FDU 高等线性代数 Homework 02

Due: Sept. 14, 2024
姓名: 雍崔扬
学号: 21307140051

Problem 1

设 m,n 为正整数且 $m\geq n$，给定列满秩矩阵 $A\in\mathbb{R}^{m\times n}$ 和向量 $b\in\mathbb{R}^m$.
试证明 $\mathbb{R}^n$ 上的函数 $f(x)=\|b-Ax\|_2^2$ 的全局最小点为 $x_\star=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b$.

Solution:

$$\begin{aligned}\nabla f(x)&=\nabla_x\{\|b-Ax\|_2^2\}\\&=\nabla_x\{b-Ax\}\cdot 2(b-Ax)\\&=-A^{\mathrm{T}}\cdot 2(b-Ax)\\&=-2(A^{\mathrm{T}}b-A^{\mathrm{T}}Ax)\\\nabla^2 f(x)&=-2\nabla_x\{A^{\mathrm{T}}b-A^{\mathrm{T}}Ax\}\\&=-2(-A^{\mathrm{T}}A)^{\mathrm{T}}\\&=2A^{\mathrm{T}}A\end{aligned}$$

注意到 $\operatorname{rank}(A)=n\leq m$ 保证了 $\nabla^2 f(x)=2A^{\mathrm{T}}A\succ 0$ $(\forall\, x\in\mathbb{R}^n)$.
因此 $f(x)$ 在 $\mathbb{R}^n$ 上是严格凸函数，其全局最小点是驻点 (即令梯度为零的点)，且若存在则必定唯一.
令 $\nabla f(x)=-2(A^{\mathrm{T}}b-A^{\mathrm{T}}A)x=0_n$ 即可解得全局最小点为 $x_\star=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b$.
(值得注意的是，A 列满秩的条件保证了 $A^{\mathrm{T}}A$ 的正定性，因此 $A^{\mathrm{T}}A$ 是非奇异的)

Problem 2

设 n 为正整数，$A\in\mathbb{R}^{n\times n}$ 是给定的实对称阵.
试求下列 $\mathbb{R}^n\mapsto\mathbb{R}$ 的函数的梯度:

$$\begin{aligned}f(x)&=\operatorname{tr}\left((A+xx^{\mathrm{T}})^2\right)\\g(x)&=\operatorname{tr}\left((A+xx^{\mathrm{T}})^3\right)\end{aligned}$$

Solution:
首先求解 f 的梯度:

$$\begin{aligned}f(x)&=\operatorname{tr}\left((A+xx^{\mathrm{T}})^2\right)\\&=\operatorname{tr}\left(A^2+xx^{\mathrm{T}}A+Axx^{\mathrm{T}}+x^{\mathrm{T}}xxx^{\mathrm{T}}\right)\\&=\operatorname{tr}\left(A^2\right)+\operatorname{tr}\left(xx^{\mathrm{T}}A\right)+\operatorname{tr}\left(Axx^{\mathrm{T}}\right)+x^{\mathrm{T}}x\operatorname{tr}\left(xx^{\mathrm{T}}\right)\\&=\operatorname{tr}\left(A^2\right)+2\operatorname{tr}\left(x^{\mathrm{T}}Ax\right)+x^{\mathrm{T}}x\operatorname{tr}\left(x^{\mathrm{T}}x\right)\\&=\operatorname{tr}\left(A^2\right)+2x^{\mathrm{T}}Ax+(x^{\mathrm{T}}x)^2\\\hline\nabla f(x)&=2\nabla_x\{x^{\mathrm{T}}Ax\}+\nabla_x\{(x^{\mathrm{T}}x)^2\}\\&=2\cdot(A+A^{\mathrm{T}})x+\nabla_x\{x^{\mathrm{T}}x\}\cdot 2(x^{\mathrm{T}}x)\quad(\text{note that } A^{\mathrm{T}}=A)\\&=4Ax+2x\cdot 2(x^{\mathrm{T}}x)\\&=4(A+xx^{\mathrm{T}})x\end{aligned}$$

接下来求解 g 的梯度:

$$
\begin{aligned}
g(x) &= \operatorname{tr}\left((A+xx^{\mathrm{T}})^3\right) \\
&= \operatorname{tr}\left(A^3 + xx^{\mathrm{T}}A^2 + Axx^{\mathrm{T}}A + x^{\mathrm{T}}xxx^{\mathrm{T}}A + A^2xx^{\mathrm{T}} + xx^{\mathrm{T}}Axx^{\mathrm{T}} + Axx^{\mathrm{T}}xx^{\mathrm{T}} + \left(x^{\mathrm{T}}x\right)^2 xx^{\mathrm{T}}\right) \\
&= \operatorname{tr}\left(A^3\right) + \operatorname{tr}\left(xx^{\mathrm{T}}A^2\right) + \operatorname{tr}\left(Axx^{\mathrm{T}}A\right) + \operatorname{tr}\left(x^{\mathrm{T}}xxx^{\mathrm{T}}A\right) + \operatorname{tr}\left(A^2xx^{\mathrm{T}}\right) + \operatorname{tr}\left(xx^{\mathrm{T}}Axx^{\mathrm{T}}\right) \\
&\quad + \operatorname{tr}\left(Axx^{\mathrm{T}}xx^{\mathrm{T}}\right) + \operatorname{tr}\left(\left(x^{\mathrm{T}}x\right)^2 xx^{\mathrm{T}}\right) \\
&= \operatorname{tr}\left(A^3\right) + 3\operatorname{tr}\left(x^{\mathrm{T}}A^2x\right) + 3(x^{\mathrm{T}}x)\operatorname{tr}\left(x^{\mathrm{T}}Ax\right) + (x^{\mathrm{T}}x)^2\operatorname{tr}\left(x^{\mathrm{T}}x\right) \\
&= \operatorname{tr}\left(A^3\right) + 3x^{\mathrm{T}}A^2x + 3(x^{\mathrm{T}}x)(x^{\mathrm{T}}Ax) + (x^{\mathrm{T}}x)^3 \\
\hline
\nabla g(x) &= 3\nabla_x\{x^{\mathrm{T}}A^2x\} + 3\nabla_x\{(x^{\mathrm{T}}x)(x^{\mathrm{T}}Ax)\} + \nabla_x\{(x^{\mathrm{T}}x)^3\} \\
&= 3(A^2+(A^2)^{\mathrm{T}})x + 3((x^{\mathrm{T}}x)\cdot\nabla_x\{x^{\mathrm{T}}Ax\} + \nabla_x\{x^{\mathrm{T}}x\}\cdot(x^{\mathrm{T}}Ax)) + 3(x^{\mathrm{T}}x)^2\nabla_x\{x^{\mathrm{T}}x\} \\
&= 3(A^2+(A^2)^{\mathrm{T}})x + 3((x^{\mathrm{T}}x)\cdot(A+A^{\mathrm{T}})x + 2x\cdot(x^{\mathrm{T}}Ax)) + 3(x^{\mathrm{T}}x)^2\cdot 2x \quad (\text{note that } A^{\mathrm{T}} = A) \\
&= 6A^2x + 6(x^{\mathrm{T}}x)Ax + 6(x^{\mathrm{T}}Ax)x + 6(x^{\mathrm{T}}x)^2x \\
&= 6(A^2 + Axx^{\mathrm{T}} + xx^{\mathrm{T}}A + xx^{\mathrm{T}}xx^{\mathrm{T}})x \\
&= 6(A+xx^{\mathrm{T}})^2x
\end{aligned}
$$

