4.2 Let $\text{span}(\{1, x, x^2\})$ be a subspace of the inner product space $L^2([0,1]);\mathbb{R}$. Let D be the derivative operator $D : V \to V$ given by $D[p](x) = p'(x)$. Recall from a previous exercise (from Wk2) that $D = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. Given that D is upper triangular, we observe that the eigenvalues are all 0. So, the algebraic multiplicity is 3. Note that D has just one eigenvector for eigenvalue 0. Thus, the eigenspace of 0 is $\text{span}(\{1\})$ and the geometric multiplicity is 1.

4.4

(i) *Proof.* We know that the determinant of a 2×2 matrix A denoted by $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ can be expressed as $\det(A) = ad - bc$. Now, given the Hermitian of A, denoted A^H, it follows that $a = \bar{a}$, $b = \bar{c}$, and $d = \bar{d}$. Thus, a and d must be real. Observe that $bc = \bar{c}c = ||c||^2$ is also real. Now, by exercise 4.3 the characteristic polynomial of any 2×2 matrix has the form

$$\begin{aligned} p(\lambda) &= \lambda^2 - \text{tr}(A)\lambda + \det(A) \\ &= \lambda^2 - (a+d)\lambda + ad - ||c||^2 \end{aligned}$$

It follows that the solutions of this are given by

$$\begin{aligned} \lambda_{\pm} &= \frac{(a+d) \pm \sqrt{(a+d)^2 - 4(ad - ||c||^2)}}{2} \\ &= \frac{(a+d) \pm \sqrt{(a-d)^2 + ||c||^2}}{2} \end{aligned}$$

But $(a-d)^2 + ||c||^2 \geq 0$, thus $\lambda_{\pm}$ is real. □

(ii) Let A be the same 2×2 matrix as defined in 4.4 (i). Hence $\det(A) = ad - bc$. Now, suppose A is a skew-Hermitian matrix (still 2×2. It follows that $a = -\bar{a}$, $b = -\bar{c}$, and $d = -\bar{d}$. Thus, a and d are imaginary. Furthermore, $bc = -\bar{c}c = -||c||^2$ and ad are both negative. By similar fashion to part (i) we use the same characteristic polynomial, whose solutions are also given by

$$\begin{aligned}\lambda_{\pm} &= \frac{(a+d) \pm \sqrt{(a+d)^2 - 4(ad - ||c||^2)}}{2} \\ &= \frac{(a+d) \pm \sqrt{(a-d)^2 + ||c||^2}}{2}\end{aligned}$$

However, this time $(a-d)^2 + ||c||^2 < 0$ (since each term is negative). Thus, for all a, b, c, d we have that $\lambda_{\pm}$ is imaginary.

4.6

Proof. Let $A \in M_n(\mathbb{F})$ be upper-triangular, and suppose $\lambda \in \mathbb{C}$. It follows that $\lambda I - A$ is also upper-triangular, and the determinant is just the product of the entries that lie along the diagonal of the upper-triangular matrix. The characteristic polynomial of A is given by

$$\begin{aligned}p(\lambda) &= \det(\lambda I - A) \\ &= \prod_{i=1}^{n} (\lambda - a_{ii})\end{aligned}$$

Note that the roots to this are the diagonal entries of A, and hence are the eigenvalues. Note that the proof for a lower-diagonal matrix is similar in that the determinant is the same as for an upper-triangular matrix. □

4.8 Let V be the span of the set $S = \{\sin(x), \cos(x), \sin(2x), \cos(2x)\}$ in the vector space $C^{\infty}(\mathbb{R}, \mathbb{R})$

(i) *Proof.* Recall from a previous exercise (from Wk2) that this set is orthonormal given the inner product $\langle \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)g(t)dt \rangle$. So, each element in the spanning is independent. Thus, they form a basis for the span. Therefore, S is a basis for V. □

(ii) Let D be the derivative operator. Observe that

$$\begin{aligned} D\sin(x) &= \cos(x) \\ D\cos(x) &= -\sin(x) \\ D\sin(2x) &= 2\cos(x) \\ D\cos(2x) &= -2\sin(x) \end{aligned}$$

