 $BC\in\mathrm{Com}(A)$；由此证明对任意多项式 $f(x)$，都有 $f(A)\in\mathrm{Com}(A)$.

证明:

1. 我们只需要证明：对任意 $B,C\in\mathrm{Com}(A), x,y\in\mathbb{F}$, 都有 $xB+yC\in\mathrm{Com}(A)$. 这是因为：

$$A(xB+yC)=xAB+yAC=xBA+yCA=(xB+yC)A.$$

2. 若 $B,C\in\mathrm{Com}(A)$, 则 $A(BC)=(AB)C=(BA)C=B(AC)=(BC)A$, 所以 $BC\in\mathrm{Com}(A)$. 那么对任意 A, 显然 $A\in\mathrm{Com}(A)$, 则对任意 $a\in\mathbb{F}, n\geq 0$ 都有 $aA^n\in\mathrm{Com}(A)$; 那么再由第 1 小题知，任意 A 的单项式的线性组合都在 $\mathrm{Com}(A)$ 中，因此 $f(A)\in\mathrm{Com}(A)$.

□

练习 7.1.10. 考虑矩阵空间 $\mathbb{F}^{n\times n}$ 的分别满足以下条件的矩阵 A 的全体，判断是否为子空间.

1. A 可逆.
2. A 不可逆.
3. $A^2=0$.
4. $AA^T=I$.
5. $\mathcal{R}(A)=\mathcal{M}$，其中 $\mathcal{M}$ 是 $\mathbb{F}^n$ 的给定子空间.
6. $\mathcal{N}(A)=\mathcal{M}$，其中 $\mathcal{M}$ 是 $\mathbb{F}^n$ 的给定子空间.
7. 存在矩阵 B,C，使得 $A=\begin{bmatrix}C & -B\\ B & C\end{bmatrix}$.

证明:

1. 不是子空间，不包含零元素，对于加法也不封闭。
2. 不是子空间，对加法不封闭，例如 $\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}$ 都不可逆，但是它们的和 $\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ 可逆。
3. 不是子空间，对加法不封闭，例如 $\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}$ 的平凡都是零矩阵，但是它们的和 $\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ 是可逆矩阵，不可能幂零。
4. 不是子空间，不包含零元素。
5. 只有 $\mathcal{M}=\{0\}$ 时，才是子空间。因为若不然，不包含零元素。
6. 只有 $\mathcal{M}=\mathbb{F}^n$ 时，才是子空间。若不然，不包含零元素。
7. 是子空间。直接做矩阵运算可以验证这个空间对线性组合封闭。

□

练习 7.1.11. 考虑多项式空间 $\mathbb{R}[x]$ 的分别满足以下条件的多项式 p 的全体，判断是否为子空间.

1. $p(2)=0$. *2.* $p(2)=1$. *3.* $p'(2)=0$. *4.* p 的次数是奇数. *5.* p 的根都是实数.

证明: 1. 是子空间。可以直接验证其对线性组合封闭。

2. 不是子空间，不包含零元素。
3. 是子空间。$(ap(x)+bq(x))'=ap'(x)+bq'(x)$, 所以对线性组合封闭。
4. 不是子空间。例如：$x^3+x^2, -x^3$ 都是次数是奇数的多项式，但是它们的和是 x^2，是偶数次的，因此对线性组合不封闭。
5. 不是子空间。例如：$x^2, x+1$ 的根都是实数，但是它们的和 x^2+x+1 没有实数根。

□

练习 7.1.12. 设 $\mathcal{M}_1=\operatorname{Span}(\boldsymbol{a}_1,\boldsymbol{a}_2), \mathcal{M}_2=\operatorname{Span}(\boldsymbol{b}_1,\boldsymbol{b}_2)$，其中

$$\boldsymbol{a}_1=\begin{bmatrix}1\\-1\\0\\1\end{bmatrix},\quad \boldsymbol{a}_2=\begin{bmatrix}-2\\3\\1\\-3\end{bmatrix},\quad \boldsymbol{b}_1=\begin{bmatrix}1\\2\\0\\-2\end{bmatrix},\quad \boldsymbol{b}_2=\begin{bmatrix}1\\3\\1\\-3\end{bmatrix}.$$

分别求 $\mathcal{M}_1+\mathcal{M}_2, \mathcal{M}_1\cap\mathcal{M}_2$ 的一组基和维数.

证明: 由于 $\mathcal{M}_1+\mathcal{M}_2=\operatorname{Span}(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{b}_1,\boldsymbol{b}_2)$，那么我们只要求向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{b}_1,\boldsymbol{b}_2$ 的一个极大线性无关组即可。基本方法还是把它们排成一个矩阵，作初等行变换，求出主元所在列。

我们对矩阵 $A=\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{b}_1 & \boldsymbol{b}_2\end{bmatrix}=\begin{bmatrix}1&-2&1&1\\-1&3&2&3\\0&1&0&1\\1&-3&-2&-3\end{bmatrix}$ 做初等行变换得行简化阶梯形矩阵为

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

主元列是第 1,2,3 列，所以 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{b}_1$ 就构成了 $\mathcal{M}_1+\mathcal{M}_2$ 的一组基。

另一方面，根据维数公式

$$\dim(\mathcal{M}_1+\mathcal{M}_2)+\dim(\mathcal{M}_1\cap\mathcal{M}_2)=\dim(\mathcal{M}_1)+\dim(\mathcal{M}_2)$$

知 $\dim(\mathcal{M}_1\cap\mathcal{M}_2)=1$. 方程组 $A\boldsymbol{x}=0$ 的一组基为 $\begin{bmatrix}-2\\-1\\-1\\1\end{bmatrix}$, 即 $-2\boldsymbol{a}_1-\boldsymbol{a}_2-\boldsymbol{a}_3+\boldsymbol{a}_4=0$，则

$$2\boldsymbol{a}_1+\boldsymbol{a}_2=-\boldsymbol{a}_3+\boldsymbol{a}_4=\begin{bmatrix}0\\1\\1\\-1\end{bmatrix}\in\mathcal{M}\cap\mathcal{N}\text{ 是一组基。}$$ □

练习 7.1.13. 设 $\mathbb{F}_0^{n\times n}$ 是矩阵空间 $\mathbb{F}^{n\times n}$ 中所有迹为零的矩阵构成的子集.

1. 证明 $\mathbb{F}_0^{n\times n}$ 是 $\mathbb{F}^{n\times n}$ 的子空间.
2. 求子空间 $\mathbb{F}_0^{n\times n}$ 和 $\mathrm{Span}(I_n):=\{kI_n\mid k\in\mathbb{F}\}$ 的交与和.

证明: 1. 对于 $A,B\in\mathbb{F}_0^{n\times n},a,b\in\mathbb{F}$, 那么

$$\mathrm{trace}(aA+bB)=a\mathrm{trace}(A)+b\mathrm{trace}(B)=0$$

所以 $\mathbb{F}_0^{n\times n}$ 在线性组合之下封闭，因此是子空间。

2. 设 $kI_n\in\mathbb{F}_0^{n\times n}$，那么 $\mathrm{trace}(kI_n)=0$, 即 $nk=0$, 所以 $k=0$, 这就说明 $\mathbb{F}_0^{n\times n}\cap\mathrm{Span}(I_n)=\{0\}$.
另外一方面，对任意 $A\in\mathbb{F}^{n\times n}$, 令 $A_0=A-\frac{\mathrm{trace}(A)}{n}I_n$, 那么 $\mathrm{trace}(A_0)=0$, 且 $A=A_0+\frac{\mathrm{trace}(A)}{n}I_n$. 因此 $\mathbb{F}_0^{n\times n}+\mathrm{Span}(I_n)=\mathbb{F}^{n\times n}$.

□

练习 7.1.14. 考虑函数空间 $C(\mathbb{R})$ 的如下子集：

$$\mathcal{V}=\{f\mid f''+3f'+2f=0\},\quad \mathcal{M}=\{f\mid f'+2f=0\},\quad \mathcal{N}=\{f\mid f'+f=0\}.$$

1. 证明 $\mathcal{V},\mathcal{M},\mathcal{N}$ 都是 $C(\mathbb{R})$ 的子空间，且 $\mathcal{M},\mathcal{N}$ 是 $\mathcal{V}$ 的子空间.
2. 描述 $\mathcal{M},\mathcal{N}$ 中的所有元素，并证明 $\mathcal{M}\cap\mathcal{N}=\{0\}$.
3. 对任意 $f\in\mathcal{V}$，证明 $f'+f\in\mathcal{M},f'+2f\in\mathcal{N}$.
4. 证明 $\mathcal{M}\oplus\mathcal{N}=\mathcal{V}$.
5. 设 $f\in\mathcal{V}$，且满足 $f(0)=1,f'(0)=2$，求 f.

证明: 1. 我们只需验证 $\mathcal{V},\mathcal{M},\mathcal{N}$ 在线性组合之下都封闭，而这一点由求导的线性性保证，我们略过。对任意 $f\in\mathcal{M}$, $f'+2f=0$, 求导有 $f''+2f'=0$, 相加就有 $f''+3f'+2f=0$, 所以 $f\in\mathcal{V}$, 这就有 $\mathcal{M}\subseteq\mathcal{V}$. 同理可得 $\mathcal{N}\subseteq\mathcal{V}$.

2. $\mathcal{M}=\mathrm{Span}(e^{-2x}),\mathcal{N}=\mathrm{Span}(e^{-x})$, 显然 $\mathcal{M}\cap\mathcal{N}=\{0\}$.

3. 对 $f\in\mathcal{V}$, 令 $g=f'+f$, 那么 $g'=f''+f'$, 那么 $g'+2g=f''+f'+2(f'+f)=f''+3f'+2f=0$, 即 $g\in\mathcal{M}$. 同理可证 $f'+2f\in\mathcal{N}$.

4. 对任意 $f\in\mathcal{V}$, $f=(-(f'+f))+(f'+2f)\in\mathcal{M}+\mathcal{N}$, 再结合第 2 小题就有 $\mathcal{M}\oplus\mathcal{N}=\mathcal{V}$

5. 由题意，$f'+f=c_1e^{-2x},f'+2f=c_2e^{-x}$. 代入初值有 $c_1=3,c_2=4$, 所以 $f=4e^{-x}-3e^{-2x}$

□

7.2 基和维数

练习 7.2.1. 把数域 $\mathbb{F}$ 看作自身上的线性空间，求它的一组基和维数.

证明: 1 是一组基，维数是 1. □

练习 7.2.2. 求练习 *7.1.1*中线性空间 $\mathbb{R}^+$ 的一组基和维数.

证明: 因为任意 $a\in\mathbb{R}^+$, 我们都有 $a=\ln(a)\otimes e$, 因此维数是 1, e 是一组基。 □

练习 7.2.3. 在练习 *7.1.2*中的线性空间 $\mathbb{Q}[\omega]=\{a+b\omega\mid a,b\in\mathbb{Q}\}$ 内：

1. 求下列向量组的秩：$S_1:\dfrac{1}{2},3,-7$，$S_2:1,\omega,\omega^2,\omega^3$，$S_3:\omega,\overline{\omega},\sqrt{3}\boldsymbol{i}$.
2. 求 $\mathbb{Q}[\omega]$ 的一组基和维数.

证明: 我们需要强调的是系数域是 $\mathbb{Q}$.

1. $\mathrm{rank}(S_1)=1,\mathrm{rank}(S_2)=2,\mathrm{rank}(S_3)=2$.
2. 它的一组基是 $1,\omega$, 维数是 2.

□

练习 7.2.4. 判断练习 *7.1.5*中线性空间的维数是否有限.

证明: 维数不有限。对任意 $k\geq 1$, 我们定义序列 $A_k=(a_{1,k},a_{2,k},a_{3,k},\cdots,)$ 满足

$$a_{1,k}=\cdots=a_{k,k}=1,\quad a_{k+1,k}=a_{k+2,k}=\cdots=0$$

显然所有的 A_k 都属于 $\mathcal{V}$, 但是它们并不线性相关。 □

练习 7.2.5. 判断 $\mathbb{R}$ 上的线性空间 $C[-\pi,\pi]$ 内的下列向量组是否线性相关，并求其秩：

1. $\cos^2 x,\sin^2 x$.
2. $\cos^2 x,\sin^2 x,1$.
3. $\cos 2x,\sin 2x$.
4. $1,\sin x,\sin 2x,\ldots,\sin nx$.
5. $1,\sin x,\sin^2 x,\ldots,\sin^n x$.

证明:

1. 设 $t_1\cos^2 x+t_2\sin^2 x=0$; 取 $x=0$ 则推出 $t_1=0$，取 $x=\frac{\pi}{2}$ 则推出 $t_2=0$, 所以 $\cos^2 x,\sin^2 x$ 线性无关，秩是 2.
2. 线性相关，因为 $\cos^x+\sin^2 x-1=0$. 那么秩小于等于 2，再由第 1 小题知秩大于等于 2，所以秩就是 2.
3. 设 $t_1\cos 2x+t_2\sin 2x=0$, 取 $x=0$ 则推出 $t_1=0$，再取 $x=\frac{\pi}{4}$ 则推出 $t_2=0$，所以秩是 2.
4. 设 $t_0+t_1\sin x+t_2\sin 2x+\cdots+t_n\sin nx=0$, 取 $x=0$ 则推出 $t_0=0$. 于是 $t_1\sin x+t_2\sin 2x+\cdots+t_n\sin nx=0$. 我们对此求两次导数即有：

$$(-1)t_1\sin x+(-2)t_2\sin 2x+\cdots+(-n)t_n\sin nx=0$$

再求 2 次，4 次，6 次... 导数，可得

$$\begin{bmatrix}1&1&\cdots&1\\-1&-2&\cdots&-n\\\vdots&&\vdots&\\(-1)^{n-1}&(-2)^{n-1}&\cdots&(-n)^{n-1}\end{bmatrix}\begin{bmatrix}t_1\sin x\\t_2\sin 2x\\\vdots\\t_n\sin nx\end{bmatrix}=0$$

系数矩阵是 Vondermonde 矩阵，可逆，等式两边左乘以矩阵的逆可得

$$t_1 \sin x = t_2 \sin 2x = \cdots = t_n \sin nx = 0,$$

故 $t_1 = \cdots = t_n = 0$，所以 $1, \sin x, \sin 2x, \ldots, \sin nx$ 线性无关，秩为 $n+1$.

5. 设 $t_0 + t_1 \sin x + t_2 \sin^2 x + \cdots + t_n \sin^n x = 0$, 取 $x = 0$ 则推出 $t_0 = 0$; 然后我们分别取 $\sin x = 1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{n}$, 则我们有

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \frac{1}{2} & \frac{1}{2^2} & \cdots & \frac{1}{2^n} \\ \vdots & & \vdots & \\ \frac{1}{n} & \frac{1}{n^2} & \cdots & \frac{1}{n^n} \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix} = 0$$

系数矩阵是 $\operatorname{diag}(1, 1/2, \cdots, 1/n)$ 和 Vondermonde 矩阵的乘积，所以可逆，因此 $t_1 = \cdots = t_n = 0$, 故 $1, \sin x, \sin^2 x, \ldots, \sin^n x$ 线性无关，秩为 $n+1$.

□

练习 7.2.6. 考虑练习 *7.1.9*中的线性空间 $\operatorname{Com}(A)$，对下列 A 求 $\operatorname{Com}(A)$ 的一组基.

1. $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

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

3. $\begin{bmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}_{n \times n}$.

4. $\operatorname{diag}(a_i)$，其中 a_i 各不相同.

5. $\operatorname{diag}(a_i)$.

证明*:* 1. $A = J_3(0)$, 直接进行矩阵运算可得 $\operatorname{Com}(A)$ 的一组基是 $I_3, J_3(0), J_3(0)^2$.

2. $A = I_3 + J_3(0)$, 那么 $\operatorname{Com}(A) = \operatorname{Com}(J_3(0))$, 所以它的一组基是 $I_3, J_3(0), J_3(0)^2$.

3. 还是直接进行矩阵运算，不过就是复杂一点。$A = J_n(0)$, 设 $B = \begin{bmatrix} b_{ij} \end{bmatrix} \in \operatorname{Com}(A)$, 那么 AB 是把去掉 B 的第一行，把下面的行往上移一行，最后一行是零行；同样地通过列来考虑 BA。那么比较 $AB = BA$ 的每一个分量，就可以得到 B 是上三角矩阵，且对角线方向上的同一条线上的元素都相等，所以 $\operatorname{Com}(A)$ 的一组基是

$$I_n, J_n(0), J_n(0)^2, \cdots, J_n(0)^{n-1}.$$

4. 设 $B = \begin{bmatrix} b_{ij} \end{bmatrix} \in \operatorname{Com}(A)$, $AB = \begin{bmatrix} a_i b_{ij} \end{bmatrix}, BA = \begin{bmatrix} a_j b_{ij} \end{bmatrix}$; 比较 $AB = BA$ 的每一个分量就知道 B 是对角矩阵，所以 $\operatorname{Com}(A)$ 的一组基是 $\{E_{ii} : 1 \leq i \leq n\}$, 其中 E_{ii} 是第 i 个对角线为 1, 其他所有对角线都为零的矩阵。

5. 我们写出 $\operatorname{Com}(A)$; 而去描述它的一组基，语言叙述稍微麻烦一些，我们略过。假设 $a_1, \cdots, a_r$ 是对角线上互不相同的元素，且分别出现了 $n_1, \cdots, n_r$ 次，那么存在置换矩阵 P 使得 $\tilde{A} := PAP = \operatorname{diag}(a_1 I_{n_1}, \cdots, a_r I_{n_r})$. 那么 $B \in \operatorname{Com}(A)$ 当且仅当 $PBP \in \operatorname{Com}(\tilde{A})$. 我们利用分块矩阵乘法可得 $\operatorname{Com}(\tilde{A}) = \operatorname{diag}(X_1, \cdots, X_r)$，其中 $X_i \in M_{n_i \times n_i}(\mathbb{F})$. 那么 $\operatorname{Com}(A) = P\operatorname{diag}(X_1, \cdots, X_r)P$，其中 $X_i \in M_{n_i \times n_i}(\mathbb{F})$

□

练习 7.2.7. 给定 $\mathbb{F}$ 中两两不等的数 $a_1, a_2, \ldots, a_n$.

1. 在线性空间 $\mathbb{F}[x]_n$ 中，令

$$f_i(x) = (x - a_1) \cdots \widehat{(x - a_i)} \cdots (x - a_n), \quad i = 1, 2, \ldots, n,$$

   其中 $\widehat{(x - a_i)}$ 表示不含该项. 证明，$f_1(x), f_2(x), \ldots, f_n(x)$ 是 $\mathbb{F}[x]_n$ 的一组基.

2. 设 $b_1, b_2, \ldots, b_n$ 是 $\mathbb{F}$ 中任意 n 个数，找出 $f(x) \in \mathbb{F}[x]_n$，使得 $f(a_i) = b_i, i = 1, 2, \ldots, n$.

证明:

设 $t_1 f_1(x) + t_2 f_2(x) + \cdots + t_n f_n(x) = 0$, 取 $x = a_i$，因为 $f_i(a_i) \neq 0, f_j(a_i) = 0 (\forall j \neq i)$, 则 $t_i = 0$, 因此 $f_1(x), \cdots, f_n(x)$ 线性无关。另外一方面，显然 $1, x, \cdots, x^{n-1}$ 是 $\mathbb{F}[x]_n$ 的一组基，所以 $\dim_{\mathbb{F}} \mathbb{F}[x]_n = n$。那么 $f_1(x), \cdots, f_n(x)$ 也是 $\mathbb{F}[x]_n$ 的一组基。

我们可以思考这个问题：如果我们能找到 $g_i(x)$ 使得 $g_i(a_i) = 1, g_i(a_j) = 0 (\forall j \neq i)$, 那么我们令 $f(x) = b_1 g_1(x) + \cdots + b_n g_n(x)$ 即可。按照这个思路，答案就明显了：

$$f(x) = \sum_{i=1}^{n} \frac{b_i}{f_i(a_i)} f_i(x).$$

□

练习 7.2.8. 证明，n 维线性空间中任意多于 n 个的向量都线性相关.

证明: 取该线性空间的一组基 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$, 设 $m > n$ 且 $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m$ 是 m 个向量，那么存在

$$\begin{aligned}
\boldsymbol{w}_1 &= a_{11}\boldsymbol{v}_1 + a_{21}\boldsymbol{v}_2 + \cdots + a_{n1}\boldsymbol{v}_n \\
\boldsymbol{w}_2 &= a_{12}\boldsymbol{v}_1 + a_{22}\boldsymbol{v}_2 + \cdots + a_{n2}\boldsymbol{v}_n \\
\vdots \quad & \quad \vdots \\
\boldsymbol{w}_m &= a_{1m}\boldsymbol{v}_1 + a_{2m}\boldsymbol{v}_2 + \cdots + a_{nm}\boldsymbol{v}_n
\end{aligned}$$

令 $A = \begin{bmatrix} a_{ij} \end{bmatrix} \in M_{n \times m}$, 因为 $m > n$, 那么 $A\boldsymbol{x} = 0$ 存在非零解 $\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$, 于是 $x_1\boldsymbol{w}_1 + x_2\boldsymbol{w}_2 + \cdots + x_m\boldsymbol{w}_m = 0$, 即 $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m$ 线性相关。 □

练习 7.2.9. 考虑练习 7.1.7中的线性空间 $P(A)$.

1. 判断其维数是否有限.
2. 证明存在次数不大于 n^2 的多项式 $f(x) \in \mathbb{F}[x]$，使得 $f(A) = 0$.
3. 令 $A = \operatorname{diag}(1, \omega, \omega^2)$，其中 $\omega = \dfrac{-1 + \sqrt{3}\boldsymbol{i}}{2}$，求 $P(A)$ 的维数和一组基.

证明:

1. 维数有限，因为它是 $M_{n \times n}(\mathbb{F})$ 的子空间。
2. 考虑 $I_n, A, \cdots, A^{n^2}$ 一共是 $n^2 + 1$ 个向量，$M_{n \times n}(\mathbb{F})$ 维数是 n^2, 所以这 $n^2 + 1$ 个向量线性相关，即存在不全为零的 $a_0, a_1, \cdots, a_{n^2}$ 使得 $a_0 I_n + a_1 A + \cdots + a_{n^2} A^{n^2} = 0$. 取 $f(x) = a_0 + a_1 x + \cdots + a_{n^2} x^{n^2}$ 即可。

3. $P(A)=\mathrm{Span}(A^k : k\geq 0)$. 注意到 $A^{k+3}=A^k$, 所以 $P(A)=\mathrm{Span}(I_3,A,A^2)$. 而 I_3,A,A^2 线性无关，所以它们是 $P(A)$ 的一组基，$P(A)$ 的维数是 3.

$\square$

练习 7.2.10. 证明连续函数空间的子集 $V:=\{f(x)=k_0+k_1\cos x+k_2\cos 2x \mid f(0)=0\}$ 是一个子空间，并求一组基.

证明: 我们验证 V 在线性组合之下封闭。设 $f_1(x)=k_0+k_1\cos x+k_2\cos 2x, f_2(x)=k_0'+k_1'\cos x+k_2'\cos 2x$ 都属于 V，$a,b\in\mathbb{R}$, $f=af_1(x)+bf_2(x)$, 那么

$$f(x)=(ak_0+bk_0')+(ak_1+bk_1')\cos x+(ak_2+bk_2')\cos 2x,$$

且 $f(0)=af_1(0)+bf_2(0)=0$, 所以 $f(0)=0$.

对 $f(x)=k_0+k_1\cos x+k_2\cos 2x\in V$, $f(0)=0$ 等价于 $k_0+k_1+k_2=0$, 所以 $f(x)=k_1(\cos x-1)+k_2(\cos 2x-1)$, 所以 $V=\mathrm{Span}(\cos x-1,\cos 2x-1)$, 即 $\cos x-1,\cos 2x-1$ 是 V 的一组生成元集；我们还需要说明它们线性无关：设 $k_1(\cos x-1)+k_2(\cos 2x-1)=0$, 分别取 $x=\frac{\pi}{2},\pi$ 可得

$$-k_1-2k_2=0,-2k_1=0$$

因此 $k_1=k_2=0$, 这说明 $\cos x-1,\cos 2x-1$ 线性无关，从而他们是 V 的一组基。 $\square$

练习 7.2.11. 设 $\mathcal{M}_1,\mathcal{M}_2$ 是 $\mathbb{R}^n$ 的两个子空间，且 $\mathcal{M}_1\subseteq\mathcal{M}_2$. 证明，如果 $\dim\mathcal{M}_1=\dim\mathcal{M}_2$，则 $\mathcal{M}_1=\mathcal{M}_2$.

