Solutions

MATH 504: Numerical Methods I

Instructor: Emre Mengi

Fall Semester 2018
Final Examination

NAME ____________________

STUDENT ID ____________________

SIGNATURE ____________________

#1	20	
#2	20	
#3	25	
#4	20	
#5	15	
Σ	100	

- Put your name, student ID and signature in the spaces provided above.
- Duration for this exam is 165 minutes.

Problem 1. (20 points) Let $A \in \mathbb{C}^{n \times n}$ be a matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{n-1}, \lambda_n$ such that

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{n-1}| > |\lambda_n|,$$

and n linearly independent eigenvectors $v_1, v_2, \ldots, v_{n-1}, v_n$, where v_j is an eigenvector corresponding to λ_j for $j = 1, \ldots, n$.

Suppose also that an LU factorization of $A \in \mathbb{C}^{n \times n}$ obtained by employing partial pivoting strategy is also given, that is a permutation matrix $P \in \mathbb{R}^{n \times n}$, a unit lower triangular matrix $L \in \mathbb{C}^{n \times n}$, and an upper triangular matrix $U \in \mathbb{C}^{n \times n}$ satisfying

$$PA = LU$$

are given.

Given A as above and $q^{(0)} \in \mathbb{C}^n$ such that $q^{(0)} \notin \text{span}\{v_1, \ldots, v_{n-1}\}$, write down a pseudocode that generates a sequence $\{q^{(k)}\}$ in $\mathbb{C}^n$ satisfying

$$\text{span}\{q^{(k)}\} \to \text{span}\{v_n\} \quad \text{as} \quad k \to \infty.$$

Your pseudocode must perform as few flops as possible.

Noting that, as $q^{(0)} \notin span\{v_1, \ldots, v_{n-1}\}$,

$$q^{(0)} = \alpha_1 v_1 + \cdots + \alpha_{n-1} v_{n-1} + \alpha_n v_n$$

for some $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ s.t. $\alpha_n \neq 0$, as well as, $|\lambda_{n-1}| > |\lambda_n|$, the sequence $\{q^{(k)}\}$ by the inverse iteration, that is

$$q^{(k+1)} = A^{-1} q^{(k)} / \|A^{-1} q^{(k)}\|_2$$

satisfies $span\{q^{(k)}\} \to span\{v_n\}$ as $k \to \infty$.

An efficient implementation

$q^{(0)} \leftarrow q^{(0)} / \|q^{(0)}\|_2$

for $k = 1, 2, 3, \ldots$

$\quad \hat{q} \leftarrow P q^{(k-1)}$

$\quad$ Solve $Ly = \hat{q}$ by forward substitution (for y)

$\quad$ Solve $Ux = y$ by back substitution

$\quad q^{(k)} \leftarrow x / \|x\|_2$

end

Every iteration requires only $O(n^2)$ flops due to forward and back substitutions.

Problem 2. Let

$$A = \begin{bmatrix} -2 & -4 & 3 \\ 4 & -1 & -6 \\ 1 & 2 & 3 \end{bmatrix}.$$

(a) **(10 points)** Compute a unit lower triangular matrix $L \in \mathbb{R}^{3\times 3}$ (that is L must be a lower triangular matrix with 1s on the diagonal) and an upper triangular matrix $U \in \mathbb{R}^{3\times 3}$ such that

$$A = LU.$$

(b) **(10 points)** Compute a unit upper triangular matrix $U \in \mathbb{R}^{3\times 3}$ (that is U must be an upper triangular matrix with 1s on the diagonal) and a lower triangular matrix $L \in \mathbb{R}^{3\times 3}$ such that

$$A = UL.$$

(a)

$$\begin{bmatrix} -2 & -4 & 3 \\ 4 & -1 & -6 \\ 1 & 2 & 3 \end{bmatrix} \xmapsto[r_3 + \frac{1}{2} r_1]{r_2 + 2r_1} \begin{bmatrix} -2 & -4 & 3 \\ -2\ \cancel{0} & -9 & 0 \\ -\frac{1}{2}\ \cancel{0} & 0 & 9/2 \end{bmatrix}$$

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1/2 & 0 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} -2 & -4 & 3 \\ 0 & -9 & 0 \\ 0 & 0 & 9/2 \end{bmatrix}}_{U}$$

(b)

$$\begin{bmatrix} -2 & -4 & 3 \\ 4 & -1 & -6 \\ 1 & 2 & 3 \end{bmatrix} \xmapsto[r_2 + 2r_3]{r_1 - r_3} \begin{bmatrix} -3 & -6 & \cancel{0}\ 1 \\ 6 & 3 & \cancel{0}\ -2 \\ 1 & 2 & 3 \end{bmatrix}$$

$$\xmapsto[r_1 + 2r_2]{} \begin{bmatrix} 9 & \overset{-2}{\cancel{0}} & 1 \\ 6 & 3 & -2 \\ 1 & 2 & 3 \end{bmatrix}$$

$$A = \underbrace{\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 9 & 0 & 0 \\ 6 & 3 & 0 \\ 1 & 2 & 3 \end{bmatrix}}_{L}$$

