

4.1

- 4.2 Let V be the subspace $V = \text{span}(1, x, x^2)$ in the inner product space $L^2([0, 1], \mathbb{R})$. Let D be the derivative operator $D : V \rightarrow V$. Find all the eigenvalues and eigenspaces of D . What are their algebraic and geometric multiplicities?

First of all $D = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

Since D is an upper right triangular matrix its eigenvalues are on its diagonal. So we have $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

Now we find the eigenvectors: $D\vec{x} = \lambda\vec{x} \Rightarrow \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ 2x_3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

So x_1 is arbitrary, $x_2 = x_3 = 0$, and an eigenvector for $\lambda_1 = 0$ is $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

And so it follows that the eigenspace of D is $\text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$

The algebraic multiplicity of λ_1 is 3 and its geometric multiplicity is 1. ♣

4.3

- 4.4 Recall that a matrix $A \in M_n(\mathbb{F})$ is called Hermitian if $A^H = A$ and skew-Hermitian if $A^H = -A$. Using exercise 4.3, prove the following:

(i) A Hermitian 2x2 matrix has only real eigenvalues.

Proof:

First we note that since $A^H = \overline{A}^T = A$ it follows that the entries of A are real ($a_{ij} = \overline{a_{ij}}$ if and only if a_{ij} is real). Since A is 2x2 we can write as $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$

where $a, b, c, d \in \mathbb{R}$ and $A = A^H$ implies $c = d$. Using our characteristic polynomial from 4.3 we get $p(\lambda) = \lambda^2 - \text{tr}(A)\lambda + \det(A) = \lambda^2 - (a+d)\lambda + (ad-bc) = \lambda^2 - (a+d)\lambda + ad - b^2$. Using the quadratic formula we have $\lambda = \frac{(a+d) \pm \sqrt{(a+d)^2 - 4(1)(ad-b^2)}}{2} = \frac{a+d}{2} \pm \frac{1}{2}\sqrt{a^2 + 2ad + d^2 - 4ad + 4b^2} = \frac{a+d}{2} \pm \frac{1}{2}\sqrt{a^2 - 2ad + d^2 + 4b^2} = \frac{a+d}{2} \pm \frac{1}{2}\sqrt{(a-d)^2 + 4b^2}$ since $a, b, d \in \mathbb{R}$ then $(a-d)^2 + 4b^2$ is real as well and is greater than or equal to zero and so the eigenvalues are real. ♣

(ii) A skew-Hermitian 2x2 matrix has only imaginary eigenvalues.

Proof:

$A^H = -A$ implies $a_{11} = -\bar{a}_{11}$ and $a_{22} = -\bar{a}_{22}$ which implies the diagonal entries of A are pure imaginary numbers. Also we have $a_{12} = -\bar{a}_{21}$. If we use x and y to denote the real part and imaginary part of a_{21} , respectively. Then we have $a_{12} = -\overline{-x + iy} = (-1)(x - iy) = -x + iy$. So we can write the matrix A as $A = \begin{bmatrix} ai & -x + iy \\ x + iy & di \end{bmatrix}$ Where $a, d, x, y \in \mathbb{R}$

And so using the characteristic polynomial from 4.3 we get $p(y) = \lambda^2 - \text{tr}(A)\lambda + \det(A) = \lambda^2 - (ai + di)\lambda - (-x + iy)(x + iy) = 0$

At this point we use the quadratic formula to solve for the zeros of the polynomial.

$$\begin{aligned} \lambda &= \frac{1}{2}(ai + di) \pm \frac{1}{2}\sqrt{(ai + di)^2 - 4(1)(aidi - (-x + iy)(x + iy))} \\ &= \frac{1}{2}(ai + di) \pm \frac{1}{2}\sqrt{(ai)^2 + 2aidi + (di)^2 - 4(-ad + x^2 + y^2)} \\ &= \frac{1}{2}(ai + di) \pm \frac{1}{2}\sqrt{(ai)^2 - 2ad + (di)^2 + 4ad - 4x^2 - 4y^2} \\ &= \frac{1}{2}(ai + di) \pm \frac{1}{2}\sqrt{-a^2 + 2ad - d^2 - 4x^2 - 4y^2} \\ &= \frac{1}{2}(ai + di) \pm \frac{1}{2}\sqrt{(-1)(a^2 - 2ad + d^2) - 4x^2 - 4y^2} \\ &= \frac{i}{2}(a + d) \pm \frac{1}{2}\sqrt{(-1)(a - d)^2 - 4x^2 - 4y^2} \\ &= \frac{i}{2}(a + d) \pm \frac{1}{2}\sqrt{(-1)((a - d)^2 + 4x^2 + 4y^2)} \\ &= i \left(\frac{1}{2}(a + d) \pm \frac{1}{2}\sqrt{(a - d)^2 + 4x^2 + 4y^2} \right) \end{aligned}$$