证明: $\dim\mathcal{M}_1=\dim\mathcal{M}_2$ 说明 $\mathcal{M}_1$ 的基也是 $\mathcal{M}_2$ 的基，所以若 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ 是 $\mathcal{M}_1$ 的基，那么 $\mathcal{M}=\mathrm{Span}(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n)=\mathcal{M}_1$. $\square$

练习 7.2.12. 设 $\mathcal{M}$ 是 $\mathbb{R}^n$ 的子空间，$\mathcal{M}^\perp$ 是其正交补空间，证明，$\mathbb{R}^n=\mathcal{M}\oplus\mathcal{M}^\perp$.

证明: 我们取 $\mathcal{M}$ 的一组基 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r$, 那么 $\boldsymbol{x}\in\mathcal{M}^\perp$ 当且仅当 $\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{x}=0$, 即 $\boldsymbol{x}$ 是方程组 $\begin{bmatrix}\boldsymbol{v}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{v}_r^{\mathrm{T}}\end{bmatrix}\boldsymbol{x}=0$ 的解，这说明 $\mathcal{M}^\perp=\mathcal{N}(A)$, 其中 $A=\begin{bmatrix}\boldsymbol{v}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{v}_r^{\mathrm{T}}\end{bmatrix}$, 根据维数公式 $\dim\mathcal{M}^\perp=n-r$.

另一方面，若 $\boldsymbol{x}\mathcal{M}\cap\mathcal{M}^\perp$, 则 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=0$, 则 $\boldsymbol{x}=0$, 所以 $\mathcal{M}$ 和 $\mathcal{M}^\perp$ 是直和。因为 $\mathcal{M}\oplus\mathcal{M}^\perp\subseteq\mathbb{R}^n$, 且 $\dim\mathcal{M}\oplus\mathcal{M}^\perp=\dim\mathcal{M}+\dim\mathcal{M}^\perp=n=\dim\mathbb{R}^n$, 所以 $\mathcal{M}\oplus\mathcal{M}^\perp=\mathbb{R}^n$. $\square$

练习 7.2.13. 证明，练习 *7.1.13*中的 $\mathbb{F}^{n\times n}=\mathbb{F}_0^{n\times n}\oplus\mathrm{Span}\, I_n$，并求 $\dim\mathbb{F}_0^{n\times n}$.

证明: 练习 7.1.13等价于在证明 $\mathbb{F}^{n\times n}=\mathbb{F}_0^{n\times n}\oplus\mathrm{Span}\, I_n$, 我们不再重复。

$\dim\mathbb{F}_0^{n\times n}=\dim\mathbb{F}^{n\times n}-\dim\mathrm{Span}(I_n)=n^2-1.$ $\square$

7.3 线性映射

练习 7.3.1. 证明命题 *7.3.4*：集合 $\mathrm{Hom}(\mathcal{U},\mathcal{V})$ 关于加法和数乘两种运算构成线性空间。

证明: 我们需要说明的是，两个映射相等等价于说它们在每个点的取值都相等。

1. 加法结合律，即要验证 $(f+g)+h=f+(g+h)$. 对任意 $\boldsymbol{u}\in\mathcal{U}$, $((f+g)+h)(\boldsymbol{u})=(f+g)(\boldsymbol{u})+h(\boldsymbol{u})=f(\boldsymbol{u})+g(\boldsymbol{u})+h(\boldsymbol{u})$, $(f+(g+h))(\boldsymbol{u})=f(\boldsymbol{u})+(g+h)(\boldsymbol{u})=f(\boldsymbol{u})+g(\boldsymbol{u})+h(\boldsymbol{u})$, 所以 $(f+g)+h=f+(g+h)$.
2. 加法交换律，即要验证 $f+g=g+f$. 事实上，上式左右两边在任意 $\boldsymbol{u}$ 上取值都等于 $f(\boldsymbol{u})+g(\boldsymbol{u})$.
3. 零元素，即存在元素 $\mathbf{0}$, 对任意 $f\in\mathrm{Hom}(\mathcal{U},\mathcal{V})$, $\mathbf{0}\oplus f=f$. 事实上，我们取 $\mathbf{0}$ 是在所有元素上取值都等于 0 的映射。
4. 负元素。对任意 $f\in\mathrm{Hom}(\mathcal{U},\mathcal{V})$, $-f$ 定义为映射 $\boldsymbol{u}\mapsto -f(\boldsymbol{u})$, 即 $(-f)(\boldsymbol{u})=-f(\boldsymbol{u})$.
5. 单位数。因为 $(1\cdot f)(\boldsymbol{u})=1\cdot f(\boldsymbol{u})=f(\boldsymbol{u})$, 所以 $1\cdot f=f$.
6. 数乘结合律，即要验证对任意 $f,g\in\mathrm{Hom}(\mathcal{U},\mathcal{V}),k,l\in\mathbb{F}$, 有 $(kl)f=k(lf)$. 事实上，上式左右两边在任意的 $\boldsymbol{u}$ 上取值都等于 $klf(\boldsymbol{u})$.
7. 数乘对数的分配律：即要验证对任意 $f\in\mathrm{Hom}(\mathcal{U},\mathcal{V}),k,l\in\mathbb{R}$, 都有 $(k+l)f=(kf)+(lf)$. 事实上，上式左右两边在任意的 $\boldsymbol{u}$ 上取值都等于 $kf(\boldsymbol{u})+lf(\boldsymbol{u})$.
8. 数乘对向量的分配律：即要验证对任意 $f,g\in\mathrm{Hom}(\mathcal{U},\mathcal{V}),k\in\mathbb{R}$, 都有 $k(f+g)=(kf)+(kg)$. 事实上，上式左右两边在任意的 $\boldsymbol{u}$ 上取值都等于 $kf(\boldsymbol{u})+kg(\boldsymbol{u})$.

□

练习 7.3.2. *证明命题 7.3.7：对 $\mathcal{U}$ 到 $\mathcal{V}$ 的线性映射 f, 有*

1. $\mathcal{N}(f)$ *是 $\mathcal{U}$ 的子空间，而且 f 是单射当且仅当* $\mathcal{N}(f)=\{0\}$.

2. $\mathcal{R}(f)$ *是 $\mathcal{V}$ 的子空间，而且 f 是满射当且仅当* $\mathcal{R}(f)=\mathcal{V}$.

证明*:* 1. 对任意 $\boldsymbol{u}_1,\boldsymbol{u}_2\in\mathcal{N}(f),a,b\in\mathbb{F}$, $f(a\boldsymbol{u}_1+b\boldsymbol{u}_2)=af(\boldsymbol{u}_1)+bf(\boldsymbol{u}_2)=0$, 所以 $a\boldsymbol{u}_1+b\boldsymbol{u}_2\in\mathcal{N}(f)$. f 是单射，当且仅当

$$\text{若 } f(\boldsymbol{u}_1)=f(\boldsymbol{u}_2) \text{ 则 } \boldsymbol{u}_1=\boldsymbol{u}_2$$

当且仅当

$$\text{若 } f(\boldsymbol{u}_1-\boldsymbol{u}_2)=0 \text{ 则 } \boldsymbol{u}_1-\boldsymbol{u}_2=0$$

当且仅当

$$\text{若 } f(\boldsymbol{u})=0 \text{ 则 } \boldsymbol{u}=0; \text{ 即 } \mathcal{N}(f)=0.$$

2. 对任意 $\boldsymbol{v}_1,\boldsymbol{v}_2\in\mathcal{R}(f),a,b\in\mathbb{F}$, 那么存在 $\boldsymbol{u}_1,\boldsymbol{u}_2\in\mathcal{U}$ 使得 $\boldsymbol{v}_1=f(\boldsymbol{u}_1),\boldsymbol{v}_2=f(\boldsymbol{u}_2)$, 那么 $a\boldsymbol{v}_1+b\boldsymbol{v}_2=f(a\boldsymbol{u}_1+b\boldsymbol{u}_2)$, 所以 $a\boldsymbol{v}_1+b\boldsymbol{v}_2\in\mathcal{R}(f)$.
   f 是满射，当且仅当对任意 $\boldsymbol{v}\in\mathcal{V}$, 都存在 $\boldsymbol{u}\in\mathcal{U}$ 使得 $\boldsymbol{v}=f(\boldsymbol{u})$, 这等价于说 $\mathcal{V}\subseteq\mathcal{R}(f)$。结合 $\mathcal{R}(f)\subseteq\mathcal{V}$ 可知：f 是满射当且仅当 $\mathcal{R}(f)=\mathcal{V}$.

□

练习 7.3.3. *考虑 $\mathbb{C}^n$ 上的变换* $\mathbf{C}(\boldsymbol{v})=\overline{\boldsymbol{v}}$.

1. $\mathbf{C}$ *是否为一个 $\mathbb{C}$ 上线性空间的线性变换？*

2. *如果将 $\mathbb{C}^n$ 看作一个 $\mathbb{R}$ 上的线性空间，那么 $\mathbf{C}$ 是否为一个 $\mathbb{R}$ 上线性空间的线性变换？*

证明: 1. 不是。我们取 $a=\boldsymbol{i}, 0\neq \boldsymbol{v}\in\mathbb{C}^n$, $\mathbf{C}(a\boldsymbol{v})=\bar{a}\mathbf{C}\boldsymbol{v}=-a\overline{\boldsymbol{v}}$, 而不等于 $a\mathbf{C}(\boldsymbol{v})$。

2. 是。对任意 $a,b\in\mathbb{R},\boldsymbol{v}_1,\boldsymbol{v}_2\in\mathbb{C}^n$,

$$\mathbf{C}(a\boldsymbol{v}_1+b\boldsymbol{v}_2)=\overline{a\boldsymbol{v}_1+b\boldsymbol{v}_2}=\overline{a\boldsymbol{v}_1}+\bar{b}\overline{\boldsymbol{v}_2}$$

此时 $\overline{a}=a,\bar{b}=b$, 所以 $\mathbf{C}(a\boldsymbol{v}_1+b\boldsymbol{v}_2)=a\mathbf{C}(\boldsymbol{v}_1)+b\mathbf{C}(\boldsymbol{v}_2)$.

□

练习 7.3.4. 给定 $a\in\mathbb{F}$，判断下面定义的 $\mathbb{F}[x]$ 上的变换 $\mathbf{T}_a$ 是否为线性变换：

$$\mathbf{T}_a(f(x))=f(x+a),\quad \forall f(x)\in\mathbb{F}[x].$$

证明: 是线性映射。对任意 $f_1(x),f_2(x)\in\mathbb{F}[x],t_1,t_2\in\mathbb{F}$,

$$\mathbf{T}_a(t_1f_1+t_2f_2)=(t_1f_1+t_2f_2)(x+a)=t_1f_1(x+a)+t_2f_2(x+a)=t_1\mathbf{T}_a(f_1)+t_2\mathbf{T}_a(f_2).$$

□

练习 7.3.5. 在光滑函数空间 $C^\infty(\mathbb{R})$ 上定义变换：$\mathbf{A}(f(x))=(f'(x))^2$. 判断 $\mathbf{A}$ 是否为线性变换.

证明: 不是线性变换，例如 $\mathbf{A}(x^2)=4x^2,\mathbf{A}(x)=1$ 但是

$$\mathbf{A}(x^2+x)=(2x+1)^2=4x^2+4x+1\neq\mathbf{A}(x^2)+\mathbf{A}(x).$$

□

练习 7.3.6. 给定 $A=\begin{bmatrix}1&0&-1\\-1&1&0\\0&-1&1\end{bmatrix}$ 和 $\mathbb{R}^{3\times3}$ 上的线性变换 $f\colon X\mapsto AX$. 分别求 $\mathcal{N}(f)$ 和 $\mathcal{R}(f)$ 的维数和一组基.

证明: 设 $X=\begin{bmatrix}\boldsymbol{x}_1&\boldsymbol{x}_2&\boldsymbol{x}_3\end{bmatrix}$,

- $X\in\mathcal{N}(f)$ 当且仅当 $\boldsymbol{x}_i\in\mathcal{N}(A),i=1,2,3$. 而 $\boldsymbol{x}_0=\begin{bmatrix}1\\1\\1\end{bmatrix}$ 是 $\mathcal{N}(A)$ 的一组基，所以 $\mathcal{N}(f)$ 的维数是 3, $\begin{bmatrix}\boldsymbol{x}_0&0&0\end{bmatrix},\begin{bmatrix}0&\boldsymbol{x}_0&0\end{bmatrix},\begin{bmatrix}0&0&\boldsymbol{x}_0\end{bmatrix}$
- $X\in\mathcal{R}(A)$ 当且仅当 $\boldsymbol{x}_i\in\mathcal{R}(A),i=1,2,3$. 而 $\boldsymbol{a}_1=\begin{bmatrix}1\\-1\\0\end{bmatrix},\boldsymbol{a}_2=\begin{bmatrix}0\\1\\-1\end{bmatrix}$ 是 $\mathcal{R}(A)$ 的一组基，因此 $\mathcal{R}(f)$ 的维数是 6,

$$\begin{bmatrix}\boldsymbol{a}_1&0&0\end{bmatrix},\begin{bmatrix}\boldsymbol{a}_2&0&0\end{bmatrix}\begin{bmatrix}0&\boldsymbol{a}_1&0\end{bmatrix}\begin{bmatrix}0&\boldsymbol{a}_2&0\end{bmatrix}\begin{bmatrix}0&0&\boldsymbol{a}_1\end{bmatrix}\begin{bmatrix}0&0&\boldsymbol{a}_2\end{bmatrix}$$

是 $\mathcal{R}(f)$ 的一组基。

□

练习 7.3.7. 计算例 *7.3.2* 中线性映射的核与像集，并求二者的维数.

证明: □

练习 7.3.8. 给定 m 阶方阵 A_1，n 阶方阵 A_2 和 $m\times n$ 矩阵 B. 证明，如果 A_1 和 A_2 没有相同的特征值，则：

1. 对 $m\times n$ 矩阵 X，$A_1X=XA_2$ 只有平凡解.
   提示. 利用 *Hamilton-Cayley* 定理.
2. 对 $m\times n$ 矩阵 X，$A_1X-XA_2=B$ 只有唯一解.

附注. 练习 5.2.23和练习 5.4.7用不同方法证明了相同结论.

证明:

1. 设 X 是 $A_1X=XA_2$ 的一个解，对任意 k, $A_1^kX=A_1^{k-1}XA_2=\cdots=XA_2^k$, 那么对任意多项式 $f(x)$, 都有 $f(A_1)X=Xf(A_2)$. 设 f_1,f_2 分别是 A_1,A_2 的特征多项式，因为 A_1,A_2 没有相同特征值，所以他们没有公共根，因此他们互素。那么存在多项式 $g_1(x),g_2(x)$ 使得 $f_1(x)g_1(x)+f_2(x)g_2(x)=1$. 根据 Hamilton-Cayley 定理，$f_2(A_1)g_2(A_1)=I_n$, 所以 $f_2(A_1)$ 可逆,$g_2(A_1)$ 是它的逆。于是 $f_2(A_1)X=Xf_2(A_2)=0$, 所以 $X=0$.
2. 考虑线性变换：$f:M_{m\times n}(\mathbb{F})\to M_{m\times n}(\mathbb{F}),X\mapsto A_1X-XA_2$, 那么上一小题说明 f 是单射，那么 f 就是双射。所以对任意 B, 矩阵方程 $A_1X-XA_2=B$ 都存在唯一解。

□

练习 7.3.9. 定义 $\mathbb{F}[x]$ 上的变换：$\mathbf{A}(f(x))=xf(x),\forall f(x)\in\mathbb{F}[x]$.

1. 证明 $\mathbf{A}$ 是 $\mathbb{F}[x]$ 上的一个线性变换.
2. 设 $\mathbf{D}$ 是求导算子，证明 $\mathbf{DA}-\mathbf{AD}=\mathrm{I}$.

证明:

1. 对任意 $f,g\in\mathbb{F}[x],a,b\in\mathbb{F}$,
$$\mathbf{A}(af+bg)=x(af+bg)(x)=axf(x)+bxg(x)=a\mathbf{A}(f)+b\mathbf{A}(g)$$
所以 $\mathbf{A}$ 是线性的.
2. 对任意 $f\in\mathbb{F}[x]$
$$(\mathbf{DA}-\mathbf{AD})(f)=\mathbf{D}(xf(x))-\mathbf{A}(f'(x))=(xf(x))-xf'(x)=f(x)$$
因此 $\mathbf{DA}-\mathbf{AD}=\mathrm{I}$.

□

练习 7.3.10. 令 $\mathcal{V}$ 为全体实数数列组成的线性空间，其中元素记为 $(a_0,a_1,\dots)$. 定义其上变换

$$\mathbf{D}\left((a_0,a_1,\dots)\right)=(0,a_0,a_1,\dots),\qquad \mathbf{M}\left((a_0,a_1,\dots)\right)=(a_1,2a_2,3a_3,\dots).$$

1. 证明 $\mathbf{D},\mathbf{M}$ 都是线性变换.
2. 证明 $\mathbf{MD}-\mathbf{DM}=\mathbf{I}$.
3. 对于任意 n 阶方阵 A,B，证明 $AB-BA\neq I_n$.

证明: 1. $\mathbf{D}(x\{a_n\}+y\{b_n\}) = \mathbf{D}(\{xa_n+yb_n\}) = (0, xa_1+yb_1, xa_2+yb_2, \cdots)$, 而

$$x\mathbf{D}(\{a_n\})+y\mathbf{D}(\{b_n\}) = x(0,a_1,a_2,\cdots)+y(0,b_1,b_2,\cdots) = (0,xa_1+yb_1,xa_2+yb_2,\cdots)$$

所以 $x\mathbf{D}(\{a_n\})+y\mathbf{D}(\{b_n\}) = \mathbf{D}(x\{a_n\}+y\{b_n\})$, 即 $\mathbf{D}$ 是线性变换。同理也可以证明 $\mathbf{M}$ 是线性变换。

2. 对任意 $\{a_n\}$,

$$\begin{aligned}(\mathbf{MD}-\mathbf{DM})(\{a_n\}) &= \mathbf{M}(0,a_0,a_1,\cdots)-\mathbf{D}(a_1,2a_2,3a_3)\\ &= (a_0,2a_1,3a_2,\cdots)-(0,a_1,2a_2,\cdots)\\ &= (a_0,a_1,a_2,\cdots)\end{aligned}$$

所以 $\mathbf{MD}-\mathbf{DM}=\mathbf{I}$.

3. 对任意 n 阶方阵 A,B 都有 $\mathrm{trace}(AB-BA)=\mathrm{trace}(AB)-\mathrm{trace}(BA)=0\neq\mathrm{trace}(I_n)$; 所以 $AB-BA\neq I_n$.

□

练习 7.3.11. 设 f 是线性空间 $\mathcal{V}$ 上的线性变换，$\boldsymbol{a}\in\mathcal{V}$. 证明，如果存在正整数 m，使得 $f^{m-1}(\boldsymbol{a})\neq 0, f^m(\boldsymbol{a})=0$，则 $\boldsymbol{a}, f(\boldsymbol{a}),\ldots,f^{m-1}(\boldsymbol{a})$ 线性无关. 由此推出 $\dim\mathcal{V}\geq m$.

证明: 设 $x_0\boldsymbol{a}+x_1f(\boldsymbol{a})+\cdots+x_{m-1}f^{m-1}(\boldsymbol{a})=0$, 则

$$0=f^{m-1}(x_0\boldsymbol{a}+x_1f(\boldsymbol{a})+\cdots+x_{m-1}f^{m-1}(\boldsymbol{a}))=x_0f^{m-1}(\boldsymbol{a})$$

所以 $x_0=0$, 所以 $x_1f(\boldsymbol{a})+\cdots+x_{m-1}f^{m-1}(\boldsymbol{a})=0$, 那么

$$0=f^{m-2}(x_1f(\boldsymbol{a})+\cdots+x_{m-1}f^{m-1}(\boldsymbol{a}))=x_1f^{m-1}(\boldsymbol{a})$$

所以 $x_1=0$, 所以 $x_2f^2(\boldsymbol{a})+\cdots+x_{m-1}f^{m-1}(\boldsymbol{a})=0$. 依次类推，既可以证明 $x_0=x_1=\cdots=x_{m-1}=0$. 所以那么 $\boldsymbol{a},f(\boldsymbol{a}),\ldots,f^{m-1}(\boldsymbol{a})$ 线性无关。 □

练习 7.3.12. 证明命题 *7.3.10*：对 $\mathcal{U}$ 上的线性变换 f，属于不同特征值的特征向量线性无关。

证明: 设 $\lambda_1,\cdots,\lambda_r$ 是 f 的不同的特征值，$\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r$ 分别是属于这些特征值的特征向量。设 $x_1\boldsymbol{v}_1+\cdots+x_r\boldsymbol{v}_r=0$, 那么 $f(x_1\boldsymbol{v}_1+\cdots+x_r\boldsymbol{v}_r)=0$, 即 $\lambda_1x_1\boldsymbol{v}_1+\cdots+\lambda_rx_r\boldsymbol{v}_r=0$, 以此类推，则有

$$\begin{bmatrix}1&\cdots&1\\ \lambda_1&\cdots&\lambda_r\\ \vdots&&\vdots\\ \lambda_1^{r-1}&\cdots&\lambda_r^{r-1}\end{bmatrix}\begin{bmatrix}x_1\boldsymbol{v}_1\\ \vdots\\ x_r\boldsymbol{v}_r\end{bmatrix}=0$$

$\lambda_1,\cdots,\lambda_r$ 互不相等，所以 Vondermonde 矩阵可逆，所以 $x_i\boldsymbol{v}_i=0$, 即 $x_i=0, i=1,2,\cdots,r$. 所以 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r$ 线性无关。 □

练习 7.3.13. 对线性空间 $\mathbb{C}^{n\times n}$，考虑例 *7.3.2* 中的 $\mathbf{L}_A$ 和 $\mathbf{R}_A$，其中 $A\in\mathbb{C}^{n\times n}$.

1. 证明 $\mathbf{L}_A\mathbf{R}_A=\mathbf{R}_A\mathbf{L}_A$，并指出这是矩阵乘法的何种性质.

2. 证明 $\mathbf{L}_A$ 的特征值必是 A 的特征值.

3. 求 $\mathbb{C}^{n\times n}$ 上可逆线性变换 $\mathbf{T}$，使得 $\mathbf{L}_A = \mathbf{T}\mathbf{R}_A\mathbf{T}^{-1}$. **提示**. 利用 A 与 A^{T} 相似，参见练习 5.4.5.

证明: 1. 对任意 $X \in \mathbb{C}^{n\times n}$, $(\mathbf{L}_A\mathbf{R}_A)(X) = \mathbf{L}_A(XA) = AXA$, $(\mathbf{R}_A\mathbf{L}_A)(X) = (\mathbf{R}_A)(AX) = AXA$, 因此 $\mathbf{L}_A\mathbf{R}_A = \mathbf{R}_A\mathbf{L}_A$，这是矩阵乘法的结合律。

2. 设 (λ, X) 是 $\mathbf{L}_A$ 的特征对，则 $\mathbf{L}_A X = \lambda X$, 即 $AX = \lambda X$. 设 $\boldsymbol{x}$ 是 X 的一个非零列向量，那么从 $AX = \lambda X$ 知 $A\boldsymbol{x} = \lambda\boldsymbol{x}$, 所以 λ 是 A 的特征值。

3. 因为 A 和 A^{T} 相似，那么存在可逆矩阵 P 使得 $A = PA^{\mathrm{T}}P^{-1}$. 我们定义 $\mathbf{T}(X) := PX^{\mathrm{T}}P^{-1}$, 易知 $\mathbf{T}$ 是可逆的。那么 $\mathbf{L_A}\mathbf{T}(X) = APX^{\mathrm{T}}P^{-1}$, 而

$$(\mathbf{T}\mathbf{R}_A)(X) = \mathbf{T}(XA) = PA^{\mathrm{T}}X^{\mathrm{T}}P^{-1} = PA^{\mathrm{T}}P^{-1}PX^{\mathrm{T}}P^{-1} = APX^{\mathrm{T}}P^{-1}.$$

□

练习 7.3.14. 设线性空间 $\mathcal{V}$ 有直和分解：$\mathcal{V} = \mathcal{M}_1 \oplus \mathcal{M}_2$. 任取 $\boldsymbol{a} \in \mathcal{V}$，都有唯一的分解式：$\boldsymbol{a} = \boldsymbol{a}_1 + \boldsymbol{a}_2$，其中 $\boldsymbol{a}_1 \in \mathcal{M}_1, \boldsymbol{a}_2 \in \mathcal{M}_2$. 定义 $\mathcal{V}$ 上的变换：

$$\mathbf{P}_{\mathcal{M}_1}(\boldsymbol{a}) = \boldsymbol{a}_1, \quad \mathbf{P}_{\mathcal{M}_2}(\boldsymbol{a}) = \boldsymbol{a}_2.$$

1. 证明，$\mathbf{P}_{\mathcal{M}_1}, \mathbf{P}_{\mathcal{M}_2}$ 都是 $\mathcal{V}$ 上的线性变换. 称 $\mathbf{P}_{\mathcal{M}_1}$ 为沿 $\mathcal{M}_2$ 向 $\mathcal{M}_1$ 的投影变换.

