高等线性代数 Homework 07

Due: Oct. 28, 2024
姓名: 雍崔扬
学号: 21307140051

Problem 1

给定正整数 m,n，$A\in\mathbb{C}^{m\times n}$ 和 $B\in\mathbb{C}^{n\times m}$.
证明: AB 和 BA 具有完全相同的非零特征值 (计代数重数).
当 $m=n$ 时，AB 和 BA 是否一定相似?

证明:
任意给定矩阵 $A\in\mathbb{C}^{m\times n},B\in\mathbb{C}^{n\times m}$ (其中 $m\geq n$)，我们都有:

$$\begin{bmatrix} I_n & -B \\ & I_m \end{bmatrix}\begin{bmatrix} BA & 0_{n\times m} \\ A & 0_{m\times m} \end{bmatrix}\begin{bmatrix} I_n & B \\ & I_m \end{bmatrix}=\begin{bmatrix} 0_{n\times n} & 0_{n\times m} \\ A & AB \end{bmatrix}.$$

注意到 $\begin{bmatrix} I_n & -B \\ & I_m \end{bmatrix}$ 的逆矩阵即为 $\begin{bmatrix} I_n & B \\ & I_m \end{bmatrix}$.

记:

$$C_1=\begin{bmatrix} BA & 0_{n\times m} \\ A & 0_{m\times m} \end{bmatrix},\quad C_2=\begin{bmatrix} 0_{n\times n} & 0_{n\times m} \\ A & AB \end{bmatrix},$$

则上述等式表明 C_1,C_2 相似，于是 C_1,C_2 的特征值完全相同.
(即特征多项式 $\det(tI_{m+n}-C_1)=\det(tI_{m+n}-C_2)$)

注意到 C_1 的特征值由 BA 的 n 个特征值和 m 个零特征值构成，
(因为特征多项式 $\det(\lambda I_{m+n}-C_1)=\lambda^m\det(\lambda I_n-BA)$)
而 C_2 的特征值由 AB 的 m 个特征值和 n 个零特征值构成.
(因为特征多项式 $\det(\lambda I_{m+n}-C_2)=\lambda^n\det(\lambda I_m-AB)$).
比较二者，即可知 AB 的 m 个特征值即 BA 的 n 个特征值附加上 $m-n$ 个零特征值.
这意味着 AB,BA 的非零特征值是完全相同的 (计代数重数)，而零特征值的个数相差 $m-n$ 个.

特殊地，当 $m=n$ 时，AB 和 BA 具有完全相同的特征值 (无论是非零特征值还是零特征值).
此时若 A,B 至少有一个是非奇异阵 (不妨设 A 非奇异)，则有 $AB=A(BA)A^{-1}$，表明 AB 和 BA 相似.
但 A,B 均为非奇异阵时，AB 不一定相似于 BA，例如:

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

这是因为特征值完全相同 (计代数重数) 并不能保证 Jordan 标准型相同.
更多内容可参考: [(On the similarity of AB and BA for normal and other matrices)](#).

TA 提供的解法:
先将第二列右乘 A 减到第一列上 (列变换是右乘)，
再将第一行左乘 A 减到第二行上 (行变换是左乘):

$$\begin{vmatrix} \lambda I_n-BA & B \\ 0_{m\times n} & \lambda I_m \end{vmatrix}=\begin{vmatrix} \lambda I_n & B \\ \lambda A & \lambda I_m \end{vmatrix}=\begin{vmatrix} \lambda I_n & B \\ 0_{m\times n} & \lambda I_m-AB \end{vmatrix}.$$

实际上就是将我的解法简化了，后面的步骤相同.

TA 提供的涉及极限思想的解法:
特殊地，当 $m=n$ 时，AB 和 BA 具有完全相同的特征值 (无论是非零特征值还是零特征值).
我们可以这样证明:

- ① 若 A, B 中至少有一个非奇异 (不妨设 A 非奇异)，则我们有 $AB = A(BA)A^{-1}$.
  这表明 AB 和 BA 相似，因而具有完全相同的特征值 (无论是非零特征值还是零特征值).
  我们有 $\det(I_n - AB) = \det(I_n - BA)$ 成立.
  (或者直接由 $\det(I_n - AB) = \det(A(BA)A^{-1}) = \det(I_n - BA)$ 说明)
- ② 若 A, B 均非奇异，我们可对 A 施加任意小的扰动 εI_n 得到非奇异阵 $A + \varepsilon I_n$.
  根据 ① 的结论我们有 $\det(I_n - (A + \varepsilon I_n)B) = \det(I_n - B(A + \varepsilon I_n))$.
  注意到 $\det(I_n - (A + \varepsilon I_n)B)$ 和 $\det(I_n - B(A + \varepsilon I_n))$ 都是关于 ε 的连续函数.
  因此我们有:

$$\begin{aligned}\det(I_n - AB) &= \lim_{\varepsilon \to 0} \det(I_n - (A + \varepsilon I_n)B) \\ &= \lim_{\varepsilon \to 0} \det(I_n - B(A + \varepsilon I_n)) \\ &= \det(I_n - BA).\end{aligned}$$

  这表明 AB 和 BA 具有完全相同的特征值 (无论是非零特征值还是零特征值).
  尽管 AB 和 BA 不一定相似.

Problem 2

给定正整数 n，$A \in \mathbb{C}^{n \times n}$，设其特征值为 $\lambda_1, \ldots, \lambda_n$.
证明: 对于复多项式 $f(t)$，$f(A)$ 的特征值为 $f(\lambda_1), \ldots, f(\lambda_n)$.

- 一个简单的想法是:

$$\begin{gathered}Ax = x\lambda \\ A^2 x = A(Ax) = A(x\lambda) = (x\lambda)\lambda = x\lambda^2 \\ A^3 x = A(A^2 x) = A(x\lambda^2) = (x\lambda)\lambda^2 = x\lambda^3 \\ \vdots \\ A^k x = x\lambda^k\end{gathered}$$

  因此对于复多项式 $f(t)$，直观上有 $f(A)x = xf(\lambda)$ 成立.
  但上述逻辑只能保到几何重数 (因为依赖于特征向量 x).

证明:
设复方阵 $A \in \mathbb{C}^{n \times n}$ 的 Jordan 标准型为:

$$S^{-1}AS = J = \begin{bmatrix} J^{(1)}(\mu_1) & & & \\ & J^{(2)}(\mu_2) & & \\ & & \ddots & \\ & & & J^{(d)}(\mu_d) \end{bmatrix},$$

其中 $\mu_1, \ldots, \mu_d$ 互不相同，$J_1(\mu_1), \ldots, J_d(\mu_d)$ 为对应的 Jordan 矩阵:

$$\begin{cases} J^{(1)}(\mu_1) = J_{n_1^{(1)}}(\mu_1) \oplus \cdots \oplus J_{n_1^{(p_1)}}(\mu_1) \text{ where } n_1 = \sum_{i=1}^{p_1} n_1^{(i)} \quad (n_1^{(1)} \geq \cdots \geq n_1^{(p_1)} \geq 1) \\ \qquad \cdots \\ J^{(d)}(\lambda_d) = J_{n_d^{(1)}}(\mu_d) \oplus \cdots \oplus J_{n_d^{(p_d)}}(\mu_d) \text{ where } n_d = \sum_{i=1}^{p_d} n_d^{(i)} \quad (n_d^{(1)} \geq \cdots \geq n_d^{(p_d)} \geq 1). \end{cases}$$

任意给定一元多项式 $f(t)$.
要证明 $f(A)$ 的特征值为 $f(\lambda_1), \ldots, f(\lambda_n)$,
只需证明 $f(J)$ 的特征值为 $f(\lambda_1), \ldots, f(\lambda_n)$,
只需证明对于任意 $\mu \in \mathbb{C}$ 和正整数 $m \in \mathbb{Z}_+$，$f(J_m(\mu))$ 的特征值为 $f(\mu)$,
其中 $J_m(\mu)$ 代表关于 μ 的 m 的 Jordan 块:

$$J_m(\mu) := \begin{bmatrix} \mu & 1 & & & \\ & \mu & 1 & & \\ & & \ddots & \ddots & \\ & & & \mu & 1 \\ & & & & \mu \end{bmatrix}_{m \times m}$$

显然对于任意自然数 $k \in \mathbb{N}$，幂 $(J_m(\mu))^k$ 均为上三角阵，且主对角元均为 μ^k.
因此其特征值为 μ^k.
于是 $f(J_m(\mu))$ 的特征值为 $f(\mu)$.
(代数重数和几何重数可能会发生改变，这取决于 f 是否会将不同的特征值映射到相同的值)
命题得证.

Problem 3

(Eigenvalues of the rank one matrix uv^{T})
给定正整数 n，$u, v \in \mathbb{C}^n$.
试求 uv^{H} 的特征多项式和极小多项式.

