Anar Rzayev MAS 365: Homework 4 20190788

2) In conjugate gradient method, we successively find

$$t_k = \frac{\langle r^{(k-1)}, r^{(k-1)} \rangle}{\langle v^{(k)}, Av^{(k)} \rangle} \qquad x^{(k)} = x^{(k-1)} + t_k v^{(k)}$$

$$r^{(k)} = r^{(k-1)} - t_k A v^{(k)}$$

$$s_k = \frac{\langle r^{(k)}, r^{(k)} \rangle}{\langle r^{(k-1)}, r^{(k-1)} \rangle} \qquad v^{(k+1)} = r^{(k)} + s_k v^{(k)}$$

starting with $r^{(0)} = b - Ax^{(0)}$ and $v^{(1)} = r^{(0)}$

a) $v^{(1)} = r^{(0)} = b - Ax^{(0)}$ and $r^{(1)} = r^{(0)} - t_1 A v^{(1)} =$

$$= r^{(0)} - \frac{\langle r^{(0)}, r^{(0)} \rangle}{\langle v^{(1)}, Av^{(1)} \rangle} \cdot Av^{(1)} = r^{(0)} - \frac{\langle v^{(1)}, v^{(1)} \rangle}{\langle v^{(1)}, Av^{(1)} \rangle} Av^{(1)}$$

This gives $\langle r^{(1)}, v^{(1)} \rangle = \langle v^{(1)}, r^{(1)} \rangle =$

$$= \left\langle v^{(1)}, r^{(0)} - \frac{\langle v^{(1)}, v^{(1)} \rangle}{\langle v^{(1)}, Av^{(1)} \rangle} Av^{(1)} \right\rangle = \langle v^{(1)}, r^{(0)} \rangle -$$

$$- \frac{\langle r^{(0)}, r^{(0)} \rangle}{\langle v^{(1)}, Av^{(1)} \rangle} \cdot \langle v^{(1)}, Av^{(1)} \rangle = \langle v^{(1)}, r^{(0)} \rangle -$$

$- \langle r^{(0)}, r^{(0)} \rangle = 0$ since we know that $v^{(1)} = r^{(0)}$ $\Rightarrow$ Hence, $\boxed{\langle r^{(1)}, v^{(1)} \rangle = 0}$

b) Let's assume that $\langle r^{(k)}, v^{(j)} \rangle = 0$ for each $k \le p$ and $j = 1, 2, \dots, k \Rightarrow \langle r^{(p+1)}, v^{(j)} \rangle = \langle r^{(p)} - t_{(p+1)} A v^{(p+1)}, v^{(j)} \rangle$

$$= \left\langle r^{(p)} - \frac{\langle r^{(p)}, r^{(p)} \rangle}{\langle v^{(p+1)}, A v^{(p+1)} \rangle} \cdot A v^{(p+1)}, v^{(j)} \right\rangle =$$

$$= \langle r^{(p)}, v^{(j)} \rangle - \frac{\langle r^{(p)}, r^{(p)} \rangle}{\langle v^{(p+1)}, A v^{(p+1)} \rangle} \cdot \langle A v^{(p+1)}, v^{(j)} \rangle$$

Since $\underline{\langle v^{(i)}, A v^{(j)} \rangle = 0}$ if $i \ne j$ for A-orthogonal set of vectors and $\underline{\langle r^{(p)}, v^{(j)} \rangle = 0}$ from definition for each $k \le p$ and $j = 1, 2, \dots, k \Rightarrow$ we find that

$$\langle r^{(p+1)}, v^{(j)} \rangle = 0 - \frac{\langle r^{(p)}, r^{(p)} \rangle}{\langle v^{(p+1)}, A v^{(p+1)} \rangle} \cdot 0 = 0 \Rightarrow$$

$$\boxed{\langle r^{(p+1)}, v^{(j)} \rangle = 0} \; \boxed{\star} \; \checkmark$$

c) Now, $\langle r^{(p+1)}, v^{(p+1)} \rangle = \langle r^{(p)} - t_{(p+1)} A v^{(p+1)}, v^{(p+1)} \rangle$

