Math, Inner Product Space#3

OSM Lab instructor, John, Van den Berghe

OSM Lab student, CHEN Anhua

Due Wednesday, July 10 at 8:00am

1. (4.2)
   D should take the form of $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$
   $p_D(z) = det(zI - D) = z^3$. If $p_D(z) = 0$, it indicates that eigenvalue of D is 0 with the algebric multiplicity of 3. Also $\mathcal{N}(0I - D) = \mathcal{N}(-D) = span\{[x, 0, 0]^T\}$. The geometric multiplicity is 1, since $dim(\mathcal{N}(-D)) = 1$.

2. (4.4)
   Using 4.3, we could show if $(tr(A)^2 - 4det(A))$ is non-negative, then the matrix only got real eigenvalues, otherwise, it only gets imaginary eigenvalues.
   i. If the matrix is Hermitian, the $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$, then $(tr(A)^2 - 4det(A)) = (a-d)^2 + 4b^2 \geq= 0$
   ii. If the matrix is skew-Hermitian, then $\begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$. Therefore, $(tr(A)^2 - 4det(A)) = -4b^2 < 0$ for $b \neq 0$

3. (4.6)
   Let $A_{n \times n}$ be an upper triangular matrix. $det(\lambda I - A) = 0 \implies \prod_{i=1}^{n}(\lambda_i - a_{ii}) = 0$. Therefore, the eigenvalues of matrix A are its diagonal elements.

4. (4.8)
   i. To prove $\{sin(x), cos(x), sin(2x), cos(2x)\}$ is the basis for V, we need to prove they are linearly independent. This is equivalent to prove that for $\forall x \in \mathbb{R}, asin(x) + bcos(x) + csin(2x) + dcos(2x) = 0 only when a = b = c = d = 0$. When $x = 0, b + d = 0$. When $x = \pi, -b + d = 0$. Therefore, $b = d = 0$. Also, when $x = \frac{\pi}{2}, a - d = 0 \implies a = 0$. Then $\forall x \in \mathbb{R}, dcos(2x) = 0 \implies d = 0$. This completes the proof that S is a basis for V.
   ii. $(asin(x)+bcos(x)+csin(2x)+dcos(2x))' = (-b)sin(x)+acos(x)+(-2d)sin(2x)+2ccos(x)$. So the matrix representation of D is $\begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 2 & 0 \end{bmatrix}$
   iii. $span\{[1, 0, 0, 0], [0, 1, 0, 0]\}$ and $span\{[0, 0, 1, 0], [0, 0, 0, 1]\}$

5. (4.13)
Let $B = P^{-1}AP$, where B is a diagonal matrix. Since P is a non-singular matrix, it indicates that A is diagonalizable. According to the Theorem 4.3.7, A is also semisimple. Then columns of P are just the eigenvectors of A and B is a diagonal marix with eigenvalues on its diagonal. $\lambda = 1$ and 0.4. Therefore, we let $P = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$

6. (4.15)
We will use Theorem 4.3.7 and Proposition 4.3.10 to prove this. According to Theorem 4.3.7, A is semisimple, therefore diagonalizable. We could write $D = P^{-1}AP$ where D is a diagonalizable matrix with eigenvalues of A on its diagonal. And columns of P form the eigenspace of A. According to Proposition 4.3.10, we know that for $i = 1, 2, \dots n, D^i = P^{-1}A^iP)$. Also, becasue P is a linear operator, we have $P^{-1}f(A)P = a_0I + a_1P^{-1}AP + \dots + a_nP^{-1}A^nP = a_0I + a_1D + \dots + a_nD^n$. Since we know that $D^i = \begin{bmatrix} \lambda_1^i & & \\ & \ddots & \\ & & \lambda_n^i \end{bmatrix}$, Then

$$P^{-1}f(A)P = \begin{bmatrix} \sum_{i=0}^n a_i\lambda_1^i & & \\ & \ddots & \\ & & \sum_{i=0}^n a_i\lambda_n^i \end{bmatrix} = \begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{bmatrix}$$

