

Math Problem Set 3

Open Source Macroeconomics Laboratory Boot Camp

Ruby Zhang

July 8, 2017

4.2 Letting the basis be $[1, x, x^2]$, then the coordinates of the basis are $[1, 0, 0]$, $[0, 1, 0]$, $[0, 0, 1]$. Thus $p(x) = a + bx + cx^2$ can be represented as $[a, b, c]$. Then the differentiation matrix is the following:

$$D[p](x) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} b \\ 2c \\ 0 \end{bmatrix} = b + 2cx = p'(x)$$

The characteristic polynomial of $D[p](x)$ given by $\det(D - \lambda I)$ is:

$$\det(D - \lambda I) = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 2 \\ 0 & 0 & -\lambda \end{vmatrix} = \lambda^3$$

Since the eigenvalues are roots of the characteristic polynomial, the only eigenvalue of $D[p](x)$ is $\lambda = 0$. Now let us find the eigenspace (which in this scenario is equivalent to $\mathcal{N}(D)$). From before, note that $D[p](x)$ maps vector $\mathbf{p} = [a, b, c]$ to $[b, 2c, 0]$. Therefore if $\mathbf{p} \in \mathcal{N}(D)$, then $b, c = 0$. This means that the eigenspace of $\lambda = 0$ is spanned by \mathbf{e}_1 since a can equal anything. Therefore, the only eigenvalue $\lambda = 0$ has an algebraic multiplicity of 3 but a geometric multiplicity of 1.

4.4 (i) Suppose we have the following general Hermitian matrix where all coefficients represent real numbers:

$$A = \begin{bmatrix} a & b + ci \\ b - ci & d \end{bmatrix}$$

Then according to 4.3, the characteristic polynomial is $\lambda^2 - \text{tr}(A)\lambda + \det(A)$. The determinant of this quadratic equation is then:

$$\begin{aligned} \Delta &= \text{tr}^2(A) - 4\det(A) = (a + d)^2 - 4(ad - (b + ci)(b - ci)) \\ &= (a + d)^2 - 4(ad - b^2 - c^2) = a^2 + d^2 + 2ad - 4ad + 4b^2 + 4c^2 \\ &= (a^2 - 2ad + d^2) + 4b^2 + 4c^2 = (a - d)^2 + 4b^2 + 4c^2 \end{aligned}$$

Then $\Delta \geq 0$ since the sum of squared values is always nonnegative. Therefore, the eigenvalues, which are the roots of the characteristic polynomial, are always real.

- (ii) Suppose we have the following general skew Hermitian matrix where all coefficients represent real numbers:

$$A = \begin{bmatrix} ai & b + ci \\ -b + ci & di \end{bmatrix}$$

Then using 4.3 again, the determinant of the characteristic polynomial is:

$$\begin{aligned} \Delta &= \text{tr}^2(A) - 4\det(A) = (ai + di)^2 - 4(-ad - (b + ci)(ci - b)) \\ &= -(a + d)^2 + 4ad + 4(-c^2 - b^2) = -((a - d)^2 + 4c^2 + 4b^2) \end{aligned}$$

Then $\Delta \leq 0$ since the sum of squared values is always nonnegative so negating it will always make it nonpositive. Therefore, the eigenvalues, which are the roots of the characteristic polynomial, are always imaginary.

- 4.6 We will first show that the determinant of an upper triangular matrix is the product of its diagonal entries. For the base case of a 2×2 matrix (the 1×1 case is obvious and trivial), let A be a general upper triangular matrix:

$$A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$$

Then $\det(A) = ad$, so the proposition holds. Now let the inductive hypothesis be that determinant of any $n - 1 \times n - 1$ upper triangular matrix is the product of its diagonal entries. Let A be an $n \times n$ upper triangular matrix. Let A_{ij} denote the $n - 1 \times n - 1$ square matrix obtained by removing the i th row and j th column of matrix A . Then by using cofactor expansion along the first row of A , we have that:

$$\begin{aligned} \det(A) &= \sum_{j=1}^n a_{1j}(-1)^{1+j} \det(A_{ij}) \\ &= a_{11} \det(A_{11}) + \sum_{j=2}^n a_{1j}(-1)^{1+j} \det(A_{ij}) \end{aligned}$$

