Clarissa Labrum A02274169

1. Define the order of accuracy of a finite difference approximation of the derivative.

first order accurate - approximation error is roughly propotional to h,

$\log|E(h)| \approx \log|c| + p \log h$

p is the order of accuracy, the higher the p the more accurate the approximation

2. Show that the one sided finite difference method, f'(x) ≈ (f(x + h) − f(x))/h is first order accurate using Taylor series expansion.

$f(x+h) = f(x) + hf'(x) + \frac{1}{2}h^2 f''(x) + \frac{1}{6}h^3 f'''(x) + O(h^4)$

$f(x) = f(x)$

$D_+ f(x) = \frac{1}{h}(f(x+h) - f(x)) = \frac{1}{h}(hf'(x) + \frac{1}{2}h^2 f''(x) + \frac{1}{6}h^3 f'''(x))$

$D_+ f(x) = f'(x) + \frac{1}{2}hf''(x) + \frac{1}{6}h^2 f'''(x)$

$D_+ f(x) - f'(x) = \frac{1}{2}hf''(x) + \frac{1}{6}h^2 f'''(x) + O(h^4)$

for sufficiently small h, the error is dominated by the first term so the method is first order accurate

3. Define local truncation error and describe the difference between the local truncation error and roundoff error/machine precision.

LTE - refers to the error in the finite differnece aproximation (some terms are left out leaving some uncertainty)

roundoff error is where the computer only has a few digits of precision and the rounding can change the final results

4. Derive the finite difference method for the approximation of f''(x), with the highest order of accuracy given the form f''(x) ≈ a_−1 f(x − h) + a_0 f(x) + a_1 f(x + h) $f''(x) \approx a_{-1} f(x-h) + a_0 f(x) + a_1 f(x+h)$

$a_{-1} f(x-h) = a_{-1} f(x) - a_{-1} hf'(x) + a_{-1} h^2 f''(x) - a_{-1} h^3 f'''(x) + O(h^4)$

$a_1 f(x+h) = a_1 f(x) + a_1 hf'(x) + a_1 h^2 f''(x) + a_1 h^3 f'''(x) + O(h^4)$

$a_0 f(x) = a_0 f(x)$

$f(x): a_{-1} + a_0 + a_1 = 0$ $\Rightarrow a_0 + \frac{1}{h^2} = 0 \Rightarrow \underline{a_0 = -\frac{1}{h^2}}$

$f'(x): -a_{-1}h + a_1 h = 0 \Rightarrow a_{-1} = a_1$

$f''(x): a_{-1}h^2 + a_1 h^2 = 1$ $\Rightarrow a_1 h^2 + a_1 h^2 = 1 \Rightarrow 2a_1 = \frac{1}{h^2} \Rightarrow \underline{a_1 = a_{-1} = \frac{1}{2h^2}}$

$$f''(x) \approx \frac{f(x-h) - 2f(x) + f(x+h)}{h^2}$$

5. For the following two-point boundary value problem u'' + σ u = f with u(a) = u_a and u(b) = u_b define a second order accurate approximation for the problem.

$u''(x_i) + \sigma u(x_i) = f(x_i)$

$\frac{U_{i-1} - 2U_i + U_{i+1}}{h^2} + \sigma U_i = F_i \Rightarrow U_{i-1} + U_i(h^2\sigma - 2) + U_{i+1} = h^2 F_i$

$U_{i-1} = u_i - hu_i' + \frac{1}{2}h^2 u''(x_i) - \frac{1}{6}h^3 u'''(x_i)$

$u_{i+1} = u(x_i) + hu'(x_i) + \frac{1}{2}h^2 u''(x_i) + \frac{1}{6}h^3 u'''(x_i)$

$\cancel{2u(x_i)} + h^2 u''(x_i) + O(h^4) + h^2\sigma u(x_i) - \cancel{2u(x_i)} = h^2 F_i$

$\Rightarrow O(h^2)$

6. Define the terms consistent finite difference method and stable finite difference method. How are these related to convergence of the approximations for a two-point boundary value problem?