And since $a, d, x, y \in \mathbb{R}$ then $x^2 \geq 0, y^2 \geq 0, (a - d)^2 \geq 0$ So the contents of the square root is a real number and $\frac{a+b}{2}$ is a real number, and i times the sum of two real numbers is a pure imaginary number. In other words, the eigenvalues of the matrix are pure imaginary. ♣

4.6 Let $A = \begin{bmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{1}{5} & \frac{3}{5} \end{bmatrix}$

Compute the transition matrix P such that $P^{-1}AP$ is diagonal.

$$\det(A - \lambda I) = \det \left(\begin{bmatrix} \frac{4}{5} - \lambda & \frac{2}{5} \\ \frac{1}{5} & \frac{3}{5} - \lambda \end{bmatrix} \right) = \left(\frac{4}{5} - \lambda \right) \left(\frac{3}{5} - \lambda \right) - \frac{2}{5} \frac{1}{5} = 0$$

$$\lambda^2 - \frac{7}{5}\lambda + \frac{10}{25} = 0$$

At this point we use the quadratic formula to solve for the eigenvalues.

$$\begin{aligned} \lambda &= \frac{\frac{7}{5} \pm \sqrt{\left(\frac{7}{5}\right)^2 - 4\left(\frac{10}{25}\right)}}{2} \\ &= \frac{7}{10} \pm \frac{1}{2}\sqrt{\frac{49}{25} - \frac{40}{25}} \\ &= \frac{7}{10} \pm \frac{1}{2}\sqrt{\frac{9}{25}} \\ &= \frac{7}{10} \pm \frac{3}{10} \\ &\Rightarrow \lambda_1 = \frac{2}{5} \text{ and } \lambda_2 = 1 \end{aligned}$$

Now we find eigenvectors that correspond to the eigenvalues.

$$\begin{bmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{1}{5} & \frac{3}{5} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \frac{2}{5} \begin{bmatrix} x \\ y \end{bmatrix} \text{ Gives } \frac{4}{5}x + \frac{2}{5}y = \frac{2}{5}x \Rightarrow x = -y \text{ and so the eigenvector for } \lambda_1 = \frac{2}{5}$$

is $\begin{bmatrix} 1 \\ \sqrt{2} \\ -1 \\ \sqrt{2} \end{bmatrix}$

Now we find the second eigenvector.

$$\begin{bmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{1}{5} & \frac{3}{5} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \text{ Gives } \frac{4}{5}x + \frac{2}{5}y = x \Rightarrow x = 2y \text{ so the eigenvector for } \lambda_2 = 1 \text{ is } \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$$

$$\Rightarrow P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{2}{\sqrt{5}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{5}} \end{bmatrix} \Rightarrow P^{-1} = \begin{bmatrix} \frac{\sqrt{2}}{3} & \frac{-2\sqrt{2}}{3} \\ \frac{\sqrt{5}}{3} & \frac{\sqrt{5}}{3} \end{bmatrix} D = \begin{bmatrix} \frac{2}{5} & 0 \\ 0 & 1 \end{bmatrix}$$

4.7

4.8 Let A be the matrix in Exercise 4.6 above.

(i) Compute $\lim_{n \rightarrow \infty} A^n$. That is find a matrix B such that for any $\epsilon > 0$ there exists an $N > 0$ with $\|A^k - B\|_1 < \epsilon$ whenever $k > N$.

$$\lim_{n \rightarrow \infty} A^n = \lim_{n \rightarrow \infty} (PDP^{-1})^n = \lim_{n \rightarrow \infty} PD^nP^{-1} = \lim_{n \rightarrow \infty} P \begin{bmatrix} (\frac{2}{5})^n & 0 \\ 0 & 1 \end{bmatrix} P^{-1} = P \begin{bmatrix} \lim_{n \rightarrow \infty} (\frac{2}{5})^n & 0 \\ 0 & 1 \end{bmatrix} P^{-1} =$$

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

First we show that A^n converges to B using the $\|\cdot\|_1$ norm.

