Math Problem Set #3: Spectral Theory

OSM Lab - University of Chicago
Geoffrey Kocks

Exercise 1: HJE 4.2.

Solution. With respect to the power basis, we seek a D such that:

$$D(1,0,0) = (0,0,0).$$

$$D(0,1,0) = (1,0,0).$$

$$D(0,0,1) = (0,2,0).$$

This is satisfied by:

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

The eigenvalues are then given by the values of λ such that $|D - \lambda I| = 0$. Thus:

$$D - \lambda I = \begin{pmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 2 \\ 0 & 0 & -\lambda \end{pmatrix}$$

Because $D - \lambda I$ is an upper triangular matrix, its determinant is given by the product of its diagonal entries, so: $|D - \lambda I| = -\lambda^3$. Setting this equal to 0 gives eigenvalues of $\lambda = 0$. This has an algebraic multiplicity of 3. The eigenspace is given by $\Sigma_0(D) = (\{a, 0, 0\})$ for all $a \in \mathbb{R}$. This has a geometric multiplicity of 1.

Exercise 2: HJE 4.4.

Solution.

(i.) Let the 2x2 matrix A be given by:

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

The matrix A^H is then given by:

$$A^H = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

For a Hermitian matrix, $A = A^H$ which occurs if and only if $b = c$. We can therefore rewrite a Hermitian matrix A as:

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$$

Matrix A then has properties $tr(A) = a + d$ and $\det(A) = ad - b^2$. From Exercise 4.3, the characteristic polynomial of any 2x2 matrix has the form $p(\lambda) = \lambda^2 - tr(A)\lambda +$

$det(A)$ so there will be real eigenvalues if and only if $(tr(A))^2 - 4\det(A) \geq 0$. For Hermitian matrix A:

$$(tr(A))^2 - 4\det(A) = (a+d)^2 - 4(ad - bc) = a^2 + d^2 - 2ad + 4bc = (a-d)^2 + 4b^2.$$

Because this is a sum of squares, it will always be at least 0, and therefore the matrix will only have real eigenvalues.

(ii.) If matrix A is skew-Hermitian, then $A^H - A$. This will only occur when A has the form:

$$A = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}$$

For skew-Hermitian matrix A:

$$(tr(A))^2 - 4\det(A) = 0 - 4(0 + b^2) = -4b^2 < 0.$$

This value will always be less than 0 if $b \neq 0$, so the matrix will only have imaginary eigenvalues.

Exercise 3: HJE 4.6.

Solution. Let A be an upper triangular square matrix. Then $A - \lambda I$ is also an upper triangular square matrix with diagonal entries each reduced by λ. For any upper triangular square matrix of size n, the determinant is the product of the diagonal entries $\{d_1, d_2, ..., d_n\}$. Therefore, the characteristic polynomial of $A - \lambda I$ is given by:

$$\|A - \lambda I\| = \Pi(d_i - \lambda) = 0.$$

This will have a solution given by $\lambda = d_i$ for each d_i when calculating the eigenvalues.

Therefore we have shown Proposition 4.1.22 that the diagonal entries of an upper-triangular (or lower-triangular) matrix are its eigenvalues.

Exercise 4: HJE 4.8.

Solution.

(i.) In Exercise 3.8 we showed that this same set S is orthonormal. Because it is orthonormal, each element must be linearly independent. Therefore, the linearly independent elements make up a basis of its span.

(ii.) The derivative operator D in the basis is given by:

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

(iii.) The two complementary D-invariant subspaces are $\{\sin(x), \cos(x)\}$ and $\{\sin(2x), \cos(2x)\}$.

Exercise 5: HJE 4.13.

