Anar Rzayev MAS 365 - HW 3 ID: 20190788

1) a) Recall the properties for which a matrix norm has to satisfy for all $n\times n$ matrices $A, B$ and $\alpha \in \mathbb{R}$:

→ $\|A\| \geq 0$ and $\|A\| = 0 \iff A = 0$ ①

→ $\|\alpha A\| = |\alpha| \, \|A\|$ ②

→ $\|A+B\| \leq \|A\| + \|B\|$ ③

→ $\|AB\| \leq \|A\| \cdot \|B\|$ ④

① $\|A\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \underbrace{|a_{ij}|^2}_{\geq 0}} \geq 0$ and if $\|A\|_F = 0$, then $\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 = 0$ and $a_{ij} = 0$ (otherwise, if there would exist nonzero term, then $\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 > 0$ would be true)

So, $a_{ij} = 0 \Rightarrow \underline{A = 0}$ Other direction, If $A = 0 \Rightarrow a_{ij} = 0$ for all $i, j$ and $\|A\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2} = 0$

② $\|\alpha A\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |\alpha a_{ij}|^2} = \sqrt{|\alpha|^2 \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2} =$

$= \sqrt{|\alpha|^2} \cdot \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2} = |\alpha| \cdot \|A\|_F$

③ Using the famous inequality Cauchy-Bunyakovsky-Schwarz, we have

$$\sqrt{(x_1^2+\dots+x_m^2)(y_1^2+\dots+y_m^2)} \geq x_1y_1+\dots+x_my_m$$

If we take correspondent values for each $x_i, y_j \Rightarrow$

$$\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2 \cdot \sum_{i=1}^{n}\sum_{j=1}^{n}|b_{ij}|^2} \geq \sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|\,|b_{ij}| \quad (*)$$

From the geometric inequality, $|a_{ij}+b_{ij}| \leq |a_{ij}|+|b_{ij}|$

So, $$\|A+B\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}+b_{ij}|^2} \leq \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}(|a_{ij}|+|b_{ij}|)^2}$$

$$= \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2 + \sum_{i=1}^{n}\sum_{j=1}^{n}|b_{ij}|^2 + 2\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|\,|b_{ij}|} \overset{(*)\ \text{ineq}}{\leq}$$

$$\leq \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2 + \sum_{i=1}^{n}\sum_{j=1}^{n}|b_{ij}|^2 + 2\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2}\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|b_{ij}|^2}}$$

$$= \sqrt{\|A\|_F^2+\|B\|_F^2+2\|A\|_F\cdot\|B\|_F} = \sqrt{(\|A\|_F+\|B\|_F)^2} =$$

$= \|A\|_F + \|B\|_F$ since from the definition, $\|A\|_F \geq 0$

$\Rightarrow \|A\|_F+\|B\|_F \geq 0$ or just $\|A+B\|_F \leq \|A\|_F + \|B\|_F$ ✓

(4) From the def of matrix multiplic, $AB = C = (c_{ij})$ where

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \Rightarrow \|AB\|_F = \|C\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |c_{ij}|^2} =$$

$$= \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right|^2} \overset{(*)}{\leq} \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \left( \sqrt{\sum_{k=1}^{n} |a_{ik}|^2} \cdot \sqrt{\sum_{k=1}^{n} |b_{kj}|^2} \right)^2}$$

$$= \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \left( \sum_{k=1}^{n} |a_{ik}|^2 \cdot \sum_{k=1}^{n} |b_{kj}|^2 \right)} = \sqrt{\sum_{i=1}^{n} \sum_{k=1}^{n} |a_{ik}|^2 \cdot \sum_{j=1}^{n} \sum_{k=1}^{n} |b_{kj}|^2}$$

$$= \sqrt{\|A\|_F^2 \cdot \|B\|_F^2} = \|A\|_F \cdot \|B\|_F$$ since each of the values $\|A\|_F, \|B\|_F \geq 0$ and $(*)$ is depicting the CBS inequality (cauchy)