Stability: $A^h u^h = F^h$, the method is stable if $(A^h)^{-1}$ exists for all small h

and there is a constant C indepent of h s.t. $\|(A^h)^{-1}\| \le C$

consistent: $\|\tau^h\| \to 0$ as $h \to 0$ $(\|\tau^h\| = O(h^p)$ $p>0$

consistency + stability $\Rightarrow$ convergence.

$$\|E^h\| \le \|(A^h)^{-1}\| \|\tau^h\| \le C\|\tau^h\| \to 0 \text{ as } h \to 0$$

7. Describe the steps needed to show that a finite difference method converges in the 2-norm. What is needed to prove stability of the finite difference method in the 2-norm?

1. $\|A\|_2 = \rho(A) = \max|\lambda_p|$
2. $\|A^{-1}\|_2 = \rho(A^{-1}) = \max|\lambda_p^{-1}| = (\min|\lambda_p|)^{-1}$
3. $\lambda_p = \frac{2}{h^2}(\cos(\pi p h) - 1)$
4. $\lambda_1 = -\pi^2 + O(h^2)$ ← bounded away from 0
5. $\|E^h\|_2 \le \|(A^h)^{-1}\|_2 \|\tau^h\|_2 \approx \frac{1}{\pi^2}\|\tau^h\|_2$

Needed: symmetric matrix, show λ are bounded away from 0 as $h \to 0$

8. What issues arise in finite difference methods in one dimension when Neumann boundary conditions are prescribed? has no solutions or infinitely many

1. one-sided approach leaves us with $O(h)$
2. centered difference (ghost point)
3. $\frac{1}{h}(\frac{3}{2}u_0 - 2u_1 + \frac{1}{2}u_2) = \sigma$ $u' = \sigma$ disturbs tridiag structure

9. Define and compare Dirichlet and Neumman boundary conditions in terms of a simple two-point boundary value problem.

Dirichlet: $u(a) = u_a$ $u(b) = u_b$ Neumman: $u'(a) = \alpha$ $u'(b) = \beta$

Dirichlet allows for m system matrix while Neumman requires adding more equations to solve for u_a + u_b using finite methods

10. In terms of linear algebra compare the finite difference schemes for the elliptic differential equations in one-dimension, two-dimensions, and three-dimensions? Use the form of the matrix in your discussions. The form and size of the matrices depends on the number of mesh points and the order of accuracy used for the finite diffence method. Typically we see one-dim as a tridiag, two-dim as penta diag and a three dimensional problem will have 7 diagonals

11. Compare three methods for ordering mesh points in a two-dimensional finite difference method for the elliptic problem. State the pros and cons of each of these orderings.

natural rowwise ordering - nice diagonal structure but the ones are separated from the main diagonal

red-black ordering - advantages for certain iterative methods but the matrix is much more sparse

12. In the solution of linear systems of equations, give a definition of the term, direct method. Give at least two examples of direct methods for the solution of a linear system of equations.

direct methods - produce a exact solution in a finite number of operations

examples:

gaussian elimination

LU decomposition

13. Define the term iterative method for the approximate solution of a linear system of differential equations. Give at least two examples of these types of methods.

iterative methods - start with an initial guess and improves it through iteration

examples:

Jacobi

Gauss-Seidel

14. Compare and contrast the 5-point stencil and 9-point stencil in the approximate solution of two-point boundary value problems.

Both discretizations have an error of $O(h^2)$ but for $f=0$ or f is harmonic then the 9-point laplacian can be 4th order accurate. This error reducing trick is easier with 9 pt but can still be done with a 5pt

15. Define the term diagonally dominant in terms of linear systems of equations. What can be said about linear solution methods like Gauss Elimination with Back substitution when the coefficient matrix is diagonally dominant?

By definition, a matrix is digionally dominant if $|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$

When using a solver method like gauss elim. or LU decomp. a diagonally dominant matrix helps simplify the problem because no pivoting is required

16. Define the 2-condition number of a square matrix. How can we use the condition number of a matrix to determine how accurate an approximate solution of a linear system is?

condition number - $\kappa(A) = \|A\| \|A^{-1}\|$. For a normal matrix the condition number is the ratio of the largest to smallest eigenvalue. The condition number is used with the convergence rate for many iterative methods. When working with the conjugate gradient the condition number helps give an upper bound on the reduction of the error

17. Define the term Toeplitz matrix. Give some examples that arise in the approximate solution of differential equations.

