Problem 2

Proof. Note that the matrix representation of D is:

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

Thus, the determinant of $D - \lambda I$ is given by:

$$|D - \lambda I| = \left| \begin{bmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 2 \\ 0 & 0 & -\lambda \end{bmatrix} \right| = -\lambda^3$$

So, the only eigenvalue is $\lambda = 0$ with eigenspace $span\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\}$. The algebraic multiplicity of λ is 3, and the geometric multiplicity of λ is 1. $\square$

Problem 4

Part (i)

Proof. Consider an Hermitian 2×2 matrix A. Define:

$$A := \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

From problem 3, we know that $p(\lambda) = \lambda^2 - tr(A)\lambda + det(A)$ Thus, the eigenvalue(s) of A are given by the forumla:

$$\lambda = \frac{tr(A) \pm \sqrt{tr(A)^2 - 4det(A)}}{2}$$

Note that all of the diagonal elements of A must be real since $A = A^H$, i.e. $a = \bar{a}$ for a a diagonal element. So, we have that $tr(A) = a + d \in \mathbb{R}$. It follows that the eigenvalue(s) of A are real iff $\sqrt{tr(A)^2 - 4det(A)}$ is real. We have that:

$$\begin{aligned} \sqrt{tr(A)^2 - 4det(A)} &= \sqrt{(a+d)^2 - 4(ad - bc)} \\ &= \sqrt{a^2 - 2ad + d^2 + 4bc} \\ &= \sqrt{(a-d)^2 + 4bc} \end{aligned}$$

Note that $(a - d)^2 \in \mathbb{R}_+$ since a and d are diagonal elements. Also, note that $c = \bar{b}$ since $A = A^H$. Define $b := \alpha + \beta i$ for some $\alpha, \beta \in \mathbb{R}$. Then $b\bar{b} = (\alpha + \beta i)(\alpha - \beta i) = \alpha^2 + \beta^2 \in \mathbb{R}_+$. So, we have that $4bc = 4b\bar{b} \in \mathbb{R}_+$. Thus, $\sqrt{(a-d)^2 + 4bc} = \sqrt{tr(A)^2 - 4det(A)}$ is real, as required. $\square$

Part (ii)

Proof. Consider a skew-Hermitian 2×2 matrix A. Define:

$$A \coloneqq \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$I \coloneqq \{\alpha + \beta i \in \mathbb{C} \mid \alpha = 0\}$$

From problem 3, we know that $p(\lambda) = \lambda^2 - tr(A)\lambda + det(A)$ Thus, the eigenvalue(s) of A are given by the forumla:

$$\lambda = \frac{tr(A) \pm \sqrt{tr(A)^2 - 4det(A)}}{2}$$

Note that all of the diagonal elements of A must be in I since $-A = A^H$, i.e. $-a = \bar{a}$ for a a diagonal element. So, we have that $tr(A) = a + d \in I$. It follows that the eigenvalue(s) of A are in I iff $\sqrt{tr(A)^2 - 4det(A)}$ is in I. We have that:

$$\begin{aligned} \sqrt{tr(A)^2 - 4det(A)} &= \sqrt{(a+d)^2 - 4(ad - bc)} \\ &= \sqrt{a^2 - 2ad + d^2 + 4bc} \\ &= \sqrt{(a-d)^2 + 4bc} \end{aligned}$$

Note that $(a-d)^2 \leq 0$ since $(a-d) \in I$ since $a, d \in I$ since they are diagonal elements. Also, note that $c = -\bar{b}$ since $A = -A^H$. Define $b \coloneqq \alpha + \beta i$ for some $\alpha, \beta \in \mathbb{R}$. Then $b(-\bar{b}) = (\alpha + \beta i)(\beta i - \alpha) = -\alpha^2 - \beta^2 \leq 0$. So, we have that $4bc = 4b(-\bar{b}) \leq 0$. Thus, $\sqrt{(a-d)^2 + 4bc} = \sqrt{tr(A)^2 - 4det(A)} \in I$ since it is the square root of a non-positive number, as required. $\square$

Problem 6

Proof. Consider $A - \lambda I$ for A an $n \times n$ upper-triangular matrix, defined as:

$$A - \lambda I \coloneqq \begin{bmatrix} \alpha_{1,1} - \lambda & * & \dots & * \\ 0 & \alpha_{2,2} - \lambda & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_{n,n} - \lambda \end{bmatrix}$$

