Fall 2016 CSE/CS 383 HW 2 Po-Cheng Pan (UT eid: pp22828)

1. Textbook 3.3

(a) $X \in \mathbb{C}^m = \begin{bmatrix} x(1) \\ x(2) \\ \vdots \\ x(m) \end{bmatrix}$

Without losing the generality, we can suppose $x(j)$ satisfying $\max_i |x(i)|$

$$\|X\|_\infty = \max_i |x(i)| = |x(j)| = (|x(j)|^2)^{1/2} \le (|x(1)|^2 + |x(2)|^2 + \cdots |x(j)|^2 + \cdots |x(m)|^2)^{1/2}$$
$$= \left(\sum_{i=1}^{m} |x(i)|^2\right)^{\frac{1}{2}}$$
$$= \|X\|_2$$

Q.E.D. #

Ex: $e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{C}^m$

$\|e_1\|_\infty = 1 = \|e_1\|_2$ #

(b) According to Cauchy's inequality,

We know that

$$(x(1)^2 + x(2)^2 + \cdots + x(m)^2)(\overbrace{1^2 + 1^2 + \cdots + 1^2}^{m}) \ge [x(1)^2 + x(2)^2 + \cdots + x(m)^2]^2$$

$$\Rightarrow [x(1)^2 + x(2)^2 + \cdots + x(m)^2]^2 \le m \times (x(1)^2 + x(2)^2 + \cdots + x(m)^2)$$
$$\le m \times (\underbrace{x(j)^2 + x(j)^2 + \cdots x(j)^2}_{m}) \quad (x(j) \text{ satisfies } \max_i |x(i)|)$$
$$= m^2 \cdot x(j)^2$$

$$\Rightarrow (x(1)^2 + x(2)^2 + \cdots x(m)^2) \le m \cdot |x(j)|$$
$$\Rightarrow [x(1)^2 + x(2)^2 + \cdots + x(m)^2]^{1/2} \le \sqrt{m}\sqrt{|x(j)|} \le \sqrt{m}\,|x(j)| = \sqrt{m}\,\|X\|_\infty$$
$$\Rightarrow \|X\|_2 \le \sqrt{m}\,\|X\|_\infty$$

Q.E.D. #

Ex: if $X = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{C}^m \Rightarrow \|X\|_2 = \sqrt{m} = \sqrt{m} \cdot 1 = \sqrt{m}\,\|X\|_\infty$ #

1. Textbook 3.3

(c) $\|A\|_\infty = \sup_{x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} \leq \sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_\infty}$ ($\because \|x\|_\infty < \|x\|_2$ from (a))

($x \in \mathbb{C}^n$ $\because A \in \mathbb{C}^{m\times n}$) $\leq \sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2/\sqrt{n}}$ ($\because x \in \mathbb{C}^n$, we know $\|x\|_2 \leq \sqrt{n}\|x\|_\infty$ from (b))

$= \sqrt{n}\|A\|_2$

Q.E.D. #

For equality, let $A \in \mathbb{C}^{m\times n}$ to be the matrix whose first row is all ones and zeros elsewhere

$\Rightarrow \|A\|_\infty = n$ $\quad \|A\|_2 = \sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}$ Let $x = \begin{bmatrix} x(1) \\ x(2) \\ \vdots \\ x(n) \end{bmatrix} \in \mathbb{C}^{n\times 1}$

$= \sup_{x\neq 0} \frac{\sqrt{(x(1)+x(2)+\cdots+x(n))^2}}{\sqrt{x(1)^2+x(2)^2+\cdots+x(n)^2}}$

By Cauchy inequality, we know

$(x(1)^2+x(2)^2+\cdots+x(n)^2)(\underbrace{1^2+\cdots+1^2}_{n}) \geq (x(1)+x(2)+\cdots+x(n))^2$

$\Rightarrow \sqrt{\frac{(x(1)+\cdots+x(n))^2}{x(1)^2+\cdots+x(n)^2}} \leq \sqrt{n}$