$$= \left\langle r^{(p)} - \frac{\langle r^{(p)}, r^{(p)} \rangle}{\langle v^{(p+1)}, A v^{(p+1)} \rangle} \cdot A v^{(p+1)}, v^{(p+1)} \right\rangle$$

$$\langle r^{(P+1)}, v^{(P+1)} \rangle = \langle r^{(P)}, v^{(P+1)} \rangle - \frac{\langle r^{(P)}, r^{(P)} \rangle}{\langle v^{(P+1)}, Av^{(P+1)} \rangle} \cdot \langle Av^{(P+1)}, v^{(P+1)} \rangle$$

$$= \langle r^{(P)}, v^{(P+1)} \rangle - \frac{\langle r^{(P)}, r^{(P)} \rangle}{\langle v^{(P+1)}, Av^{(P+1)} \rangle} \cdot \langle v^{(P+1)}, Av^{(P+1)} \rangle$$

$$= \langle r^{(P)}, v^{(P+1)} \rangle - \langle r^{(P)}, r^{(P)} \rangle = \langle r^{(P)}, v^{(P+1)} - r^{(P)} \rangle$$

Since $v^{(P+1)} = r^{(P)} + S_P v^{(P)} \Rightarrow v^{(P+1)} - r^{(P)} = S_P v^{(P)}$

and $\langle r^{(P+1)}, v^{(P+1)} \rangle = \langle r^{(P)}, S_P \cdot v^{(P)} \rangle$ or just

$\langle r^{(P)}, v^{(P+1)} \rangle = \langle r^{(P)}, r^{(P)} + S_P v^{(P)} \rangle = \langle r^{(P)}, r^{(P)} \rangle$

$+ S_P \cdot \langle r^{(P)}, v^{(P)} \rangle = \langle r^{(P)}, r^{(P)} \rangle$ since we know

$\underline{\langle r^{(k)}, v^{(j)} \rangle = 0}$ was true $\underline{\text{for each } k \leq P,\ j = 1, 2, \dots, k}$

In particular, $\langle r^{(P)}, v^{(P)} \rangle = 0 \Rightarrow \langle r^{(P)}, v^{(P+1)} \rangle =$

$= \langle r^{(P)}, r^{(P)} \rangle$ and $\langle r^{(P+1)}, v^{(P+1)} \rangle = \langle r^{(P)}, v^{(P+1)} \rangle$

$- \langle r^{(P)}, r^{(P)} \rangle = 0 \Rightarrow \boxed{\langle r^{(P+1)}, v^{(P+1)} \rangle = 0}$ ✓

Hence, from the Mathematical Induction, we get

$\boxed{\langle r^{(k)}, v^{(j)} \rangle = 0 \text{ for each } j = 1, 2, \dots, k}$ ✓ ∎

1) a) From Theorem 5 in Section 7.2, we know that if A is an $n \times n$ matrix $\Rightarrow \rho(A) \le \|A\|$ for any natural norm $\|\cdot\| \Rightarrow$ since $\rho(A) = \max|\lambda|$ where λ is an eigenvalue of $A \Rightarrow$ for $\forall \lambda$-eigenvalue of A,

$$|\lambda| \le \rho(A) \le \|A\|$$

$\boxed{\forall \lambda\text{-eigenvalue of } A \Rightarrow 0 \le |\lambda| \le \|A\| \text{ for any natural norm } \|\cdot\|}$ Since we know that if A is an $n \times n$ matrix, then $\det A = \prod_{i=1}^{n} \lambda_i$, where λ_i-eigenvalues of A

(proof is based on the consideration of $p(0)$ where $p(\lambda) = \det(A - \lambda I)$) We can imply $\boxed{A\text{-singular} \Leftrightarrow \lambda = 0 \text{ is an eigenvalue of } A}$

As $\|\cdot\|$ is any natural norm, then $\|\cdot\|$ is semi multiplicative; that is, $\boxed{\|A^T A\| \le \|A^T\| \cdot \|A\|}$

A-nonsingular $\Rightarrow A^T A$ is also nonsingular (consider $A^T A X = 0$; here, AX-an element in the range of A, is in the null space of A^T. However, null space of A^T and the range of A are orthogonal complements, so $AX = 0$ ✓) (since $A^T A$-square matrix $\Rightarrow A^T A$ is invertible)

