Homework 1

1. **Binary Representation** Write the following base 10 numbers "in binary".

   [i] 27

   [ii] 15.140625

2. **Floating Point Arithmetic** *For these exercises assume that numbers are expressed in the form* $r = \pm 0.d_1 d_2 ....d_k \times 10^t$*, with* $k = 5$ *and* $t \in [-5, 5]$*). The* d_i *are integers in* $[0, 9]$.

   [i] What is the largest number that can be expressed in this system? the smallest in magnitude?

   [ii] If chopping is used (rather than rounding), the number $.12300 \times 10^3$ represents what interval of real numbers.

   [iii] If $a = .12300 \times 10^3$ and $b = .12300 \times 10^{-2}$, calculate the relative and absolute error the computer will make when evaluating the following expressions: $a + b, a - b, (a + b) - a, a/b, b/a, ((a + b) - a)/b$.

3. **Bisection:** Calculate $\sqrt{2}$ by hand to five decimal places using bisection. Show your computation in a table of the successive intervals.

Homework 2

4. **Fixed Point Algorithm**

   [i] Show that the fixed point algorithm converges locally when $|f'(x)| < k < 1$ in the neighborhood of the fixed point.

   [ii] Under what circumstance will the fixed point iteration converge faster than linearly.

   [iii] Consider the fixed point algorithm $x_{i+1} = 1/3(x_i^2 + 2)$ as a method to find the roots of $x^2 - 3x + 2$. Is this iteration locally convergent near $x = 2$? Near $x = 1$?

   [iv] Write your own fixed point algorithm to find the roots of $x^2 - 3x + 2$, such that it converges at $x = 2$.

Homework 3

5. **Newton's Method**

   [i] Show that Newton's method converges quadratically to a root p of a twice differentiable function f, if $f'(p) \neq 0$ and $f''(p) \neq 0$.

   [ii] Consider the function $f(x) = xe^{-x^2}$.

   [ii.a] Plot this function

   [ii.b] Find all roots.

   [ii.c] Find the first two iterates of Newton's method when the initial guess is $x_0 = 1$. Will this guess converge? Illustrate your conclusion on the graph.

   [ii.d] Find a guess which will converge to $x = 0$.

   [iii] Implement Newton's method for $f(x) = xe^{-x^2}$. For a convergent initial guess, confirm the error prediction by plotting the log of the error at the n+1 iteration vs the log of the error at the nth iteration.

Homework 4

6. **Newton's Method** Does Newton's method ever converge faster than quadratically? Explain by writing the appropriate power series expansions. What does this imply about the associated fixed point method?
7. **Secant Method** Implement the secant method for $f(x) = xe^{-x}$. For a convergent initial guess, observe the convergence rate by plotting the log of the error at the n+1 iteration vs the log of the error at the nth iteration. Compare this rate to that of Newton's method and the fixed point method on the same function.
8. **Regula Falsi** Draw a diagram illustrating the method of false position (Regula Falsi).
9. **Müller's Method** Draw a diagram illustrating Müller's Method.

Homework 5

10. Fix all problems you got incorrect on Quiz 1 and Quiz 2. Turn in correct solutions.
11. Study for exam.
12. Begin thinking about your course project.

Homework 6

13. Fix all problems you got incorrect on Exam 1. Turn in correct solutions.

Homework 7

14. Find the determinant, using cofactors, of

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 0 & 3 & 3 \\ 2 & 4 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 3 & 3 \\ 0 & 3 & 3 \\ 8 & 1 & 1 \end{pmatrix}$$

15. Solve the system

$$\begin{pmatrix} 1 & 3 & 5 \\ 0 & 3 & 3 \\ 2 & 4 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 8 \end{pmatrix}$$

16. Find a solution to the system

$$\begin{pmatrix} 1 & 3 & 3 \\ 0 & 3 & 3 \\ 8 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 8 \end{pmatrix}$$

17. Evaluate

    [i] $\sum_{i=1}^{n}(1 - 3i + n)$

    [ii] $\sum_{i=1}^{n} 2i^2 - 5$

Homework 8

18. Use an appropriate pivoting algorithm to solve $Gx = f$. Compare the exact answer, the floating point answer without pivoting, and the floating point answer with appropriate pivoting. Use a floating point system where eps is the smallest number which when added to one is distinguishable from 1, and $100/eps < INF$.

$$G = \begin{pmatrix} 1 & \frac{3}{eps} & 3 \\ eps & 1 & 4 \\ 4 & 5 & 2 \end{pmatrix} \qquad f = \begin{pmatrix} 1 \\ 5 \\ 3 \end{pmatrix}$$

19. Find an upper triangular matrix U and a lower triangular matrix L where the diagonal of L is all ones, with $A = LU$.