$$\sum_{k=1}^{n} a_{ik} b_{kj} \leq \sqrt{\sum_{k=1}^{n} a_{ik}^2 \cdot \sum_{k=1}^{n} b_{kj}^2} = \sqrt{\sum_{k=1}^{n} |a_{ik}|^2 \cdot \sum_{k=1}^{n} |b_{kj}|^2} =$$

$$= \sqrt{\sum_{k=1}^{n} |a_{ik}|^2} \cdot \sqrt{\sum_{k=1}^{n} |b_{kj}|^2}$$ or just $$\left| \sum_{k=1}^{n} a_{ik} b_{kj} \right|^2 \leq$$

$$\leq \sum_{k=1}^{n} |a_{ik}|^2 \cdot \sum_{k=1}^{n} |b_{kj}|^2 = \left( \sqrt{\sum_{k=1}^{n} |a_{ik}|^2} \cdot \sqrt{\sum_{k=1}^{n} |b_{kj}|^2} \right)^2$$ so

we found $\boxed{\|AB\|_F \leq \|A\|_F \cdot \|B\|_F}$ ✓ In conclusion, we obtained that $\boxed{\|\cdot\|_F - \text{matrix norm}}$ ✓

b) By def, $\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 =$

$$= \max_{\|x\|_2=1} \left\| \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right\|_2 = \max_{\|x\|_2=1} \left\| \begin{bmatrix} \sum_{i=1}^{n} x_i a_{1i} \\ \sum_{i=1}^{n} x_i a_{2i} \\ \vdots \\ \sum_{i=1}^{n} x_i a_{ni} \end{bmatrix} \right\|_2$$

$$= \max_{\|x\|_2=1} \sqrt{\left(\sum_{i=1}^{n} x_i a_{1i}\right)^2 + \left(\sum_{i=1}^{n} x_i a_{2i}\right)^2 + \dots + \left(\sum_{i=1}^{n} x_i a_{ni}\right)^2}$$

$$\overset{\star}{\leq} \max_{\|x\|_2=1} \sqrt{\left(\sqrt{\sum_{i=1}^{n} x_i^2 \cdot \sum_{i=1}^{n} a_{1i}^2}\right)^2 + \dots + \left(\sqrt{\sum_{i=1}^{n} x_i^2 \cdot \sum_{i=1}^{n} a_{ni}^2}\right)^2}$$

$$= \max_{\|x\|_2=1} \sqrt{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} a_{1i}^2 + \dots + \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} a_{ni}^2} = \max_{\|x\|_2=1} \sqrt{\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} a_{1i}^2 + \dots + \sum_{i=1}^{n} a_{ni}^2\right)}$$

$$= \underbrace{\max_{\|x\|_2=1} \|x\|_2}_{=1} \sqrt{\sum_{j=1}^{n}\sum_{i=1}^{n} a_{ji}^2} = \|A\|_F$$

where $\star$ describes the CBS-inequality applied to all the each of n terms. Hence, we found

$\boxed{\|A\|_2 \leq \|A\|_F}$ $\star$ is true $\boxed{V}$ ■ +

2) a) Since $\|A\|_2 = \sqrt{\rho(A^TA)}$ and $A$-symmetric ($A = A^T$) we get $\boxed{\|A\|_2 = \sqrt{\rho(A^2)}}$ Let $\lambda$-eigenvalue of $A$ with associated eigenvector $x$ ($Ax = \lambda x$) $\Rightarrow A^2x = A(Ax) = A(\lambda x) = \lambda(Ax) = \lambda^2 x$; Hence, if $\lambda_1, \dots, \lambda_n$-eigenvalues of $A \Rightarrow \lambda_1^2, \dots, \lambda_n^2$-eigenvalues of $A^2$.