Note that every A_{ij} is also an upper triangular matrix so the inductive hypothesis can be invoked. From the inductive hypothesis, note that $\det(A_{11}) = \prod_{i=2}^n a_{ii}$. However, the first diagonal entry for all other A_{ij} is $a_{21} = 0$, therefore $\det(A_{ij}) = 0$. Then $\det(A) = a_{11} \det(A_{11}) = \prod_{i=1}^n a_{ii}$, which satisfies the inductive hypothesis.

Note that a similar inductive reasoning can be applied for lower triangular matrices by since the base case is satisfied and the same inductive hypothesis can be reached by using the cofactor expansion along the last row.

For a triangular matrix A , the matrix $A - \lambda I$ is also a triangular matrix with diagonal entries $a_{ii} - \lambda$. Therefore, the characteristic polynomial of A is $\det(A - \lambda I) = \prod_{i=1}^n (a_{ii} - \lambda)$. As seen, the roots of the characteristic polynomial, which correspond to the eigenvalues of A , are all the diagonal entries of A .

- 4.8 (i) Since V is the span of set S , we just need to show that the set S is linearly independent. Suppose that S is linearly dependent, then there exists scalars a, b, c, d such that $a \cos(x) + b \sin(x) + c \cos(2x) + d \sin(2x) = 0$ for all values of $x \in \mathbb{R}$. Plugging in $x = 0$ and $x = \pi$, we see that $a + c = 0$ and $-a + c = 0$, respectively. Therefore, $a = c = 0$. Plugging in $x = \frac{\pi}{2}$ we see that $b + c = 0$, so $b = 0$ since we concluded that $c = 0$. Since $d \sin(2x) = 0$ for all values of x , then $d = 0$. Therefore, $a = b = c = d = 0$ so the set S is linearly independent. Since it spans V , it is also the basis of V .

- (ii) Letting the basis be set $S = \{\cos(x), \sin(x), \cos(2x), \sin(2x)\}$, then the coordinates of the basis are $[1, 0, 0, 0], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1]$. Thus $p(x) = a \cos(x) + b \sin(x) + c \cos(2x) + d \sin(2x)$ can be represented as $[a, b, c, d]$. Then the differentiation matrix is the following:

$$D[p](x) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & -2 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} b \\ -a \\ 2d \\ -2c \end{bmatrix} = -a \sin(x) + b \cos(x) - 2c \sin(2x) + 2d \cos(2x) = p'(x)$$

- (iii) Let $U = \text{span}\{\cos(x), \sin(x)\}$ and $W = \text{span}\{\cos(2x), \sin(2x)\}$. Then U and W are complementary subspaces of V since their combined basis forms the basis of V . Note that both U and W are D -invariant since for any $g(x) = a \cos(x) + b \sin(x) \in U$, we have that $D[g](x) = -a \sin(x) + b \cos(x) \in U$ and for any $f(x) = c \cos(2x) + d \sin(2x) \in W$, we have that $D[f](x) = -2c \sin(2x) + 2d \cos(2x) \in W$.

4.13 To find the diagonalization of A , let us find its eigenvalues using its characteristic polynomial given by question 4.3:

$$\begin{aligned} p(\lambda) &= \lambda^2 - (0.8 + 0.6)\lambda + (0.8 \times 0.6 - 0.4 \times 0.2) = \lambda^2 - 1.4\lambda + 0.4 \\ &= (\lambda - 1)(\lambda - 0.4) \end{aligned}$$

Therefore, the eigenvalues of A are $\lambda = \{1, \frac{2}{5}\}$. From simple calculation, we can see that the corresponding eigenvectors are $[2, 1]^T$ and $[-1, 1]^T$, respectively (note that they are linearly independent). Then the columns of transition matrix P are simply the corresponding eigenvectors:

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

One can check that $P^{-1}AP = D$ as desired.

