

### 1 Error analysis (20 pts)

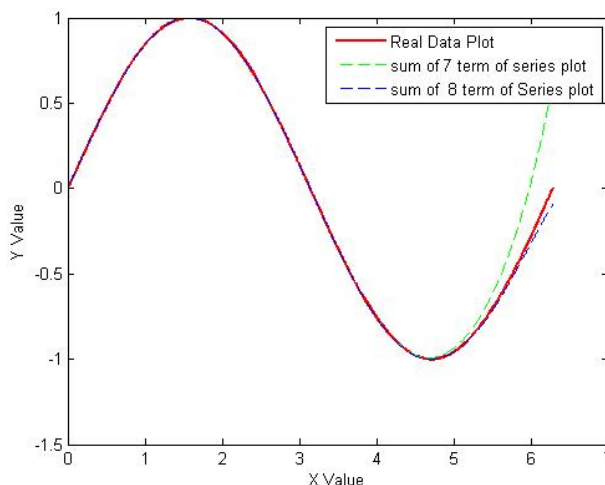
Consider the function  $\sin(x)$  and the following Taylor series:

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \dots$$

(a) Use MATLAB/OCTAVE to plot :

- The real  $\sin(x)$  function,
- The sum of the first 7 terms of the series,
- The sum of the first 8 terms of the series

on the same graph. What change do you observe as you use more terms from the series?



(b) Use MATLAB/OCTAVE to sum the first 7 terms of the series for  $x = 1.4 \pi$ . What is the relative true error?

$$y = -0.9511 = \sin(1.4 \pi)$$

When using this formula to sum the first 7 terms of the series for  $x = 1.4 \pi$  is

When  $x = 1.4 \pi$

$$y^* = -0.9479 = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \dots$$

- True Error = True Value - Approximate Value :  
 $|y - y^*| = |-0.9511 - 0.9479| = 0.0032$
- Relative true error

$$\frac{|y - y^*|}{|y|} = \frac{0.0032}{-0.9511} = 0.0033$$

- Percent error :  $0.0033 \times 100 = 0.33\%$

(c) How many terms do you need to include in the sum for the relative true error to become less than 1%?

The terms do you need to include in the sum for the relative true error to become less than 1% is 2.

(d) How many terms do you need to include in the sum for the relative approximation error to become less than 1%?

The terms do you need to include in the sum for the relative approximation error to become less than 1% is 4.

## 2 Error measurement on floating point number (15 pts)

Consider a computer that uses 16 bits to represent floating-point numbers (Table 1). 1 bit for sign of number, 1 bit for sign of exponent, 4 bit for exponent, 10 bit for mantissa

Table 1: 16 bits to represent floating-point numbers

sn	se	e	e	e	e	m	m	m	m	m	m	m	m	m	m
----	----	---	---	---	---	---	---	---	---	---	---	---	---	---	---

(a) What is the machine epsilon for this computer?

The machine epsilon is as follow.

$$2^{-10} = 9.7656 \times 10^{-4}$$

(b) How would  $(571, 632)_{10}$  be represented in this system? What is the relative true error in this representation?

$$(571, 632)_{10} = (10001011100011110000)_2$$

$$\text{In this system: } (10001011100000000000)_2 = 571392$$

$$\text{Thus, the relative true error} = \frac{571632 - 571392}{571632} = \frac{240}{571632} = 0.00041985$$

## 3 Linear Equations (10 pts)

Solve the system below using LU decomposition

$$x_1 + 2x_2 + x_3 = 0$$

$$8x_2 + 6x_3 = 10$$

$$2x_1 + 5x_3 = 11$$

The linear equation can be written as follow.

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 8 & 6 \\ 2 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 10 \\ 11 \end{bmatrix}$$

Step 1: finding the U matrix

$$\text{Row3-Row1*2} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 8 & 6 \\ 0 & -4 & 3 \end{bmatrix}$$

$$\text{Row3-Row2*(-1/2)} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 8 & 6 \\ 0 & 0 & 6 \end{bmatrix}$$

$$\text{Thus, [U]} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 8 & 6 \\ 0 & 0 & 6 \end{bmatrix}$$

Step 2: finding the L matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2/3 & 0 \\ 0 & 8 & 6 \\ 2 & 0 & 5 \end{bmatrix} : l_{31} = \frac{a_{31}}{a_{11}} = 2$$

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 8 & 6 \\ 0 & -4 & 3 \end{bmatrix} : l_{32} = \frac{a_{32}}{a_{22}} = -\frac{1}{2}$$

$$\text{Thus, [L]} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -1/2 & 1 \end{bmatrix}$$

$$\text{Set [L][Z]=[C]} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -1/2 & 1 \end{bmatrix} \begin{bmatrix} z1 \\ z2 \\ z3 \end{bmatrix} = \begin{bmatrix} 0 \\ 10 \\ 11 \end{bmatrix}$$

$$z1=0$$

$$z2=10$$

$$z1*2+z2*(-1/2)+z3=11$$

$$z3=16$$

$$[Z] = \begin{bmatrix} 0 \\ 10 \\ 16 \end{bmatrix}$$

$$\text{Set [U][X]=[Z]}$$

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 8 & 6 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} 0 \\ 10 \\ 16 \end{bmatrix}$$

$$x3=8/3$$

$$x2*8+x3*6=10$$

$$x2=-3/4$$

$$x1+x2*2+x3=0$$

$$x1=-7/6$$

Thus, the solution vector

$$\begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} -7/6 \\ -3/4 \\ 8/3 \end{bmatrix}$$

#### 4 Interpolation (30 pts)

Scientists try to understand following observation data and come with a formulation to determine the relationship between x and y in Table 2.