Notice that symmetric matrices have real eigenvalues $\Rightarrow \lambda_i^2$-positive reals. Assume $\lambda_{max}^2$-largest eigenvalue of $A^2$

Then, $\lambda_{max}^2 = |\lambda_{max}^2| = \rho(A^2)$ and $\lambda_{max}^2 = |\lambda_{max}|^2 = \rho(A)^2 \Rightarrow$ Thus, $\boxed{\rho(A^2) = \rho(A)^2}$ and $\|A\|_2 = \sqrt{\rho(A^2)} = \sqrt{\rho(A)^2} = |\rho(A)| = \rho(A)$ since $\rho(A) = \max \underbrace{|\lambda_i|}_{\geq 0}$ where $\lambda_i$-eigenvalue of $A \Rightarrow \rho(A) \geq 0$ becomes true (for symmetric $A$, $\lambda_i \in \mathbb{R}$) with $\boxed{\|A\|_2 = \rho(A)}$ ☆ $\boxed{V}$ ∎

($u$-eigenvalue of $A^2 \Rightarrow u > 0$ and $\sqrt{u}$-eigenvalue of $A$. So, if $\lambda_{max} \to$ max eigenvalue of $A \Rightarrow \lambda_{max}^2 \to$ max eigenvalue of $A^2$)

b) Recall the property that $\rho(A) \leq \|A\|$ for any natural norm $\|\cdot\|$ where $\rho(A) = \max\limits_{\lambda\text{-eigenvalue}} |\lambda|$

From these, we infer that $|\lambda| \leq \rho(A) \leq \|A\|$ for

Scanned with CamScanner

any eigenvalue $\lambda \Rightarrow |\lambda| \le ||A||$ holds ✓ (since $\rho(A)$ is max among the values $|\lambda|$, it should be $\ge |\lambda|$ for $\forall\lambda$)

Meanwhile, notice that if $\lambda$ - eigenvalue of $A$, then $\frac{1}{\lambda}$ - eigenvalue of $A^{-1}$ (considering that $A$ - nonsingular)

Then, we get $\frac{1}{|\lambda|} \le \rho(A^{-1}) \le ||A^{-1}||$ or $|\lambda|^{-1} \le \rho(A^{-1}) \le ||A^{-1}||$ implies $\frac{1}{||A^{-1}||} = ||A^{-1}||^{-1} \le |\lambda|$ Combining a results yields $\boxed{||A^{-1}||^{-1} \le |\lambda| \le ||A|| \text{ for any } \lambda\text{-eigenvalue of nonsingular matrix } A}$

$\boxed{V}$ ■ $\oplus$

<u>Note:</u> Since $\rho(A) = \max\limits_{\lambda\text{-eigenvalue of } A} |\lambda| \Rightarrow \rho(A)$ will be $\ge$ for any value $|\lambda|$; meaning $\rho(A) \ge |\lambda|$ for any $\lambda$-eigenvalue is true (and obvious) from the definition itself ✓

<u>Note:</u> If $Ax = \lambda x \Rightarrow |\lambda| \cdot ||x|| = ||Ax|| \le ||A|| \cdot ||x||$ which also implies $|\lambda| \le ||A||$ ✓ Also, $\left(\frac{1}{\lambda}\right) x = A^{-1}x$, concluding $\lambda^{-1} \rightarrow$ eigenvalue of $A^{-1}$ (where $A$ - nonsingular matrix)

Thus, using previous result for matrix $A^{-1}$ and eigenvalue $\lambda^{-1} \Rightarrow |\lambda^{-1}| = \frac{1}{|\lambda|} \le ||A^{-1}||$ or $\frac{1}{||A^{-1}||} = ||A^{-1}||^{-1} \le |\lambda|$ ✓

Implying $||A^{-1}||^{-1} \le |\lambda| \le ||A||$ $\boxed{V}$

Scanned with CamScanner

3) The linear system can be denoted as $Ax=b$, where

$$A=\begin{bmatrix}1 & 2 & -2\\ 1 & 1 & 1\\ 2 & 2 & 1\end{bmatrix} \text{ and } b=\begin{bmatrix}7\\ 2\\ 5\end{bmatrix}$$

Jacobi

a) From the given concepts, we know $T_j = D^{-1}(L+U)$

with $D=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$, $L=\begin{bmatrix}0&0&0\\-1&0&0\\-2&-2&0\end{bmatrix}$, $U=\begin{bmatrix}0&-2&2\\0&0&-1\\0&0&0\end{bmatrix}$