$$\text{We have } A^n = P \begin{bmatrix} (\frac{2}{5})^n & 0 \\ 0 & 1 \end{bmatrix} P^{-1} \text{ And so } A^n - B = P \begin{bmatrix} (\frac{2}{5})^n & 0 \\ 0 & 1 \end{bmatrix} P^{-1} - P \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} P^{-1} =$$


$$P \left(\begin{bmatrix} (\frac{2}{5})^n & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right) P^{-1} = P \begin{bmatrix} (\frac{2}{5})^n & 0 \\ 0 & 0 \end{bmatrix} P^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{2}{\sqrt{5}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} (\frac{2}{5})^n & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{3} & \frac{-2\sqrt{2}}{3} \\ \frac{\sqrt{5}}{3} & \frac{\sqrt{5}}{3} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sqrt{2}}(\frac{2}{5})^n & 0 \\ \frac{-1}{\sqrt{2}}(\frac{2}{5})^n & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{3} & \frac{-2\sqrt{2}}{3} \\ \frac{\sqrt{5}}{3} & \frac{\sqrt{5}}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3}(\frac{2}{5})^n & \frac{-2}{3}(\frac{2}{5})^n \\ \frac{-1}{3}(\frac{2}{5})^n & \frac{2}{3}(\frac{2}{5})^n \end{bmatrix} = \frac{1}{3}(\frac{2}{5})^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}$$

$$\text{So, } \|A^n - B\|_1 = \|\frac{1}{3}(\frac{2}{5})^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}\|_1 = |\frac{2}{3}(\frac{2}{5})^n| + |\frac{-2}{3}(\frac{2}{5})^n| + |\frac{1}{3}(\frac{2}{5})^n| + |\frac{-1}{3}(\frac{2}{5})^n| =$$

$$6|\frac{1}{3}(\frac{2}{5})^n| = 2|(\frac{2}{5})^n|$$

For each $\epsilon > 0$ we can choose an $N \in \mathbb{N}$ large enough so that $n > N \Rightarrow 2|(\frac{2}{5})^n| < \epsilon$

Hence for each $\epsilon > 0$ we can choose an $N \in \mathbb{N}$ large enough so that $n > N \Rightarrow \|A^n - B\|_1 < \epsilon$. The proof is finished. 

We now show that $\|A^n - B\|_\infty$ converges to zero.

Proof:

Much of the initial calculations and logic in this sub-proof are identical to the preceding proof. So we pick up from the spot in the last proof where we started using the ∞ -norm.

$$\text{That is we have } \|A^n - B\|_\infty = \|\frac{1}{3}(\frac{2}{5})^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}\|_\infty = |\frac{2}{3}(\frac{2}{5})^n|$$

For each $\epsilon > 0$ we can choose an $N \in \mathbb{N}$ large enough so that $n > N \Rightarrow |\frac{2}{3}(\frac{2}{5})^n| < \epsilon$

Hence for each $\epsilon > 0$ we can choose an $N \in \mathbb{N}$ large enough so that $n > N \Rightarrow \|A^n - B\|_\infty < \epsilon$. The proof is finished. ♣

We now prove the convergence of A^n to B using the 2-norm.

Proof:

$$\begin{aligned} \text{We have } \|A^n - B\|_2 &= \left\| \frac{1}{3} \left(\frac{2}{5} \right)^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} \right\|_2 \\ &= \left(\left| \frac{1}{3} \left(\frac{2}{5} \right)^n (1) \right|^2 + \left| \frac{1}{3} \left(\frac{2}{5} \right)^n (-1) \right|^2 + \left| \frac{2}{3} \left(\frac{2}{5} \right)^n (1) \right|^2 + \left| \frac{-2}{3} \left(\frac{2}{5} \right)^n (1) \right|^2 \right)^{\frac{1}{2}} \\ &= (10 \left| \frac{1}{3} \left(\frac{2}{5} \right)^n (1) \right|^2)^{\frac{1}{2}} = \sqrt{10} \left| \frac{1}{3} \left(\frac{2}{5} \right)^n (1) \right| \end{aligned}$$

For each $\epsilon > 0$ we can choose an $N \in \mathbb{N}$ large enough so that $n > N \Rightarrow \sqrt{10} \left| \frac{1}{3} \left(\frac{2}{5} \right)^n (1) \right| < \epsilon$

Hence for each $\epsilon > 0$ we can choose an $N \in \mathbb{N}$ large enough so that $n > N \Rightarrow \|A^n - B\|_2 < \epsilon$. The proof is finished. ♣

We now prove the convergence of A^n to B using the Frobenius norm.

$$\begin{aligned} \text{We have } \|A^n - B\|_{Frobenius} &= \left\| \frac{1}{3} \left(\frac{2}{5} \right)^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} \right\|_{Frobenius} \\ &= \sqrt{\text{tr} \left(\left(\frac{1}{3} \left(\frac{2}{5} \right)^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} \right) \left(\frac{1}{3} \left(\frac{2}{5} \right)^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} \right)^H \right)} \\ &= \sqrt{\text{tr} \left(\frac{1}{3} \left(\frac{2}{5} \right)^n \frac{1}{3} \left(\frac{2}{5} \right)^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix} \right)} \\ &= \sqrt{\frac{1}{9} \left(\frac{2}{5} \right)^{2n} \text{tr} \left(\begin{bmatrix} 5 & -5 \\ -5 & 5 \end{bmatrix} \right)} = \frac{1}{3} \left(\frac{2}{5} \right)^n \sqrt{10} \end{aligned}$$

For each $\epsilon > 0$ we can choose an $N \in \mathbb{N}$ large enough so that $\frac{1}{3} \left(\frac{2}{5} \right)^n \sqrt{10} < \epsilon$ when $n > n$. ♣

It seems that whether it converged or not was independent of which norm we used.