By Laplace expansion along the first column, we get that:

$$det(A - \lambda I) = (\alpha_{1,1} - \lambda) det \left(\begin{bmatrix} \alpha_{2,2} - \lambda & * & \dots & * \\ 0 & \alpha_{3,3} - \lambda & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_{n,n} - \lambda \end{bmatrix} \right)$$

By induction, then, we get that:

$$det(A - \lambda I) = \left(\prod_{i=1}^{n-1} \alpha_{i,i} - \lambda\right) det([\alpha_{n,n} - \lambda]) = \prod_{i=1}^{n} \alpha_{i,i} - \lambda$$

Thus, we have that the eigenvalues of A are $\{\alpha_{1,1}, \alpha_{2,2}, \ldots, \alpha_{n,n}\}$, as required. The same proof applies for lower-triangular matrices since $det(A - \lambda I) = det((A - \lambda I)^T) = det(A^T - \lambda I^T) = det(A^T - \lambda I)$. □

Problem 8

Part (i)

Proof. In order to prove that S is a basis for V, we need only show that the elements of S are linearly independent, since S spans V by definition. Computing the Wronskian, we get:

$$W = det\left(\begin{bmatrix} sin(x) & cos(x) & sin(2x) & cos(2x) \\ cos(x) & -sin(x) & 2cos(2x) & -2sin(2x) \\ -sin(x) & -cos(x) & -4sin(2x) & -4cos(2x) \\ -cos(x) & sin(x) & -8cos(2x) & 8sin(2x) \end{bmatrix}\right) = 18 \neq 0$$

Thus, the elements of S are linearly independent, so S is a basis for V □

Part (ii)

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

Part (iii)

Proof. Consider $U := span\{sin(x), cos(x)\}$ and $W := span\{sin(2x), cos(2x)\}$. Then U and W are complementary subspaces of V since $U \cap W = \{0\}$ and $V = U \oplus W$. Also, U and W are clearly D-invariant, since the derivative of any element of U will be some linear combination of $sin(x)$ and $cos(x)$, and the derivative of any element in W will be some linear combination of $sin(2x)$ and $cos(2x)$. □

Problem 13

Consider the determinant of the matrix $A - \lambda I$:

$$det\left(\begin{bmatrix} 0.8 - \lambda & 0.4 \\ 0.2 & 0.6 - \lambda \end{bmatrix}\right) = (0.8 - \lambda)(0.6 - \lambda) - 0.08$$

Solving for λ, we get eigenvalues 0.4 and 1. The corresponding eigenvectors are of the form $x = -y$ and $x = 2y$, respectively. We will pick eigenvectors $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, for simplicity. Note that these eigenvectors form an eigenbasis of A, so A is semisimple. So, we will define P to be the matrix of the eigenvectors:

$$P \coloneqq \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$$

Note that:

$$P^{-1} = \begin{bmatrix} 1/3 & -2/3 \\ 1/3 & 1/3 \end{bmatrix}$$

So, we have that:

$$P^{-1}AP = \begin{bmatrix} 1/3 & -2/3 \\ 1/3 & 1/3 \end{bmatrix} \cdot \begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0.4 & 0 \\ 0 & 1 \end{bmatrix}$$

Problem 15

Proof. Suppose $A \in M_n(\mathbb{F})$ semisimple. Then $\exists$ a nonsingular matrix P and a diagonal matrix D s.t. $D = P^{-1}AP$. It follows then that $PDP^{-1} = A$, and therefore that $PD^nP^{-1} = A^n$ $\forall n \in \mathbb{N}$. Note that in particular, we have that:

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

It follows that:

$$\begin{aligned} f(A) &= a_0 I + a_1 A + \cdots + a_n A^n \\ &= a_0 I + a_1 PDP^{-1} + \cdots + a_n PD^nP^{-1} \\ &= P(a_0 + a_1 D + \cdots + a_n D^n)P^{-1} \\ &= P \cdot \begin{bmatrix} f(\lambda_1) & 0 & \dots & 0 \\ 0 & f(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f(\lambda_n) \end{bmatrix} \cdot P^{-1} \\ &\sim \begin{bmatrix} f(\lambda_1) & 0 & \dots & 0 \\ 0 & f(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f(\lambda_n) \end{bmatrix} \end{aligned}$$

