Math, Problem Set #3, Spectral Theory

OSM Lab, John Van den Berghe

Due Monday, July 10 at 8:00am

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

Solution From Problem Set 2 (assuming the basis is ordered as above), we recall that $D = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, and D is in upper-triangular form, which makes it easy to see that all the eigenvalues are 0. Thus, the algebraic multiplicity of 0 is 3.

However, the (non-generalized) eigenspace of 0 is only $span(\{1\})$, i.e. it has geometric multiplicity 1. This comes from the observation that D only has one eigenvector for eigenvalue 0. $\square$

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

(i) An Hermitian 2×2 matrix has only real eigenvalues.

Proof Note that for a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $det(A) = ad - bc$. From A being Hermitian, we know that $a = \overline{a}$, $d = \overline{d}$, $b = \overline{c}$, so a, d are real, and $bc = \overline{c}c = \|c\|^2$ is also real.

Then by 4.3, the characteristic polynomial has the form

$$p(\lambda) = \lambda^2 - tr(A)\lambda + det(A) = \lambda^2 - (a+d)\lambda + ad - \|c\|^2$$

Now the solutions to this equation are

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

Now note that $(a-d)^2 + \|c\|^2) \geq 0$ for all values, so that $\lambda_{\pm}$ must be real. $\square$

(ii) A Skew-Hermitian 2×2 matrix has only imaginary eigenvalues.

Proof As above, $det(A) = ad - bc$. However, A is now Skew-Hermitian, so $a = -\overline{a}$, $d = -\overline{d}$, $b = -\overline{c}$, so a, d are purely imaginary, and both $bc = -\overline{c}c = -\|c\|^2$ and ad are negative.

Then by 4.3, the characteristic polynomial has the form

$$p(\lambda) = \lambda^2 - tr(A)\lambda + det(A) = \lambda^2 - (a+d)\lambda + ad + \|c\|^2$$

Now the solutions to this equation are

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

Now note that each of the terms in $(a-d)^2 + \|c\|^2)$ is negative, so their sum is as well, so for all values of a, b, c, and d, $\lambda_{\pm} = \frac{(a+d) \pm \sqrt{(a-d)^2 + \|c\|^2)}}{2}$ must purely imaginary. $\square$

4.6 Prove that "The diagonal entries of an upper-triangular (or a lower-triangular matrix) are its eigenvalues.

Proof Assume the matrix A is upper-triangular, then the matrix $\lambda I - A$ will also be upper triangular, with the values of the diagonal being $\{\lambda - d_i\}_{i=1}^n$, where d_i is the ith diagonal entry of A.

We claim $p(\lambda) = \prod_{i=1}^n (\lambda - d_i)$.

Induction start: n = 1 If $A = d \in M_1(\mathbb{F})$ (a scalar and thus trivially upper-triangular), the characteristic polynomial is just $p(\lambda) = det(\lambda - A) = \lambda - d = \prod_{i=1}^1 (\lambda - d_i)$

Now suppose the results holds for all matrices up to size $A \in M_{n-1}(\mathbb{F})$, i.e. for all $k < n$, we have $p(\lambda) = \prod_{i=1}^{n-1} (\lambda - d_i)$.

Induction step: Suppose $A \in M_n(\mathbb{F})$, then by Thm 2.9.16,

$$det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} det(A_{ij}) \qquad \forall j \in \{1, 2, \cdots, n\},$$

where A_{ij} is the matrix obtained by removing row i and column j. Now observe that $a_{i1} = \begin{cases} d_1 & i = 1 \\ 0 & i \neq 1 \end{cases}$, and that $A_{11} \in M_{n-1}(\mathbb{F})$ is an upper-triangular matrix. Hence, by induction,

$$p(\lambda) = det(A) = \sum_{i=1}^n (-1)^{i+1} a_{i1} det(A_{i1}) = a_{11} det(A_{11})$$

$$= (\lambda - d_1)\prod_{i=1}^{n-1}(\lambda - (d_{A_{11}})_i) = (\lambda - d_1)\prod_{i=2}^{n}(\lambda - d_i) = \prod_{i=1}^{n}(\lambda - d_i),$$

We have shown by the principles of mathematical induction, that for an upper-triangular matrix, $p(\lambda) = det(A) = \prod_{i=1}^{n}(\lambda - d_i)$.$\square$

The proof for lower-triangular matrices follows the same procedure, where we now would choose the nth (instead of the first) column in the induction step.