- **Lemma (Matrix Analysis 定理 $3.3.6$)**
  设复方阵 $A \in \mathbb{C}^{n\times n}$ 的 Jordan 标准型为:

$$S^{-1}AS = J = \begin{bmatrix} J^{(1)}(\lambda_1) & & & \\ & J^{(2)}(\lambda_2) & & \\ & & \ddots & \\ & & & J^{(d)}(\lambda_d) \end{bmatrix},$$

  其中 $\lambda_1, \dots, \lambda_d$ 互不相同，$J_1(\lambda_1), \dots, J_d(\lambda_d)$ 为对应的 Jordan 矩阵:

$$\begin{cases} J^{(1)}(\lambda_1) = J_{n_1^{(1)}}(\lambda_1) \oplus \cdots \oplus J_{n_1^{(p_1)}}(\lambda_1) \text{ where } n_1 = \sum_{i=1}^{p_1} n_1^{(i)} \quad (n_1^{(1)} \geq \cdots \geq n_1^{(p_1)} \geq 1) \\ \qquad \cdots \\ J^{(d)}(\lambda_d) = J_{n_d^{(1)}}(\lambda_d) \oplus \cdots \oplus J_{n_d^{(p_d)}}(\lambda_d) \text{ where } n_d = \sum_{i=1}^{p_d} n_d^{(i)} \quad (n_d^{(1)} \geq \cdots \geq n_d^{(p_d)} \geq 1). \end{cases}$$

  定义:

$$\begin{aligned} r_i &:= \min\{k \in \mathbb{Z}_+ : (J^{(i)}(\lambda_i) - \lambda_i I_{n_i})^k = 0_{n_i\times n_i}\} \\ &= \max\{n_i^{(1)}, \dots, n_i^{(p_i)}\}, \end{aligned} \quad (i = 1, \dots, d)$$

  即 r_i 为 A 关于特征值 λ_i 的所有 Jordan 块 $J_{n_i^{(1)}}(\lambda_i), \dots, J_{n_i^{(p_i)}}(\lambda_i)$ 的最大的阶，
  则 $J^{(i)}(\lambda_i)$ 的极小多项式 (即最小次数的首一零化多项式) 为 $m_{J^{(i)}(\lambda_i)}(t) = (t - \lambda_i)^{r_i}$.
  因此 A 的极小多项式为:

$$m_A(t) = m_J(t) = \prod_{i=1}^{d} m_{J^{(i)}(\lambda_i)}(t) = \prod_{i=1}^{d} (t - \lambda_i)^{r_i}.$$

Solution:

- **Case** 1:
  若 u, v 至少有一个是零向量，则 $A := uv^{\mathrm{H}} = 0_{n\times n}$.
  特征多项式 $p_A(t) = \det(tI - A) = \det(tI) = t^n$,
  表明 $A = 0_{n\times n}$ 的特征值均为 0 (代数重数为 n).
  考虑到:

$$\begin{aligned} \mathrm{Ker}(A) &= \mathrm{Range}(A^{\mathrm{H}})^{\perp} \\ \dim(\mathrm{Ker}(A)) &= n - \mathrm{rank}(A^{\mathrm{H}}) = n - 0 = n, \end{aligned}$$

  因此 0 特征值的特征子空间 $\mathrm{Ker}(A)$ 的维数为 n，表明 0 特征值的几何重数为 n.
  故 Jordan 标准型中 0 特征值对应的 Jordan 块的最大阶数为 1.
  根据 **Lemma** 可知极小多项式 $m_A(t) = t$.

- **Case** 2:
  若 u, v 均不为零向量，则 $A := uv^{\mathrm{H}}$ 为秩 1 矩阵.
  根据 $(uv^{\mathrm{H}})u = u(v^{\mathrm{H}}u)$ 可知 $v^{\mathrm{H}}u$ 是 $A = uv^{\mathrm{H}}$ 的特征值，u 是对应的特征向量.
  考虑到:

$$\begin{aligned} \mathrm{Ker}(A) &= \mathrm{Range}(A^{\mathrm{H}})^{\perp} \\ \dim(\mathrm{Ker}(A)) &= n - \mathrm{rank}(A^{\mathrm{H}}) = n - 1, \end{aligned}$$

因此 0 特征值的特征子空间 $\mathrm{Ker}(A)$ 的维数为 $n-1$，表明 0 特征值的几何重数为 $n-1$.
于是 0 特征值的代数重数至少为 $n-1$，表明 A 至多有 1 个非零特征值.

- ① 若 $v^{\mathrm{H}}u \neq 0$，则 A 的 0 特征值的代数重数为 $n-1$.
  此时特征值 $v^{\mathrm{T}}u$ 是简单的 (代数重数 = 几何重数 = 1),
  特征值 0 是半简单的 (代数重数 = 几何重数 = $n-1$).
  特征多项式为 $p_A(t) = t^{n-1}(t - v^{\mathrm{H}}u)$
  注意到 Jordan 标准型中特征值 $v^{\mathrm{T}}u$ 和 0 对应的 Jordan 块的最大阶数都为 1.
  根据 **Lemma** 可知极小多项式 $m_A(t) = t(t - v^{\mathrm{T}}u)$
- ② 若 $v^{\mathrm{H}}u = 0$，则 A 的 0 特征值的代数重数为 n，大于其几何重数 $n-1$
  特征多项式为 $p_A(t) = t^n$.
  注意到 Jordan 标准型中 0 特征值对应的 Jordan 块的最大阶数为 2.
  根据 **Lemma** 可知极小多项式 $m_A(t) = t^2$.

Problem 4

(带有周期性边界条件的环形三对角矩阵, 具体来说是 Toeplitz 矩阵)
给定正整数 n，试求以下 n 阶方阵的所有特征值及特征向量:

$$\widetilde{A}_n := \begin{bmatrix} 0 & 1 & & & 1 \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 1 \\ 1 & & & 1 & 0 \end{bmatrix}_{n\times n}$$

- 处理了上述问题，我们可以离散化 Laplace 算子:

$$y'' \approx \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} \text{ where } h > 0 \text{ is the step size: } \begin{cases} x_k = x_{k-1} + h \\ x_{k+1} = x_k + h \end{cases}$$

$$L_n := \begin{bmatrix} -2 & 1 & & & 1 \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ 1 & & & 1 & -2 \end{bmatrix}_{n\times n}$$

Laplace 算子 L_n 的特征值和 $\widetilde{A}_n$ 的特征值只差一个平移.
可以证明 Laplace 算子有一个很好的性质:
其特征向量恰好是 $\widetilde{A}_n$ 的特征向量的离散化 (其他算子没有这样的性质).

邵老师提供的解法:

$$\begin{aligned} A_n &= \begin{bmatrix} 0 & 1 & & & 1 \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 1 \\ 1 & & & 1 & 0 \end{bmatrix}_{n\times n} \\ &= \begin{bmatrix} 0 & & & & 1 \\ 1 & 0 & & & \\ & \ddots & \ddots & & \\ & & 1 & 0 & \\ & & & 1 & 0 \end{bmatrix}_{n\times n} + \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ 1 & & & & 0 \end{bmatrix}_{n\times n} \\ &= U_n + U_n^{n-1}. \end{aligned}$$

考虑 U_n 的特征值和特征向量:
注意到 U_n 是一个 Frobenius 友阵，其特征多项式 $p(t) = t^n - 1$,
因此特征值为 $1, \omega, \ldots, \omega^{n-1}$ (其中 $\omega = \exp\left(\mathrm{i}\frac{2\pi}{n}\right)$).
其中 ω^k $(k = 0, 1, \ldots, n-1)$ 的特征向量为 $[1, \omega^k, \ldots, \omega^{(n-1)k}]^{\mathrm{T}}$.
这是 $\text{Homework 01 Problem 01}$ 的结论.

相应地，U_n^{n-1} 的特征值为 $\omega^{(n-1)k} = \omega^{-k}\ (k = 0, 1, \dots, n-1)$，特征向量也为 $[1, \omega^k, \dots, \omega^{(n-1)k}]^{\mathrm{T}}$.
因此 $A_n = U_n + U_n^{n-1}$ 的特征值为 $\omega^k + \omega^{-k} = 2\cos\left(\frac{2k\pi}{n}\right)\ (k = 0, 1, \dots, n-1)$,
特征向量为 $[1, \omega^k, \dots, \omega^{(n-1)k}]^{\mathrm{T}}$.
将相同的特征值合并在一起，便可说得更详细一点:

- ① $\lambda = 2$ 是 $\widetilde{A}_n$ 的特征值，代数重数和几何重数都是 1.
  特征向量为 $[1, 1, \dots, 1]^{\mathrm{T}}$.
- ② $\lambda = \omega^k + \omega^{-k} = 2\cos\left(\frac{2k\pi}{n}\right)\ (k = 1, 2, \dots, \lfloor\frac{n-1}{2}\rfloor)$ 是 $\widetilde{A}_n$ 的特征值，代数重数和几何重数都是 2.
  特征向量为 $[1, \omega^k, \dots, \omega^{(n-1)k}]^{\mathrm{T}}$ 和 $[1, \omega^{-k}, \cdots, \omega^{-(n-1)k}]^{\mathrm{T}}$.
- ③ 当 n 为偶数时，$\lambda = -2 = 2\cos(\pi) = \omega^{\frac{n}{2}} + \omega^{-\frac{n}{2}}$ 是 $\widetilde{A}_n$ 的特征值，代数重数和几何重数都是 1.
  特征向量为 $[1, -1, \dots, 1, -1]^{\mathrm{T}}$.

My Solution:
给定 $\lambda \in \mathbb{C}$ 我们记:

$$
A_n := \begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 1 \\ & & & 1 & 0 \end{bmatrix}_{n\times n} \qquad D_n := \det(\lambda I - A_n) = \begin{vmatrix} \lambda & -1 & & & \\ -1 & \lambda & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & \lambda & -1 \\ & & & -1 & \lambda \end{vmatrix}_{n\times n}
$$
$$
\widetilde{A}_n = \begin{bmatrix} 0 & 1 & & & 1 \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 1 \\ 1 & & & 1 & 0 \end{bmatrix}_{n\times n} \qquad \widetilde{D}_n := \det\left(\lambda I - \widetilde{A}_n\right) = \begin{vmatrix} \lambda & -1 & & & -1 \\ -1 & \lambda & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & \lambda & -1 \\ -1 & & & -1 & \lambda \end{vmatrix}_{n\times n}
$$

其中 A_n 可称为**非周期性边界条件的三对角矩阵**,
而 $\widetilde{A}_n$ 可称为**带有周期性边界条件的环形三对角矩阵**.
邵老师说 A_n 显然没有重特征值，因为它是不可约三对角阵.
可以发现 $\lambda I_n - A_n$ 的左下 $n-1$ 分块是非奇异的，因此 $\operatorname{rank}(\lambda I_n - A_n) \geq n-1$.
所以几何重数 $\operatorname{null}(\lambda I_n - A_n) \leq 1$，表明没有重特征值.
类似地，不可约上 Hessenberg 矩阵也没有重特征值，这是一个重要的性质.