Problem 3

设 n 为正整数，$a, b \in \mathbb{R}^n$ 为给定的向量.
试求下列 $\mathbb{R}^{n\times n} \mapsto \mathbb{R}$ 的函数的梯度:

$$
f(X) = a^{\mathrm{T}}X^{\mathrm{T}}Xb.
$$

邵老师提供的解法:

> 一般地，我们有:
>
> $$\nabla_X \operatorname{tr}(AX) = \nabla_X \operatorname{tr}(XA) = \nabla_X \operatorname{tr}\left(A^{\mathrm{T}}X^{\mathrm{T}}\right) = \nabla_X \operatorname{tr}\left(X^{\mathrm{T}}A^{\mathrm{T}}\right) = A^{\mathrm{T}}.$$
>
> 特殊地，我们有 $\nabla_X \operatorname{tr}(X) = I$.
> 据此可推出:
>
> $$\nabla_X a^{\mathrm{T}}Xb = \nabla_X \operatorname{tr}\left(a^{\mathrm{T}}Xb\right) = \nabla_X \operatorname{tr}\left(ba^{\mathrm{T}}X\right) = \nabla_X \operatorname{tr}\left(X^{\mathrm{T}}ab^{\mathrm{T}}\right) = \nabla_X b^{\mathrm{T}}X^{\mathrm{T}}a = ab^{\mathrm{T}}.$$
>
> 一般来说，我们有:
>
> $$\nabla_X \operatorname{tr}\left(AX^k\right) = \nabla_X \operatorname{tr}\left(X^kA\right) = \sum_{i=0}^{k-1}(X^iAX^{k-1-i})^{\mathrm{T}}.$$

现在我们计算 $f(X)$ 的梯度:

$$
\begin{aligned}
\nabla_X f(X) &= \nabla_X\{a^{\mathrm{T}}X^{\mathrm{T}}Xb\} \\
&= \nabla_X\{\operatorname{tr}\left(a^{\mathrm{T}}X^{\mathrm{T}}Xb\right)\} \\
&= \nabla_X\{\operatorname{tr}\left(X^{\mathrm{T}}Xba^{\mathrm{T}}\right)\} \\
&= \left(\nabla_{X_1}\{\operatorname{tr}(X_1^{\mathrm{T}}X_2ba^{\mathrm{T}})\} + \nabla_{X_2}\{\operatorname{tr}(ba^{\mathrm{T}}X_1^{\mathrm{T}}X_2)\}\right)\Big|_{X_1=X_2=X} \\
&= \left(X_2ba^{\mathrm{T}} + (ba^{\mathrm{T}}X_1^{\mathrm{T}})^{\mathrm{T}}\right)\Big|_{X_1=X_2=X} \\
&= X(ba^{\mathrm{T}} + ab^{\mathrm{T}}).
\end{aligned}
$$

My Solution:
首先我们证明一个引理:

> **Lemma:** $\nabla_Y\{a^{\mathrm{T}}Yb\} = ab^{\mathrm{T}}$ $(\forall\, Y = [y_{ij}] \in \mathbb{C}^{n\times n})$
> **Proof:**
>
> $$
> \begin{aligned}
> \nabla_Y\{a^{\mathrm{T}}Yb\} &= \left[\frac{\partial(a^{\mathrm{T}}Yb)}{\partial y_{ij}}\right]_{i,j=1}^n \\
> &= \left[\frac{\partial}{\partial y_{ij}}\sum_{k,l=1}^n a_ky_{kl}b_l\right]_{i,j=1}^n \\
> &= [a_ib_j]_{i,j=1}^n \\
> &= ab^{\mathrm{T}}
> \end{aligned}
> $$

现在我们计算 $f(X)$ 的梯度:

$$
\begin{aligned}
\nabla_X f(x) &= \nabla_X\{a^{\mathrm{T}}X^{\mathrm{T}}Xb\} \\
&= \left[\nabla_{X_1}\{a^{\mathrm{T}}X_1^{\mathrm{T}}X_2b\} + \nabla_{X_2}\{a^{\mathrm{T}}X_1^{\mathrm{T}}X_2b\}\right]\Big|_{X_1=X_2=X} \\
&= \left[\nabla_{X_1}\{(X_2b)^{\mathrm{T}}X_1a\} + \nabla_{X_2}\{(X_1a)X_2b\}\right]\Big|_{X_1=X_2=X} \quad (\text{utilizing Lemma:}\nabla_Y\{a^{\mathrm{T}}Yb\} = ab^{\mathrm{T}}) \\
&= \left[(X_2b)a^{\mathrm{T}}\} + \{(X_1a)b^{\mathrm{T}}\}\right]\Big|_{X_1=X_2=X} \\
&= Xba^{\mathrm{T}} + Xab^{\mathrm{T}} \\
&= X(ba^{\mathrm{T}} + ab^{\mathrm{T}})
\end{aligned}
$$

下面是另一种更繁难的解法:
(不幸的是，这是我最开始想到的解法)

记 X 的行向量为 $x_1^{\mathrm{T}},\dots,x_n^{\mathrm{T}}$，记 X 的 (i,j) 位置元素为 x_{ij}.
定义 Kronecker δ-函数为:

$$
\delta_{ij} := \begin{cases} 1 & \text{if } i=j \\ 0 & \text{otherwise,} \end{cases}
$$

则我们有:

$$
\begin{aligned}
\frac{\partial (X^{\mathrm{T}}X)_{kl}}{\partial x_{ij}} &= \frac{\partial}{\partial x_{ij}}\{x_k^{\mathrm{T}}x_l\} \\
&= \frac{\partial}{\partial x_{ij}}\left\{\sum_{m=1}^n x_{km}x_{lm}\right\} \\
&= \delta_{ik}x_{lj} + \delta_{il}x_{kj}
\end{aligned}
$$

