Question 1

[10 marks]

Consider $A, B, C \in \mathcal{R}^{n\times n}$ and $x, y, z \in \mathcal{R}^n$. Let A, B and C be non-singular (invertible) matrices, and let $I \in \mathcal{R}^{n\times n}$ represent the $n \times n$ identity matrix.

a. Show how to compute

$$z = B^{-1}(2A + I)(C^{-1} + A)x$$

without explicitly inverting either B or C. You may use vector addition, matrix-vector multiplication, and/or a routine for solving a system of linear equations. (You may assume the existence of such a routine—you do **not** need to give the details of Gaussian elimination and the $PA = LU$ factorization; i.e., you may assume such a factorization exists and can be computed.)

Soln:

3

$$\vec{z} = B^{-1}(2A+I)(C^{-1}+A)\vec{x}$$

$$B\vec{z} = B\cdot B^{-1}(2A+I)(C^{-1}+A)\vec{x}$$

$$B\vec{z} = (2A+I)(C^{-1}+A)\vec{x} \checkmark$$

$$(2A+I)^{-1}B\vec{z} = (2A+I)^{-1}(2A+I)(C^{-1}+A)\vec{x}$$

$$(2A+I)^{-1}B\vec{z} = (C^{-1}+A)\vec{x}$$

$$C(2A+I)^{-1}B\vec{z} = C(C^{-1}+A)\vec{x}$$

$$C(2A+I)^{-1}B\vec{z} = (I+CA)\vec{x}$$

unknown

Solve $C\vec{a} = (I+CA)\vec{x}$ for $\vec{a}$ how ?

then solve $(2A+I)^{-1}B\vec{z} = \vec{a}$ for $\vec{z}$.

Question 5

[15 marks]

Consider the functions $f(x) = 1 - 1/(2x)$ and $g(x) = 2x(1-x)$.

15

a. How many roots does f have? Are the roots of f fixed-points of g? Are there more fixed points of g than roots of f? **Justify your answers.**

$f(x) = 1 - 1/2x = 0$ $\quad 1 = \frac{1}{2x} \quad 2x = 1 \quad x = \frac{1}{2}$

f have only one root $x = \frac{1}{2}$

$g(\frac{1}{2}) = 2 \cdot \frac{1}{2}(1 - \frac{1}{2}) = \frac{1}{2}$

so the root of f is a fixed-point of g.

Suppose we have more fixed points of g $x_0 = \frac{1}{2}, x_1, \cdots$

$g(x) = x + 2x(1-x) - x = x + (x - 2x^2)$

Let $h(x) = x - 2x^2$

by FPT, $x_1 = h(x_0) = h(\frac{1}{2}) = \frac{1}{2} - 2 \cdot (\frac{1}{2})^2 = \frac{1}{2} - \frac{1}{2} = 0$

so there are more fixed points of g than roots of f

b. Using an appropriate theorem proven in lecture, determine the region of local convergence of the fixed-point iteration $x_{k+1} = g(x_k)$, $k = 0, 1, \ldots$, with $g(x)$ as defined above. In other words, find the *largest* interval on the x-axis for which the iteration is *guaranteed* to converge.

$x_{k+1} = g(x_k)$

$g(x) = 2x(1-x) = 2x - 2x^2$

by observation if $x \in [0, \frac{1}{2}]$ $\quad g(x) \in [0, \frac{1}{2}]$ $\quad$ ①

$g'(x) = 2 - 4x$

$|2 - 4x| \le L < 1$

$-1 < 2 - 4x < 1 \quad \Rightarrow \quad \begin{cases} 4x > 1 & x > \frac{1}{4} \\ 4x < 3 & x < \frac{3}{4} \end{cases} \quad x \in (\frac{1}{4}, \frac{3}{4})$ $\quad$ ②

Combining ① and ②, by FPT

we know that $g(x)$ has a unique fixed-point on $[\frac{1}{4}, \frac{1}{2}]$

and the iteration is guaranteed to converge.

[Continue your answer on the next page if necessary ...]