首先考虑 D_n，我们有:

$D_1 = \lambda$
$D_2 = \lambda^2 - 1$

$$
\begin{aligned}
D_n &= \begin{vmatrix}
\lambda & -1 & & & & & \\
-1 & \lambda & -1 & & & & \\
 & -1 & \lambda & -1 & & & \\
 & & \ddots & \ddots & \ddots & & \\
 & & & -1 & \lambda & -1 & \\
 & & & & -1 & \lambda & -1 \\
 & & & & & -1 & \lambda
\end{vmatrix}_{n\times n} \\
&= \lambda \begin{vmatrix}
\lambda & -1 & & & & \\
-1 & \lambda & -1 & & & \\
 & \ddots & \ddots & \ddots & & \\
 & & -1 & \lambda & -1 & \\
 & & & -1 & \lambda & -1 \\
 & & & & -1 & \lambda
\end{vmatrix}_{(n-1)\times(n-1)} + (-1)(-1)^{1+2} \begin{vmatrix}
-1 & -1 & & & & \\
0 & \lambda & -1 & & & \\
 & -1 & \ddots & \ddots & & \\
 & & \ddots & \lambda & -1 & \\
 & & & -1 & \lambda & -1 \\
 & & & & -1 & \lambda
\end{vmatrix}_{(n-1)\times(n-1)} \\
&= \lambda \begin{vmatrix}
\lambda & -1 & & & & \\
-1 & \lambda & -1 & & & \\
 & \ddots & \ddots & \ddots & & \\
 & & -1 & \lambda & -1 & \\
 & & & -1 & \lambda & -1 \\
 & & & & -1 & \lambda
\end{vmatrix}_{(n-1)\times(n-1)} + (-1) \begin{vmatrix}
\lambda & -1 & & & \\
-1 & \ddots & \ddots & & \\
 & \ddots & \lambda & -1 & \\
 & & -1 & \lambda & -1 \\
 & & & -1 & \lambda
\end{vmatrix}_{(n-2)\times(n-2)} \\
&= \lambda D_{n-1} - D_{n-2} \ (\forall\, n \geq 3).
\end{aligned}
$$

设 ϕ_1, ϕ_2 为极限方程 $\phi^2 = \lambda\phi - 1$ 的两根，则我们有:

$$
\begin{cases}
\phi_1 + \phi_2 = \lambda \\
\phi_1\phi_2 = 1.
\end{cases}
$$

进而有:

$$
\begin{aligned}
D_n - \phi_1 D_{n-1} &= \lambda D_{n-1} - D_{n-2} - \phi_1 D_{n-1} \\
&= (\phi_1 + \phi_2) D_{n-1} - \phi_1\phi_2 D_{n-1} - \phi_1 D_{n-1} \\
&= \phi_2 (D_{n-1} - \phi_1 D_{n-2}) \\
&= \cdots \\
&= \phi_2^{n-2} (D_2 - \phi_1 D_1) \\
&= \phi_2^{n-2} (\lambda^2 - 1 - \phi_1 \lambda) \\
&= \phi_2^{n-2} [(\phi_1 + \phi_2)^2 - \phi_1\phi_2 - \phi_1(\phi_1 + \phi_2)] \\
&= \phi_2^{n-2} \cdot \phi_2^2 \\
&= \phi_2^n.
\end{aligned}
$$

类似地，我们有 $D_n - \phi_2 D_{n-1} = \phi_1^n$.

- 当 $\lambda = 2$ 时，我们有 $\phi_1 = \phi_2 = 1$，于是有:

  $$
  \begin{cases}
  D_1 = 2 \\
  D_n - D_{n-1} = 1 \ (n \geq 2).
  \end{cases}
  $$

  进而有 $D_n = n + 1 \ (\forall\, n \geq 1)$ 成立.

- 当 $\lambda = -2$ 时，我们有 $\phi_1 = \phi_2 = -1$，于是有:

  $$
  \begin{cases}
  D_1 = -2 \\
  D_n + D_{n-1} = (-1)^n \ (n \geq 2).
  \end{cases}
  $$

  进而有 $D_n = (-1)^n (n + 1) \ (\forall\, n \geq 1)$ 成立.

- 当 $\lambda \neq \pm 2$ 时，我们有 $\phi_1 \neq \phi_2$，联立:

  $$
  \begin{cases}
  D_n - \phi_1 D_{n-1} = \phi_2^n \\
  D_n - \phi_2 D_{n-1} = \phi_1^n \\
  \phi_1\phi_2 = 1.
  \end{cases}
  $$

解得:

$$D_n = \frac{\phi_1^{n+1} - \phi_2^{n+1}}{\phi_1 - \phi_2} \ (\forall\, n \geq 1).$$

可以证明 A_n 的特征值 (即 $D_n = 0$ 的解) 为 $\lambda = 2\cos\left(\frac{k\pi}{n+1}\right)$ $(k = 1, \dots, n)$.
令 $\widetilde{D}_n = 0$，即有 $\phi_1^{n+1} - \phi_2^{n+1} = 0$ $(\phi_1 \neq \phi_2)$.
设 $\phi_1 = e^{\mathrm{i}\theta}$ $(\theta \in (0, \pi))$，根据

$$\begin{cases} \phi_1 + \phi_2 = \lambda \\ \phi_1\phi_2 = 1 \end{cases}$$

可知:

$$\begin{cases} \phi_2 = 1/\phi_1 = 1/e^{\mathrm{i}\theta} = e^{-\mathrm{i}\theta} \\ \lambda = \phi_1 + \phi_2 = e^{\mathrm{i}\theta} + e^{-\mathrm{i}\theta} = 2\cos(\theta). \end{cases}$$

于是我们有:

$$\begin{aligned} \phi_1^{n+1} - \phi_2^{n+1} &= (e^{\mathrm{i}\theta})^{n+1} - (e^{-\mathrm{i}\theta})^{n+1} \\ &= e^{\mathrm{i}(n+1)\theta} - e^{-\mathrm{i}(n+1)\theta} \\ &= 2\mathrm{i} \cdot \sin((n+1)\theta) \\ &= 0. \end{aligned}$$

因此我们有:

$$(n+1)\theta = k\pi \ (k \in \mathbb{Z}_+).$$

结合 $\theta \in (0, \pi)$ 的假设可知 $\theta = \frac{k\pi}{n+1}$ $(k = 1, 2, \dots, n)$.
因此 A_n 的特征值为 $\lambda = 2\cos\left(\frac{k\pi}{n+1}\right)$ $(k = 1, \dots, n)$ (显然它们的代数重数均为 1).

对于特征值 $\lambda = 2\cos\left(\frac{k\pi}{n+1}\right)$ $(k = 1, \dots, n)$，求解 $(2\cos\left(\frac{k\pi}{n+1}\right)I - A_n)x = 0_n$:

$$\left(2\cos\left(\frac{k\pi}{n+1}\right)I - A_n\right)x = \begin{bmatrix} 2\cos\left(\frac{k\pi}{n+1}\right) & -1 & & & \\ -1 & 2\cos\left(\frac{k\pi}{n+1}\right) & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2\cos\left(\frac{k\pi}{n+1}\right) & -1 \\ & & & -1 & 2\cos\left(\frac{k\pi}{n+1}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} = 0_n.$$

可以验证上述线性方程组具有实向量解 (可作为实特征向量):
(基于 $2\cos(\theta)\sin(m\theta) = \sin((m-1)\theta) + \sin((m+1)\theta)$ 的事实，同时 $\theta = \frac{k\pi}{n+1}$ 能满足边界条件)

$$x = \begin{bmatrix} \sin\left(\frac{k\pi}{n+1}\right) \\ \sin\left(\frac{2k\pi}{n+1}\right) \\ \vdots \\ \sin\left(\frac{(n-1)k\pi}{n+1}\right) \\ \sin\left(\frac{nk\pi}{n+1}\right) \end{bmatrix}.$$

其次考虑 $\widetilde{D}_n$，我们有:

$\widetilde{D}_1 = \lambda$

$\widetilde{D}_2 = \lambda^2 - 1$

$$
\begin{aligned}
\widetilde{D}_n &= \begin{vmatrix}
\lambda & -1 & & & & & -1 \\
-1 & \lambda & -1 & & & & \\
 & -1 & \lambda & -1 & & & \\
 & & \ddots & \ddots & \ddots & & \\
 & & & -1 & \lambda & -1 & \\
 & & & & -1 & \lambda & -1 \\
-1 & & & & & -1 & \lambda
\end{vmatrix}_{n\times n} \\
&= \lambda D_{n-1} + \begin{vmatrix}
-1 & -1 & & & & \\
0 & \lambda & -1 & & & \\
 & -1 & \ddots & \ddots & & \\
 & & \ddots & \lambda & -1 & \\
 & & & -1 & \lambda & -1 \\
-1 & & & & -1 & \lambda
\end{vmatrix}_{(n-1)\times(n-1)} + (-1)(-1)^{1+n} \begin{vmatrix}
-1 & \lambda & -1 & & & \\
 & -1 & \lambda & -1 & & \\
 & & -1 & \ddots & \ddots & \\
 & & & \ddots & \lambda & -1 \\
 & & & & -1 & \lambda \\
-1 & & & & & -1
\end{vmatrix}_{(n-1)\times(n-1)} \\
&= \lambda D_{n-1} + (-1)D_{n-2} + (-1)(-1)^{(n-1)+1} \begin{vmatrix}
-1 & & & & \\
\lambda & -1 & & & \\
-1 & \ddots & \ddots & & \\
 & \ddots & \lambda & -1 & \\
 & & -1 & \lambda & -1
\end{vmatrix}_{(n-2)\times(n-2)} \\
&\qquad + (-1)^{n+2} \left\{ (-1) \begin{vmatrix}
-1 & \lambda & -1 & & \\
 & -1 & \ddots & \ddots & \\
 & & \ddots & \lambda & -1 \\
 & & & -1 & \lambda \\
 & & & & -1
\end{vmatrix}_{(n-2)\times(n-2)} + (-1)(-1)^{(n-1)+1} D_{n-2} \right\} \\
&= \lambda D_{n-1} - D_{n-2} + (-1)^{n+1}(-1)^{n-2} + (-1)^{n+2}\{(-1)(-1)^{n-2} + (-1)^{n+1} D_{n-2}\} \\
&= \lambda D_{n-1} - D_{n-2} - 1 - 1 - D_{n-2} \\
&= \lambda D_{n-1} - 2D_{n-2} - 2 \quad (\text{note that } D_n = \lambda D_{n-1} - D_{n-2}) \\
&= D_n - D_{n-2} - 2 \ (\forall\, n \geq 3).
\end{aligned}
$$

首先考虑 $n = 1, 2$ 时 $\widetilde{A}_n$ 的特征值和特征向量:

- $n = 1$ 时 $\widetilde{A}_1 := [0]$，特征值为 0，特征向量为 $[1]$.
- $n = 2$ 时 $\widetilde{A}_2 := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$，特征值为 $-1, 1$，特征向量分别为 $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ 和 $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

其次我们考虑 $n \geq 3$ 时 $\widetilde{A}_n$ 的特征值 (即 $\widetilde{D}_n = 0$ 的解):

- 当 $\lambda = 2$ 时，我们有 $D_n = n + 1 \ (\forall\, n \geq 1)$ 成立.
  因此我们有:

$$
\begin{aligned}
\widetilde{D}_n &= D_n - D_{n-2} - 2 \\
&= (n+1) - (n-2+1) - 2 \\
&= 0.
\end{aligned}
$$

  因此 $\lambda = 2$ 是 $\widetilde{A}_n \ (n \geq 3)$ 的一个特征值.

