Problem Set 3

Yung-Hsu Tsui*†

July 9 , 2018

Ex 4.2

Answer : The eigenvalue of this linear differential operator $D[p](x) = p^{'}(x)$ is 0. Also, the eigenspace is $\sum_{\lambda}(D) = \{a + bx + cx^2 \in V | b = c = 0\}$
Algebraic and geometric multiplicities of D are the dimension of $\sum_{\lambda}(D)$, which is 1 and the algebraic multiplicity of $\lambda_i = 0$ is 1.

Ex 4.4

Proof of (i)
Note that by the definition of eigenvalues, eigenvalues in 2 by 2 matrix satisfy the following;

$$p(\lambda) = \lambda^2 - (a + d)\lambda + (ad - bc) = 0$$

where

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

Using quadratic formula,

$$(a + d)^2 - 4(ad - bc) = (a - d)^2 + 4bc$$

Note that the matrix A is hermitian, meaning that $A^H = A$. This implies that a and d are real number, and the multiplication of $b = \bar{c}$ and $c = \bar{b}$ results in positive number(this is really easy but tedious work, so calculation is up to you!). Thus, $(a - d)^2 + 4bc > 0$, implying that the solutions of characteristic equations are all real.

*University of Chicago, Master of Art Program in Social Science, 1126 E. 59th Street, Chicago, Illinois, 60637, (773) 702-5079, yhtsui@uchicago.edu.

†I thank Jayhyung Kim for his great comments.

Q.E.D

Proof if (ii)
This is same task as above. If we find the quadratic formula of this 2nd order polynomial equation, then

$$D = (a-d)^2 + 4bc = -(a_1 - d_1)^2 - 4(abs(b)^2) < 0$$

(Also, the calculation is up to you!) Q.E.D

Ex 4.6

Proof. Let A be the upper triangular matrix. Consider $det(\lambda I - A) = 0$. This results in the following equation;

$$\prod_{i=1}^{n} (\lambda - a_i i)$$

, where a_{ii} is the ith diagonal entry. This expression equals to zero, iff $\lambda = a_{ii}$ for some i. Thus, the diagonal entries of the matrix are exactly the eigenvalues. Q.E.D

Ex 4.8

Proof of (i). Set the following matrix;

$$\begin{bmatrix} sin(t_1) & cos(t_1) & sin(2t_1) & cos(2t_1) \\ sin(t_2) & cos(t_2) & sin(2t_2) & cos(2t_2) \\ sin(t_3) & cos(t_3) & sin(2t_3) & cos(2t_3) \\ sin(t_4) & cos(t_4) & sin(2t_4) & cos(2t_4) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

with $t_1 = 0, t_2 = \pi/2, t_3 = \pi, t_4 = 3/2\pi$. This results in $c_1 = c_2 = c_3 = c_4 = 0$. Q.E.D

Proof of (ii). Let $D[p](x) := p^{'}(x)$. By calculation, the matrix that represents for this differential operator is the following;

$$\begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 2 & 0 \end{bmatrix}$$

Q.E.D

Answer of (iii). Let $V_1 = span(\{sin(x), cos(x)\})$ and $V_1 = span(\{sin(2x), cos(2x)\})$.

Ex 4.13

Answer.

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 0.4 \end{bmatrix}$$

and

$$P = \begin{bmatrix} 0.7454 & 0.7454 \\ -0.4714 & 0.9428 \end{bmatrix}$$

Ex 4.15

Proof. Due to the assumption that A is semi-simple, A is diagonalizable, i.e. $\exists P\ s.t. P^{-1}AP = D$, where D is diagonal matrix. Then, $A^k = PD^kP^{-1}$. Then,

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

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

Ex 4.16

Proof of (i). The markov chain which this matrix A^T represents for is irreducible and aperiodic. Thus, there exists distribution π such that $A\pi = \pi$. If we solve this, then $\pi = (2/3, 1/3)$, which is exactly the same with the first and the second columns of $\lim_{n\to\infty} A^n$ Q.E.D

Answer of (ii). Yes, $\|\lim A^n\|_\infty = 4/3$, and $\|\lim A^n\|_F = \sqrt{10}/3$
Answer of (iii). By the Theorem 4.3.12, $f(\lambda_1) = 3 + 5 * \lambda_1 + \lambda_1^3 = 9$, and $f(\lambda_2) = 3 + 5 * \lambda_2 + \lambda_2^3 = 5.0640$

Ex 4.18

Proof. Note that

$$det(A^T - \lambda I) = det((A - \lambda I)^T) = det(A - \lambda I) = 0$$

Thus, $A^T x = \lambda x$, and with trnasposition, $x^T A = \lambda x^T$

Ex. 4.20

Proof. Note that $A^H = A$. Using the notations in the Definition 4.4.1,

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

Q.E.D

Ex. 4.24

Proof. Due to the assumption that the matrix A is hermitian, then all the eigenvalues of A are real, because;

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

Thus, $\lambda = \bar{\lambda}$, meaning that the eigenvalues are real. Also, due to this fact, the matrix A is positive semi-definite. Thus, by Proposition 4.5.6, the eigenvectors of

A corresponding to distinct eigenvalues are orthogonal. Then, any vector x can be expressed in the following;

$$x = \sum_{j=1}^{n} c_j v_j = Vc$$

, where v_js are orthonormal eigenvectors.