Question 4

[10 marks]

10

Let $\hat{x}$ be a computed solution to $Ax = b$, $A \in \mathcal{R}^{n \times n}$. The following bound for the relative error in $\hat{x}$ was derived in class:

$$\frac{\|x - \hat{x}\|}{\|x\|} \le \text{cond}(A) \frac{\|r\|}{\|b\|},$$

where $r = b - A\hat{x}$. Starting with the equations $Ax = b$ and $A\hat{x} = b - r$, derive a *lower* bound for $\|x - \hat{x}\|/\|x\|$. What do these bounds tell us about the reliability of $\hat{x}$?

$$A\hat{x} = b - r$$
$$A\hat{x} = Ax - r$$
$$r = Ax - A\hat{x}$$
$$r = A(x - \hat{x})$$

$$Ax = b$$
$$A^{-1}Ax = A^{-1}b$$
$$x = A^{-1}b$$

② $\|x\| = \|A^{-1}b\| \le \|A^{-1}\| \, \|b\|$

① $\|r\| = \|A(x - \hat{x})\| \le \|A\| \, \|x - \hat{x}\|$

Combining ① and ②

$$\|r\| \, \|x\| \le \|A\| \, \|x - \hat{x}\| \, \|A^{-1}\| \, \|b\|$$

$$\frac{\|r\|}{\|b\|} \le \|A\| \, \|A^{-1}\| \frac{\|x - \hat{x}\|}{\|x\|}$$

$$\frac{1}{\|A\| \, \|A^{-1}\|} \frac{\|r\|}{\|b\|} \le \frac{\|x - \hat{x}\|}{\|x\|}$$

(recall $\text{cond}(A) = \|A\| \, \|A^{-1}\|$)

so the lower bound for $\frac{\|x - \hat{x}\|}{\|x\|}$ is $\frac{1}{\text{cond}(A)} \cdot \frac{\|r\|}{\|b\|}$

these bounds tell us that if relative error is small, it does not mean that relative residual is small.

And they depend on cond(A).

CONTINUED ...

5

b. Use the factorization computed in (a) to solve the system.

$PA = LU$

$A\vec{x} = \vec{b}$

$PA\vec{x} = P\vec{b}$

$LU\vec{x} = P\vec{b}$

Let $U\vec{x} = \vec{d}$ $\quad \vec{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$

Solve $L\vec{d} = P\vec{b}$ first

$$\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{4} & 1 & 0 \\ \frac{3}{4} & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 40 \\ 24 \\ 26 \end{bmatrix} = \begin{bmatrix} 24 \\ 26 \\ 40 \end{bmatrix}$$

$d_1 = 24$

$\frac{1}{4}d_1 + d_2 = 26 \quad d_2 = 20$

$\frac{3}{4}d_1 + \frac{1}{2}d_2 + d_3 = 40 \quad d_3 = 12$

Solve $U\vec{x} = \vec{d}$

$$\begin{bmatrix} 4 & 4 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 24 \\ 20 \\ 12 \end{bmatrix}$$

$4x_1 + 4x_2 + 4x_3 = 24 \quad x_1 = 1$

$4x_2 + 4x_3 = 20 \quad x_2 = 2$

$4x_3 = 12 \quad x_3 = 3$

$\therefore \vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ ✓

c. Why is Gaussian Elimination usually implemented as in this question (i.e., $PA = LU$ is computed separately, and then the factorization is used to solve $Ax = b$)?

2

Because if we compute $PA = LU$ separately, we can store the value of P, L, U for further use, such as use them to solve $A\vec{y} = \vec{c}$, $A\vec{z} = \vec{d}$, ...

CONTINUED ...