- 当 $\lambda = -2$ 时，我们有 $D_n = (-1)^n(n+1) \ (\forall\, n \geq 1)$ 成立.
  因此我们有:

$$
\begin{aligned}
\widetilde{D}_n &= D_n - D_{n-2} - 2 \\
&= (-1)^n(n+1) - (-1)^{n-2}(n-2+1) - 2 \\
&= 2 \cdot [(-1)^n - 1] \\
&= \begin{cases} 0, & \text{if } n \text{ is even} \\ -4, & \text{if } n \text{ is odd.} \end{cases}
\end{aligned}
$$

因此 $\lambda=-2$ 是 $\widetilde{A}_n\ (n\geq 3)$ 的一个特征值，当且仅当 n 是偶数.

- 当 $\lambda\neq\pm 2$ 时，我们有:

$$D_n=\frac{\phi_1^{n+1}-\phi_2^{n+1}}{\phi_1-\phi_2}\ (\forall\ n\geq 1).$$

因此我们有:

$$\begin{aligned}\widetilde{D}_n&=D_n-D_{n-2}-2\\&=\frac{\phi_1^{n+1}-\phi_2^{n+1}}{\phi_1-\phi_2}-\frac{\phi_1^{n-1}-\phi_2^{n-1}}{\phi_1-\phi_2}-2\quad(\text{note that }a^n-b^n=(a-b)(a^{n-1}+a^{n-2}b+\cdots+ab^{n-2}+b^{n-1}))\\&=\sum_{k=0}^{n}\phi_1^k\phi_2^{n-k}-\sum_{k=0}^{n-2}\phi_1^k\phi_2^{n-2-k}-2\\&=\phi_1^n+\phi_2^n+\sum_{k=1}^{n-1}\phi_1^k\phi_2^{n-k}-\sum_{k=0}^{n-2}\phi_1^k\phi_2^{n-2-k}-2\\&=\phi_1^n+\phi_2^n+(\phi_1\phi_2)\sum_{k=1}^{n-1}\phi_1^{k-1}\phi_2^{n-1-k}-\sum_{k=0}^{n-2}\phi_1^k\phi_2^{n-2-k}-2\quad(\text{note that }\phi_1\phi_2=1)\\&=\phi_1^n+\phi_2^n+\sum_{l=0}^{n-2}\phi_1^{l-1}\phi_2^{n-2-l}-\sum_{k=0}^{n-2}\phi_1^k\phi_2^{n-2-k}-2\\&=\phi_1^n+\phi_2^n-2.\end{aligned}$$

令 $\widetilde{D}_n=0$，即有 $\phi_1^n+\phi_2^n-2=0\ (\phi_1\neq\phi_2)$.
设 $\phi_1=e^{\mathrm{i}\theta}\ (\theta\in(0,\pi))$，根据

$$\begin{cases}\phi_1+\phi_2=\lambda\\\phi_1\phi_2=1\end{cases}$$

可知:

$$\begin{cases}\phi_2=1/\phi_1=1/e^{\mathrm{i}\theta}=e^{-\mathrm{i}\theta}\\\lambda=\phi_1+\phi_2=e^{\mathrm{i}\theta}+e^{-\mathrm{i}\theta}=2\cos(\theta).\end{cases}$$

于是我们有:

$$\begin{aligned}\phi_1^n+\phi_2^n-2&=(e^{\mathrm{i}\theta})^n+(e^{-\mathrm{i}\theta})^n-2\\&=e^{in\theta}+e^{-in\theta}-2\\&=2\cos(n\theta)-2\\&=0.\end{aligned}$$

因此我们有:

$$n\theta=2k\pi\ (k\in\mathbb{Z}_+).$$

结合 $\theta\in(0,\pi)$ 的假设可知 $\theta=\frac{2k\pi}{n}\ (k=1,2,\dots,\lfloor\frac{n}{2}\rfloor)$ (其中 $\lfloor\cdot\rfloor$ 代表下取整).
因此 $\lambda=2\cos(\theta)=2\cos\left(\frac{2k\pi}{n}\right)\ (k=1,2,\dots,\lfloor\frac{n}{2}\rfloor)$.
根据 θ 的 $\lfloor\frac{n}{2}\rfloor$ 个解的等价性可知这 $\lfloor\frac{n}{2}\rfloor$ 个特征值的代数重数均为 2.
由于 $\widetilde{A}_n$ 是对称阵 (自然是正规矩阵)，故一定存在谱分解，
这意味着其所有特征值都是半简单的 (几何重数 $=$ 代数重数).
因此 $\lambda=2\cos(\theta)=2\cos\left(\frac{2k\pi}{n}\right)\ (k=1,2,\dots,\lfloor\frac{n}{2}\rfloor)$ 的代数重数和几何重数都是 2.

> **注:** 严格来说，$\lambda=2\cos(\theta)=2\cos\left(\frac{2k\pi}{n}\right)\ (k=1,2,\dots,\lfloor\frac{n}{2}\rfloor)$ 代数重数均为 2 的事实不能这样简单得出.
> 更严谨的说明方法是它们每一个都对应两个线性无关的特征向量 (参见后面的证明).
> 因此它们每一个的几何重数都是 2，进而每一个的代数重数都至少是 2.
> 但由于它们总共的代数重数是 $2\cdot\lfloor\frac{n}{2}\rfloor$，故它们每一个的代数重数都是 2.

综上所述，当 $n\geq 3$ 时，关于 $\widetilde{A}_n$ 的特征值我们有如下结论:

- ① $\lambda=2=2\cos(0)=2\cos\left(\frac{0\cdot 2\pi}{n}\right)$ 是 $\widetilde{A}_n$ 的特征值，代数重数和几何重数都是 1
- ② $\lambda=2\cos\left(\frac{2k\pi}{n}\right)\ (k=1,2,\dots,\lfloor\frac{n}{2}\rfloor)$ 是 $\widetilde{A}_n$ 的特征值，代数重数和几何重数都是 2
- ③ 当 n 为偶数时，$\lambda=2=2\cos(\pi)=2\cos\left(\frac{n\cdot 2\pi}{n}\right)$ 是 $\widetilde{A}_n$ 的特征值，代数重数和几何重数都是 1

最后我们考虑 $n\geq 3$ 时 $\widetilde{A}_n$ 的特征向量:

- ① 对于特征值 $\lambda=2$，求解 $(2I-\widetilde{A}_n)x=0_n$:

$$(2I-\widetilde{A}_n)x=\begin{bmatrix}2&-1&&&-1\\-1&2&-1&&\\&\ddots&\ddots&\ddots&\\&&-1&2&-1\\-1&&&-1&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_{n-1}\\x_n\end{bmatrix}=\begin{bmatrix}0\\0\\\vdots\\0\\0\end{bmatrix}=0_n$$
$$\Downarrow$$
$$x=\begin{bmatrix}x_1\\x_2\\\vdots\\x_{n-1}\\x_n\end{bmatrix}=\begin{bmatrix}1\\1\\\vdots\\1\\1\end{bmatrix}=1_n$$

  因此可以取特征向量 $x=1_n$.

- ② 对于特征值 $\lambda=2\cos\left(\frac{2k\pi}{n}\right)$ $(k=1,2,\dots,\lfloor\frac{n}{2}\rfloor)$，求解 $(2\cos\left(\frac{2k\pi}{n}\right)I-\widetilde{A}_n)x=0_n$:

$$\left(2\cos\left(\frac{2k\pi}{n}\right)I-\widetilde{A}_n\right)x=\begin{bmatrix}2\cos\left(\frac{2k\pi}{n}\right)&-1&&&-1\\-1&2\cos\left(\frac{2k\pi}{n}\right)&-1&&\\&\ddots&\ddots&\ddots&\\&&-1&2\cos\left(\frac{2k\pi}{n}\right)&-1\\-1&&&-1&2\cos\left(\frac{2k\pi}{n}\right)\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_{n-1}\\x_n\end{bmatrix}=\begin{bmatrix}0\\0\\\vdots\\0\\0\end{bmatrix}=0_n.$$

  可以验证上述线性方程组具有两个线性无关的实向量解 (可作为实特征向量):

$$x^{(1)}=\begin{bmatrix}1\\\cos\left(\frac{2k\pi}{n}\right)\\\cos\left(\frac{4k\pi}{n}\right)\\\vdots\\\cos\left(\frac{2(n-1)k\pi}{n}\right)\end{bmatrix},\quad x^{(2)}=\begin{bmatrix}1\\-\sin\left(\frac{2k\pi}{n}\right)\\-\sin\left(\frac{4k\pi}{n}\right)\\\vdots\\-\sin\left(\frac{2(n-1)k\pi}{n}\right)\end{bmatrix}.$$

  前者基于 $2\cos(\theta)\cos(m\theta)=\cos((m-1)\theta)+\cos((m+1)\theta)$ 的事实，同时 $\theta=\frac{2k\pi}{n}$ 能满足边界条件.
  后者基于 $2\cos(\theta)\sin(m\theta)=\sin((m-1)\theta)+\sin((m+1)\theta)$ 的事实，同时 $\theta=\frac{2k\pi}{n}$ 能满足边界条件.
  更深刻地，可以验证上述线性方程组具有两个线性无关的复向量解 (可作为复特征向量):

$$x^{(1)}=\begin{bmatrix}1\\\exp\left(i\frac{2k\pi}{n}\right)\\\exp\left(i\frac{4k\pi}{n}\right)\\\vdots\\\exp\left(i\frac{2(n-1)k\pi}{n}\right)\end{bmatrix},\quad x^{(2)}=\begin{bmatrix}1\\\exp\left(-i\frac{2k\pi}{n}\right)\\\exp\left(-i\frac{4k\pi}{n}\right)\\\vdots\\\exp\left(-i\frac{2(n-1)k\pi}{n}\right)\end{bmatrix},$$

  其中 $k=1,2,\dots,\lfloor\frac{n}{2}\rfloor$.

