

Problem Set 3

OSM Lab-Math

Sophia Mo

Problem 1 (4.2)

In Problem 3.38 we have shown that the derivative operator is

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

To find the eigenvalue, we can calculate the root to the characteristic polynomial

$$\det(\lambda I - D) = \lambda^3$$

Therefore, the only eigenvalue is $\lambda = 0$, with algebraic multiplicity 3 and geometric multiplicity 1. Any vector in the form of $(a, 0, 0)$, $a \in \mathbb{R}$ is an eigenvector and belongs to the eigenspace.

Problem 2 (4.4)

(i)

Suppose the A is a Hermitian matrix. Let λ be any eigenvalue of A , with associated eigenvector v . Then by definition,

$$\begin{aligned} Av &= \lambda v \\ (Av)^H &= \bar{\lambda} v^H \\ v^H A^H &= \bar{\lambda} v^H \\ v^H A^H v &= \bar{\lambda} v^H v \\ v^H A v &= \bar{\lambda} v^H v \\ v^H \lambda v &= \bar{\lambda} v^H v \\ \lambda v^H v &= \bar{\lambda} v^H v \\ \lambda &= \bar{\lambda} \end{aligned}$$

For the last equation to hold, λ must be real.

(ii)

Suppose the A is a skew-Hermitian matrix. Let λ be any eigenvalue of A , with

associated eigenvector v . Then by definition,

$$\begin{aligned}
 Av &= \lambda v \\
 (Av)^H &= \bar{\lambda} v^H \\
 v^H A^H &= \bar{\lambda} v^H \\
 v^H A^H v &= \bar{\lambda} v^H v \\
 v^H (-A)v &= \bar{\lambda} v^H v \\
 v^H (-\lambda)v &= \bar{\lambda} v^H v \\
 -\lambda v^H v &= \bar{\lambda} v^H v \\
 \lambda &= -\bar{\lambda}
 \end{aligned}$$

For the last equation to hold, λ must be imaginary.

Problem 3 (4.6)

For any $n \times n$ upper triangular matrix A , we know that its eigenvalues are the solution to $p(\lambda) = \det(A - \lambda I)$. $\forall a_{i,i}$ on the diagonal of A , we know that it is a solution to $p(\lambda)$ since the determinant of an upper-triangular matrix is the product of all entries on its diagonal. The sum of the algebraic multiplicity of each diagonal entry will sum to n , and we know, by proposition 4.1.11, that they are all the eigenvalues A has.

Problem 4 (4.8)

(i)

Since V is the span of the set S , we only need to show that the elements in S are linearly independent. However, in Problem 3.8(i), we have already shown that S is an orthonormal set, so every two elements in S are orthogonal to each other. Linear independence follows, and S is a basis for V .

(ii)

Apply D to each element of the basis, we get:

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

The transformations of the basis elements form the columns of the matrix D , so

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

(iii)

From our calculations in (ii) we could see that $\{\sin(x), \cos(x)\}$ and $\{\sin(2x), \cos(2x)\}$ are two D -invariant subspaces in V .

Problem 5 (4.13)

We know the columns of P are the right eigenvectors of A . Denote a column by $(x, y)^T$, an eigenvalue by λ , then

$$0.8x + 0.4y = \lambda x$$

$$0.2x + 0.6y = \lambda y$$

So $\lambda = 1$ or 0.4 , $(x, y) = (2, 1)$ or $(1, -1)$

The transition matrix $P = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$.

Problem 6 (4.15)

Let A be a semi-simple matrix. We know that A is diagonalizable, so $A = P^{-1}DP$, for some nonsingular matrix P and diagonal matrix D . Thus,

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

Since λ_i 's, as the eigenvalues of A , are the diagonal entries of D , from the equation above we see that $f(A)$ is also a diagonalizable matrix, and $f((\lambda_i))$'s, are the diagonal entries of $f(D)$. Therefore, $f(\lambda_i)$'s are eigenvalues of $f(A)$.

Problem 7 (4.16)

(i)

$$\begin{aligned} A &= \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.4 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} \end{pmatrix} \\ A^k &= \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.4^k \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{3}(2 + 0.4^k) & \frac{1}{3}(2 - 2 \times 0.4^k) \\ \frac{1}{3}(1 - 0.4^k) & \frac{1}{3}(1 + 2 \times 0.4^k) \end{pmatrix} \end{aligned}$$

$$\lim_{n \rightarrow \infty} A^n = \begin{pmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

$$\|A^k - B\|_1 = \frac{4}{3} \times 0.4^k \rightarrow 0.$$

(ii)

$\|A^k - B\|_\infty = 0.4^k \rightarrow 0$
 $\|A^k - B\|_F = \frac{\sqrt{10}}{3} \times 0.4^k \rightarrow 0$
 So the result do not change (B is the same).