(ii)

4.9

4.10 For any semi-simple matrix $A \in M_n(\mathbb{F})$, let $p(x) = a_1x + a_2x^2 + \dots + a_nx^n$ be the characteristic polynomial of A. Prove that $p(A) = 0_{n \times n}$

Proof:

We make use of Theorem 4.2.6 which states that a matrix is diagonalizable if and only if it is semi-simple. So the matrix A can be written as $A = PDP^{-1}$ Where D is a diagonal matrix whose entries along the diagonal are the the n eigenvalues of A. So

$$\begin{aligned}
p(A) &= a_1 A + a_2 A^2 + \dots + a_n A^n = a_1 (PDP^{-1}) + a_2 (PDP^{-1})^2 + \dots + a_n (PDP^{-1})^n \\
&\text{which by proposition 4.2.9 is equal to} \\
&a_1 PDP^{-1} + a_2 PD^2P^{-1} + \dots + a_n PD^nP^{-1} \\
&= a_1 P \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} P^{-1} + a_2 P \begin{pmatrix} \lambda_1^2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n^2 \end{pmatrix} P^{-1} + \dots + a_n P \begin{pmatrix} \lambda_1^n & & 0 \\ & \ddots & \\ 0 & & \lambda_n^n \end{pmatrix} P^{-1} \\
&= P \left(a_1 \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} + a_2 \begin{pmatrix} \lambda_1^2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n^2 \end{pmatrix} + \dots + a_n \begin{pmatrix} \lambda_1^n & & 0 \\ & \ddots & \\ 0 & & \lambda_n^n \end{pmatrix} \right) P^{-1} \\
&= P \begin{pmatrix} a_1 \lambda_1 + \dots + a_n \lambda_1^n & & 0 \\ & \ddots & \\ 0 & & a_1 \lambda_n + \dots + a_n \lambda_n^n \end{pmatrix} P^{-1} \\
&\text{and } p(\lambda_i) = 0 \text{ for } i = 1, \dots, n \text{ so we have} \\
p(A) &= P \begin{pmatrix} a_1 \lambda_1 + \dots + a_n \lambda_1^n & & 0 \\ & \ddots & \\ 0 & & a_1 \lambda_n + \dots + a_n \lambda_n^n \end{pmatrix} P^{-1} \\
&= P \begin{pmatrix} 0 & & 0 \\ & \ddots & \\ 0 & & 0 \end{pmatrix} P^{-1} = 0_{n \times n} \quad \clubsuit
\end{aligned}$$

4.11

4.12 Give an example to show that the power method may fail if there are two distinct dominant eigenvalues of the same magnitude. What happens if the dominant eigenvalue has two linearly independent eigenvectors?

Consider the following matrix:

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

Its eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = -2$, with corresponding eigenvectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 4 \end{bmatrix}$

Note that the moduli of the two eigenvalues are equal, so one of the hypotheses of the power method is not satisfied and so the power method may have problems.

Using the power method with the initial guess of $[1, 1]^T$ we get:

$$\mathbf{x}_1 = A\mathbf{x}_0 = \begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

$$\mathbf{x}_2 = A\mathbf{x}_1 = \begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 5 \\ -2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{x}_3 = A\mathbf{x}_2 = \begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 20 \\ -8 \end{bmatrix} = 4 \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

$$\mathbf{x}_4 = A\mathbf{x}_3 = \begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 20 \\ -8 \end{bmatrix} = \begin{bmatrix} 16 \\ 16 \end{bmatrix} = 16 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

\vdots

So we see that in this case the power method does not converge since it alternates between the two eigenvectors.

Now we consider a case where one dominant eigenvalue has two eigenvectors:

$$\begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$$

Which has only one eigenvalue, 5. It also has two eigenvectors from this, which are:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Starting with our same initial guess, we get:

$$\mathbf{x}_1 = A\mathbf{x}_0 = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

$$\mathbf{x}_2 = A\mathbf{x}_1 = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 5 \\ 5 \end{bmatrix} = 25 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

\vdots

it only stays on the one vector. ♣

4.13

4.15

4.16 Given an nxn matrix A, we define the Rayleigh quotient:

$$p(x) := \frac{\langle x, Ax \rangle}{\|x\|^2}$$

Show that the Rayleigh Quotient can only take on real values for Hermitian matrices and only imaginary values for skew-Hermitian matrices.