定义单位基矩阵 E_{ij} (它在 (i,j) 位置上为 1，在其余位置为零).
注意到 $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$，因此我们有:

$$
\begin{aligned}
\nabla_{x_{ij}}\{X^{\mathrm{T}}X\} &= \left[\frac{\partial (X^{\mathrm{T}}X)_{kl}}{\partial x_{ij}}\right]_{k,l=1}^n \\
&= [\delta_{ik}x_{lj} + \delta_{il}x_{kj}]_{k,l=1}^n \\
&= [\delta_{ik}x_{lj}]_{k,l=1}^n + [\delta_{il}x_{kj}]_{k,l=1}^n \\
&= X^{\mathrm{T}}E_{ij} + E_{ji}X.
\end{aligned}
$$

记 e_i 为 $\mathbb{C}^n$ 的标准单位基向量，记 $a = [a_1,\dots,a_n]^{\mathrm{T}}, b = [b_1,\dots,b_n]^{\mathrm{T}}$.
回忆起 X 的行向量为 $x_1^{\mathrm{T}},\dots,x_n^{\mathrm{T}}$，于是我们有:

$$
\begin{aligned}
\nabla f(X) &= \nabla_X\{a^{\mathrm{T}}X^{\mathrm{T}}Xb\} \quad (\text{utilizing Lemma:}\nabla_Y\{a^{\mathrm{T}}Yb\} = ab^{\mathrm{T}}) \\
&= \nabla_X\{X^{\mathrm{T}}X\}\cdot(ab^{\mathrm{T}}) \quad (\text{note that } \nabla_X\{X^{\mathrm{T}}X\} \in \mathbb{R}^{n\times n\times n\times n} \text{ is a four dimensional tensor}) \\
&= [a^{\mathrm{T}}\nabla_{x_{ij}}\{X^{\mathrm{T}}X\}b]_{i,j=1}^n \\
&= [a^{\mathrm{T}}(X^{\mathrm{T}}E_{ij} + E_{ji}X)b]_{i,j=1}^n \\
&= [a^{\mathrm{T}}X^{\mathrm{T}}E_{ij}b]_{i,j=1}^n + [a^{\mathrm{T}}E_{ji}Xb]_{i,j=1}^n \\
&= [a^{\mathrm{T}}X^{\mathrm{T}}b_je_i]_{i,j=1}^n + [a_je_i^{\mathrm{T}}Xb]_{i,j=1}^n \\
&= [b_ja^{\mathrm{T}}x_i]_{i,j=1}^n + [a_jx_i^{\mathrm{T}}b]_{i,j=1}^n \\
&= [(x_i^{\mathrm{T}}a)b_j]_{i,j=1}^n + [(x_i^{\mathrm{T}}b)a_j]_{i,j=1}^n \\
&= (Xa)b^{\mathrm{T}} + (Xb)a^{\mathrm{T}} \\
&= X(ab^{\mathrm{T}} + ba^{\mathrm{T}}).
\end{aligned}
$$

Problem 4

定义 $\mathbb{C}^2$ 上的实值函数 $\|\cdot\|_{1/2}$ 为 $\|x\|_{1/2} := (|x_1|^{\frac{1}{2}} + |x_2|^{\frac{1}{2}})^2$ $(\forall\, x \in \mathbb{C}^2)$.
试证明 $\|\cdot\|_{1/2}$ 不是 $\mathbb{C}^2$ 上的范数.

- 事实上，$\|\cdot\|_p$ $(0<p<1)$ 都不是 $\mathbb{C}^n$ 上的范数，参见本次作业的 **Problem 10**.
- **(反 Minkowski 不等式)**
  给定 $0<p<1$ 和 $x,y \in \mathbb{C}^n$.
  若存在 $\lambda_1,\dots,\lambda_n \ge 0$ 使得 $y_i = \lambda_i x_i$ $(\forall\, i = 1,\dots,n)$,
  则我们有 $\|x+y\|_p \ge \|x\|_p + \|y\|_p$ 成立.
  这说明由 $\|x\|_p := (\sum_{i=1}^n |x_i|^p)^{1/p}$ $(\forall\, x \in \mathbb{C}^n)$ 定义的函数 $\|\cdot\|_p$ 在 $p \in (0,1)$ 时不是范数.

证明过程参见 **"FDU 高等线性代数 2. 范数 & 内积"**，搜索关键词 "反 Minkowski 不等式".

Solution:
我们只要举例说明 $\|\cdot\|_{1/2}$ 不满足三角不等式即可.
根据反 Minkowski 不等式的指导 (需要避开取等条件)，我们可以构造如下反例:

$$
x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad x + y = \begin{bmatrix} 1 \\ 1 \end{bmatrix}
$$

$$
\begin{aligned}
\|x+y\|_{1/2} - \|x\|_{1/2} - \|y\|_{1/2} &= (\sqrt{1} + \sqrt{1})^2 - (1+0)^2 - (0+1)^2 \\
&= 4 - 1 - 1 \\
&= 2 \\
&> 0
\end{aligned}
$$

因此三角不等式 $\|x+y\|_{1/2} \le \|x\|_{1/2} + \|y\|_{1/2}$ 并不对任意 $x, y \in \mathbb{C}^2$ 成立.
这表明 $\|\cdot\|_{1/2}$ 不是 $\mathbb{C}^2$ 上的范数.