Since $D$ is an identity matrix $\Rightarrow D^{-1}=D=I$ and

$$L+U=\begin{bmatrix}0&-2&2\\-1&0&-1\\-2&-2&0\end{bmatrix} \Rightarrow \boxed{T_j=\begin{bmatrix}0&-2&2\\-1&0&-1\\-2&-2&0\end{bmatrix}}$$

The eigenvalues of $T_j$ can be found from the

$$\det(T_j-\lambda I)=\det\left(\begin{bmatrix}-\lambda&-2&2\\-1&-\lambda&-1\\-2&-2&-\lambda\end{bmatrix}\right)=-\lambda\cdot\det\left(\begin{bmatrix}-\lambda&-1\\-2&-\lambda\end{bmatrix}\right)$$

$$-(-2)\cdot\det\left(\begin{bmatrix}-1&-1\\-2&-\lambda\end{bmatrix}\right)+2\cdot\det\left(\begin{bmatrix}-1&-\lambda\\-2&-2\end{bmatrix}\right)=$$

$$=-\lambda(\lambda^2-2)+2(\lambda-2)+2(2-2\lambda)=-\lambda^3\cancel{+2\lambda}\cancel{+2\lambda}-$$

$$-4+4\cancel{-4\lambda}=-\lambda^3=0\Rightarrow \boxed{\lambda=0}$$ is the only eigenvalue of $T_j$

Thus, $\rho(T_j)$ - maximum value of the absolute value of eigenvalues of $T_j$ and since $\lambda=0$ is the only eigenvalue

$\Rightarrow \boxed{\rho(T_j)=0}$ Using Jacobi's method, $x^{(k)}=(x_1^{(k)},x_2^{(k)},x_3^{(k)})$

$X^{(k)} = \left(X_1^{(k)}, X_2^{(k)}, X_3^{(k)}\right)^t$ can be calculated from $X^{(k-1)}$ by:

$$\left.\begin{aligned} X_1^{(k)} &= -2X_2^{(k-1)} + 2X_3^{(k-1)} + 7 \\ X_2^{(k)} &= -X_1^{(k-1)} - X_3^{(k-1)} + 2 \\ X_3^{(k)} &= -2X_1^{(k-1)} - 2X_2^{(k-1)} + 5 \end{aligned}\right\}$$

where $c_j = D^{-1}b = b = \begin{bmatrix} 7 \\ 2 \\ 5 \end{bmatrix}$ and $T_j = \begin{bmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix}$

Using $X^{(0)} = 0$, we obtain

$X_1^{(1)} = 7,\ X_2^{(1)} = 2,\ X_3^{(1)} = 5 \Rightarrow \boxed{X^{(1)} = (7, 2, 5)^t}$ Similarly,

$$\left.\begin{aligned} X_1^{(2)} &= -2\cdot 2 + 2\cdot 5 + 7 = -4 + 17 = 13 \\ X_2^{(2)} &= -7 - 5 + 2 = -10 \\ X_3^{(2)} &= -2\cdot 7 - 2\cdot 2 + 5 = -14 - 4 + 5 = -13 \end{aligned}\right\} \boxed{\begin{aligned} X^{(1)} &= (7.0, 2.0, 5.0)^t \\ X^{(2)} &= (13.0, -10.0, -13.0)^t \end{aligned}}$$

| $k$ | $X_1^{(k)}$ | $X_2^{(k)}$ | $X_3^{(k)}$ |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 1 | 7.0 | 2.0 | 5.0 |
| 2 | 13.0 | -10.0 | -13.0 |

$$\boxed{\begin{aligned} X^{(1)} &= (7.0, 2.0, 5.0)^t \\ X^{(2)} &= (13.0, -10.0, -13.0)^t \end{aligned}}$$

(accurate solution occurs after 4 steps)

b) The matrix $T_g$ is calculated as $(D-L)^{-1}U$ where $D, L, U$ were revealed in the previous exercise