2. 证明，$\mathcal{N}(\mathbf{P}_{\mathcal{M}_1}) = \mathcal{M}_2, \mathcal{R}(\mathbf{P}_{\mathcal{M}_1}) = \mathcal{M}_1$.

3. 证明，$\mathbf{P}_{\mathcal{M}_1}^2 = \mathbf{P}_{\mathcal{M}_1}, \mathbf{P}_{\mathcal{M}_1} + \mathbf{P}_{\mathcal{M}_2} = \mathbf{I}, \mathbf{P}_{\mathcal{M}_1}\mathbf{P}_{\mathcal{M}_2} = \mathbf{O}$.

4. 分别求 $\mathbf{P}_{\mathcal{M}_1}, \mathbf{P}_{\mathcal{M}_2}$ 的特征值和特征子空间.

证明: 1. 我们需要证明对任意 $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{V}, x, y \in \mathbb{F}$, 我们有 $\mathbf{P}_{\mathcal{M}_1}(x\boldsymbol{a} + y\boldsymbol{b}) = x\mathbf{P}_{\mathcal{M}_1}(\boldsymbol{a}) + y\mathbf{P}_{\mathcal{M}_1}(\boldsymbol{b})$. 设 $\boldsymbol{a} = \boldsymbol{a}_1 + \boldsymbol{a}_2, \boldsymbol{b} = \boldsymbol{b}_1 + \boldsymbol{b}_2$，其中 $\boldsymbol{a}_1, \boldsymbol{b}_1 \in \mathcal{M}_1, \boldsymbol{a}_2, \boldsymbol{b}_2 \in \mathcal{M}_2$. 那么 $x\boldsymbol{a} + y\boldsymbol{b} = (x\boldsymbol{a}_1 + y\boldsymbol{b}_1) + (x\boldsymbol{a}_2 + y\boldsymbol{b}_2)$, 其中 $x\boldsymbol{a}_1 + y\boldsymbol{b}_1 \in \mathcal{M}_1, x\boldsymbol{a}_2 + y\boldsymbol{b}_2 \in \mathcal{M}_2$, 所以

$$\mathbf{P}_{\mathcal{M}_1}(x\boldsymbol{a} + y\boldsymbol{b}) = x\boldsymbol{a}_1 + y\boldsymbol{b}_1 = \mathbf{P}_{\mathcal{M}_1}(\boldsymbol{a}) + y\mathbf{P}_{\mathcal{M}_1}(\boldsymbol{b})$$

所以 $\mathbf{P}_{\mathcal{M}_1}$ 是 $\mathcal{V}$ 上的线性变换. 同理可证 $\mathbf{P}_{\mathcal{M}_2}$ 也是 $\mathcal{V}$ 上的线性变换.

2. $\boldsymbol{a} \in \mathcal{N}(\mathbf{P}_{\mathcal{M}_1})$ 当且仅当把 $\boldsymbol{a}$ 分解为 $\boldsymbol{a}_1 + \boldsymbol{a}_2$，其中 $\boldsymbol{a}_1 \in \mathcal{M}_1, \boldsymbol{a}_2 \in \mathcal{M}_2$ 时 $\boldsymbol{a}_1 = 0$，当且仅当 $\boldsymbol{a} \in \mathcal{M}_2$，所以 $\mathcal{N}(\mathbf{P}_{\mathcal{M}_1}) = \mathcal{M}_2$. 根据定义，$\mathcal{R}(\mathbf{P}_{\mathcal{M}_1}) \subseteq \mathcal{M}_1$; 另一方面，对任意 $\boldsymbol{a} \in \mathcal{M}_1$, $\mathbf{P}_{\mathcal{M}_1}(\boldsymbol{a}) = \boldsymbol{a}$, $\mathcal{M}_1 \subseteq \mathcal{R}(\mathbf{P}_{\mathcal{M}_1})$. 所以 $\mathcal{R}(\mathbf{P}_{\mathcal{M}_1}) = \mathcal{M}_1$。

3. 对任意 $\boldsymbol{a} \in \mathcal{V}$，它唯一的分解式：$\boldsymbol{a} = \boldsymbol{a}_1 + \boldsymbol{a}_2$，其中 $\boldsymbol{a}_1 \in \mathcal{M}_1, \boldsymbol{a}_2 \in \mathcal{M}_2$，那么

$$\mathbf{P}_{\mathcal{M}_1}^2\boldsymbol{a} = \mathbf{P}_{\mathcal{M}_1}\boldsymbol{a}_1 = \boldsymbol{a}_1 = \mathbf{P}_{\mathcal{M}_1}\boldsymbol{a},$$

所以 $\mathbf{P}_{\mathcal{M}_1}^2 = \mathbf{P}_{\mathcal{M}_1}$.

$$(\mathbf{P}_{\mathcal{M}_1} + \mathbf{P}_{\mathcal{M}_2})\boldsymbol{a} = \mathbf{P}_{\mathcal{M}_1}\boldsymbol{a} + \mathbf{P}_{\mathcal{M}_2}\boldsymbol{a} = \boldsymbol{a}_1 + \boldsymbol{a}_2 = \boldsymbol{a},$$

所以 $\mathbf{P}_{\mathcal{M}_1} + \mathbf{P}_{\mathcal{M}_1} = \mathbf{I}$.

$$\mathbf{P}_{\mathcal{M}_1}\mathbf{P}_{\mathcal{M}_2}\boldsymbol{a} = \mathbf{P}_{\mathcal{M}_1}\boldsymbol{a}_2 = 0$$

所以 $\mathbf{P}_{\mathcal{M}_1}\mathbf{P}_{\mathcal{M}_2} = \mathbf{O}$.

4. 根据第 2 小题，$\mathbf{P}_{\mathcal{M}_1}$ 特征值为 $0,1$, 对应的特征子空间分别为 $\mathcal{M}_2,\mathcal{M}_1$.

$\square$

练习 7.3.15. 考虑例 *7.3.11* 中的线性变换 $\mathbb{F}^{n\times n}$ 上的线性变换 $\mathbf{S}$，试问它是否为投影变换.

证明: $\square$

练习 7.3.16. 考虑练习 *7.1.1*中的实线性空间 $\mathbb{R}^+$，给定 $a>0$，判断映射

$$\begin{aligned}\log_a\colon\ \mathbb{R}^+ &\to \mathbb{R},\\ x &\mapsto \log_a x,\end{aligned}$$

是否为线性映射. 如果是，进一步分析当 a 取何值时，该映射是同构.

证明: $a=1$ 时，$\log_a$ 没有定义，所以我们考虑 $a\neq 1$.

$\log_a$ 是线性映射，当且仅当对任意 $x,y\in\mathbb{R}^+,k\in\mathbb{R}$, $\log_a((k\otimes x)\oplus(l\otimes y))=k\log_a x+l\log_a y$. 因为 $(k\otimes x)\oplus(l\otimes y)=x^k\oplus y^l=x^ky^l$, 所以

$$\log_a((k\otimes x)\oplus(l\otimes y))=\log_a(x^ky^l)=k\log_a x+l\log_a y.$$

因此 $\log_a$ 是线性映射。

$a\neq 1$ 时，$\log_a$ 是满射：对任意 $t\in\mathbb{R},\log_a(a^t)=1$. $\log_a$ 也是单射：若 $\log_a x=0$ 则 $t=a^0=1$, 而 1 是 $\mathbb{R}^+$ 的零元素。所以对任意 $a>0,a\neq 1$, $\log_a$ 都是同构。 $\square$

练习 7.3.17. 任取二阶实方阵 A，满足 $A^2=-I_2$. 考虑练习 *7.1.7*中的线性空间 $P(A)$. 证明 $f:\mathbb{C}\to P(A), f(a+b\boldsymbol{i})=aI_2+bA$ 是这两个线性空间的同构. 请举出至少两种可能的 A.

证明: 由于 $A^2=-I_2$，对任意 $f(x)=a_0+a_1x+\cdots+a_nx^n$, 都有 $f(A)=a_0I_2+a_1A-a_2I_2-a_3A+\cdots$, 因此 $P(A)=\mathrm{Span}(I_2,A)$。另一方面，若 $xI_2+yA=0$, 那么乘以 A 就有 $xA-yI_2=0$, 消去 A 就有 $(x^2+y^2)I_2=1$，所以 $x=y=0$, 因此 I_2,A 是 $P(A)$ 的一组基。所以 $\dim_{\mathbb{R}}\mathbb{C}=\dim_{\mathbb{R}}P(A)=2$.

对任意 $x,y\in\mathbb{R},z_1=a_1+b_1\boldsymbol{i},z_2=a_2+b_2\boldsymbol{i}\in\mathbb{C}$,

$$f(xz_1+yz_2)=f((xa_1+ya_2)+(xb_1+yb_2)\boldsymbol{i})=(xa_1+ya_2)I_2+(xb_1+yb_2)=xf(z_1)+yf(z_2),$$

所以 f 是 $\mathbb{R}$ 上线性空间的线性映射。因为 I_2,A 线性无关，所以 $f(a+b\boldsymbol{i})=0$ 当且仅当 $a=b=0$, 因此 f 是单射，又因为两个线性空间维数相同，从而是同构。

例如 $A=\begin{bmatrix}0&1\\-1&0\end{bmatrix},\begin{bmatrix}0&-1\\1&0\end{bmatrix}$ $\square$

以下三个问题是同一个思路，因为涉及的线性空间维数是相同的，我们只需构造一个单射或者满射的线性映射即可，我们给出构造，具体的验证忽略，请读者自行完成。

练习 7.3.18. 证明矩阵空间 $\mathbb{F}^{m\times n}$ 与 $\mathbb{F}^{mn}$ 同构.

$$\mathbb{F}^{m\times n}\longrightarrow\mathbb{F}^{mn},\quad \begin{bmatrix}\boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n\end{bmatrix}\mapsto\begin{bmatrix}\boldsymbol{a}_1\\ \vdots\\ \boldsymbol{a}_n\end{bmatrix}$$

注意，这个映射不是转置。矩阵是把列向量排成一个长方形，而这个像是把列向量排成行数更多的列向量。

练习 7.3.19. 证明多项式空间 $\mathbb{F}[x]_n$ 与 $\mathbb{F}^n$ 同构.

$$\mathbb{F}^n\longrightarrow\mathbb{F}[x]_n,\quad \begin{bmatrix}a_1\\ a_2\\ \vdots\\ a_n\end{bmatrix}\mapsto a_1+a_2x+\cdots+a_nx^{n-1}.$$

练习 7.3.20. 证明练习 *7.1.2*中的线性空间 $\mathbb{Q}[\omega]$ 与练习 *7.1.4*中的线性空间 $\mathbb{Q}[\boldsymbol{i}]$ 同构.

$$\mathbb{Q}[\omega]\longrightarrow\mathbb{Q}[\boldsymbol{i}],\quad a+b\omega\mapsto a+b\boldsymbol{i}.$$

7.4 向量的坐标表示

练习 7.4.1. 证明命题 *7.4.1*：向量组 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_s$ 线性无关，如果 $\boldsymbol{b}$ 可以被其线性表示，则表示法唯一

证明: 若 $\boldsymbol{b}=x_1\boldsymbol{a}_1+\cdots+x_s\boldsymbol{a}_s=y_1\boldsymbol{a}_1+\cdots+y_s\boldsymbol{a}_s$, 那么

$$(x_1-y_1)\boldsymbol{a}_1+\cdots+(x_s-y_s)\boldsymbol{a}_s=0,$$

$\boldsymbol{a}_1,\cdots,\boldsymbol{a}_s$ 线性无关则 $x_1-y_1=\cdots=x_s-y_s=0$, 即 $x_1=y_1,x_2=y_2,\cdots,x_s=y_s$, 所以表示法唯一。 □

练习 7.4.2. 求 $\mathbb{F}^4$ 中由基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4$ 到基 $\boldsymbol{t}_1,\boldsymbol{t}_2,\boldsymbol{t}_3,\boldsymbol{t}_4$ 的过渡矩阵，并分别求向量 $\boldsymbol{a}$ 在两组基下的坐标.

1. $$\boldsymbol{e}_1=\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\ \boldsymbol{e}_2=\begin{bmatrix}0\\1\\0\\0\end{bmatrix},\ \boldsymbol{e}_3=\begin{bmatrix}0\\0\\1\\0\end{bmatrix},\ \boldsymbol{e}_4=\begin{bmatrix}0\\0\\0\\1\end{bmatrix};$$
$$\boldsymbol{t}_1=\begin{bmatrix}2\\1\\-1\\1\end{bmatrix},\ \boldsymbol{t}_2=\begin{bmatrix}0\\3\\1\\0\end{bmatrix},\ \boldsymbol{t}_3=\begin{bmatrix}5\\3\\2\\1\end{bmatrix},\ \boldsymbol{t}_4=\begin{bmatrix}6\\6\\1\\3\end{bmatrix};\quad \boldsymbol{a}=\begin{bmatrix}a_1\\a_2\\a_3\\a_4\end{bmatrix}.$$

2. $\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}, \boldsymbol{e}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{e}_3 = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{e}_4 = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$;

$$\boldsymbol{t}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{t}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix}, \boldsymbol{t}_3 = \begin{bmatrix} -2 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \boldsymbol{t}_4 = \begin{bmatrix} 1 \\ 3 \\ 1 \\ 2 \end{bmatrix}; \quad \boldsymbol{a} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

证明*:* 这个问题有特殊性，因为考虑的都是 $\mathbb{R}^4$ 的基（若是一般的 $\mathbb{R}^n$ 也类似），我们来分析一下。假设 M 是基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$ 到基 $\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3, \boldsymbol{t}_4$ 的过渡矩阵，即

$$\begin{bmatrix} \boldsymbol{t}_1 & \boldsymbol{t}_2 & \boldsymbol{t}_3 & \boldsymbol{t}_4 \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{e}_4 \end{bmatrix} M$$

令 $E = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{e}_4 \end{bmatrix}, T = \begin{bmatrix} \boldsymbol{t}_1 & \boldsymbol{t}_2 & \boldsymbol{t}_3 & \boldsymbol{t}_4 \end{bmatrix}$，那么 E, T 都是 4 阶可逆方阵，所以 $M = E^{-1}T$.

另一方面，若 $\boldsymbol{a}$ 在基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$ 之下的坐标是 x_1, x_2, x_3, x_4，那么

$$\boldsymbol{a} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3 + x_4\boldsymbol{e}_4 = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{e}_4 \end{bmatrix} = E \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

那么 $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = E^{-1}\boldsymbol{a}$. $\boldsymbol{a}$ 在基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$ 之下的坐标表示的讨论也类似。

下面我们就写出最终答案，不再写具体结果。

1. $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$ 到基 $\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3, \boldsymbol{t}_4$ 的过渡矩阵是 $T = \begin{bmatrix} 2 & 0 & 5 & 6 \\ 1 & 3 & 3 & 6 \\ -1 & 1 & 2 & 1 \\ 1 & 0 & 1 & 3 \end{bmatrix}$ $\boldsymbol{a}$ 在基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$ 下的坐标为 a_1, a_2, a_3, a_4, 在基 $\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3, \boldsymbol{t}_4$ 下的坐标是 $\frac{4}{9}a_1 + \frac{1}{3}a_2 - a_3 - \frac{11}{9}a_4, \frac{1}{29}a_1 + \frac{4}{9}a_2 - \frac{1}{3}a_3 - \frac{23}{27}a_4, \frac{1}{3}a_1 - \frac{2}{3}a_4, -\frac{7}{27}a_1 - \frac{1}{9}a_1 + \frac{1}{2}a_3 + \frac{26}{27}a_4$.

2. $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$ 到基 $\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3, \boldsymbol{t}_4$ 的过渡矩阵是 $T = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ $\boldsymbol{a}$ 在基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$ 下的坐标为 $\frac{3}{13}, \frac{5}{13}, -\frac{2}{13}, \frac{-3}{13}$, 在基 $\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3, \boldsymbol{t}_4$ 下的坐标是 $\frac{4}{13}, \frac{2}{13}, -\frac{3}{13}, -\frac{1}{13}$

□

练习 7.4.3. 考虑函数空间的子空间 $\mathrm{Span}(\sin^2 x, \cos^2 x)$.

1. 证明 $\sin^2 x, \cos^2 x$ 和 $1, \cos 2x$ 分别是子空间的一组基.

2. 分别求从 $\sin^2 x, \cos^2 x$ 到 $1, \cos 2x$，和从 $1, \cos 2x$ 到 $\sin^2 x, \cos^2 x$ 的过渡矩阵.

3. 分别求 1 和 $\sin^2 x$ 在两组基下的坐标.

证明: 1. 设 $t_1 \sin^2 x + t_2 \cos^2 x = 0$, 分别取 $x = 0, \frac{\pi}{2}$, 就有 $t_1 = t_2 = 0$, 所以 $\sin^2 x, \cos^2 x$, 那么它们就是一组基。而 $\begin{bmatrix} 1 & \cos 2x \end{bmatrix} = \begin{bmatrix} \cos^2 x & \sin^2 x \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, 因为 $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ 可逆, 所以 $1, \cos 2x$ 也是一组基。

2. 我们已经说明了从 $1, \cos 2x$ 到 $\sin^2 x, \cos^2 x$ 的过渡矩阵是 $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$,那么从 $\sin^2 x, \cos^2 x$ 到 $1, \cos 2x$ 的过渡矩阵是 $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$,

3.
- $1 = 1 \cdot 1 + 0 \cdot \cos 2x = \cos^2 x + \sin^2 x$, 所以 1 在基 $1, \cos 2x$ 之下的坐标是 $(1, 0)$；在基 $\cos^2 x, \sin^2 x$ 之下的坐标是 $(1, 1)$.
- $\sin^2 x = \frac{1}{2} - \frac{1}{2} \cos 2x = 1 \cdot \sin^2 x + 0 \cdot \cos^2 x$, 所以 $\sin^2 x$ 基 $1, \cos 2x$ 之下的坐标是 $(1/2, -1/2)$；在基 $\cos^2 x, \sin^2 x$ 之下的坐标是 $(0, 1)$.

□

练习 7.4.4. 考虑多项式空间 $\mathbb{R}[x]_4$.

1. 求如下这三组基之间的过渡矩阵：

$$1, x, x^2, x^3; \qquad 1, x+1, (x+1)^2, (x+1)^3; \qquad 1, x+2, (x+2)^2, (x+2)^3.$$

2. 求多项式 $p(x) = x^3 + 6x^2 + 12x + 8$ 在上述三组基下的坐标.

3. 设多项式 $q(x)$ 在基 $1, x+1, (x+1)^2, (x+1)^3$ 下的坐标为 $\begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$. 求 $q(1), q(0), q'(0)$.

证明: 1. $\begin{bmatrix} 1 & x+1 & (x+1)^2 & (x+1)^3 \end{bmatrix} = \begin{bmatrix} 1 & x & x^2 & x^3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, 所以从 $1, x, x^2, x^3$ 到 $1, x+1, (x+1)^2, (x+1)^3$ 的过度矩阵是 $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

$\begin{bmatrix} 1 & x+2 & (x+2)^2 & (x+2)^3 \end{bmatrix} = \begin{bmatrix} 1 & x & x^2 & x^3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 & 8 \\ 0 & 1 & 4 & 12 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, 所以从 $1, x, x^2, x^3$

到 $1, x+2, (x+2)^2, (x+2)^3$ 的过度矩阵是 $\begin{bmatrix} 1 & 2 & 4 & 8 \\ 0 & 1 & 4 & 12 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

$\begin{bmatrix} 1 & x+2 & (x+2)^2 & (x+2)^3 \end{bmatrix} = \begin{bmatrix} 1 & x+1 & (x+1)^2 & (x+1)^3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, 所以

从 $1, x+1, (x+1)^2, (x+1)^3$ 到 $1, x+2, (x+2)^2, (x+2)^3$ 的过度矩阵是 $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

2. $p(x) = \begin{bmatrix} 1 & x & x^2 & x^3 \end{bmatrix} \begin{bmatrix} 8 \\ 12 \\ 6 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & x+1 & (x+1)^2 & (x+1)^3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 8 \\ 12 \\ 6 \\ 1 \end{bmatrix} =$ $\begin{bmatrix} 1 & x+2 & (x+2)^2 & (x+2)^3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 & 8 \\ 0 & 1 & 4 & 12 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 8 \\ 12 \\ 6 \\ 1 \end{bmatrix}$。所以，它在基 $1, x, x^2, x^3$ 之下的坐标是 $8, 12, 6, 1$；在 $1, x+1, (x+1)^2, (x+1)^3$ 之下的坐标是 $1, 3, 3, 1$；在 $1, x+2, (x+2)^2, (x+2)^3$ 之下的坐标是 $0, 0, 0, 1$.

3. 按题意，$q(x) = 4 + 3(x+1) + 2(x+1)^2 + (x+1)^3$. 那么 $q(1) = 4+6+8+8 = 26$, $q(0) = 4+3+2+1 = 10$. $q'(x) = 3 + 4(x+1) + 3(x+1)^2$, 所以 $q'(0) = 3+4+3 = 10$.

$\square$

练习 7.4.5. 考虑多项式空间 $\mathbb{R}[x]_4$ 中的的一组基 p_0, p_1, p_2, p_3，其中每个多项式 p_n 满足 $p_n(\cos\theta) = \cos(n\theta)$. 求这组基和基 $1, x, x^2, x^3$ 之间的过渡矩阵.

这样的多项式叫 *Chebyshev* 多项式，有很多特殊性质，参见??.

证明: $p_0(\cos\theta) = \cos(0\theta) = 1$, $p_1(\cos\theta) = \cos\theta$, 那么 $p_1(x) = x$. $\cos(2\theta) = 2\cos^2\theta - 1$, $\cos(3\theta) = 4\cos^3\theta - 3\cos\theta$. 所以 $p_2(x) = 2x^2 - 1$, $p_3(x) = 4x^3 - 3x$. 那么

$$\begin{bmatrix} p_0 & p_1 & p_2 & p_3 \end{bmatrix} = \begin{bmatrix} 1 & x & x^2 & x^3 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

$\square$

练习 7.4.6. 考虑练习 7.2.7中的线性空间 $\mathbb{F}[x]_n$，考虑 $n = 3$ 的情形.

1. 给定 $\mathbb{F}$ 中两两不等的数 a_1,a_2,a_3，求由基 $1,x,x^2$ 到基 $f_1(x),f_2(x),f_3(x)$ 的过渡矩阵.

2. 考虑练习 7.2.7中的 $f(x)$，给定 $\mathbb{F}$ 中任意三个数 b_1,b_2,b_3，求 $f(x)$ 在两组基下的坐标.

证明: 1. $\begin{bmatrix} f_1(x) & f_2(x) & f_3(x) \end{bmatrix} = \begin{bmatrix} 1 & x & x^2 \end{bmatrix} \begin{bmatrix} a_2a_3 & a_1a_3 & a_1a_2 \\ -a_2-a_3 & -a_1-a_3 & -a_1-a_2 \\ 1 & 1 & 1 \end{bmatrix}$, 所以由基 $1,x,x^2$ 到基 $f_1(x),f_2(x),f_3(x)$ 的过渡矩阵是 $\begin{bmatrix} a_2a_3 & a_1a_3 & a_1a_2 \\ -a_2-a_3 & -a_1-a_3 & -a_1-a_2 \\ 1 & 1 & 1 \end{bmatrix}$.

2. $f(x) = \dfrac{b_1}{f_1(a_1)}f_1(x) + \dfrac{b_2}{f_2(a_2)}f_2(x) + \dfrac{b_3}{f_3(a_3)}f_3(x)$, 所以 $f(x)$ 在基 $f_1(x),f_2(x),f_3(x)$ 之下的坐标是 $\dfrac{b_1}{(a_1-a_2)(a_1-a_3)}, \dfrac{b_2}{(a_2-a_1)(a_2-a_3)}, \dfrac{b_3}{(a_3-a_1)(a_3-a_2)}$.

我们利用过渡矩阵计算它在 $1,x,x^2$ 之下的坐标，即计算

$$\begin{bmatrix} a_2a_3 & a_1a_3 & a_1a_2 \\ -a_2-a_3 & -a_1-a_3 & -a_1-a_2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{b_1}{f_1(a_1)} \\ \frac{b_2}{f_2(a_2)} \\ \frac{b_3}{f_3(a_3)} \end{bmatrix},$$

结果请自行计算。

□

练习 7.4.7. 矩阵空间 $\mathbb{F}^{2\times 2}$ 有两组基 $\boldsymbol{e}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \boldsymbol{e}_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \boldsymbol{e}_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \boldsymbol{e}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, 和 $\boldsymbol{t}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \boldsymbol{t}_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \boldsymbol{t}_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \boldsymbol{t}_4 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. 求从基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4$ 到基 $\boldsymbol{t}_1,\boldsymbol{t}_2,\boldsymbol{t}_3,\boldsymbol{t}_4$ 的过渡矩阵.