> 如果不使用反 Minkowski 不等式，则可以基于如下推理构造反例.
> 假设 $x, y \in \mathbb{R}^2_+$ (即它们的元素都是非负实数)，则我们有:
>
> $$
> \begin{aligned}
> &\|x+y\|_{1/2} - \|x\|_{1/2} - \|y\|_{1/2} \\
> &= \left((x_1+y_1)^{\frac{1}{2}} + (x_2+y_2)^{\frac{1}{2}}\right)^2 - (x_1^{\frac{1}{2}} + x_2^{\frac{1}{2}})^2 - (y_1^{\frac{1}{2}} + y_2^{\frac{1}{2}})^2 \\
> &= 2\left(\left(\frac{1}{2}x_1 + \frac{1}{2}y_1\right)^{\frac{1}{2}} + \left(\frac{1}{2}x_2 + \frac{1}{2}y_2\right)^{\frac{1}{2}}\right)^2 - (x_1^{\frac{1}{2}} + x_2^{\frac{1}{2}})^2 - (y_1^{\frac{1}{2}} + y_2^{\frac{1}{2}})^2 \quad (\text{use concavity of } f(x) = \sqrt{x}\ (x \ge 0)) \\
> &\ge 2\left(\left(\frac{1}{2}x_1^{\frac{1}{2}} + \frac{1}{2}y_1^{\frac{1}{2}}\right) + \left(\frac{1}{2}x_2^{\frac{1}{2}} + \frac{1}{2}y_2^{\frac{1}{2}}\right)\right)^2 - (x_1^{\frac{1}{2}} + x_2^{\frac{1}{2}})^2 - (y_1^{\frac{1}{2}} + y_2^{\frac{1}{2}})^2 \\
> &= \frac{1}{2}\left((x_1^{\frac{1}{2}} + x_2^{\frac{1}{2}}) + (y_1^{\frac{1}{2}} + y_2^{\frac{1}{2}})\right)^2 - (x_1^{\frac{1}{2}} + x_2^{\frac{1}{2}})^2 - (y_1^{\frac{1}{2}} + y_2^{\frac{1}{2}})^2 \\
> &= -\frac{1}{2}\left((x_1^{\frac{1}{2}} + x_2^{\frac{1}{2}}) - (y_1^{\frac{1}{2}} + y_2^{\frac{1}{2}})\right)^2
> \end{aligned}
> $$
>
> 注意到最末端的式子在 $x_1^{\frac{1}{2}} + x_2^{\frac{1}{2}} = y_1^{\frac{1}{2}} + y_2^{\frac{1}{2}}$ 时等于零，
> 为要求凹性保证的不等式严格成立，我们还要求 $x \ne y$.
> 因此我们只需取满足上面两个条件的 x, y，便可构造反例.

Problem 5

($\mathbb{C}^n$ 上范数的等价性, Matrix Analysis 推论 $5.4.5$**)**
设 n 为正整数，$\|\cdot\|_\alpha, \|\cdot\|_\beta$ 为 $\mathbb{C}^n$ 上的范数.
试证明存在正实数 $C_{\min}$ 和 $C_{\max}$ 使得:

$$
C_{\min}\|x\|_\alpha \le \|x\|_\beta \le C_{\max}\|x\|_\alpha \ (\forall\ x \in \mathbb{C}^n).
$$

- 简单起见，在下面的证明中我们直接使用了范数的连续性，
  但这里存在循环论证的风险，因为在笔记中，范数的连续性是通过范数的等价性证明的.
  事实上，在 *Matrix Analysis* 一书中，范数等价性的证明依赖于引理 $5.4.3$.

Solution:
在 Euclid 单位球面 $S := \{x \in \mathbb{C}^n : \|x\|_2 = 1\}$ (它是一个紧集) 上定义 $h(x) = \|x\|_\beta / \|x\|_\alpha$.
根据范数的连续性可知 h 是紧集 S 上的连续函数.
根据 **Weierstrass 定理**可知 h 在 S 上存在有限的最小值 $C_{\min}$ 和最大值 $C_{\max}$，
即对于任意 $x \in S$ (即 $x \in \mathbb{C}^n$ 且 $\|x\|_2 = 1$) 我们都有 $C_{\min} \le h(x) = \|x\|_\beta / \|x\|_\alpha \le C_{\max}$.
根据范数的非负性和正定性，并结合 $0_n \notin S$ 可知 $C_{\min}, C_{\max} > 0$.

根据范数的齐次性可知，对于任意非零向量 $x \in \mathbb{C}^n \backslash \{0_n\}$ 都有:

$$
\begin{gathered}
C_{\min} \le h\left(\frac{x}{\|x\|_2}\right) = \frac{\|x/\|x\|_2\|_\beta}{\|x/\|x\|_2\|_\alpha} = \frac{\|x\|_\beta/\|x\|_2}{\|x\|_\alpha/\|x\|_2} = \frac{\|x\|_\beta}{\|x\|_\alpha} \le C_{\max} \\
\Updownarrow \\
C_{\min}\|x\|_\alpha \le \|x\|_\beta \le C_{\max}\|x\|_\alpha.
\end{gathered}
$$

显然 $x = 0_n$ 也满足不等式 $C_{\min}\|x\|_\alpha \le \|x\|_\beta \le C_{\max}\|x\|_\alpha$.
因此我们有:

$$C_{\min}\|x\|_\alpha \le \|x\|_\beta \le C_{\max}\|x\|_\alpha \ \ (\forall\, x\in\mathbb{C}^n).$$

命题得证.

Problem 6 (optional)

设 n 为正整数.
试求下列函数的梯度:

$$f(X)=\log\left(\det\left(X\right)\right) \text{ where } \mathrm{dom}(f)=\{X\in\mathbb{R}^{n\times n}:\det\left(X\right)>0\}.$$

Solution:
根据伴随矩阵的定义我们有:

$$\begin{aligned}(\mathrm{adj}(A))^{\mathrm{T}} &= [(-1)^{i+j}\det\left(A[\{i\}^c,\{j\}^c]\right)]_{i,j=1}^n \\ &= \left[\frac{\partial}{\partial a_{ij}}\det\left(A\right)\right]_{i,j=1}^n,\end{aligned}$$

其中 $A[\{i\}^c,\{j\}^c]$ 代表 A 删去第 i 行和第 j 列所有元素构成的余子阵.
因此我们有:

$$\frac{\partial}{\partial A}\det\left(A\right)=\mathrm{adj}(A).$$

容易验证:

$$[\mathrm{adj}(A)A]_{i,j}=\begin{cases}\det\left(A\right), & \text{if } i=j\\ 0, & \text{otherwise,}\end{cases}$$

这表明 $\mathrm{adj}(A)A=\det\left(A\right)I_n$.
当 A 非奇异时，我们有 $\mathrm{adj}(A)=\det\left(A\right)A^{-1}$ 成立.
因此对于任意满足 $\det\left(X\right)>0$ 的 $X\in\mathbb{R}^{n\times n}$，我们都有:

$$\begin{aligned}\frac{\partial}{\partial X}f(X) &= \frac{\partial}{\partial X}\{\log\left(\det\left(X\right)\right)\}\\ &= \frac{1}{\det\left(X\right)}\cdot\frac{\partial}{\partial X}\det\left(X\right)\\ &= \frac{1}{\det\left(X\right)}\mathrm{adj}(X) \quad (\text{note that } \mathrm{adj}(X)=\det\left(X\right)X^{-1})\\ &= X^{-1}\end{aligned}$$

转置即得:

$$\nabla_X\log\left(\det\left(X\right)\right)=\left(\frac{\partial}{\partial X}f(X)\right)^{\mathrm{T}}=X^{-\mathrm{T}}.$$

Problem 7 (optional)

考虑闭区间 $[0,1]$ 上的实值连续函数构成的空间 $C([0,1])$.
试证明下列定义的 $C([0,1])\mapsto\mathbb{R}$ 的泛函 $\|\cdot\|_2$ 是范数:

$$\|f\|_2:=\left(\int_0^1|f(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}$$

Solution:

- ① 非负性显然成立.
- ② 正定性:

$$\begin{gathered}\|f\|_2=\left(\int_0^1|f(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}=0\\ \Updownarrow\\ f(x)=0 \text{ almost everywhere on } [0,1]\end{gathered}$$

  因此如果我们将 $C([0,1])$ 空间中的几乎处处相等的函数视为同一个函数，那么泛函 $\|\cdot\|_2$ 满足正定性.