- ③ 当 n 为偶数时，对于特征值 $\lambda=-2$，求解 $(-2I-\widetilde{A}_n)x=0_n$:

$$
(-2I - \widetilde{A}_n)x = \begin{bmatrix} -2 & -1 & & & -1 \\ -1 & -2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & -2 & -1 \\ -1 & & & -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} = 0_n
$$

$$
\Downarrow
$$

$$
x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ \vdots \\ 1 \\ -1 \end{bmatrix} = \sum_{k=1}^{n/2} (e_{2k-1} - e_{2k}).
$$

因此可以取特征向量 $x = \sum_{k=1}^{n/2} (e_{2k-1} - e_{2k})$，
其中 n 为偶数，e_k 代表 $\mathbb{R}^n$ 空间的第 k 个标准单位基向量.

(what a waste of parchment)

Problem 5

给定:

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

利用 Cayley-Hamilton 定理计算 $\sum_{k=1}^{8} A^k$.

- **(Cayley-Hamilton 定理, Matrix Analysis 定理 $2.4.3.2$)**
  设 $p_A(t) := \det(tI_n - A)$ 是 $A \in \mathbb{C}^{n\times n}$ 的特征多项式，则我们有 $p_A(A) = 0_{n\times n}$ 成立.
  换言之，任意复方阵都满足其特征方程.

Solution:
A 的特征多项式为 $p_A(t) = \det(tI - A) = (\lambda - 1)\lambda - 1 = \lambda^2 - \lambda - 1 = 0$.
根据 **Cayley-Hamilton 定理**可知 $p_A(A) = A^2 - A - I = 0_{2\times 2}$.
于是我们有:

$$
\begin{aligned}
&A^2 = A + I \\
&A^3 = A(A+I) = A^2 + A = (A+I) + A = 2A + I \\
&A^4 = A(2A+I) = 2A^2 + A = 2(A+I) + A = 3A + 2I \\
&A^5 = A(3A+2I) = 3A^2 + 2A = 3(A+I) + 2A = 5A + 3I \\
&A^6 = A(5A+3I) = 5A^2 + 3A = 5(A+I) + 3A = 8A + 5I \\
&A^7 = A(8A+5I) = 8A^2 + 5A = 8(A+I) + 5A = 13A + 8I \\
&\underline{A^8 = A(13A+8I) = 13A^2 + 8A = 13(A+I) + 8A = 21A + 13I} \\
&A^n = f(n)A + f(n-1)I \; (\forall\, n \geq 1),
\end{aligned}
$$

其中 $\{f(n)\}_{n=1}^{\infty}$ 代表 **Fibonacci 数列**:
($f(n)$ 通项公式的推导与 Problem 4 中 D_n 的通项公式的推导类似，这里不作证明)

$$
\begin{aligned}
f(n) &:= \begin{cases} 0 & n = 0 \\ 1 & n = 1, 2 \\ f(n-1) + f(n-2) & n \geq 3 \end{cases} \\
&= \frac{\phi_1^n - \phi_2^n}{\sqrt{5}},
\end{aligned}
$$

其中 $\phi_{1,2} = (1 \pm \sqrt{5})/2$ 为特征方程 $\phi^2 = \phi + 1$ 的根.
于是我们有:

$$
\begin{aligned}
\sum_{k=1}^{8} A^k &= A^8 + A^7 + A^6 + A^5 + A^4 + A^3 + A^2 + A \\
&= (21A + 13I) + (13A + 8I) + (8A + 5I) + (5A + 3I) + (3A + 2I) + (2A + I) + (A + I) + A \\
&= 54A + 33I \\
&= \begin{bmatrix} 87 & 54 \\ 54 & 33 \end{bmatrix}
\end{aligned}
$$

实际上我们有:

$$
\begin{array}{c}
S_0 := 0 \\
\hline
\begin{aligned}
S_n &:= \sum_{i=1}^{n} f(n) \\
&= 1 + f(1) + f(2) + f(3) + \cdots + f(n) - 1 \\
&= f(2) + f(1) + f(2) + f(3) + \cdots + f(n) - 1 \\
&= f(3) + f(2) + f(3) + \cdots + f(n) - 1 \\
&= \cdots \\
&= f(n) + f(n-1) + f(n) - 1 \\
&= f(n+1) + f(n) - 1 \\
&= f(n+2) - 1
\end{aligned} \quad (\forall\, n \geq 1) \\
\hline
S_8 = f(10) - 1 = 55 - 1 = 54 \\
S_7 = f(9) - 1 = 34 - 1 = 33
\end{array}
$$

因此我们有:

$$
\begin{array}{c}
\begin{aligned}
\sum_{k=1}^{n} A^k &= \sum_{k=1}^{n} [f(k)A + f(k-1)I] \\
&= \sum_{k=1}^{n} f(k)A + \sum_{k=1}^{n} f(k-1)I \\
&= S_n A + S_{n-1} I
\end{aligned} \quad (n \geq 1) \\
\hline
\sum_{k=1}^{8} A^k = S_8 A + S_7 I = 54A + 33I = \begin{bmatrix} 87 & 54 \\ 54 & 33 \end{bmatrix}
\end{array}
$$

Insight: [Linear Recurrence Relations.pdf]()
将 Fibonacci 数列 $\{f(n)\}_{n=1}^{\infty}$ 的递推公式变为矩阵形式能让我们看得更清楚:

$$
\begin{bmatrix} f(k+2) \\ f(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} f(k+1) \\ f(k) \end{bmatrix}
$$

$$
\begin{bmatrix} f(n+1) \\ f(n) \end{bmatrix} = \left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right)^n \begin{bmatrix} f(1) \\ f(0) \end{bmatrix} = \left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right)^n \begin{bmatrix} 1 \\ 0 \end{bmatrix}
$$

我们对系数矩阵 $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ 进行谱分解:

$$
\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1 & \\ & \phi_2 \end{bmatrix} \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 1 \end{bmatrix}^{-1}
$$

$$
\left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right)^n = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1^n & \\ & \phi_2^n \end{bmatrix} \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 1 \end{bmatrix}^{-1}
$$

其中 $\phi_{1,2} = (1 \pm \sqrt{5})/2$ 为特征方程 $\phi^2 = \phi + 1$ 的根.
于是我们有:

$$
\begin{aligned}
\begin{bmatrix} f(n+1) \\ f(n) \end{bmatrix} &= \left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\right)^n \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1^n & \\ & \phi_2^n \end{bmatrix} \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1^n & \\ & \phi_2^n \end{bmatrix} \left(\frac{1}{\phi_1 - \phi_2} \begin{bmatrix} 1 & -\phi_2 \\ -1 & \phi_1 \end{bmatrix}\right) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} \phi_1^{n+1} & \phi_2^{n+1} \\ \phi_1^n & \phi_2^n \end{bmatrix} \left(\frac{1}{\phi_1 - \phi_2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) \\
&= \frac{1}{\phi_1 - \phi_2} \begin{bmatrix} \phi_1^{n+1} - \phi_2^{n+1} \\ \phi_1^n - \phi_2^n \end{bmatrix}
\end{aligned}
$$

即得到 $f(n) = \frac{\phi_1^n - \phi_2^n}{\phi_1 - \phi_2}$，其中 ϕ_1^n, ϕ_2^n 的计算复杂度通过二进制求幂法可以优化为 $\log(n)$ 级别.

变体: 如何处理 $x_{k+1} = \frac{2x_k+3}{3x_k-1}$ 的迭代公式?
考虑递归公式 $x_{k+1} = \frac{ax_k+b}{cx_k+d}$ $(\forall\ k \in \mathbb{Z}_+)$,
我们可以将其等价写作 $x_{k+1} = u_{k+1}/v_{k+1}$，其中:

$$
\begin{gathered}
\begin{cases} u_{k+1} = au_k + bv_k \\ v_{k+1} = cu_k + dv_k \end{cases} \\
\Updownarrow \\
\begin{bmatrix} u_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} u_k \\ v_k \end{bmatrix}
\end{gathered}
$$

详细内容参见 **Homework 08 Problem 06**.

Problem 6 (optional)

给定正整数 n，试证明:
若 $A, B \in \mathbb{C}^{n\times n}$ 可交换 (即 $AB = BA$),
则存在酉矩阵 $U \in \mathbb{C}^{n\times n}$ 使得 $U^{\mathrm{H}}AU$ 和 $U^{\mathrm{H}}BU$ 都是上三角阵.

- **(Schur 分解定理, Matrix Analysis 定理 $2.3.1$)**
  设 $A = [a_{ij}] \in \mathbb{C}^{n\times n}$ 的特征值为 $\lambda_1, \dots, \lambda_n$ (按任意指定的次序排列).
  则存在一个酉矩阵 $U \in \mathbb{C}^{n\times n}$ 使得 $T := U^{\mathrm{H}}AU = [t_{ij}]$ 是以 $\lambda_1, \dots, \lambda_n$ 为对角元的上三角阵.