Question 3

[15 marks]

Consider the linear system $Ax = b$ where

$$A = \begin{bmatrix} 3 & 5 & 9 \\ 4 & 4 & 4 \\ 1 & 5 & 5 \end{bmatrix}, \quad b = \begin{bmatrix} 40 \\ 24 \\ 26 \end{bmatrix}.$$

8

a. Compute the $PA = LU$ factorization of A. Use exact arithmetic. Show all intermediate calculations, including Gauss transforms and permutation matrices.

$$A = \begin{bmatrix} 3 & 5 & 9 \\ 4 & 4 & 4 \\ 1 & 5 & 5 \end{bmatrix} \quad P_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad P_1A = \begin{bmatrix} 4 & 4 & 4 \\ 3 & 5 & 9 \\ 1 & 5 & 5 \end{bmatrix} \quad L_1 = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{3}{4} & 1 & 0 \\ -\frac{1}{4} & 0 & 1 \end{bmatrix} \quad L_1P_1A = \begin{bmatrix} 4 & 4 & 4 \\ 0 & 2 & 6 \\ 0 & 4 & 4 \end{bmatrix}$$

$$P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad P_2L_1P_1A = \begin{bmatrix} 4 & 4 & 4 \\ 0 & 4 & 4 \\ 0 & 2 & 6 \end{bmatrix} \quad L_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{bmatrix} \quad L_2P_2L_1P_1A = \begin{bmatrix} 4 & 4 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix} = U$$

$$P = P_2P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

$$L_2P_2L_1P_1A = U$$
$$L_2P_2L_1P_2P_2P_1A = U$$
$$L_2\tilde{L}_1P_2P_1A = U$$
$$P_2P_1A = \tilde{L}_1^{-1}L_2^{-1}U$$
$$PA = LU$$

$$\left(P_2L_1 = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{4} & 0 & 1 \\ -\frac{3}{4} & 1 & 0 \end{bmatrix} \right)$$

$$\Downarrow$$

$$L_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{bmatrix} \quad \tilde{L}_1 = P_2L_1P_2 = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{4} & 1 & 0 \\ -\frac{3}{4} & 0 & 1 \end{bmatrix} \quad \tilde{L}_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{4} & 1 & 0 \\ \frac{3}{4} & 0 & 1 \end{bmatrix}$$

$$L = \tilde{L}_1^{-1}L_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{4} & 1 & 0 \\ \frac{3}{4} & \frac{1}{2} & 1 \end{bmatrix}$$

CONTINUED ...

Question 2

[10 marks]

When each of the following expressions is evaluated using floating-point arithmetic, poor results are obtained for a certain range of values of x. In each instance, identify this range and provide an alternate expression that can be used for such values of x.

a. $\sqrt{1+x} - \sqrt{1-x}$

this is not what question asked for

Assume $\hat{x} = x(1+\delta_1)$ $\quad \hat{1} = 1(1+\delta_2)$

$$fl(1+x) = (1(1+\delta_2) + x(1+\delta_1))(1+\delta_3)$$
$$= (1+\delta_2+x+\delta_1 x)(1+\delta_3)$$
$$= 1+x+\delta_2+\delta_1 x+\delta_3+\delta_2\delta_3+x\delta_3+\delta_1\delta_3 x$$

($\delta^2 \ll \delta$, insignificant) $= (1+\delta_2+\delta_3)+(1+\delta_1+\delta_3)x$

$$= (1+x)(1+2\delta_4)$$

$$fl(1-x) = (1(1+\delta_2) - x(1+\delta_1))(1+\delta_4)$$

similarly $= (1-x)(1+2\delta)$

$$fl(\sqrt{1+x}) = \sqrt{(1+x)(1+2\delta_4)}\,(1+\delta_5)$$
$$= \sqrt{(1+x)(1+2\delta_4)(1+\delta_5)^2}$$

($\delta^3 \ll \delta$, also insignificant) $= \sqrt{(1+x)(1+2\delta_4+2\delta_5+4\delta_4\delta_5+\delta_5^2+2\delta_4\delta_5^2)}$

$$= \sqrt{(1+x)(1+4\delta_{5'})}$$

similarly $fl(\sqrt{1-x})$
$$= \sqrt{(1-x)(1+4\delta_{5'})}$$

$$fl(\sqrt{1+x} - \sqrt{1-x}) = \left(\sqrt{(1+x)(1+4\delta_{5'})} - \sqrt{(1-x)(1+4\delta_{5'})}\right)(1+\delta_6)$$
$$= \left(\sqrt{1+x}\,(\sqrt{1+4\delta_5}) - \sqrt{1-x}\,(\sqrt{1+4\delta_5})\right)(1+\delta_6)$$
$$= (1+\delta_6)\sqrt{1+4\delta_5}\,(\sqrt{1+x}) - (1+\delta_6)\sqrt{1+4\delta_5}\,(\sqrt{1-x})$$
$$= (\sqrt{1+x} - \sqrt{1-x})\sqrt{1+6\delta}$$

