CSE 383C: Numerical Analysis: Linear Algebra – Final Exam – Solutions

10:30am – 12:30pm, December 13, 2024.

Question 1: (20p) Please provide answers only. Motivations will not be graded.

(a) (5p) Mark as true or false:

	TRUE	FALSE
If $\mathbf{A}$ is tridiagonal and positive definite, then its LU factors are bidiagonal.	X	
If $\mathbf{A}$ is tridiagonal and invertible, then $\mathbf{A}^{-1}$ is tridiagonal.		X
Every square matrix $\mathbf{A}$ admits a factorization $\mathbf{A} = \mathbf{QTQ}^*$ where $\mathbf{Q}$ is unitary and $\mathbf{T}$ is upper triangular.	X	

(b) (5p) Given a square $m \times m$ matrix $\mathbf{A}$, an $m \times 1$ vector $\mathbf{b}$, and a positive integer n, give the definition of the *Krylov space* $\mathcal{K}_n(\mathbf{A}, \mathbf{b})$:

$$\mathcal{K}_n(\mathbf{A}, \mathbf{b}) = \text{span}\{\mathbf{b},\ \mathbf{Ab},\ \mathbf{A}^2\mathbf{b},\ \dots,\ \mathbf{A}^{n-1}\mathbf{b}\}$$

(c) (5p) Given the function $f(x) = 2x^2 + x^3$, specify its *relative* condition number at $x = 2$:

$$\kappa_f(2) = |f'(x)|\frac{|x|}{|f(x)|} = \{\text{Set } x = 2\} = |4 \cdot 2 + 3 \cdot 4|\frac{|2|}{|2 \cdot 4 + 8|} = 20\frac{2}{16} = \frac{5}{2}$$

(d) (5p) The matrix $\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}$ has the QR factorization $\mathbf{A} = \mathbf{QR}$. Specify $\mathbf{Q}$ and $\mathbf{R}$

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{R} = \begin{bmatrix} 3 & 3 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

Question 2: (10p) Let $\mathbf{A}$ be symmetric and positive definite, and suppose that it has the Cholesky factorization $\mathbf{A} = \mathbf{R}^*\mathbf{R}$, where $\mathbf{R}$ is upper triangular. Set $\mathbf{B} = \mathbf{RR}^*$. Do $\mathbf{A}$ and $\mathbf{B}$ necessarily have the same eigenvalues? If you answer yes, then briefly motivate why. If you answer no, then provide a counter example.

Solution: Yes. To see why, observe that

$$\mathbf{A} = \mathbf{RR}^* = \{\text{Use that } \mathbf{R}^* = \mathbf{AR}^{-1}\} = \mathbf{RAR}^{-1}.$$

We see that $\mathbf{A}$ and $\mathbf{B}$ are SIMILAR. Consequently, they have the same eigenvalues.

Alternative solution: Suppose that λ is an eval of $\mathbf{A}$. Then for some non-zero vector $\mathbf{v}$:

$$\mathbf{Av} = \lambda\mathbf{v} \quad \Rightarrow \quad \mathbf{R}^*\mathbf{Rv} = \lambda\mathbf{v} \quad \Rightarrow \quad \mathbf{RR}^*\mathbf{Rv} = \lambda\mathbf{Rv} \quad \Rightarrow \quad \mathbf{BRv} = \lambda\mathbf{Rv},$$

so $\mathbf{Rv}$ is an evec of $\mathbf{B}$ with eval λ. Then analogously prove: λ is an eval of $\mathbf{B} \ \Rightarrow\ \lambda$ is eval of $\mathbf{A}$.

Question 3: (20p)

(a) (10p) For a positive integer n, define

$$s_n = \sum_{j=1}^{n} \frac{1}{j^2} \qquad \text{and} \qquad s = \sum_{j=1}^{\infty} \frac{1}{j^2}.$$

It is known that $s = \pi^2/6$, and we clearly have $s_n \to s$ as $n \to \infty$. Moreover, as we showed in class, $s - s_n \approx \int_{n+1}^{\infty} x^{-2}\,dx = (n+1)^{-1}$. For $n = 10^{10}$, we would expect $s - s_n \approx 10^{-10}$, yet the code

```
nmax = 1e10;
s    = 0;
for j = 1:nmax
  s = s + 1/(j*j);
end
fprintf(1,'Error          = %20.16f\n',pi*pi/6-s)
fprintf(1,'Expected error = %20.16f\n',1/(nmax+1))
```

produces the output:

```
Error          =   0.0000000090136514
Expected error =   0.0000000001000000
```

How would you explain the discrepancy? How would you compute s_n more accurately?