Solution. Let

$$A = \begin{pmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{pmatrix}$$

Then its eigenvalues λ satisfy the characteristic polynomial:

$$(0.8 - \lambda)(0.6 - \lambda) - 0.08 = 0 \implies (\lambda - 1)(\lambda - 0.4) = 0.$$

Thus there are two eigenvalues: $\lambda_1 = 1$ and $\lambda_2 = 0.4$. The associated eigenvector for $\lambda_1 = 1$ is given by:

$$\mathbf{x_1} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

since $A\mathbf{x_1} = \mathbf{x_1}$. The associated eigenvector for $\lambda_2 = 0.4$ is given by:

$$\mathbf{x_2} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

since $A\mathbf{x_2} = 0.4\mathbf{x_2}$. Therefore the matrix P is given by:

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

This results in the diagonal matrix:

$$P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & 0.4 \end{pmatrix}$$

Exercise 6: HJE 4.15.

Solution. Let $(\lambda_i)_{i=1}^n$ be the eigenvalues of a semisimple matrix A and let $f(x)$ be a polynomial $a_0 + a_1 x + ... + a_n x^n$. It was proved in the lecture notes that if an n x n matrix is semisimple, then it is diagonizable such that $A = PDP^{-1}$ where the diagonal entries of the diagonal matrix D are the eigenvalues of A.

$$\begin{aligned} f(A) = f(PDP^{-1}) &= a_0 + a_1(PDP^{-1}) + ... + a_n(PDP^{-1})^n \\ &= a_0 PP^{-1} + a_1 PDP^{-1} + ... + a_n PD^n P^{-1} \\ &= P(a_0 + a_1 D + ... + a_n D^n)P^{-1} \\ &= P(f(D))P^{-1} \end{aligned}$$

$f(D)$ will also be a diagonal matrix, which each of the diagonal entries being an eigenvalue of the matrix $f(A)$ by the theorem that a matrix is semisimple if it is diagonizable. These diagonal entries will exactly be given by $f(\lambda_i)_{i=1}^n$.

Therefore $f(\lambda_i)_{i=1}^n$ are the eigenvalues of $f(A)$.

Exercise 7: HJE 4.16.

Solution.

(i). Let

$$A = \begin{pmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{pmatrix}$$

We note that by Proposition 4.3.10, if matrices A and B are similar with $A = P^{-1}BP$, then $A^k = P^{-1}B^kP$. Given the diagonalization of A in Exercise 4.13, this means that:

$$A^k = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1^k & 0 \\ 0 & 0.4^k \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 \\ -1/3 & 2/3 \end{pmatrix}$$

$$A^k = \begin{pmatrix} 2 & -(.4)^k \\ -.4^k & 2(.4)^k \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 \\ -1/3 & 2/3 \end{pmatrix} = \begin{pmatrix} 1/3(2+.4^k) & 1/3(2-4(.4^k)) \\ 1/3(1-.4^k) & 1/3(1+2(.4^k)) \end{pmatrix}$$

Then the limit of A^n will be given by the matrix B:

$$B = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 \\ -1/3 & 2/3 \end{pmatrix} = \begin{pmatrix} 2/3 & 2/3 \\ -1/3 & 2/3 \end{pmatrix}$$

Taking the induced 1-norm:

$$A^k - B = \begin{pmatrix} (1/3).4^k & (-4/3)(.4^k) \\ (-1/3).4^k & (2/3)(.4^k) \end{pmatrix}$$

$\|A^k - B\|_1 = 2(.4)^k < \epsilon \implies k < \frac{\ln \epsilon/2}{\ln .4}$. Therefore the limit exists for the 1-norm.

(ii). Taking the ∞-norm on the matrix $A^k - B$:

$$\|A^k - B\|_\infty = \frac{5}{3}(.4)^k < \epsilon \implies k < \frac{\ln 3\epsilon/5}{\ln .4}$$

Taking the Frobenius norm on the matrix $A^k - B$:

$$\|A^k - B\|_{Fr} = \sqrt{\langle A^k - B, A^k - B\rangle} = \sqrt{tr((A^k - B)^H(A^k - B))} = \frac{22}{9}(.4)^k.$$

$$\frac{22}{9}(.4)^k < \epsilon \implies k < \frac{\ln 9\epsilon/22}{\ln .4}$$

Thus, the matrix will still converge on the ∞-norm and Frobenius norm but it will impact the rate of convergence.