This results in the following;

$$\begin{aligned}\rho(x) =& \frac{x^H A x}{x^H x} \\ =& \frac{c^H V^H A V c}{c^H V^H V c} \\ =& \frac{c^H \mathbf{A} c}{c^H c}\end{aligned}$$

where $A = V\mathbf{A}V^H$, and $\mathbf{A}$ is diagonal matrix. Thus,

$$\rho(x) = \frac{\lambda_1 |c_1|^2 + \cdots \lambda_n |c_n|^2}{|c_1|^2 + \cdots |c_n|^2 +}$$

Thus, $\rho(x)$ are real with hermitian matrix, A. We can do the same proof for Skewed matrix too! Q.E.D

Ex. 4.25

Proof of (i). Note that the following holds due to the assumption that $[x_1, ..., x_n]$ are orthonormal vectors; Thus,

$$(x_1 x_1^H + \cdots + x_n x_n^H) x_j = x_j = I x_j$$

for all j. Thus, the statement holds!

Proof of (ii). Note that by the Thoerem 4.4.14, A is orthonormally diagonalizable, i,e.

$$A = UTU^H$$

with T being diagonal, and U being orhonormal. This results in;

$$A = \sum_{i=1}^{n} t_{ii} u_i u_i^H$$

Q.E.D Please, prove more!

Ex. 4.27

Note that the positive-definite matrix A satisfies the following;

$$\forall x \neq 0, x^H A x > 0$$

If we feed the standard basis vector to x, then

$$e_i^H A e_i = a_{ii} > 0$$

Thus, all the diagonal entries are real and positive. Q.E.D

Ex. 4.28

Proof. Note that by the same logics of Ex. 4.27, one can show that the diagonal entries of any semi-positive definite matrix are non-negative. For the first inequality, we need to show that AB is a semi-positive definite matrix.

Note that $\forall x \neq 0, x^T A x, x^T B x \geq 0$. Then,

$$(x^T A x)(x^T B x) = (x^T A)(x x^T)(B x) \geq 0$$

Note that $x x^T$ is positive scalar when $x \neq 0$. Thus, $x^T A B x \geq 0$, implying that AB is positive semi-definite. This leads to the result that

$$tr(AB) \geq 0$$

Lastly, by using Cauchy-Scwartz inequality, you can show that

$$tr(AB) \leq tr(A) tr(B)$$

Please, prove more!

Ex. 4.31

Proof of (i). Note that we want to show $\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}(\mathbf{A}^H\mathbf{A})}$
We can first simply prove when $\mathbf{P}$ is hermitian,

$$\lambda_{\max} = \max_{\|\mathbf{x}\|_2=1} \mathbf{x}^H\mathbf{P}\mathbf{x}$$

That's because when $\mathbf{P}$ is Hermitian, there exists one and only one unitary matrix $\mathbf{U}$ that can diagonalize $\mathbf{P}$ as $\mathbf{U^HPU} = \mathbf{D}$ (so $\mathbf{P} = \mathbf{UDU^H}$), where $\mathbf{D}$ is a diagonal matrix with eigenvalues of $\mathbf{P}$ on the diagonal, and the columns of $\mathbf{U}$ are the corresponding eigenvectors. Let $\mathbf{y} = \mathbf{U^H x}$ and substitute $\mathbf{x} = \mathbf{Uy}$ to the optimization problem, we obtain

$$\max_{\|\mathbf{x}\|_2=1} \mathbf{x}^H\mathbf{P}\mathbf{x} = \max_{\|\mathbf{y}\|_2=1} \mathbf{y}^H\mathbf{D}\mathbf{y} = \max_{\|\mathbf{y}\|_2=1} \sum_{i=1}^{n} \lambda_i |y_i|^2 \leq \lambda_{\max} \max_{\|\mathbf{y}\|_2=1} \sum_{i=1}^{n} |y_i|^2 = \lambda_{\max}$$

Thus, just by choosing $\mathbf{x}$ as the corresponding eigenvector to the eigenvalue $\lambda_{\mathbf{max}}$, $\max_{\|\mathbf{x}\|_2=1} \mathbf{x}^H\mathbf{P}\mathbf{x} = \lambda_{\max}$. This proves $\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}(\mathbf{A}^H\mathbf{A})}$

Proof of (ii). Note that if the matrix A is invertible, then;

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

Note that $\hat{U}$ and $\hat{V}$ are all trivially orthornormal. Also, inverted diagonal matrix Σ^{-1} takes the inverse values of its diagonal entries on its diagonal line. Thus, $\|A^{-1}\|_2 = \sigma_n^{-1}$

Proof of (iii).