证明: $\begin{bmatrix} \boldsymbol{t}_1 & \boldsymbol{t}_2 & \boldsymbol{t}_3 & \boldsymbol{t}_4 \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{e}_4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \end{bmatrix}$, 那么从基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4$ 到基 $\boldsymbol{t}_1,\boldsymbol{t}_2,\boldsymbol{t}_3,\boldsymbol{t}_4$ 的过渡矩阵是 $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \end{bmatrix}$. □

练习 7.4.8. 设 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4$ 是线性空间 $\mathcal{V}$ 的一组基，求下列向量组的一个极大线性无关部分组:

$$\boldsymbol{t}_1 = \boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3 + 3\boldsymbol{e}_4, \qquad \boldsymbol{t}_2 = -\boldsymbol{e}_1 - 3\boldsymbol{e}_2 + 5\boldsymbol{e}_3 + \boldsymbol{e}_4,$$
$$\boldsymbol{t}_3 = 3\boldsymbol{e}_1 + 2\boldsymbol{e}_2 - \boldsymbol{e}_3 + 4\boldsymbol{e}_4, \quad \boldsymbol{t}_4 = -2\boldsymbol{e}_1 - 6\boldsymbol{e}_2 + 10\boldsymbol{e}_3 + 2\boldsymbol{e}_4.$$

证明: $\begin{bmatrix} \boldsymbol{t}_1 & \boldsymbol{t}_2 & \boldsymbol{t}_3 & \boldsymbol{t}_4 \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{e}_4 \end{bmatrix} \begin{bmatrix} 1 & -1 & 3 & -2 \\ 1 & -3 & 2 & -6 \\ 1 & 5 & -1 & 10 \\ 3 & 1 & 4 & 2 \end{bmatrix}$ 根据命题 7.4.4，我们只需要求

$\begin{bmatrix}1 & -1 & 3 & -2\\ 1 & -3 & 2 & -6\\ 1 & 5 & -1 & 10\\ 3 & 1 & 4 & 2\end{bmatrix}$ 的一个极大线性无关组。我们作初等行变换

$$\begin{bmatrix}1 & -1 & 3 & -2\\ 1 & -3 & 2 & -6\\ 1 & 5 & -1 & 10\\ 3 & 1 & 4 & 2\end{bmatrix} \longrightarrow \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 2\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}$$

前 3 列是个极大线性无关组，所以 $\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3$ 是一个极大线性无关组。 □

练习 7.4.9. 设 $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ 是线性空间 $\mathcal{V}$ 的一组基.

1. 判断 $\boldsymbol{t}_1 = \boldsymbol{e}_1, \boldsymbol{t}_2 = \boldsymbol{e}_1 + \boldsymbol{e}_2, \ldots, \boldsymbol{t}_n = \boldsymbol{e}_1 + \boldsymbol{e}_2 + \cdots + \boldsymbol{e}_n$ 是否也是 $\mathcal{V}$ 的一组基.
2. 判断 $\boldsymbol{t}_1 = \boldsymbol{e}_1 + \boldsymbol{e}_2, \boldsymbol{t}_2 = \boldsymbol{e}_2 + \boldsymbol{e}_3, \ldots, \boldsymbol{t}_n = \boldsymbol{e}_n + \boldsymbol{e}_1$ 是否也是 $\mathcal{V}$ 的一组基.

证明: 1. $\begin{bmatrix}\boldsymbol{t}_1 & \boldsymbol{t}_2 & \cdots & \boldsymbol{t}_n\end{bmatrix} = \begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n\end{bmatrix} \begin{bmatrix}1 & 1 & \cdots & 1\\ 0 & 1 & \cdots & 1\\ \vdots & & \ddots & \vdots\\ 0 & \cdots & & 1\end{bmatrix}$. 矩阵 $\begin{bmatrix}1 & 1 & \cdots & 1\\ 0 & 1 & \cdots & 1\\ \vdots & & \ddots & \vdots\\ 0 & \cdots & & 1\end{bmatrix}$ 可逆，所以 $\boldsymbol{t}_1, \cdots, \boldsymbol{t}_n$ 也是一组基。

2. $\begin{bmatrix}\boldsymbol{t}_1 & \boldsymbol{t}_2 & \cdots & \boldsymbol{t}_n\end{bmatrix} = \begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n\end{bmatrix} \begin{bmatrix}1 & 1 & & \\ & 1 & 1 & \\ & & \ddots & \\ 1 & & & 1\end{bmatrix}$. 矩阵 $\begin{bmatrix}1 & 1 & & \\ & 1 & 1 & \\ & & \ddots & \\ 1 & & & 1\end{bmatrix}$ 可逆当且仅当 n 是奇数，所以当且仅当 n 是奇数时，$\boldsymbol{t}_1, \cdots, \boldsymbol{t}_n$ 也是一组基。

□

练习 7.4.10. 设 $a_1, a_2, \ldots, a_n$ 是 $\mathbb{F}$ 中两两不等的数，$\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ 是线性空间 $\mathcal{V}$ 的一组基，令 $\boldsymbol{t}_i = \boldsymbol{e}_1 + a_i\boldsymbol{e}_2 + \cdots + a_i^{n-1}\boldsymbol{e}_n, i = 1, 2, \ldots, n$. 证明 $\boldsymbol{t}_1, \boldsymbol{t}_2, \ldots, \boldsymbol{t}_n$ 也是 $\mathcal{V}$ 的一组基.

证明: $\begin{bmatrix}\boldsymbol{t}_1 & \boldsymbol{t}_2 & \cdots & \boldsymbol{t}_n\end{bmatrix} = \begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n\end{bmatrix} V(a_1, a_2, \cdots, a_n)$, 其中 $V(a_1, a_2, \cdots, a_n)$ 是 Vondermonde 矩阵。$a_1, \cdots, a_n$ 两两不等，所以 $V(a_1, a_2, \cdots, a_n)$ 可逆，所以 $\boldsymbol{t}_1, \cdots, \boldsymbol{t}_n$ 也是一组基。 □

练习 7.4.11. 设 (I): $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$, (II): $\boldsymbol{t}_1, \boldsymbol{t}_2, \ldots, \boldsymbol{t}_n$ 和 (III): $\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_n$ 是线性空间 $\mathcal{V}$ 的三组基，如果从 (I) 到 (II) 的过渡矩阵是 P，从 (II) 到 (III) 的过渡矩阵是 Q，证明：

1. 从 (II) 到 (I) 的过渡矩阵是 P^{-1}.
2. 从 (I) 到 (III) 的过渡矩阵是 PQ.

证明: $\begin{bmatrix}\boldsymbol{t}_1 & \boldsymbol{t}_2 & \cdots & \boldsymbol{t}_n\end{bmatrix} = \begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n\end{bmatrix} P$, 则

$$\begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n\end{bmatrix} = \begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n\end{bmatrix} PP^{-1} = \begin{bmatrix}\boldsymbol{t}_1 & \boldsymbol{t}_2 & \cdots & \boldsymbol{t}_n\end{bmatrix} P^{-1}$$

所以从 (II) 到 (I) 的过渡矩阵是 P^{-1}. $\begin{bmatrix} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \cdots & \boldsymbol{s}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{t}_1 & \boldsymbol{t}_2 & \cdots & \boldsymbol{t}_n \end{bmatrix} Q$, 则

$$\begin{bmatrix} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \cdots & \boldsymbol{s}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{t}_1 & \boldsymbol{t}_2 & \cdots & \boldsymbol{t}_n \end{bmatrix} Q = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n \end{bmatrix} PQ$$

所以从 (I) 到 (III) 的过渡矩阵是 PQ. □

7.5 线性映射的矩阵表示

练习 7.5.1. 设 $\mathbf{A}$ 是 $\mathbb{F}^3$ 上的一个线性变换:

$$\mathbf{A}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ x_3 - x_2 \\ x_2 - x_3 \end{bmatrix}.$$

求 $\mathbf{A}$ 在标准基下的矩阵，并分别求 $\mathcal{N}(\mathbf{A})$ 和 $\mathcal{R}(\mathbf{A})$ 的一组基和维数.

证明:

$$\begin{aligned} \mathbf{A}(\boldsymbol{e}_1) &= \mathbf{A}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \boldsymbol{e}_1 \\ \mathbf{A}(\boldsymbol{e}_2) &= \mathbf{A}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} = 2\boldsymbol{e}_1 - \boldsymbol{e}_2 + \boldsymbol{e}_3 \\ \mathbf{A}(\boldsymbol{e}_3) &= \mathbf{A}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \boldsymbol{e}_2 - \boldsymbol{e}_3 \end{aligned}$$

所以 $\mathbf{A}$ 在标准基下的矩阵是 $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$. $\mathcal{N}(A)$ 的一组基是 $\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, $\mathcal{R}(A)$ 的一组基是 $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$. 所以 $\mathcal{N}(\mathbf{A})$ 的维数是 1, $-2\boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3$ 是 $\mathcal{N}(\mathbf{A})$ 的一组基; $\mathcal{R}(\mathbf{A})$ 的维数是 2, $\boldsymbol{e}_1, \boldsymbol{e}_2 - \boldsymbol{e}_3$ 是 $\mathcal{R}(\mathbf{A})$ 的一组基。 □

练习 7.5.2. 设 $\mathcal{V} = \operatorname{Span}(f_1, f_2)$ 是函数空间的子空间，其中 $f_1 = e^{ax}\cos bx, f_2 = e^{ax}\sin bx$. 证明求导算子 $\boldsymbol{D}$ 是 $\mathcal{V}$ 上的线性变换，并求其在基 f_1, f_2 下的矩阵.

证明:

$$\begin{aligned} & \boldsymbol{D}(t_1 f_1 + t_2 f_2) \\ = \ & t_1 a e^{ax}\cos bx - t_1 b e^{ax}\sin bx + t_2 a e^{ex}\sin bx + t_2 b e^{ax}\cos bx \ , \\ = \ & (t_1 a + t_2 b) f_1 + (t_2 a - t_1 b) f_2 \end{aligned}$$

所以 $\boldsymbol{D}$ 的像集包含在 $\mathcal{V}$ 中。$\boldsymbol{D}$ 显然是线性的，所以 $\boldsymbol{D}$ 是 $\mathcal{V}$ 上的线性变换。

$$\boldsymbol{D}\begin{bmatrix} f_1 & f_2 \end{bmatrix} = \begin{bmatrix} f_1 & f_2 \end{bmatrix}\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

所以 $\boldsymbol{D}$ 在基 f_1, f_2 下的矩阵是 $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$. □

练习 7.5.3. 设 $\mathbb{F}^{2\times 2}$ 上的线性变换 $\mathbf{L}_A\colon X \mapsto AX$，其中 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. 求 $\mathbf{L}_A$ 在基 E_{11}，E_{12}，E_{21}，E_{22} 下的矩阵.

证明:

$$AE_{11} = \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix}, AE_{12} = \begin{bmatrix} 0 & a \\ 0 & c \end{bmatrix}, AE_{21} = \begin{bmatrix} b & 0 \\ d & 0 \end{bmatrix}, AE_{22} = \begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}$$

于是

$$\mathbf{L}_A \begin{bmatrix} E_{11} & E_{12} & E_{21} & E_{22} \end{bmatrix} = \begin{bmatrix} E_{11} & E_{12} & E_{21} & E_{22} \end{bmatrix} \begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix}$$

$\mathbf{L}_A$ 在基 E_{11}，E_{12}，E_{21}，E_{22} 下的矩阵是 $\begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix}$. □

练习 7.5.4. 设 $\mathcal{V}$ 是所有二阶对称矩阵构成的线性空间，f 是其上的线性变换：$f(X) = A^{\mathrm{T}} X A$，其中 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. 求 f 在基 $E_{11}, E_{22}, E_{12} + E_{21}$ 下的矩阵.

证明:

$$\begin{aligned}
f(E_{11}) &= A^{\mathrm{T}} E_{11} A &&= \begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix} &&= a^2 E_{11} + b^2 E_{22} + ab(E_{12} + E_{21}) \\
f(E_{22}) &= A^{\mathrm{T}} E_{22} A &&= \begin{bmatrix} c^2 & cd \\ cd & d^2 \end{bmatrix} &&= c^2 E_{11} + d^2 E_{22} + cd(E_{12} + E_{21}) \\
f(E_{12} + E_{21}) &= A^{\mathrm{T}} (E_{12} + E_{21}) A &&= \begin{bmatrix} 2ac & bc + ad \\ bc + ad & 2bd \end{bmatrix} &&= 2ac E_{11} + 2bd E_{22} + (ad + bc)(E_{12} + E_{21})
\end{aligned}$$

即

$$f \begin{bmatrix} E_{11} & E_{22} & E_{12} + E_{21} \end{bmatrix} = \begin{bmatrix} E_{11} & E_{22} & E_{12} + E_{21} \end{bmatrix} \begin{bmatrix} a^2 & c^2 & 2ac \\ b^2 & d^2 & 2bd \\ ad & cd & ad + bc \end{bmatrix}$$

所以 f 在基 $E_{11}, E_{22}, E_{12} + E_{21}$ 下的矩阵是 $\begin{bmatrix} a^2 & c^2 & 2ac \\ b^2 & d^2 & 2bd \\ ad & cd & ad + bc \end{bmatrix}$. □

练习 7.5.5. 设 $\mathbf{A}$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 $\mathcal{V}$ 上的一个线性变换，证明，存在 $\mathbb{F}[x]$ 中一个次数不超过 n^2 的多项式 $f(x)$，使得 $f(\mathbf{A}) = \mathbf{O}$.

证明: $\mathrm{Hom}(\mathcal{V},\mathcal{V})$ 是维数为 n^2 的线性空间，那么 $\boldsymbol{I},\mathbf{A},\cdots,\mathbf{A}^{n^2}$ 线性相关，于是存在不全为零的 $a_0,a_1,\cdots,a_{n^2}$ 使得

$$a_0\boldsymbol{I}+a_1\mathbf{A}+\cdots+a_{n^2}\mathbf{A}^{n^2}=\mathbf{O}.$$

我们取 $f(x)=a_0+a_1x+\cdots+a_{n^2}x^{n^2}$ 即可 □

练习 7.5.6. 证明命题 7.5.5. 给定线性空间 $\mathcal{U},\mathcal{V}$ 及各自的一组基 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n$; $\boldsymbol{i}_1,\cdots,\boldsymbol{i}_m$. 设 $f\in\mathrm{Hom}(\mathcal{U},\mathcal{V})$, 而 F 是 f 在给定基之下的矩阵，则

$$\sigma_{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n}:\mathcal{N}(f)\to\mathcal{N}(F),\quad \sigma_{\boldsymbol{i}_1,\cdots,\boldsymbol{i}_m}:\mathcal{R}(f)\to\mathcal{R}(F)$$

都是同构。特别地，$\dim\mathcal{N}(f)=\dim\mathcal{N}(F),\dim\mathcal{R}(f)=\dim\mathcal{R}(F)$. 因此，$\dim\mathcal{N}(f)+\dim\mathcal{R}(f)=\dim\mathcal{U}$.

证明: 这是一个验证定义的题，所以我们先梳理一下定义。按照定义，我们有

- $\sigma_{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n}\left(\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix}\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\right)=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$,
- $\sigma_{\boldsymbol{i}_1,\cdots,\boldsymbol{i}_n}\left(\begin{bmatrix}\boldsymbol{i}_1 & \cdots & \boldsymbol{i}_m\end{bmatrix}\begin{bmatrix}y_1\\ \vdots\\ y_m\end{bmatrix}\right)=\begin{bmatrix}y_1\\ \vdots\\ y_m\end{bmatrix}$,
- $f\left(\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix}\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\right)=\begin{bmatrix}\boldsymbol{i}_1 & \cdots & \boldsymbol{i}_m\end{bmatrix}F\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$。

那么

- $\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\in\mathcal{N}(F)\Leftrightarrow\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix}\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\in\mathcal{N}(f)$,
- $\begin{bmatrix}y_1\\ \vdots\\ y_m\end{bmatrix}\in\mathcal{R}(F)\Leftrightarrow\exists\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}s.t.\begin{bmatrix}y_1\\ \vdots\\ y_m\end{bmatrix}=F\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\Leftrightarrow\begin{bmatrix}\boldsymbol{i}_1 & \cdots & \boldsymbol{i}_m\end{bmatrix}\begin{bmatrix}y_1\\ \vdots\\ y_m\end{bmatrix}\in\mathcal{R}(f)$.

所以 $\sigma_{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n}:\mathcal{N}(f)\to\mathcal{N}(F),\quad \sigma_{\boldsymbol{i}_1,\cdots,\boldsymbol{i}_m}:\mathcal{R}(f)\to\mathcal{R}(F)$ 都是良定义的。容易验证这两个映射是线性的，我们略过。我们还需要证明它们即单又满。它们是单射这一点也是明显的，由 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n$; $\boldsymbol{i}_1,\cdots,\boldsymbol{i}_m$ 可得。下面验证是满射。

设 $\begin{bmatrix}x_1\\ v\ldots\\ x_n\end{bmatrix}\in\mathcal{N}(F)$, 那么 $f\left(\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix}\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\right)=\begin{bmatrix}\boldsymbol{i}_1 & \cdots & \boldsymbol{i}_m\end{bmatrix}F\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}=0$，所以 $\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix}\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\in\mathcal{N}(f)$，此时 $\sigma_{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n}\left(\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix}\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\right)=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$, 这就证明了 $\sigma_{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n}:\mathcal{N}(f)\to\mathcal{N}(F)$ 是满射。

设 $\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \in \mathcal{R}(F)$，那么存在 $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ 使得 $\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = F \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, 那么

$$f\left(\begin{bmatrix} \boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = \begin{bmatrix} \boldsymbol{i}_1 & \cdots & \boldsymbol{i}_m \end{bmatrix} F \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{i}_1 & \cdots & \boldsymbol{i}_m \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix},$$

即 $\boldsymbol{i} = \begin{bmatrix} \boldsymbol{i}_1 & \cdots & \boldsymbol{i}_m \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \in \mathcal{R}(f)$，此时 $\sigma_{\boldsymbol{i}_1,\cdots,\boldsymbol{i}_m}(\boldsymbol{i}) = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$, 这就证明了 $\sigma_{\boldsymbol{i}_1,\cdots,\boldsymbol{i}_n} : \mathcal{R}(f) \to \mathcal{R}(F)$ 是满射。

□

练习 7.5.7. 设 $\mathcal{V}$ 是数域 $\mathbb{F}$ 上的 n 维线性空间，$\mathbb{F}$ 可以看作自身上的线性空间，而 $\mathcal{V}$ 到 $\mathbb{F}$ 的线性映射称为 $\mathcal{V}$ 上的**线性函数**. 令 $\mathcal{V}^* = \mathrm{Hon}(\mathcal{V}, \mathbb{F})$，称为 $\mathcal{V}$ 的**对偶空间**. 证明，$\mathcal{V}^*$ 和 $\mathcal{V}$ 同构.

证明: 我们取 $\mathcal{V}$ 的一组基 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$, 我们定义 $f_1, \cdots, f_n \in \mathcal{V}^*$ 如下。由线性性，f_i 由它在 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 上的取值完全确定，我们定义

$$f_i(\boldsymbol{v}_j) = \delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

那么 $f_i(a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n) = a_i$.

- 对任意 $f \in \mathcal{V}^*$, 我们断言 $f = f(\boldsymbol{v}_1)f_1 + \cdots + f(\boldsymbol{v}_n)f_n$。要证明这一点，只需要证明对任意 $\boldsymbol{v}_i (1 \le i \le n)$ 我们有 $f(\boldsymbol{v}_i) = (f(\boldsymbol{v}_1)f_1 + \cdots + f(\boldsymbol{v}_n)f_n)(\boldsymbol{v}_i)$. 事实上

  $$(f(\boldsymbol{v}_1)f_1 + \cdots + f(\boldsymbol{v}_n)f_n)(\boldsymbol{v}_i) = f(\boldsymbol{v}_1)f_1(\boldsymbol{v}_i) + \cdots + f(\boldsymbol{v}_i)f_i(\boldsymbol{v}_i) + \cdots + f(\boldsymbol{v}_n)f_n(\boldsymbol{v}_i) = f(\boldsymbol{v}_i).$$

□

练习 7.5.8. 求导算子 $\boldsymbol{D}$ 定义了多项式空间 $\mathbb{F}[x]_n$ 上的线性变换,给定 $\mathbb{F}[x]_n$ 的两组基 $1, x, \ldots, x^{n-1}$ 和 $1, x, \ldots, \dfrac{1}{(n-1)!}x^{n-1}$.

1. 求两组基之间的过渡矩阵.
2. 分别求 $\boldsymbol{D}$ 在两组基下的矩阵.
3. 通过过渡矩阵验证这两个不同基下的矩阵相似.
4. 是否存在一组基，使得 $\boldsymbol{D}$ 在该组基下的矩阵是对角矩阵?

证明: 1. $\begin{bmatrix} 1 & x & \ldots & \frac{1}{(n-1)!}x^{n-1} \end{bmatrix} = \begin{bmatrix} 1 & x & \ldots & x^{n-1} \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & \frac{1}{(n-1)!} \end{bmatrix}$ 所以从

$1, x, \ldots, x^{n-1}$ 到 $1, x, \ldots, \dfrac{1}{(n-1)!}x^{n-1}$ 的过渡矩阵是 $T = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & \frac{1}{(n-1)!} \end{bmatrix}$, 这是对角矩阵，第 k 个对角元是 $\frac{1}{(k-1)!}$

2.
$$\begin{aligned} \boldsymbol{D}\begin{bmatrix} 1 & x & \ldots & x^{n-1} \end{bmatrix} &= \begin{bmatrix} 0 & 1 & \ldots & (n-1)x^{n-2} \end{bmatrix} \\ &= \begin{bmatrix} 1 & x & \ldots & x^{n-1} \end{bmatrix} \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & n-1 \\ & & & 0 \end{bmatrix} \end{aligned}$$

所以 $\boldsymbol{D}$ 在基 $1, x, \ldots, x^{n-1}$ 之下的矩阵是 $A_1 = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & n-1 \\ & & & 0 \end{bmatrix}$.

$$\begin{aligned} \boldsymbol{D}\begin{bmatrix} 1 & x & \ldots & \frac{x^{n-1}}{(n-1)!} \end{bmatrix} &= \begin{bmatrix} 0 & 1 & \ldots & \frac{x^{n-2}}{(n-2)!} \end{bmatrix} \\ &= \begin{bmatrix} 1 & x & \ldots & \frac{x^{n-1}}{(n-1)!} \end{bmatrix} \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} \end{aligned}$$

所以 $\boldsymbol{D}$ 在基 $1, x, \ldots, \frac{x^{n-1}}{(n-1)!}$ 之下的矩阵是 $A_2 = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}$.

3. 通过过渡矩阵，我们有 $A_2 = T^{-1}A_1T$, 直接通过矩阵乘法，计算也可得。

4. 不存在。否则，A_1 可以相似对角化。但是 A_1 只有一个特征值 0, 其代数重数是 n, 几何重数是 1, 所以不能相似对角化。

□

练习 7.5.9. 设 $\mathcal{V}$ 是 n 维线性空间，f 是其上的线性变换，又设存在向量 $\boldsymbol{a} \in \mathcal{V}$，使得 $f^{n-1}(\boldsymbol{a}) \neq 0$，且 $f^n(\boldsymbol{a}) = 0$. 证明，$\mathcal{V}$ 存在一组基，使得 f 在该组基下的矩阵是

$$J_n = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}.$$

证明: 令 $\boldsymbol{v}_1=\boldsymbol{a},\boldsymbol{v}_2=f(\boldsymbol{a}),\cdots,\boldsymbol{v}_n=f^{n-1}(\boldsymbol{a})$. 我们断言 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 线性无关，从而构成 $\mathcal{V}$ 的一组基。设 $a_1\boldsymbol{v}_1+\cdots+a_n\boldsymbol{v}_n=0$, 那么

$$0=f^{n-1}(a_1\boldsymbol{v}_1+\cdots+a_n\boldsymbol{v}_n)=a_1f^{n-1}(\boldsymbol{a})$$

于是 $a_1=0$, 代回就有设 $a_2\boldsymbol{v}_2+\cdots+a_n\boldsymbol{v}_n=0$, 那么

$$0=f^{n-2}(a_2\boldsymbol{v}_2+\cdots+a_n\boldsymbol{v}_n)=a_2f^{n-1}(\boldsymbol{a})$$

于是 $a_2=0$. 依次类推我们可以证明 $a_1=a_2=\cdots=a_n=0$.

$$f\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n\end{bmatrix}=\begin{bmatrix}\boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n & 0\end{bmatrix}=\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n\end{bmatrix}\begin{bmatrix}0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1\\ & & & 0\end{bmatrix}$$

□

这就证明了结论。

练习 7.5.10. 考虑练习 7.5.9中的方阵 J_n，判断 J_n 与 J_n^{T} 是否相似.

