Homework 4

Monday, March 1, 2021 2:07 PM

1)

1. The Schur decomposition theorem states that every square matrix $\mathbf{A} \in \mathbb{C}^{m\times m}$ has a Schur decomposition, $\mathbf{A} = \mathbf{QUQ}^*$, where Q is unitary and T is upper-triangular. Use this theorem to prove that, for an arbitrary norm $||\cdot||$,

$$\lim_{n\to\infty} ||\mathbf{A}^n|| = 0 \iff \rho(\mathbf{A}) < 1. \quad (4)$$

Since Q is unitary and unitary matrices preserve length and angle

$$\lim_{n\to\infty} ||A^n||_2 = \lim_{n\to\infty} ||U^n||_2$$

$(\rightarrow)$ $\lim_{n\to\infty} ||A^n||_2 = 0 \Rightarrow \lim_{n\to\infty} ||U^n||_2 = 0 \Rightarrow U^n = 0_{m\times m}$

$U_{ii} < 1$ for $i = 1, \dots, m$. And by Schur decomposition

$\rho(A) = \rho(U) = \sqrt{U_{kk}} < 1 \Rightarrow \rho(A) < 1.$

$(\leftarrow)$ assume $\rho(A) < 1$. By Schur decomposition

$A = QUQ^*$ and since A is similar to

U we know they share the same spectral radius

$\rho(A) = \rho(U) < 1$ thus we know $U_{ii} < 1$ for $i = 1, \dots, m$

Since the diagonal elements of U are less then $\underline{1}$

$\lim_{n\to\infty} U^n = 0_{m\times m} \Rightarrow \lim_{n\to\infty} ||U^n||_2 = \lim_{n\to\infty} ||A^n||_2 = 0.$

2.

2. **(a)** Let $\mathbf{A} \in \mathbb{C}^{m\times m}$ be tridiagonal and hermitian, with all its sub- and super-diagonal entries nonzero. Prove that the eigenvalues of $\mathbf{A}$ are distinct. (Hint: Show that for any $\lambda \in \mathbb{C}$, $\mathbf{A} - \lambda\mathbf{I}$ has rank at least $m - 1$.)

(b) On the other hand, let $\mathbf{A}$ be upper-Hessenberg, with all its sub-diagonal entries nonzero. Give an example that shows that the eigenvalues of $\mathbf{A}$ are not necessarily distinct.

a) Assume A is tridiagonal and hermitian if A is hermitian then A is diagonalizable thus for an eigenvalue of A alegbraic multiplicity is equal to geometric multiplicity

Next assume λ is an eigen value of A then

$$A-\lambda I = \begin{bmatrix} a_{11}-\lambda & a_{12} & 0 & \cdots & 0 \\ a_{21} & a_{22}-\lambda & a_{23} & 0 \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & a_{m-1\,m} \\ \vdots & & & & \\ 0 & 0 & & a_{m,m-1} & a_{mm}-\lambda \end{bmatrix}$$

if we remove the first row and last column of the above matrix then we have an uppertriangular matrix with all diagonal entries being non-zero. thus $\text{Rank}(A-\lambda I) \geq m-1$ and since λ is an eigen value $\text{Rank}(A-\lambda I) = m-1$

thus all eigen values are distinct.

Thus all eigen values are distinct.

b) $A = \begin{bmatrix} -1 & 2 \\ -2 & 3 \end{bmatrix}$ $\quad \det(A - \lambda I)$

$$\left| \begin{bmatrix} -1-\lambda & 2 \\ -2 & 3-\lambda \end{bmatrix} \right| = (-1-\lambda)(3-\lambda) + 4 = -3 - 3\lambda + \lambda + \lambda^2 + 4$$

$$= 1 - 2\lambda + \lambda^2 = (\lambda - 1)^2 \Rightarrow \lambda = 1 \text{ with}$$

multiplicity 2.