(b) (10p) Consider the function

$$f(x) = 1 - e^{-x^2}.$$

Suppose that for some number x such that $10^{-11} \le x \le 10^{-10}$ you want to evaluate $f(x)$ with several correct digits in *relative* precision. How would you proceed if you worked in an environment such as Matlab, and you could only use standard double precision arithmetic?

Solution:

(a) Once j becomes large, j^{-2} becomes smaller than $\epsilon_{\text{mach}} \times s$, so nothing gets added to the sum at each step. To fix this problem, sort the elements by order of magnitude, and start the summation with the smallest terms first.

(b) The easiest solution is to rewrite the function:

$$f(x) = 1 - e^{-x^2} = e^{-x^2/2}\left(e^{x^2/2} - e^{-x^2/2}\right) = 2e^{-x^2/2}\frac{e^{x^2/2} - e^{-x^2/2}}{2} = 2e^{-x^2/2}\sinh(x^2/2).$$

Alternatively, you could use a Taylor expansion

$$f(x) = 1 - e^{-x^2} = 1 - \left(1 - x^2 + O(x^4)\right) = x^2 + O(x^4).$$

It follows that the approximation

$$f(x) \approx x^2$$

will have sixteen correct digits in the interval specified.

Question 4: (20p) The matrix $\mathbf{A}$ has the LU factorization

$$\underbrace{\begin{bmatrix} 3 & 1 & -2 \\ -3 & -2 & a_{23} \\ 6 & a_{32} & -1 \end{bmatrix}}_{=\mathbf{A}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix}}_{=\mathbf{L}} \underbrace{\begin{bmatrix} 3 & 1 & u_{13} \\ 0 & -1 & 1 \\ 0 & 0 & 2 \end{bmatrix}}_{=\mathbf{U}}.$$

For this problem, no motivation is required. Just give the answers.

(a) (5p) Specify the missing entries a_{23}, a_{32}, and u_{13}.

(b) (5p) Specify the determinant of $\mathbf{A}$:

(c) (5p) Specify the solution $\mathbf{x}$ to the linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ where $\mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$:

(d) (5p) Set $\mathbf{V} = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ and $\mathbf{B} = \mathbf{AV}$. Suppose that $\mathbf{B}$ has the LU factorization $\mathbf{B} = \mathbf{L}'\mathbf{U}'$. Specify $\mathbf{L}'$ and $\mathbf{U}'$.

Solution:

(a)

$$a_{23} = 3 \qquad a_{32} = 1 \qquad u_{13} = -2$$

(b)

$$\det(\mathbf{A}) = \det(\mathbf{U})\det(\mathbf{L}) = (1 \cdot 1 \cdot 1)\,(3 \cdot (-1) \cdot 2) = -6$$

(c) We have

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \mathbf{U}^{-1}\mathbf{L}^{-1}\mathbf{b} = \{\text{Set } \mathbf{y} := \mathbf{L}^{-1}\mathbf{b}\} = \mathbf{U}^{-1}\mathbf{y}.$$

Now immediately observe that $\mathbf{y} = \mathbf{b}$. It remains to solve $\mathbf{U}\mathbf{y} = \mathbf{b}$:

$$\left[\begin{array}{ccc|c} 3 & 1 & -2 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|c} 3 & 1 & 0 & 1 \\ 0 & -1 & 0 & -1/2 \\ 0 & 0 & 2 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|c} 3 & 0 & 0 & 1/2 \\ 0 & -1 & 0 & -1/2 \\ 0 & 0 & 2 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1/6 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 1 & 1/2 \end{array}\right],$$

so

$$\mathbf{x} = \begin{bmatrix} 1/6 \\ 1/2 \\ 1/2 \end{bmatrix},$$

(d) We have

$$\mathbf{B} = \mathbf{AV} = \mathbf{LUV} =$$

Since $\mathbf{UV}$ is upper triangular, we find that $\mathbf{A} = \mathbf{L}\,(\mathbf{UV})$ is the LU factorization of $\mathbf{B}$. In other words:

$$\mathbf{L}' = \mathbf{L} \qquad \text{and} \qquad \mathbf{U}' = \mathbf{UV} = \begin{bmatrix} 3 & 4 & -7 \\ 0 & -1 & 2 \\ 0 & 0 & 4 \end{bmatrix}.$$

Question 5: (20p) Consider an $m \times n$ matrix $\mathbf{A}$, and an $m \times 1$ vector $\mathbf{b}$.