证明: 我们令 $\mathcal{V},f,\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 如上题。我们考虑 f 在基 $\boldsymbol{v}_n,\boldsymbol{v}_{n-1},\cdots,\boldsymbol{v}_1$ 之下的矩阵：

$$f\begin{bmatrix}\boldsymbol{v}_n & \boldsymbol{v}_{n-1} & \cdots & \boldsymbol{v}_1\end{bmatrix}=\begin{bmatrix}0 & \boldsymbol{v}_n & \cdots & \boldsymbol{v}_2\end{bmatrix}=\begin{bmatrix}\boldsymbol{v}_n & \boldsymbol{v}_{n-1} & \cdots & \boldsymbol{v}_1\end{bmatrix}\begin{bmatrix}0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0\end{bmatrix}$$

这就是说，J_n 与 J_n^{T} 是同一个线性映射在不同基之下的矩阵，所以它们是相似的。 □

练习 7.5.11. 已知 $\mathbb{F}^3$ 上的线性变换 f 在标准基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 下的矩阵是 $A=\begin{bmatrix}15 & -11 & 5\\ 20 & -15 & 8\\ 8 & -7 & 6\end{bmatrix}$. 设 $\boldsymbol{t}_1=\begin{bmatrix}2\\3\\1\end{bmatrix},\boldsymbol{t}_2=\begin{bmatrix}3\\4\\1\end{bmatrix},\boldsymbol{t}_3=\begin{bmatrix}1\\2\\2\end{bmatrix}$ 是 $\mathbb{F}^3$ 的另一组基，求 f 在这组基下的矩阵.

证明: 从标准基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 到 $\boldsymbol{t}_1,\boldsymbol{t}_2,\boldsymbol{t}_3$ 的过渡矩阵是 $\begin{bmatrix}2 & 3 & 1\\ 3 & 4 & 2\\ 1 & 1 & 2\end{bmatrix}$, 所以 f 在基 $\boldsymbol{t}_1,\boldsymbol{t}_2,\boldsymbol{t}_3$ 下的矩阵是

$$\begin{bmatrix}2 & 3 & 1\\ 3 & 4 & 2\\ 1 & 1 & 2\end{bmatrix}^{-1}\begin{bmatrix}15 & -11 & 5\\ 20 & -15 & 8\\ 8 & -7 & 6\end{bmatrix}\begin{bmatrix}2 & 3 & 1\\ 3 & 4 & 2\\ 1 & 1 & 2\end{bmatrix}=\begin{bmatrix}1 & & \\ & 2 & \\ & & 3\end{bmatrix}$$

. □

练习 7.5.12. 证明命题 *7.5.12.* 设 $\mathbf{F}$ 上 n 维线性空间 $\mathcal{V}$, f 是其上的线性变换，它在某基下的矩阵是 F, 则 F 可对角化，当且仅当 f 有 n 个线性无关的特征向量。

证明: 设 f 在基 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 之下的矩阵是 F，则 $f\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix} = \begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix} F$

- 若 F 可对角化，那么存在可逆矩阵 T 使得 $T^{-1}FT$ 是对角矩阵 $\operatorname{diag}(\lambda_1, \cdots, \lambda_n)$。令 $\begin{bmatrix}\boldsymbol{t}_1 & \cdots & \boldsymbol{t}_n\end{bmatrix} = \begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix} T$, T 可逆，所以 $\boldsymbol{t}_1, \cdots, \boldsymbol{t}_n$ 也是一组基，且 T 是从 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 到 $\boldsymbol{t}_1, \cdots, \boldsymbol{t}_n$ 的过渡矩阵。根据命题 5.4.6，f 在基 $\boldsymbol{t}_1, \cdots, \boldsymbol{t}_n$ 之下的矩阵就是 $\operatorname{diag}(\lambda_1, \cdots, \lambda_n)$，即 $f(\boldsymbol{t}_i) = \lambda_i \boldsymbol{t}_i (1 \le i \le n)$.
- 若 f 有 n 个线性无关的特征向量 $\boldsymbol{t}_1, \cdots, \boldsymbol{t}_n$，那么 f 在基 $\boldsymbol{t}_1, \cdots, \boldsymbol{t}_n$ 之下的矩阵是 $\operatorname{diag}(\lambda_1, \cdots, \lambda_n)$. 设从 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 到 $\boldsymbol{t}_1, \cdots, \boldsymbol{t}_n$ 的过渡矩阵是 T。根据命题 5.4.6 就有 $T^{-1}FT = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$.

□

练习 7.5.13. 证明命题 *7.5.13.*

证明: 这是命题 7.5.6 的直接推论。 □

练习 7.5.14. 设三维线性空间 $\mathcal{V}$ 有一组基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$，其上的线性变换 f 在该组基下的矩阵是

$$A = \begin{bmatrix} 2 & 3 & 2 \\ 1 & 8 & 2 \\ -2 & -14 & -3 \end{bmatrix}.$$

1. 求 f 的全部特征值和特征向量.

2. 判断是否存在一组基，使得 f 在该组基下的矩阵是对角矩阵. 如果存在，写出这组基及对应的对角矩阵.

证明:

1. $\det(\lambda I_3 - A) = (\lambda - 1)(\lambda - 3)^2$, 因此 A 有两个特征值 $1, 3$, 特征值 1 的几何重数是 1, $\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$ 是一个属于 1 特征向量。特征值 3 的几何重数是 1, $\begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$ 是一个属于 3 的特征向量。所以 f 有两个特征值 $1, 2$, $-2\boldsymbol{e}_1 + \boldsymbol{e}_3$ 和 $-\boldsymbol{e}_1 + \boldsymbol{e}_1 - 2\boldsymbol{e}_2$ 分别是属于这两个特征值的特征子空间的一组基。
2. 不存在。若存在，则 A 可相似对角化。但是 A 的特征值 3 的代数重数大于几何重数，所以 A 不可相似对角化。

□

练习 7.5.15. 设 4 维线性空间 $\mathcal{V}$ 有一组基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$，其上的线性变换 f 在该组基下的矩阵是

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

1. 求 f 的全部特征值和特征向量.

2. 判断是否存在一组基，使得 f 在该组基下的矩阵是对角矩阵. 如果存在，写出这组基及对应的对角矩阵.

证明: 1. $\det(\lambda I_4 - A) = \lambda^2(\lambda-1)^2$. A 有两个特征值：$0, 1$. 属于特征值 0 的特征子空间的一组基是 $\begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}$; 属于特征值 1 的特征子空间的一组基是 $\begin{bmatrix}1\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$. 那么 f 的特征值是 $0, 1$, $\boldsymbol{e}_2, \boldsymbol{e}_3$ 是属于特征值 0 的特征子空间的一组基，$\boldsymbol{e}_1 + \boldsymbol{e}_3, \boldsymbol{e}_4$ 是属于特征值 1 的特征子空间的一组基。

2. 存在，$\boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_1 + \boldsymbol{e}_3, \boldsymbol{e}_4$ 是这样的一组基，f 在这组基下的矩阵是 $\mathrm{diag}(0, 0, 1, 1)$

□

练习 7.5.16. 设 $B = \begin{bmatrix}-1 & -1\\ 2 & 1\end{bmatrix}$，在 $\mathbb{F}^{2\times 2}$ 中定义如下变换：

$$f(X) = B^{-1}XB, \quad \forall X \in \mathbb{F}^{2\times 2}.$$

1. 证明 f 是线性变换.
2. 求 f 的全部特征值和特征向量.

证明: 1. $f(aX + bY) = B^{-1}(aX + bY)B = aB^{-1}XB + bB^{-1}YB = af(X) + bf(Y)$. 所以 f 是线性映射。

2. 我们取 $\mathbb{F}^{2\times 2}$ 的一组基 $E_{11}.E_{12}, E_{21}, E_{22}$,

$$\begin{aligned}
f(E_{11}) &= \begin{bmatrix}-1 & -1\\ 2 & 2\end{bmatrix}\\
f(E_{12}) &= \begin{bmatrix}2 & 1\\ -4 & -2\end{bmatrix}\\
f(E_{21}) &= \begin{bmatrix}-1 & -1\\ 1 & 1\end{bmatrix}\\
f(E_{22}) &= \begin{bmatrix}2 & 1\\ -2 & -1\end{bmatrix}
\end{aligned}$$

f 在这组基下的矩阵是

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

$\det(\lambda I_4 - A) = (\lambda-1)^2(\lambda+1)^2$, 所以 A 有两个特征值：$1, -1$. $\begin{bmatrix}-2\\-1\\2\\0\end{bmatrix}, \begin{bmatrix}1\\0\\0\\1\end{bmatrix}$ 是属于

特征值 1 的特征子空间的一组基；$\begin{bmatrix}0\\1\\2\\0\end{bmatrix},\begin{bmatrix}-1\\-1\\0\\1\end{bmatrix}$ 是属于特征值 -1 的特征子空间的一组基。所以 f 有两个特征值 $1,-1$. $-2E_{11}-E_{12}+2E_{21}, E_{11}+E_{22}$ 是属于特征值 1 的特征子空间的一组基；$E_{12}+2E_{21}, -E_{11}-E_{12}+E_{22}$ 是属于特征值 -1 的特征子空间的一组基。

□

阅读 7.5.17 (矩阵函数).

第八章　内积空间

8.1　欧 氏 空 间

练习 8.1.1. 在 $\mathbb{R}^2$ 中，对任意 $\boldsymbol{a}=\begin{bmatrix}a_1\\a_2\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}b_1\\b_2\end{bmatrix}$，定义二元函数

$$f(\boldsymbol{a},\boldsymbol{b})=a_1b_1-a_1b_2-a_2b_1+2a_2b_2.$$

判断 $f(\boldsymbol{a},\boldsymbol{b})$ 是否为 $\mathbb{R}^2$ 上的一个内积.

证明:

- 对称性：$f(\boldsymbol{b},\boldsymbol{a})=b_1a_1-b_1a_2-b_2a_1+2b_2a_2=f(\boldsymbol{b},\boldsymbol{a})$, 所以 f 是对称的。
- 双线性：对任意 $\boldsymbol{a}=\begin{bmatrix}a_1\\a_2\end{bmatrix},\boldsymbol{a}'=\begin{bmatrix}a_1'\\a_2'\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}b_1\\b_2\end{bmatrix},\boldsymbol{b}'=\begin{bmatrix}b_1'\\b_2'\end{bmatrix},k,k'\in\mathbb{R}$,

$$\begin{aligned}f(k\boldsymbol{a}+k'\boldsymbol{a}',\boldsymbol{b})&=(ka_1+k'a_1')b_1-(ka_1+k'a_1')b_1-(ka_2+k'a_2')b_2+2(ka_2+k'a_2')b_2\\&=k(a_1b_1-a_1b_2-a_2b_1+2a_2b_2)+k'(a_1'b_1-a_1'b_2-a_2'b_1+2a_2'b_2)\\&=kf(\boldsymbol{a},\boldsymbol{b})+k'f(\boldsymbol{a}',\boldsymbol{b})\end{aligned}$$

  结合对称性，

$$f(\boldsymbol{a},k\boldsymbol{b}+k'\boldsymbol{b}')=f(k\boldsymbol{b}+k'\boldsymbol{b}',\boldsymbol{a})=kf(\boldsymbol{b},\boldsymbol{a})+k'f(\boldsymbol{b}',\boldsymbol{a})=kf(\boldsymbol{a},\boldsymbol{b})+k'f(\boldsymbol{a},\boldsymbol{b}').$$

- 正定性：$f(\boldsymbol{a},\boldsymbol{a})=a_1^2-2a_1a_2+2a_2^2=(a_1-a_2)^2+a_2^2\geq 0$，取等号，当且仅当 $a_1-a_2=a_2=0$, 即 $\boldsymbol{a}=0$.

□

练习 8.1.2. 证明，$\mathbb{R}^n$ 上的二元函数 $\boldsymbol{b}^{\mathrm{T}}A\boldsymbol{a}$ 定义了一个内积，当且仅当 A 是对称正定矩阵.

证明:

- 若 $f(\boldsymbol{a},\boldsymbol{b})=\boldsymbol{b}^{\mathrm{T}}A\boldsymbol{a}$ 定义了一个内积, 由对称性,对任意 $\boldsymbol{a},\boldsymbol{b}$,都有 $\boldsymbol{b}^{\mathrm{T}}A\boldsymbol{a}=\boldsymbol{a}^{\mathrm{T}}A\boldsymbol{b}$. 对任意 i,j, 我们取 $\boldsymbol{a}=\boldsymbol{e}_i,\boldsymbol{b}=\boldsymbol{e}_j$, 那么上式说明 $a_{ji}=a_{ij}$, 所以 A 是对称矩阵。再由 $f(\boldsymbol{a},\boldsymbol{a})$ 的正定性知 A 是正定矩阵。
- 反之人, 若 A 是实对称正定矩阵；首先显然 $f(\boldsymbol{a},\boldsymbol{b})$ 满足双线性性；A 对称则 $\boldsymbol{b}^{\mathrm{T}}A\boldsymbol{a}=(\boldsymbol{b}^{\mathrm{T}}A\boldsymbol{a})^{\mathrm{T}}=\boldsymbol{a}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{b}=\boldsymbol{a}^{\mathrm{T}}A\boldsymbol{b}$, 所以 $f(\boldsymbol{a},\boldsymbol{b})=f(\boldsymbol{b},\boldsymbol{a})$；$A$ 正定则显然 f 满足正定性，所以 f 是一个内积。

□

练习 8.1.3. 证明，矩阵空间 $\mathbb{R}^{m\times n}$ 关于如下二元函数构成欧氏空间：$\langle A,B\rangle=\mathrm{trace}(B^{\mathrm{T}}A)$.

证明: • 对称性: $\mathrm{trace}(B^{\mathrm{T}}A) = \mathrm{trace}((B^{\mathrm{T}}A)^{\mathrm{T}}) = \mathrm{trace}(A^{\mathrm{T}}B)$, 即 $\langle A, B\rangle = \langle B, A\rangle$.

- 双线性:

$$\begin{aligned}\langle k_1A_1 + k_2A_2, B\rangle &= \mathrm{trace}(B^{\mathrm{T}}(k_1A_1 + k_2A_2)) \\ &= k_1\mathrm{trace}(B^{\mathrm{T}}A_1) + k_2\mathrm{trace}(B^{\mathrm{T}}A_2) \\ &= k_1\langle A_1, B\rangle + k_2\langle A_2, B\rangle\end{aligned}$$

  结合对称性，可得关于第二个变量的线性性。

- 正定性: 设 $A = \begin{bmatrix} a_{ij} \end{bmatrix}$, 那么

$$\langle A, A\rangle = \mathrm{trace}(A^{\mathrm{T}}A) = \sum_{i,j} a_{ij}^2 \geq 0;$$

  且 $\langle A, A\rangle = 0$ 当且仅当 $a_{ij} = 0, \forall i, j$, 当且仅当 $A = 0$.

□

练习 8.1.4. 证明正交向量组线性无关.

证明: 设 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_r$ 是正交向量组，$a_1\boldsymbol{v}_1 + \cdots + a_r\boldsymbol{v}_r = 0$, 对任意 $1 \leq i \leq r$,

$$a_i\langle \boldsymbol{v}_i, \boldsymbol{v}_i\rangle = \langle a_1\boldsymbol{v}_1 + \cdots + a_r\boldsymbol{v}_r, \boldsymbol{v}_i\rangle = 0$$

所以 $a_i = 0$. □

练习 8.1.5. 设三维欧氏空间 $\mathcal{V}$ 的一组基是 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$，它的度量矩阵是 $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 10 & -2 \\ 1 & -2 & 2 \end{bmatrix}$. 求 $\mathcal{V}$ 的一组标准正交基.

证明: 令 $\boldsymbol{b}_1 = \boldsymbol{a}_1, \boldsymbol{b}_2 = \boldsymbol{a}_2 - \dfrac{\langle \boldsymbol{a}_2, \boldsymbol{b}_1\rangle}{\langle \boldsymbol{b}_1, \boldsymbol{b}_1\rangle} = \boldsymbol{a}_2$,

$$\boldsymbol{a}_3 = \boldsymbol{a}_3 - \frac{\langle \boldsymbol{a}_3, \boldsymbol{b}_1\rangle}{\langle \boldsymbol{b}_1, \boldsymbol{b}_1\rangle}\boldsymbol{b}_1 - \frac{\langle \boldsymbol{a}_3, \boldsymbol{b}_2\rangle}{\langle \boldsymbol{b}_2, \boldsymbol{b}_2\rangle}\boldsymbol{b}_2 = a_3 - a_1 + \frac{1}{5}a_2$$

那么 $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$ 就是两两正交的向量，我们再将他们单位化，就可以得到一组标准正交基了。

- $\langle \boldsymbol{b}_1, \boldsymbol{b}_1\rangle = 1$, 所以 $\tilde{\boldsymbol{b}}_1 = \boldsymbol{b}_1 = \boldsymbol{a}_1$.
- $\langle \boldsymbol{b}_2, \boldsymbol{b}_2\rangle = 10$, 所以 $\tilde{\boldsymbol{b}}_2 = \dfrac{\boldsymbol{a}_2}{\sqrt{10}}$.
- $\langle \boldsymbol{b}_3, \boldsymbol{b}_3\rangle = \frac{3}{5}$, 所以 $\tilde{\boldsymbol{b}}_3 = \sqrt{\dfrac{5}{3}}\left(a_3 - a_1 + \dfrac{1}{5}a_2\right)$

那么 $\tilde{\boldsymbol{b}}_1, \tilde{\boldsymbol{b}}_2, \tilde{\boldsymbol{b}}_3$ 就是一组标准正交基。 □

练习 8.1.6. 设 $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3$ 是三维欧氏空间 $\mathcal{V}$ 的一组标准正交基，令

$$\boldsymbol{b}_1 = \frac{1}{3}(2\boldsymbol{q}_1 - \boldsymbol{q}_2 + 2\boldsymbol{q}_3), \quad \boldsymbol{b}_2 = \frac{1}{3}(2\boldsymbol{q}_1 + 2\boldsymbol{q}_2 - \boldsymbol{q}_3), \quad \boldsymbol{b}_3 = \frac{1}{3}(\boldsymbol{q}_1 - 2\boldsymbol{q}_2 - 2\boldsymbol{q}_3).$$

证明，$\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$ 也是 $\mathcal{V}$ 的一组标准正交基.

证明: $\begin{bmatrix}\boldsymbol{b}_1 & \boldsymbol{b}_2 & \boldsymbol{b}_3\end{bmatrix} = \begin{bmatrix}\boldsymbol{q}_1 & \boldsymbol{q}_2 & \boldsymbol{q}_3\end{bmatrix}\begin{bmatrix}\frac{2}{3} & \frac{2}{3} & \frac{1}{3}\\ -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3}\\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3}\end{bmatrix}$, 令 $T = \begin{bmatrix}\frac{2}{3} & \frac{2}{3} & \frac{1}{3}\\ -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3}\\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3}\end{bmatrix}$. 由于 $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3$ 是标准正交基，因此内积在 $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3$ 之下的度量矩阵是 I_3, 那么内积在 $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$ 之下的度量矩阵就是 $T^{\mathrm{T}} I_3 T = I_3$, 所以 $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$ 也是 $\mathcal{V}$ 的一组标准正交基。 □

练习 8.1.7. 设 $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3, \boldsymbol{q}_4, \boldsymbol{q}_5$ 是 5 维欧氏空间 $\mathcal{V}$ 的一组标准正交基，令

$$\boldsymbol{a}_1 = \boldsymbol{q}_1 + \boldsymbol{q}_5, \quad \boldsymbol{a}_2 = \boldsymbol{q}_1 - \boldsymbol{q}_2 + \boldsymbol{q}_4, \quad \boldsymbol{a}_3 = 2\boldsymbol{q}_1 + \boldsymbol{q}_2 + \boldsymbol{q}_3.$$

求 $\mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$ 的一组标准正交基.

证明: $\boldsymbol{b}_1 = \boldsymbol{a}_1, \boldsymbol{b}_2 = \boldsymbol{a}_2 - \dfrac{\langle \boldsymbol{a}_2, \boldsymbol{b}_1\rangle}{\langle \boldsymbol{b}_1, \boldsymbol{b}_1\rangle}\boldsymbol{b}_1 = \frac{1}{2}\boldsymbol{q}_1 - \boldsymbol{q}_2 + \boldsymbol{q}_4 - \frac{1}{2}\boldsymbol{q}_5$, $\boldsymbol{b}_3 = \boldsymbol{a}_3 - \dfrac{\langle \boldsymbol{a}_3, \boldsymbol{b}_1\rangle}{\langle \boldsymbol{b}_1, \boldsymbol{b}_1\rangle}\boldsymbol{b}_1 - \dfrac{\langle \boldsymbol{a}_3, \boldsymbol{b}_2\rangle}{\langle \boldsymbol{b}_2, \boldsymbol{b}_2\rangle}\boldsymbol{b}_2 = \boldsymbol{q}_1 + \boldsymbol{q}_2 + \boldsymbol{q}_3 - \boldsymbol{q}_5$. 然后把它们单位化

- $\langle \boldsymbol{b}_1, \boldsymbol{b}_2\rangle = 2$, 则 $\tilde{\boldsymbol{b}}_1 = \sqrt{\dfrac{1}{2}}(\boldsymbol{q}_1 + \boldsymbol{a}_5)$.
- $\langle \boldsymbol{b}_2, \boldsymbol{b}_2\rangle = \frac{5}{2}$，则 $\tilde{\boldsymbol{b}}_2 = \sqrt{\dfrac{2}{5}}\left(\dfrac{1}{2}\boldsymbol{q}_1 - \boldsymbol{q}_2 + \boldsymbol{q}_4 - \dfrac{1}{2}\boldsymbol{q}_5\right)$.
- $\langle \boldsymbol{b}_3, \boldsymbol{b}_3\rangle = 4$, 则 $\tilde{\boldsymbol{b}}_3 = \dfrac{1}{2}(\boldsymbol{q}_1 + \boldsymbol{q}_2 + \boldsymbol{q}_3 - \boldsymbol{q}_5)$

那么 $\tilde{\boldsymbol{b}}_1, \tilde{\boldsymbol{b}}_2, \tilde{\boldsymbol{b}}_3$ 就是 $\mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$ 的一组标准正交基。 □

练习 8.1.8. 计算例 *8.1.11* 中 *Lengendre* 多项式作为多项式空间中向量的范数，并说明该向量组不是正交单位向量组.