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

and if we denote inverse of $D-L$

$$(D-L)^{-1} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \Rightarrow (D-L)(D-L)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

or just equivalently

$a_1 = 1,\ a_2 = a_3 = 0,\ b_1 + b_2 + 2b_3 = 0,\ b_2 + 2b_3 = 1,\ b_3 = 0$

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \left.\begin{array}{l} a_1 + a_2 + 2a_3 = 1 \\ b_1 + b_2 + 2b_3 = 0 \\ c_1 + c_2 + 2c_3 = 0 \end{array}\right\}$$

$$\left.\begin{array}{l} a_2 + 2a_3 = 0 \\ b_2 + 2b_3 = 1 \\ c_2 + 2c_3 = 0 \end{array}\right\} \quad \boxed{\begin{array}{l} a_3 = 0 \\ b_3 = 0 \\ c_3 = 1 \end{array}} \quad \boxed{\begin{array}{l} a_2 = 0 \\ b_2 = 1 \\ c_2 = -2 \end{array}} \quad \begin{array}{l} a_1 = 1 \Rightarrow \\ b_1 + 1 = 0 \Rightarrow \\ c_1 - 2 + 2 = 0 \Rightarrow \end{array} \boxed{\begin{array}{l} a_1 = 1 \\ b_1 = -1 \\ c_1 = 0 \end{array}}$$

Thus, $(D-L)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$ and $T_g = (D-L)^{-1} U =$

$$= \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0 & -2 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -2 & 2 \\ 0 & 2 & -3 \\ 0 & 0 & 2 \end{bmatrix} \Rightarrow \boxed{T_g = \begin{bmatrix} 0 & -2 & 2 \\ 0 & 2 & -3 \\ 0 & 0 & 2 \end{bmatrix}}$$

Then, eigenvalues will be found by $\det(T_g - \lambda I) = 0$

$$\det\left( \begin{bmatrix} -\lambda & -2 & 2 \\ 0 & 2-\lambda & -3 \\ 0 & 0 & 2-\lambda \end{bmatrix} \right) = -\lambda \cdot \det\left( \begin{bmatrix} 2-\lambda & -3 \\ 0 & 2-\lambda \end{bmatrix} \right) +$$

$$+ 2 \cdot \det\left( \begin{bmatrix} 0 & -3 \\ 0 & 2-\lambda \end{bmatrix} \right) - 2 \cdot \det\left( \begin{bmatrix} 0 & 2-\lambda \\ 0 & 0 \end{bmatrix} \right) = 0 \Rightarrow$$

$-\lambda(2-\lambda)^2 = 0$ or just $\boxed{\lambda = 0 \text{ and } \lambda = 2}$ are the only eigenvalues of $T_g$

Therefore, we get $\rho(T_g) = \max\limits_{\lambda\text{-eigenvalue}} \{|\lambda|\} = 2$; $\boxed{\rho(T_g) = 2}$

$$c_g = (D-L)^{-1} b = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 7 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 7 \\ -5 \\ 1 \end{bmatrix}$$

From the method of Gauss-Seidel, we know that

$X^{(k)}$ can be found from $X^{(k-1)}$ by using the formula

$$X^{(k)} = T_g X^{(k-1)} + C_g \Rightarrow \begin{cases} X_1^{(k)} = -2X_2^{(k-1)} + 2X_3^{(k-1)} + 7 \\ X_2^{(k)} = 2X_2^{(k-1)} - 3X_3^{(k-1)} - 5 \\ X_3^{(k)} = 2X_3^{(k-1)} + 1 \end{cases}$$

where $X^{(0)} = 0$

Iteratively computing gives

$$X_1^{(1)} = 7,\ X_2^{(1)} = -5,\ X_3^{(1)} = 1 \Rightarrow \boxed{X^{(1)} = (7, -5, 1)^t}$$ and

$$X_1^{(2)} = -2(-5) + 2 + 7 = 10 + 9 = 19$$

$$X_2^{(2)} = 2(-5) - 3 - 5 = -10 - 8 = -18$$

$$X_3^{(2)} = 2 + 1 = 3$$

$$\boxed{X^{(2)} = (19, -18, 3)^t}$$

$$\boxed{\begin{aligned} X^{(1)} &= (7.0, -5.0, 1.0)^t \\ X^{(2)} &= (19.0, -18.0, 3.0)^t \end{aligned}}$$