- ③ 齐次性:
  对于任意 $\alpha \in \mathbb{R}$ 都有:

$$\begin{aligned}\|\alpha f\|_2 &= \left(\int_0^1 |\alpha f(x)|^2 \mathrm{d}x\right)^{\frac{1}{2}} \\ &= |\alpha|\left(\int_0^1 |f(x)|^2 \mathrm{d}x\right)^{\frac{1}{2}} \\ &= |\alpha|\|f\|_2.\end{aligned}$$

- ④ 次可加性 (三角不等式):
  (这实际上是积分形式的 Minkowski 不等式在 $p=2$ 时的特殊情况)
  任意给定 $f,g\in C([0,1])$，只要 $f+g$ 在 $[0,1]$ 上不几乎处处为零，我们就有:

$$\begin{aligned}\int_0^1 |f(x)+g(x)|^2\mathrm{d}x &= \int_0^1 |f(x)+g(x)|\cdot|f(x)+g(x)|\mathrm{d}x \\ &\leq \int_0^1 |f(x)+g(x)||f(x)|\mathrm{d}x + \int_0^1 |f(x)+g(x)||g(x)|\mathrm{d}x \quad \text{(Cauchy–Schwarz inequality)} \\ &\leq \left(\int_0^1 |f(x)+g(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}\left(\int_0^1 |f(x)|^2\mathrm{d}x\right)^{\frac{1}{2}} + \left(\int_0^1 |f(x)+g(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}\left(\int_0^1 |g(x)|^2\mathrm{d}x\right)^{\frac{1}{2}} \\ &= \left(\int_0^1 |f(x)+g(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}\left(\left(\int_0^1 |f(x)|^2\mathrm{d}x\right)^{\frac{1}{2}} + \left(\int_0^1 |g(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}\right).\end{aligned}$$

  左右同除 $(\int_0^1 |f(x)+g(x)|^2\mathrm{d}x)^{\frac{1}{2}}$ 可得:

$$\begin{gathered}\left(\int_0^1 |f(x)+g(x)|^2\mathrm{d}x\right)^{\frac{1}{2}} \leq \left(\int_0^1 |f(x)|^2\mathrm{d}x\right)^{\frac{1}{2}} + \left(\int_0^1 |g(x)|^2\mathrm{d}x\right)^{\frac{1}{2}} \\ \Updownarrow \\ \|f+g\|_2 \leq \|f\|_2 + \|g\|_2\end{gathered}$$

  事实上，当 $f+g$ 在 $[0,1]$ 上几乎处处为零时，上述不等式仍然成立.
  因此对于任意 $f,g\in C_{[0,1]}$ 都有 $\|f+g\|_2 \leq \|f\|_2+\|g\|_2$ 成立.

  值得注意的是，上述证明中存在一个潜在的逻辑风险.
  严格来说，我们需要先证明由 $\langle f,g\rangle_2 := \int_0^1 f(x)g(x)\mathrm{d}x$ 定义的泛函 $\langle\cdot,\cdot\rangle_2$ 是 $C([0,1])$ 上的内积，
  才能在上述证明中使用 Cauchy–Schwarz 不等式.

  邵老师提供了关于次可加性的另一种证法:
  (这个证法规避了上一种证法的分类讨论，但没有解决潜在的逻辑风险)

$$\begin{aligned}\|f+g\|_2^2 &= \int_0^1 |f(x)+g(x)|^2\mathrm{d}x \\ &\leq \int_0^1 (|f(x)|^2+2|f(x)||g(x)|+|g(x)|^2)\mathrm{d}x \\ &= \int_0^1 |f(x)|^2\mathrm{d}x + \int_0^1 |g(x)|^2\mathrm{d}x + 2\int_0^1 |f(x)||g(x)|\mathrm{d}x \quad \text{(Cauchy–Schwarz inequality)} \\ &\leq \int_0^1 |f(x)|^2\mathrm{d}x + \int_0^1 |g(x)|^2\mathrm{d}x + 2\left(\int_0^1 |f(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}\left(\int_0^1 |g(x)|^2\mathrm{d}x\right)^{\frac{1}{2}} \\ &= \|f\|_2^2+\|g\|_2^2+2\|f\|_2\|g\|_2 \\ &= (\|f\|_2+\|g\|_2)^2\end{aligned}$$

  因此对于任意 $f,g\in C_{[0,1]}$ 都有 $\|f+g\|_2 \leq \|f\|_2+\|g\|_2$ 成立.

综上所述，泛函 $\|\cdot\|_2$ 是 $C_{[0,1]}$ 空间上的范数.

Problem 8

设 m,n 为正整数，$A\in\mathbb{C}^{m\times m}, B\in\mathbb{C}^{n\times n}$.
试用 $\det(A)$ 和 $\det(B)$ 表示 $\det(A\otimes B)$.

- Kronecker 乘积的性质:

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD) \text{ if the matrix products } AB \text{ and } CD \text{ are well-defined}$$

$$\begin{cases} (A \otimes B)^{\mathrm{T}} = A^{\mathrm{T}} \otimes B^{\mathrm{T}} \\ (A \otimes B)^{-1} = A^{-1} \otimes B^{-1} \\ (A \otimes B)^{\dagger} = A^{\dagger} \otimes B^{\dagger} \end{cases}$$