证明: 根据 Legendre 函数的求导表达式 $P_n(x) = \dfrac{1}{2^n n!}\dfrac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n$, 那么

$$\langle P_n, P_n\rangle = \int_{-1}^{1}\left(\frac{1}{2^n n!}\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n\right)^2\mathrm{d}x = \frac{1}{(2^n n!)^2}\int_{-1}^{1}\left(\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n\right)^2\mathrm{d}x$$

利用分部积分：

$$\begin{aligned}\int_{-1}^{1}\left(\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n\right)^2\mathrm{d}x &= \int_{-1}^{1}\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n\mathrm{d}\left(\frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}}(x^2-1)^n\right)\\ &= \frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n\frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}}(x^2-1)^n\bigg|_{-1}^{1} - \int_{-1}^{1}\frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}}(x^2-1)^n\mathrm{d}\left(\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n\right)\end{aligned}$$

注意到 $(x^2-1)^n = (x-1)^n(x+1)^n$, 所以 $\dfrac{\mathrm{d}^k}{\mathrm{d}x^k}(x^2-1)^n\bigg|_{-1} = \dfrac{\mathrm{d}^k}{\mathrm{d}x^k}(x^2-1)^n\bigg|_{1} = 0$, 所以

$$\begin{aligned}\int_{-1}^{1}\left(\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n\right)^2\mathrm{d}x &= -\int_{-1}^{1}\frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}}(x^2-1)^n\mathrm{d}\left(\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n\right)\\ &= -\int_{-1}^{1}\frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}}(x^2-1)^n\frac{\mathrm{d}^{n+1}}{\mathrm{d}x^{n+1}}(x^2-1)^n\mathrm{d}x\end{aligned}$$

继续分部积分，最终可得

$$\int_{-1}^{1}\left(\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n\right)^2\mathrm{d}x = (-1)^n\int_{-1}^{1}(x^2-1)^n\mathrm{d}x$$

换元 $x=\cos\theta$:

$$\int_{-1}^{1}(x^2-1)^n\mathrm{d}x=\int_{\pi}^{0}(-\sin^2\theta)^n(-\sin\theta\mathrm{d}\theta)=\int_0^{\pi}\sin^{2n+1}\theta\mathrm{d}\theta.$$

做这个三角积分的方法如下：

$$\int_0^{\pi}\sin^{2n+1}\theta\mathrm{d}\theta=-\int_0^{\pi}\sin^{2n}\mathrm{d}\cos\theta=-\sin^{2n}\theta\cos\theta\Big|_0^{\pi}+\int_0^{\pi}\cos\theta\mathrm{d}\sin^{2n}\theta$$

$-\sin^{2n}\theta\cos\theta\Big|_0^{\pi}=0$，所以

$$\int_0^{\pi}\sin^{2n+1}\theta\mathrm{d}\theta=2n\int_0^{\pi}\sin^{2n-1}\theta\cos^2\theta\mathrm{d}\theta=2n\int_0^{\pi}\sin^{2n-1}\theta\mathrm{d}\theta-2n\int_0^{\pi}\sin^{2n+1}\theta\mathrm{d}\theta$$

所以

$$\int_0^{\pi}\sin^{2n+1}\mathrm{d}\theta=\frac{2n}{2n+1}\int_0^{\pi}\sin^{2n-1}\theta\mathrm{d}\theta$$

逐步递推即可。具体结果留给读者自行计算。

因为范数不是 1，所以 Legendre 多项式不是单位向量。 □

练习 8.1.9. *设 f 是 n 维欧氏空间 $\mathcal{V}$ 内的线性函数，证明，存在唯一的固定向量 $\boldsymbol{b}\in\mathcal{V}$，使得对任意 $\boldsymbol{a}\in\mathcal{V}$，都有 $f(\boldsymbol{a})=\langle\boldsymbol{a},\boldsymbol{b}\rangle$.*

证明*:* 取 $\mathcal{V}$ 的一组标准正交基 $\boldsymbol{q}_1,\cdots,\boldsymbol{q}_n$, 令 $\boldsymbol{b}=f(\boldsymbol{q}_1)\boldsymbol{q}_1+\cdots+f(\boldsymbol{q}_n)\boldsymbol{q}_n$, 那么对任意 $1\le i\le n$,

$$f(\boldsymbol{q}_i)=\langle\boldsymbol{q}_i,f(\boldsymbol{q}_1)\boldsymbol{q}_1+\cdots+f(\boldsymbol{q}_n)\boldsymbol{q}_n\rangle=\langle\boldsymbol{q}_i,\boldsymbol{b}\rangle;$$

对任意 $\boldsymbol{a}$, $\boldsymbol{a}=x_1\boldsymbol{q}_1+\cdots+x_n\boldsymbol{q}_n$, 那么

$$f(\boldsymbol{a})=\sum_{i=1}^{n}x_if(\boldsymbol{q}_i)=\sum_{i=1}^{n}x_i\langle\boldsymbol{q}_i,\boldsymbol{b}\rangle=\langle\boldsymbol{a},\boldsymbol{b}\rangle.$$

□

练习 8.1.10. *证明命题 8.1.15.*

证明*:* 这即是要证明把内积限制在 $\mathcal{M}$ 之后也满足：对称性、双线线性、正定性。这是明显的，没什么好证明的。 □

练习 8.1.11. *证明命题 8.1.17 第 2 条和第 3 条。*

证明*:*

2. 根据维数公式有 $\dim(\mathcal{V})=\dim\mathcal{M}+\dim\mathcal{M}^{\perp}-\dim\mathcal{M}\cap\mathcal{M}^{\perp}=\dim\mathcal{M}+\dim\mathcal{M}^{\perp}$.

3 根据正交补的定义，我们有 $\mathcal{M}\subseteq(\mathcal{M}^{\perp})^{\perp}$. 而它们的维数都等于 $\dim\mathcal{V}-\dim\mathcal{M}^{\perp}$, 就有 $\mathcal{M}=(\mathcal{M}^{\perp})^{\perp}$

□

练习 8.1.12. *考虑欧氏空间 $C[-1,1]$ 及其内积 $\langle f,g\rangle=\int_{-1}^{1}f(x)g(x)\mathrm{d}x$. 证明奇函数组成的子空间和偶函数组成的子空间互为正交补.*

证明: 设奇函数空间为 $\mathcal{V}$, 偶函数空间为 $\mathcal{W}$. 对任意 $f \in C[-1,1]$, 记 $\bar{f}$ 为函数 $\bar{f}(x) = f(-x)$, 那么 $f = \frac{f+\bar{f}}{2} + \frac{f-\bar{f}}{2}$, 而 $\frac{f+\bar{f}}{2} \in \mathcal{V}$, $\frac{f-\bar{f}}{2} \in \mathcal{W}$, 所以 $C[-1,1] = \mathcal{V} + \mathcal{W}$. 又 $\mathcal{V} \cap \mathcal{W} = 0$, 所以 $C[-1,1] = \mathcal{V} \oplus \mathcal{W}$. 对任意 $f \in \mathcal{V}, g \in \mathcal{W}$,

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\mathrm{d}x = \int_{1}^{-1} f(-x)g(-x)\mathrm{d}(-x) = -\langle f, g\rangle$$

所以 $\mathcal{V} \subseteq \mathcal{W}^{\perp}, \mathcal{W} \subseteq \mathcal{V}^{\perp}$

对任意 $f \in \mathcal{W}^{\perp}$, 令 $f = f_1 + f_2$, 其中 $f_1 \in \mathcal{V}, f_2 \in \mathcal{W}$, 那么 $\langle f, f_2\rangle = 0$, 即 $\langle f_1, f_2\rangle + \langle f_2, f_2\rangle = 0$. 而 $\langle f_1, f_2\rangle = 0$, 所以 $\langle f_2, f_2\rangle = 0$, 则 $f_2 = 0$. 因此 $f = f_1 \in \mathcal{V}$, 这就证明了 $\mathcal{V} \subseteq \mathcal{W}^{\perp}$. 同理可证 $\mathcal{W} = \mathcal{V}^{\perp}$. □

练习 8.1.13. 设 $\mathcal{M}$ 是 n 维欧氏空间 $\mathcal{V}$ 的子空间, 在 $\mathcal{M}$ 中取一组标准正交基 $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_r$. 证明, $\mathcal{V}$ 中任意向量 $\boldsymbol{a}$ 在 $\mathcal{M}$ 上的正交投影 $\boldsymbol{a}_1 = \sum_{i=1}^{r} \langle \boldsymbol{a}, \boldsymbol{q}_i\rangle \boldsymbol{q}_i$.

证明: 因为 $\boldsymbol{a}_1 \in \mathcal{M}$, 只需证明 $\boldsymbol{a} - \boldsymbol{a}_1 \in \mathcal{M}^{\perp}$, 或者等价地只需证明 $\forall \boldsymbol{b} \in \mathcal{M}, \langle \boldsymbol{a} - \boldsymbol{a}_1, \boldsymbol{b}\rangle = 0$, 或者等价地只需证明 $\forall \boldsymbol{b} \in \mathcal{M}, \langle \boldsymbol{a}, \boldsymbol{b}\rangle = \langle \boldsymbol{a}_1, \boldsymbol{b}\rangle$, 或者等价地只需证明 $\forall 1 \leq i \leq r, \langle \boldsymbol{a}, \boldsymbol{q}_i\rangle = \langle \boldsymbol{a}_1, \boldsymbol{q}_i\rangle$. 直接由定义知: $\langle \boldsymbol{a}, \boldsymbol{q}_i\rangle = \langle \boldsymbol{a}_1, \boldsymbol{q}_i\rangle$. □

练习 8.1.14. 设 $\mathcal{M}$ 是 n 维欧氏空间 $\mathcal{V}$ 的子空间, $\mathbf{P}_{\mathcal{M}}$ 是 $\mathcal{V}$ 到 $\mathcal{M}$ 上的正交投影. 证明

$$\langle \mathbf{P}_{\mathcal{M}}(\boldsymbol{a}), \boldsymbol{b}\rangle = \langle \boldsymbol{a}, \mathbf{P}_{\mathcal{M}}(\boldsymbol{b})\rangle, \quad \forall \boldsymbol{a}, \boldsymbol{b} \in \mathcal{V}.$$

证明: 设 $\boldsymbol{a} = \boldsymbol{a}_1 + \boldsymbol{a}_2, \boldsymbol{b} = \boldsymbol{b}_1 + \boldsymbol{b}_2$, 其中 $\boldsymbol{a}_1, \boldsymbol{b}_1 \in \mathcal{M}, \boldsymbol{a}_2, \boldsymbol{b}_2 \in \mathcal{M}^{\perp}$, 那么 $\mathbf{P}_{\mathcal{M}}(\boldsymbol{a}) = \boldsymbol{a}_1, \mathbf{P}_{\mathcal{M}}(\boldsymbol{b}) = \boldsymbol{b}_1$, 那么

$$\begin{aligned} \langle \mathbf{P}_{\mathcal{M}}(\boldsymbol{a}), \boldsymbol{b}\rangle &= \langle \boldsymbol{a}_1, \boldsymbol{b}_1 + \boldsymbol{b}_2\rangle = \langle \boldsymbol{a}_1, \boldsymbol{b}_1\rangle \\ \langle \boldsymbol{a}, \mathbf{P}_{\mathcal{M}}(\boldsymbol{b})\rangle &= \langle \boldsymbol{a}_1 + \boldsymbol{a}_2, \boldsymbol{b}_1\rangle = \langle \boldsymbol{a}_1, \boldsymbol{b}_1\rangle \end{aligned}$$

□

阅读 8.1.15 (无限维欧氏空间中的正交投影).

8.2 欧氏空间上的线性映射

练习 8.2.1 (线性变换的范数). 类似于定义 *6.3.6*, 欧氏空间 $\mathcal{V}$ 上向量的范数可以自然地诱导出 $\mathcal{V}$ 上线性变换的范数, 即 $\sup_{\boldsymbol{x} \neq 0} \frac{\|f(\boldsymbol{x})\|}{\|\boldsymbol{x}\|}$ 称为 f 的**范数**, 记为 $\|f\|$. 注意 $\|f\| = +\infty$ 有可能成立.

对练习 *7.3.9*中的线性变换 $\mathbf{D}$ 和 $\mathbf{A}$, 证明 $\|\mathbf{D}\| \|\mathbf{A}\| \geq \frac{1}{2}$.

这是量子力学中 *Heisenberg* **不确定性原理**的一个简化模型.

证明:

$$1 = \|\boldsymbol{I}\| = \|\mathbf{AD} - \mathbf{DA}\| \leq \|\mathbf{AD}\| + \|\mathbf{DA}\| \leq 2 \|\mathbf{D}\| \|\mathbf{A}\|$$

所以 $\|\mathbf{D}\| \|\mathbf{A}\| \geq \frac{1}{2}$. □

练习 8.2.2. 证明伴随映射的基本结论:

1. 若 $g=f^*$，则 $f=g^*$.　　　　　　　　*2.* $(g\circ f)^*=f^*\circ g^*$.

证明: 1. 设 $f\in\mathrm{Hom}(\mathcal{U},\mathcal{V}),g\in\mathrm{Hom}(\mathcal{V},\mathcal{U})$ $f=g^*\Leftrightarrow\forall\boldsymbol{u}\in\mathcal{U},\boldsymbol{v}\in\mathcal{V}$, $\langle g(\boldsymbol{v}),\boldsymbol{u}\rangle_{\mathcal{U}}=\langle\boldsymbol{v},f(\boldsymbol{u})\rangle_{\mathcal{V}}$. 由内积的对称性，这等价于 $\langle\boldsymbol{u},g(\boldsymbol{v})\rangle_{\mathcal{U}}=\langle f(\boldsymbol{u}),\boldsymbol{v}\rangle_{\mathcal{V}}$, 这等价于 $g=f^*$.

2. 设 $f\in\mathrm{Hom}(\mathcal{U},\mathcal{V}),g\in\mathrm{Hom}(\mathcal{V},\mathcal{W})$ $(g\circ f)^*=f^*\circ g^*$ 等价于 $\forall\boldsymbol{u}\in\mathcal{U},\boldsymbol{w}\in\mathcal{W}$, $\langle(g\circ f)(\boldsymbol{u}),\boldsymbol{w}\rangle_{\mathcal{W}}=\langle\boldsymbol{u},(f^*\circ g^*)(\boldsymbol{w})\rangle_{\mathcal{U}}$. 事实上我们有

$$\langle\boldsymbol{u},(f^*\circ g^*)(\boldsymbol{w})\rangle_{\mathcal{U}}=\langle f(\boldsymbol{u}),g^*(\boldsymbol{w})\rangle_{\mathcal{V}}=\langle g(f(\boldsymbol{u})),\boldsymbol{w}\rangle_{\mathcal{U}}.$$

□

练习 8.2.3. 给定对称矩阵 $A=\begin{bmatrix}1&1&1\\1&2&3\\1&3&4\end{bmatrix}$，它在标准欧氏空间 $\mathbb{R}^n$ 上定义的线性变换 $\mathbf{A}$ 是一个对称变换. 给 $\mathbb{R}^n$ 找一组基，使得 $\mathbf{A}$ 在新的基下的矩阵不再是对称矩阵. 这是否意味着 $\mathbf{A}$ 不再是对称变换了？

证明: 例如，我们取 $\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}$, 那么 $\mathbf{A}$ 在这一组基下的矩阵是

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

这个矩阵不再是对称矩阵了，但是 $\mathbf{A}$ 依然是对称变换，因为变换本身不依赖于基。 □

练习 8.2.4. 给定多项式空间 $\mathbb{R}[x]$ 及其内积 $\langle p,q\rangle=\int_0^1 p(x)q(x)\mathrm{d}x$，考虑其上线性变换 $\mathbf{A}$，满足 $(\mathbf{A}p)(x)=xp(x)$. 证明 $\mathbf{A}$ 是对称变换，但没有特征值和特征向量.

证明: 要证明 $\mathbf{A}$ 是对称变换，等价于证明对任意 $p,q\in\mathbb{R}$, $\langle\mathbf{A}(p),q\rangle=\langle p,\mathbf{A}(q)\rangle$.

$$\langle\mathbf{A}(p),q\rangle=\int_0^1\mathbf{A}(p)(x)q(x)\mathrm{d}x=\int_0^1 xp(x)q(x)\mathrm{d}x=\langle p,\mathbf{A}(q)\rangle.$$

若 (λ,p) 是 $\mathbf{A}$ 的特征对，即 $0\neq p(x)\in\mathbb{R}[x]$ 满足 $\mathbf{A}(p)=\lambda p$, 那么 $xp(x)=\lambda p(x)$. 但这是不可能的：因为 $xp(x)$ 的次数比 $p(x)$ 大 1. 所以 $\mathbf{A}$ 没有特征值和特征向量. □

练习 8.2.5. 考虑例 *8.2.9* 中的子空间 $\mathcal{V}_{a,b}$ 上和其上对称变换 $\mathbf{D}^2$. 这个对称变换对应的 *Rayleigh* 商 $\dfrac{\langle f,\mathbf{D}^2f\rangle}{\langle f,f\rangle}$ 对非零的 f，取值最大是多少？此时 f 是什么函数？

证明: 根据例 8.2.12 中的介绍，通过平移和伸缩，把区间 $[-\pi,\pi]$ 变换到 $[a,b]$ 可知 $\mathcal{V}_{a,b}$ 有一组标准正交基：$\left\{\sqrt{\frac{2}{b-a}}\sin\left(n\left(\frac{2\pi}{b-a}x-\frac{b+a}{b-a}\pi\right)\right):n\geq 1\right\}$. 令 $f_n=\sqrt{\frac{2}{b-a}}\sin\left(n\left(\frac{2\pi}{b-a}x-\frac{b+a}{b-a}\pi\right)\right)$, 若 $f=\sum_{n\geq 1}a_nf_n$, 则 Rayleigh 商

$$\frac{\langle f,\mathbf{D}^2f\rangle}{\langle f,f\rangle}=\frac{-\sum_{n\geq 1}n^2a_n^2}{\sum_{n\geq 1}a_n^2}=-1+\frac{\sum_{n\geq 1}(1-n^2)a_n^2}{\sum_{n\geq 1}a_n^2}\leq -1$$

取等号当且仅当对任意 $n \geq 2$ 时 $a_n = 0$, 即 f 是 $f_1 = \sqrt{\frac{2}{b-a}} \sin\left(\frac{2\pi}{b-a}x - \frac{b+a}{b-a}\pi\right)$ 的非零倍。 □

练习 8.2.6. 证明命题 *8.2.14.* 给定 n 维欧式空间 $\mathcal{V}$ 上的线性变换 f, 以下叙述等价:

1. f 是正交变换，即对任意 $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$, 都有 $\langle f(\boldsymbol{x}), f(\boldsymbol{y})\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$;
2. f 是保距变换，即对任意 $\boldsymbol{x} \in \mathcal{V}$, 都有 $\|f(\boldsymbol{x})\| = \|\boldsymbol{x}\|$;
3. f 把 $\mathcal{V}$ 的标准正交基映射成标准正交基。

证明:

(1)⇒(3). 一组向量是标准正交基，当且仅当每个向量是单位向量且不同向量之间内积为 0。设 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 是标准正交基，那么

$$\langle f(\boldsymbol{v}_i), f(\boldsymbol{v}_j)\rangle = \langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

因此 $f(\boldsymbol{v}_1), \cdots, f(\boldsymbol{v}_n)$ 也是标准正交基。

(3)⇒(2). 设 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 是标准正交基且 $\boldsymbol{x} = a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n$, 则

$$\|\boldsymbol{x}\|^2 = \langle a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n, a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n\rangle = \sum_{i,j} a_i a_j \langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle = \sum_{i=1}^{n} a_i^2$$

由于 $f(\boldsymbol{v}_1), \cdots, f(\boldsymbol{v}_n)$ 也是标准正交基, $f(\boldsymbol{x}) = a_1 f(\boldsymbol{v}_1) + \cdots + a_n f(\boldsymbol{v}_n)$, 同样的计算可知 $\|f(\boldsymbol{x})\|^2 = \sum_{i=1}^n a_i^2$, 因此 $\|f(\boldsymbol{x})\| = \|\boldsymbol{x}\|$.

(2)⇒(1). 对任意 $\boldsymbol{x}, \boldsymbol{y}$, $\|f(\boldsymbol{x}+\boldsymbol{y})\| = \|\boldsymbol{x}+\boldsymbol{y}\|$ 平方展开就有 $\|f(\boldsymbol{x})\|^2 + 2\langle f(\boldsymbol{x}), f(\boldsymbol{y})\rangle + \|f(\boldsymbol{y})\|^2 = \|\boldsymbol{x}\|^2 + 2\langle \boldsymbol{x}, \boldsymbol{y}\rangle + \|\boldsymbol{y}\|^2$, 因此 $\langle f(\boldsymbol{x}), f(\boldsymbol{y})\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$.

□

练习 8.2.7. 证明命题 *8.2.15.* 给定 n 维欧式空间 $\mathcal{V}$ 上的线性变换 f, 以下叙述等价:

1. f 是正交变换，当且仅当 $ff^* = f^*f = \mathrm{id}_{\mathcal{V}}$;
2. f 是正交变换，当且仅当 f 在任意标准正交基下的矩阵都是正交矩阵。

证明: f 是正交变换, 则对任意 $\boldsymbol{x}, \boldsymbol{y}$, 都有 $\langle f(\boldsymbol{x}), f(\boldsymbol{y})\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$. 而 $\langle f(\boldsymbol{x}), f(\boldsymbol{y})\rangle = \langle \boldsymbol{x}, f^*f(\boldsymbol{y})\rangle$, 所以 $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, f^*f(\boldsymbol{y})\rangle$. 取 $\boldsymbol{x} = y - f^*f(\boldsymbol{y})$, 则有 $\|f^*f(\boldsymbol{y}) - \boldsymbol{y}\| = 0$, 因此 $f^*f(\boldsymbol{y}) = \boldsymbol{y}$. 由 $\boldsymbol{y}$ 的任意性知 $f^*f = \mathrm{id}_{\mathcal{V}}$. 于是 f 是 $\mathcal{V}$ 上的单射，那么有限维推出 f 是同构，因此 $f^* = f^{-1}$, 进而有 $ff^* = \mathrm{id}_{\mathcal{V}}$.

反之，若 $ff^* = f^*f = \mathrm{id}_{\mathcal{V}}$，则对任意 $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}, \langle \boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, f^*f(\boldsymbol{y})\rangle = \langle f(\boldsymbol{x}), f(\boldsymbol{y})\rangle$。所以 f 是正交变换。

任意给定一组标准正交基，设 f 在这组基下的矩阵是 A, 由命题 8.2.3，f^* 在这组基下的矩阵是 A^T; 在有命题 7.5.4,ff^*, f^*f 在这组基下的矩阵分别是 $AA^{\mathrm{T}}, A^{\mathrm{T}}A$. 由第 1 小题即得:

$$f \text{ 是正交变换，当且仅当 } AA^{\mathrm{T}} = A^{\mathrm{T}}A = I_n, \text{ 即 } A \text{ 是正交矩阵。}$$

□

练习 8.2.8. 回忆阅读 *3.1.23* 和阅读 *3.2.3* 中的小波基和 *Hadamard* 矩阵 $A=\frac{1}{2}\begin{bmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\\1&-1&-1&1\end{bmatrix}$，它是正交矩阵．求其正交相似标准形．

证明: $\det(\lambda I_4-A)=(\lambda-1)^2(\lambda+1)^2$, 特征值为 $1,1,-1,-1$. 由于它是实对称矩阵，一定可以正交相似对角化，因此它的正交相似标准型是 $\operatorname{diag}(1,1,-1,-1)$ □

练习 8.2.9. 求正交矩阵 $\frac{1}{3}\begin{bmatrix}2&-1&2\\2&2&-1\\-1&2&2\end{bmatrix}$ 的正交相似标准形．

证明: $\det(\lambda I_3-A)=(\lambda-1)(\lambda^2-\lambda+1)$, 所以 $1,\cos\frac{\pi}{3}\pm\boldsymbol{i}\sin\frac{\pi}{3}$ 是 A 的特征值，所以 A 的正交相似标准型是

$$\begin{bmatrix}1&0&0\\0&\cos\frac{\pi}{3}&-\sin\frac{\pi}{3}\\0&\sin\frac{\pi}{3}&\cos\frac{\pi}{3}\end{bmatrix}$$

□

练习 8.2.10. 设 $\boldsymbol{q}$ 是 n 维欧氏空间 $\mathcal{V}$ 中的单位向量，$\mathbf{P}$ 是 $\mathcal{M}$ 向 $\operatorname{Span}(\boldsymbol{q})$ 上的正交投影．令 $\mathbf{Q}=\mathbf{I}-2\mathbf{P}$ 是 $\mathcal{V}$ 上的线性变换．求证：

1. $\mathbf{P}(\boldsymbol{a})=\langle\boldsymbol{a},\boldsymbol{q}\rangle\boldsymbol{q}$.
2. $\mathbf{Q}$ 是正交变换，且其正交相似标准形是 $\operatorname{diag}(-1,1,1,\ldots,1)$。称 $\mathbf{Q}$ 为关于超平面 $\operatorname{Span}(\boldsymbol{q})^{\perp}$ 的反射.

证明: 1. 我们只需要验证 $\boldsymbol{a}-\mathbf{P}(\boldsymbol{a})\in\boldsymbol{q}^{\perp}$, 即 $\langle\boldsymbol{a}-\mathbf{P}(\boldsymbol{a}),\boldsymbol{q}\rangle=0$. 这是明显的。

2. 对任意 $\boldsymbol{a},\boldsymbol{b}$,

$$\begin{aligned}\langle\mathbf{Q}(\boldsymbol{a}),\mathbf{Q}(\boldsymbol{b})\rangle&=\langle\boldsymbol{a}-2\langle\boldsymbol{a},\boldsymbol{q}\rangle\boldsymbol{q},\boldsymbol{b}-2\langle\boldsymbol{b},\boldsymbol{q}\rangle\boldsymbol{q}\rangle\\&=\langle\boldsymbol{a},\boldsymbol{b}\rangle-2\langle\boldsymbol{a},\langle\boldsymbol{b},\boldsymbol{q}\rangle\boldsymbol{q}\rangle-2\langle\langle\boldsymbol{a},\boldsymbol{q}\rangle\boldsymbol{q},\boldsymbol{b}\rangle+2\langle\langle\boldsymbol{a},\boldsymbol{q}\rangle\boldsymbol{q},\langle\boldsymbol{b},\boldsymbol{q}\rangle\boldsymbol{q}\rangle\\&=\langle\boldsymbol{a},\boldsymbol{b}\rangle\end{aligned}$$

因此 $\mathbf{Q}$ 是正交矩阵。我们取 $\boldsymbol{q}^{\perp}$ 的一组标准正交基 $\boldsymbol{q}_2,\cdots,\boldsymbol{q}_n$，那么 $\boldsymbol{q},\boldsymbol{q}_2,\cdots,\boldsymbol{q}_n$ 是 $\mathcal{V}$ 的一组标准正交基，在这组基下，$\mathbf{Q}$ 的矩阵是是 $\operatorname{diag}(-1,1,1,\ldots,1)$，那么它就是 $\mathbf{Q}$ 的正交标准型。

□

练习 8.2.11. 设 $\mathbf{Q}$ 是 n 维欧氏空间 $\mathcal{V}$ 上的正交变换，$\lambda_0=1$ 是一个特征值，对应的特征子空间 $\mathcal{N}(\mathbf{I}-\mathbf{Q})$ 的维数是 $n-1$. 证明，$\mathbf{Q}$ 是反射.

