

The midterm exam is on Thursday February 28 in class. It will cover all the class material up to and including Thursday February 21. You may use a calculator to do arithmetic and one sheet of handwritten notes (i.e. one side of one page, 8.5 in \times 11 in). You must justify your answers to receive full credit. All vector and matrix norms are the ∞ -norm.

1. True or False? Give a reason to justify your answer.

a) $(10101.01)_2 = (21.25)_{10}$ b) $D_+D_-f(x) = D_-D_+f(x)$ c) $D_+D_+f(x) = f''(x) + O(h^2)$

d) When the derivative $f'(x)$ is approximated by the forward difference approximation $D_+f(x)$ with step size h in finite precision arithmetic, for large h the roundoff error dominates the truncation error, but for small h the truncation error dominates the roundoff error.

e) The central difference approximation defined by $D_0f(x) = \frac{f(x+h)-f(x-h)}{2h}$ is a 2nd order accurate approximation to the first derivative $f'(x)$.

f) The approximation $D_0f(x) \approx f'(x)$ is exact when $f(x)$ is a polynomial of degree less than or equal to three.

g) In solving a nonlinear equation $f(x) = 0$ by fixed-point iteration $x_{n+1} = g(x_n)$, the method will converge when the initial guess x_0 is sufficiently close to the root r and $g'(r) = 0$.

h) If A is invertible, then there exists a nonzero vector x such that $Ax = 0$.

i) In solving a linear system with three equations and three unknowns, in step 1 of Gaussian elimination, variable x_1 is eliminated from equations 1, 2, and 3.

j) In solving a general $n \times n$ linear system by Gaussian elimination, if the dimension n is doubled, then the operation count is also approximately doubled.

k) In solving a linear system of equations by Gaussian elimination, pivoting is recommended even if the pivots are nonzero, in order to reduce the operation count.

l) If x is the input to a system and Ax is the output, then the norm of the output is less than or equal to $\|A\|$ times the norm of the input.

m) Gaussian elimination is an unstable method for solving $Ax = b$ because it can replace an ill-conditioned matrix A by a well-conditioned matrix U .

n) In solving a linear system $Ax = b$ by a numerical method, if the residual norm $\|r\|$ is small, then the error norm $\|e\|$ is also small.

o) In solving a two-point boundary value problem using a 2nd order accurate finite-difference scheme, if the mesh size h is reduced by one half, then the norm of the error is reduced by approximately one fourth.

p) Consider an iterative method $x_{k+1} = Bx_k + c$ for solving $Ax = b$. If $\|B\| < 1$, then the method converges for all initial vectors x_0 .

2. Matlab gives $\sqrt{5} = 2.236067977499790$. Express $\sqrt{5}$ in normalized floating point form, $\pm(0.d_1 \dots d_n)_\beta \cdot \beta^e$, with $d_1 \neq 0$, on a computer with $\beta = 2, n = 4, M = 3$, and then express the result in decimal form.

3. Let $f(x) = \sqrt{1+x} - \sqrt{1-x}$ and $g(x) = 2x/(\sqrt{1+x} + \sqrt{1-x})$. Show that $f(x) = g(x)$ for all x such that $|x| \leq 1$. If you are using finite precision arithmetic, which expression is better to use when $x \approx 0$? Explain.

4. Consider the finite-difference approximation $f'(x) \approx \frac{af(x+h) + bf(x) + cf(x-h)}{h}$, where a, b, c are constants. The forward approximation D_+f has $(a, b, c) = (1, -1, 0)$ and is 1st order

accurate. The central approximation D_0f has $(a, b, c) = (\frac{1}{2}, 0, -\frac{1}{2})$ and is 2nd order accurate. Are there any values of (a, b, c) that yield 3rd order accuracy?

5. Below is the algorithm for the bisection method. Find and correct any bugs.

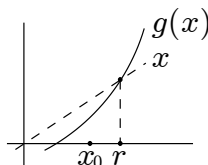
bisection method (assume $f(a) \cdot f(b) < 0$)

1. $n = 0$, $a_0 = a$, $b_0 = b$
2. $x_n = \frac{1}{2}(a_n + b_n)$
3. if $f(x_n) \cdot f(a_n) < 0$, then $a_{n+1} = a_n$, $b_{n+1} = x_n$
4. else $a_{n+1} = x_n$, $b_{n+1} = b_n$
5. set $n = n + 1$ and go to line 1

6. Consider solving $f(x) = 0$. (a) State one advantage of Newton's method over the bisection method; (b) State one advantage of the bisection method over Newton's method.

7. Show that $f(x) = x^2 - 3x + 2 = 0$ is equivalent to $x = g(x) = \frac{1}{3}x^2 + \frac{2}{3}$. Suppose fixed-point iteration $x_{n+1} = g(x_n)$ is applied with initial guess $x_0 = 0$. Find $\lim_{n \rightarrow \infty} x_n$. Justify your answer.

8. Consider fixed-point iteration $x_{n+1} = g(x_n)$. The figure shows the function $y = g(x)$, the line $y = x$, the fixed point r , and the initial guess x_0 . Does the sequence x_n converge to r in this case? Explain.



9. The screened Coulomb potential is defined by $\phi(x) = \frac{e^{-\kappa x}}{4\pi\epsilon x}$, where x is the distance from a charged particle to a point in space, ϵ is the dielectric constant, and κ controls the screening effect. Let $\epsilon = 2, \kappa = \frac{1}{2}$. Apply Newton's method to find the value of x for which $\phi(x) = 0.005$. Let $x_0 = 2$ be the starting value and take two steps, x_1, x_2 . How many digits in x_1 are correct?

10. Consider the nonlinear system, $f(x, y) = (x - 1)^2 + y^2 - 4 = 0, g(x, y) = xy - 1 = 0$, which gives the intersection of a circle and a hyperbola in the xy -plane. Find an approximate solution using Newton's method for systems. Take one step starting from $(x_0, y_0) = (3, 0)$.

11. Solve $2x_1 - x_2 + x_3 = -1, 4x_1 + 2x_2 + x_3 = 4, 6x_1 - 4x_2 + 2x_3 = -2$ by Gaussian elimination.

12. Solve $Ax = b$ by Gaussian elimination with partial pivoting.

a) $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ b) $A = \begin{pmatrix} 0 & 4 & -15 \\ 10 & 0 & 15 \\ 1 & -1 & -1 \end{pmatrix}, b = \begin{pmatrix} -12 \\ 100 \\ 0 \end{pmatrix}$

13. Let $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$. Find a vector x such that $\frac{\|Ax\|}{\|x\|} = \|A\|$.

14. Fill in the blanks. In solving a linear system $Ax = b$, the _____ of the matrix A controls the relative error in the solution x due to _____ in the right hand side b .

15. Suppose $Ax = b$ and $A\tilde{x} = \tilde{b}$, where $A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and $\|b - \tilde{b}\| \leq 10^{-2}$. Find the maximum value that $\|x - \tilde{x}\|$ can attain.

16. Consider the linear system $2x_1 - x_2 = 1, -x_1 + 2x_2 - x_3 = 0, -x_2 + 2x_3 = 1$, with solution $x_1 = x_2 = x_3 = 1$. a) Write Jacobi's method in component form. Take two steps starting from the zero vector. Compute the error norms $\|e_k\|, k = 0 : 2$. b) Repeat for Gauss-Seidel.

17. Let $A_1 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. a) For which of these does Jacobi's method converge? b) For which of these does Gauss-Seidel converge?