A^TA is also symmetric ($B = A^TA \Rightarrow B^T = (A^TA)^T = A^T(A^T)^T = A^TA = B$) Thus, $\boxed{A^TA \rightarrow \text{symmetric, nonsingular}}$

Considering A^TA to be positive semidefinite, we know for any conforming v, it follows that $v^TA^TAv = |Av|^2 \geq 0 \Rightarrow$ thus, the eigenvalues of A^TA are all non-negative (we can actually prove this directly by looking at $A^TAv = \lambda v \Rightarrow 0 \leq |Av|^2 = \lambda |v|^2$)

If A^TA has an eigenvalue equal to zero, then its determinant will be zero (rule mentioned in previous page) or A^TA will be singular; however, this is impossible $\boxtimes$

$\boxed{\text{All eigenvalues of } A^TA \text{ are positive}}$ ☆

$\boxed{\lambda\text{-eigenvalue of } A^TA \Rightarrow 0 < \lambda}$ ☆ Since we proved in the previous page that $\forall \lambda$-eigenvalue of $A^TA \Rightarrow |\lambda| \leq \|A^TA\|$ for any natural norm $\|\cdot\|$

Then, $\lambda > 0 \Rightarrow \lambda \leq \|A^TA\| \leq \|A^T\| \cdot \|A\|$ becomes true

Hence, $\boxed{\lambda\text{-eigenvalue of } A^TA \Rightarrow 0 < \lambda \leq \|A^T\| \cdot \|A\|}$ ☆

☑ ■

(Note: There exist x such that $A^TAx = \lambda x$ and $\|x\| = 1$

$0 < \lambda = \|\lambda x\| = \|A^TAx\| \leq \|A^TA\| \cdot \|x\| = \|A^TA\| \leq \|A^T\| \cdot \|A\|$)

Scanned with CamScanner

c) For $n \times n$ matrix A, we know $\|A\|_2 = \sqrt{\rho(A^TA)}$
and for any natural norm $\|\cdot\| \Rightarrow \rho(A) \leq \|A\|$
for each $n \times n$ matrix A

Then, $\rho(A^TA) \leq \|A^TA\| \leq \|A^T\| \cdot \|A\|$ (as discussed earlier)

$\|A\|_2 = \sqrt{\rho(A^TA)} \leq \sqrt{\|A^T\| \cdot \|A\|}$ Since we mentioned $\rho(A) \leq \|A\|$ for any natural norm $\|\cdot\|$, then we could get $\rho(A^TA) \leq \|A^TA\|_\infty \leq \|A^T\|_\infty \cdot \|A\|_\infty$ since we know $\|AB\| \leq \|A\| \cdot \|B\|$ for any natural matrix norm

Thus, $\|A\|_2 = \sqrt{\rho(A^TA)} \leq \sqrt{\|A^T\|_\infty \cdot \|A\|_\infty}$

<u>Claim:</u> $\|A^T\|_\infty = \|A\|_1$

<u>Pf:</u> From definition, $\|A\|_1 = \max\limits_{1 \leq j \leq n} \sum\limits_{i=1}^{n} |a_{ij}|$ where

$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$ where $\|A\|_1 = \max\limits_{\|x\|_1 = 1} \|Ax\|_1$ means we just look at the sums of absolute values of elements in each column and take largest

Moreover, from Theorem 4 at 7.2 $\Rightarrow$ If $A = (a_{ij})$ is an $n \times n$ matrix, then $\|A\|_\infty = \max\limits_{1 \leq i \leq n} \sum\limits_{j=1}^{n} |a_{ij}|$

Simply put, $||A||_\infty$ tells us to look at the sums of absolute values of elements in each row and take the biggest value.

