THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
MATH3310 2022-2023
Homework Assignment 4 Suggested Solution

1. Consider a $n \times n$ tridiagonal linear system $A\mathbf{x} = \mathbf{b}$, where:

$$A = \begin{pmatrix} \alpha & -\beta & & \\ -\beta & \alpha & -\beta & \dots \\ & \ddots & \ddots & \ddots \\ & & -\beta & \alpha \end{pmatrix}$$

where $\alpha \geq \beta$.

(a) Prove that the eigenvectors of A are given by

$$q_j = \begin{pmatrix} \sin(j\theta) \\ \sin(2j\theta) \\ \vdots \\ \sin(nj\theta) \end{pmatrix}$$

for $j = 1, 2, \dots, n$ and $\theta = \frac{\pi}{n+1}$.

(b) Suppose $\alpha = 2$ and $\beta = 1$. Prove that the Jacobi method to solve $A\mathbf{x} = \mathbf{b}$ converges by looking at the spectral radius of a suitable matrix. Please explain your answer with details.

(c) Suppose $\alpha = 2$ and $\beta = 1$. Using the Housholder-John theorem, prove that the Gauss-Seidel method to solve $A\mathbf{x} = \mathbf{b}$ converges. Please explain with details.

(d) Suppose $\alpha = 2$ and $\beta = 1$. Explain why the SOR method converges for $0 < \omega < 2$. What is the optimal parameter ω_{opt} in the SOR method to obtain the fastest convergence. Please explain your answer with details.

Solution:

(a) Consider $A\mathbf{q}_j$:

$$\begin{aligned}(A\mathbf{q}_j)_1 &= \alpha \sin(j\theta) - \beta \sin(2j\theta) \\ &= \alpha \sin(j\theta) - 2\beta \sin(j\theta)\cos(j\theta) \\ &= (\alpha - 2\beta\cos(j\theta))\sin(j\theta) \\ &= (\alpha - 2\beta cos(j\theta))(\mathbf{q}_j)_1\end{aligned}$$

For $k = 2, 3, \dots, n-1$,

$$(A\mathbf{q}_j)_k = -\beta \sin((k-1)j\theta) + \alpha \sin(kj\theta) - \beta \sin((k+1)j\theta)$$

Note :

$$\sin((k-1)j\theta) = \sin(kj\theta)\cos(j\theta) - \sin(j\theta)\cos(kj\theta)$$
$$\sin((k+1)j\theta) = \sin(kj\theta)\cos(j\theta) + \sin(j\theta)\cos(kj\theta)$$

Then we have

$$(A\mathbf{q}_j)_k = (\alpha - 2\beta\cos(j\theta))\sin(kj\theta) = (\alpha - 2\beta\cos(j\theta))(\mathbf{q}_j)_k$$

Finally,

$$(A\mathbf{q}_j)_n = -\beta \sin((n-1)j\theta) + \alpha \sin(nj\theta)$$

Note

$$\begin{aligned}
\sin((n-1)j\theta) &= \sin((n+1)j\theta - 2j\theta) \\
&= \sin\left((n+1)j\frac{\pi}{n+1} - 2j\theta\right) \\
&= \sin(j\pi)\cos(2j\theta) - \cos(j\pi)\sin(2j\theta) \\
&= -\cos(j\pi) \times 2\sin(j\theta)\cos(j\theta) \\
&= -2\cos(j\pi)\cos(j\theta)\sin(j\pi - nj\theta) \\
&= -2\cos(j\pi)\cos(j\theta)\big(\sin(j\pi)\cos nj\theta - \cos(j\pi)\sin(nj\theta)\big) \\
&= -2\cos^2(j\pi)\cos(j\theta)\sin(nj\theta) \\
&= -2\cos(j\theta)\sin(nj\theta)
\end{aligned}$$

So,

$$(A\mathbf{q}_j)_n = (-2\beta\cos(j\theta) + \alpha)\sin(nj\theta) = (\alpha - 2\beta\cos(j\theta))(\mathbf{q}_j)_n$$

Hence, $\mathbf{q}_j$ is an eigenvector with eigenvalue $\alpha - 2\beta\cos(j\theta)$.