3) **3.** Let $\mathbf{A}$ be $m \times n$ and $\mathbf{B}$ be $n \times m$. Show that the matrices $\begin{bmatrix} \mathbf{AB} & 0 \\ B & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \\ \mathbf{B} & \mathbf{BA} \end{bmatrix}$ have the same eigenvalues.

we know that non zero e-values of AB is equal to nonzero e-values of BA

$$\det\left(\begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} - \lambda I \right) = 0$$

$$= \det\left(\begin{bmatrix} AB - \lambda I & 0 \\ B & -\lambda I \end{bmatrix} \right) = 0$$

$$= \det(AB - \lambda I)\det(-\lambda I) = 0$$

$$\det\left(\begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix} - \lambda I\right) = 0$$

$$= \det\left(\begin{bmatrix} -\lambda I & 0 \\ B & BA - \lambda I \end{bmatrix}\right) = 0$$

$$= \det(-\lambda I)\det(BA - \lambda I) = 0$$

From here we see that

$\det(-\lambda I)$ is only 0 eigenvalues and

since we know BA and AB have the same eigenvalues $\det(BA - \lambda I) = \det(AB - \lambda I)$

Thus the above matrices have the same eigenvalues.

4) **4.** Show that the Gerschgorin theorem also holds with the bounds r_i which are given by the partial column sums (instead of the partial row sums):

$$r_i = \sum_{i=1, i\neq j}^{m} |a_{i,j}|, \quad i = 1, \ldots, m. \tag{5}$$

Since A and A^T have the same eigen values the the proof from lecture notes can be followed and the indices swapped

Let $Ax = \lambda x$ let k be the subscript of a component of x s.t. $|x_k| = \max_i |x_i| = \|x\|$

then we see that the k-th component satisfies

$$\lambda x_k = \sum_{j=1}^{n} a_{jk} x_j$$

So that

$$(\lambda - a_{kk}) x_k = \sum_{\substack{j=1 \\ j\neq k}}^{n} a_{jk} x_j .$$

Therefore

$$|\lambda - a_{kk}||x_k| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} |a_{jk}||x_j| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} |a_{kj}| \cdot \|x\| .$$

5)

5. Use the Gerschgorin theorem to show that the matrix

$$\mathbf{A} = \begin{bmatrix} 1.0 & 0.3 & 0.1 & 0.4 \\ 0.0 & 2.0 & 0.0 & 0.1 \\ 0.0 & 0.4 & 3.0 & 0.0 \\ 0.1 & 0.0 & 0.0 & 4.0 \end{bmatrix} \tag{6}$$

has exactly one eigenvalue in each of the four circles

$$|z - k| \leq 0.1, \quad k = 1, 2, 3, 4. \tag{7}$$

Using Gerschgorin thm we can take the smaller of the column sum or row sum. giving us the following results.

$$r_1 = \sum_{\substack{j=1 \\ j\neq 1}}^{4} |A_{j1}| = 0 + 0 + .1 = .1$$

$$r_2 = \sum_{\substack{j=1 \\ j\neq 2}}^{4} |A_{2j}| = 0 + 0 + .1 = .1$$

$$r_3 = \sum_{\substack{j=1 \\ j\neq 3}}^{4} |A_{j3}| = .1 + 0 + 0 = .1$$

$$r_4 = \sum_{\substack{j=1 \\ j\neq 4}}^{4} |A_{4j}| = .1 + 0 + 0 = .1$$

$$\Rightarrow \quad |Z - k| \leq .1 \quad , k = 1, 2, 3, 4$$

6)

6. Let $\mathbf{A} \in \mathbb{R}^{m\times m}$ be strictly diagonally dominant matrix, namely,

$$|a_{ii}| > \sum_{j\neq i} |a_{ij}|, \quad 1 \leq i \leq m. \tag{8}$$

Show that $\mathbf{A}$ is nonsingular. Also, find a lower bound for $||\mathbf{A}^{-1}||_2$ using the Gerschgorin theorem.