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

(iii) Two complementary D-invariant subspaces of V are $\operatorname{span}(\{\sin(x), \cos(x)\})$ and $\operatorname{span}(\{\sin(2x), \cos(2x)\})$

4.13 Let $A = \begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix}$. Observe that

$$\det(\lambda I - A) = \lambda^2 - 1.4\lambda + 0.4$$

The roots of which are 1 and 0.4. Hence the eigenvalues are 1 and 0.4 The corresponding eigenvector for $\lambda = 1$ is the null space of $\begin{bmatrix} -0.2 & 0.4 \\ 0.2 & -0.4 \end{bmatrix}$ and has a solution of $[2 \;\; 1]^T$. Now, the eigenvector for $\lambda = 0.4$ is the null space of $\begin{bmatrix} 0.4 & 0.4 \\ 0.2 & 0.2 \end{bmatrix}$ and has a solution of $[1 \;\; -1]^T$. Thus

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

4.15

Proof. Let $(\lambda_i)_{i=1}^n$ be the eigenvalues of a semisimple matrix $A \in M_n(\mathbb{F})$ and $f(x) = a_0 + a_1 x + ... + a_n x^n$ be a polynomial. Now by Theorem 4.3.7, A can be diagonalized as PBP^{-1}. Observe that

$$\begin{aligned} f(A) &= a_0 I + a_1 A + ... + a_n A^n \\ &= a_0 PP^{-1} + a_1 PBP^{-1} + ... + a_n PB^n P^{-1} \\ &= Pf(B)P^{-1} \end{aligned}$$

but each term in $f(B)$ is a diagonal matrix. Hence, each diagonal entry is $(f(\lambda_i))_{i=1}^n$, and since $f(B)$ is similar to $f(A)$, it follows that the eigenvalues are the same denoted by $(f(\lambda_i))_{i=1}^n$ □

4.16 Let A be the matrix in Exercise 4.13, namely $A = \begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix}$

(i) Observe that $A^n = PC^nP^{-1}$ with

$$\begin{aligned} C^n &= \begin{bmatrix} 1^n & 0 \\ 0 & 0.4^n \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0.4^n \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} A^k &= \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0.4^k \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{-2}{3} \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} 2 + 0.4^k & 2 - 2 \cdot 0.4^k \\ 1 - 0.4^k & 1 + 2 \cdot 0.4^k \end{bmatrix} \end{aligned}$$

Note also that the limit $B = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{-2}{3} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$ Now, observe that $A^k - B = \frac{1}{3} \begin{bmatrix} 0.4^k & -2 \cdot 0.4^k \\ -0.4^k & 2 \cdot 0.4^k \end{bmatrix}$ and each term converges. Thus, it converges with respect to the 1-norm.

(ii) Note that for a given matrix, the ∞-norm is just the largest row sum. Note also that $-0.4^k + 2 \cdot 0.4^k \to 0$ as $k \to \infty$.

Now, observe that

$$\begin{aligned}||A^k - B||_F =& \sqrt{\operatorname{tr}\left(\begin{bmatrix} 0.4^k & -0.4^k \\ -2 \cdot 4^k & 2 \cdot 0.4^k \end{bmatrix} \begin{bmatrix} 0.4^k & -2 \cdot 0.4^k \\ -0.4^k & 2 \cdot 0.4^k \end{bmatrix} \right)} \\ =& \sqrt{\operatorname{tr}\left(\begin{bmatrix} 2 \cdot 0.4^{2k} & -4 \cdot 0.4^{2k} \\ -4 \cdot 4^{2k} & 8 \cdot 0.4^{2k} \end{bmatrix} \right)} \\ =& \sqrt{10 \cdot 0.4^{2k}} \to 0 \text{ as } k \to \infty \end{aligned}$$

Hence, $||A^k - B||_F \to 0$. Therefore, the answer does not depend on the choice of norm.