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) Prove that S is a basis for V.

Proof From Problem Set 2, the set is orthonormal under the inner product $\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)dt$, i.e. all elements in the spanning set are independent, so they are a basis of the span.$\square$

(ii) Let D be the derivative operator. Write the matrix representation of D in the basis S.

Solution We have $Dsin(x) = cos(x), Dcos(x) = -sin(x), Dsin(2x) = 2cos(x)$, and $Dcos(2x) = -2sin(x)$. Hence $D = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 2 & 0 \end{bmatrix}$ $\square$

(iii) Find two complementary D-invariant subspaces in V.

Solution Without even needing to compute the eigenvalues and vectors, we can easily see that $span(\{sin(x), cos(x)\})$ and $span(\{sin(2x), cos(2x)\})$ are D-invariant. $\square$

4.13 Let $A = \begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix}$. Compute the transition matrix P such that $P^{-1}AP$ is diagonal.

Proof Observe from $det(\lambda I - A) = \lambda^2 - 1.4\lambda + 0.4$, we can see that the eigenvalues are 1 and 0.4 . As discussed in the notes, the transition matrix can be formed from the eigenvectors, i.e. $P = \begin{bmatrix} v_1 & v_{0.4} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$. (If the orthonormal P is desired, we would just normalize column 1 by $\sqrt{5}$, and column 2 by $\sqrt{2}$. $\square$

4.15 Prove: "If $(\lambda_i)_{i=1}^n$ are the eigenvalues of a semisimple matrix $A \in M_n(\mathbb{F})$ and $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ is a polynomial, then $(f(\lambda_i))_{i=1}^n$ are the eigenvalues of $f(A) = a_0 I + a_1 A + \cdots + a_n A^n$.

Proof By Theorem 4.3.7, we can diagonalize the matrix A as $P\Lambda P^{-1}$. Now note that $f(A) = a_0 I + a_1 A + \cdots + a_n A^n = a_0 PP^{-1} + a_1 P\Lambda P^{-1} + \cdots + a_n P\Lambda^n P^{-1} = Pf(\Lambda)P^{-1}$, but now note that every term in $f(\Lambda)$ is a diagonal matrix, so the diagonal entries are exactly $(f(\lambda_i))_{i=1}^n$. Since $f(\Lambda)$ is similar to $f(A)$, they have the same eigenvalues, namely $(f(\lambda_i))_{i=1}^n$. $\square$

4.16 Let $A = \begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix}$

(i) Compute $lim_{n\to\infty} A^n$ with respect to the 1-norm;

Computation Note that $A^n = P\Lambda^n P^{-1}$, where $\Lambda^n = \begin{bmatrix} 1^n & 0 \\ 0 & 0.4^n \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0.4^n \end{bmatrix}$, such that $A^k = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0.4^k \end{bmatrix} \frac{1}{3} \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 + 0.4^k & 2 - 2 * 0.4^k \\ 1 - 0.4^k & 1 + 2 * 0.4^k \end{bmatrix}$ with the limit $B = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$

Finally, note that $A^k - B = \frac{1}{3} \begin{bmatrix} 0.4^k & -2 * 0.4^k \\ -0.4^k & 2 * 0.4^k \end{bmatrix}$, which clearly converges with 1-norm since all the terms converge. $\square$

(ii) Repeat (i) for the ∞-norm and the Frobenius norm. Does the answer depend on the choice of norm?

Analysis For ∞-norm, nothing changes, as the largest entry goes to zero. For the Frobenius norm, $\|A^k - B\| = \frac{1}{3}\sqrt{tr(\begin{bmatrix} 0.4^k & -0.4^k \\ -2 * 0.4^k & 2 * 0.4^k \end{bmatrix} \begin{bmatrix} 0.4^k & -2 * 0.4^k \\ -0.4^k & 2 * 0.4^k \end{bmatrix})} = \frac{1}{3}\sqrt{tr(\begin{bmatrix} 2 * 0.4^{2k} & -4 * 0.4^{2k} \\ -4 * 0.4^{2k} & 8 * 0.4^{2k} \end{bmatrix})} = \sqrt{10 * 0.4^{2k}}$, which also clearly goes to zero. Hence the convergence does not depend on the norm. $\square$

(iii) Find all the eigenvalues of the matrix $3I + 5A + A^3$.