(iii). By Theorem 4.3.13 (Semisimple Spectral Mapping), if $(\lambda_i)_{i=1}^n$ are the eigenvalues of a semisimple matrix A then if $f(x) = 3+5x+x^3$, the matrix $3I+5A+A^3$ will have eigenvalues $(f(\lambda_i))_{i=1}^n$. We note that matrix A used throughout this problem is a semisimple matrix because it is a simple matrix (which is a stronger condition). The matrix is simple because its eigenvalues are distinct. Therefore the two eigenvalues of $3I + 5A + A^3$ are:

$$f(1) = 3 + 5 + 1 = 9$$

$$f(0.4) = 3 + 2 + .064 = 5.064.$$

Exercise 8: HJE 4.18.

Solution. If λ is an eigenvalue of A, then by definition there exists $\mathbf{y}$ such that $A\mathbf{y} = \lambda\mathbf{y}$. This implies that:

$$(A\mathbf{y})^T = (\lambda\mathbf{y})^T \implies (\mathbf{y}^T A^T) = \lambda\mathbf{y}^T \implies \mathbf{y}^T(\lambda I - A^T) = 0.$$

λ is then an eigenvalue of A^T so there exists another vector $\mathbf{x}$ such that:

$$(A^T\mathbf{x}) = \lambda\mathbf{x} \implies (\mathbf{x}^T A) = \lambda\mathbf{x}^T.$$

Therefore we have shown that if λ is an eigenvalue of A, then there exists a nonzero row vector $\mathbf{x}^T$ such that $\mathbf{x}^T A = \lambda\mathbf{x}^T$.

Exercise 9: HJE 4.20.

Solution. Let A be a Hermitian matrix such that $A^H = A$. Let matrices A and B be orthonormally similar, so that there is an orthonormal matrix U such that $B = U^H AU$. Then:

$$B^H = (U^H AU)^H = U^H A^H U = U^H AU = B.$$

$B^H = B$ so B is also a Hermitian matrix. Therefore we have proven Lemma 4.4.2 that if A is Hermitian and orthonormally similar to B, then B is also Hermitian.

Exercise 10: HJE 4.24.

Solution. Let matrix A be Hermitian. Then $A^H = A$. The numerator of the Rayleigh quotient will then be given by:

$$\langle \mathbf{x}, A\mathbf{x}\rangle = \mathbf{x}^H A\mathbf{x} = \mathbf{x}^H A^H\mathbf{x} = \langle A\mathbf{x}, \mathbf{x}\rangle = \overline{\langle \mathbf{x}, A\mathbf{x}\rangle}$$

Because the numerator is equal to its own conjugate, this means that the numerator will be real. The denominator of the Rayleigh quotient will always be real. Therefore the Rayleigh quotient can only take on real values for Hermitian matrices.

Now let matrix B be skew-Hermitian so that $A^H = -A$. The numerator of the Rayleigh quotient will then be given by:

$$\langle \mathbf{x}, B\mathbf{x}\rangle = \mathbf{x}^H B\mathbf{x} = -\mathbf{x}^H B^H\mathbf{x} = -\langle B\mathbf{x}, \mathbf{x}\rangle = -\overline{\langle \mathbf{x}, B\mathbf{x}\rangle}$$

Because the numerator is equal to the negative of its conjugate, this means that the numerator will only be imaginary. Therefore the Rayleigh quotient can only take on imaginary values for skew-Hermitian matrices.

Exercise 11: HJE 4.25.

Solution.