(a) (10p) Let $\mathbf{x}_\star$ denote the *least squares solution* to the linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$. Mark the following statements as true or false:

	TRUE	FALSE
If $\text{rank}(\mathbf{A}) = n$, then $\mathbf{x}_\star = \left(\mathbf{A}^*\mathbf{A}\right)^{-1}\mathbf{A}^*\mathbf{b}$.	X	
Regardless of the rank of $\mathbf{A}$, we have $\mathbf{A}^*\mathbf{A}\mathbf{x}_\star = \mathbf{A}^*\mathbf{b}$.	X	
$\lVert\mathbf{A}\mathbf{x}_\star - \mathbf{b}\rVert = \inf\{\lVert\mathbf{b} - \mathbf{y}\rVert \,:\, \mathbf{y} \in \text{col}(\mathbf{A})\}$.	X	
If $\text{rank}(\mathbf{A}) = m$, then the minimization problem $\inf\{\lVert\mathbf{A}\mathbf{y} - \mathbf{b}\rVert \,:\, \mathbf{y} \in \mathbb{R}^n\}$ has a unique solution.		X
It is always the case that $\mathbf{x}_\star \in \text{null}(\mathbf{A})^\perp$.	X	

(b) (10p) Consider the matrix $\mathbf{A}$ and the vector $\mathbf{b}$ defined by

$$\mathbf{A} = \underbrace{\begin{bmatrix} -2 & -2 \\ 2 & 4 \\ -2 & -2 \\ 2 & 4 \end{bmatrix}}_{=\mathbf{A}} = \underbrace{\begin{bmatrix} -1 & 1 \\ 1 & 1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix}}_{=\mathbf{W}} \underbrace{\begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}}_{=\mathbf{S}} \qquad \text{and} \qquad \mathbf{b} = \begin{bmatrix} -1 \\ 5 \\ -1 \\ 1 \end{bmatrix}$$

Specify the least squares solution $\mathbf{x}_\star$ to $\mathbf{A}\mathbf{x} = \mathbf{b}$. Please motivate your answer.

Solution to (b) via normal equations: The normal equations take the form

$$\mathbf{A}^*\mathbf{A}\mathbf{x} = \mathbf{A}^*\mathbf{b}$$

Now observe that $\mathbf{A}^*\mathbf{A} = \mathbf{S}^*\mathbf{W}^*\mathbf{W}\mathbf{S} = \mathbf{S}^*(4\mathbf{I})\mathbf{S} = 4\mathbf{S}^*\mathbf{S}$, so the normal equations simplify to

$$4\mathbf{S}^*\mathbf{S}\mathbf{x} = \mathbf{S}^*\mathbf{W}^*\mathbf{b}.$$

It follows that

$$\mathbf{x} = \frac{1}{4}\mathbf{S}^{-1}\mathbf{W}^*\mathbf{b} = \frac{1}{4}\frac{1}{2\cdot 1 - 0\cdot 3}\begin{bmatrix} 1 & -3 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 8 \\ 4 \end{bmatrix} = \frac{1}{4}\frac{1}{2}\begin{bmatrix} -4 \\ 8 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1 \end{bmatrix}.$$

Solution to (b) via QR: The factorization of $\mathbf{A}$ that is given is *almost* the QR factorization, since

$$\mathbf{W}^*\mathbf{W} = 4\mathbf{I}.$$

It follows that $\frac{1}{2}\mathbf{W}$ is an ON matrix, and so the QR factorization is

$$\mathbf{A} = \mathbf{W}\mathbf{S} = \left(\frac{1}{2}\mathbf{W}\right)(2\mathbf{S}) = \mathbf{Q}\mathbf{R}, \qquad \text{where} \qquad \mathbf{Q} = \frac{1}{2}\mathbf{W}, \qquad \text{and} \qquad \mathbf{R} = 2\mathbf{S}.$$