Solution:
设 A 的 Jordan 分解为 $A = P_1 J_1 P_1^{-1}$，B 的 Jordan 分解为 $B = P_2 J_2 P_2^{-1}$.
设 A 的特征值是 $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ (于是 J_1 的对角元为 $\lambda_1, \dots, \lambda_m$).
设 B 的特征值是 $\mu_1, \dots, \mu_n \in \mathbb{C}$ (于是 J_2 的对角元为 $\mu_1, \dots, \mu_n$).
则我们有:

$$\begin{aligned}
\det(A \otimes B) &= \det\left((P_1 J_1 P_1^{-1}) \otimes (P_2 J_2 P_2^{-1})\right) \\
&= \det\left((P_1 \otimes P_2)(J_1 \otimes J_2)(P_1^{-1} \otimes P_2^{-1})\right) \\
&= \det\left((P_1 \otimes P_2)(J_1 \otimes J_2)(P_1 \otimes P_2)^{-1}\right) \\
&= \det\left((J_1 \otimes J_2)(P_1 \otimes P_2)^{-1}(P_1 \otimes P_2)\right) \\
&= \det(J_1 \otimes J_2) \\
&= \prod_{i=1}^{m} \left\{ \lambda_i \prod_{j=1}^{n} \mu_j \right\} \\
&= \prod_{i=1}^{m} \lambda_i \cdot \prod_{j=1}^{n} \mu_j \\
&= \det(A) \cdot \det(B).
\end{aligned}$$

Problem 9

构造非奇异矩阵 P 使得

$$P^{-1}AP = P^{-1} \begin{bmatrix} 1 & 1 & 5 & 6 \\ & 1 & 7 & 8 \\ & & 2 & 3 \\ & & & 4 \end{bmatrix} P = \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 2 & \\ & & & 4 \end{bmatrix} = J.$$

- **Observation:** 等式右侧的矩阵 J 是原矩阵 A 的 Jordan 标准型.

Solution:
注意到相似变换前后的矩阵都是上三角阵,
因此对称行列消元是不可取的，因为这样会破坏上三角的性状.
一个朴素的想法是将 P 取为上三角阵 (特别地，单位上三角阵):

$$P = \begin{bmatrix} 1 & a & b & c \\ & 1 & d & e \\ & & 1 & f \\ & & & 1 \end{bmatrix}.$$

根据等式:

$$\begin{bmatrix} 1 & 1 & 5 & 6 \\ & 1 & 7 & 8 \\ & & 2 & 3 \\ & & & 4 \end{bmatrix} \begin{bmatrix} 1 & a & b & c \\ & 1 & d & e \\ & & 1 & f \\ & & & 1 \end{bmatrix} = \begin{bmatrix} 1 & a & b & c \\ & 1 & d & e \\ & & 1 & f \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 2 & \\ & & & 4 \end{bmatrix},$$

把两边相乘并逐项比较上三角元素，得到以下方程组:

$$\begin{aligned}
&(1,2): a + 1 = 1 + a \\
&(1,3): b + d + 5 = 2b \\
&(1,4): c + e + 5f + 6 = 4c \\
&(2,3): d + 7 = 2d \\
&(2,4): e + 7f + 8 = 4e \\
&(3,4): 2f + 3 = 4f.
\end{aligned}$$

解得:

$$
\begin{aligned}
a \in \mathbb{R}, \quad b = 12, \quad c = \frac{59}{9}, \\
d = 7, \quad e = \frac{37}{6}, \quad f = \frac{3}{2}.
\end{aligned}
$$

因此 P 可以取为:

$$
P = \begin{bmatrix} 1 & a & 12 & \frac{59}{9} \\ & 1 & 7 & \frac{37}{6} \\ & & 1 & \frac{3}{2} \\ & & & 1 \end{bmatrix},
$$

其中 $a \in \mathbb{R}$ 可以是任意实数.
(为方便手工求逆，我们可以取 $a = 0$)

Problem 10

设 n 是大于 1 的整数，$p \in (0,1)$，定义 $\mathbb{C}^n$ 上的函数 $\|\cdot\|_p$ 如下:

$$
\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \quad (\forall\, x \in \mathbb{C}^n).
$$

证明: $\|\cdot\|_p$ 不是 $\mathbb{C}^n$ 上的范数.

- **Note:** 此命题是本次作业 **Problem 04** 的推广.

Solution:
任意给定 $n \geq 2$ 和 $p \in (0,1)$.
在反 Minkowski 不等式的指导下，我们构造如下反例:

$$
x = e_1, \quad y = e_2, \quad x + y = e_1 + e_2,
$$

其中 e_i 代表 $\mathbb{C}^n$ 的第 i 个单位基向量 (即第 i 个元素为 1，其余元素为 0 的 n 维向量).
我们有:

$$
\begin{aligned}
\|x+y\|_p &= (1^p + 1^p)^{\frac{1}{p}} \\
&= 2^{\frac{1}{p}} \quad (\text{note that } \frac{1}{p} > 1) \\
&> 2 \\
&= (1^p)^{\frac{1}{p}} + (1^p)^{\frac{1}{p}} \\
&= \|x\|_p + \|y\|_p.
\end{aligned}
$$

这说明当 $p \in (0,1)$ 时，函数 $\|\cdot\|_p$ 不满足次可加性，因而不是一个范数.

Problem 11 (optional)

设 m, n 为正整数，$A \in \mathbb{C}^{m\times m}, B \in \mathbb{C}^{n\times n}$.
证明: 若 A 和 B 有公共特征值，则存在 $X \in \mathbb{C}^{m\times n} \backslash \{0_{m\times n}\}$ 使得 $AX = XB$.

- **(Sylvester 定理, Matrix Analysis 定理** $2.4.4.1$**)**
  设 $A \in \mathbb{C}^{m\times m}, B \in \mathbb{C}^{n\times n}, C \in \mathbb{C}^{m\times n}$.
  当且仅当 A, B 没有公共特征值 (即 $\text{eig}(A) \cap \text{eig}(B) = \emptyset$) 时，
  Sylvester 方程 $AX - XB = C$ 有唯一解 $X \in \mathbb{C}^{m\times n}$.
  特别地，若 C 是全零矩阵，则此唯一解就是 $X = 0_{m\times n}$.

  证明过程参见 **"FDU 高等线性代数 1. 复数与复方阵"**，搜索关键词 "Sylvester 定理".

Solution:
这个命题是 Sylvester 定理 ($C = 0_{m\times n}$ 的特殊情况) 的反面，证明从略.

Problem 12 (optional)

设 $p\in[1,+\infty)$.
对于闭区间 $[0,1]$ 上的实值连续函数 $f\in C([0,1])$ 定义泛函

$$\|f\|_p:=\left(\int_0^1|f(x)|^p\mathrm{d}x\right)^{1/p},\quad \|f\|_\infty:=\max_{0\leq x\leq 1}|f(x)|.$$

证明:

- ① 泛函 $\|\cdot\|_1,\|\cdot\|_2,\|\cdot\|_\infty$ 是 $C([0,1])$ 上的范数.
- ② 对于任意 $f\in C([0,1])$ 都有 $\lim_{p\to\infty}\|f\|_p=\|f\|_\infty$ 成立.