(b) For $\alpha = 2$ and $\beta = 1$,

$$M_J = \begin{pmatrix} 0 & \frac{1}{2} & & \\ \frac{1}{2} & 0 & \frac{1}{2} & \cdots \\ & \ddots & \ddots & \ddots \\ & & \frac{1}{2} & 0 \end{pmatrix}$$

By (a), M_J has eigenvalues $-\cos(j\theta), j = 1, 2, \ldots, n$.
Since $0 < j\theta < \pi$, we have $-1 < -\cos(j\theta) < 1$ for all j. And so the spectral radius is less than 1 and Jacobi method converges.

(c) Clearly, A is symmetric real matrix and self-adjoint, and so as $N^* + N - A$.
Note

$$A = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \cdots \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 \end{pmatrix}$$

has eigenvalues $2 - 2\cos(j\theta)$ and $0 < 2 - 2\cos(j\theta)$ as $0 < j\theta < \pi$. Also, $N = D + L$ for Gauss-Seidel method,

$$N^* + N - A = D^* + L^* + D + L - (L + D + U) = D$$

, $D = diag(2, 2, \ldots, 2)$. So A and $N^* + N - A$ are positive definite. Then Gauss-Seidel method converges by Householder-John Theorem.

(d) Note Jacobi method converges and $-1 < -\cos(\theta) \leq -\cos(j\theta) \leq -\cos(n\theta) < 1$ and $\cos(\theta) = -\cos(\pi - \theta) = -\cos(n\theta) > 0$. So,

$$\rho(M_J) = \cos(\theta) = -\cos(n\theta)$$

A is consistently ordered as it is tridiagonal. Then by D. Young's Theorem, SOR method converges when $0 < \omega < 2$ and $\omega_{opt} = \frac{2}{1+\sqrt{1-\cos^2\theta}} = \frac{2}{1+\sin\theta}$.

2. Consider:

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

Find the QR factorization of A by Gram-Schmidt process. Compute the first iteration in QR method. Please show all your steps.

Solution:

$$\mathbf{q}_1 = \frac{\mathbf{a}_1}{||\mathbf{a}_1||} = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}$$

$$\begin{aligned}
\tilde{\mathbf{q}}_2 &= \mathbf{a}_2 - (\mathbf{q}_1^T \mathbf{a}_2)\mathbf{q}_1 \\
&= \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} -1/2 \\ -1/2 \\ -1 \end{pmatrix} \\
\mathbf{q}_2 &= \frac{\tilde{\mathbf{q}}_2}{||\tilde{\mathbf{q}}_2||} = \begin{pmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ -\sqrt{2/3} \end{pmatrix} \\
\tilde{\mathbf{q}}_3 &= \mathbf{a}_3 - (\mathbf{q}_1^T \mathbf{a}_3)\mathbf{q}_1 - (\mathbf{q}_2^T \mathbf{a}_3)\mathbf{q}_2 \\
&= \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}} \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix} - (-\frac{1}{\sqrt{6}}) \begin{pmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ -\sqrt{2/3} \end{pmatrix} \\
&= \begin{pmatrix} -2/3 \\ -2/3 \\ 2/3 \end{pmatrix} \\
\mathbf{q}_3 &= \frac{\tilde{\mathbf{q}}_3}{||\tilde{\mathbf{q}}_3||} = \begin{pmatrix} -1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}
\end{aligned}$$

So,

$$Q = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{6} & -1/\sqrt{3} \\ -1/\sqrt{2} & -1/\sqrt{6} & -1/\sqrt{3} \\ 0 & -\sqrt{2/3} & 1/\sqrt{3} \end{pmatrix} \text{ and } R = \begin{pmatrix} \sqrt{2} & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{3/2} & -1/\sqrt{6} \\ 0 & 0 & 2/\sqrt{3} \end{pmatrix}$$

3. Consider:

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

Suppose an initial vector is given as $\mathbf{x}^{(0)} = (1, -1, 1)^T$. Calculate the first iteration of power method. Find the eigenvalue and the normalised eigenvector associated to it.