Since similar matrices have the same eigenvalues, we have that $(f(\lambda_i))_{i=1}^n$ are the eigenvalues of $f(A)$. $\square$

Problem 16

Part (i)

Proof. As in problem 13, we may define a nonsingular matrix P and a diagonal matrix D s.t. $P^{-1}AP = D$:

$$P := \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \qquad P^{-1} = \begin{bmatrix} 1/3 & -2/3 \\ 1/3 & 1/3 \end{bmatrix} \qquad D := \begin{bmatrix} 0.4 & 0 \\ 0 & 1 \end{bmatrix}$$

It follows that $PDP^{-1} = A$, and therefore that $PD^nP^{-1} = A^n$. Note then that we have:

$$\begin{aligned} \lim_{n\to\infty} A^n &= \lim_{n\to\infty} PD^nP^{-1} \\ &= P(\lim_{n\to\infty} D^n)P^{-1} \\ &= P \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} P^{-1} \\ &= \begin{bmatrix} 2/3 & 2/3 \\ 1/3 & 1/3 \end{bmatrix} \end{aligned}$$

□

Part (ii)

The answer does not depend on the choice of norm, since $\lim_{n\to\infty} D^n = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ in every case.

Part (iii)

By theorem 4.3.12, the eigenvalues of $3I + 5A + A^3$ are $3 + 5 \cdot 1 + 0 \cdot 1^2 + 1 \cdot 1^3 = 9$ and $3 + 5 \cdot 0.4 + 0 \cdot 0.4^2 + 1 \cdot 0.4^3 = 5.064$

Problem 18

Proof. Define $A \in M_n(\mathbb{F})$ as:

$$A := \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,n} \end{bmatrix}$$

In order to have a nonzero row vector x^T s.t. $x^TA = \lambda x^T$, we must have that $\alpha_{1,i}x_1^T, \alpha_{2,i}x_2^T, \ldots, \alpha_{n,i}x_n^T = \lambda x_i^T \ \forall\ 1 \le i \le n$, with $\{x_1^T, x_2^T, \ldots, x_n^T\}$ the components of x^T. Equivalently, we must have that the null space of the matrix B is nonzero, for B defined as:

$$B := \begin{bmatrix} \alpha_{1,1} - \lambda & \alpha_{2,1} & \dots & \alpha_{n,1} \\ \alpha_{1,2} & \alpha_{2,2} - \lambda & \dots & \alpha_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1,n} & \alpha_{2,n} & \dots & \alpha_{n,n} - \lambda \end{bmatrix}$$

Note that $B = (A - \lambda I)^T$. Since λ is an eigenvalue of A, we have that $det(A - \lambda I) = 0$. It follows that $det((A - \lambda I)^T) = det(A - \lambda I) = 0$, i.e. that the null space of $(A - \lambda I)^T = B$ is nonzero, as required. □

Problem 20

Proof. Suppose A is Hermitian and orthonormally similar to B. Then we have that $B = U^H AU$ for U an orthonormal matrix, and $A = A^H$. So, we have that:

$$\begin{aligned} B &= U^H AU \\ &= U^H A^H U \\ &= (U^H AU)^H \\ &= B^H \end{aligned}$$

Thus, B is Hermitian. □

Problem 24

Proof. Suppose A is Hermitian, so $A = A^H$. Then we have that:

$$\begin{aligned} \rho(x) &= \frac{\langle x, Ax \rangle}{\|x\|^2} \\ &= \frac{x^H Ax}{\|x\|^2} \\ &= \frac{x^H A^H x}{\|x\|^2} \\ &= \frac{(x^H Ax)^H}{\|x\|^2} \\ &= \frac{(\langle x, Ax \rangle)^H}{\|x\|^2} \end{aligned}$$

So, we must have that $\langle x, Ax\rangle$ is real, so $\frac{\langle x, Ax\rangle}{\|x\|^2} = \rho(x)$ is real since $\|x\|^2$ is real.