Using a standard formula for the pseudo-inverse, we get

$$\mathbf{A}^\dagger = \mathbf{R}^{-1}\mathbf{Q}^* = (2\mathbf{S})^{-1}((1/2)\mathbf{W})^* = \frac{1}{4}\mathbf{S}^{-1}\mathbf{W}^*.$$

Finally,

$$\mathbf{x} = \mathbf{R}^{-1}\mathbf{Q}^*\mathbf{b} = \frac{1}{8}\begin{bmatrix} 2 & -6 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} -1 \\ 5 \\ -1 \\ 1 \end{bmatrix} = \frac{1}{8}\begin{bmatrix} 2 & -6 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \frac{1}{8}\begin{bmatrix} -4 \\ 8 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1 \end{bmatrix}.$$

Solution to (b) via pseudoinverse: You can ignore QR and go directly for the formula

$$\mathbf{A}^\dagger = \left(\mathbf{A}^*\mathbf{A}\right)^{-1}\mathbf{A}^* = \left(\mathbf{S}^*\mathbf{W}^*\mathbf{W}\mathbf{S}\right)^{-1}\mathbf{S}^*\mathbf{W}^* = \left(4\mathbf{S}^*\mathbf{S}\right)^{-1}\mathbf{S}^*\mathbf{W}^* = \frac{1}{4}\mathbf{S}^{-1}\mathbf{S})^{-*}\mathbf{S}^*\mathbf{W}^* = \frac{1}{4}\mathbf{S}^{-1}\mathbf{W}^*.$$

Question 6: (10p) Let $\mathbf{A}$ be a symmetric and positive definite matrix of size $n \times n$, with eigenvalues $\{\lambda_j\}_{j=1}^n$ that are ordered by modulus, so that

$$0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots \leq \lambda_n.$$

You know that

$$\lambda_1 = 1 \qquad \text{and} \qquad \lambda_n = 10^5.$$

Given a positive integer p and a vector $\mathbf{b}$, we seek to solve the linear system

$$\mathbf{A}^p \mathbf{x} = \mathbf{b}.$$

We work in Matlab, using standard double precision floating point arithmetic. Our first attempt is:

```
T = A;
for i = 1:(p-1)
  T = T*A;
end
x = inv(T)*b;
```

We find that for $p = 2$, the code works fine. We next try $p = 5$, and find that we then run into problems, and get no accurate digits in the answer.

Explain what problem occurred, and then describe how you would proceed instead. For maximal credit, pay attention to computational efficiency in the case where p and n are large.

Explanation: We observe that $\kappa(\mathbf{A}) = \lambda_n/\lambda_1 = 10^5$. This means that if we compute $\mathbf{A}^p$ for $p > 3$, then the matrix will be singular to floating point precision. So plain inversion will not work. What is more, the modes that get lost are precisely the modes you need in order to solve the system, so if you form $\mathbf{A}^p$ for $p > 3$, then nothing you do after that will fix the problem.

An excellent solution: Compute the eigenvalue decomposition

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*.$$

We can then write down the solution explicitly as

$$\mathbf{x} = \mathbf{A}^{-p}\mathbf{b} = \mathbf{U}\mathbf{D}^{-p}\mathbf{U}^*\mathbf{b}.$$

This formula can be evaluated accurately (and rapidly) for any positive p.

Total cost of this approach is $O(n^3)$, *with no dependence on* p.

An accurate but more expensive solution: Iterate on $\mathbf{A}^{-1}$ instead of $\mathbf{A}$. Then the modes that get lost due to round-off errors are the ones associated with large eigenvalues of $\mathbf{A}$, which is what we want. In other words, set

$$\mathbf{x}_1 = \mathbf{A}^{-1}\mathbf{b},$$

and then evaluate iteratively

$$\mathbf{x}_j = \mathbf{A}^{-1}\mathbf{x}_{j-1}, \qquad \text{for } j = 2, 3, \dots, p.$$

Then $\mathbf{x}_p$ will be the solution $\mathbf{x}$.

Total cost is $O(p\,n^3)$ if you evaluate $\mathbf{x}_p$ via Gaussian elimination from scratch at every step.

The cost can be reduced to $O(n^3 + p\,n^2)$ if you precompute $\mathbf{A}^{-1}$. (Or do a Cholesky factorization.)