  Proof:

  设 x 为 A 关于特征值 λ_1 的单位特征向量，即满足 $\begin{cases} Ax = \lambda_1 x \\ \|x\|_2 = 1 \end{cases}$
  任取一个第一列为 x 的酉矩阵 $U_1 = [x, u_2, \dots, u_n] \in \mathbb{C}^{n\times n}$，则我们有:

$$
\begin{aligned}
U_1^{\mathrm{H}} A U_1 &= \begin{bmatrix} x^{\mathrm{H}} \\ u_2^{\mathrm{H}} \\ \vdots \\ u_n^{\mathrm{H}} \end{bmatrix} A \begin{bmatrix} x & u_2 & \cdots & u_n \end{bmatrix} \\
&= \begin{bmatrix} x^{\mathrm{H}} \\ u_2^{\mathrm{H}} \\ \vdots \\ u_n^{\mathrm{H}} \end{bmatrix} \begin{bmatrix} \lambda_1 x & A u_2 & \cdots & A u_n \end{bmatrix} \\
&= \left[\begin{array}{c|ccc} \lambda_1 x^{\mathrm{H}} x & x^{\mathrm{H}} A u_2 & \cdots & x^{\mathrm{H}} A u_n \\ \hline \lambda_1 x^{\mathrm{H}} u_2 & u_2^{\mathrm{H}} A u_2 & \cdots & u_2^{\mathrm{H}} A u_n \\ \vdots & \vdots & & \vdots \\ \lambda_1 x^{\mathrm{H}} u_n & u_n^{\mathrm{H}} A u_2 & \cdots & u_n^{\mathrm{H}} A u_n \end{array}\right] \\
&= \left[\begin{array}{c|ccc} \lambda_1 & x^{\mathrm{H}} A u_2 & \cdots & x^{\mathrm{H}} A u_n \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right]
\end{aligned}
$$

由于 $u_2, \ldots, u_n$ 是标准正交的,
故子矩阵 $A_1 = [u_i^{\mathrm{H}} A u_j]_{i,j=2}^{n} = [u_2, \ldots, u_n]^{\mathrm{H}} A [u_2, \ldots, u_n]$ 的特征值是 $\lambda_2, \ldots, \lambda_n$

对 A_1 重新执行上述过程,
可得到一个酉矩阵 $\widetilde{U}_2 \in \mathbb{C}^{(n-1)\times(n-1)}$ 使得 $\widetilde{U}_2^{\mathrm{H}} A_1 \widetilde{U}_2 = \begin{bmatrix} \lambda_2 & * \\ 0_{n-2} & A_2 \end{bmatrix}$ (其中 A_2 的特征值是 $\lambda_3, \ldots, \lambda_n$)
记 $U_2 = \begin{bmatrix} 1 & \\ & \widetilde{U}_2 \end{bmatrix} \in \mathbb{C}^{n\times n}$ 则我们有:

$$
\begin{aligned}
U_2^{\mathrm{H}} (U_1^{\mathrm{H}} A U_1) U_2 &= U_2^{\mathrm{H}} \begin{bmatrix} \lambda_1 & * \\ & A_1 \end{bmatrix} U_2 \\
&= \begin{bmatrix} 1 & \\ & \widetilde{U}_2 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \lambda_1 & * \\ & A_1 \end{bmatrix} \begin{bmatrix} 1 & \\ & \widetilde{U}_2 \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 & * \\ & \widetilde{U}_2^{\mathrm{T}} A_1 \widetilde{U}_2 \end{bmatrix} \\
&= \left[\begin{array}{c|cc} \lambda_1 & * & * \\ \hline & \lambda_2 & * \\ & & A_2 \end{array}\right]
\end{aligned}
$$

依此类推，我们最终得到 $n-1$ 个酉矩阵 $\{\widetilde{U}_i\}_{i=1}^{n-1}$ (其中 $\widetilde{U}_i \in \mathbb{C}^{(n-i+1)\times(n-i+1)}$)
记 $U_1 = \widetilde{U}_1$ 和 $U_i = \begin{bmatrix} I_{i-1} & \\ & \widetilde{U}_i \end{bmatrix}$ $(i = 2, \ldots, n-1)$
取 $U = U_1 \cdots U_{n-1}$ 即得 A 的 **Schur 分解**:

$$
U^{\mathrm{H}} A U = U_{n-1}^{\mathrm{H}} \cdots U_1^{\mathrm{H}} A U_1 \cdots U_{n-1} = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n \end{bmatrix} \stackrel{\Delta}{=} T
$$

定理得证.

Proof:
现在我们要证明的是: 非空交换族 $\mathcal{F} \subseteq \mathbb{C}^{n\times n}$ 中的所有方阵可同时酉上三角化.
我们注意到三个事实:

- **① 相似变换能够保留交换性:**
  若 $A, B \in \mathbb{C}^{n\times n}$ 可交换 (即 $AB = BA$)，则对于任意非奇异阵 $S \in \mathbb{C}^{n\times n}$ 我们都有:

$$
\begin{aligned}
(S^{-1}AS)(S^{-1}BS) &= S^{-1}ABS \quad (\text{note that } AB = BA) \\
&= S^{-1}BAS \\
&= (S^{-1}BS)(S^{-1}AS)
\end{aligned}
$$

- **② (Matrix Analysis 引理 $1.3.19$) 非空交换族 $\mathcal{F} \subseteq \mathbb{C}^{n\times n}$ 中的所有方阵一定存在公共特征向量:**
  我们这里只证明这个结论的简单版本:
  设 $A, B \in \mathbb{C}^{n\times n}$ 可交换 (即 $AB = BA$)，且 (λ, x) 为 A 的一个特征对 (其中非零向量 $x \in \text{Ker}(A - \lambda I_n)$)
  则我们有:

$$
\begin{gathered}
Ax = x\lambda \\
\Rightarrow \\
ABx = BAx = B(x\lambda)
\end{gathered}
$$

  因此 Bx 也是 A 关于特征值 λ 的特征向量，即 $Bx \in \text{Ker}(A - \lambda I_n)$
  这表明 $\text{Ker}(A - \lambda I_n)$ 是线性变换 B 的不变子空间.

  设 $\text{Ker}(A - \lambda I_n)$ 的维数是 d，一组基为 $v_1, \dots, v_d$
  则 $Bv_1, \dots, Bv_d$ 均能表示为基 $v_1, \dots, v_d$ 的线性组合:

$$
B[v_1, \dots, v_d] = [v_1, \dots, v_d]C
$$

  其中 $C \in \mathbb{C}^{d\times d}$ 是非奇异的系数矩阵.
  设 C 的 Schur 分解为 $C = U^{\text{H}}TU$ (其中 $U \in \mathbb{C}^{d\times d}$ 是酉矩阵，而 $T \in \mathbb{C}^{n\times n}$ 是上三角阵)
  则我们有:

$$
\begin{gathered}
B[v_1, \dots, v_d] = [v_1, \dots, v_d]C = [v_1, \dots, v_d]UTU^{\text{H}} \\
\Leftrightarrow \\
B[v_1, \dots, v_d]U = [v_1, \dots, v_d]UT \\
\Leftrightarrow \\
B[q_1, \dots, q_d] = [q_1, \dots, q_d]T
\end{gathered}
$$

  其中 $q_1, \dots, q_d$ 是不变子空间 $\text{Ker}(A - \lambda I_n)$ 的新的一组基.
  记 T 在 $(1,1)$ 位置上的元素为 μ，则我们有 $Bq_1 = q_1\mu$ 成立 (μ 就是 B 的一个特征值)
  注意到 $q_1 \in \text{Ker}(A - \lambda I_n)$，故 $Aq_1 = q_1\lambda$
  这样我们就找到了 A, B 的公共特征向量 q_1

- **③ 若两个划分相同的分块上三角阵可交换，则其对角分块也可交换:**
  以 2×2 分块为例:
  若 $A = \begin{bmatrix} A_{11} & A_{12} \\ 0_{k\times k} & A_{22} \end{bmatrix} \in \mathbb{C}^{n\times n}$ 和 $B = \begin{bmatrix} B_{11} & B_{12} \\ 0_{k\times k} & B_{22} \end{bmatrix} \in \mathbb{C}^{n\times n}$ 可交换，则我们有:

$$
\begin{aligned}
\begin{bmatrix} A_{11}B_{11} & A_{11}B_{12} + A_{12}B_{22} \\ 0_{k\times k} & A_{22}B_{22} \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} \\ 0_{k\times k} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ 0_{k\times k} & B_{22} \end{bmatrix} \\
&= AB \\
&= BA \\
&= \begin{bmatrix} B_{11} & B_{12} \\ 0_{k\times k} & B_{22} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0_{k\times k} & A_{22} \end{bmatrix} \\
&= \begin{bmatrix} B_{11}A_{11} & B_{11}A_{12} + B_{12}A_{22} \\ 0_{k\times k} & B_{22}A_{22} \end{bmatrix} \\
\hline
\Rightarrow &\begin{cases} A_{11}B_{11} = B_{11}A_{11} \\ A_{22}B_{22} = B_{22}A_{22} \end{cases}
\end{aligned}
$$

基于以上事实，回到 **Schur 分解定理**的证明，
我们断言关于酉矩阵 $U \in \mathbb{C}^{n\times n}$ 的所有组成成分都可对非空交换族 $\mathcal{F} \subseteq \mathbb{C}^{n\times n}$ 中的所有方阵以同样的方式选取，
因此它们可以同时酉上三角化.

Problem 7 (optional)

设 $\{A_k\}_{k=1}^{\infty}$ 是收敛于 $A \in \mathbb{C}^{n\times n}$ 的复方阵序列，$A_k = U_kT_kU_k^{\text{H}}$ 是 A_k 的 Schur 分解.
试证明: 存在 $\{U_k\}_{k=1}^{\infty}$ 的收敛子列 $\{U_{k_i}\}_{i=1}^{\infty}$ 使得 $U := \lim_{i\to\infty} U_{k_i}$ 是酉矩阵且 $U^{\text{H}}AU$ 是上三角矩阵.