A Matrix is Toeplitz if the value along each diagonal is constant.

$$u''(x) = f(x) \Rightarrow \begin{bmatrix} -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \cdots & \\ 0 & 1 & -2 & 1 & \vdots \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & & 1 & -2 \end{bmatrix}$$

$$u''(x,y) = f(x,y) \Rightarrow \begin{bmatrix} -4 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 1 & -4 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots 0 \\ 0 & 1 & -4 & 1 & 0 & \cdots & 0 & 1 & 0 \cdots \\ & & \ddots & \ddots & \ddots & & & \ddots & \\ 1 & 0 & \cdots & 0 & 1 & -4 & 1 & 0 & \cdots \\ & \ddots & & & & \ddots & \ddots & \ddots & \end{bmatrix}$$

18. Give examples of when we might consider the use of LU-factorization instead of Gaussian Elimination with Backsubstitution.

LU-factorization is useful for problems where just the rhs changes but the matrix A stays the same. This way you already have a factorized A and the order of operations is reduced.

19. Define the term vector-norm and give at least three useful vector norms. Make sure you list the properties that must be satisfied to be a norm.

defined to be mapping vectors x to nonnegative real numbers

Properties: 1. $\|x\| \geq 0$ for any $x \in \mathbb{R}^s$ and $\|x\| = 0$ if and only if $x = \vec{0}$

2. If a is any scaler, then $\|ax\| = |a|\|x\|$

3. If $x, y \in \mathbb{R}^m$, then $\|x+y\| \leq \|x\| + \|y\|$

- $\|e\|_\infty = \max_{1 \leq i \leq s} |e_i|$
- $\|e\|_1 = \sum_{i=1}^{s} |e_i|$
- $\|e\|_2 = \sqrt{\sum_{i=1}^{s} |e_i|^2}$

20. Define the term matrix-norm and give at least three examples of matrix norms and the properties that must be satisfied.

A matrix A can be consider as a particular kind of vector an it's norm is an function that maps A to a real number.

Properties: 1: $\|A\| \geq 0$, $\|A\| = 0$ iff $A = 0$

2. $\|aA\| = |a|\|A\|$

3. $\|A+B\| \leq \|A\| + \|B\|$

4. $\|AB\| \leq \|A\| \cdot \|B\|$

- $\|A\|_1 = \max_{1 \leq j \leq s} \sum_{i=1}^{s} |a_{ij}|$
- $\|A\|_\infty = \max_{1 \leq i \leq s} \sum_{j=1}^{s} |a_{ij}|$
- $\|A\|_2 = \sqrt{\rho(A^T A)}$

21. What is a consistent matrix-norm? Give an example of a consistent matrix norm. Give an example of where a consistency norm is needed in our analysis.

A matrix norm $\|\cdot\|$ on $K^{m \times n}$ is called consistent with a vector norm $\|\cdot\|_a$ on K^n and a vector norm $\|\cdot\|_b$ on K^m if $\|Ax\|_b \leq \|A\|\|x\|_a$ for all $A \in K^{m \times n}$

All induced matrices are considered consistent matrix-noms. A consistency norm would be needed when dealing with vectors of different size to see if the solutions are consistent

22. What is an induced matrix norm? Use the 2-matrix norm to illustrate the parts of your discussion?

Induced norm: $\|A\| = \max\{\|Ax\| : x \in K^n \text{ with } \|x\| = 1\} = \max\{\frac{\|Ax\|}{\|x\|} : x \in K^n \text{ with } x \neq 0\}$

The matrix norm is induced by the vector norm.

This means that using the definition the matrix norm can be induced.

For the 2-matrix norm we use the vector 2 norm $\|\cdot\|_2 = \sqrt{\sum |\cdot|^2}$ in the induced norm definition and the derive the matrix norm definition.

23. Discuss how one can perform a computational convergence analysis of a sequence of approximations to determine the rate of convergence of a particular finite difference method.

with a sequence of approximations we get $X = \{\log h_0, \log h_1, \log h_2, \dots \log h_m\}$

$y = \{\log\|E^{h_0}\|, \log\|E^{h_1}\|, \dots \log\|E^{h_m}\|\}$

representing the system of equations as

$$\begin{bmatrix} \text{length of } X & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

We solve for b which gives us the rate of convergence.

4/4