Table 2: Observation data.

X(cm) y(cm)

2.00 7.2

4.25 7.1

5.25 6.0

(a) Using quadratic Lagrange (second order Lagrange polynomial) method what is the value of y at x = 4.00

(b) Plot the resulting functions in MATLAB/OCTAVE

(a) Using quadratic Lagrange (second order Lagrange polynomial) method what is the value of y at x = 4.00

$$y = a_0 + a_1x + a_2x^2$$

$$7.2 = a_0 + 2a_1 + 4a_2$$

$$7.1 = a_0 + 4.25a_1 + 4.25^2a_2$$

$$6.0 = a_0 + 5.25a_1 + 5.25^2a_2$$

Solving the above three equations gives

$$a_0 = 4.5282$$

$$a_1 = 1.9855$$

$$a_2 = -0.3248$$

Thus,

$$y = 4.5282 + 1.9855x - 0.3248x^2$$

The value of y at x=4.00

$$y = 4.5282 + 1.9855 * 4 - 0.3248 * 16 = 7.2734$$

(b) Plot the resulting functions in MATLAB/OCTAVE

The plot of resulting function is as follow.



## 5. Root calculation (25 pts)

Write a MATLAB/OCTAVE code that calculates the root of the function:

$$f(x) = x \cdot \cos(x) - \sin(x)$$

using the Newton-Raphson method. Conduct three iterations to estimate the root of the above equation. Calculate and plot the absolute relative approximate error at the end of each iteration using initial condition  $x_0 = 1$ .

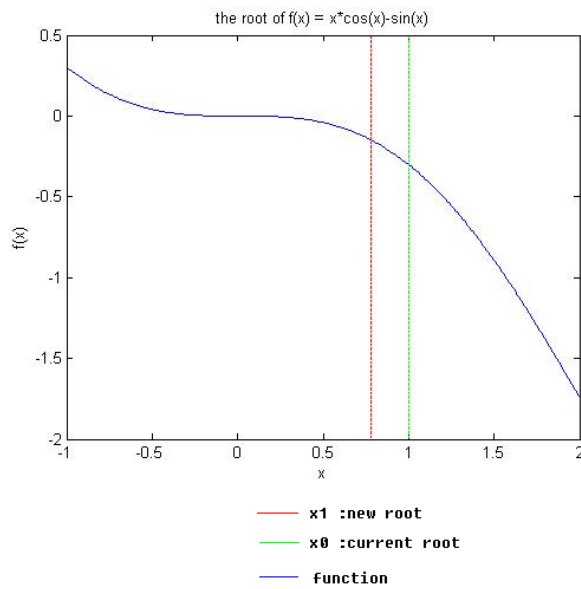
$$f(x) = x \cdot \cos(x) - \sin(x)$$

$$f'(x) = -x \cdot \sin(x) - \cos(x)$$

Iteration 1:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 1 - \frac{1 \cdot \cos(1) - \sin(1)}{-1 \cdot \sin(1) - \cos(1)} = 0.7820$$

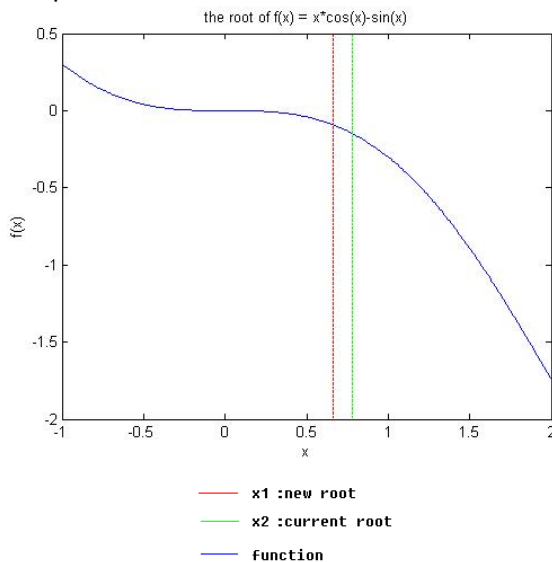
$$\epsilon_a = \left| \frac{x_1 - x_0}{x_1} \right| \times 100 = 27.9$$



Iteration 2:

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 0.7820 - \frac{0.7820 * \cos(0.7820) - \sin(0.7820)}{-0.7820 * \sin(0.7820) - \cos(0.7820)} = 0.6631$$

$$\epsilon_a = \left| \frac{0.6631 - 0.7820}{0.6631} \right| \times 100 = 17.9$$



Iteration 3:

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 0.6631 - \frac{0.6631 * \cos(0.6631) - \sin(0.6631)}{-0.6631 * \sin(0.6631) - \cos(0.6631)} = 0.5854$$

$$\epsilon_a = \left| \frac{0.5854 - 0.6631}{0.5854} \right| \times 100 = 13.3$$