Problem 3. (25 points) For every matrix $A \in \mathbb{C}^{n\times n}$, there exist unitary matrices $U, V \in \mathbb{C}^{n\times n}$ such that

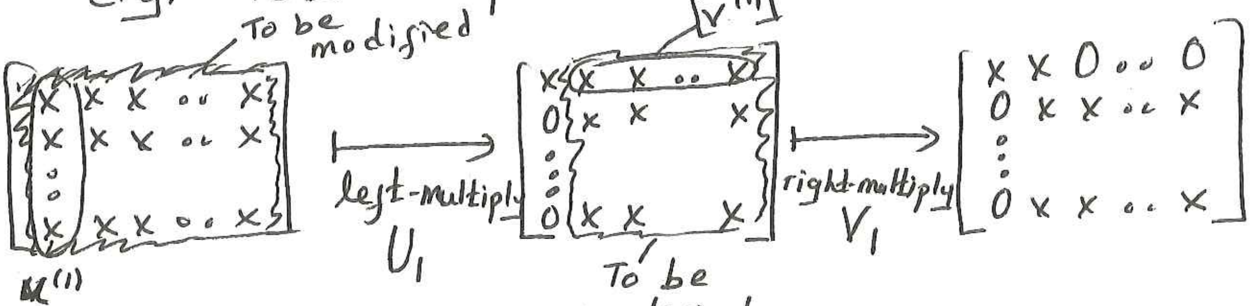
$$UAV = B \tag{1}$$

is bidiagonal (that is B is such that $b_{ij} = 0$ if $j \neq i$ and $j \neq i+1$).

Write down a pseudocode that, for a given $A \in \mathbb{C}^{n\times n}$, computes a bidiagonal matrix $B \in \mathbb{C}^{n\times n}$ such that (1) holds for some unitary matrices $U, V \in \mathbb{C}^{n\times n}$. Your pseudocode does not have to return U and V, it suffices if it only returns B.

Use the Householder reflectors from left & right.

e.g. 1st step



Next proceed 2nd column & 2nd row,
3rd column & 3rd row, so on

Pseudocode

for $j = 1, \ldots, n-1$

Left Multiply:
- $v \leftarrow A(j:n, j)$
- $q \leftarrow \{v - \|v\|_2 e_1\} / \|v - \|v\|_2 e_1\|_2$ — e_1: 1st col of $(j+1)\times(j+1)$ identity
- $A(j:n, j:n) \leftarrow A(j:n, j:n) - 2q\{q^* A(j:n, j:n)\}$

Right Multiply:
- if $(j < n-1)$
- $v \leftarrow A(j, j+1:n)^*$ — e_1: 1st col of $j\times j$ identity
- $q \leftarrow \{v - \|v\|_2 e_1\} / \|v - \|v\|_2 e_1\|_2$
- $A(j:n, j+1:n) \leftarrow A(j:n, j+1:n) - \{A(j:n, j+1:n) q\}\{2q^*\}$
- end

end

$B \leftarrow A$

return

Problem 4. The Hadamard product $\odot$ of two matrices $A, B \in \mathbb{C}^{n\times n}$ is defined by

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix}}_{A} \odot \underbrace{\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & & b_{2n} \\ \vdots & & & \\ b_{n1} & b_{n2} & & b_{nn} \end{bmatrix}}_{B} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \dots & a_{1n}b_{1n} \\ a_{21}b_{21} & a_{22}b_{22} & & a_{2n}b_{2n} \\ \vdots & & & \\ a_{n1}b_{n1} & a_{n2}b_{n2} & & a_{nn}b_{nn} \end{bmatrix}.$$

For a given $D \in \mathbb{C}^{n\times n}$, define

$$f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}, \quad f(X) = D \odot X. \tag{2}$$

(a) **(10 points)** Assuming D and X in (2) have exact representations in IEEE floating point arithmetic, show that $\widehat{f}(X)$, that is the computed $f(X)$ in IEEE floating point arithmetic, satisfies

$$\widehat{f}(X) = f(X + \delta X), \qquad \exists\, \delta X \in \mathbb{C}^{n\times n} \text{ s.t. } \|\delta X\|_F / \|X\|_F = O(\epsilon_{\text{mach}}).$$

(b) **(10 points)** Let κ denote the absolute condition number of $f(X)$ defined as in (2) for a given $D \in \mathbb{C}^{n\times n}$ when the Frobenius norm $\|\cdot\|_F$ is used on the input and output spaces of $f(X)$. Show that $\kappa = \|D\|_F$.

(Note: You can make use of the inequality

$$\|A \odot B\|_F \leq \|A\|_F \|B\|_F$$

that is satisfied for every $A, B \in \mathbb{C}^{n\times n}$.)