Proof:

Assume that A is hermitian. Then $\langle x, Ax \rangle = x^H Ax = x^H A^H x = (x^H Ax)^H = \overline{\langle x, Ax \rangle}$. The fact that this number is equal to its conjugate implies that it is real (*let* $z \in \mathbb{C} = x + iy$ for $x, y \in \mathbb{R}$ then $\bar{z} = \overline{x + iy} = x - iy$ and $x + iy = x - iy \Rightarrow y = 0$). The denominator in the Rayleigh quotient is a norm, which by definition is in the reals, and a real number divided by a real number is a real number and the proof is complete. ♣

Proof:

Let A be a skew-Hermitian. Then $\langle x, Ax \rangle = x^H Ax = -x^H A^H x = -(x^H Ax)^H = -\overline{\langle x, Ax \rangle}$. The fact that this number is equal to the negative of its conjugate implies that it is pure imaginary (let $z \in \mathbb{C} = x + iy$ for $x, y \in \mathbb{R}$ then $\bar{z} = -(x + iy) = -x + iy$ and $x + iy = x - iy \Rightarrow x = 0$). So this pure imaginary number can be written as bi where $b \in \mathbb{R}$. The denominator in the Rayleigh quotient is a norm, which by definition is in the reals, call it $a \in \mathbb{R}$. So the Rayleigh quotient $= \frac{bi}{a} = i \frac{b}{a}$ and $\frac{b}{a} \in \mathbb{R}$ so it follows that the Rayleigh quotient is a pure imaginary number. \clubsuit

4.19

4.20 Assume $A, B \geq 0$. Prove that $0 \leq \text{tr}(AB) \leq \text{tr}(A)\text{tr}(B)$ and use this to prove that $\|\cdot\|_F$ is a matrix norm.

Proof:

In this proof we will use Proposition 4.4.5 and we also will make use of the fact that in general $\text{tr}(CM) = \text{tr}(MC)$ for two square matrices M and N . So we can write $A = UD_1U^H$ and $B = VD_2V^H$ where U and V are orthonormal matrices and D_1 and D_2 are diagonal matrices and where d_{ij} and δ_{ij} are the (i, j) entries of D_1 and D_2 respectively. So it follows that $\text{tr}(A)\text{tr}(B) = \text{tr}(UD_1U^H)\text{tr}(VD_2V^H) = \text{tr}(U^H U D_1)\text{tr}(V^H V D_2) = \text{tr}(D_1)\text{tr}(D_2) = \sum_{i=1}^n d_{ii} \sum_{i=1}^n \delta_{ii} \geq \sum_{i=1}^n d_{ii} \delta_{ii} = \text{tr}(D_1 D_2)$.

At this point we need to show that $\text{tr}(D_1 D_2) \geq \text{tr}(UD_1U^H V D_2 V^H)$

$$\begin{aligned} A = UD_1U^H &= \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \dots & u_{nn} \end{bmatrix} \begin{bmatrix} d_{11} & & & 0 \\ & d_{22} & & \\ & & \ddots & \\ 0 & & & d_{nn} \end{bmatrix} = \begin{bmatrix} u_{11}d_{11} & u_{12}d_{12} & \dots & u_{1n}d_{1n} \\ u_{21}d_{11} & u_{22}d_{12} & \dots & u_{2n}d_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1}d_{11} & u_{n2}d_{12} & \dots & u_{nn}d_{1n} \end{bmatrix} U^H \\ &= \begin{bmatrix} u_{11}d_{11} & u_{12}d_{12} & \dots & u_{1n}d_{1n} \\ u_{21}d_{11} & u_{22}d_{12} & \dots & u_{2n}d_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1}d_{11} & u_{n2}d_{12} & \dots & u_{nn}d_{1n} \end{bmatrix} \begin{bmatrix} \overline{u_{11}} & \overline{u_{21}} & \dots & \overline{u_{n1}} \\ \overline{u_{12}} & \overline{u_{22}} & \dots & \overline{u_{n2}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{u_{1n}} & \overline{u_{2n}} & \dots & \overline{u_{nn}} \end{bmatrix} \end{aligned}$$