(iii) By Theorem 4.3.12, since A is semi-simple, the eigenvalues for $f(A) = 3I+5A+A^3$ are $f(1) = 3+5+1 = 9$ and $f(0.4) = 3+5\cdot 0.4+0.4^3 = 5.064$, where 1 and 0.4 are the eigenvalues from exercise 4.13.

4.18

Proof. Let λ be an eigenvalue of $A \in M_n(\mathbb{F})$ Now, A and A^T have the same characteristic polynomial, hence it follows that λ is also an eigenvalue of A^T. Thus, there exists some $\mathbf{x}$ such that $A^T\mathbf{x} = \lambda\mathbf{x}$ which implies that $(A^T\mathbf{x})^T = (\lambda\mathbf{x})^T$. Therefore $\mathbf{x}^T A = \lambda\mathbf{x}^T$ □

4.20

Proof. Let A be Hermitian and orthonormally similar to B. Observe that

$$\begin{aligned} B &= PAP^H \\ &= PA^H P^H \\ &= (PAP^H)^H \\ &= B^H \end{aligned}$$

Therefore, B is Hermitian □

4.24

Proof. Let $A \in M_n(\mathbb{C})$. Define the *Rayleigh quotient* as

$$\rho(\mathbf{x}) = \frac{\langle \mathbf{x}, A\mathbf{x} \rangle}{||\mathbf{x}||^2},$$

where $\langle \cdot, \cdot \rangle$ is the usual inner product on $\mathbb{F}^n$. It must be that $||\mathbf{x}||^2$ is always real, so we determine what happens with the numerator of the *Rayleigh quotient*.
So, suppose A is Hermitian. Observe that

$$\begin{aligned} \langle \mathbf{x}, A\mathbf{x} \rangle &= \mathbf{x}^H A\mathbf{x} \\ &= \mathbf{x}^H A^H \mathbf{x} \\ &= \langle A\mathbf{x}, \mathbf{x} \rangle \\ &= \overline{\langle \mathbf{x}, A\mathbf{x} \rangle} \end{aligned}$$

Hence the numerator of the *Rayleigh quotient* is real, so the *Rayleigh quotient* must take real values.
Now suppose A is Skew-Hermitian. Observe that

$$\begin{aligned} \langle \mathbf{x}, A\mathbf{x} \rangle &= \mathbf{x}^H A\mathbf{x} \\ &= -\mathbf{x}^H A^H \mathbf{x} \\ &= -\langle A\mathbf{x}, \mathbf{x} \rangle \\ &= -\overline{\langle \mathbf{x}, A\mathbf{x} \rangle} \end{aligned}$$

Hence the numerator of the *Rayleigh quotient* is imaginary, so the *Rayleigh quotient* must also take imaginary values. □

4.25 Let $A \in M_n(\mathbb{C})$ be a normal matrix with eigenvalues $\lambda_1, ..., \lambda_n$ and corresponding eigenvectors $[\mathbf{x}_1, ..., \mathbf{x}_n]$.

(i) Observe that $\langle \mathbf{x}_j, \mathbf{x}_j \rangle = \mathbf{x}_j^H \mathbf{x}_j = 1$, and also that $\langle \mathbf{x}_i, \mathbf{x}_j \rangle = \mathbf{x}_i^H \mathbf{x}_j = 0$ for all $i \neq j$. So, $(\mathbf{x}_1\mathbf{x}_1^H + ... + \mathbf{x}_n\mathbf{x}_n^H)\mathbf{x}_j = \mathbf{x}_j\mathbf{x}_j^H\mathbf{x}_j = I\mathbf{x}_j$. Hence $I = \mathbf{x}_1\mathbf{x}_1^H + ... + \mathbf{x}_n\mathbf{x}_n^H$.

(ii) Observe that $(\lambda_1\mathbf{x}_1\mathbf{x}_1^H + ... + \lambda_n\mathbf{x}_n\mathbf{x}_n^H)\mathbf{x}_j = \lambda\mathbf{x}_j\mathbf{x}_j^H\mathbf{x}_j = \lambda\mathbf{x}_j = A\mathbf{x}_j$. Hence $A = \lambda_1\mathbf{x}_1\mathbf{x}_1^H + ... + \lambda_n\mathbf{x}_n\mathbf{x}_n^H$.