Solution: $Ax^{(0)} = (2, -2, 2)^T$. Then normalised eigenvector is $(1, -1, 1)^T$. Thus the eigenvalue is 2.

4. Let $A \in M_{n\times n}(\mathbb{C})$ be a $n \times n$ complex-valued matrix. Suppose the characteristic polynomial of A is given by: $f_A(t) = (-1)^n(t-\lambda_1)(t-\lambda_2)...(t-\lambda_n)$, where $\lambda_1, ..., \lambda_n$ are eigenvalues of A. Assuming that

$$|\lambda_1| = |\lambda_2| = ... = |\lambda_k| > |\lambda_{k+1}| \geq ... \geq |\lambda_n|,$$

where $k < n$. Suppose $A = QJQ^{-1}$, where J is the Jordan canonical form of A and Q is an invertible matrix. Assuming that the diagonal entries of J are arranged in descending order in terms of their magnitudes. Denote the j-th column of Q by $\mathbf{q}_j$, where $\mathbf{q}_1$, $\mathbf{q}_2$, ..., $\mathbf{q}_k$ are eigenvectors of A associated to λ_1, λ_2,..., λ_k respectively.

Let $\mathbf{x}_0$ be the initial vector defined as $\mathbf{x}_0 = a_1\mathbf{q}_1 + a_2\mathbf{q}_2 + ... + a_n\mathbf{q}_n$, where $a_j \in \mathbb{C}$ for $1 \leq j \leq n$ and $a_i \neq 0$ for $i = 1, 2, ..., k$. Consider the iterative scheme:

$$\mathbf{x}_{j+1} = \frac{A\mathbf{x}_j}{||A\mathbf{x}_j||_\infty} \text{ for } j = 0, 1, 2, ...$$

(a) Suppose $\lambda_1 = \lambda_2 = ... = \lambda_k \in \mathbb{R}$. will $||A\mathbf{x}_j||_\infty$ always converge as $j \to \infty$. If yes, what will it converge to? If not, please give a counter-example and explain your answer with details. Please show the full details of your proof.

(b) In general, if $|\lambda_1| = |\lambda_2| = ... = |\lambda_k|$, will $||A\mathbf{x}_j||_\infty$ always converge $j \to \infty$? If yes, what will it converge to? If not, please give a counter-example and explain your answer with details. Please show the full details of your proof.

Solution: It's easy to find for all $m \in \mathbb{N}^+$

$$\mathbf{x}_m = \frac{A\mathbf{x}_{m-1}}{\|A\mathbf{x}_{m-1}\|_\infty} = \frac{A^2\mathbf{x}_{m-2}}{\|A\mathbf{x}_{m-1}\|_\infty\|A\mathbf{x}_{m-2}\|_\infty} = \cdots = \frac{A^m\mathbf{x}_0}{\prod_{i=0}^{m-1}\|A\mathbf{x}_i\|_\infty}.$$

On the other side, we have $\|\mathbf{x}_m\|_\infty = 1$, so $\prod_{i=0}^{m-1}\|A\mathbf{x}_i\|_\infty = \|A^m\mathbf{x}_0\|_\infty$ and then

$$\mathbf{x}_m = \frac{A^m\mathbf{x}_0}{\|A^m\mathbf{x}_0\|_\infty}.$$

From the definition of $\mathbf{x}_0$,

$$A^m\mathbf{x}_0 = \sum_{i=1}^{n} a_i\lambda_i^m\mathbf{q}_i.$$

(a) Yes. Given $\lambda_1 = \lambda_2 = \cdots = \lambda_k \in \mathbb{R}$ and $|\lambda_1| > |\lambda_{k+1}| \geq |\lambda_{k+2}| \geq \cdots \geq |\lambda_n| > 0$, we can split $A^m\mathbf{x}_0$ into 2 parts,

$$A^m\mathbf{x}_0 = \lambda_1^m\sum_{i=1}^{k} a_i\mathbf{q}_i + \sum_{i=k+1}^{n} a_i\lambda_i^m\mathbf{q}_i = \lambda_1^m\mathbf{y} + \mathbf{z}_m.$$

