MATHEMATICS 2T

18 FEBRUARY 2011 MIDTERM TEST

DR. N. KEVLAHAN

DURATION OF EXAM: 50 MINUTES

PLEASE SHOW ALL WORK AND CALCULATIONS

Answers should be as brief as possible

THIS TEST PAPER HAS 1 PAGE AND 3 QUESTIONS.

---

1. Consider a normalized floating point number system with base 10, precision 3, and exponent range $[-2, 2]$.

[2] (a) Are the floating point numbers evenly distributed along the real line?

[4] (b) What are the largest and smallest positive numbers representable in this number system?

[2] (c) What is the maximum possible relative error in representing a non-zero floating point number in this system (with rounding by chopping)?

[4] (d) If $x = 3.33 \times 10^1$ and $y = 3.33 \times 10^{-2}$, what is the result of $x - y$ in this number system?

2. Consider numerical solutions to the linear system $Ax = b$, where A is non-singular.

[4] (a) What is partial pivoting and why is it necessary for a *stable* Gaussian elimination algorithm?

[2] (b) Does every matrix have an LU factorization?

[4] (c) Is the residual *always* a good estimate of the actual error? Why or why not?

[4] (d) Consider

$$A = \begin{bmatrix} 2 & -3 & 3 \\ 2 & 1 & -2 \\ 6 & -2 & 1 \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} -1/2 & -1/2 & 1/2 \\ -7/3 & -8/3 & 5/3 \\ -5/3 & -7/3 & 4/3 \end{bmatrix}.$$

How many digits of accuracy would the computed solution of $Ax = b$ (for arbitrary b) lose compared with the input data accuracy? [Hint: you can use any convenient norm.]

[2] 3. (a) When are iterative methods preferable to direct methods for solving linear systems?

[2] (b) Why does Gauss–Seidel iteration usually converge faster than Jacobi iteration?

[4] (c) Will Jacobi iterations converge for the following matrix? You must justify your answer.

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

[6] (d) Suppose that the spectral radius $\rho(G) = 0.8912$, where G is the Jacobi iteration matrix for a particular matrix A. Estimate how large A must be for the Jacobi method to be more efficient than LU decomposition (assume the maximum tolerable error is 10^{-6}). You must justify your answer.

[Total: 40]

THE END

Math 2T test

1. $\beta = 10, \ p = 3, \ L = -2, \ U = 2$

(a) No.

(b) Smallest $= \beta^{L} = 10^{-2} = 0.01$
largest $= \beta^{U+1}(1-\beta^{-p}) = 10^{3}\underbrace{(1-10^{-3})}_{0.999} = 999$

(c) $\varepsilon = \beta^{1-p} = 10^{-2} = 0.01$

(d)
$$\begin{array}{r} 3.33 \times 10^{1} \\ -\ 0.00 \times 10^{1} \\ \hline 3.33 \times 10^{1} \end{array}$$
No change!

2.
(a) In partial pivoting one permutates rows to ~~bring~~ ~~largest~~ use element ~~in~~ of largest magnitude on or below diagonal as pivot. This ensures rounding error is minimized.

(b) No. (Although the rows of any matrix may be permuted to give a matrix which does have an LU factorization $PA = LU$)

(c) The residual $\|A\hat{x} - b\|$ is only a good estimate of the actual error $\|x - \hat{x}\|$ ~~is~~ if cond $\|A\|$ is not too large.

2. (d) use ~~L_∞~~ or L_∞ norm :

$\|A\|_\infty = 9 \quad \|A^{-1}\|_\infty = 20/3$

$\therefore \operatorname{cond}_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty = 180/3 = 60$

$\rightarrow$ lose $\log_{10} \operatorname{cond}_\infty(A)$ digits $= \log_{10} 60 = 1.78$

$\sim$ 2 digits.

3. (a) Iterative methods are preferable for large (especially sparse systems), since complexity is $\sim K n^2$, where K is number of iterations for required accuracy, compared with $\sim \frac{2}{3} n^3$ for LU.

$\rightarrow$ iterative methods win if $K \ll n$.

(b) G-S usually converges faster than Jacobi since updated entries of vector x are used immediately, rather than waiting until next iteration, i.e.

new value.

$$x_i^{(k+1)} = \frac{b_i - \sum_{j<i} a_{ij} x_j^{(k+1)} - \sum_{j>i} a_{ij} x^{(k)}}{a_{ii}}$$

(c) Not strictly diagonally dominant → need to check spectral radius of G.

$$\rightarrow G = -D^{-1}(L+U) = \begin{pmatrix} -1/2 & 0 \\ 0 & -1/2 \end{pmatrix}\begin{pmatrix} 0 & 3 \\ 1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -3/2 \\ -1/2 & 0 \end{pmatrix}$$ need e-vals. → $\lambda^2 = 3/4$

→ $\lambda = \pm \frac{\sqrt{3}}{2}$

$$\rightarrow \boxed{\rho(G) = \frac{\sqrt{3}}{2} \approx 0.866 < 1 \rightarrow \text{Converges}}$$

(d) ~~as~~ Gain $-\log_{10}\rho(G)$ ~~digits~~ $= 0.05$ digits per iteration (or 20 iterations for one digit)

→ $K \sim 120$ iterations for tolerance 10^{-6}.

Now need $K < 2/3\, N$ or $N > 3/2\, K$ for Jacobi to be more efficient → $N > 3/2 \cdot 120$

or $\boxed{N > 180}$ ∴ Jacobi will be more efficient than LU for matrices larger than 180×180.