$\delta < \sqrt{6}$ eps

b. $e^x - 1$

Assume $\hat{e} = e(1+\delta_e)$

$$fl(e\cdot e) = (e(1+\delta_e)\cdot e(1+\delta_e))(1+\delta_1)$$
$$= e\cdot e\,(1+2\delta_e+\delta_e^2)(1+\delta_1)$$

($\delta^2, \delta^3 \ll \delta$, insignificant) $= e\cdot e\,(1+2\delta_e+\delta_e^2+\delta_1+2\delta_e\delta_1+\delta_e^2\delta_1)$

$$= e\cdot e\,(1+3\delta_2)$$
$$fl(e^3) = (e^2(1+3\delta_2)\cdot e(1+\delta_e))(1+\delta_1)$$
$$= e^3(1+3\delta_2+2\delta_1\delta_e+6\delta_2\delta_e\delta_1+\delta_e\delta_1+3\delta_2\delta_e\delta_1)$$
$$= e^3(1+5\delta_3)$$
$$\vdots$$
$$fl(e^x) = e^x(1+(2x-1)\delta)$$
$$fl(e^x-1) = (e^x(1+(2x-1)\delta) - 1(1+\delta_0))(1+\delta_-)$$
$$= e^x(1+(2x-1)\delta+\delta_-+(2x-1)\delta\delta_-) - (1+\delta_0+\delta_-+\delta_0\delta_-)$$

CONTINUED ...

$$= e^x(1+2x\delta) - (1+2\delta')$$

$\delta < (2x)$ eps $\quad \delta' < 2$ eps

b. Show how to compute $y = A^{-6}x$ *without* explicitly inverting A. (Note: $A^{-6} = (A^{-1})^6$.) The same assumptions as in **(a)** apply.

5

Soln:

$$\vec{y} = A^{-6}\vec{x}$$
$$A\vec{y} = A \cdot A^{-6}\vec{x}$$
$$A\vec{y} = A^{-5}\vec{x}$$
$$\vdots$$
$$A^6\vec{y} = \vec{x}$$

Assume $PA = LU$ exists.

$A^6\vec{y} = \vec{x}$ ~~$P \cdot P = I$, so $P^6 = (P^2)^3 = (I)^3 = I$~~

$PA \cdot A^5\vec{y} = P\vec{x}$

$LU \cdot A^5\vec{y} = P\vec{x}$

let $U \cdot A^5\vec{y} = \vec{x_1}$ solve $L\vec{x_1} = P\vec{x}$ for $\vec{x_1}$

Let $A^5\vec{y} = \vec{y_1}$ solve $U \cdot \vec{y_1} = \vec{x_1}$ for $\vec{y_1}$

$PA \cdot A^4\vec{y} = P\vec{y_1}$

$LU \cdot A^4\vec{y} = P\vec{y_1}$

let $U \cdot A^4\vec{y} = \vec{x_2}$ solve $L\vec{x_2} = P\vec{y_1}$ for $\vec{x_2}$

Let $A^4\vec{y} = \vec{y_2}$ solve $U \cdot \vec{y_2} = \vec{x_2}$ for $\vec{y_2}$

$\vdots$

$A\vec{y} = \vec{y_5}$

$PA\vec{y} = P\vec{y_5}$

$LU\vec{y} = P\vec{y_5}$

Let $U\vec{y} = \vec{x_6}$ solve $L\vec{x_6} = P\vec{y_5}$ for $\vec{x_6}$

$\vec{y}$ solve $U\vec{y} = \vec{x_6}$ for $\vec{y}$.

CONTINUED ...