证明: 我们取 $\mathcal{N}(\mathbf{I}-\mathbf{Q})$ 的一组标准正交基 $\boldsymbol{q}_1,\cdots,\boldsymbol{q}_{n-1}$，把它扩张成 $\mathcal{V}$ 的一组标准正交基 $\boldsymbol{q}_1,\cdots,\boldsymbol{q}_{n-1},\boldsymbol{q}_n$。那么在这组基之下，$\mathbf{Q}$ 的矩阵为 $A=\begin{bmatrix}I_{n-1}&\boldsymbol{v}\\\boldsymbol{u}^{\mathrm{T}}&a\end{bmatrix}$. 那么 A 是正交矩阵，那么 $\boldsymbol{u}=\boldsymbol{v}=0$ 且 $a^2=1$. 若 $a=1$，那么 $\mathbf{Q}$ 的特征值 1 对应的特征子空间维数是 n，矛盾，因此 $a=-1$. 那么显然 $\mathbf{Q}$ 关于超平面 $\operatorname{Span}(\boldsymbol{q}_n)^{\perp}$ 的反射 □

练习 8.2.12. 对欧氏空间 $\mathcal{V}$ 的子空间 $\mathcal{W}$，定义包含映射 $\mathbf{J}\colon \mathcal{W}\to\mathcal{V}, \mathbf{J}\boldsymbol{v}=\boldsymbol{v}$. 证明 $\mathbf{J}\mathbf{J}^*$ 是到 $\mathcal{W}$ 的正交投影.

证明: 我们只需证明：对任意 $\boldsymbol{v}\in\mathcal{V}$, $\mathbf{J}\mathbf{J}^*\boldsymbol{v}\in\mathcal{W}$ 且 $\boldsymbol{v}-\mathbf{J}\mathbf{J}^*\boldsymbol{v}\in\mathcal{W}^{\perp}$. $\mathbf{J}^*$ 是 $\mathcal{V}$ 到 $\mathcal{W}$ 的映射，所以 $\mathbf{J}^*\boldsymbol{v}\in\mathcal{W}$, 因此 $\mathbf{J}\mathbf{J}^*\boldsymbol{v}\in\mathcal{W}$.

要证明：$\boldsymbol{v}-\mathbf{J}\mathbf{J}^*\boldsymbol{v}\in\mathcal{W}^{\perp}$ 即证明对任意的 $\boldsymbol{w}\in\mathcal{W}, \langle\boldsymbol{v}-\mathbf{J}\mathbf{J}^*\boldsymbol{v},\boldsymbol{w}\rangle=0$, 或者等价地 $\langle\boldsymbol{v},\boldsymbol{w}\rangle=\langle\mathbf{J}\mathbf{J}^*\boldsymbol{v},\boldsymbol{w}\rangle=\langle\boldsymbol{v},\mathbf{J}\mathbf{J}^*\boldsymbol{w}\rangle$. 那么我们只需要证明：对任意 $\boldsymbol{w}\in\mathcal{W}$, 都有 $\mathbf{J}^*\boldsymbol{w}=\boldsymbol{w}$. 对任意 $\boldsymbol{w}$' $\in\mathcal{W}$, 我们有

$$\langle\mathbf{J}^*\boldsymbol{w},\boldsymbol{w}'\rangle=\langle\boldsymbol{w},\mathbf{J}\boldsymbol{w}'\rangle=\langle\boldsymbol{w},\boldsymbol{w}'\rangle.$$

那么我们取 $\boldsymbol{w}'=\mathbf{J}^*\boldsymbol{w}-\boldsymbol{w}$, 就有 $\|\mathbf{J}^*\boldsymbol{w}-\boldsymbol{w}\|=0$. 这就证明了对任意 $\boldsymbol{w}\in\mathcal{W}$, $\mathbf{J}^*\boldsymbol{w}=\boldsymbol{w}$ 成立。 □

练习 8.2.13. 欧氏空间 $\mathcal{V}$ 上的线性变换 f，如果满足 $ff^*=f^*f$，则称 f 是一个**正规变换**。验证对称变换和正交变换都是正规变换.

实矩阵 A，如果满足 $AA^{\mathrm{T}}=A^{\mathrm{T}}A$，则称 A 是一个**正规矩阵**。验证对称、反对称、正交矩阵都是正规矩阵.

证明: 若 f 是对称变换，则 $f=f^*$, 显然 $ff^*=f^*f$; 若 f 是正交变换，则由命题 8.1.5，有 $ff^*=f^*f=\mathrm{id}_{\mathcal{V}}$.

若 A 是对称矩阵，则 $A=A^{\mathrm{T}}$, 所以 $AA^{\mathrm{T}}=A^{\mathrm{T}}A=A^2$; 若 A 是反对称矩阵，则 $A^{\mathrm{T}}=-A$, 所以 $AA^{\mathrm{T}}=A^{\mathrm{T}}A=-A^2$; 若 A 是正交矩阵，则 $AA^{\mathrm{T}}=A^{\mathrm{T}}A=I_n$; 所以它们都是正规矩阵。 □

8.3 酉 空 间

练习 8.3.1. 在 $\mathbb{C}[x]_n$ 中定义二元函数 $\langle f,g\rangle=\sum\limits_{k=1}^{n}f(k)\overline{g(k)}$.

1. 证明，它定义了 $\mathbb{C}[x]_n$ 上的一个内积.
2. 当 $n=3$ 时，求它的一组标准正交基.

证明: 1.
- 共轭对称：$\langle f,g\rangle=\sum_{k=1}^{n}f(k)\overline{g(k)}=\overline{\sum_{k=1}^{n}g(k)\overline{f(k)}}=\overline{\langle g,f\rangle}$.
- 线性和共轭线性：对任意 $f_1,f_2,g_1,g_2\in\mathbb{C}[x]_n, k_1,k_2\in\mathbb{C}$,

$$\begin{aligned}\langle k_1f_1+k_2f_2,g\rangle &=\textstyle\sum_{k=1}^{n}(k_1f_1+k_2f_2)(k)\overline{g_1(k)} &&=k_1\langle f_1,g_1\rangle+k_2\langle f_2,g_1\rangle\\ \langle f_1,k_1g_1+k_2g_2\rangle &=\textstyle\sum_{k=1}^{n}f_1(k)\overline{(k_1g_1+k_2g_2)(k)} &&=\overline{k}_1\langle f_1,g_1\rangle+\overline{k}_2\langle f_2,g_1\rangle\end{aligned}$$

- 正定：对任意 $f\in\mathbb{C}[x]_n$, $\langle f,f\rangle=\sum\limits_{k=1}^{n}f(k)\overline{f(k)}\geq 0$. 且 $\langle f,f\rangle=0$ 当且仅当 $f(1)=\cdots=f(n)=0$. 根据代数基本定理，一个 m 次非零多项式恰好（计算重数）有 m 个根；然而 f 次数小于 n 且有 n 个不同的根，因此 $f=0$

2. 我们先找 $\mathbb{C}[x]_n$ 的一组基 $f_1=1, f_2=x, f_3=x^2$, 再做 Gram-Schmidt 正交化：

$$\begin{aligned} g_1 &= f_1 \\ g_2 &= f_2 - \frac{\langle f_2, g_1\rangle}{\langle g_1, g_1\rangle} g_1 && = x-2 \\ g_3 &= f_3 - \frac{\langle f_3, g_1\rangle}{\langle g_1, g_1\rangle} g_1 - \frac{\langle f_3, g_2\rangle}{\langle g_2, g_2\rangle} g_1 && = x^2-4x+\tfrac{10}{3} \end{aligned}$$

单位化即得 $\mathbb{C}[x]_3$ 的一组标准正交基

$$\frac{g_1}{\sqrt{3}},\ \frac{g_2}{\sqrt{2}},\ \frac{3g_3}{\sqrt{6}}.$$

□

练习 8.3.2. 在具有标准内积的酉空间 $\mathbb{C}^3$ 中，设 $\boldsymbol{a}_1=\begin{bmatrix}1\\-1\\1\end{bmatrix}, \boldsymbol{a}_2=\begin{bmatrix}1\\0\\\boldsymbol{i}\end{bmatrix}$，求与之线性等价的一个正交向量组.

证明: 令 $\boldsymbol{b}_2=\boldsymbol{a}_2-\dfrac{\langle \boldsymbol{a}_2,\boldsymbol{a}_1\rangle}{\langle \boldsymbol{a}_1,\boldsymbol{a}_1\rangle}\boldsymbol{a}_1=\begin{bmatrix}\frac{2-i}{3}\\\frac{1+i}{3}\\\frac{-1+2i}{3}\end{bmatrix}$. 那么 $\boldsymbol{a}_1,\boldsymbol{b}_2$ 正交，且和 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 线性等价。 □

练习 8.3.3. 证明命题 8.3.9.

证明: 设 $A=\begin{bmatrix}\boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n\end{bmatrix}=\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix}U$, 根据命题 8.3.10, $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n$ 是标准正交基，当且仅当 $A^HA=I_n$, 当且仅当 $U^H\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix}^H\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix}U=I_n$. 因为 $\boldsymbol{e}_1\cdots,\boldsymbol{e}_n$ 是标准正交基，所以 $\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix}^H\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{bmatrix}=I_n$. 所以 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n$ 是标准正交基，当且仅当 $U^HU=I_n$ □

练习 8.3.4. 证明命题 8.3.10.

证明: 设 $U=\begin{bmatrix}\boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n\end{bmatrix}$, 那么 $U^HU=\begin{bmatrix}\overline{\boldsymbol{u}}_i^T\boldsymbol{u}_j\end{bmatrix}$. U 是酉矩阵当且仅当 $U^HU=I_n$, 当且仅当 $\overline{\boldsymbol{u}}_i^T\boldsymbol{u}_j=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}$, 当且仅当 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n$ 是标准正交基. □

练习 8.3.5. 利用 Schur 分解证明命题 8.3.13.

证明:

1. 设 A 是 Hermite 矩阵，我们对它使用 Schur 分解：$U^HAU=T$, 其中 U 是酉矩阵，T 是上三角矩阵，那么 $T^H=U^HA^HU=T$, 所以 T 是对角矩阵，那么 $T^H=\overline{T}$, 所以 T 的对角线都是实数。
2. 设 A 是酉矩阵，我们对它使用 Schur 分解：$U^HAU=T$, 其中 U 是酉矩阵，T 是上三角矩阵，T 也是酉矩阵。T 是酉矩阵则 T 的列向量（行向量）是 $\mathbb{C}^n$ 的一组标准正交基，我们考虑第一列，就推出 T 的 $(1,1)$ 元素是绝对值为 1 的复数，再考虑第一行，那么第一行除了 $(1,1)$ 元都是 0, 再考虑第二列，依次同理可证 T 是对角元都是绝对值为 1 的复数的对角矩阵。

□

练习 8.3.6. 证明，酉矩阵的行列式的绝对值为 1；酉矩阵的特征值的绝对值为 1.

证明: 因为 $\det(A^H) = \overline{\det(A)}$. 若 A 是酉矩阵，则 $\det(A)\overline{\det(A)} = \det(AA^H) = 1$, 所以 $|\det(A)| = 1$.

若 $(\lambda, \boldsymbol{v})$ 是酉矩阵 A 的特征对：$A\boldsymbol{v} = \lambda\boldsymbol{v}$, 那么 $(\lambda\boldsymbol{v})^H(\lambda\boldsymbol{v}) = \boldsymbol{v}^H A^H A\boldsymbol{v} = \boldsymbol{v}^H\boldsymbol{v}$, 所以 $\overline{\lambda}\lambda = 1$, 即 $|\lambda| = 1$. □

练习 8.3.7. 设 f 是酉空间 $\mathcal{V}$ 上的 *Hermite* 变换，证明，对任意 $\boldsymbol{a} \in \mathcal{V}$，$\langle f(\boldsymbol{a}), \boldsymbol{a}\rangle$ 都是实数.

证明: $\overline{\langle f(\boldsymbol{a}), \boldsymbol{a}\rangle} = \langle \boldsymbol{a}, f(\boldsymbol{a})\rangle = \langle f(\boldsymbol{a}), \boldsymbol{a}\rangle$, 所以 $\langle f(\boldsymbol{a}), \boldsymbol{a}\rangle$ 是实数。 □

练习 8.3.8. 如果矩阵的共轭转置等于其自身的负矩阵，那么该矩阵称为**反 *Hermite* 矩阵**或**斜 *Hermite* 矩阵**。证明，反 *Hermite* 矩阵是正规矩阵，且其非零特征值都是纯虚数.

证明: 设 A 是反 Hermite 矩阵，即 $A^H = -A$, 则 $AA^H = A^H A = -A^2$，所以反 Hermite 矩阵是正规矩阵。

设 $(\lambda, \boldsymbol{v})$ 是反 Hermite 矩阵矩阵 A 的特征对：$A\boldsymbol{v} = \lambda\boldsymbol{v}$, 且 $\lambda \neq 0$. 那么

$$\overline{\lambda}\boldsymbol{v}^H\boldsymbol{v} = (\lambda\boldsymbol{v})^H\boldsymbol{v} = \boldsymbol{v}^H A^H\boldsymbol{v} = -\lambda\boldsymbol{v}^H\boldsymbol{v},$$

所以 $\overline{\lambda} = -\lambda$, 因此 λ 是纯虚数 □

练习 8.3.9. 如果一个 *Hermite* 矩阵的特征值都是正实数，则称其为**正定** *Hermite* 矩阵. 证明，二元函数 $\overline{\boldsymbol{b}}^{\mathrm{T}}A\boldsymbol{a}$ 定义了 $\mathbb{C}^n$ 上的一个内积，当且仅当 A 是正定 *Hermite* 矩阵.

证明:

- 若二元函数 $\overline{\boldsymbol{b}}^{\mathrm{T}}A\boldsymbol{a}$ 定义了 $\mathbb{C}^n$ 上的一个内积，有共轭对称性知：那么 $\overline{\overline{\boldsymbol{b}}^{\mathrm{T}}A\boldsymbol{a}} = \overline{\boldsymbol{a}}^{\mathrm{T}}A\boldsymbol{b}$, 即 $\boldsymbol{b}^{\mathrm{T}}\overline{A\boldsymbol{a}} = \overline{\boldsymbol{a}}^{\mathrm{T}}A\boldsymbol{b}$. 取 $\boldsymbol{a} = \boldsymbol{e}_i, \boldsymbol{b} = \boldsymbol{e}_j$，则有 $\bar{a_{ji}} = a_{ij}$, 因此 A 是 Hermitian 矩阵。设 $(\lambda, \boldsymbol{v})$ 是 A 的一个特征对，那么由内积的正定性 $\overline{\boldsymbol{v}}^{\mathrm{T}}A\boldsymbol{v} > 0$, 即 $\lambda\overline{\boldsymbol{v}}^{\mathrm{T}}\boldsymbol{v} > 0$. 所以 λ 是正实数。
- 若 A 是正定 Hermite 矩阵，我们需要验证二元函数 $\overline{\boldsymbol{b}}^{\mathrm{T}}A\boldsymbol{a}$ 满足共轭对称性、线性和共轭线性以及正定性。线性和共轭线性是明显的。因为 A 是 Heimite 矩阵，

  $$\overline{\overline{\boldsymbol{a}}^{\mathrm{T}}A\boldsymbol{b}} = \boldsymbol{a}^{\mathrm{T}}\overline{A\boldsymbol{b}} = \left(\boldsymbol{a}^{\mathrm{T}}\overline{A\boldsymbol{b}}\right)^{\mathrm{T}} = \overline{\boldsymbol{b}}^{\mathrm{T}}A\boldsymbol{a},$$

  所以 $\overline{\boldsymbol{b}}^{\mathrm{T}}A\boldsymbol{a}$ 满足共轭对称性。Hermite 矩阵都有意酉相似对角化，而 A 是正定 Hermite 矩阵，所以存在酉矩阵 U, 使得 $U^H AU = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$, 其中 $\lambda_1, \cdots, \lambda_n$ 都是正实数。那么多对任意 $\boldsymbol{v}$, 记 $U\boldsymbol{v} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$, 则 $\overline{\boldsymbol{v}}^{\mathrm{T}}A\boldsymbol{v} = \lambda_1 |z_1|^2 + \cdots + \lambda_m |z_n|^2 \geq 0$. 取等号时 $|z_1| = \cdots = |z_n| = 0$, 即 $U\boldsymbol{v} = 0$, 所以 $\boldsymbol{v} = 0$.

□

练习 8.3.10. 利用酉矩阵的酉相似标准形证明定理 *8.2.18*

证明*:* 设 A 是正交矩阵，U 是酉矩阵，Λ 是对角矩阵使得 $AU=U\Lambda$. 我们可以重排 Λ 的对角元，使得 $\Lambda=\operatorname{diag}(I_a,-I_b,\lambda_1,\bar{\lambda}_1,\cdots,\lambda_r,\bar{\lambda}_r)$. 这是因为 A 是实矩阵，它的特征多项式是实系数的，因此非实根都是成对出现的。我们还需要注意到，$|\lambda_1|=\cdots=|\lambda_r|=1$，所以我们可以假设 $\lambda_j=\cos\theta_j+\boldsymbol{i}\sin\theta_j$.

首先,我们可以取 U 的前 $a+b$ 列都是实向量。再注意到若 $A\boldsymbol{u}=\lambda_j\boldsymbol{u}$, 那么 $A\bar{\boldsymbol{u}}=\bar{\lambda}_j\bar{\boldsymbol{u}}$, 所以 U 的后 $2r$ 列也可以按照这种方式去选取。也就是说,$U=\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_{a+b} & \boldsymbol{u}_1 & \bar{\boldsymbol{u}}_1 & \cdots & \boldsymbol{u}_r & \bar{\boldsymbol{u}}_r\end{bmatrix}$, 其中 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{a+b}$ 是正交的实向量。对 $1\le j\le r$, 令 $\boldsymbol{v}_j=\frac{\boldsymbol{u}_j+\bar{\boldsymbol{u}}_j}{\sqrt{2}},\boldsymbol{v}_j'=\frac{\boldsymbol{i}(\boldsymbol{u}_j-\bar{\boldsymbol{u}}_j)}{\sqrt{2}}$, 把它们也排成一个矩阵：

$$V=\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_{a+b} & \boldsymbol{v}_1 & \boldsymbol{v}_1' & \cdots & \boldsymbol{v}_r & \boldsymbol{v}_r'\end{bmatrix}=U\operatorname{diag}\left(I_{a+b},\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{\boldsymbol{i}}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{\boldsymbol{i}}{\sqrt{2}}\end{bmatrix},\cdots,\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{\boldsymbol{i}}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{\boldsymbol{i}}{\sqrt{2}}\end{bmatrix}\right)$$

那么 V 也是酉矩阵，因为它是实矩阵，所以 V 是正交矩阵。下一步，我们就是要计算 $V^{\mathrm{T}}AV$. 事实上

$$\begin{aligned}
A\begin{bmatrix}\boldsymbol{v}_j & \boldsymbol{v}_j'\end{bmatrix} &= A\begin{bmatrix}\boldsymbol{u}_j & \bar{\boldsymbol{u}}_j\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{\boldsymbol{i}}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{\boldsymbol{i}}{\sqrt{2}}\end{bmatrix}\\
&= \begin{bmatrix}\boldsymbol{u}_j & \bar{\boldsymbol{u}}_j\end{bmatrix}\begin{bmatrix}\lambda_j & \\ & \bar{\lambda}_j\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{\boldsymbol{i}}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{\boldsymbol{i}}{\sqrt{2}}\end{bmatrix}\\
&= \begin{bmatrix}\boldsymbol{v}_j & \boldsymbol{v}_j'\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{\boldsymbol{i}}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{\boldsymbol{i}}{\sqrt{2}}\end{bmatrix}^H\begin{bmatrix}\lambda_j & \\ & \bar{\lambda}_j\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{\boldsymbol{i}}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{\boldsymbol{i}}{\sqrt{2}}\end{bmatrix}\\
&= \begin{bmatrix}\boldsymbol{v}_j & \boldsymbol{v}_j'\end{bmatrix}R_{\theta_j}
\end{aligned}$$

所以 $V^{\mathrm{T}}AV=\operatorname{diag}(I_a,-I_b,R_{\theta_1},\cdots,R_{\theta_r})$

$\square$

练习 8.3.11. *判断以下矩阵是否为酉矩阵、Hermite 矩阵、反 Hermite 矩阵、正规矩阵，再计算谱分解.*

1. $\begin{bmatrix}0 & 1-\boldsymbol{i}\\ i+1 & 1\end{bmatrix}$.
2. $\begin{bmatrix}2 & 1+\boldsymbol{i}\\ \boldsymbol{i}-1 & 3\end{bmatrix}$.
3. $\dfrac{1}{\sqrt{3}}\begin{bmatrix}1 & 1-\boldsymbol{i}\\ 1+\boldsymbol{i} & -1\end{bmatrix}$.
4. $\begin{bmatrix}0 & \boldsymbol{i} & 0\\ 0 & 0 & \boldsymbol{i}\\ \boldsymbol{i} & 0 & 0\end{bmatrix}$.
5. $A=\begin{bmatrix}\boldsymbol{i} & 1 & \boldsymbol{i}\\ 1 & \boldsymbol{i} & 1\end{bmatrix}$, *考虑* $A^{\mathrm{H}}A$ *和* AA^{H}.

证明*:* 谱分解的方法同实矩阵一样，先算特征多项式，解出特征值，然后计算每个特征值的特征子空间的一组基即可。我们略去具体过程。需要说明的是，此处的谱分解并没有追求用酉矩阵作酉相似对角化；比如说对 Hermite 矩阵作酉相似对角化，我们只需要对解答中的矩阵列向量单位化即可。

1. $A=A^H$, 所以 A 是 Hermite 矩阵. $A=\begin{bmatrix}1 & 1-i\\ 1+i & -1\end{bmatrix}\begin{bmatrix}2 & \\ & -1\end{bmatrix}\begin{bmatrix}1 & 1-i\\ 1+i & -1\end{bmatrix}^{-1}$.