(i.) Because Matrix A, is normal that means that its eigenvectors form an orthonormal matrix. Denote this matrix B whose columns are the eigenvectors $[x_1, ..., x_n]$ of A. B is also orthonormal so $BB^H = B^H B = I$ By definition this product is equivalent to:

$$BB^H = x_1x_1^H + ... + x_nx_n^H = I.$$

(ii.) We note that for any $\mathbf{x_j}$ in the group of eigenvectors:

$$(\lambda_1 \mathbf{x_1}\mathbf{x_1}^H + ... + \lambda_n \mathbf{x}_n \mathbf{x}_n^H)\mathbf{x}_j = \lambda_j \mathbf{x}_j \mathbf{x}_j^H \mathbf{x}_j = \lambda_j \mathbf{x}_j = A\mathbf{x}_j.$$

Therefore: $A = (\lambda_1 \mathbf{x_1}\mathbf{x_1}^H + ... + \lambda_n \mathbf{x}_n \mathbf{x}_n^H)\mathbf{x}_j$.

Exercise 12: HJE 4.27.

Solution. If matrix A is positive definite, then $\langle \mathbf{x}, A\mathbf{x} \rangle > 0$ for all $\mathbf{x} \neq 0$. This means that $\mathbf{x}^H A\mathbf{x} > 0$. We can take as $\mathbf{x}$ each $\mathbf{e}_{i=1}^n$ where n is the number of diagonal entries in the matrix and $\mathbf{e}_i$ is a column vector that has a 1 in row i and a 0 everywhere else. Taking the product $\mathbf{e}_i^H A\mathbf{e}_i$ will give the i-th diagonal of the matrix and this product has to be nonzero. Therefore, all of the diagonal entries of the positive definite matrix are real and positive.

Exercise 13: HJE 4.28.

Solution. Assume that A, B are positive semi-definite. Then by Proposition 4.5.7 we can write them as $A = S^H S$ and $B = T^H T$. Then:

$$tr(AB) = tr(S^H S T^H T) = tr(S T^H T S^H) = tr(QQ^H)$$

where $Q = ST^H$. The diagonal entries of the matrix formed from multiplying a matrix by its Hermitian will always be positive because they are the sum of squared terms. Thus $tr(QQ^H) \geq 0$ and $tr(AB) \geq 0$.

Next, by the Cauchy-Schwartz inequality: $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|\|\mathbf{y}\|$ so $tr(AB) \leq \sqrt{tr(A^2)tr(B^2)}$. We note that for a positive semi-definite matrix A and B: $tr(A^2) \leq (tr(A))^2$ and $tr(B^2) \leq (tr(B))^2$. In order to show this, we use the result that the sum of the eigenvalues for a square matrix is the same as its trace. Therefore for A (with an identical argument for B):

$$(tr(A^2)) = \Sigma_i \lambda_i^2 \leq (\Sigma_i \lambda_i)^2 = (tr(A))^2.$$

Thus $tr(AB) \leq tr(A)tr(B)$ and we have shown that $0 \leq tr(AB) \leq tr(A)tr(B)$.

Finally, we show that the Frobenius norm is a matrix norm. In order to be a matrix norm, it must satisfy positivity, scale preservation, and the triangle inequality. We define the Frobenius norm as $\|A\|_F = \sqrt{tr(A^H A)}$. Then positivity follows directly by the definition of a square root, and the square root will only equal 0 if $A = 0$. Scale preservation also follows directly, as:

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

We use the result from the first part of the problem to prove that the triangle inequality holds:

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

Therefore the triangle equality holds for the norm when A and B are positive semi-definite so the Frobenius norm is a matrix norm.

Exercise 14: HJE 4.31.

Solution.

(i.) Assume that A not identically 0 has an SVD decomposition given by $U\Sigma V^H$. Then its induced 2-norm is given by:

$$\sup_{x\neq 0}\frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2}=\sup_{x\neq 0}\frac{\|U\Sigma V^H\mathbf{x}\|_2}{\|\mathbf{x}\|_2}=\sup_{x\neq 0}\frac{\|\Sigma V^H\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$$

since U is an orthonormal matrix and using the result of part (iv) of this problem. Define $\mathbf{y}=V^H\mathbf{x}$. Then the induced 2-norm above is equivalent to:

$$\sup_{y\neq 0}\frac{\|\Sigma\mathbf{y}\|_2}{\|V\mathbf{y}\|_2}=\sup_{y\neq 0}\frac{\|\Sigma\mathbf{y}\|_2}{\|\mathbf{y}\|_2}=\sup_{y\neq 0}\frac{\sqrt{\sum_{i=1}^n\sigma_i^2|y_i|^2}}{\sqrt{\sum_{i=1}^n|y_i|^2}}$$