(1) Part 1

证明泛函 $\|\cdot\|_1,\|\cdot\|_2,\|\cdot\|_\infty$ 是 $C([0,1])$ 上的范数.

Solution:
本次作业的 **Problem 07** 已经证明了泛函 $\|\cdot\|_2$ 是 $C([0,1])$ 上的范数，
其中最难证明的是次可加性 (即三角不等式).
事实上，这一性质可以借助积分形式的 Minkowski 不等式来直接推导，
但引入积分形式的 Minkowski 不等式是有代价的，需要一些泛函分析的知识，
因此我们还是采用直接证明的方法.

下面证明泛函 $\|\cdot\|_1$ 是 $C([0,1])$ 上的范数.

- ① 非负性显然成立.
- ② 正定性:

$$\begin{gathered}\|f\|_1=\int_0^1|f(x)|\mathrm{d}x=0\\ \Updownarrow\\ f(x)=0\text{ almost everywhere on }[0,1]\end{gathered}$$

  因此如果我们将 $C([0,1])$ 空间中的几乎处处相等的函数视为同一个函数，那么泛函 $\|\cdot\|_1$ 满足正定性.

- ③ 齐次性:
  对于任意 $\alpha\in\mathbb{R}$ 都有:

$$\begin{aligned}\|\alpha f\|_1&=\int_0^1|\alpha f(x)|\mathrm{d}x\\&=|\alpha|\int_0^1|f(x)|\mathrm{d}x\\&=|\alpha|\|f\|_1.\end{aligned}$$

- ④ 次可加性 (三角不等式):
  (这实际上是积分形式的 Minkowski 不等式在 $p=1$ 时的特殊情况)
  任意给定 $f,g\in C([0,1])$，我们都有:

$$\begin{aligned}\|f+g\|_1&=\int_0^1|f(x)+g(x)|\mathrm{d}x\\&\leq\int_0^1|f(x)|\mathrm{d}x+\int_0^1|g(x)|\mathrm{d}x\\&\leq\|f\|_1+\|g\|_1.\end{aligned}$$

下面证明泛函 $\|\cdot\|_\infty$ 是 $C([0,1])$ 上的范数.

- ① 非负性显然成立.
- ② 正定性:

$$\begin{gathered}\|f\|_\infty=\max_{0\leq x\leq 1}|f(x)|=0\\ \Updownarrow\\ f(x)=0\text{ almost everywhere on }[0,1]\end{gathered}$$

  因此如果我们将 $C([0,1])$ 空间中的几乎处处相等的函数视为同一个函数，那么泛函 $\|\cdot\|_\infty$ 满足正定性.

- ③ 齐次性:
  对于任意 $\alpha\in\mathbb{R}$ 都有:

$$\begin{aligned}\|\alpha f\|_\infty &= \max_{0\le x\le 1}|\alpha f(x)| \\ &= |\alpha|\max_{0\le x\le 1}|f(x)| \\ &= |\alpha|\|f\|_\infty.\end{aligned}$$

- ④ 次可加性 (三角不等式):
  任意给定 $f, g \in C([0,1])$，我们都有:

$$\begin{aligned}\|f+g\|_\infty &= \max_{0\le x\le 1}|f(x)+g(x)| \\ &\le \max_{0\le x\le 1}|f(x)| + \max_{0\le x\le 1}|g(x)| \\ &\le \|f\|_\infty + \|g\|_\infty.\end{aligned}$$

(2) Part 2

证明对于任意 $f \in C([0,1])$ 都有 $\lim_{p\to\infty}\|f\|_p = \|f\|_\infty$ 成立.

Solution:
任意给定 $f \in C([0,1])$.
一方面，对于任意 $p \ge 1$，我们都有:

$$\begin{aligned}\|f\|_p^p &= \int_0^1 |f(x)|^p \mathrm{d}x \\ &\le \int_0^1 \left(\max_{0\le x\le 1}|f(x)|\right)^p \mathrm{d}x \\ &= \int_0^1 \|f\|_\infty^p \mathrm{d}x \\ &= \|f\|_\infty^p.\end{aligned}$$

因此我们有:

$$\limsup_{p\to\infty}\|f\|_p \le \|f\|_\infty.$$

另一方面，对于任意 $p \ge 1$ 和 $\varepsilon \in (0, \|f\|_\infty)$，我们都有:

$$\begin{aligned}\|f\|_p^p &= \int_0^1 |f(x)|^p \mathrm{d}x \\ &\ge \int_{\{x:|f(x)|\ge \|f\|_\infty - \varepsilon\}} |f(x)|^p \mathrm{d}x \\ &\ge \int_{E_\varepsilon} (\|f\|_\infty - \varepsilon)^p \mathrm{d}x \quad (\text{denote } E_\varepsilon := \{x : |f(x)| \ge \|f\|_\infty - \varepsilon\}) \\ &= (\|f\|_\infty - \varepsilon)^p m(E_\varepsilon),\end{aligned}$$

其中 $m(E_\varepsilon) > 0$ 代表集合 E_ε 的 Lebesgue 测度.
于是我们有:

$$\|f\|_p \ge (\|f\|_\infty - \varepsilon)(m(E_\varepsilon))^{\frac{1}{p}} \to (\|f\|_\infty - \varepsilon) \quad (p \to \infty).$$

因此我们有:

$$\liminf_{p\to\infty}\|f\|_p \ge (\|f\|_\infty - \varepsilon),$$

令 $\varepsilon \to 0$ 即得:

$$\liminf_{p\to\infty}\|f\|_p \ge \|f\|_\infty.$$

综上所述，我们有:

$$\|f\|_\infty \le \liminf_{p\to\infty}\|f\|_p \le \limsup_{p\to\infty}\|f\|_p \le \|f\|_\infty.$$

于是极限 $\lim_{p\to\infty}\|f\|_p$ 存在，且 $\lim_{p\to\infty}\|f\|_p = \|f\|_\infty$.

Problem 13 (optional)

构造次数最低的首一有理系数多项式 $p(\cdot)$，使得 $p(\sqrt{2}+\sqrt{3}+\sqrt{6})=0$.

- **Observation:**
  注意到 $\sqrt{2}+\sqrt{3}+\sqrt{6}$ 是 $\{1,\sqrt{2},\sqrt{3},\sqrt{6}\}$ 的有理系数组合，
  换言之，$(\sqrt{2}+\sqrt{3}+\sqrt{6})\in\mathbb{Q}[\sqrt{2},\sqrt{3}]=\{q_1+q_2\sqrt{2}+q_3\sqrt{3}+q_4\sqrt{6}:q_1,q_2,q_3,q_4\in\mathbb{Q}\}$.
  可以证明 $\mathbb{Q}[\sqrt{2},\sqrt{3}]$ 是包含 $\sqrt{2}+\sqrt{3}+\sqrt{6}$ 的最小数域，参见 **Homework 04 Problem 04**，
  因此能够零化 $\sqrt{2}+\sqrt{3}+\sqrt{6}$ 的首一有理系数多项式的次数至少为 4.