4.27

Proof. Let $A \in M_n\mathbb{F}$ be positive definite. From definition 4.5.1 it follows that $\langle \mathbf{x}, A\mathbf{x} \rangle = \mathbf{x}^H A\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$. Now take e_i which the the i^{th} vector of the standard basis. Thus $0 < e_i^H A e_i = a_{ii}$ must be real and positive by definition. $\square$

4.28

Proof. Let $A, B \in M_n(\mathbb{F})$ be positive semidefinite. So, it follows that there exist some matrices S, T such that $A = S^H S$ and $B = T^H T$. Observe that

$$\begin{aligned}
\operatorname{tr}(AB) &= \operatorname{tr}(S^H S T^H T) \\
&= \operatorname{tr}(TS^H (TS^H)^H) \\
&= \operatorname{tr}((TS^H)^H TS^H)
\end{aligned}$$

where in the inside of the last expression is positive semidefinite, and has nonnegative entries along the diagonal (see exercise 4.26). Hence, $\operatorname{tr}(AB)7 = \operatorname{tr}((TS^H)^H TS^H) \geq 0$. Now, let $A = PDP^P - 1$ and $B = QEQ^{-1}$. Observe that

$$\begin{aligned}
\operatorname{tr}(AB) &= \operatorname{tr}(PDP^{-1}QEQ^{-1}) \\
&= \operatorname{tr}(PP^{-1}QDEQ^{-1}) \\
&= \operatorname{tr}(QQ^{-1}DE) \\
&= \text{DE} \\
&= \sum_i \lambda_i \xi_i \\
&\leq \sum_i \lambda_i \sum_i \xi_i \\
&= \operatorname{tr}(A)\operatorname{tr}(B)
\end{aligned}$$

where λ_i and ξ_i represent the eigenvalues of both A and B respectively.
Now we must show that $||\cdot||_F$ is a matrix norm.
Positivity. Observe that $||A||_F = \operatorname{tr}(A^H A)$ where $A^H A$ is positive semi

definite. So, it follows that $\text{tr}(A^H A) \geq 0$. Hence $||A||_F \geq 0$. Now suppose $||A||_F = 0$. Given that all the elements along the diagonal of $A^H A$ are weakly positive, it follows that each one must be 0 so that $||A||_F = 0$. Thus, each singular value of A is 0, and A is the 0 matrix.

Scale Preservation. Suppose we have a scalar $c \in \mathbb{R}$. Observe that

$$\begin{aligned} ||cA||_F &= \sqrt{\text{tr}((cA)^H (cA))} \\ &= \sqrt{c^2 \text{tr}(A^H A)} \\ &= c\sqrt{\text{tr}(A^H A)} \\ &= c||A||_F \end{aligned}$$

Triangle inequality. Observe that

$$\begin{aligned} ||A+B||_F^2 = \text{tr}((A+B)^H (A+B)) &= \text{tr}(A^H A + A^H B + B^H A + B^H B) \\ &= \text{tr}(A^H A) + 2\text{tr}(A^H B) + \text{tr}(B^H B) \\ &\leq ||A||_F^2 + 2||A||\,||B|| + ||B||_F^2 \\ &= (||A||_F + ||B||_F)^2 \end{aligned}$$

(Note that the inequality comes from applying the Cauchy-Schwartz Inequality)

Submultiplicative Property. This follows from the inequality we proved above; namely $0 \leq \text{tr}(AB) \leq \text{tr}(A)\text{tr}(B)$.