$= \sqrt{n}$

$\Rightarrow \|A\|_\infty = n = \sqrt{n}\cdot\sqrt{n} = \sqrt{n}\|A\|_2$ #

(d) $\|A\|_2 = \sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} \leq \sup_{x\neq 0} \frac{\sqrt{m}\|Ax\|_\infty}{\|x\|_2}$ ($\because Ax \in \mathbb{C}^m$, we know $\|Ax\|_2 \leq \sqrt{m}\|Ax\|_\infty$ from (b))

$\leq \sup_{x\neq 0} \frac{\sqrt{m}\|Ax\|_\infty}{\|x\|_\infty}$ ($\because \|x\|_\infty \leq \|x\|_2$ from (a))

$\leq \sqrt{m}\|A\|_\infty$ Q.E.D. #

For equality, let $A \in \mathbb{C}^{m\times n}$ to be the matrix whose first colum is all ones, and zero elsewhere

$\Rightarrow \|A\|_\infty = 1$ $\quad \|A\|_2 = \sup_{x\neq 0} \frac{\sqrt{\overbrace{x(1)^2+x(1)^2+\cdots+x(1)^2}^{m}}}{\sqrt{x(1)^2+x(2)^2+\cdots+x(n)^2}}$

To maximize it, we let $x = \begin{bmatrix} x(1) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{C}^{n\times 1}$

$\Rightarrow \|A\|_2 = \sqrt{\frac{m\times x(1)^2}{x(1)^2}}$

$= \sqrt{m}$

$\Rightarrow \|A_2\| = \sqrt{m} = 1\times\sqrt{m} = \sqrt{m}\|A\|_\infty$

#

1. Textbook 3.4

$A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{\mu\times\nu}$ $(\mu \le m, \nu \le n)$

(a) $I_m \in \mathbb{C}^{m\times m} = \begin{bmatrix} - e_1^* - \\ - e_2^* - \\ \vdots \\ - e_m^* - \end{bmatrix}$, $e_i \in \mathbb{C}^m$ for $1 \le i \le m = \begin{array}{c} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \\ \leftarrow i\text{th} \\ \\ \\ \\ \end{array}$

Consider $I_m \cdot A = I_m \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} e_1^* a_1 & e_1^* a_2 & \cdots & e_1^* a_n \\ & \vdots & & \\ e_m^* a_1 & e_m^* a_2 & \cdots & e_m^* a_n \end{bmatrix} = \begin{bmatrix} - r_1 - \\ - r_2 - \\ \vdots \\ - r_m - \end{bmatrix} = A$

If we consider to delete row 1、row 2 from A

we can make a matrix $D_r \in \mathbb{C}^{m-2\times n} = \begin{bmatrix} - e_3^* - \\ - e_4^* - \\ \vdots \\ - e_n^* - \end{bmatrix}$

$\Rightarrow DA = \begin{bmatrix} e_3^* a_1 & e_3^* a_2 & \cdots & e_3^* a_n \\ & \vdots & & \\ e_m^* a_1 & e_m^* a_2 & \cdots & e_m^* a_n \end{bmatrix} = \begin{bmatrix} - r_3 - \\ - r_4 - \\ \vdots \\ - r_m - \end{bmatrix}$

Hence, if we want to delete some rows from A to make its row number to be μ

we can build a matrix $D_r \in \mathbb{C}^{\mu\times m} = \begin{bmatrix} - d_{r1} - \\ - d_{r2} - \\ \vdots \\ - d_{r\mu} - \end{bmatrix}$ where $d_{rj} = e_{k_j}^*$ for $1 \le j \le \mu$, $1 \le k_j \le m$

& $k_1 < k_2 < k_3 < \cdots < k_{\mu-1} < k_\mu$

D_r is the "row deletion matrices" #

By the similar method

Consider $A \cdot I_n = \begin{bmatrix} - r_1 - \\ - r_2 - \\ - r_3 - \\ \vdots \\ - r_m - \end{bmatrix} \cdot \begin{bmatrix} | & | & & | \\ e_1 & e_2 & \cdots & e_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} r_1 e_1 & r_1 e_2 & \cdots & r_1 e_n \\ r_2 e_1 & r_2 e_2 & \cdots & r_2 e_n \\ \vdots & \vdots & \ddots & \vdots \\ r_m e_1 & r_m e_2 & \cdots & r_m e_n \end{bmatrix} = A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix}$