  事实上，$\sqrt{2}+\sqrt{3}+\sqrt{6}$ 的次数最低的首一有理系数零化多项式为:

  $$\begin{aligned}p(x)&=x^4-22x^2-48x-23\\&=(x-(\sqrt{2}+\sqrt{3}+\sqrt{6}))(x-(\sqrt{2}-\sqrt{3}-\sqrt{6}))(x-(-\sqrt{2}+\sqrt{3}-\sqrt{6}))(x-(-\sqrt{2}-\sqrt{3}+\sqrt{6})).\end{aligned}$$

  其四个根对应对 $\sqrt{2}$ 和 $\sqrt{3}$ 做独立的 Galois 共轭置换，$\sqrt{6}$ 的符号随之确定.

(1) Solution 1

记 $x=\sqrt{2}+\sqrt{3}+\sqrt{6}$, $\alpha=\sqrt{6}$，则我们有:

$$\begin{gathered}x-\alpha=\sqrt{2}+\sqrt{3}\\\Downarrow\\(x-\alpha)^2=(\sqrt{2}+\sqrt{3})^2\end{gathered}$$

展开左右两式:

$$\begin{array}{rl}\text{LHS}&=(x-\alpha)^2\\&=x^2-2\alpha x+\alpha^2\\&=x^2-2\alpha x+6\\\hline\text{RHS}&=(\sqrt{2}+\sqrt{3})^2\\&=5+2\sqrt{6}\\&=5+2\alpha\end{array}$$

因此我们有:

$$x^2-2\alpha x+6=5+2\alpha,$$

解得:

$$\alpha=\frac{x^2+1}{2(x+1)}.$$

代入 $\alpha^2=6$ 可得:

$$\left(\frac{x^2+1}{2(x+1)}\right)^2=6.$$

展开并化简可得:

$$x^4-22x^2-48x-23=0.$$

由于其次数为 4，最高次项的系数为 1，
故 $p(x)=x^4-22x^2-48x-23$ 是 $\sqrt{2}+\sqrt{3}+\sqrt{6}$ 的次数最低的首一有理系数零化多项式.

(2) Solution 2

- **Lemma 1:**
  考虑关于 $x\in\mathbb{F}$ 的 n 次首一复系数多项式 $p(x)=x^n+\alpha_{n-1}x^{n-1}+\alpha_{n-2}x^{n-2}+\cdots+\alpha_1x+\alpha_0$.
  方程 $p(x)=0$ 可等价表示为

  $$\begin{bmatrix}0&1&&&\\&0&1&&\\&&\ddots&\ddots&\\&&&0&1\\-\alpha_0&-\alpha_1&\cdots&-\alpha_{n-2}&-\alpha_{n-1}\end{bmatrix}\begin{bmatrix}1\\x\\\vdots\\x^{n-2}\\x^{n-1}\end{bmatrix}=\begin{bmatrix}1\\x\\\vdots\\x^{n-2}\\x^{n-1}\end{bmatrix}x.$$

记上述线性方程组的系数矩阵为 A，
则 $p(x)$ 的根 λ 一定是 A 的特征值，对应的特征向量为 $[1,\lambda,\cdots,\lambda^{n-2},\lambda^{n-1}]^{\mathrm{T}}$.
我们称 A 为 $p(x)$ 的 **Frobenius 友阵** (Frobenius companion matrix).
可以证明 $p(x)$ 既是其 Frobenius 友阵 A 的特征多项式，又是 A 的极小多项式.

- **Lemma 2:**
  设 (λ,x) 是 $A\in\mathbb{C}^{m\times m}$ 的特征对，(μ,y) 是 $B\in\mathbb{C}^{n\times n}$ 的特征对.
  我们有:

$$\begin{array}{rl}(A\otimes B)(x\otimes y) &= (Ax)\otimes(By)\\ &= (x\lambda)\otimes(y\mu)\\ &= (x\otimes y)(\lambda\mu)\\ \hline (A\otimes I_n+I_m\otimes B)(x\otimes y) &= (Ax)\otimes(I_ny)+(I_mx)\otimes(By)\\ &= (x\lambda)\otimes y+x\otimes(y\mu)\\ &= (x\otimes y)(\lambda+\mu).\end{array}$$

  这说明 $(\lambda\mu,x\otimes y)$ 是 $A\otimes B$ 的特征对，$(\lambda+\mu,x\otimes y)$ 是 $(A\otimes I_n+I_m\otimes B)$ 的特征对.
  因此 $(\lambda+\mu+\lambda\mu,x\otimes y)$ 是 $(A\otimes I_n+I_m\otimes B+A\otimes B)$ 的特征对.

Solution:
显然 $\sqrt{2}$ 和 $\sqrt{3}$ 的次数最低的首一零化多项式为:

$$\begin{aligned}p_1(x)&=x^2-2\\ p_2(x)&=x^2-3.\end{aligned}$$

根据 **Lemma 1** 我们可以构造 $p_1(x),p_2(x)$ 的 Frobenius 友阵如下:

$$A=\begin{bmatrix}0&1\\2&0\end{bmatrix},\quad B=\begin{bmatrix}0&1\\3&0\end{bmatrix}.$$

根据 **Lemma 2** 可知 $\sqrt{2}+\sqrt{3}+\sqrt{6}$ 是如下矩阵的特征值:

$$\begin{aligned}&A\otimes I_2+I_2\otimes B+A\otimes B\\ &=\begin{bmatrix}0&0&1&0\\0&0&0&1\\2&0&0&0\\0&2&0&0\end{bmatrix}+\begin{bmatrix}0&1&0&0\\3&0&0&0\\0&0&0&1\\0&0&3&0\end{bmatrix}+\begin{bmatrix}0&0&0&1\\0&0&3&0\\0&2&0&0\\6&0&0&0\end{bmatrix}\\ &=\begin{bmatrix}0&1&1&1\\3&0&3&1\\2&2&0&1\\6&2&3&0\end{bmatrix}.\end{aligned}$$

容易验证，上述矩阵的特征多项式为 $p(x)=x^4-22x^2-48x-23$.
由于其次数为 4，最高次项的系数为 1，
故 $p(x)$ 是 $\sqrt{2}+\sqrt{3}+\sqrt{6}$ 的次数最低的首一有理系数零化多项式.

The End