This supremum is equivalent to finding the max of $\|\Sigma\mathbf{y}\|$ when $\|\mathbf{y}\|=1$ by the properties of the 2-norm, and this maximum will occur exactly when $\mathbf{y}=[1,0,...0]^T$. Therefore $\|A\|_2=\sigma_1$ where σ_1 is the largest singular value of A.

(ii.) Assume that A has SVD decomposition given by $U\Sigma V^H$. Then the induced 2-norm of A^{-1} is given by:

$$\|A^{-1}\|_2=\sup_{x\neq 0}\frac{\|A^{-1}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}=\sup_{x\neq 0}\frac{\|(V^H)^{-1}\Sigma^{-1}U^{-1}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}=\sup_{x\neq 0}\frac{\|\Sigma^{-1}U^{-1}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}.$$

Define $\mathbf{y}=U^{-1}\mathbf{x}$. Then the induced 2-norm above is equivalent to:

$$\sup_{y\neq 0}\frac{\|\Sigma^{-1}\mathbf{y}\|_2}{\|U\mathbf{y}\|_2}=\sup_{y\neq 0}\frac{\|\Sigma^{-1}\mathbf{y}\|_2}{\|\mathbf{y}\|_2}=\sup_{y\neq 0}\frac{\sqrt{\sum_{i=1}^n(1/\sigma_i)^2|y_i|^2}}{\sqrt{\sum_{i=1}^n|y_i|^2}}$$

By the same reasoning as in part (ii), we will obtain the supremum when $\mathbf{y}=[0,0,...,1]^T$. Therefore $\|A^{-1}\|_2=\sigma_n^{-1}$.

(iii.) We note from part (i) of this problem that $\|A\|_2^2=\sigma_1^2$ where σ_1 is the largest singular value of A. First: $A^HA=V\Sigma^HU^HU\Sigma V^H=V\Sigma^H\Sigma V^H=V\Sigma^2V^H$. Then:

$$\|A^HA\|_2=\|V\Sigma^2V^H\|_2=\sup_{x\neq 0}\frac{\|V\Sigma^2V^H\mathbf{x}\|_2}{\|\mathbf{x}\|_2}=\sup_{x\neq 0}\frac{\|\Sigma^2V^H\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$$

Define $\mathbf{y}=V^H\mathbf{x}$. Then the induced 2-norm above is equivalent to:

$$\sup_{y\neq 0}\frac{\|\Sigma^2\mathbf{y}\|_2}{\|V\mathbf{y}\|_2}=\sup_{y\neq 0}\frac{\|\Sigma^2\mathbf{y}\|_2}{\|\mathbf{y}\|_2}=\sup_{y\neq 0}\frac{\sqrt{\sum_{i=1}^n(\sigma_i^2)^2|y_i|^2}}{\sqrt{\sum_{i=1}^n|y_i|^2}}$$

By the same reasoning as in part (i), we will obtain the supremum when $\mathbf{y}=[1,0,..,0]^T$. Therefore $\|A^HA\|_2=\|A\|_2^2$.

Next, we note that by Theorem 4.5.10, the singular values of a matrix A are all positive real numbers. Therefore $A^H = A^T$ so $\|A^H\|2^2 = \|A^T\|_2^2$.