(iii)

By Theorem 4.3.12, the eigenvalues are $3 + 5 + 1 = 9$ and $3 + 5 \times 0.4 + (0.4)^3 = 5.064$

Problem 8 (4.18)

By Proposition 4.1.23, A and A^T have the same characteristic polynomial, so they have the same eigenvalues. Therefore, if λ is an eigenvalue of A , then it is an eigenvalue of A^T , so \exists associated eigenvector x s.t. $A^T x = \lambda x$. Taking the transpose of both sides, we get $x^T A = \lambda x^T$.

Problem 9 (4.20)

Suppose A is Hermitian and orthonormally similar to B . Then \exists orthonormal matrix U such that $B = U^H A U$. So $B^H = U^H A^H U$. Since A is Hermitian, $A^H = A$, so $B^H = U^H A U = B$. So B is Hermitian.

Problem 10 (4.24)

Suppose A is Hermitian, then

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

For a number to equal its complex conjugate, the number must be real. Since $\|x\|^2$ is real, we know that $\rho(x)$ is real.

Suppose A is skew-Hermitian, then

$$\langle x, Ax \rangle = x^H A^H x = x^H (-A) x = \langle -Ax, x \rangle = -\langle Ax, x \rangle = -\overline{\langle x, Ax \rangle}$$

For a number to be the opposite of its complex conjugate, the number must be imaginary. Since $\|x\|^2$ is real, we know that $\rho(x)$ is imaginary.

Problem 11 (4.25)

(i)

We know $[x_1, \dots, x_n]$ are orthonormal eigenvectors.

\Rightarrow Let $Q = [x_1, x_2, \dots, x_n]$ (orthonormal matrix), Problem 3.10 shows that $Q^H Q = Q Q^H = I$. But $Q Q^H = x_1 x_1^H + \dots + x_n x_n^H$, so $x_1 x_1^H + \dots + x_n x_n^H = I$.

(ii)

$$\lambda_1 x_1 x_1^H + \dots + \lambda_n x_n x_n^H = A x_1 x_1^H + \dots + A x_n x_n^H = A (x_1 x_1^H + \dots + x_n x_n^H) = A I = A$$

Problem 12 (4.27)

Suppose that A is positive definite. Then by Proposition 4.5.7, we know A can be written as $S^H S$, where S is non-singular. Notice that the i the diagonal element of A is then just the square of the (i, i) the element of S , which must be real and positive.

Problem 13 (4.28)

Since A, B are positive semidefinite matrices, we know $A = S^H S, B = T^H T$
 $\Rightarrow \text{tr}(AB) = \text{tr}(S^H S T^H T) = \text{tr}((TS^H)(ST^H)) = \text{tr}(U^H U)$, if we let $U = ST^H$.
 $\Rightarrow \text{tr}(AB)$ is equal to the trace of some semidefinite matrix (by Proposition 4.5.9).
 Following the proof for Problem 4.27, we know that a semi-definite matrix has non-negative diagonal entries, so its trace ≥ 0 . The first inequality holds.

In the space of semi-definite matrices, we can define the inner product by $\langle A, B \rangle = \text{tr}(AB)$. Note that trace is indeed a norm since

$$\begin{aligned}\text{tr}(AA) &= \sum (\lambda_i)^2 \geq 0 \\ \text{tr}(A(bB + cC)) &= b\text{tr}(AB) + c\text{tr}(AC) \\ \text{tr}(AB) &= \overline{\text{tr}(BA)}\end{aligned}$$

Therefore, by Cauchy-Schwartz Inequality, we know

$$\text{tr}(AB) \leq \sqrt{\text{tr}(A^2)\text{tr}(B^2)}$$

Also, $\text{tr}(A^2) = \sum (\lambda_i)^2 \leq (\sum \lambda_i)^2 = [\text{tr}(A)]^2$. The same reasoning applies to B .
 So $\sqrt{\text{tr}(A^2)\text{tr}(B^2)} \leq \text{tr}(A)\text{tr}(B)$ (using the fact that the trace of a matrix = sum of its eigenvalues), and the second inequality holds.

Problem 14 (4.31)

(i)

$$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sup_{\|x\|=1} \|Ax\|_2$$

So we want to

$$\max f = x^H A^H A x \text{ s.t. } g = x^H x = 1$$

Introducing the lagrange multiplier λ , we know the solutions occur when $\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0$.

$$\Rightarrow (A^H A - \lambda I)x = 0 \Rightarrow (A^H A - \lambda I) = 0(*) \text{ since } x \neq 0.$$

$$\Rightarrow x^H A^H A x = x^H \lambda I x = \lambda x^H x = \lambda$$

So we want to find the maximal value that satisfies equation (*), which is equivalent to finding the largest eigenvalue of $A^H A$.

It then follow that $\|A\|_2 = \sigma_1$, the largest singular value of A .

(ii)

Using SVD, A could be decomposed into $U\Sigma V^H$, where U, V are orthonormal matrices. So $A^{-1} = (U\Sigma V^H)^{-1} = V\Sigma^{-1}U^{-1}$.

