School of Mathematics, Thapar Institute of Engineering & Technology, Patiala

UMA007 : Numerical Analysis
Assignment 1
Floating Point Arithmetic and Errors

1. Compute the absolute error and relative error in approximations of x by x^*.
a. $x = \pi,\ x^* = 22/7$ b. $x = \sqrt{2},\ x^* = 1.414$ c. $x = 8!,\ x^* = 39900$.

2. Find the largest interval in which x^* must lie to approximate x with relative error at most 10^{-4} for each value of x.
a. π b. e c. $\sqrt{3}$ d. $\sqrt[3]{7}$.

3. Use three-digit rounding arithmetic to perform the following calculations. Compute the absolute error and relative error with the exact value determined to at least five digits.
a. $\sqrt{3} + (\sqrt{5} + \sqrt{7})$ b. $(121 - 0.327) - 119$ c. $-10\pi + 6e - \dfrac{3}{62}$ d. $\dfrac{\pi - 22/7}{1/17}$.

4. The error in the measurement of area of a circle is not allowed to exceed 0.5%. How accurately the radius should be measured.

5. Calculate the value of $x^2 + 2x - 2$ and $(2x - 2) + x^2$ where $x = 0.7320e0$, using normalized point arithmetic and proves that they are not the same. Compare with the value of $(x^2 - 2) + 2x$.

6. Use four-digit rounding arithmetic and the formula to find the most accurate approximations to the roots of the following quadratic equations. Compute the absolute errors and relative errors.

$$\frac{1}{3}x^2 + \frac{123}{4}x - \frac{1}{6} = 0.$$

7. Verify that the functions $f(x)$ and $g(x)$ are identical functions.

$$f(x) = 1 - \sin x, \quad g(x) = \frac{\cos^2 x}{1 + \sin x}.$$

a. Which function should be used for computations when x is near $\pi/2$? Why?
b. Which function should be used for computations when x is near $3\pi/2$? Why?

8. a. Evaluate the polynomial $y = x^3 - 7x^2 + 8x - 0.35$ at $x = 1.37$ using 3-digit arithmetic with chopping. Evaluate the percent relative error.
b. Repeat part (a) but express y as $y = ((x - 7)x + 8)x - 0.35$. Evaluate the error and compare with part (a).

9. The derivative of $f(x) = \dfrac{1}{(1 - 3x^2)}$ is given by $\dfrac{6x}{(1 - 3x^2)^2}$. Do you expect to have difficulties evaluating this derivative at $x = 0.577$? Try it using 3- and 4-digit arithmetic with chopping.

10. A rectangular parallelepiped has sides of length 3 cm, 4 cm, and 5 cm, measured to the nearest centimeter. What are the best upper and lower bounds for the volume of this parallelepiped? What are the best upper and lower bounds for the surface area?

11. Suppose two points (x_0, y_0) and (x_1, y_1) are on a straight line with $y_1 \neq y_0$. Two formulas are available to find the x-intercept of the line:

$$x = \frac{x_0 y_1 - x_1 y_0}{y_1 - y_0}, \text{ and } x = x_0 - \frac{(x_1 - x_0)y_0}{y_1 - y_0}.$$

Use the data $(x_0, y_0) = (1.31, 3.24)$ and $(x_1, y_1) = (1.93, 4.76)$ and three-digit rounding arithmetic to compute the x-intercept both ways. Which method is better and why?

CONTINUED

12. a. How many multiplications and additions are required to determine a sum of the form

$$\sum_{i=1}^{n}\sum_{j=1}^{i} a_i\, b_j \ ?$$

b. Modify the sum in part (a) to an equivalent form that reduces the number of computations.

13. Let $P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial, and let x_0 be given. Construct an algorithm to evaluate $P(x_0)$ using nested multiplication.

14. Construct an algorithm that has as input an integer $n \geq 1$, numbers $x_0, x_1, \cdots, x_n$, and a number x and that produces as output the product $(x - x_0)(x - x_1)\cdots(x_0 - x_n)$.

15. a. Consider the stability (by calculating the condition number) of $\sqrt{1+x} - 1$ when x is near 0. Rewrite the expression to rid it of subtractive cancellation.

b. Rewrite $e^x - \cos x$ to be stable when x is near 0.

16. Suppose that a function $f(x) = \ln(x+1) - \ln(x)$, is computed by the following algorithm for large values of x using six digit rounding arithmetic

$$\begin{aligned} x_0 : &= x = 12345 \\ x_1 : &= x_0 + 1 \\ x_2 : &= \ln x_1 \\ x_3 : &= \ln x_0 \\ f(x) := x_4 : &= x_2 - x_3. \end{aligned}$$

By considering the condition $\kappa(x_3)$ of the subproblem of evaluating the function, show that such a function evaluation is not stable. Also propose the modification of function evaluation so that algorithm will become stable.