And so it follows that

$$a_{11} = d_{11}u_{11}\overline{u_{11}} + d_{22}u_{12}\overline{u_{12}} + d_{33}u_{13}\overline{u_{13}} + \dots + d_{nn}u_{1n}\overline{u_{1n}}$$

$$a_{22} = d_{11}u_{21}\overline{u_{21}} + d_{22}u_{22}\overline{u_{22}} + d_{33}u_{23}\overline{u_{23}} + \dots + d_{nn}u_{2n}\overline{u_{2n}}$$

$$a_{ii} = d_{11}u_{i1}\overline{u_{i1}} + d_{22}u_{i2}\overline{u_{i2}} + d_{33}u_{i3}\overline{u_{i3}} + \dots + d_{nn}u_{in}\overline{u_{in}}$$

$$a_{nn} = d_{11}u_{n1}\overline{u_{n1}} + d_{22}u_{n2}\overline{u_{n2}} + d_{33}u_{n3}\overline{u_{n3}} + \dots + d_{nn}u_{nn}\overline{u_{nn}}$$

So $\text{tr}(A) =$

$$\begin{aligned}
& d_{11}(u_{11}\overline{u_{11}} + u_{21}\overline{u_{21}} + \dots + u_{i1}\overline{u_{i1}} + \dots u_{n1}\overline{u_{n1}}) \\
& + d_{22}(u_{12}\overline{u_{12}} + u_{22}\overline{u_{22}} + \dots + u_{i2}\overline{u_{i2}} + \dots u_{n2}\overline{u_{n2}}) \\
& \vdots \\
& + d_{nn}(u_{1n}\overline{u_{1n}} + u_{2n}\overline{u_{2n}} + \dots + u_{in}\overline{u_{in}} + \dots u_{nn}\overline{u_{nn}})
\end{aligned}$$

And since U is orthonormal we have $UU^H = I$ and $U^H U = I$

WE NEED HELP FIGURING THIS ONE OUT



Now we show that the Frobenius is a norm.

Proof:

Positivity:

Proposition 4.4.7 shows that AA^H is positive semi-definite. $\|A\|_F = \sqrt{\text{tr}(AA^H)} = \sqrt{\text{tr}(AA^H)}$. I is also positive semi-definite. So by our proof of the first part of this problem we know that the trace of the product of two positive semidefinite matrices is nonnegative.

Scale Preservation:

$$\|\alpha A\|_F = \sqrt{\text{tr}(\alpha A \alpha A^H)} = \sqrt{\alpha^2 \text{tr}(AA^H)} = \sqrt{\alpha^2} \sqrt{\text{tr}(AA^H)} = |\alpha| \|A\|_F.$$

Triangle Inequality:

$$\begin{aligned}
\|A+B\|_F &= \sqrt{\text{tr}((A+B)(A+B)^H)} = \sqrt{\text{tr}((A+B)(A^H+B^H))} = \sqrt{\text{tr}(AA^H + AB^H + BA^H + BB^H)} \\
&= \sqrt{\text{tr}(AA^H) + \text{tr}(BB^H) + 2\text{tr}(AB^H)}
\end{aligned}$$

$$\text{So } \|A+B\|^2 = \text{tr}(AA^H) + \text{tr}(BB^H) + 2\text{tr}(AB^H)$$

$$\|A\|_F + \|B\|_F = \sqrt{\text{tr}(AA^H)} + \sqrt{\text{tr}(BB^H)}$$

$$(\|A\|_F + \|B\|_F)^2 = \|A\|_F^2 + 2\|A\|_F\|B\|_F + \|B\|_F^2 = \text{tr}(AA^H) + 2\sqrt{\text{tr}(AA^H)\text{tr}(BB^H)} + \text{tr}(BB^H)$$

In order to show that the triangle inequality holds we only need to show that $2\sqrt{\text{tr}(AA^H)\text{tr}(BB^H)} \geq 2\text{tr}(AB^H)$, or equivalently that $\text{tr}(AA^H)\text{tr}(BB^H) \geq \text{tr}(AB^H)^2$

$$\begin{aligned}
\text{tr}(AA^H)\text{tr}(BB^H) &= \|A\|_F^2 \|B\|_F^2 \text{ and by Cauchy-Schwarz we have} \\
\|A\|_F^2 \|B\|_F^2 &\geq \langle A, B \rangle^2 = \text{tr}(AB^H)^2
\end{aligned}$$

So the Triangle Inequality holds.



4.21

4.22

4.24 Assume $A \in M_{m \times n}(\mathbb{F})$ is of rank r . Prove the following:

(i). $\|A\|_F = (\sigma_1^2 + \dots + \sigma_r^2)^{\frac{1}{2}}$, where $\sigma_1, \dots, \sigma_r$ are the singular values of A .