4.15 Since A is semisimple, then it is diagonalizable so $A = P^{-1}DP$ such that the eigenvalues of A make up the diagonal of D and the corresponding eigenvectors are the columns of P . Then we have that:

$$\begin{aligned} f(A) &= f(P^{-1}DP) = a_0I + a_1P^{-1}DP + \cdots + a_nP^{-1}D^nP \\ &= a_0P^{-1}IP + a_1P^{-1}DP + \cdots + a_nP^{-1}D^nP = P^{-1}(a_0I + a_1D + \cdots + a_nD^n)P \\ &= P^{-1}\bar{D}P \end{aligned}$$

where \bar{D} is the diagonal matrix with $f(\lambda_i)$ as its entries. Therefore, the matrix $f(A)$ is diagonalizable, where the resulting diagonal matrix, \bar{D} , contains its eigenvalues as the diagonal elements, which are $f(\lambda_i)$.

4.16 (i) Note that by proposition 4.3.10, we have that $\lim_{n \rightarrow \infty} A^n = \lim_{n \rightarrow \infty} P^{-1}D^nP = P^{-1}(\lim_{n \rightarrow \infty} D^n)P$. We know that D^n is simply the diagonal entries to the n^{th} power, thus, $\lim_{n \rightarrow \infty} D^n$ is the matrix with diagonal entries 1 and 0. Therefore, the matrix B is given by the following:

$$B = P^{-1}(\lim_{n \rightarrow \infty} D^n)P = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

Note that the 1-norm is the maximum column sum of a matrix. Therefore, we have that:

$$\|A^k - B\|_1 = \|P^{-1}D^kP - B\|_1 = \left\| \begin{bmatrix} \frac{2}{3} + \frac{0.4^k}{3} & -\frac{1}{3} + \frac{0.4^k}{3} \\ -\frac{2}{3} + \frac{2(0.4)^k}{3} & \frac{1}{3} + \frac{2(0.4)^k}{3} \end{bmatrix} - B \right\|_1 = \left\| \begin{bmatrix} \frac{0.4^k}{3} & \frac{0.4^k}{3} \\ \frac{2(0.4)^k}{3} & \frac{2(0.4)^k}{3} \end{bmatrix} \right\|_1 = 0.4^k$$

Note that for any $\epsilon > 0$, we can always find N such that $0.4^k < \epsilon$ for all $k > N$. Therefore we have verified that $B = \lim_{n \rightarrow \infty} A^n$ with respect to the 1-norm.

- (ii) We will show that B is the limit of A^n regardless of norm. We already know the matrix of $A^k - B$, thus let us compute the ∞ -norm (row sum) and Frobenius norm:

$$\|A^k - B\|_\infty = \left\| \begin{bmatrix} \frac{0.4^k}{3} & \frac{0.4^k}{3} \\ \frac{2(0.4)^k}{3} & \frac{2(0.4)^k}{3} \end{bmatrix} \right\|_\infty = \frac{4}{3}0.4^k$$

Note that for any $\epsilon > 0$, we can always find N such that $\frac{4}{3}0.4^k < \epsilon$ for all $k > N$. Therefore we have verified that $B = \lim_{n \rightarrow \infty} A^n$ with respect to the ∞ -norm.

$$\|A^k - B\|_F = \left\| \begin{bmatrix} \frac{0.4^k}{3} & \frac{0.4^k}{3} \\ \frac{2(0.4)^k}{3} & \frac{2(0.4)^k}{3} \end{bmatrix} \right\|_F = \sqrt{10 \frac{0.4^{2k}}{9}} = \frac{\sqrt{10}}{3}0.4^k$$

Note that for any $\epsilon > 0$, we can always find N such that $\frac{\sqrt{10}}{3}0.4^k < \epsilon$ for all $k > N$. Therefore we have verified that $B = \lim_{n \rightarrow \infty} A^n$ with respect to the Frobenius norm. Then we have that the answer does NOT depend on the choice of norm.