Therefore, given that these properties are satisfied, it follows that $||\cdot||_F$ is a matrix norm. $\square$

4.31 Let $A \in M_{m\times n}(\mathbb{F})$ and A not identically zero.

(i) *Proof.* Let $A = U\Sigma V^H$, $y = V^H x$, and σ_1 be the largest singular value of A. Observe that

$$\begin{aligned}
||A||_2 &= \sup_{x \neq 0} \frac{||Ax||_2}{||x||_2} \\
&= \sup_{x \neq 0} \frac{||U\Sigma V^H x||_2}{||x||_2} \\
&= \sup_{x \neq 0} \frac{||\Sigma V^H x||_2}{||x||_2} \\
&= \sup_{y \neq 0} \frac{||\Sigma y||_2}{||Vy||_2} \\
&= \sup_{y \neq 0} \frac{||\Sigma y||_2}{||y||_2} \\
&= \sup_{||y||=1} ||\Sigma y||_2 \\
&= \sigma_1
\end{aligned}$$

□

(ii) Let $A = U\Sigma V^H$ and σ_n be the smallest singular value of A. Suppose A is invertible. Observe that

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

Note that $\frac{1}{\sigma_1}, ..., \frac{1}{\sigma_n}$ are the diagonal entries of Σ^{-1}. Given that σ_n is the smallest singular value of A, it follows that $\frac{1}{\sigma_n}$ is the largest singular value of A^{-1}. By part(i), $||A^{-1}||_2 = \frac{1}{\sigma_n}$

(iii) Let $A = U\Sigma V^H$. Note then that $A^T = V\Sigma^T U^T$, and also that $A^H = V\Sigma^H U^H$. Now, the singular values of A are each positive and real, and Σ is diagonal comprised of singular values. Hence, it is equivalent to its transpose as well as its Hermitian. So by part (i), we have that

$$\begin{aligned}
||A||_2^2 &= ||A^T||_2^2 \\
&= ||A^H||_2^2 \\
&= \sigma_1^2
\end{aligned}$$

Observe that

$$\begin{aligned} A^H A &= V\Sigma^H U^H U \Sigma V^H \\ &= V\Sigma^H \Sigma V^H \\ &= V\Sigma^2 V^H \end{aligned}$$

Given that multiplication is preserved with diagonal matrices, it follows that Σ^2 is also diagonal, with σ_i^2 as the diagonal entries. So by part(i) it follows that $||A^H A||_2 = \sigma_1^2 = ||A||_2^2$

(iv) Let $U \in M_m(\mathbb{F})$ and $V \in M_n(\mathbb{F})$ be orthonormal. Observe that

$$\begin{aligned} ||UAV||_2^2 &= ||(UAV)^H UAV||_2 \\ &= ||V^H A^H AV||_2 \\ &= ||A^H AVV^H||_2 \\ &= ||A^H A||_2 \\ &= ||A||_2^2 \end{aligned}$$

This follows from norm properties and by part(iii).

4.32 Let $A \in M_{m\times n}()\mathbb{F}$ be of rank r.

(i) *Proof.* Let $U \in M_m(\mathbb{F})$ and $V \in M_n(\mathbb{F})$ be orthonormal. Observe that

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

□

(ii) *Proof.* By SVD and part (i), we have that

$$\begin{aligned}
||A||_F &= ||U\Sigma V^H||_F \\
&= ||\Sigma||_F \\
&= \sqrt{\operatorname{tr}(\Sigma^H \Sigma)} \\
&= \left(\sum_{i=1}^{r} \sigma_i^2\right)^{\frac{1}{2}} \\
&= (\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_r^2)^{\frac{1}{2}}
\end{aligned}$$

□

4.33

Proof. Let $A \in M_n(\mathbb{F})$. By exercise 4.31, we have that $||A||_2 = \sigma_1$, with σ_1 being the largest singular value of A. Observe that

$$\begin{aligned}
\sup_{||x||_2=1, ||y||_2=1} |\mathbf{y}^H A\mathbf{x}| &\leq \sup_{||x||_2=1, ||y||_2=1} ||\mathbf{y}||_2 ||\Sigma \mathbf{x}||_2 \\
&= \sup_{||x||_2=1} ||\Sigma \mathbf{x}||_2 \\
&\leq \sigma_1 \text{ by exercise 4.31}
\end{aligned}$$