- **Lemma (Matrix Analysis 定理 $2.1.7$)**
  $\mathbb{C}^{n\times n}$ 中酉矩阵的全体构成的集合和矩阵乘法构成一个**群** (group):

- **封闭性:** 对于任意酉矩阵 $U_1, U_2 \in \mathbb{C}^{n\times n}$，$U_1 U_2 \in \mathbb{C}^{n\times n}$ 也是酉矩阵
- **可结合:** 对于任意酉矩阵 $U_1, U_2, U_3 \in \mathbb{C}^{n\times n}$ 我们有 $(U_1U_2)U_3 = U_1(U_2U_3)$
- **单位元:** I_n 是一个酉矩阵，且对于任意酉矩阵 $U \in \mathbb{C}^{n\times n}$ 都有 $UI_n = U$
- **逆元:** 对于任意酉矩阵 $U \in \mathbb{C}^{n\times n}$，$U^{\mathrm{H}}$ 都是一个酉矩阵，且满足 $UU^{\mathrm{H}} = I_n$

值得注意的是，n 阶酉矩阵的集合是 $\mathbb{C}^{n\times n}$ 的**闭子集**.
也就是说，酉矩阵构成的序列如果收敛，则极限一定是一个酉矩阵.
同时我们发现酉矩阵的所有元素的模长都小于等于 1，因此 n 阶酉矩阵的集合是**有界的**.
于是 n 阶酉矩阵的集合是有限维空间 $\mathbb{C}^{n\times n}$ 的**有界闭子集**，因而是**紧集**.
也就是说，酉矩阵构成的序列一定存在收敛子列，这称为**酉矩阵的选择原理**.

Proof:
证明过程按以下步骤进行: (其中收敛性是逐元素收敛)

- ① **Schur 分解定理**保证了 $T_k := U_k^{\mathrm{H}} A_k U_k\ (\forall\ k \in \mathbb{Z}_+)$ 都是上三角阵
  且 T_k 的对角元为 A_k 的特征值，可按任意预先指定的次序排列.
- ② 根据**酉矩阵的选择原理**可知序列 $\{U_k\}$ 存在一个收敛的子列 $\{U_{k_i}\}$，其极限 $U := \lim_{i\to\infty} U_{k_i}$ 存在且是酉矩阵
- ③ 根据 $\begin{cases} \lim_{k\to\infty} A_k = A \ \Rightarrow\ \lim_{i\to\infty} A_{k_i} = A \\ \lim_{i\to\infty} U_{k_i} = U \end{cases}$ 可知 $T_{k_i} = U_{k_i}^{\mathrm{H}} A_{k_i} U_{k_i}$ 收敛于极限 $T := U^{\mathrm{H}} A U$
  因为每个 T_{k_i} 都是上三角阵，故 T 也是上三角阵.

命题得证.

Problem 8 (optional)

设 $A \in \mathbb{C}^{n\times n}$ 的极小多项式的次数为 s，定义 s 阶方阵为:

$$B := \begin{bmatrix} \mathrm{tr}\left(A^0\right) & \mathrm{tr}\left(A^1\right) & \cdots & \mathrm{tr}\left(A^{s-1}\right) \\ \mathrm{tr}\left(A^1\right) & \mathrm{tr}\left(A^2\right) & \cdots & \mathrm{tr}\left(A^{s}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{tr}\left(A^{s-1}\right) & \mathrm{tr}\left(A^s\right) & \cdots & \mathrm{tr}\left(A^{2s-2}\right) \end{bmatrix}$$

试证明: A 可对角化的充要条件是 B 非奇异.

- **Lemma 1 (Matrix Analysis 定理 $3.3.6$)**
  设复方阵 $A \in \mathbb{C}^{n\times n}$ 的 Jordan 标准型为:

$$S^{-1}AS = J = \begin{bmatrix} J^{(1)}(\lambda_1) & & & \\ & J^{(2)}(\lambda_2) & & \\ & & \ddots & \\ & & & J^{(d)}(\lambda_d) \end{bmatrix}$$

  其中 $\lambda_1, \ldots, \lambda_d$ 互不相同，$J_1(\lambda_1), \ldots, J_d(\lambda_d)$ 为对应的 Jordan 矩阵:

$$\begin{cases} J^{(1)}(\lambda_1) = J_{n_1^{(1)}}(\lambda_1) \oplus \cdots \oplus J_{n_1^{(p_1)}}(\lambda_1) \text{ where } n_1 = \sum_{i=1}^{p_1} n_1^{(i)} \quad (n_1^{(1)} \geq \cdots \geq n_1^{(p_1)} \geq 1) \\ \qquad\cdots \\ J^{(d)}(\lambda_d) = J_{n_d^{(1)}}(\lambda_d) \oplus \cdots \oplus J_{n_d^{(p_d)}}(\lambda_d) \text{ where } n_d = \sum_{i=1}^{p_d} n_d^{(i)} \quad (n_d^{(1)} \geq \cdots \geq n_d^{(p_d)} \geq 1) \end{cases}$$

  定义:

$$\begin{aligned} r_i &:= \min\{k \in \mathbb{Z}_+ : (J^{(i)}(\lambda_i) - \lambda_i I_{n_i})^k = 0_{n_i\times n_i}\} \\ &= \min\{n_i^{(1)}, \ldots, n_i^{(p_i)}\} \end{aligned} \ (i = 1, \ldots, d)$$

  即 r_i 为 A 关于特征值 λ_i 的所有 Jordan 块 $J_{n_i^{(1)}}(\lambda_i), \ldots, J_{n_i^{(p_i)}}(\lambda_i)$ 的最大的阶.
  则 $J^{(i)}(\lambda_i)$ 的极小多项式 (即最小次数的首一零化多项式) 为 $m_{J^{(i)}(\lambda_i)}(t) = (t - \lambda_i)^{r_i}$
  因此 A 的极小多项式为:

$$m_A(t) = m_J(t) = \prod_{i=1}^{d} m_{J^{(i)}(\lambda_i)}(t) = \prod_{i=1}^{d} (t - \lambda_i)^{r_i}$$

- **Lemma 2 (Matrix Analysis 推论 $3.3.8$)**
  复方阵 $A \in \mathbb{C}^{n\times n}$ 可相似对角化，当且仅当其极小多项式没有重根，即其所有特征值对应的所有 Jordan 块的阶数都是 1
  换言之，当且仅当 $\prod_{i=1}^d (t-\lambda_i)$ 可以零化 A，即 $(A-\lambda_1 I_n)\cdots(A-\lambda_d I_n) = 0_{n\times n}$

Proof:
注意到 A 的极小多项式的次数为 s
设 A 存在 d 个互不相同的特征值 (根据 **Lemma 1** 可知 $1\leq d\leq s$)，记为 $\lambda_1,\dots,\lambda_d$
设其代数重数分别为 $n_1,\dots,n_d$ (满足 $n_1+\cdots+n_d = n$)

设复方阵 A 的 Jordan 标准型为:

$$S^{-1}AS = J = \begin{bmatrix} J^{(1)}(\lambda_1) & & & \\ & J^{(2)}(\lambda_2) & & \\ & & \ddots & \\ & & & J^{(d)}(\lambda_d) \end{bmatrix}$$

其中 $J_1(\lambda_1),\dots,J_d(\lambda_d)$ 为 $\lambda_1,\dots,\lambda_d$ 对应的 Jordan 矩阵:

$$\begin{cases} J^{(1)}(\lambda_1) = J_{n_1^{(1)}}(\lambda_1)\oplus\cdots\oplus J_{n_1^{(p_1)}}(\lambda_1) \text{ where } n_1 = \sum_{i=1}^{p_1} n_1^{(i)} \quad (n_1^{(1)}\geq\cdots\geq n_1^{(p_1)}\geq 1) \\ \qquad\cdots \\ J^{(d)}(\lambda_d) = J_{n_d^{(1)}}(\lambda_d)\oplus\cdots\oplus J_{n_d^{(p_d)}}(\lambda_d) \text{ where } n_d = \sum_{i=1}^{p_d} n_d^{(i)} \quad (n_d^{(1)}\geq\cdots\geq n_d^{(p_d)}\geq 1) \end{cases}$$

则我们有:

$$\begin{aligned} \operatorname{tr}\left(A^k\right) &= \operatorname{tr}\left(\left(SJS^{-1}\right)^k\right) \\ &= \operatorname{tr}\left(SJ^kS^{-1}\right) \\ &= \operatorname{tr}\left(J^kS^{-1}S\right) \qquad (k=0,1,\dots) \\ &= \operatorname{tr}\left(J^k\right) \\ &= n_1\lambda_1^k+\cdots+n_d\lambda_d^k \end{aligned}$$