- (iii) Since A is semisimple (eigenvectors span \mathbf{R}), we can use theorem 4.3.12, which we proved in 4.15. According to the theorem, the eigenvalues of the matrix $f(A) = 3I + 5A + A^3$ are $f(1) = 9$ and $f(0) = 3$.

4.18 If λ is an eigenvalue of square matrix A , then $0 \nmid \dim(\mathcal{N}(A - \lambda I)) = \dim(\mathcal{N}(A^T - \lambda I))$. Then there exists a nonzero vector \mathbf{x} such that $A^T \mathbf{x} = \lambda \mathbf{x}$. Taking the transpose of both sides, we have the desired result: $\mathbf{x}^T A = \lambda \mathbf{x}^T$.

4.20 Suppose matrix A is Hermitian ($A = A^H$) and orthonormally similar to B . Then there exists orthonormal matrix U such that $UAU^H = B$. We have that:

$$B^H = (UAU^H)^H = UA^H U^H = UA U^H = B$$

Therefore, B is also hermitian.

4.24 From the second spectral theorem, we know that both hermitian and skew-hermitian matrices have an orthonormal eigenbasis since they are normal matrices are thus orthonormally diagonalizable. Suppose $\lambda_1, \dots, \lambda_n$ are the eigenvalues with corresponding orthonormal eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ that span \mathbb{F}^n . Then any $\mathbf{x} \in \mathbb{F}^n$ can be written as $\mathbf{x} = a_1 \mathbf{v}_1 + \dots + a_n \mathbf{v}_n$ with $a_i \in \mathbb{F}$ for all $1 \leq i \leq n$. Then we have

that:

$$\begin{aligned}
\rho(\mathbf{x}) &= \frac{\langle \mathbf{x}, A\mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle} \\
&= \frac{\langle (a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n), (Aa_1\mathbf{v}_1 + \cdots + Aa_n\mathbf{v}_n) \rangle}{\langle (a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n), (a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) \rangle} \\
&= \frac{\langle (a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n), (\lambda_1 a_1\mathbf{v}_1 + \cdots + \lambda_n a_n\mathbf{v}_n) \rangle}{\langle (a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n), (a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) \rangle} \\
&= \frac{\lambda_1 |a_1|^2 \|\mathbf{v}_1\|^2 + \cdots + \lambda_n |a_n|^2 \|\mathbf{v}_n\|^2}{|a_1|^2 \|\mathbf{v}_1\|^2 + \cdots + |a_n|^2 \|\mathbf{v}_n\|^2}
\end{aligned}$$

Note that all $|a_i|^2 \|\mathbf{v}_i\|^2$ terms are real, positive numbers. From exercise 4.4, since hermitian and skew-hermitian matrices have all real and all imaginary eigenvalues, respectively, then the Rayleigh quotients for hermitian and skew-hermitian matrices take on only real and imaginary values, respectively.

4.25 (i) For orthonormal eigenvectors $[\mathbf{x}_1, \dots, \mathbf{x}_n]$, we have that

$$(\mathbf{x}_1 \mathbf{x}_1^H + \cdots + \mathbf{x}_n \mathbf{x}_n^H) \mathbf{x}_j = \mathbf{x}_j \mathbf{x}_j^H \mathbf{x}_j = \mathbf{x}_j \|\mathbf{x}_j\|^2 = \mathbf{x}_j = I \mathbf{x}_j$$

Therefore, $\mathbf{x}_1 \mathbf{x}_1^H + \cdots + \mathbf{x}_n \mathbf{x}_n^H = I$.