Proof By Theorem 4.3.12, the eigenvalues of $f(A)$ are $(f(\lambda_i))_{i=1}^n$, where the $(\lambda_i)_{i=1}^n$ are the original eigenvalues of A. Now since $f(x) = 3 + 5x + x^3$, and the original eigenvalues were 1 and 0.4, we have new eigenvalues $f(1) = 9$, and $f(0.4) = 5.064$. $\square$

4.18 Prove: If λ is an eigenvalue of the $A \in M_n(\mathbb{F})$, then there exists a nonzero row vector x^T such that $x^T A = \lambda x^T$.

Proof Note that $det(\lambda I - A) = det(\lambda I - A^T)$, so any λ that is an eigenvalue for A must also be an eigenvalue of A^T. Hence there exists v such that $A^T v = \lambda v$. Once again taking a transpose of this equation, we see that such v also satisfies $v^T A = \lambda v^T$.$\square$

4.20 Prove: If A is Hermitian and orthonormally similar to B, then B is also Hermitian.

Proof From orthonormal similarity, there is an orthonormal P such that $B = PAP^{-1}$. Now observe that

$$B^H = (PAP^{-1})^H = (PAP^H)^H = (P^H)^H A^H P^H = PAP^H = B,$$

since A is Hermitian, and P is orthonormal ($PP^H = I$). Hence B is Hermitian.$\square$

4.24 Given $A \in M_n(\mathbb{C})$, define the *Rayleigh quotient* as

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

where $\langle \cdot, \cdot \rangle$ is the usual inner product on $\mathbb{F}^n$. Show that the Rayleigh quotient can only take on real values for Hermitian operators and only imaginary values for skew-Hermitian matrices.

Proof Since $\|x\|^2 \in \mathbb{R}$, we only need to show the statement for $\langle x, Ax \rangle = x^H Ax$. Now note that

$$\overline{x^H Ax} = (x^H Ax)^H = x^H A^H x = \begin{cases} x^H Ax & , A \text{ Hermitian} \\ -x^H Ax & , A \text{ Skew-Hermitian} \end{cases}$$

This shows the statement, since in the first case $\langle x, Ax \rangle$ must be purely real, and in the second, purely imaginary.$\square$

4.25 Let $A \in M_n(\mathbb{C})$ be a normal matrix with eigenvalues $(\lambda_i)_{i=1}^n$, and corresponding orthonormal eigenvectors $\{x_i\}_{i=1}^n$

(i) Show that the identity matrix can be written $I = \sum_{i=1}^n x_i x_i^H$.

Proof We let $\sum_{i=1}^n x_i x_i^H$ act on an arbitrary vector. Observe that any vector can be written as $\sum_{i=1}^n \alpha_i x_i$, and recall that by orthonormality, $x_j^H x_i = 0$ if $i \neq j$. So $(\sum_{i=1}^n x_i x_i^H)(\sum_{i=1}^n \alpha_i x_i) = \sum_{i=1}^n x_i x_i^H x_i \alpha_i = \sum_{i=1}^n x_i \alpha_i = \sum_{i=1}^n \alpha_i x_i = v$. Hence $(\sum_{i=1}^n x_i x_i^H)v = v$, i.e. $\sum_{i=1}^n x_i x_i^H = I$.$\square$

(ii) Show that A can be written as $A = \sum_{i=1}^n \lambda_i x_i x_i^H$

Proof $Av = A(\sum_{i=1}^n \alpha_i x_i) = \sum_{i=1}^n \alpha_i A x_i = \sum_{i=1}^n \alpha_i \lambda_i x_i$

and

$(\sum_{i=1}^n \lambda_i x_i x_i^H)v = (\sum_{i=1}^n \lambda_i x_i x_i^H)(\sum_{i=1}^n \alpha_i x_i) = \sum_{i=1}^n \lambda_i x_i x_i^H x_i \alpha_i = \sum_{i=1}^n \alpha_i \lambda_i x_i$

so $A = \sum_{i=1}^n \lambda_i x_i x_i^H$ $\square$

4.27 Assume $A \in M_n(\mathbb{F})$ is positive definite. Prove that all its diagonal entries are real and positive.

Proof Take the e_i from the orthonormal basis the matrix is written in. Then since A is positive definite, $0 < e_i^H A e_i = a_{ii}$, where real-valuedness was clear by definition 4.5.1. $\square$

4.28 Assume $A, B \in M_n(\mathbb{F})$ are positive semidefinite. Prove that $0 \leq tr(AB) \leq tr(A)tr(B)$, and use this result to prove that $\|\cdot\|_F$ is a matrix norm.