By (i), we know that $\|A^{-1}\|_2$ is equal to the maximal diagonal entry of Σ^{-1} , which is equal to σ_n^{-1} , where σ_n is the smallest singular value of A .

(iii)

Let λ be an eigenvalue of $A^H A$, then $\exists x$ s.t. $A^H A x = \lambda x \Rightarrow A A^H (A x) = A \lambda x = \lambda (A x)$. So we know λ is an eigenvalue of $A A^H$. Conversely, let λ be an eigenvalue of $A A^H$, then $A A^H x = \lambda x \Rightarrow A^H A (A^H x) = \lambda (A^H x)$, so λ is an eigenvalue of $A^H A$. Therefore, we know $A^H A$ and $A A^H$ have the same eigenvalues, which indicates that A^H and A have the same singular values. By (i), we then know that $\|A\|_2 = \|A^H\|_2$.

Since $(A^T)^H A^T = A A^H$, A^T and A^H have the singular eigenvalues, so $\|A^H\|_2 = \|A^T\|_2$

$(A^H A)^H (A^H A) = A^H A A^H A$. Let λ be a singular value for $A^H A$, then $\exists x$ s.t. $(A^H A)^H (A^H A) x = \lambda^2 x \Rightarrow (A^H A) (A^H A x) = \lambda (\lambda x)$, so λ is an eigenvalue for $A^H A$. Conversely, let λ be an eigenvalue for $A^H A$, then $(A^H A) x = \lambda x \Rightarrow (A^H A)^H (A^H A) x = \lambda^2 x$, so λ is a singular value of $A^H A$. Thus the eigenvalues of $A^H A$ are the same as its singular values. Since $\|A^H A\|_2$ is its maximal singular value, it is also its maximal eigenvalue, which is equal to $\|A\|_2^2$

(iv)

As proved in Problem 3.29, orthonormal matrices have norm 1. Moreover, we know \forall matrix A, B , $\|AB\|_2 \leq \|A\|_2 \|B\|_2$.

So $\|UAV\|_2 \leq \|U\|_2 \|A\|_2 \|V\|_2 = \|A\|_2$.

As proved in 3.10, the inverse of an orthonormal matrix is also orthonormal. So $\|A\|_2 \leq \|U^{-1} U A V V^{-1}\|_2 = \|U^{-1}\|_2 \|U A V\|_2 \|V^{-1}\|_2 = \|U A V\|_2$

$\Rightarrow \|U A V\|_2 = \|A\|_2$

Problem 15 (4.32)

(i)

$$\|U A V\|_F = \sqrt{\text{tr}[(U A V)^H (U A V)]} = \sqrt{\text{tr}(V^H A^H U^H U A V)} = \sqrt{V^H A^H A V} = \sqrt{A^H A V^H V} = \sqrt{A^H A} = \|A\|_F$$

(ii)

$$\|A\|_F = \sqrt{\text{tr}(A^H A)} = \sqrt{\sum \lambda_i}, \text{ where } \lambda_i \text{'s are the eigenvalues of } A^H A.$$

$$\Rightarrow \|A\|_F = (\sigma_1^2 + \dots + \sigma_r^2)^{\frac{1}{2}}.$$

Problem 16 (4.33)

We know $\|A\|_2 = \sigma_1$, the maximal diagonal element of Σ , and that $A = U \Sigma V^H \Rightarrow U^H A V = \Sigma$, where U, U^H, V, V^H are all orthonormal matrices.

Note that each column of an orthonormal matrix has norm 1, so the diagonal elements of Σ could be written as $y^H A x$, for some y, x with norm one.

Finding the maximal σ is therefore equivalent to finding $\sup_{\|x\|_2=\|y\|_2=1} y^H A x$.

Problem 17 (4.36)

For example, $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ has eigenvalues 1 and 3, but its singular value $2\sqrt{10} + 7$, is not equal to any of the eigenvalues.

Problem 18 (4.38)

(i)

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

(ii)

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

(iii)

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

(iv)

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

(v)

Denote the transformation AA^\dagger by P . Since $P^2 = AA^\dagger AA^\dagger = AA^\dagger = P$, by Problem 3.47 we know that P is a projection. Also notice that $PA = AA^\dagger A = A$.

$\forall y \in \mathcal{R}(A), \exists x$ s.t. $Ax = y \Rightarrow Ax = PAx = Py = y$.

If $Py = y$, then $y = AA^\dagger y$, clearly $y \in \mathcal{R}(A)$.

So $P = AA^\dagger = \text{proj}_{\mathcal{R}(A)}$

(vi)

Following the previous reasoning, we first notice that $Q^2 = Q$, so Q is a projection matrix, and $QA^\dagger = A^\dagger$.

$\forall y \in \mathcal{R}(A^H), \exists x$ s.t. $A^H x = y \Rightarrow A^H x = QA^H x = Qy = y$.

If $Qy = y$, then $y = A^\dagger A y$, clearly $y \in \mathcal{R}(A^H)$.

So $Q = A^\dagger A = \text{proj}_{\mathcal{R}(A^H)}$