Proof:

We make use of theorem 4.4.8 The Singular Value decomposition. Since A has rank r we can write A as $A = U\Sigma V^H$ Where U and V are orthonormal matrices and Σ is a diagonal matrix whose entries along the diagonal are the singular values of A . So we have $\|A\|_F = \sqrt{\text{tr}(AA^H)} = \sqrt{\text{tr}(U\Sigma V^H(U\Sigma V^H)^H)} = \sqrt{\text{tr}(U\Sigma V^H V \Sigma^H U^H)} = \sqrt{\text{tr}(U\Sigma \Sigma^H U^H)}$ And in general $\text{tr}(BC) = \text{tr}(CB)$ so if $B = U\Sigma \Sigma^H$ and $C = U^H$ then $\sqrt{\text{tr}(U\Sigma \Sigma^H U^H)} = \sqrt{\text{tr}(U^H U \Sigma \Sigma^H)} = \sqrt{\text{tr}(\Sigma \Sigma^H)}$ and since Σ is diagonal then $\Sigma^H = \Sigma$ so $\sqrt{\text{tr}(\Sigma \Sigma^H)} = \sqrt{\text{tr}(\Sigma^2)} = \sqrt{\sigma_1^2 + \dots + \sigma_r^2}$ which is the desired result. ♣

(ii) If $U \in M_{m \times n}(\mathbb{F})$ and $V \in M_{m \times n}(\mathbb{F})$ are orthonormal matrices then $\|UAV\|_F = \|A\|_F$

Proof:

$$\|UAV\|_F = (\text{tr}(UAV(UAV)^H))^{\frac{1}{2}} = (\text{tr}(UAVV^H A^H U^H))^{\frac{1}{2}} = (\text{tr}(U A A^H U^H))^{\frac{1}{2}} = (U^H U A A^H)^{\frac{1}{2}} = (\text{tr}(A A^H))^{\frac{1}{2}} = \|A\|_F \quad \clubsuit$$

4.26 Let A be an $n \times n$ matrix. Prove that $|\det(A)| = \prod_{i=1}^n \sigma_i$, or in other words, the modulus of the determinant is the product of the singular values.

Proof:

We make use of theorem 4.4.8 The Singular Value decomposition. We can write A as $A = U\Sigma V^H$ Where U and V are orthonormal matrices and Σ is a diagonal matrix whose entries along the diagonal are the singular values of A $\sigma_1, \sigma_2, \dots, \sigma_n$ (if A has rank $r < n$ then the diagonal entries $\sigma_{n-r+1} = \sigma_{n-r+2} = \dots = \sigma_n$ of Σ are zeros. So $|\det(A)| = |\det(U\Sigma V^H)| = |\det(U)| |\det(\Sigma)| |\det(V^H)| = (1) |(\sigma_1)(\sigma_2) \dots (\sigma_n)| (1) = |\prod_{i=1}^n \sigma_i|$ and since the singular values are all nonnegative $|\prod_{i=1}^n \sigma_i| = \prod_{i=1}^n \sigma_i$. and the proof is complete. ♣

4.27

4.28 Let A be the matrix in Exercise 4.21. Find the Moore-Penrose inverse A^\dagger of A . Compare $A^\dagger A$ to $A^H A$.

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

The good news is, the work for this problem is mostly done from problem 21, we know that:

$$U = \begin{bmatrix} \frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} \\ 0 & 0 & 1 \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sqrt{15} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad V = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

So, we're clipping the Σ , but there isn't anything else to really clip, so it's going to

end up being the transpose. So now using:

$$\begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{15}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} \\ 0 & 0 & 1 \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \end{bmatrix}^H$$

The result is:

$$A^\dagger = \begin{bmatrix} \frac{1}{15} & 0 & \frac{2}{15} \\ \frac{1}{15} & 0 & \frac{2}{15} \\ \frac{1}{15} & 0 & \frac{2}{15} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\text{Then } A^\dagger A \text{ is: } \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } A^H A \text{ is: } \begin{bmatrix} 5 & 5 & 5 & 0 \\ 5 & 5 & 5 & 0 \\ 5 & 5 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

These two are very similar, they just differ by a scalar multiple.

$$4.34 \text{ Consider the matrix } V_n = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \dots & x_{n-1}^n \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{bmatrix}$$

Prove that $\det(V_n) = \prod_{i < j} (x_j - x_i)$

Proof:

To begin with we know that $\det(V_n) = \det(V_n^T)$

$$\det(V_n^T) = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ x_0 & x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_0^{n-1} & x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \\ x_0^n & x_1^n & x_2^n & \dots & x_n^n \end{vmatrix}$$

Adding or subtracting a multiple of a row to another doesn't affect the determinant.