(a) $$\hat{f}(X) = \begin{bmatrix} fl(d_{11}x_{11}) & \dots & fl(d_{1n}x_{1n}) \\ \vdots & & \\ fl(d_{n1}x_{n1}) & \dots & fl(d_{nn}x_{nn}) \end{bmatrix}$$

$$= \begin{bmatrix} d_{11}x_{11}(1+\varepsilon_{11}) & \dots & d_{1n}x_{1n}(1+\varepsilon_{1n}) \\ \vdots & & \vdots \\ d_{n1}x_{n1}(1+\varepsilon_{n1}) & \dots & d_{nn}x_{nn}(1+\varepsilon_{nn}) \end{bmatrix} \quad \exists\, \varepsilon_{ij} \text{ s.t. } |\varepsilon_{ij}| \leq \varepsilon_{mach} \quad i,j = 1,\dots,n$$

$$= \begin{bmatrix} d_{11}(x_{11} + x_{11}\varepsilon_{11}) & \dots & d_{1n}(x_{1n} + x_{1n}\varepsilon_{1n}) \\ \vdots & & \vdots \\ d_{n1}(x_{n1} + x_{n1}\varepsilon_{n1}) & \dots & d_{nn}(x_{nn} + x_{nn}\varepsilon_{nn}) \end{bmatrix}$$

$$= D \odot (X + \delta X) \text{ where } \delta X = \begin{bmatrix} \varepsilon_{11}x_{11} & \dots & \varepsilon_{1n}x_{1n} \\ \vdots & & \\ \varepsilon_{n1}x_{n1} & \dots & \varepsilon_{nn}x_{nn} \end{bmatrix}$$

Hence,

$$\hat{f}(X) = f(X + \delta X) \quad \exists\, \delta X$$

s.t.

$$\|\delta X\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} \varepsilon_{ij}^2 x_{ij}^2} \leq \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} \epsilon_{mach}^2 x_{ij}^2} = \varepsilon_{mach}\|X\|_F$$

as desired.

(b) $$\kappa = \lim_{\delta \to 0^+} \sup_{0 < \|\delta X\|_F \leq \delta} \frac{\|D \odot (X + \delta X) - D \odot X\|_F}{\|\delta X\|_F}$$

$$= \lim_{\delta \to 0^+} \sup_{0 < \|\delta X\|_F \leq \delta} \frac{\|D \odot \delta X\|_F}{\|\delta X\|_F}$$

For every δX s.t. $0 < \|\delta X\|_F \leq \delta$

$$\frac{\|D \odot \delta X\|_F}{\|\delta X\|_F} \leq \frac{\|D\|_F \|\delta X\|_F}{\|\delta X\|_F} = \|D\|_F.$$

Furthermore, for $\delta X_* = \delta \cdot I$, we have

$$\frac{\|D \odot \delta X_*\|_F}{\|\delta X_*\|_F} = \frac{\delta \, \|D \odot I\|_F}{\delta} = \|D\|_F,$$

that is

$$\sup_{0 < \|\delta X\|_F \leq \delta} \frac{\|D \odot \delta X\|_F}{\|\delta X\|_F} = \|D\|_F.$$

Consequently,

$$\kappa = \lim_{\delta \to 0^+} \|D\|_F = \|D\|_F.$$

Problem 5. (15 points) A pseudocode is provided below for the basic QR algorithm without shifts to compute the eigenvalues of a matrix $A \in \mathbb{C}^{n \times n}$.

Algorithm 1 The QR Algorithm without Shifts

```
A_0 ← A
for k = 0, 1, ... do
    Compute a QR factorization A_k = Q_{k+1} R_{k+1}
    A_{k+1} ← R_{k+1} Q_{k+1}
end for
```

Prove that the iterates Q_k, R_k by this algorithm satisfy

$$A^k = (Q_1 Q_2 \dots Q_k)(R_k \dots R_2 R_1)$$

for all integer $k \geq 1$.

By induction on k.

Base case, $k=1$
By definition of the QR algorithm

$$A = A_0 = Q_1 R_1$$

as desired.

Inductive case,
Assume

$$A^{k-1} = (Q_1 \cdots Q_{k-1})(R_{k-1} \cdots R_1)$$

as the inductive hypothesis.

Now

$$(*) \quad A^k = A(Q_1 \cdots Q_{k-1})(R_{k-1} \cdots R_1).$$

By defn of the QR Algorithm

$$A_{j+1} = Q_{j+1}^T A_j Q_{j+1} \iff A_j Q_{j+1} = Q_{j+1} A_{j+1}$$

for all j, so from $(*)$

$$\begin{aligned} A^k &= (A_0 Q_1)(Q_2 \cdots Q_{k-1})(R_{k-1} \cdots R_1) \\ &= Q_1 (A_1 Q_2)(Q_3 \cdots Q_{k-1})(R_{k-1} \cdots R_1) \\ &= Q_1 Q_2 (A_2 Q_3) \cdots Q_{k-1} (R_{k-1} \cdots R_1) \\ &= Q_1 Q_2 \cdots Q_{k-1} A_{k-1} (R_{k-1} \cdots R_1) \\ &= (Q_1 Q_2 \cdots Q_{k-1} Q_k)(R_k R_{k-1} \cdots R_1) \end{aligned}$$

as $A_{k-1} = Q_k R_k$

completing the proof.