2. $A^H=\begin{bmatrix}2 & -\boldsymbol{i}-1\\ 1-\boldsymbol{i} & 3\end{bmatrix}, AA^H=\begin{bmatrix}6 & \boldsymbol{i}+1\\ 1-\boldsymbol{i} & 11\end{bmatrix}, A^HA=\begin{bmatrix}6 & -\boldsymbol{i}-1\\ \boldsymbol{i}-1 & 11\end{bmatrix}$, 所以上述的四种矩阵 A 都不是。A 的谱分解为 $A=\begin{bmatrix}\frac{1-\sqrt{7}\boldsymbol{i}}{2} & \frac{1+\sqrt{7}\boldsymbol{i}}{2}\\ 1-\boldsymbol{i} & 1-\boldsymbol{i}\end{bmatrix}\begin{bmatrix}\frac{5+\sqrt{7}\boldsymbol{i}}{2} & \\ & \frac{5-\sqrt{7}\boldsymbol{i}}{2}\end{bmatrix}\begin{bmatrix}\frac{1-\sqrt{7}\boldsymbol{i}}{2} & \frac{1+\sqrt{7}\boldsymbol{i}}{2}\\ 1-\boldsymbol{i} & 1-\boldsymbol{i}\end{bmatrix}^{-1}$

3. $A=A^H, AA^H=I_2$, 所以 A 既是 Hermite 矩阵，也是酉矩阵。A 的谱分解是 $A=\begin{bmatrix}\frac{1}{\sqrt{3}}+1 & \frac{1}{\sqrt{3}}-1\\ \frac{1+\boldsymbol{i}}{\sqrt{3}} & \frac{1+\boldsymbol{i}}{\sqrt{3}}\end{bmatrix}\begin{bmatrix}1 & \\ & -1\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{3}}+1 & \frac{1}{\sqrt{3}}-1\\ \frac{1+\boldsymbol{i}}{\sqrt{3}} & \frac{1+\boldsymbol{i}}{\sqrt{3}}\end{bmatrix}^{-1}$

4. A 的列向量是两两正交的单位向量,所以 A 是酉矩阵,自然是正规矩阵。$A^H=\begin{bmatrix}0 & 0 & -\boldsymbol{i}\\ -\boldsymbol{i} & 0 & 0\\ 0 & -\boldsymbol{i} & 0\end{bmatrix}$, $A^H\neq\pm A$, 所以 A 不是（反）Hermite 矩阵。A 的谱分解是

$$A=\begin{bmatrix}1 & -e^{-\frac{\pi}{3}\boldsymbol{i}} & -e^{\frac{\pi}{3}\boldsymbol{i}}\\ 1 & -e^{\frac{\pi}{3}\boldsymbol{i}} & -e^{-\frac{\pi}{3}\boldsymbol{i}}\\ 1 & 1 & 1\end{bmatrix}\begin{bmatrix}e^{\frac{\pi}{2}\boldsymbol{i}} & & \\ & e^{\frac{7\pi}{6}\boldsymbol{i}} & \\ & & e^{\frac{11\pi}{6}\boldsymbol{i}}\end{bmatrix}\begin{bmatrix}1 & -e^{-\frac{\pi}{3}\boldsymbol{i}} & -e^{\frac{\pi}{3}\boldsymbol{i}}\\ 1 & -e^{\frac{\pi}{3}\boldsymbol{i}} & -e^{-\frac{\pi}{3}\boldsymbol{i}}\\ 1 & 1 & 1\end{bmatrix}^{-1}$$

5. $A^H=\begin{bmatrix}-\boldsymbol{i} & 1\\ 1 & -\boldsymbol{i}\\ -\boldsymbol{i} & 1\end{bmatrix}, A^HA=\begin{bmatrix}2 & 0 & 2\\ 0 & 2 & 0\\ 2 & 0 & 2\end{bmatrix}, AA^H=\begin{bmatrix}3 & \boldsymbol{i}\\ -\boldsymbol{i} & 3\end{bmatrix}$, 它们都是 Hermite 矩阵，其他种类都不属于。它们的谱分解是

$$A^HA=\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ -1 & 0 & 1\end{bmatrix}\begin{bmatrix}0 & & \\ & 2 & \\ & & 4\end{bmatrix}\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ -1 & 0 & 1\end{bmatrix}^{-1}, AA^H=\begin{bmatrix}-\boldsymbol{i} & \boldsymbol{i}\\ 1 & 1\end{bmatrix}\begin{bmatrix}2 & \\ & 4\end{bmatrix}\begin{bmatrix}-\boldsymbol{i} & \boldsymbol{i}\\ 1 & 1\end{bmatrix}^{-1}$$

□

练习 8.3.12. 证明或举出反例.

1. 既是 Hermite 矩阵又是酉矩阵的矩阵是对角阵.

2. 如果 H 是 Hermite 矩阵，则 $\boldsymbol{i}H$ 是反 Hermite 矩阵.

3. 如果 A 是 n 阶实方阵，则 $A+\boldsymbol{i}I_n$ 可逆.

4. 如果 A 是 n 阶酉矩阵，则 $A+\boldsymbol{i}I_n$ 可逆.

5. Hermite 矩阵的行列式是实数.

6. 反 Hermite 矩阵的行列式是纯虚数.

证明: 1. 反例：$A=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & \boldsymbol{i}\\ -\boldsymbol{i} & 1\end{bmatrix}$

2. 正确：$(\boldsymbol{i}H)^H=\bar{\boldsymbol{i}}H^H=-\boldsymbol{i}H$.

3. 反例：$A=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$

4. 反例：$A=\begin{bmatrix}\boldsymbol{i} & 0\\ 0 & -\boldsymbol{i}\end{bmatrix}$

5. 正确，因为 $\det(A^H)=\overline{\det(A)}$, 所以 $A=A^H$ 推出 $\det(A)=\overline{\det(A)}$, 因此 $\det(A)$ 是实数。

6. 错误，例如 $A=\begin{bmatrix}\boldsymbol{i} & 0\\ 0 & -\boldsymbol{i}\end{bmatrix}$, 它的行列式是 1.

$\square$

练习 8.3.13. 给定 n 阶实方阵 A,B，我们考虑复矩阵 $A+\boldsymbol{i}B$ 和实矩阵 $\begin{bmatrix}A & -B\\ B & A\end{bmatrix}$ 的关系.

1. 证明 $A+\boldsymbol{i}B$ 是 *Hermite* 矩阵当且仅当 $\begin{bmatrix}A & -B\\ B & A\end{bmatrix}$ 是实对称矩阵.

2. 证明 $A+\boldsymbol{i}B$ 是酉矩阵当且仅当 $\begin{bmatrix}A & -B\\ B & A\end{bmatrix}$ 是正交矩阵.

3. 证明对 n 维实向量 $\boldsymbol{v},\boldsymbol{w}$ 和实数 a,b，我们有 $(A+\boldsymbol{i}B)(\boldsymbol{v}+\boldsymbol{i}\boldsymbol{w})=(\boldsymbol{v}+\boldsymbol{i}\boldsymbol{w})(a+\boldsymbol{i}b)$ 当且仅当 $\begin{bmatrix}A & -B\\ B & A\end{bmatrix}\begin{bmatrix}\boldsymbol{v} & -\boldsymbol{w}\\ \boldsymbol{w} & \boldsymbol{v}\end{bmatrix}=\begin{bmatrix}\boldsymbol{v} & -\boldsymbol{w}\\ \boldsymbol{w} & \boldsymbol{v}\end{bmatrix}\begin{bmatrix}a & -b\\ b & a\end{bmatrix}$.

4. 证明 $(A_1+\boldsymbol{i}B_1)(A_2+\boldsymbol{i}B_2)=A_3+\boldsymbol{i}B_3$ 当且仅当 $\begin{bmatrix}A_1 & -B_1\\ B_1 & A_1\end{bmatrix}\begin{bmatrix}A_2 & -B_2\\ B_2 & A_2\end{bmatrix}=\begin{bmatrix}A_3 & -B_3\\ B_3 & A_3\end{bmatrix}$, 其中 $A_i,B_i,\ i=1,2,3$ 是 n 阶实方阵.

5. $\operatorname{trace}(A+\boldsymbol{i}B)$ 和 $\operatorname{trace}\left(\begin{bmatrix}A & -B\\ B & A\end{bmatrix}\right)$ 有何关系? $\det(A+\boldsymbol{i}B)$ 和 $\det\left(\begin{bmatrix}A & -B\\ B & A\end{bmatrix}\right)$ 如何?

6. 求可逆矩阵 C 使得矩阵 $\begin{bmatrix}A+\boldsymbol{i}B & 0\\ 0 & A-\boldsymbol{i}B\end{bmatrix}=C\begin{bmatrix}A & -B\\ B & A\end{bmatrix}C^{-1}$.

附注. 因此两矩阵相似.

证明*:* 1. $A+\boldsymbol{i}B$ 是 Hermite 矩阵, 当且仅当 $A+\boldsymbol{i}B=A^T-\boldsymbol{i}B^T$, 当且仅当 $A^T=A, B^T=-B$, 当且仅当 $\begin{bmatrix}A & -B\\ B & A\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}A & -B\\ B & A\end{bmatrix}$.

2. $A+\boldsymbol{i}B$ 是酉矩阵，当且仅当 $(A+\boldsymbol{i}B)(A^T-\boldsymbol{i}B^T)=I_n$, 当且仅当 $AA^T+BB^T=I_n, BA^T-AB^T=0$, 当且仅当 $\begin{bmatrix}A & -B\\ B & A\end{bmatrix}\begin{bmatrix}A & -B\\ B & A\end{bmatrix}^{\mathrm{T}}=I_{2n}$.

3. $(A+\boldsymbol{i}B)(\boldsymbol{v}+\boldsymbol{i}\boldsymbol{w})=(\boldsymbol{v}+\boldsymbol{i}\boldsymbol{w})(a+\boldsymbol{i}b)$ 当且仅当 $(A\boldsymbol{v}-B\boldsymbol{w})+\boldsymbol{i}(B\boldsymbol{v}+A\boldsymbol{w})=(a\boldsymbol{v}-b\boldsymbol{w})+\boldsymbol{i}(a\boldsymbol{w}+b\boldsymbol{v})$, 当且仅当 $A\boldsymbol{v}-B\boldsymbol{w}=a\boldsymbol{v}-b\boldsymbol{w}, A\boldsymbol{w}+B\boldsymbol{v}=a\boldsymbol{w}+b\boldsymbol{v}$.
另一方面, $\begin{bmatrix}A & -B\\ B & A\end{bmatrix}\begin{bmatrix}\boldsymbol{v} & -\boldsymbol{w}\\ \boldsymbol{w} & \boldsymbol{v}\end{bmatrix}=\begin{bmatrix}\boldsymbol{v} & -\boldsymbol{w}\\ \boldsymbol{w} & \boldsymbol{v}\end{bmatrix}\begin{bmatrix}a & -b\\ b & a\end{bmatrix}$ 即 $\begin{bmatrix}A\boldsymbol{v}-B\boldsymbol{w} & -A\boldsymbol{w}-B\boldsymbol{v}\\ B\boldsymbol{v}+A\boldsymbol{w} & -B\boldsymbol{w}+A\boldsymbol{v}\end{bmatrix}=\begin{bmatrix}a\boldsymbol{v}-b\boldsymbol{w} & -a\boldsymbol{w}-b\boldsymbol{v}\\ b\boldsymbol{v}+a\boldsymbol{w} & -b\boldsymbol{w}+a\boldsymbol{v}\end{bmatrix}$, 当且仅当 $A\boldsymbol{v}-B\boldsymbol{w}=a\boldsymbol{v}-b\boldsymbol{w}, A\boldsymbol{w}+B\boldsymbol{v}=a\boldsymbol{w}+b\boldsymbol{v}$.

4. $(A_1+\boldsymbol{i}B_1)(A_2+\boldsymbol{i}B_2)=(A_1A_2-B_1B_2)+\boldsymbol{i}(A_1B_2+B_1A_2)$, $(A_1+\boldsymbol{i}B_1)(A_2+\boldsymbol{i}B_2)=A_3+\boldsymbol{i}B_3$ 当且仅当 $A_1A_2-B_1B_2=A_3, A_1B_2+B_1A_2=B_3$.
另一方面，$\begin{bmatrix}A_1 & -B_1\\ B_1 & A_1\end{bmatrix}\begin{bmatrix}A_2 & -B_2\\ B_2 & A_2\end{bmatrix}=\begin{bmatrix}A_1A_2-B_1B_2 & -A_1B_2-B_1A_2\\ B_aA_2+A_1B_2 & -B_1B_2+A_1A_2\end{bmatrix}$，所以

$\begin{bmatrix} A_1 & -B_1 \\ B_1 & A_1 \end{bmatrix}\begin{bmatrix} A_2 & -B_2 \\ B_2 & A_2 \end{bmatrix} = \begin{bmatrix} A_3 & -B_3 \\ B_3 & A_3 \end{bmatrix}$ 当且仅当 $A_1A_2 - B_1B_2 = A_3, A_1B_2 + B_1A_2 = B_3$.

5. $\operatorname{trace}(A+\boldsymbol{i}B) = \operatorname{trace}(A)+\boldsymbol{i}\operatorname{trace}(B)$, 所以 $\operatorname{trace}\left(\begin{bmatrix} A & -B \\ B & A \end{bmatrix}\right) = 2\operatorname{trace}(A) = \operatorname{trace}(A+\boldsymbol{i}B) + \overline{\operatorname{trace}(A+\boldsymbol{i}B)}$. 我们做初等行变换 $\begin{bmatrix} A & -B \\ B & A \end{bmatrix} \to \begin{bmatrix} A+\boldsymbol{i}B & \boldsymbol{i}(A+\boldsymbol{i}B) \\ B & A \end{bmatrix}$, 若 $A+\boldsymbol{i}B$ 不可逆，则 $\det(A+\boldsymbol{i}B) = \det\left(\begin{bmatrix} A & -B \\ B & A \end{bmatrix}\right) = 0$. 若 $A+\boldsymbol{i}B$ 可逆，则

$$\begin{aligned}
\det\left(\begin{bmatrix} A & -B \\ B & A \end{bmatrix}\right) &= \det(A+\boldsymbol{i}B)\det\left(\begin{bmatrix} I_n & \boldsymbol{i}I_n \\ B & A \end{bmatrix}\right) \\
&= \det(A+\boldsymbol{i}B)\det\left(\begin{bmatrix} I_n & \boldsymbol{i}I_n \\ 0 & A-\boldsymbol{i}B \end{bmatrix}\right) \\
&= \det(A+\boldsymbol{i}B)\det(A-\boldsymbol{i}B)
\end{aligned}$$

所以我们始终有 $\det\left(\begin{bmatrix} A & -B \\ B & A \end{bmatrix}\right) = \det(A+\boldsymbol{i}B)\overline{\det(A+\boldsymbol{i}B)}$.

6. 这个问题，我们通过：观察现象——猜测——验证这三步曲来做，或者根据：观察特殊情况——猜测——验证这三步。例如，当 A, B 都是 1 阶方阵，即是一个实数的情形，对 $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ 作相似对角化，可得 $C = \begin{bmatrix} 1 & \boldsymbol{i} \\ 1 & -\boldsymbol{i} \end{bmatrix}$. 对于 A, B 是 n 阶方阵，那么我们可以猜测 $C = \begin{bmatrix} I_n & \boldsymbol{i}I_n \\ I_n & -\boldsymbol{i}I_n \end{bmatrix}$, 那么很容易看出 C 可逆，计算矩阵乘法：$\begin{bmatrix} A+\boldsymbol{i}B & 0 \\ 0 & A-\boldsymbol{i}B \end{bmatrix} C$ 和 $C\begin{bmatrix} A & -B \\ B & A \end{bmatrix}$ 就完成了证明。

□

练习 8.3.14. 考虑酉矩阵 $U = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$，其中 $a, b, c, d \in \mathbb{C}$.

1. 证明存在 $\phi \in \mathbb{R}$ 使得 $c = -e^{\boldsymbol{i}\phi}\overline{b}, d = e^{\boldsymbol{i}\phi}\overline{a}$. 计算 $\det(U)$，将结果用 ϕ 表示.
2. 由 $|a|^2 + |b|^2 = 1$ 可设 $|a| = \cos\theta, |b| = \sin\theta$，其中 $\theta \in \mathbb{R}$. 求 $\phi_1, \phi_2 \in \mathbb{R}$ 满足 $U = e^{\boldsymbol{i}\frac{\phi}{2}}\begin{bmatrix} e^{\boldsymbol{i}\phi_1}\cos\theta & e^{\boldsymbol{i}\phi_2}\sin\theta \\ -e^{-\boldsymbol{i}\phi_2}\sin\theta & e^{-\boldsymbol{i}\phi_1}\cos\theta \end{bmatrix}$.

   附注. 这说明一个二阶酉矩阵被四个实系数 $\phi, \phi_1, \phi_2, \theta$ 决定.

   求 $\psi_1, \psi_2 \in \mathbb{R}$，使得 $U = e^{\boldsymbol{i}\frac{\phi}{2}}D_1R_\theta D_2$，其中 R_θ 是旋转矩阵 $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$，而 $D_j = \begin{bmatrix} e^{\boldsymbol{i}\psi_j} & 0 \\ 0 & e^{-\boldsymbol{i}\psi_j} \end{bmatrix}$.

   附注. 这也说明一个二阶酉矩阵被四个实系数 $\phi, \psi_1, \psi_2, \theta$ 决定.

证明: 1. 由 $U^H=U^{-1}$ 可得 $\begin{bmatrix}\bar a & \bar c\\ \bar b & \bar d\end{bmatrix}=\det(U)^{-1}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}$, 酉矩阵行列式的绝对值是 1，因此存在 $\phi\in\mathbb{R}$ 使得 $\det(U)=e^{i\phi}$, 代入即得 $c=-e^{i\phi}\overline{b}, d=e^{i\phi}\overline{a}$。

2. 我们可以用待定系数法，把 D_1,D_2 代入计算。但是事实上，稍微敏感一些，直接观察即可得

$$D_1=\begin{bmatrix}e^{i\frac{\phi_1-\phi_2}{2}} & 0\\ 0 & e^{-i\frac{\phi_1-\phi_2}{2}}\end{bmatrix}, D_2=\begin{bmatrix}e^{i\frac{\phi_1+\phi_2}{2}} & 0\\ 0 & e^{-i\frac{\phi_1+\phi_2}{2}}\end{bmatrix}$$

即 $\psi_1=\frac{\phi_1-\phi_2}{2},\psi_2=\frac{\phi_1+\phi_2}{2}$.

□

练习 8.3.15. 设 n 阶酉矩阵 A 满足 $A=A^{\mathrm{T}}$（A 未必是 *Hermite* 矩阵）.

1. 证明 $\overline{A}=A^{-1}$ 是 A 的逆矩阵.

2. 证明 A 的每个特征值都有实特征向量.

3. 证明存在对角阵 D 和（实）正交矩阵 P，使得 $A=PDP^{-1}$.

注意 $\dfrac{1}{\sqrt{n}}F_n$ 就满足条件，其中 F_n 是 n 阶 *Fourier* 矩阵.

证明: 1. 因为 $A^{-1}=A^H=\overline{A^T}=\overline{A}$.

2. 设 $\boldsymbol{v}\neq 0$ 使得 $A\boldsymbol{v}=\lambda\boldsymbol{v}$, 因为 A 的酉矩阵，所以 A 的特征值模长是 1, 那么 $\lambda^{-1}=\bar\lambda$. 我们对等式 $A\boldsymbol{v}=\lambda\boldsymbol{v}$ 取复共轭就有 $\bar A\bar{\boldsymbol{v}}=\bar\lambda\bar{\boldsymbol{v}}$, 所以 $A\bar{\boldsymbol{v}}=\lambda\bar{\boldsymbol{v}}$. 由于 $\boldsymbol{v}\neq 0$, 要么 $\boldsymbol{v}_1=\boldsymbol{v}+\bar{\boldsymbol{v}}\neq 0$, 要么 $\boldsymbol{v}_2=\boldsymbol{i}(\boldsymbol{v}-\bar{\boldsymbol{v}})\neq 0$, 这两个向量都是实向量, 且 $A\boldsymbol{v}_i=\lambda\boldsymbol{v}_i(i=1,2)$. 所以 A 的特征值 λ 有实特征向量。

3. 我们对阶数做归纳。$n=1$ 时，命题是平凡的。假设结论对 n 成立，A 是 $n+1$ 阶对称酉矩阵。任取 A 的一个特征值 λ_1 和属于它的实单位特征向量 $\boldsymbol{v}_1$，把 $\boldsymbol{v}_1$ 扩张成 $\mathbb{R}^n$ 的一组标准正交基，那么 $U=\begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2,\cdots,\boldsymbol{v}_{n+1}\end{bmatrix}$ 是实正交矩阵。那么

$$AU=\begin{bmatrix}\lambda_1\boldsymbol{v}_1 & A\boldsymbol{v}_2 & \cdots & A\boldsymbol{v}_{n+1}\end{bmatrix}=U\begin{bmatrix}\lambda_1 & \boldsymbol{u}^{\mathrm{T}}\\ 0 & B\end{bmatrix}$$

那么 $\begin{bmatrix}\lambda_1 & \boldsymbol{u}^{\mathrm{T}}\\ 0 & B\end{bmatrix}=U^{\mathrm{T}}AU$ 是酉矩阵, 也是对称矩阵, 因此 B 也是对称的酉矩阵。由归纳假设, 存在 n-阶实正交矩阵 V 施舍 $V^{\mathrm{T}}BV$ 是对称矩阵, 因此 $\left(U\begin{bmatrix}1 & \\ & V\end{bmatrix}\right)^{\mathrm{T}}A\left(U\begin{bmatrix}1 & \\ & V\end{bmatrix}\right)$ 是对角矩阵，并且 $\left(U\begin{bmatrix}1 & \\ & V\end{bmatrix}\right)$ 是实对称矩阵，这就完成了归纳。

□

练习 8.3.16. 给定 n 阶 *Fourier* 矩阵 F_n，计算 $\det(F_n)$ 和 F_n^2.

证明: F_n 是范德蒙行列式，所以 $\det(F_n)=\prod\limits_{0\le i<j\le n-1}(\omega_n^j-\omega_n^i)$.

直接计算矩阵乘法，我们就有 F_n^2 的第 (i,j) 元是

$$\sum_{k=0}^{n-1}\omega_n^{k(i+j-2)}=\begin{cases} n & i+j=2\text{或者}n+2\\ 0 & \text{其他}\end{cases}$$

所以 $F_n^2=\begin{bmatrix} n & 0 & \cdots & 0\\ 0 & 0 & \cdots & n\\ \vdots & & \iddots & \vdots\\ 0 & n & \cdots & 0\end{bmatrix}$. 从这个等式可以看出 $\det(F_n)^2=(-1)^{\frac{(n-1)(n-2)}{2}}n^n$ □

练习 8.3.17. 证明命题 *8.3.18.*

证明*:* 因为 F_n 是对称矩阵，所以 $F_n^H=\overline{F_n^{\mathrm{T}}}=\overline{F_n}$.

考虑 $F_n\overline{F_n}$, 它的第 (i,j) 元是

$$\sum_{k=1}^{n}\omega_n^{k(i-1)}\overline{\omega}_n^{k(j-1)}=\sum_{k=0}^{n-1}\omega_n^{k(n+i-j)}=\begin{cases} n & i=j\\ 0 & \text{其他}\end{cases}$$

所以 $F_n\overline{F_n}=nI_n$.

□

练习 8.3.18. 证明定理 *8.3.20.*

证明*:* 这里的验证仅仅是按部就班的矩阵运算而已，唯一需要注意的一点是：$\omega_{2n}^2=\omega_n$. □

练习 8.3.19. 对任意复数 $c_1,c_2,\ldots,c_n$，形如 $\begin{bmatrix} c_0 & c_1 & \cdots & c_{n-2} & c_{n-1}\\ c_{n-1} & c_0 & c_1 & & c_{n-2}\\ \vdots & c_{n-1} & c_0 & \ddots & \vdots\\ c_2 & & \ddots & \ddots & c_1\\ c_1 & c_2 & \cdots & c_{n-1} & c_0\end{bmatrix}$ 的方阵称为**循环矩阵**.

1. 证明，对任意 n 阶循环矩阵 C，$F_n^{-1}CF_n$ 是对角矩阵，其中 F_n 是 n 阶 *Fourier* 矩阵.

2. 设计一个计算 $C\boldsymbol{x}$，其中 C 为循环矩阵的快速算法.

3. 设计一个求解 $C\boldsymbol{x}=\boldsymbol{b}$，其中 C 为循环矩阵的快速算法.

证明*:* 1. 令 $\xi_i=\begin{bmatrix} 1\\ \omega_n^{i-1}\\ \vdots\\ \omega_n^{(n-1)(i-1)}\end{bmatrix}$ 是 F_n 的第 i 个列向量，那么

$$C\xi_i=\begin{bmatrix} c_0+c_1\omega_n^{i-1}++\cdots+c_{n-1}\omega_n^{(n-1)(i-1)}\\ c_{n-1}+c_0\omega_n^{i-1}++\cdots+c_{n-2}\omega_n^{(n-1)(i-1)}\\ \vdots\\ c_1+c_2\omega_n^{i-1}++\cdots+c_0\omega_n^{(n-1)(i-1)}\end{bmatrix}$$

注意到 $(\omega_n^{i-1})^n = 1$, 令 $f(x) = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}$, 那么 $C\xi_i = f(\omega_n^{i-1})\xi_i$. 所以

$$CF_n = F_n \begin{bmatrix} 1 & & & \\ & f(\omega_n) & & \\ & & \ddots & \\ & & & f(\omega_n^{n-1}) \end{bmatrix}, \quad \text{即} F_n^{-1}CF_n = \begin{bmatrix} 1 & & & \\ & f(\omega_n) & & \\ & & \ddots & \\ & & & f(\omega_n^{n-1}) \end{bmatrix}$$

2. 令 $\Lambda = \begin{bmatrix} 1 & & & \\ & f(\omega_n) & & \\ & & \ddots & \\ & & & f(\omega_n^{n-1}) \end{bmatrix}$, 则 $C\boldsymbol{x} = F_n \Lambda F_n^{-1}\boldsymbol{x} = n^{-1}F_n\Lambda\overline{F}_n\boldsymbol{x}$, 因此我们利用快速算法计算 $\overline{F}_n\boldsymbol{x}$，然后做一个矩阵乘法：左乘 Λ, 然后再用快速算法计算左乘 F_n 即可。
3. 等价于求解 $\Lambda\boldsymbol{y} = F_n\boldsymbol{b}$ 和 $\boldsymbol{x} = n^{-1}\overline{F}_n\boldsymbol{y}$; 这两者都有快速算法。

□