现假设标量 $\alpha_1,\dots,\alpha_s$ 使得:

$$\begin{cases} \alpha_1\operatorname{tr}\left(A^0\right)+\alpha_2\operatorname{tr}\left(A^1\right)+\cdots+\alpha_s\operatorname{tr}\left(A^{s-1}\right) = 0 \\ \alpha_1\operatorname{tr}\left(A^1\right)+\alpha_2\operatorname{tr}\left(A^2\right)+\cdots+\alpha_s\operatorname{tr}\left(A^{s}\right) = 0 \\ \qquad\cdots \\ \alpha_1\operatorname{tr}\left(A^{s-1}\right)+\alpha_2\operatorname{tr}\left(A^s\right)+\cdots+\alpha_s\operatorname{tr}\left(A^{2s-2}\right) = 0 \end{cases}$$

$$\Leftrightarrow$$

$$\begin{cases} \alpha_1(n_1+\cdots+n_d)+\alpha_2(n_1\lambda_1+\cdots+n_d\lambda_d)+\cdots+\alpha_s(n_1\lambda_1^{s-1}+\cdots+n_d\lambda_d^{s-1}) = 0 \\ \alpha_1(n_1\lambda_1+\cdots+n_d\lambda_d)+\alpha_2(n_1\lambda_1^2+\cdots+n_d\lambda_d^2)+\cdots+\alpha_s(n_1\lambda_1^{s}+\cdots+n_d\lambda_d^{s}) = 0 \\ \qquad\cdots \\ \alpha_1(n_1\lambda_1^{s-1}+\cdots+n_d\lambda_d^{s-1})+\alpha_2(n_1\lambda^s+\cdots+n_d\lambda_d^s)+\cdots+\alpha_s(n_1\lambda^{2s-2}+\cdots+n_d\lambda_d^{2s-2}) = 0 \end{cases}$$

$$\Leftrightarrow$$

$$\begin{cases} [\alpha_1,\alpha_2,\dots,\alpha_s]\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_d \\ \vdots & \vdots & & \vdots \\ \lambda_1^{s-1} & \lambda_2^{s-1} & \cdots & \lambda_d^{s-1} \end{bmatrix}\begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_d \end{bmatrix} = 0 \\ [\alpha_1,\alpha_2,\dots,\alpha_s]\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_d \\ \vdots & \vdots & & \vdots \\ \lambda_1^{s-1} & \lambda_2^{s-1} & \cdots & \lambda_d^{s-1} \end{bmatrix}\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_d \end{bmatrix}\begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_d \end{bmatrix} = 0 \\ \qquad\cdots \\ [\alpha_1,\alpha_2,\dots,\alpha_s]\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_d \\ \vdots & \vdots & & \vdots \\ \lambda_1^{s-1} & \lambda_2^{s-1} & \cdots & \lambda_d^{s-1} \end{bmatrix}\begin{bmatrix} \lambda_1^{s-1} & & & \\ & \lambda_2^{s-1} & & \\ & & \ddots & \\ & & & \lambda_d^{s-1} \end{bmatrix}\begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_d \end{bmatrix} = 0 \end{cases}$$

于是我们有:

$$\begin{cases}\alpha^{\mathrm{T}}V\gamma=0\\ \alpha^{\mathrm{T}}VD\gamma=0\\ \cdots\\ \alpha^{\mathrm{T}}VD^{s-1}\gamma=0\end{cases}\quad\text{where}\quad\begin{cases}\alpha=[\alpha_1,\ldots,\alpha_s]^{\mathrm{T}}\in\mathbb{C}^s\\ V=\begin{bmatrix}1&1&\cdots&1\\ \lambda_1&\lambda_2&\cdots&\lambda_d\\ \vdots&\vdots&&\vdots\\ \lambda_1^{s-1}&\lambda_2^{s-1}&\cdots&\lambda_d^{s-1}\end{bmatrix}\in\mathbb{C}^{s\times d}\\ D=\mathrm{diag}\{\lambda_1,\ldots,\lambda_d\}\in\mathbb{C}^{d\times d}\\ \gamma=[n_1,\cdots,n_d]^{\mathrm{T}}\in\mathbb{C}^d\end{cases}$$

因此对于任意 $s-1$ 次多项式 $p(t)$，我们都有 $(\alpha^{\mathrm{T}}V)p(D)\gamma=0$
由于 $\gamma=[n_1,\ldots,n_d]^{\mathrm{T}}$ 是非零向量，而 $p(D)$ 理论上可以是任意对角阵，
故 $\alpha=[\alpha_1,\ldots,\alpha_s]^{\mathrm{T}}$ 一定满足 $V^{\mathrm{T}}\alpha=0_d$

方阵 $B:=\begin{bmatrix}\mathrm{tr}\left(A^0\right)&\mathrm{tr}\left(A^1\right)&\cdots&\mathrm{tr}\left(A^{s-1}\right)\\ \mathrm{tr}\left(A^1\right)&\mathrm{tr}\left(A^2\right)&\cdots&\mathrm{tr}\left(A^{s}\right)\\ \vdots&\vdots&\ddots&\vdots\\ \mathrm{tr}\left(A^{s-1}\right)&\mathrm{tr}\left(A^{s}\right)&\cdots&\mathrm{tr}\left(A^{2s-2}\right)\end{bmatrix}$ 非奇异就等价于上述标量 $\alpha_1,\ldots,\alpha_s$ 一定全部为零，
即等价于 Vandermonde 矩阵 V 的 s 个 d 维行向量线性无关，
即等价于 $d=s$ (考虑到我们的假设天然限定了 $d\le s$，因此 $d>s$ 的情况被排除了)
即等价于 A 的极小多项式没有重根 (因为它是一个 s 次多项式，且 A 的互不相同的特征值 $\lambda_1,\ldots,\lambda_d$ 都是它的根)
根据 **Lemma 2** 可知即等价于 A 可相似对角化.
命题得证.

TA 的解法: 将 B 分解为 VDV^{T} (其中 V 是一个 Vandermonde 矩阵)
我们检查 D 满秩的充要条件.

$$\begin{aligned}B&=\begin{bmatrix}\mathrm{tr}\left(A^0\right)&\mathrm{tr}\left(A^1\right)&\cdots&\mathrm{tr}\left(A^{s-1}\right)\\ \mathrm{tr}\left(A^1\right)&\mathrm{tr}\left(A^2\right)&\cdots&\mathrm{tr}\left(A^{s}\right)\\ \vdots&\vdots&\ddots&\vdots\\ \mathrm{tr}\left(A^{s-1}\right)&\mathrm{tr}\left(A^{s}\right)&\cdots&\mathrm{tr}\left(A^{2s-2}\right)\end{bmatrix}\\
&=\begin{bmatrix}(n_1+\cdots+n_d)&(n_1\lambda_1+\cdots+n_d\lambda_d)&\cdots&(n_1\lambda_1^{s-1}+\cdots+n_d\lambda_d^{s-1})\\ (n_1\lambda_1+\cdots+n_d\lambda_d)&(n_1\lambda_1^2+\cdots+n_d\lambda_d^2)&\cdots&(n_1\lambda_1^{s}+\cdots+n_d\lambda_d^{s})\\ \vdots&\vdots&\ddots&\vdots\\ (n_1\lambda_1^{s-1}+\cdots+n_d\lambda_d^{s-1})&(n_1\lambda_1^{s}+\cdots+n_d\lambda_d^{s})&\cdots&(n_1\lambda_1^{2s-2}+\cdots+n_d\lambda_d^{2s-2})\end{bmatrix}\\
&=\sum_{i=1}^d n_i\begin{bmatrix}1\\ \lambda_i\\ \vdots\\ \lambda_i^{s-1}\end{bmatrix}\begin{bmatrix}1\\ \lambda_i\\ \vdots\\ \lambda_i^{s-1}\end{bmatrix}^{\mathrm{T}}\\
&=\begin{bmatrix}1&1&\cdots&1\\ \lambda_1&\lambda_2&\cdots&\lambda_d\\ \vdots&\vdots&\ddots&\vdots\\ \lambda_1^{s-1}&\lambda_2^{s-1}&\cdots&\lambda_d^{s-1}\end{bmatrix}\begin{bmatrix}n_1&&&\\ &n_2&&\\ &&\ddots&\\ &&&n_d\end{bmatrix}\begin{bmatrix}1&1&\cdots&1\\ \lambda_1&\lambda_2&\cdots&\lambda_d\\ \vdots&\vdots&\ddots&\vdots\\ \lambda_1^{s-1}&\lambda_2^{s-1}&\cdots&\lambda_d^{s-1}\end{bmatrix}^{\mathrm{T}}\\
&=VDV^{\mathrm{T}}\end{aligned}$$

其中 $V\in\mathbb{C}^{s\times d}$ 为 Vandermonde 矩阵，而 $D\in\mathbb{C}^{d\times d}$ 为对角阵.
注意到 $\lambda_1,\lambda_2,\ldots,\lambda_d$ 互不相同且 $d\le s$，故 V 一定列线性无关.
因此 $\mathrm{rank}(B)=\mathrm{rank}(D)=d$
注意到 $B\in\mathbb{C}^{s\times s}$ 非奇异当且仅当 $\mathrm{rank}(B)=s$，即当且仅当 $d=s$
这说明 A 存在 s 个互不相同的特征值.
由于 A 的极小多项式的次数为 s，故 A 的极小多项式由 s 个不同的因式构成.
根据 **Lemma 2** 可知 $d=s$ 等价于 A 可对角化.
综上所述，$B\in\mathbb{C}^{s\times s}$ 非奇异当且仅当 A 可对角化.

Problem 9

(Matrix Analysis 定理 $3.3.15$)
给定正整数 n 和关于 t 的 n 次首一多项式 $p(t)$.
设 $F \in \mathbb{C}^{n\times n}$ 为 $p(t)$ 的 Frobenius 友阵.
试证明 F 的特征多项式和极小多项式都是 $p(t)$.

Solution:
给定 n 次首一多项式 $p(t) := t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$,
其对应的 Frobenius 友阵为:

$$F := \begin{bmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & \vdots \\ & & \ddots & 0 & -a_{n-2} \\ & & & 1 & -a_{n-1} \end{bmatrix}$$

记 e_i 为第 i 个元素为 1，其余元素为 0 的 n 维列向量.
假设存在一个次数为 $s < n$ 的首一多项式 $m(t) = t^s + b_{s-1}t^{s-1} + \cdots + b_1 t + b_0$ 满足 $m(F) = 0_{n\times n}$,
则我们有:

$$\begin{aligned} m(F)e_1 &= F^s e_1 + b_{s-1}F^{s-1}e_1 + \cdots + b_1 F e_1 + b_0 e_1 \\ &= e_{s+1} + b_{s-1}e_s + \cdots + b_1 e_2 + b_0 e_1 \\ &= 0_n. \end{aligned}$$

但 $e_{s+1}, e_s, \ldots, e_2, e_1$ 是线性无关的，上式显然与之矛盾.
因此任意一个次数为 $s < n$ 的首一多项式 $m(t)$ 都不能零化 F.
故 $p(t)$ 就是 F 的极小多项式.
由于 $p(t)$ 次数为 n，故它也是 F 的特征多项式.

The End