By Gerschgorin theorem

$$r_i = \sum_{\substack{j=1 \\ j\neq i}}^{m} |a_{ij}| \quad \text{then there is}$$

a disk centered at a_{ii} with

radius r_i in which there is an eigenvalue

since $|a_{ii}| > r_i$ the disk can never overlap with zero which implies

zero is never an eigenvalue.

From HW 1 #6 we know

$\rho(A) \leq \|A\| \Rightarrow \|A^{-1}\| \geq \rho(A^{-1})$

7) **7.** Let $\mathbf{A} \in \mathbb{R}^{m\times m}$ be non-defective (so that there exist eigenvectors $\mathbf{v}_i$ corresponding to λ_i that form a basis) and let $\{\lambda_i\}_{i=1}^m$ be its eigenvalues. Assume $\lambda_1 = \cdots = \lambda_r$ for some $1 < r < m$ and

$$|\lambda_r| > |\lambda_{r+1}| \geq \cdots \geq |\lambda_m|. \tag{9}$$

Consider the Power Iteration method given by

$$\mathbf{x}^{(k+1)} = \mathbf{A}\mathbf{x}^{(k)}, \quad \mathbf{y}^{(k+1)} = \frac{\mathbf{x}^{(k+1)}}{||\mathbf{x}^{(k+1)}||}, \tag{10}$$

with any arbitrary norm $||\cdot||$. Show that, for almost all possible choices of an initial vector $\mathbf{x}^{(0)}$, $\mathbf{y}^{(k+1)}$ will converge to an eigenvector associated to λ_1. For which values of $\mathbf{x}_{(0)}$ does the method *not* converge to such an eigenvector?

Since A is non-defective we know that the eigenvectors $\{V_i\}$, $i=1,\ldots,m$ form an orthonormal basis where vector X can be written as

$$X = \sum_{i=1}^{m} \alpha_i V_i$$

This implies

$$Ax = \sum_{i=1}^{m} \alpha_i \lambda_i V_i \Rightarrow A^n X = \sum_{i=1}^{m} \lambda_i^n \alpha_i V_i$$

Since we are given $|\lambda_1| = \cdots = |\lambda_r| > |\lambda_{r+1}| \geq \cdots \geq |\lambda_m|$

8)

8. Let $\mathbf{A} \in \mathbb{R}^{m\times m}$ be a non-defective matrix with its eigenvalues $\{\lambda_i\}_{i=1}^m$ and its singular values $\{\sigma_i\}_{i=1}^m$, satisfying

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_m|, \quad (11)$$
$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m. \quad (12)$$

Let $\rho(\mathbf{A})$ be the spectral radius of $\mathbf{A}$ and $\mathrm{cond}(\mathbf{A}) = ||\mathbf{A}||_2||\mathbf{A}^{-1}||_2$ be the condition number of $\mathbf{A}$. Show that:

(a) $\sigma_i = |\lambda_i|$, $1 \leq i \leq m$, if $\mathbf{A}$ is normal, i.e., $\mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A}^T$.

(b) $||\mathbf{A}||_2 = |\lambda_1| = \rho(\mathbf{A})$.

By spectral theorem we know that

A can be diagonalize by the following

$A = UDU^*$ where U is unitary and

D is diagonal (both with complex entries)

$$D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{bmatrix}$$

$$A^T A = A^* A = (UDU^*)^* UDU^* = UD^* DU^*$$

$$= U \begin{bmatrix} |\lambda_1|^2 & & 0 \\ & \ddots & \\ 0 & & |\lambda_m|^2 \end{bmatrix} U^*$$

This implies the eigen values of $A^T A$ are $|\lambda_i|^2$

Thus $\sigma_i(A) = \sqrt{\lambda_i(A^T A)} = \sqrt{|\lambda_i|^2} = |\lambda_i|$

b) The spectral radius of a matrix is $\max\{|\lambda_1| \dots |\lambda_n|\}$

thus since $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_m| \Rightarrow \rho(A) = |\lambda_1|$

and $\|A\|_2 = \sigma_{max}(A)$ from part a) $|\lambda_1| = |\sigma_1|$ where

$|\sigma_1| \geq |\sigma_2| \geq \cdots \geq |\sigma_m|$ thus $|\lambda_1| = \|A\|_2 \Rightarrow$