Suppose A is skew-Hermitian, so $A = -A^H$. Then we have that:

$$\begin{aligned}\rho(x) &= \frac{\langle x, Ax\rangle}{\|x\|^2} \\ &= \frac{x^H A x}{\|x\|^2} \\ &= \frac{-x^H A^H x}{\|x\|^2} \\ &= \frac{-(x^H A x)^H}{\|x\|^2} \\ &= \frac{-(\langle x, Ax\rangle)^H}{\|x\|^2}\end{aligned}$$

So, we must have that $\langle x, Ax\rangle$ is imaginary, so $\frac{\langle x, Ax\rangle}{\|x\|^2} = \rho(x)$ is imaginary since $\|x\|^2$ is real. $\square$

Problem 25

Part (i)

Proof. Note that $\forall\, j$ s.t. $1 \leq j \leq n$:

$$\begin{aligned}(x_1x_1^H + \cdots + x_jx_j^H + \cdots + x_nx_n^H)x_j &= x_1x_1^Hx_j + \cdots + x_jx_j^Hx_j + \cdots + x_nx_n^Hx_j \\ &= x_1\langle x_1, x_j\rangle + \cdots + x_j\langle x_j, x_j\rangle + \cdots + x_n\langle x_n, x_j\rangle \\ &= 0 + \cdots + x_j \cdot 1 + \cdots + 0 \\ &= x_j\end{aligned}$$

Thus, since $\{x_1, \ldots, x_n\}$ is an orthonormal basis, we have that since $(x_1x_1^H + \cdots + x_nx_n^H)x_j = x_j\ \forall\, j$ s.t. $1 \leq j \leq n$, the matrix $(x_1x_1^H + \cdots + x_nx_n^H)$ must be the identity matrix I. $\square$

Part (ii)

Proof. Note that since $(x_1x_1^H + \cdots + x_nx_n^H)$ is the identity matrix I by part (i), we may write:

$$\begin{aligned}A &= A(x_1x_1^H + \cdots + x_nx_n^H) \\ &= Ax_1x_1^H + \cdots + Ax_nx_n^H \\ &= \lambda_1x_1x_1^H + \cdots + \lambda_nx_nx_n^H\end{aligned}$$

$\square$

Problem 27

Proof. Define:

$$A := \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \dots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \dots & \alpha_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \dots & \alpha_{n,n} \end{bmatrix}$$

Note that the diagonal elements of A must be real since A is Hermitian. Consider the unit vector $e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{F}^n$. Then we must have that:

$$\langle e_1, Ae_1 \rangle = e_1^H A e_1 = \alpha_{1,1} > 0$$

This can be applied for every e_i, so it follows that $\alpha_{i,i} > 0 \ \forall \ i$ s.t. $1 \leq i \leq n$. Thus, we have that all of the diagonal elements of A are real and positive. □

Problem 28

Proof. Replacing positive definite with positive semidefinite in the previous proof, we get that all of the diagonal elements of A and B are real and nonnegative. So, $0 \leq tr(A)$ and $0 \leq tr(B)$, so $0 \leq tr(A)tr(B)$. Also, note that we must have $A = S^H S$ and $B = T^H T$ for some matrices S and T. So, we have that:

$$tr(AB) = tr(S^H S T^H T) = tr((TS^H)(ST^H)) = tr((ST^H)^H(ST^H))$$

Thus, since $\forall \ a \in \mathbb{C}$, $a\bar{a} \in \mathbb{R}$ and $0 \leq a\bar{a}$, we have that every diagonal element of $(ST^H)^H(ST^H)$ is real and nonnegative. Thus, $0 \leq tr(AB)$

Note that since A is Hermitian, we may define $A = U^H DU$ for U an orthonormal matrix and D a diagonal matrix. Thus,

$$tr(AB) = tr(U^H DUB) = tr(DUBU^H)$$

Since $tr(UBU^H = tr(B)$, we may assume that $A = diag(\lambda_1, \dots, \lambda_n)$. Let $\{\beta_{1,1}, \dots \beta_{n,n}\}$ be the diagonal elements of B. Then, we have that:

$$\begin{aligned} tr(AB) &= \lambda_1\beta_{1,1}, \dots, \lambda_n\beta_{n,n} \\ &\leq (\lambda_1, \dots, \lambda_n)(\beta_{1,1}, \dots \beta_{n,n}) \\ &= tr(A)tr(B) \end{aligned}$$

□

Problem 31

Part (i)