From these definitions, we can easily observe that

$A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ & \vdots & & \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix}$ and $||A^T||_\infty = \max_{1 \le q \le n} \sum_{p=1}^{n} |a_{pq}|$

However, this is indeed the exact definition of $||A||_1 \rightsquigarrow \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}|$ where $A = (a_{ij})$ is $n \times n$ matrix $\Rightarrow$ So, $\boxed{||A^T||_\infty = ||A||_1}$ is true for any $n \times n$ matrix A ✓

$$||A||_2 \le \sqrt{||A^T||_\infty \cdot ||A||_\infty} = \sqrt{||A||_1 \cdot ||A||_\infty}$$

Since A-nonsingular, A^{-1} exist and doing the exact same application for A^{-1} yields $(||A||_2 \ge 0, ||A^{-1}||_2 \ge 0)$

$$||A^{-1}||_2 \le \sqrt{||A^{-1}||_1 \cdot ||A^{-1}||_\infty}$$

Multiplying the last 2 inequalities (since LHS are all ≥ 0)

$$0 \le ||A||_2 \cdot ||A^{-1}||_2 = K_2(A) \le \sqrt{(||A||_1 \cdot ||A^{-1}||_1)(||A||_\infty \cdot ||A^{-1}||_\infty)}$$

$= \sqrt{K_1(A) \cdot K_\infty(A)}$ Hence, $\boxed{K_2(A) \le \sqrt{K_1(A) K_\infty(A)}}$ ✓ ■

b) Considering the eigenvalues λ_j' of A, we know that $\rho(A) = \max |\lambda_j'|$ from the definition. From the statem.

Take the set of eigenvalues of matrix A^TA, where λ_1-smallest, λ_n-largest (set of eigenvalues is $\{\lambda_i\}$)

From part a), we proved $\lambda_i > 0$ and since A is an $n\times n$ matrix, $||A||_2 = \sqrt{\rho(A^TA)}$ where $\rho(A^TA)$ is equal to the $\underbrace{\max |\lambda_i|}_{\lambda_i\text{-eigenvalue of } A^TA}$ $\Rightarrow$ since $\lambda_i > 0$, it is sufficient to mention (λ_1-smallest, λ_n-max)

$\underbrace{\max \lambda_i}_{\lambda_i\text{-eignvl. of } A^TA} = \lambda_n$ from the statement given above

So, $\rho(A^TA) = \lambda_n$ and $\boxed{||A||_2 = \sqrt{\lambda_n}}$ * Note that we can easily see $\boxed{A^TA \text{ and } AA^T \text{ have the same eigenvalues}}$ Considering that the given matrix A is nonsingular, it means A^{-1} exists and

$$||A^{-1}||_2 = \sqrt{\rho\left((A^{-1})^T \cdot A^{-1}\right)} = \sqrt{\rho\left((A^T)^{-1} \cdot A^{-1}\right)} =$$

$$= \sqrt{\rho\left((AA^T)^{-1}\right)}$$ where we used the fact that if square matrix A is invertible $\Rightarrow (A^{-1})^T = (A^T)^{-1}$ and the relation $(AB)^{-1} = B^{-1}A^{-1}$ for invertible matrices A, B

A - nonsingular $\Rightarrow A^T$ - invertible ✓ and $\|A^{-1}\|_2 = \sqrt{\rho([AA^T]^{-1})}$

$$\|A^{-1}\|_2 = \sqrt{\rho((AA^T)^{-1})}$$

From the general rule,

$\boxed{\lambda\text{ - eigenvalue of } B \Rightarrow \frac{1}{\lambda}\text{ - eigenvalue of } B^{-1}}$

$AA^T \to$ eigenvalues are exactly same with those of A^TA and $(AA^T)^{-1} \to$ eigenvalues of this matrix will be the collection $\left\{\frac{1}{\lambda_i}\right\}$ where $\frac{1}{\lambda_1}$ - largest of them
(all λ_i's are positive)
and $\frac{1}{\lambda_n}$ - smallest of them. Thus, $\rho((AA^T)^{-1}) = \frac{1}{\lambda_1}$

with $\boxed{\|A^{-1}\|_2 = \frac{1}{\sqrt{\lambda_1}}}$ ☆ Therefore, $K_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2$

$$= \sqrt{\lambda_n} \cdot \frac{1}{\sqrt{\lambda_1}} = \sqrt{\frac{\lambda_n}{\lambda_1}}$$

In conclusion, $\boxed{K_2(A) = \sqrt{\frac{\lambda_n}{\lambda_1}}}$ ☆ ☑ ▪ ⊕