(ii) For orthonormal eigenvectors $[\mathbf{x}_1, \dots, \mathbf{x}_n]$, we have that

$$(\lambda_1 \mathbf{x}_1 \mathbf{x}_1^H + \cdots + \lambda_n \mathbf{x}_n \mathbf{x}_n^H) \mathbf{x}_j = \lambda_j \mathbf{x}_j \mathbf{x}_j^H \mathbf{x}_j = \lambda_j \mathbf{x}_j \|\mathbf{x}_j\|^2 = \lambda_j \mathbf{x}_j = A \mathbf{x}_j$$

Therefore, $\lambda_1 \mathbf{x}_1 \mathbf{x}_1^H + \cdots + \lambda_n \mathbf{x}_n \mathbf{x}_n^H = A$.

4.27 Since A is positive definite, then it is hermitian so its diagonal entries are obviously real. Furthermore, for any $\mathbf{x} \neq 0$, we have that $\mathbf{x}^H A \mathbf{x} > 0$. Then this condition is also satisfied for all standard basis vectors \mathbf{e}_i .

$$0 < \mathbf{e}_i^H A \mathbf{e}_i = [a_{i1}, \dots, a_{in}] \mathbf{e}_i = a_{ii}$$

Therefore, $a_{ii} > 0$ for all $1 \leq i \leq n$ so the diagonal entries of a positive definite matrix are real and positive.

4.28 Since A is normal and orthonormally diagonalizable, we can rewrite A and B in terms of the orthonormal eigenbasis of A . Therefore, $A = \text{diag}(\lambda_1, \dots, \lambda_n)$. Since all diagonal entries of A and B are nonnegative (exercise 4.27), then we have that:

$$\begin{aligned}
\text{tr}(AB) &= \lambda_1 b_{11} + \cdots + \lambda_n b_{nn} \\
&\leq (\lambda_1 + \cdots + \lambda_n)(b_{11} + \cdots + b_{nn}) = \text{tr}(A)\text{tr}(B)
\end{aligned}$$

4.31 (i) Note that from exercise 3.29, $\|Q\|_2 = 1$ for any orthonormal matrix Q . Let $A = U \Sigma V^H$ be the

singular value decomposition of A where σ_1 is the largest singular value of A . Then we have that:

$$\begin{aligned}
\|A\|_2 &= \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sup_{\mathbf{x} \neq 0} \frac{\|U\Sigma V^H \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \\
&= \sup_{\mathbf{x} \neq 0} \frac{\|U\|_2 \|\Sigma V^H \mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sup_{\mathbf{x} \neq 0} \frac{\|\Sigma V^H \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \\
&= \sup_{\mathbf{y} \neq 0} \frac{\|\Sigma \mathbf{y}\|_2}{\|V \mathbf{y}\|_2} = \sup_{\mathbf{y} \neq 0} \frac{\|\Sigma \mathbf{y}\|_2}{\|\mathbf{y}\|_2} \\
&= \sup_{\mathbf{y} \neq 0} \frac{\|\Sigma \mathbf{y}\|_2}{\|V\|_2 \|\mathbf{y}\|_2} = \sup_{\mathbf{y} \neq 0} \frac{\|\Sigma \mathbf{y}\|_2}{\|\mathbf{y}\|_2} \\
&= \sup_{\mathbf{y} \neq 0} \frac{\sqrt{\sum_{i=1}^r \sigma_i^2 |y_i^2|}}{\sqrt{\sum_{i=1}^r |y_i^2|}} \leq \sup_{\mathbf{y} \neq 0} \frac{\sqrt{\sum_{i=1}^r \sigma_1^2 |y_i^2|}}{\sqrt{\sum_{i=1}^r |y_i^2|}} \\
&= \sigma_1
\end{aligned}$$

Therefore, we have the inequality $\|A\|_2 \leq \sigma_1$. Note that equality is achieved by \mathbf{x} being the corresponding eigenvector to eigenvalue σ_1^2 . Being the supremum, $\|A\|_2 = \sigma_1$.