Proof. Define $\|A\|_2 \coloneqq \max\limits_{\|x\|_2=1} \|Ax\|_2$. As in theorem 4.5.10, we can find an orthonormal matrix V s.t.

$$A^H A = V \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^H$$

where $D = diag(\lambda_1, \dots, \lambda_r)$ and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$. So, we have that, for x s.t. $\|x\|_2 = 1$:

$$\|Ax\|_2^2 = x^H A^H Ax = x^H V \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^H x = y^H \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} y = \sum_{i=1}^{r} \lambda_i \bar{y}_i y_i$$

where $y \coloneqq V^H x$. It follows that:

$$\|Ax\|_2^2 \leq \lambda_{max} y^H y = \lambda_{max} x^H V V^H x = \lambda_{max} x^H x = \lambda_{max}$$

Since λ_{max} is an eigenvalue of $A^H A$, we know that $\exists$ x s.t. $\|x\|_2 = 1$ and $A^H Ax = \lambda_{max} x$. For such an x, we have that:

$$\|Ax\|_2^2 = x^H A^H Ax = \lambda_{max} x^H x = \lambda_{max}$$

Thus, we have that $\|Ax\|_2 = \sqrt{\lambda_{max}}$, with $\sqrt{\lambda_{max}}$ being the largest singular value of A, by definition. □

Part (ii)

Proof. Since A is invertible, it does not have any 0 eigenvalues, so we have that $\lambda_n = \lambda_{min}$. Consider the SVD of A, $A = U\Sigma V^H$. It follows that:

$$A^{-1} = (U\Sigma V^H)^{-1} = (V^H)^{-1} \Sigma^{-1} U^{-1} = V \Sigma^{-1} U^H$$

since U, V orthonormal. So, we have constructed an SVD of A^{-1}. By part (i), then, we have that $\|A^{-1}\|_2 = max\{diag(\Sigma^{-1})\}$. Since the diagonal elements of Σ^{-1} are simply the multiplicative inverses of the diagonal elements of Σ, we have that $\|A^{-1}\|_2 = max\{diag(\Sigma^{-1})\} = \sigma_{min}^{-1} = \sigma_n^{-1}$ □

Part (iii)

By part (i), $\|A^H A\|_2$ is given by the largest singular value of $A^H A$. This is the square root of the largest nonzero eigenvalue of $(A^H A)^H A^H A = (A^H A)^2$. By Theorem 4.3.12, this is simply the largest nonzero eigenvalue of $A^H A$, i.e. σ_1^2, where σ_1 is the largest singular value of A. So, we have that $\|A^H A\|_2 = \|A\|_2^2$ since $\|A\|_2 = \sigma_1$.

$\|A^H\|_2$ is given by the largest singular value of A^H. This is the square root of the largest nonzero eigenvalue of $(A^H)^H A^H = AA^H$. Since the nonzero eigenvalues of AA^H are the same as those of $A^H A$, we have that $\|A^H\|_2 = \sigma_1$, i.e. $\|A^H\|_2^2 = \|A\|_2^2$.

$\|A^T\|_2$ is given by the largest singular value of A^T. This is the square root of the largest nonzero eigenvalue of $(A^T)^H A^T = (A^H)^T A^T = (AA^H)^T$. Since the nonzero eigenvalues of AA^H are the same as those of $A^H A$, and since the nonzero eigenvalues of AA^H are the same as those of $(AA^H)^T$, we have that $\|A^T\|_2 = \sigma_1$, i.e. $\|A^T\|_2^2 = \|A\|_2^2$.

Part (iv)

Proof. $\|UAV\|_2$ is given by the largest singular value of UAV. This is the square root of the largest nonzero eigenvalue of $(UAV)^H UAV = V^H A^H U^H UAV = V^H A^H AV$. Since V orthonormal, this is the same as the square root of the largest nonzero eigenvalue of $A^H A$, i.e. the largest singular value of A. This is equivalent to $\|A\|_2$ by part (i). So, we have that $\|UAV\|_2 = \|A\|_2$ □

Problem 32

Part (i)

Proof.

$$\begin{aligned}
\|UAV\|_F &= \sqrt{tr((UAV)^H UAV)} \\
&= \sqrt{tr(V^H A^H U^H UAV)} \\
&= \sqrt{tr(V^H A^H AV)} \\
&= \sqrt{tr(A^H AVV^H)} \\
&= \sqrt{tr(A^H A)} \\
&= \|A\|_F
\end{aligned}$$