Finally, we note that $A^H = V\Sigma U^H$ so A^H has the induced 2-norm:

$$\|A^H\|_2 = \sup_{x\neq 0} \frac{\|A^H\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sup_{x\neq 0} \frac{\|V\Sigma U^H\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sup_{x\neq 0} \frac{\|\Sigma U^H\mathbf{x}\|_2}{\|\mathbf{x}\|_2}.$$

Define $\mathbf{y} = U^H\mathbf{x}$. Then this supremum is equivalent to:

$$\sup_{y\neq 0} \frac{\|\Sigma\mathbf{y}\|_2}{\|U\mathbf{y}\|_2} = \sup_{y\neq 0} \frac{\|\Sigma\mathbf{y}\|_2}{\|\mathbf{y}\|_2} = \sup_{y\neq 0} \frac{\sqrt{\sum_{i=1}^n (\sigma_i^2)|y_i|^2}}{\sqrt{\sum_{i=1}^n |y_i|^2}}.$$

As in part (i), this will obtain the supremum when $\mathbf{y} = [1, 0, ..., 0]^T$. Therefore $\|A^H\|_2 = \sigma_1$ and $\|A^H\|_2^2 = \|A\|_2^2$.

Therefore we have shown the four way equality:

$$\|A^H\|_2^2 = \|A^T\|_2^2 = \|A^H A\|_2 = \|A\|_2^2.$$

(iv.) Since U and V are both orthonomal, $U^{-1} = U^H$ and $V^{-1} = V^H$. Then for the 2-norm:

$$\begin{aligned} \|UAV\| &= \sqrt{\langle UAV, UAV\rangle} = \sqrt{\langle U^H UAV, AV\rangle} \\ &= \sqrt{\langle AV, AV\rangle} = \sqrt{\langle A, AVV^H\rangle} \\ &= \|A\|. \end{aligned}$$

Therefore $\|UAV\|_2 = \|A\|_2$.

Exercise 15: HJE 4.32.

Solution.

(i.) Since U and V are both orthonomal, $U^{-1} = U^H$ and $V^{-1} = V^H$. With respect to the Frobenius norm:

$$\|UAV\|_F = \sqrt{tr(V^H A^H U^H UAV)} = \sqrt{tr(V^H A^H AV)} = \sqrt{tr(A^H AV^H V)} = \|A\|_F.$$

Therefore $\|UAV\|_F = \|A\|_F$.

(ii.) Let matrix A have the decomposition $U\Sigma V^H$. Then the induced Frobenius norm is given by:

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

Exercise 16: HJE 4.33.

Solution. In part (i) of Exercise 4.31, we showed that $\|A\|_2 = \sigma_1$ where σ_1 is the largest singular value of A. Therefore

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

We show the equivalence as follows:

$$\sup_{\|x\|_2=1,\|y\|_2=1} |\mathbf{y}^H A\mathbf{x}| = \sup_{\|x\|_2=1,\|y\|_2=1} |\langle \mathbf{y}, A\mathbf{x}\rangle| = \sup_{\|x\|_2=1,\|y\|_2=1} |\langle \mathbf{y}, \Sigma\mathbf{x}\rangle|$$
$$\leq \sup_{\|x\|_2=1,\|y\|_2=1} \|\mathbf{y}\|_2\|\Sigma\mathbf{x}\|_2 = \sup_{\|x\|_2=1} \|\Sigma\mathbf{x}\|_2 = \sigma_1$$

We note that this holds as an equality when $\mathbf{y}$ and $\mathbf{x}$ both equal the corresponding eigenvector to σ_1. Therefore this gives us the supremum and we have shown that:

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

Exercise 17: HJE 4.36.

Solution. An example of a matrix with a nonzero determinant whose singular values are not equal to any eigenvalues is given by:

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

This matrix has singular value decomposition with singular values 2 and 1:

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

However its eigenvalues solve:

$$\lambda^2 - 2 = 0 \implies \lambda = \sqrt{2}, \lambda = -\sqrt{2}.$$

We note that the determinant is also nonzero.

Exercise 18: HJE 4.38.

Solution.