- (ii) If $A = U\Sigma V^H$, then $A^{-1} = V\Sigma^{-1}U^H$. Note that this is a singular value decomposition of A^{-1} , where the largest singular value is σ_n^{-1} . Therefore, $\|A^{-1}\|_2 = \sigma_n^{-1}$.
- (iii) If $A = U\Sigma V^H$, then we get the following singular value decompositions: $A^H = V\Sigma^H U^H$, $A^T = \bar{V}\Sigma^T U^T$, $A^H A = U\Sigma^H \Sigma U^H$. Note that $\Sigma^2, (\Sigma^H)^2, (\Sigma^T)^2, \Sigma^H \Sigma$ have all the same diagonal entries: the square of the singular values. Therefore we have that $\sigma_1^2 = \|A^H A\|_2 = \|A^H\|_2^2 = \|A^T\|_2^2 = \|A\|_2^2$.
- (iv) Recall that using the standard inner product used for $\|\cdot\|_2$, the adjoint of a matrix is its Hermitian conjugate: $B^* = B^H$. Furthermore, $\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2$ for all orthonormal matrices Q which was proved in exercise 3.29. Then for any $\|UAV\|_2$, we have that

$$\begin{aligned}
\|UAV\|_2 &= \sup_{\mathbf{x} \neq 0} \frac{\|UAV\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sup_{\mathbf{x} \neq 0} \frac{\sqrt{\langle UAV\mathbf{x}, UAV\mathbf{x} \rangle}}{\|\mathbf{x}\|_2} \\
&= \sup_{\mathbf{x} \neq 0} \frac{\sqrt{\langle U^H UAV\mathbf{x}, AV\mathbf{x} \rangle}}{\|\mathbf{x}\|_2} = \sup_{\mathbf{x} \neq 0} \frac{\sqrt{\langle AV\mathbf{x}, AV\mathbf{x} \rangle}}{\|\mathbf{x}\|_2} \\
&= \sup_{V\mathbf{x} \neq 0} \frac{\|AV\mathbf{x}\|_2}{\|V\mathbf{x}\|_2} = \sup_{\mathbf{y} \neq 0} \frac{\|A\mathbf{y}\|_2}{\|\mathbf{y}\|_2} \\
&= \|A\|_2
\end{aligned}$$

4.32 (i)

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

- (ii) Let $A = U\Sigma V^H$ be the singular value decomposition of A . From part (i), we have that $\|A\|_F = \|U\Sigma V^H\|_F = \|\Sigma\|_F = (\text{tr}(\Sigma^H \Sigma))^{\frac{1}{2}} = (\sigma_1^2 + \dots + \sigma_r^2)^{\frac{1}{2}}$.

4.33 From problem 4.31, we have that $\sigma_1 = \|A\|_2 = \|U\Sigma V^H\|_2 = \|\Sigma_2\|$. Therefore, it suffices to show that $\sup_{\substack{\|\mathbf{x}\|_2=1 \\ \|\mathbf{y}\|_2=1}} |\mathbf{y}^H \Sigma \mathbf{x}| = \sigma_1$:

$$\begin{aligned} \sup_{\substack{\|\mathbf{x}\|_2=1 \\ \|\mathbf{y}\|_2=1}} |\mathbf{y}^H \Sigma \mathbf{x}| &= \sup_{\substack{\|\mathbf{x}\|_2=1 \\ \|\mathbf{y}\|_2=1}} |\langle \mathbf{y}, \Sigma \mathbf{x} \rangle| \\ &\leq \sup_{\substack{\|\mathbf{x}\|_2=1 \\ \|\mathbf{y}\|_2=1}} \|\mathbf{y}\|_2 \|\Sigma \mathbf{x}\|_2 \text{ from Cauchy-Schwarz} \\ &= \sup_{\|\mathbf{x}\|_2=1} \|\Sigma \mathbf{x}\|_2 \leq \sup_{\|\mathbf{x}\|_2=1} \sigma_1 \|\mathbf{x}\|_2 = \sigma_1 \end{aligned}$$

Note that equality holds when $\mathbf{y} = \mathbf{x} = \mathbf{v}_1$, the corresponding eigenvector to σ_1^2 .