If we want to delete colum 1 & 2 from A

we can make a matrix $D_c \in \mathbb{C}^{n\times n-2} = \begin{bmatrix} | & | & & | \\ e_3 & e_4 & \cdots & e_n \\ | & | & & | \end{bmatrix}$

$\Rightarrow A D_c = \begin{bmatrix} r_1 e_3 & \cdots & r_1 e_n \\ r_2 e_3 & \cdots & r_2 e_n \\ \vdots & & \vdots \\ r_m e_3 & \cdots & r_m e_n \end{bmatrix} = \begin{bmatrix} | & | & & | \\ a_3 & a_4 & \cdots & a_n \\ | & | & & | \end{bmatrix}$

Hence, if we want to delete some colums of A to make its column number to be ν

we can build a matrix $D_c \in \mathbb{C}^{n\times\nu} = \begin{bmatrix} | & | & & | \\ d_{c1} & d_{c2} & \cdots & d_{c\nu} \\ | & | & & | \end{bmatrix}$ where $d_{ci} = e_{k_i}$ for $1 \le i \le \nu$, $1 \le k_i \le n$

& $k_1 < k_2 < \cdots < k_{\nu-1} < k_\nu$

D_c is the "column deletion matrices" #

That is

$D_r A D_c = B$ #

1. Textbook 3.4 $\quad D_r \in \mathbb{C}^{\mu \times m} \quad D_c \in \mathbb{C}^{n \times \nu}$

(b) $\|B\|_p = \|D_r A D_c\|_p \le \|D_r\|_p \|A D_c\|_p \quad -①$

$\|D_r\|_p = \max_{x \in \mathbb{C}^m} \frac{\|D_r x\|_p}{\|x\|_p} \le 1 \quad -②$ $\quad \because D_r x \in \mathbb{C}^{\mu}$ which is the subvector of x $\Rightarrow \|D_r x\|_p \le \|x\|_p$

From ① & ②; $\|B\|_p \le \|A D_c\|_p \le \|A\|_p \|D_c\|_p \quad -③$

$\because D_c$ is a subspace of the identity matrix $I_n \in \mathbb{C}^{n \times n}$

$\therefore \|D_c\|_p \le \|I_n\|_p = 1 \quad -④$ $\left(\because \sup_{x \ne 0} \frac{\|I_n x\|_p}{\|x\|_p} = \sup_{x \ne 0} \frac{\|x\|_p}{\|x\|_p} = 1 \right)$

By ③ & ④ $\|B\|_p \le \|A\|_p$ for any p with $1 \le p \le \infty$

Q.E.D.

#

2. Textbook 6.2

$Ex = (x + Fx)/2 \qquad E, F \in \mathbb{C}^{m \times m}, \; x \in \mathbb{C}^m$

$= \left(\frac{I+F}{2}\right) x \qquad F$ is a flip matrix $= \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & & 1 & 0 \\ 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{bmatrix}$

$\Rightarrow \quad E = \frac{I+F}{2}$

If E is a projector, $F^2 = I$

E^2 should be E

$E^2 = \left(\frac{I+F}{2}\right)^2$

$= \frac{I^2 + 2F + F^2}{4}$

$= \frac{I^2 + 2F + I^2}{4}$

$= \left(\frac{I+F}{2}\right) = E$

$\Rightarrow$ So, E is a projector #

$E^* = \left(\frac{I+F}{2}\right)^*$

$= \frac{I+F}{2}$

$= E$

$\Rightarrow P$ is orthogonal #

Its entries:

$\frac{I+F}{2}$

if m is odd

$E_{ij} = \begin{cases} 1/2, & \text{if } i=j \text{ but } i \ne \frac{m+1}{2} \\ & \text{or } i+j=m+1 \text{ but } i \ne \frac{m+1}{2} \\ 1, & \text{if } i=j=\frac{m+1}{2} \\ 0, & \text{others} \end{cases}$ for $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, m$ #

if m is even,

$E_{ij} = \begin{cases} 1/2, & \text{if } i=j \text{ or } i+j=m+1 \\ 0, & \text{others} \end{cases}$ for $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, m$ #

2. Textbook 6.3

$A \in \mathbb{C}^{m \times n}$ with $m \geq n$

If A^*A is nonsigular

From Theorem 1.3 on textbook

A^*A has an inverse $(A^*A)^{-1}$

rank $(A^*A) = n$ ($\because A^*A \in \mathbb{C}^{n \times n}$)

null $(A^*A) = \{0\}$

that means :

$A^*AX=0$ only when $X=0 \in \mathbb{C}^n$

So we know that $AX=0$ is also only when $X=0 \in \mathbb{C}^n$

$\Rightarrow \dim(\text{Null}(A))=0$

By Rank-nullity theorem, we know $\dim(\text{rank}(A)) + \underbrace{\dim(\text{Null}(A))}_{=0} = \dim(\mathbb{C}^n) = n$

$\Rightarrow \dim(\text{rank}(A)) = n \Rightarrow A$ is full-rank

Q.E.D. #

If A has full rank ($\because \dim(\text{Range}(A)) = n$)

$\Rightarrow \text{null}(A) = \{0\}$

$\Rightarrow AX=0$ only when $X=0 \in \mathbb{C}^n$

$\Rightarrow A$ has n nonzero eigenvalues

According to Theorem 5.4 on the textbook, the nonzero eigenvalues of A^*A are the square of the nonzero eigenvalues of A

We know that

Hence, A^*A also has n nonzero eigenvalues

& $A^*A \in \mathbb{C}^{n \times n}$

So, A^*A is full rank. That is, it's nonsingular

Q.E.D.

#

2. Textbook 6.5

$P \in \mathbb{C}^{m \times n}$, a nonzero projector

① If $\|P\|_2 = 1$

Suppose $u \in Range(P)$ $v \in Null(P)$ & $\|v\|_2 = 1$

Make an orthogonal projection of u onto v

s.t. $u = v^* u v + w$

$$\Rightarrow \|u\|_2^2 = \|w\|_2^2 + \|v^* u v\|_2^2 + 0 \quad (\because \text{It's a orthogonal projection} \therefore v^* u v \perp w)$$

$$= \|w\|_2^2 + \|v^* u\|_2^2 \|v\|_2^2 \quad (\because v^* u \in \mathbb{C})$$

$$= \|w\|_2^2 + \|v^* u\|_2^2 \quad (\because \|v\|_2^2 = 1)$$

$$Pw = P(u - v^* u v) \quad (\because u \in Range(P) \text{ \& } v \in Null(P))$$

$$= Pu = u$$

$$\Rightarrow \frac{\|Pw\|_2^2}{\|w\|_2^2} = \frac{\|u\|_2^2}{\|w\|_2^2} = \frac{\|w\|_2^2 + \|v^* u\|_2^2}{\|w\|_2^2} \le 1 \quad (\because \|P\|_2 = 1)$$

So, $\|v^* u\|_2^2$ must to be zero.

$\Rightarrow Range(P) \perp Null(P) \Rightarrow P$ is an orthogonal projector.

Q.E.D.

#

② If P is an orthogonal projection

Suppose $u \in Range(P)$ $v \in Null(P)$ $w = u + v$

$$\frac{\|Pw\|_2^2}{\|w\|_2^2} = \frac{\|P(u+v)\|_2^2}{\|w\|_2^2} = \frac{\|Pu\|_2^2}{\|u\|_2^2 + \|v\|_2^2 + u^* v + v^* u} = \frac{\|Pu\|_2^2}{\|u\|_2^2 + \|v\|_2^2} \quad (\because u \perp v)$$

$$= \frac{\|u\|_2^2}{\|u\|_2^2 + \|v\|_2^2} \quad (\because Pu = u)$$

$$\le 1 \Rightarrow \|Pw\|_2^2 \le \|w\|_2^2 \quad ⓐ$$

Besides, $\|Pw\| = \|w\| \le \|P\| \|w\|$ ⓑ

To satisfy ⓐ & ⓑ, $\|P\|$ must be 1

Q.E.D.