When m is big enough, it's clear that $|\lambda_1|^m\|\mathbf{y}\|_\infty > \|\mathbf{z}_m\|_\infty$, $\lim_{m\to\infty}\frac{\|\mathbf{z}_m\|_\infty}{|\lambda_1|^m\|\mathbf{y}\|_\infty} = 0$ and

$$|\lambda_1|^m\|\mathbf{y}\|_\infty - \|\mathbf{z}_m\|_\infty \leq \|A^m\mathbf{x}_0\|_\infty \leq |\lambda_1|^m\|\mathbf{y}\|_\infty + \|\mathbf{z}_m\|_\infty.$$

Therefore, for such big m, we have

$$\frac{|\lambda_1|^{m+1}\|\mathbf{y}\|_\infty - \|\mathbf{z}_{m+1}\|_\infty}{|\lambda_1|^m\|\mathbf{y}\|_\infty + \|\mathbf{z}_m\|_\infty} \leq \|A\mathbf{x}_m\|_\infty = \frac{\|A^{m+1}\mathbf{x}_0\|_\infty}{\|A^m\mathbf{x}_0\|_\infty} \leq \frac{|\lambda_1|^{m+1}\|\mathbf{y}\|_\infty + \|\mathbf{z}_{m+1}\|_\infty}{|\lambda_1|^m\|\mathbf{y}\|_\infty - \|\mathbf{z}_m\|_\infty}.$$

For the left one,

$$\lim_{m\to\infty}\frac{|\lambda_1|^{m+1}\|\mathbf{y}\|_\infty - \|\mathbf{z}_{m+1}\|_\infty}{|\lambda_1|^m\|\mathbf{y}\|_\infty + \|\mathbf{z}_m\|_\infty} = |\lambda_1| \cdot \frac{1 - |\lambda_1| \cdot \lim_{m\to\infty}\frac{\|\mathbf{z}_{m+1}\|_\infty}{|\lambda_1|^{m+1}\|\mathbf{y}\|_\infty}}{1 + \lim_{m\to\infty}\frac{\|\mathbf{z}_m\|_\infty}{|\lambda_1|^m\|\mathbf{y}\|_\infty}} = |\lambda_1|.$$

Similarly, $\lim_{m\to\infty}\frac{|\lambda_1|^{m+1}\|\mathbf{y}\|_\infty + \|\mathbf{z}_{m+1}\|_\infty}{|\lambda_1|^m\|\mathbf{y}\|_\infty - \|\mathbf{z}_m\|_\infty} = |\lambda_1|$, which means $\lim_{m\to\infty}\|A\mathbf{x}_m\|_\infty = |\lambda_1|$.

(b) No. Suppose

$$J = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, Q = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, A = QJQ^{-1} = \begin{pmatrix} \frac{3}{2} & \frac{1}{2} & -\frac{1}{2} \\ 2 & 0 & -2 \\ \frac{3}{2} & -\frac{3}{2} & -\frac{1}{2} \end{pmatrix},$$

Then for A, we have $\mathbf{q}_1 = (1,1,0)^T$, $\mathbf{q}_2 = (0,1,1)^T$, $\mathbf{q}_3 = (1,0,1)^T$, $\lambda_1 = 2$, $\lambda_2 = -2$ and $\lambda_3 = 1$.

Let $\mathbf{x}_0 = \mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3$,

$$A^m \mathbf{x}_0 = (2^m + 1, 2^m + (-2)^m, (-2)^m + 1)^T.$$

When m is odd, $\|A^m \mathbf{x}_0\|_\infty = 2^m + 1$ and when m is even, $\|A^m \mathbf{x}_0\|_\infty = 2^{m+1}$, hence

$$\|A\mathbf{x}_m\|_\infty = \frac{\|A^{m+1}\mathbf{x}_0\|_\infty}{\|A^m \mathbf{x}_0\|_\infty} = \begin{cases} 2 - \frac{2}{2^m+1}, m \text{ is odd} \\ \frac{1}{2} + \frac{1}{2^{m+1}}, m \text{ is even} \end{cases}$$

which means $\|A\mathbf{x}_m\|_\infty$ diverges.