

Mid Term exam: CMSC/AMSC/MAPL 460 Computational Methods 03/15/2007

1. Write down true or false: (2 points/ each)
 - a. Solving a least-squares problem via the normal equations rather than the QR decomposition leads to a better conditioned algorithm False
 - b. The matrix that results on rearranging the rows of an identity matrix is called a permutation matrix. True
 - c. The spacing between floating point numbers is uniform False
2. Write brief descriptions of the following. In each case provide any pertinent details such as formulae, but be brief:
 - a. Why is solving a system of equations by computing the inverse and then multiplying it by the right hand side usually a bad idea? (4 points)
Costlier, and more prone to error.
 - b. A cubic spline requires four pieces of information over every piece over which the function is approximated. Write down what these are for i) an internal section and, ii) an end section for the natural cubic spline (6 points)
 - i) value at the two endpoints, continuity of 1st and 2nd derivatives at end points
 - ii) value at the two endpoints, continuity of 1st and 2nd derivatives at interior point, and specified 2nd derivative (=0) at end point
3. Assume you have a base 2 computer that stores floating point numbers using 10 bits, which are divided into a sign bit, a 5 bit normalized mantissa (1.xxxxx), and a 4 bit exponent the 1st bit of which is used for a sign. Assume that all numbers are chopped rather than rounded.
 - a) Draw a small representation of the numbers and what the values of each field can be. Give the machine representation and a base 10 representation for machine epsilon, the smallest nonzero positive number which, added to 1, gives a number different from 1. Also provide the largest number that can be stored in this representation in decimal is represented (20 points)

A representation could be as [s|eeee|ffff], or as $(-1)^s \times 1.xxxxx \times 2^{(-1)xxx}$

s can take value 0 or 1

e can take value from -111 to 111 or (-7 to 7 in decimal)

f can take value from 00000 to 11111 (0 to $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32}$)

machine epsilon is $\frac{1}{32}$ (or $\frac{1}{2}$ that 0.03125 or 0.015625)

b) Which machine number is closest to π ? ($\pi=3.1415926535897\dots$)

$11.00100 = 1.10010 \times 2^{+1} = 2(1 + \frac{1}{2} + \frac{1}{16}) = 3.125$ is the right answer

(5 points)

4. Consider the data points (x_i, y_i) given by $(-3, 1), (-1, -1), (1, 1), (3, 4)$. We wish to approximate it using the parabola

$$y = a + bx + cx^2$$

and must determine the coefficients a, b and c .

- a. Write down the normal equations for solving this problem via least squares
Our equation for the coefficients can be written as

$$\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}$$

Substituting for (x_i, y_i)

$$\begin{bmatrix} 1 & -3 & 9 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 3 & 9 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 4 \end{bmatrix}$$

This equation has the form $\mathbf{Ax}=\mathbf{b}$. The normal equations are $\mathbf{A}^t\mathbf{Ax}=\mathbf{A}^t\mathbf{b}$

- b. Solve the resulting system of equations via LU decomposition

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \\ 9 & 1 & 1 & 9 \\ 1 & 3 & 9 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \\ 9 & 1 & 1 & 9 \\ 1 & 3 & 9 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 0 & 20 \\ 0 & 20 & 0 \\ 20 & 0 & 164 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \\ 45 \end{bmatrix}$$

So we need to solve the above system via LU decomposition. We could do this with pivoting or without. For this example let us do it without.

Eliminate the 20 in the first column by subtracting 5 times 1st row.

$$\begin{bmatrix} 4 & 0 & 20 \\ 0 & 20 & 0 \\ 0 & 0 & 64 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \\ 20 \end{bmatrix}$$

and we are done !

$$c = 20/64 \quad b = 11/20 \quad a = (5 - 20 \times 20/64)/4 = 5 - 25/16$$

$$U = \begin{bmatrix} 4 & 0 & 20 \\ 0 & 20 & 0 \\ 0 & 0 & 64 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix}$$

Check $LU = A$

So the fit function is

$$y = 45/16 + 11x/20 + 20x^2/64$$

5. For the set of data points in Q. 4 determine, write down the Lagrange form of the *interpolating* polynomial. (15 points)

There are 4 points so they will be fit by a 3rd order polynomial

$$y = \sum_{i=1}^N y_i \prod_{j=1, j \neq i}^N \frac{(x - x_j)}{(x_i - x_j)}$$

$$y = \frac{(x+1)(x-1)(x-3)}{(-3+1)(-3-1)(-3-3)} - \frac{(x+3)(x-1)(x-3)}{(-1+3)(-1-1)(-1-3)} + \frac{(x+3)(x+1)(x-3)}{(1+3)(1+1)(1-3)} + \frac{4(x+3)(x+1)(x-1)}{(3+3)(3+1)(3-1)}$$

$$y = \frac{(x+1)(x-1)(x-3)}{(-2)(-4)(-6)} - \frac{(x+3)(x-1)(x-3)}{(2)(-2)(-4)} + \frac{(x+3)(x+1)(x-3)}{(4)(2)(-2)} + \frac{4(x+3)(x+1)(x-1)}{(6)(4)(2)}$$

$$y = -\frac{(x+1)(x-1)(x-3)}{48} - \frac{(x+3)(x-1)(x-3)}{16} - \frac{(x+3)(x+1)(x-3)}{16} + \frac{(x+3)(x+1)(x-1)}{12}$$

6. Let $f(x) = x^2 - 5 = 0$. Perform three iterations each using the

a. Bisection method, with a starting interval of [2,3]

(7 points)

$$f(2) = -1; f(3) = 4$$

$$f(2.5) = 1.25$$

New interval [2, 2.5]

$$f(2.25) = 81/16 - 5 = 1/16 > 0$$

New interval [2, 2.25]

$$f(2.125) = 289/64 - 5 = -31/320$$

New interval = [2.125, 2.25]

Final guess = 2.1875

b. Newton's method with a guess of $x=2$

(7 points)

$$f'(x) = 2x$$

$$x_{k+1} = x_k - f(x_k)/f'(x_k)$$

$$1^{\text{st}} \text{ iteration: } x_1 = 2 - (-1)/4 = 2.25$$

$$2^{\text{nd}} \text{ iteration: } x_2 = 2.25 - (1/16)/4.5 = 2.25 - 1/72 = 9/4 - 1/72 = 161/72 = 2.23611$$

$$3^{\text{rd}} \text{ iteration: } x_3 = 161/72 - ((161/72)^2 - 5)/(322/72) = 161/72 - (161^2 - 5 \times 72^2)/(72 \times 322) = 2.2360679779158$$

Actual answer is 2.23606797749979