#

3. $P \in \mathbb{R}^{m \times m}$, $Range(P) = Range(A)$, $Range(B) \perp Null(P)$, A, B are full rank

We want to find a matrix P s.t. $P^2 = P$ & $PV = V$ for $V \in Range(P)$

$PV = 0$ for $V \in Null(P)$

$\because$ $Range(B) \perp Null(P)$

$\therefore$ For every $v \in Null(P)$, $B^T V = 0$

Hence, any $P = P_1 B^T$ also has the same property, $PV = P_1 B^T V = \underline{0}$ $\forall \bar{v} \in Null(P)$ — ①

Next, consider any $V \in Range(P)$

$AX = V$ for some X

$\Rightarrow B^T A X = B^T V$ ($\because$ $B^T A$ is invertible $\because$ B^T & A both are full rank)

$\Rightarrow X = (B^T A)^{-1} B^T V$

If P is a projection, $PV = V$

So, $V = PV = A(B^T A)^{-1} B^T V$

$\Rightarrow P = A(B^T A^{-1}) B^T$ # (from ①, we know P will satisfy $PV = 0$ $\forall V \in Null(P)$)

4. $A \in \mathbb{R}^{m\times n}$, $m \ge n$, full rank, Q & R are its QR factors

(a) $A = QR$ Suppose R is an upper-triangular matrix, Q is an orthonormal matrix

→ If we take this as reduced QR factorization

$\Rightarrow Q \in \mathbb{R}^{m\times n}$ $R \in \mathbb{R}^{n\times n}$

We can find a reduced SVD for R

s.t. $R = U_R \Sigma_R V_R^T$

Suppose A has reduced SVD $\Rightarrow A = U\Sigma V^T$

$\Rightarrow A = QR = QU_R\Sigma_R V_R^T$ $\left(\begin{array}{l} QU_R = U \\ \Sigma_R = \Sigma \\ V_R^T = V^T \end{array}\right)$

#

$\because (QU_R)^* \cdot (QU_R)$

$= U_R^* Q^* Q U_R$ ($\because U_R$ & Q are both unitary)

$= I$

$\therefore$ we can use $QU_R = U$, $\Sigma_R = \Sigma$, $V_R^T = V_T$ for A

(b) $[I;A] = \begin{bmatrix} I \\ A \end{bmatrix} \in \mathbb{R}^{(n+m)\times n}$.

$= A_I$

From Theorem 5.4 on the textbook, we know that the nonzero singular values of A are the square roots of the nonzero eigenvalues of $A_I^T A_I$ — ②

$A_I^T A_I = [I \; A^T] \begin{bmatrix} I \\ A \end{bmatrix}$

$= I^2 + A^TA$

$= I + A^TA$

$= I + (QR)^T QR$

$= I + R^TR$ ($\because Q^TQ = I$)

We want to find $\lambda \in \mathbb{R}$ s.t. $(I+R^TR)x = \lambda x$ for some $x \in \mathbb{R}^n$

$\Rightarrow R^TRx = (\lambda - 1)x$

$= \lambda_R x$ ($\lambda_R = \lambda - 1$)

From Theorem 5.4, we know the nonzero eigenvalues of R^TR are the square of the nonzero eigenvalues of R. — ①

$\because \lambda_R = \lambda - 1 \Rightarrow \lambda = \lambda_R + 1$ & ①

$\therefore$ The nonzero eigenvalues of $A_I^TA_I$ are one plus the square of the nonzero eigenvalues of R

Then from ② we know that the nonzero singular values of A_I are the square roots of one plus the square of the nonzero eigenvalues of R

That is

for λ are nonzero eigenvalues of R

$\Rightarrow$ nonzero singular values of A_I are $\sqrt{\lambda^2+1}$

#