Proof Note that by positive-semidefiniteness, there exist matrices such that $S_A^H S_A = A$ and $S_B^H S_B = B$. Now observe that $tr(AB) = tr(S_A^H S_A S_B^H S_B) = tr(S_B S_A^H S_A S_B^H) = tr((S_A S_B^H)^H S_A S_B^H) \geq 0$, since $S_A S_B^H$ is Hermitian and the multiplication of positive eigenvalues yields positive eigenvalues.

Note next that we can also diagonalize A and B as $A = P_A D_A P_A^{-1}$, and recall observe

$$tr(A) = tr(P_A D_A P_A^{-1}) = tr(P_A^{-1} P_A D_A) = tr(D_A) = \sum_i \lambda_{Ai}.$$

Now note that by invariance of the trace under even permutations,

$$tr(AB) = tr(P_A D_A P_A^{-1} P_B D_B P_B^{-1}) = tr(P_A P_A^{-1} P_B D_A D_B P_B^{-1}) = tr(P_B^{-1} P_B D_A D_B)$$

$$= tr(D_A D_B) = \sum_i \lambda_{Ai}\lambda_{Bi} \leq (\sum_i \lambda_{Ai})(\sum_i \lambda_{Bi}) = tr(A)tr(B)$$

To confirm that $\|\cdot\|_F$ is a matrix norm, we check the three conditions as in Problem Set 2. Firstly, clearly $\|A\| = \sqrt{tr(A^H A)} \geq 0$ with equality only if all diagonal entries of $A^H A$, i.e. all singular values of A being 0, which can only occur for the zero matrix. Next, we can see that

$$\|\alpha A\| = \sqrt{tr(\alpha^H A^H A \alpha} = \alpha \sqrt{tr(A^H A} = \alpha \|A\|$$

Finally, we still need to check whether the triangle inequality holds.

$$\|A + B\|_F^2 = tr((A + B)^H (A + B)) = tr(A^H A + B^H B + A^H B + A^H B)$$

$$= tr(A^H A) + tr(B^H B) + tr(A^H B + A^H B) \leq tr(A^H A) + tr(B^H B) + 2\|A\|\|B\|$$

$$= \|A\|^2 + \|B\|^2 + 2\|A\|\|B\| = (\|A\| + \|B\|)^2$$

$\square$

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

(i) $\|A\|_2 = \sigma_1$, where σ_1 is the largest singular value of A.

Proof Recall that if $A^H A$ is normal, there is an orthonormal set of eigenvectors $\{v_i\}_{i=1}^n$ with respective eigenvalues $\{\sigma_i^2\}_{i=1}^n$.

$$\|A\|_2 = sup_{\|x\|=1}\|Ax\| = sup_{\|x\|=1}\sqrt{\langle Ax, Ax\rangle} = sup_{\|x\|=1}\sqrt{\langle x, A^H Ax\rangle}$$

$$= sup_{\|x\|=1}\sqrt{\langle(\sum_{i=1}^n \alpha_i v_i), (\sum_{i=1}^n \alpha_i \sigma_i^2 v_i)\rangle} = sup_{\|x\|=1}\sqrt{\langle(\sum_{i=1}^n \alpha_i v_i), (\sum_{i=1}^n \alpha_i \sigma_i^2 v_i)\rangle}$$

$= sup_{\|x\|=1}\sqrt{\sum_{i=1}^n |\alpha_i|^2\sigma_i^2} = \sqrt{\sigma_1^2} = \sigma_1$ by choosing $x = v_1$. $\square$

(ii) If A is invertible, then $\|A^{-1}\|_2 = \sigma_n^{-1}$.