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

4.36 Take the following matrix:

$$A = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}, A^H A = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}$$

$$\det(A) = 6 \neq 0$$

$$\lambda_1, \lambda_2 = -3, -2$$

$$\sigma_1, \sigma_2 = 3, 2$$

4.38 (i)

$$AA^\dagger A = A(V_1 \Sigma_1^{-1} U_1^H U_1 \Sigma_1 V_1^H) = AV_1 V_1^H = U_1 \Sigma_1 V_1^H V_1 V_1^H = U_1 \Sigma_1 V_1^H = A$$

(ii)

$$A^\dagger AA^\dagger = (V_1 \Sigma_1^{-1} U_1^H U_1 \Sigma_1 V_1^H) A^\dagger = V_1 V_1^H A^\dagger = V_1 V_1^H V_1 \Sigma_1^{-1} U_1^H = V_1 \Sigma_1^{-1} U_1^H = A^\dagger$$

(iii) On the left side we have:

$$(AA^\dagger)^H = (V_1 \Sigma_1^{-1} U_1^H)^H (U_1 \Sigma_1 V_1^H)^H = U_1 (\Sigma_1^{-1})^H V_1^H V_1 \Sigma_1^H U_1^H = U_1 \Sigma_1^{-1} \Sigma_1 U_1^H = U_1 U_1^H$$

On the right side we have:

$$AA^\dagger = U_1 \Sigma_1 V_1^H V_1 \Sigma_1^{-1} U_1^H = U_1 \Sigma_1^{-1} \Sigma_1 U_1^H = U_1 U_1^H$$

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

(iv) On the left side we have:

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

On the right side we have:

$$A^\dagger A = V_1 \Sigma_1^{-1} U_1^H U_1 \Sigma_1 V_1^H = V_1 \Sigma_1 \Sigma_1^{-1} V_1^H = V_1 V_1^H$$

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

- (v) Denote the columns of $U_1 = [\mathbf{u}_1, \dots, \mathbf{u}_r]$, which forms an orthonormal basis for $\mathcal{R}(A)$ by construction of the SVD decomposition. From part (iii), we have that:

$$AA^\dagger \mathbf{x} = U_1 U_1^H \mathbf{x} = U_1 [\mathbf{u}_1^H \mathbf{x}, \dots, \mathbf{u}_r^H \mathbf{x}]^T = \sum_{i=1}^r \mathbf{u}_i^H \mathbf{x} \mathbf{u}_i = \sum_{i=1}^r \langle \mathbf{u}_i, \mathbf{x} \rangle \mathbf{u}_i = \text{proj}_{\mathcal{R}(A)} \mathbf{x}$$

- (vi) Denote the columns of $V_1 = [\mathbf{v}_1, \dots, \mathbf{v}_r]$, which forms an orthonormal basis for $\mathcal{R}(A^H)$ by construction of the SVD decomposition. From part (iv), we have that:

$$A^\dagger A \mathbf{x} = V_1 V_1^H \mathbf{x} = V_1 [\mathbf{v}_1^H \mathbf{x}, \dots, \mathbf{v}_r^H \mathbf{x}]^T = \sum_{i=1}^r \mathbf{v}_i^H \mathbf{x} \mathbf{v}_i = \sum_{i=1}^r \langle \mathbf{v}_i, \mathbf{x} \rangle \mathbf{v}_i = \text{proj}_{\mathcal{R}(A^H)} \mathbf{x}$$