7. (4.16)
i. We have in 4.13 that $D = P^{-1}AP, where D = \begin{bmatrix} 1 & 0 \\ 0 & 0.4 \end{bmatrix}$ and $P = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$.
According to proposition 4.3.10, $A^n = PD^nP^{-1}$. So $\lim_{n\to\infty} A^n = \lim_{n\to\infty} PD^nP^{-1} = P(\lim_{n\to\infty} D^n)P^{-1} = P \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} P^{-1} = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$
ii. It doesn't matter.
iii. Since the eigenvalues of A are 1 and 0.4, according to Theorem 4.3.12, the eigenvalues are $3 + 5 * 1 + 1^3 = 9$ and $3 + 5 * 0.4 + 0.4^3 = 5.064$.

8. (4.18)
Since the characteristic polynomial of A and A^T is the same, therefore, A and A^T have the same eigenvalues. If λ is a eigenvalue of A, it's Also an eigenvalue of A^T. Therefore, $\exists$non-zero x, such that $A^TX = \lambda x \implies x^TA = \lambda x^T$.

9. (4.20)
Since A is Hermitian, $B = U^HAU \implies B^H = (U^HAU)^H = (U^HA^HU) = (U^HAU) = B$.

10. (4.24)
i. If A is Hermitian, according to Corollary 4.4.9, it must have an orthonormal eigen basis V. We represent x in V as $x = Va$, where a is the coefficient vector of x in basis V. Then $\rho(x) = \frac{<x,Ax>}{||x||^2}$
$= \frac{a^H V^H A V a}{a^H a}$
$= \frac{a^H D a}{a^H a}$ (where D is a diagonal matrix, the diagonal elements of which are real eigenvalues of A)
$= \frac{\sum_{i=1}^{n} \lambda_i |a_i|^2}{\sum_{i=1}^{n} |a_i|^2}$ (which has to be real given all its elements are real value)
Since all eigenvlaues of a skew-Hermitian matrix are all imaginary, we could use the same proof techniques above, therefore omitted.

11. (4.25)
i. Since $X = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}$ is orthonormal, therefore, $XX^H = I \implies \sum_{i=1}^{n} x_i x_i^H$
ii. Given the result in i., left-multiplying both side by A will give us that $A = \sum_i^n A x_i x_i^H = \sum_i^n \lambda_i x_i x_i^H$

12. (4.27)
Since A is positive definite, it's also Hermitian. Therefore, all its diagonal elements are real. If any of its diagonal element is not positive, W.L.O.G., we assue $a_{ii} \leq 0$. Then, we let $y = e_i$. $< y, Ay >= y^H A y = a_{ii} \leq 0$, contradicting the fact that A is a positive definite matrix.

13. (4.28)
Since Cauchy Schwartz need to hold for any inner product space, we have $tr(AB) \leq \sqrt{tr(A^2)tr(B^2)}$
$= \sqrt{(\sum_{i=1}^{n} \lambda_{a,i}^2)(\sum_{i=1}^{n} \lambda_{b,i}^2)}$
$\leq \sqrt{(\sum_{i=1}^{n} \lambda_{a,i})^2(\sum_{i=1}^{n} \lambda_{b,i})^2}$
$= tr(A)tr(B)$

14. (4.31)
i. $||A||_2 = \sup_{x \neq 0} \frac{||Ax||_2}{||x||_2}$
$= \sup_{x \neq 0} \frac{||U\Sigma V^H x||_2}{||x||_2}$
$= \sup_{x \neq 0} \frac{||\Sigma V^H x||_2}{||x||_2}$ (invariant under multiplication by orthonormal matrix)
$= \sup_{x \neq 0} \frac{||\Sigma y||_2}{||Vy||_2}$ (let $y = V^H x$)
$= \sup_{x \neq 0} \frac{||\Sigma y||_2}{||y||_2}$ (invariant under multiplication by orthonormal matrix)
$= \sup_{x \neq 0} \frac{(\sum_{i=1}^{n} |\sigma_i y_i|^2)^{\frac{1}{2}}}{(\sum_{i=1}^{n} |y_i|^2)^{\frac{1}{2}}}$
We let y $= [1, 0, \dots 0]^H$, then $||A||_2 = \sigma_1$
ii. Since, based on singular value decomposition, $A^{-1} = (U\Sigma V^H)^{-1} = V\Sigma^{-1}U^H$. The diagonal elements of Σ^{-1} are $\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_n}$, the biggest of which will be $\frac{1}{\sigma_n}$.