(i.) Let $A^\dagger$ be the Moore-Penrose pseudoinverse of A defined by $A^\dagger = V\Sigma^{-1}U^H$.

$$AA^\dagger A = U\Sigma V^H V\Sigma^{-1}U^H U\Sigma V^H = U\Sigma\Sigma^{-1}\Sigma U^H = U\Sigma V^H \implies AA^\dagger A = A.$$

(ii.) $A^\dagger AA^\dagger = V\Sigma^{-1}U^H U\Sigma V^H V\Sigma^{-1}U^H = V\Sigma^{-1}\Sigma\Sigma^{-1}U^H = V\Sigma^{-1}U^H$. Therefore $A^\dagger AA^\dagger = A^\dagger$.

(iii.) $(AA^\dagger)^H = (U\Sigma V^H V\Sigma^{-1}U^H)^H = U\Sigma V^H V\Sigma^{-1}U^H$. Therefore $(AA^\dagger)^H = AA^\dagger$.

(iv.) $(A^\dagger A)^H = (V\Sigma^{-1}U^H U\Sigma V^H)^H = V\Sigma^{-1}U^H U\Sigma V^H$. Therefore $(A^\dagger A)^H = A^\dagger A$.

(v.) In order for $AA^\dagger$ to be the orthogonal projection onto $\mathcal{R}(A)$ it must satisfy three properties: it must map itself to itself, its projection must be in $\mathcal{R}(A)$, and its

residuals must be orthogonal to $\mathcal{R}(A)$. We first show that it maps to itself. For any vector $\mathbf{x}$:

$$AA^\dagger AA^\dagger = AA^\dagger$$

Next we note that for any $\mathbf{x} : AA^\dagger\mathbf{x} \in \mathcal{R}(A)$. Finally, let $AA^\dagger\mathbf{x} = \mathbf{b}$. Then the residual is given by $\mathbf{x} - \mathbf{b} = \mathbf{x} - AA^\dagger\mathbf{x}$. Then:

$$\begin{aligned}
\langle A\mathbf{x}, \mathbf{x} - AA^\dagger\mathbf{x}\rangle &= \mathbf{x}^H A^H \mathbf{x} - \mathbf{x}^H A^H AA^\dagger\mathbf{x} \\
&= \mathbf{x}^H (V\Sigma U^H\mathbf{x} - V\Sigma U^H U\Sigma V^H \Sigma^{-1} U^H \mathbf{x}) \\
&= \mathbf{x}^H (V\Sigma U^H\mathbf{x} - V\Sigma U^H\mathbf{x}) = 0
\end{aligned}$$

Then the projection is orthogonal. Therefore we have shown that $AA^\dagger$ is the orthogonal projection onto $\mathcal{R}(A)$.

(vi.) Now take the projection $A^\dagger A$. We first show that it is a projection because it maps to itself:

$$A^\dagger AA^\dagger A = A^\dagger A.$$

We then note that if $\mathbf{x} = A^\dagger\mathbf{b}$ then $\mathbf{x} \in \mathcal{R}(V) = \mathcal{R}(A^H)$ as noted in the proof of Theorem 4.6.1. Therefore $A^\dagger A\mathbf{x} \in \mathcal{R}(A^H)$. Finally, let $A^\dagger A\mathbf{x} = \mathbf{b}$. Then the residual is given by $\mathbf{x} - \mathbf{b} = \mathbf{x} - AA^\dagger\mathbf{x}$. Then:

$$\begin{aligned}
\langle A\mathbf{x}, \mathbf{x} - A^\dagger A\mathbf{x}\rangle &= \mathbf{x}^H A\mathbf{x} - \mathbf{x}^H AA^\dagger A\mathbf{x} \\
&= \mathbf{x}^H (U\Sigma V^H\mathbf{x} - U\Sigma V^H V\Sigma^{-1} U^H U\Sigma V^H\mathbf{x}) \\
&= \mathbf{x}^H (U\Sigma V^H\mathbf{x} - U\Sigma V^H\mathbf{x}) = 0.
\end{aligned}$$

Then the projection is orthogonal. Therefore we have shown that $A^\dagger A$ is the orthogonal projection onto $\mathcal{R}(A^H)$.