Proof Note that if $Av = \lambda v$, then $v = A^{-1}\lambda v$, i.e. $A^{-1}v = \frac{1}{\lambda}v$

$$\|A^{-1}\|_2 = sup_{\|x\|=1}\|A^{-1}x\| = sup_{\|x\|=1}\sqrt{\langle A^{-1}x, A^{-1}x\rangle} = sup_{\|x\|=1}\sqrt{\langle x, (A^H A)^{-1}x\rangle}$$

and now note that the largest eigenvalue is the square of the inverse of the smallest singular value of A, such that as before

$$= sup_{\|x\|=1}\sqrt{\sum_{i=1}^n |\alpha_i|^2 \frac{1}{\sigma_i^2}} = \sigma_n^{-1}$$

by choosing maximally $\alpha_n = 1$, $x = v_n$. $\square$

(iii) $\|A^H\|_2^2 = \|A^T\|_2^2 = \|A^H A\|_2 = \|A\|_2^2$.

Proof By singular value decomposition, $A = U\Sigma V^H$, where both U and V are orthonormal. Then we have

$$A^H = (U\Sigma V^H)^H = V\Sigma^H U^H = V\Sigma U^H,$$

$$A^T = (U\Sigma V^H)^T = \overline{V}\Sigma^T U^T = \overline{V}\Sigma U^T,$$

where both $\overline{V}$ and $U.T$ can still be checked to be orthonormal, and

$$A^H A = (U\Sigma V^H)^H U\Sigma V^H = V\Sigma^H U^H U\Sigma V^H = V\Sigma^H U^H U\Sigma V^H = V\Sigma'^2 V^H$$

where Σ' is the diagonal square matrix with just the singular values. Note that from the first two observations, and (iv) below, we have $\|A^H\|_2^2 = \|A^T\|_2^2 = \|A\|_2^2$. Finally,

$$\|A^H A\|_2 = \|V\Sigma'^2 V^H\| = sup_{\|x\|=1}\|V\Sigma'^2 V^H x\| = sup_{\|V^H x\|=1}\|V^H V\Sigma'^2 V^H x\|$$

$$= sup_{\|x'\|=1}\|\Sigma^2 x'\| = sup_{\|x\|=1}\sqrt{\sum_{i=1}^{n}|\alpha_i|^2\sigma_i^4} = \sigma_1^2 = \|A\|_2^2$$

$\square$

(iv) If $U \in M_m(\mathbb{F})$ and $V \in M_n(\mathbb{F})$ are orthonormal, then $\|UAV\|_2 = \|A\|_2$.

Proof Orthonormality of W implies that $W^{-1} = W^H$, and W has full rank, i.e. is both injective and surjective. Note that $\|UAV\|_2 = sup_{\|x\|=1}\sqrt{\langle UAVx, UAVx\rangle} = sup_{\|x\|=1}\sqrt{\langle AVx, U^H UAVx\rangle} = sup_{\|x\|=1}\sqrt{\langle AVx, AVx\rangle} = sup_{\|x'\|=1}\sqrt{\langle Ax', Ax'\rangle} = \|A\|_2$ $\square$

4.32 Assume $A \in M_{m\times n}(\mathbb{F})$ is of rank r. Prove that

(i) If $U \in M_m(\mathbb{F})$ and $V \in M_n(\mathbb{F})$ are orthonormal, then $\|UAV\|_F = \|A\|_F$.

Proof With the Frobenius norm, and invariance of the trace under even permutations, we have $\|UAV\|_F = \sqrt{tr(V^H A^H U^H UAV)} = \sqrt{tr(V^H A^H AV)} = \sqrt{tr(A^H AVV^H)} = \sqrt{tr(A^H A)} = \|A\|_F$ $\square$

(ii) $\|A\|_F = (\sum_{i=1}^n \sigma_i^2)^{\frac{1}{2}}$, where $\sigma_1 \geq \sigma_2 \geq \cdot \geq \sigma_r > 0$ are the singular values of A.