$\|A\|_2 = |\lambda|_1 = \rho(A)$

9)

9. Let $\mathbf{A} \in \mathbb{R}^{m\times m}$ be symmetric and positive definite. Suppose that you have found the following algorithm that is known to produce a sequence of $\mathbf{A}_i$. The algorithm claims that $\mathbf{A}_i$ converges to a diagonal matrix $\Lambda = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_m)$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m > 0$ as $i \to \infty$. In each iteration, the algorithm uses the Cholesky decomposition to factorize $\mathbf{A}_i$ to an upper triangular matrix $\mathbf{U}_i$ whose diagonal elements are nonzero. The algorithm proceeds as follows:

Algorithm:

$\mathbf{A}_0 = \mathbf{A}$
for $i = 1, \cdots$
$\mathbf{U}_i^T\mathbf{U}_i = \mathbf{A}_{i-1}$ **[Cholesky Decomposition]**
$\mathbf{A}_i = \mathbf{U}_i\mathbf{U}_i^T$

$\mathbf{A}_0 = \mathbf{A}$
for $i = 1, \cdots$
$\mathbf{U}_i^T \mathbf{U}_i = \mathbf{A}_{i-1}$ **[Cholesky Decomposition]**
$\mathbf{A}_i = \mathbf{U}_i \mathbf{U}_i^T$
endfor

Prove that the iteration is well-defined by showing the following steps:

(a) Show that $\mathbf{A}_i$ is also symmetric and positive definite to justify the use of Cholesky.

(b) Show that $\mathbf{A}_i$ is similar to $\mathbf{A}$ (i.e., $\mathbf{A}_i = \mathbf{B}^{-1}\mathbf{A}\mathbf{B}$ for some non-singular matrix $\mathbf{B}$).

(c) Explain why you can use this iteration method as a valid eigenvalue revealing algorithm for $\mathbf{A}$.

a) $U_i^T U_i = A_{i-1}$

$A_i = U_i U_i^T$

$A_i^T = (U_i U_i^T)^T = U_i^{T^T} U_i^T = U_i U_i^T$

Thus A_i is symmetric next to show

Positive definiteness

for $v \neq 0$

$v^T A v > 0 \Rightarrow v^T U_1 U_1^T v = (U_1^T v)^T (U_1^T v)$

$\Rightarrow \|U_1^T v\|^2$ since U_1 is a full rank

matrix $\|U_1^T v\|^2 > 0$. $\blacksquare$

b) W.T.S. A is similar to A_1

$A = U_1^T U_1$

$A_1 = U_1 U_1^T$

$A \quad U_1 \; U_1$

$A_1 = U_1 U_1^T$

$U_1^T U_1 = P(U_1 U_1^T)P^{-1}$ let $P = U_1^T$

$U_1^T U_1 = U_1^T (U_1 U_1^T) U_1^{T^{-1}} = U_1^T U_1$ ∎

Next we show A_i is similar to A_{i-1} thus A_i is similar to A_0 or A

$A_{i-1} = U_i^T U_i$

$A_i = U_i U_i^T$

$U_i^T U_i = P(U_i U_i^T)P^{-1}$ let $P = U_i^T$ $P^{-1} = U_i^{T^{-1}}$

$\Rightarrow U_i^T U_i = U_i^T (U_i U_i^T) U_i^{T^{-1}} = U_i^T U_i$

thus we have shown A_i is similar to A_{i-1}

Apply the same logic over and over again we see A_i is similar to A.

c) Since A_i converges to a diagonal matrix

and A_i is similar to A from part b)

Thus A_i has the eigenvalues of A.