We row reduce the nth row by subtracting x_0 times the nth row to the (n+1)th row.

$$\begin{aligned} \det(V_n^T) &= \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ x_0 & x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_0^{n-1} & x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \\ 0 & x_1^n - x_0 x_1^{n-1} & x_2^n - x_0 x_2^{n-1} & \dots & x_n^n - x_0 x_n^{n-1} \end{vmatrix} \\ &= \det(V_n^T) = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ x_0 & x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_0^{n-1} & x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \\ 0 & (x_1 - x_0)x_1^{n-1} & (x_2 - x_0)x_2^{n-1} & \dots & (x_n - x_0)x_n^{n-1} \end{vmatrix} \end{aligned}$$

We subtract x_0 times the $(n-1)$ th row to the n th row.

$$\begin{aligned}
\det(V_n^T) &= \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ x_0 & x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_0^{n-2} & x_1^{n-2} & x_2^{n-2} & \dots & x_n^{n-2} \\ 0 & x_1^{n-1} - x_0 x_1^{n-2} & x_2^{n-1} - x_0 x_2^{n-2} & \dots & x_n^{n-1} - x_0 x_n^{n-2} \\ 0 & (x_1 - x_0)x_1^{n-1} & (x_2 - x_0)x_2^{n-1} & \dots & (x_n - x_0)x_n^{n-1} \end{vmatrix} \\
&= \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ x_0 & x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_0^{n-2} & x_1^{n-2} & x_2^{n-2} & \dots & x_n^{n-2} \\ 0 & (x_1 - x_0)x_1^{n-2} & (x_2 - x_0)x_2^{n-2} & \dots & (x_n - x_0)x_n^{n-2} \\ 0 & (x_1 - x_0)x_1^{n-1} & (x_2 - x_0)x_2^{n-1} & \dots & (x_n - x_0)x_n^{n-1} \end{vmatrix}
\end{aligned}$$

We continue subtracting x_0 times the $(p-1)$ th row from the p th row. Which gives

$$\begin{aligned}
\det(V_n^T) &= \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & (x_1 - x_0) & (x_2 - x_0) & \dots & (x_n - x_0) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & (x_1 - x_0)x_1^{n-3} & (x_1 - x_0)x_2^{n-3} & \dots & (x_1 - x_0)x_n^{n-3} \\ 0 & (x_1 - x_0)x_1^{n-2} & (x_2 - x_0)x_2^{n-2} & \dots & (x_n - x_0)x_n^{n-2} \\ 0 & (x_1 - x_0)x_1^{n-1} & (x_2 - x_0)x_2^{n-1} & \dots & (x_n - x_0)x_n^{n-1} \end{vmatrix} \\
&= \begin{vmatrix} (x_1 - x_0) & (x_2 - x_0) & \dots & (x_n - x_0) \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-3}(x_1 - x_0) & x_2^{n-3}(x_1 - x_0) & \dots & x_n^{n-3}(x_1 - x_0) \\ x_1^{n-2}(x_1 - x_0) & x_2^{n-2}(x_2 - x_0) & \dots & x_n^{n-2}(x_n - x_0) \\ x_1^{n-1}(x_1 - x_0) & x_2^{n-1}(x_2 - x_0) & \dots & x_n^{n-1}(x_n - x_0) \end{vmatrix}
\end{aligned}$$

We can transpose this matrix and divide the first row by $(x_1 - x_0)$, divide the second row by $(x_2 - x_0)$, ..., and divide the n th row by $(x_n - x_0)$.

$$\text{So } \det(V_n) = (x_1 - x_0)(x_2 - x_0)\dots(x_n - x_0) \begin{vmatrix} 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-3} & x_2^{n-3} & \dots & x_n^{n-3} \\ x_1^{n-2} & x_2^{n-2} & \dots & x_n^{n-2} \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix}$$

However this new smaller determinant on the right is equal to the original larger determinant we set out to solve but simply with new indices which implies

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-3} & x_2^{n-3} & \dots & x_n^{n-3} \\ x_1^{n-2} & x_2^{n-2} & \dots & x_n^{n-2} \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = (x_2 - x_1)(x_3 - x_1)\dots(x_{n-1} - x_1) \begin{vmatrix} 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ x_2^{n-4} & x_3^{n-4} & \dots & x_{n-1}^{n-4} \\ x_2^{n-3} & x_3^{n-3} & \dots & x_{n-1}^{n-3} \\ x_2^{n-2} & x_3^{n-2} & \dots & x_{n-1}^{n-2} \end{vmatrix}$$

Proceeding recursively this implies that $\det(V_n) = ((x_1 - x_0)(x_2 - x_0) \dots (x_n - x_0))((x_2 - x_1)(x_3 - x_1) \dots (x_{n-1} - x_1)) \dots = \prod_{i < j} (x_j - x_i)$ ♣

4.35

4.36 Find all possible Jordan forms for a transformation with characteristic polynomial: $(x + 5)^2(x - 3)^2$:

The eigenvalues are $\lambda_1 = -5, \lambda_2 = 3$ where both algebraic multiplicity 2.

All of the possible Jordan forms:

$$\begin{bmatrix} -5 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

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

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

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

4.37

4.38