Proof From (i), and SVD, we have that $\|A\|_F = \|U\Sigma V^H\|_F = \|\Sigma\|_F = \sqrt{tr(\Sigma^H\Sigma)} = (\sum_{i=1}^n \sigma_i^2)^{\frac{1}{2}}$ $\square$

4.33 Assume $A \in M_n(\mathbb{F})$. Prove that

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

Proof From 4.32, we can easily see the definitions are equivalent, since $\sup_{\|x\|_2=\|y\|_2=1} |y^H Ax| = \sup_{\|x\|_2=\|y\|_2=1} |y^H U\Sigma V^H x| = \sup_{\|x'\|_2=\|y'\|_2=1} |y'^H \Sigma x'| = \sigma_1 = \|A\|_2$ where we simply observed that the maximum would be achieved by $x' = y' = e_1$ in standard bases.$\square$

4.36 Give an example of a 2×2 matrix whose determinant is nonzero and whose singular values are not equal to any of its eigenvalues.

Example Most simply, take a 2×2 matrix with values only on the off-diagonal: $\begin{bmatrix} 0 & \sigma_1 \\ \sigma_2 & 0 \end{bmatrix}$ with $\sigma_1 > \sigma_2 > 0$, then one possible SVD is $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, but the eigenvalues of A are $\lambda_\pm = \pm\sqrt{\sigma_1\sigma_2}$, which are clearly not equal to the singular values.$\square$

4.38 Prove: If $A \in M_{m\times n}(\mathbb{F})$, then the Moore-Penrose pseudoinverse of A satisfies:

(i) $AA^\dagger A = A$

Proof Note first that since now U, V only have full column rank, $U^H U = I = V^H V$, but $VV^H \neq I$. Observe $A^\dagger A = V_1\Sigma_1^{-1}U_1^H U_1 \Sigma_1 V_1^H = V_1\Sigma_1^{-1}\Sigma_1 V_1^H = V_1 V_1^H$, so $AA^\dagger A = AV_1V_1^H = U_1\Sigma_1 V_1^H V_1 V_1^H = U_1\Sigma_1 V_1^H = A$ □

(ii) $A^\dagger A A^\dagger = A^\dagger$

Proof From the observation in (i), $A^\dagger A A^\dagger = V_1 V_1^H V_1 \Sigma_1^{-1} U_1^H = V_1\Sigma_1^{-1}U_1^H = A^\dagger$. □

(iii) $(AA^\dagger)^H = AA^\dagger$

Proof $(AA^\dagger)^H = (U_1U_1^H)^H = (U_1^H)^H U_1^H = U_1 U_1^H = AA^\dagger$ □

(iv) $(A^\dagger A)^H = A^\dagger A$.

Proof $(A^\dagger A)^H = (V_1V_1^H)^H = (V_1^H)^H V_1^H = V_1V_1^H = A^\dagger A$ □

(v) $AA^\dagger = proj_{\mathscr{R}(A)}$ is the orthogonal projection onto $\mathscr{R}(A)$.

Proof We begin by showing orthogonality to $\mathscr{R}(A)$. Let x be any vector, and compute

$$\begin{aligned}\langle Ax, x - AA^\dagger x\rangle &= \langle x, (A^H - A^H AA^\dagger)x\rangle = \langle x, (A^H - V_1\Sigma_1 U_1^H U_1 U_1^H)x\rangle \\ &= \langle x, (A^H - V_1\Sigma_1 U_1^H)x\rangle = \langle x, (A^H - A^H)x\rangle = \langle x, 0\rangle = 0\end{aligned}$$

Having shown orthogonality, we observe that $AA^\dagger$ is a projection since $(AA^\dagger)^2 = AA^\dagger$ by (i). The mapping is clearly at least onto $\mathscr{R}(A)$, since $x \in \mathscr{R}(A)$ remain fixed by the mapping, as seen in (i). Finally, the projection does not project on a larger space, which can be seen by finite dimensionality arguments: $dim(\mathscr{R}(A)) = rank(U_1U_1^H) = rank(AA^\dagger)$. □

(vi) $A^\dagger A = proj_{\mathscr{R}(A^H)}$ is the orthogonal projection onto $\mathscr{R}(A^H)$.

Proof Observe that $A^H(A^H)^\dagger = V_1\Sigma_1 U_1^H U_1 (\Sigma_1^H)^{-1} V_1^H = V_1\Sigma_1\Sigma_1^{-1}V_1^H = V_1V_1^H = A^\dagger A$, so we can just apply (v) with A^H in the place of A. □