□

Part (ii)

Consider the SVD of A, $A = U\Sigma V^H$. Then, we have that:

Proof.

$$\begin{aligned}
\|A\|_F &= \|U\Sigma V^H\|_F \\
&= \|\Sigma\|_F && \text{by part (i)} \\
&= \sqrt{tr(\Sigma^H \Sigma)} \\
&= (\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2)^{1/2}
\end{aligned}$$

□

Problem 33

Proof. By problem 31, we have that $\|A\|_2 = \sigma_1$. Consider the SVD of A, $A = U\Sigma V^H$. Then, we have that:

$$\begin{aligned}
\sup_{\substack{\|x\|_2=1 \\ \|y\|_2=1}} |y^H Ax| &= \sup_{\substack{\|x\|_2=1 \\ \|y\|_2=1}} |y^H U\Sigma V^H x| \\
&= \sup_{\substack{\|a\|_2=1 \\ \|b\|_2=1}} |a^H U^H U\Sigma V^H V b| && \text{with } x := Vb \text{ and } y := Ua \\
&= \sup_{\substack{\|a\|_2=1 \\ \|b\|_2=1}} |a^H \Sigma b| \\
&= \sup_{\substack{\|a\|_2=1 \\ \|b\|_2=1}} \left| \sum_{i=1}^{r} \bar{a}_i \sigma_i b_i \right|
\end{aligned}$$

Clearly, the above sup is attained at $a = b = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, since this puts all the weight on the largest singular value σ_1. So, we have:

$$\sup_{\substack{\|x\|_2=1 \\ \|y\|_2=1}} |y^H Ax| = |\sigma_1| = \sigma_1 = \|A\|_2$$

□

Problem 36

Proof. Consider the matrix:

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

Then the determinant of the matrix is 1, the eigenvalues of the matrix are both -1, and the singular values of the matrix are both 1. ☐

Problem 38

Note that for this problem, we will denote U_1, Σ_1 and V_1 as U, Σ and V.

Part (i)

$$\begin{aligned} AA^\dagger A &= U\Sigma V^H V\Sigma^{-1}U^H U\Sigma V^H \\ &= U\Sigma\Sigma^{-1}\Sigma V^H \\ &= U\Sigma V^H \\ &= A \end{aligned}$$

Part (ii)

$$\begin{aligned} A^\dagger AA^\dagger &= V\Sigma^{-1}U^H U\Sigma V^H V\Sigma^{-1}U^H \\ &= V\Sigma^{-1}\Sigma\Sigma^{-1}U^H \\ &= V\Sigma^{-1}U^H \\ &= A^\dagger \end{aligned}$$

Part (iii)

$$\begin{aligned} (AA^\dagger)^H &= (U\Sigma V^H V\Sigma^{-1}U^H)^H \\ &= (UU^H)^H \\ &= UU^H \\ &= U\Sigma V^H V\Sigma^{-1}U^H \\ &= AA^\dagger \end{aligned}$$

Part (iv)

$$\begin{aligned}(A^\dagger A)^H &= (V\Sigma^{-1}U^H U\Sigma V^H)^H \\ &= (VV^H)^H \\ &= VV^H \\ &= V\Sigma^{-1}U^H U\Sigma V^H \\ &= A^\dagger A\end{aligned}$$

Part (v)

Proof. Consider $(AA^\dagger)^2 = AA^\dagger AA^\dagger = AA^\dagger$ by part (i). So, we have that $A^\dagger$ is a projection matrix. By part (iii), it is orthogonal since $(AA^\dagger)^H = AA^\dagger$. Clearly, $AA^\dagger x \in \mathcal{R}(A)$. □

Part (vi)

Proof. Consider $(A^\dagger A)^2 = A^\dagger AA^\dagger A = A^\dagger A$ by part (ii). So, we have that $A^\dagger A$ is a projection matrix. By part (iv), it is orthogonal since $(A^\dagger A)^H = A^\dagger A$. By the proof of theorem 4.6.1, we know that for any $b \in \mathbb{F}^m$, $A^\dagger b \in \mathcal{R}(V) = \mathcal{R}(A^H)$ So, we have that $A^\dagger Ax \in \mathcal{R}(A^H)$. □