Therefore, according to the result of i., we know that $||A^{-1}||_2 = \frac{1}{\sigma_n}$.
iii. It's easy to prove that $||A||_2^H = ||A||_2^T = ||A||_2$ since the Σ of their SVD is the same. $A^HA = V\Sigma^HU^HU\Sigma V^H = V\Sigma^H\Sigma V^H$. Since the diagonal elements of $\Sigma^H\Sigma$ are $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$. By using the results of i., we could know that $||A^HA|| = \sigma_1^2$.
iv. $||UAV||_2^2 = ||UAVV^HA^HU^H||_2 = ||UAA^HU^H||_2 = ||AA^HU^H||_2 = ||(AA^HU^H)^H||_2 = ||UAA^H||_2 = ||AA^H||_2 = ||A||_2^2$

15. (4.32)
i. $||UAV||_F = (tr((UAV)^HUAV))^{\frac{1}{2}} = (tr(V^HA^HAV))^{\frac{1}{2}} = (tr(VV^HA^HA))^{\frac{1}{2}} = (tr(A^HA))^{\frac{1}{2}} = ||A||_F$
ii. Accoding to SVD Theorem, let $A = U\Sigma V^H$, so $||A||_F = tr(A^HA)^{\frac{1}{2}} = tr(V\Sigma^H\Sigma V^H)^{\frac{1}{2}} = tr(V^HV\Sigma^H\Sigma)^{\frac{1}{2}} = tr(\Sigma^H\Sigma)^{\frac{1}{2}} = (\sum_{i=1}^r \sigma_i)^{\frac{1}{2}}$

16. (4.33)
By 4.31, we know that $||A||_2 = \sigma_1$. According to SVD Theorem, let $A = U\Sigma V^H$. Let $x = \sum_{i=1}^n a_iv_i = Va$, where we expressed x in the orthonormal basis of V. Similarly, we let $y = Ub$. Since $x^Hx, y^Hy = 1 \implies a^HV^HVa = 1 \implies a^Ha = 1$ and $b^Hb = 1$.
$\sup_{||x||_2=1,||y||_2=1} |y^HAx|$
$= \sup_{||a||=1,||b||=1} |b^HU^HU\Sigma V^HVa|$
$= \sup_{||a||=1,||b||=1} |b^H\Sigma a|$
$= \sup_{||a||=1,||b||=1} |\sum_{i=1}^n \sigma_ia_ib_i|$
$= \sigma_1$ (According to Cauchy Schwartz inequality)
$= ||A||_2$

17. (4.36)
$A = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$

18. (4.38)
Let $A = U_1\Sigma_1V_1^H$
i. $AA^\dagger A$
$= U_1\Sigma_1V_1^HV_1\Sigma_1^{-1}U_1^HU_1\Sigma_1V_1^H$
$= U_1\Sigma_1V_1^H = A$
ii. $A^\dagger AA^\dagger$
$= V_1\Sigma_1^{-1}U_1^HU_1\Sigma_1V_1^HV_1\Sigma_1^{-1}U_1^H$
$= V_1\Sigma_1^{-1}U_1^H = A^\dagger$
iii.$(AA^\dagger)^H$
$= (V_1\Sigma_1^{-1}U_1^H)^H(U_1\Sigma_1V_1^H)^H$
$= U_1\Sigma_1V_1^HV_1\Sigma_1^{-1}U_1^H$
$= AA^\dagger$
iv.$(A^\dagger A)^H$

$$
\begin{aligned}
&= (U_1\Sigma_1V_1^H)^H(V_1\Sigma_1^{-1}U_1^H)^H \\
&= V_1\Sigma_1^{-1}U_1^HU_1\Sigma_1V_1^H \\
&= A^\dagger A
\end{aligned}
$$