Now, if we let $\mathbf{x}$ and $\mathbf{y}$ be e_1 the standard eigenvector, it follows that

$$\begin{aligned}
\sup_{||x||_2=1, ||y||_2=1} |\mathbf{y}^H A\mathbf{x}| &\geq |\mathbf{y}^H A\mathbf{x}| \\
&= \sigma_1
\end{aligned}$$

Therefore $\sup_{||x||_2=1, ||y||_2=1} |\mathbf{y}^H A\mathbf{x}| = \sigma_1 = ||A||_2$. □

4.36 Let $A = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}$. So, $A^H A = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$. Now $\det(A) = -2$, and the singular values of A are 1 and 2, with eigenvalues $\pm\sqrt{2}$.

4.38

Proof. Let $A \in M_{m\times n}(\mathbb{F})$. Then the Moore-Penrose pseudoinverse of A satisfies the following:

(i)

$$\begin{aligned} AA^\dagger A &= U_1\Sigma_1 V_1^H V_1 \Sigma_1^{-1} U_1^H U_1 \Sigma_1 V_1^H \\ &= U_1\Sigma_1\Sigma_1^{-1}\Sigma_1 V_1^H \\ &= U_1\Sigma_1 V_1^H \\ &= A \end{aligned}$$

(ii)

$$\begin{aligned} A^\dagger A A^\dagger &= V_1\Sigma_1^{-1} U_1^H U_1 \Sigma_1 V_1^H V_1 \Sigma_1^{-1} U_1^H \\ &= V_1\Sigma_1^{-1}\Sigma_1\Sigma_1^{-1} U_1^H \\ &= V_1\Sigma_1^{-1} U_1^H \\ &= A^\dagger \end{aligned}$$

(iii)

$$\begin{aligned} (AA^\dagger)^H &= (U_1\Sigma_1 V_1^H V_1 \Sigma^{-1} U_1^H)^H \\ &= U_1\Sigma_1 V_1^H V_1 \Sigma_1^{-1} U_1^H \\ &= AA^\dagger \end{aligned}$$

(iv)

$$\begin{aligned} (A^\dagger A)^H &= (V_1\Sigma_1^{-1} U_1^H U_1 \Sigma_1 V_1^H)^H \\ &= V_1\Sigma_1^{-1} U_1^H U_1 \Sigma_1 V_1^H \\ &= A^\dagger A \end{aligned}$$

(v) By part (i), we have that $(AA^\dagger A)A^\dagger = A^\dagger A$. Hence $(AA^\dagger)(AA^\dagger) = A^\dagger A$. Now, let $U_1 = [u_1, ..., u_n]$. So, U_1 is an orthonormal basis for $\mathscr{R}(A)$ by

SVD. By part (iii), we have that

$$\begin{aligned}
AA^{\dagger} &= U_1 U_1^H x \\
&= U_1[u_1^H x, ..., u_n^H x] \\
&= \sum_{i=1}^{n} u_i^H x u_i \\
&= \sum_{i=1}^{n} \langle u_i, x \rangle u_i \\
&= \mathrm{proj}_{\mathscr{R}(A)} x
\end{aligned}$$

where n denotes the number of singular values of A. Therefore, $AA^{\dagger}$ is an orthogonal projection onto $\mathscr{R}(A)$ by definition.

(vi) This follows in similar fashion as part (v). Let $V_1 = [v_1, ..., v_n]$. So, V_1 is an orthonormal basis for $\mathscr{R}(A^H)$ by SVD. By part(iv), we have that

$$\begin{aligned}
A^{\dagger} A x &= V_1 V_1^H x \\
&= V_1[v_1^H x, ..., v_n^H x] \\
&= \sum_{i=1}^{n} v_i^H x v_i \\
&= \sum_{i=1}^{n} \langle v_i, x \rangle v_i \\
&= \mathrm{proj}_{\mathscr{R}(A^H)} x
\end{aligned}$$

where n again denotes the number of singular values of A. Therefore, $A^{\dagger}A$ is an orthogonal projection onto $\mathscr{R}(A^H)$ by definition.

$\square$