Note that according to the property of singular values(positive and real), the following holds;

$$\Sigma = \Sigma^T = \Sigma^H$$

Thus, by (i), $\|A\|_2 = \|A^H\|_2 = \|A^T\|_2$
Also,

$$A^HA = V\Sigma^H U^H U\Sigma V^H = V\Sigma^H\Sigma V^H = V\Sigma^2 V^H$$

Thus, $\|A^H A\|_2 = \|A\|_2^2$

Proof of (iv).

Note that

$$UAV = UU_1\Sigma V_1^H V = \hat{U}\Sigma\hat{V}^H$$

Note that $\hat{U}$ and $\hat{V}$ are orthonormal. For example,

$$\hat{U}^H\hat{U} = (UU_1)^H UU_1 = U_1^H U^H UU_1 = U_1^H U_1 = I$$

You can use this same argument to prove that $\hat{V}$ is orthonormal. Thus, $\|UAV\|_2 = \|A\|_2$. Q.E.D

Ex. 4.32

We first prove (ii), and then use the result to prove (i).

Proof of (ii).

$$\begin{aligned}
\|A\|_F^2 =& tr(AA^H) = tr(U\Sigma V^H V\Sigma^H U^H) \\
=& tr(U\Sigma\Sigma^H U^H) \\
=& tr(\Sigma\Sigma^H U^H U) \\
=& tr(\Sigma\Sigma^H) \\
=& \sigma_1^2 + ... + \sigma_n^2
\end{aligned}$$

Proof of (i).

$$\begin{aligned}\|U_1AV_1\|_F^2 =& tr((U_1AV_1)(U_1AV_1)^H)\\ =& tr(U_1AV_1V_1^HA^HU_1^H)\\ =& tr(U_1AA^HU_1^H)\\ =& tr(AA^HU_1^HU_1)\\ =& tr(AA^H)\\ =& tr(\Sigma\Sigma^H)\\ =& \sigma_1^2+...+\sigma_n^2\end{aligned}$$

Thus, $\|A\|_F = \|U_1AV_1^H\|$ Q.E.D

Ex. 4.33

Proof.

$$\begin{aligned}|y^HAx| =& |y^H(U\Sigma V^H)x|\\ =& |y^H(\sum_{i=1}^r \sigma_i u_i v_i^H|)x\\ \leq& \sigma_{max}|\sum_{i=1}^r y^H u_i v_i^H x|\end{aligned}$$

Note that $\|y^Hu_iv_i^H\|_2 \leq \|y^H\|\|u_i\|\|v_i^H\| \leq 1\times 1\times 1 = 1$. Thus,

$$\sigma_{max}|\sum_{i=1}^r y^H u_i v_i^H x| \leq \sigma_{max}|\sum_{i=1}^r x_u| \leq \sigma_{max}$$

We can attain equality when $\|y^Hu_iv_i^H\|_2 = 1$, and $\sum_{i=1}^r (x_i)^2 = 1$, which is possiblely chosen due to the assumption that x and y are free variable of supremam, and U and V, which are the matrices of SVD of A and orthonormal, are arbitrary. Q.E.D

Ex. 4.36

Answer. Try any non-symmetric matrix. For example,

$$A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$$

gives $\lambda_1 = 1, \lambda_2 = 5$ as eigenvalues, but it gives $\sigma_1 = 0.9262, \sigma_2 = 5.3983$ as singular values.

Ex. 4.38

Proof of (i).

$$\begin{aligned} AA^+A =&(U\Sigma V^H)(V\Sigma^{-1}U^H)(U\Sigma V^H) \\ =&U\Sigma\Sigma^{-1}U^H U\Sigma V^H \\ =&U\Sigma V^H \\ =&A \end{aligned}$$

Proof of (ii).

$$\begin{aligned} A^+AA^+ =&(V\Sigma^{-1}U^H)(U\Sigma V^H)(V\Sigma^{-1}U^H) \\ =&U\Sigma^{-1}V^H \\ =&A^+ \end{aligned}$$

Proof of (iii).

$$\begin{aligned} (AA^+)^H =&(A^+)^H A^H \\ =&(V\Sigma^{-1}U^H)^H(U\Sigma V^H)^H \\ =&UU^H \\ =&AA^+ \end{aligned}$$

Proof of (v).

We use the facts above and the fact that if $X = YZ$, then $\mathcal{R}(X) \subseteq \mathcal{R}(Y)$. We need only to show that the matrices AA^+ and A^+A are Hermitian, idempotent, and their ranges are equal to the subspaces on which they are supposed to project.

Both AA^+ and A^+A are obviously Hermitian; see (iii) and (iv). In addition, (i) and (ii) imply that they are idempotent. It remains to show that $\mathcal{R}(AA+) = R(A)$ and $\mathcal{R}(A^+A) = \mathcal{R}(A^H)$. Clearly, $\mathcal{R}(AA^+) \subseteq \mathcal{R}(A)$; $\mathcal{R}(A) \subseteq \mathcal{R}(AA^+)$ follows from (i). From (iv), we have $A^+A = A^H(A^+)^H$, so $\mathcal{R}(A^+A) \subseteq \mathcal{R}(A^H)$. From (i) and (iv), $A^H = A^+AA^H$, so $\mathcal{R}(A^H) \subseteq \mathcal{R}(A^+A)$.