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 3 & 3 \\ 2 & 3 & 5 \end{pmatrix}$$

20. For what z does the Jacobi method converge when applied to S.

$$S = \begin{pmatrix} 1 & 2 \\ 1 & z \end{pmatrix}$$

21. For what z does the Gauss-Seidel method converge when applied to M.

$$M = \begin{pmatrix} 1 & 2 \\ -1 & z \end{pmatrix}$$

Homework 9

22. Consider the matrices,

$$A_1 = \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 1 & -3/4 \\ -1/12 & 1 \end{pmatrix}$$

[i] Are these matrices diagonal dominant? For which matrix is the ratio of the off-diagonal entries to the diagonal entries smaller? For which matrix do you think the Jacobi Method converges faster?

[ii] If the Jacobi method is written $x^{k+1} = Mx^k + f$, calculate M for both A_1 and A_2. Compute the eigenvalues of each M.

[iii] The spectral radius $\rho(M)$ is the magnitude of the largest eigenvalue of M, find the spectral radius of each M. For which matrix does the Jacobi method converges faster?

23. Consider the matrix

$$S = \begin{pmatrix} 1 & \rho \\ -\rho & 1 \end{pmatrix}$$

[i] For what ρ does the Gauss-Seidel method converge for this matrix.

[ii] For which ω does the SOR method converge fastest.

24. Values of a function are known at 5 x values, $x_i = \{-2, 0, 3, 5, 7\}$, where it takes values $f(x_i) = \{3, 0, 3, -1, 2\}$. Find the polynomial of degree 4 which pass through these points using the Lagrange polynomial and Newton's divided difference.

Homework 10

25. Observe Runge-phenomenon by sampling $f(x) = e^{-x^2}$ at 20 equally spaced points in $[-1, 1]$, and comparing the polynomial interpolant of degree 19 to the actual function. *This problem should be done using a computer, turn in only the plot.*

26. Observe the error in a polynomial interpolant by plotting the $E(x) = C\prod(x - x_i)$, with $C = 1$, for

[i] $\{x_i\}$ nine equally spaced points in $[-1, 1]$

[ii] $\{x_i\}$ nine Chebyshev points in $[-1, 1]$

27. Find the line which best fits $(1, 0), (2, 3), (4, 5), (2, 0), (1, 1), (-1, -2)$, with error measured in the the vector 2 norm.

Homework 11

28. Fix your mistakes from Exam 2; turn in corrected problems.

29. Find the truncation error for the centered difference discretization of the derivative
$$f_c'(x) \equiv \frac{f(x+h) - f(x-h)}{2h}.$$

Homework 12

30. In class we showed that for forward difference the best step size was $\sqrt{eps}$.

   [i] Find the best step size for centered/leapfrog difference
$$f_c'(x) \equiv \frac{f(x+h) - f(x-h)}{2h}$$

   [ii] If we use the best step size for each method, which has smaller error (truncation and precision)?

31. In class we derived the following approximation of the first derivative
$$f_R'(x) \equiv \frac{4f(x+h/2) - 3f(x) - f(x+h)}{h}$$
   This approximation acts on step size $x_{i+1} - x_i = h/2$.

   [i] Rewrite the method so that it can be used on discretization with step size h.

   [ii] Find the truncation error of this method. Show that this method is second order, $O(h^2)$, and find the coefficient of the h^2 term.

32. Consider a general approximation $B(x,h)$ to a function $b(x)$ with truncation error
$$E = B(x,h) - b(x) = c_1 h + c_2 h^2$$

   [i] Use Richardson Extrapolation to find a second order approximation to $b(x)$.

   [ii] Use Richardson Extrapolation again to get an exact approximation for $b(x)$ in terms of B, (one with no truncation error).

Homework 13

33. Write a code which implements both the midpoint rule and simpson's rule (with weights 1 4 2 4 ... 2 4 1) for evaluating integrals. Plot the log of the error as a function of the log of h when approximating $I_1 = \int_0^\pi \sin(x)$ and $I_2 = \int_0^2 x$.

34. Find the coefficients a_i of a quadrature rule $\int f(x)dx = \sum a_i f(x_i)$ which uses cubics to approximate $f(x)$ on the interval (x_i, x_{i+3}).

35. In class we derived the gaussian quadrature points x_i and coefficients a_i when $n = 2$. Derive the weights a_i and the points x_i for a gaussian quadrature scheme when $n = 3$.