Lab Session 2

MA423 : Matrix Computations Lab — R. Alam — 2023

Finite precision arithmetic

1. The purpose of this exercise is to illustrate anomaly in automatic computation. On my computer, MATLAB produces

$$\begin{aligned}
(\frac{4}{3}-1)*3-1 &= -2.2204\times 10^{-16}\\
5\times\frac{(1+\exp(-50))-1}{(1+\exp(-50))-1} &= \textbf{NaN}\\
\frac{\log(\exp(750))}{100} &= \textbf{Inf}
\end{aligned}$$

   Try on your machine. Can you explain the reason behind these anomalies?

2. The machine epsilon `eps` of a floating point system is the distance from 1 to the next floating number bigger than 1 and `u = eps/2` is the unit roundoff (default). You can compute `eps` and `u` in MATLAB by writing a small script. What is it that the following MATLAB script computes?

```
x = 2;
while x > 1
x = x/2
end
```

   On the other hand, if the condition $x > 1$ in the above script is replaced with $1 + x > 1$ then what will be the output?

3. Consider $(\beta, t, e_{\min}, e_{\max}) = (10, 8, -99, 99)$ and the normalized floating-point numbers

$$\begin{aligned}
x &= 0.23371258\times 10^{-4}\\
y &= 0.33678429\times 10^{2}\\
z &= -0.33677811\times 10^{2}
\end{aligned}$$

   For arithmetic with $t = 8$, use MATLAB command `round(x, 8)` to calculate $\mathrm{fl}(x+\mathrm{fl}(y+z))$ and $\mathrm{fl}(\mathrm{fl}(x+y)+z)$. Is $\mathrm{fl}(x+\mathrm{fl}(y+z)) = \mathrm{fl}(\mathrm{fl}(x+y)+z)$?

   Next, calculate $x + y + z$ in MATLAB and determine the relative errors in calculating $\mathrm{fl}(x+\mathrm{fl}(y+z))$ and $\mathrm{fl}(\mathrm{fl}(x+y)+z)$.

4. A person has a nightmare in which the following equation flies past him:

$$1782^{12} + 1841^{12} = 1922^{12}. \tag{1}$$

   Note that this equation, if true, would contradict Fermat's last theorem, which states: For $n \geq 3$, there do not exist any natural numbers x, y and z that satisfy the equation $x^n + y^n = z^n$. Did the person dream up a counterexample to Fermat's last theorem?

(a) Compute $(1782^{12} + 1841^{12})^{1/12}$ by typing the following into MATLAB:

```
>> format short
>> (1782^12 + 1841^12)^(1/12)
```

What result does MATLAB report? Now look at the answer using `format long.`

(b) Determine the absolute and relative error in the approximation

$$1782^{12} + 1841^{12} \simeq 1922^{12}.$$

Such an example is called a Fermat near miss because of the small relative error. Note that the right-hand side of equation (1) is even, which can be used to prove that the equation cannot be true.

(c) Later the person writes the equation

$$3987^{12} + 4365^{12} = 4472^{12}.$$

Can you debunk this equation?

5. The roots of a quadratic polynomial $p(x) := ax^2+bx+c$ is given by $(-b \pm \sqrt{b^2 - 4ac})/2a$. Write a MATLAB function that implements the above formula to compute the roots. Your function will look like this:

```
function [x1, x2] = quadroot1( a, b, c)
d = sqrt( b^2 - 4 * a* c );
x1 = (-b + d) / (2*a);
x2 = (-b - d) / (2*a);
```

The largest (in magnitude) root of p can be computed as $x_1 = (-b - \mathrm{sign}(b)\sqrt{b^2 - 4ac})/2a$ and the second root x_2 can be computed from the identity $x_1 x_2 = c/a$. Write a MATLAB function that implements the modified method to compute the roots. Your function will look like this:

```
function [x1, x2] = quadroot2(a, b, c)
d = sqrt( b^2 - 4 * a* c );
x1 = (-b - sign(b) * d) / (2*a);
x2 = c/( a * x1 );
```

Find the roots of $x^2 - (10^7 + 10^{-7})x + 1)$ using `quadroot1` and `quadroot2`. Do you observe any difference? Which method is better and why?

6. Consider the power series for $\sin x$ given by

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

Here is a Matlab function that uses this series to compute $\sin x$.

```
function s = powersin(x)
%
% POWERSIN(x) tries to comptue sin(x) from a power series
```

```
%
  s = 0;
  t = x;
  n = 1;
  while s + t ~= s;
   s = s + t;
   t = -x.^2/((n+1)*(n+2)).*t;
   n = n + 2;
  end
```

When does the while loop terminate? Answer each of the following questions for $x = \pi/2, 11\pi/2, 21\pi/2$ and $31\pi/2$.

- How accurate is the computed result?
- How many terms are required?
- What is the largest term in the series?

What is your conclusion about the use of floating point arithmetic and power series to evaluate functions?

7. Show how to rewrite the following expressions to avoid cancellation for the indicated arguments. Evaluate these expressions using 5-digits decimal system (use MATLAB command `round`).

   (a) $\sqrt{x+1} - \sqrt{x}$; $x \approx 0$.

   (b) $\sin x - \sin y; x \approx y$.

   (c) $(1 - \cos x)/\sin x$; $x \approx 0$.

8. Consider the function $f(x) = (e^x - 1)/x$, which arises in various applications. By L'Hopital's rule, we know that $\lim_{x\to 0} f(x) = 1$.

   (a) Compute the values of $f(x)$ for $x = 10^{-n}$ for $n = 1, 2, \ldots, 16$. Do your results agree with theoretical expectations? Explain why.

   (b) Now perform the computation in part (a) again, this time using the mathematically equivalent formulation $f(x) = (e^x - 1)/\log(e^x)$ (evaluate as indicated without simplification). If this works any better, can you explain why?

9. Consider the recurrence $x_{k+1} = 111 - (1130 - 3000/x_{k-1})/x_k,\ x_0 = 11/2,\ x_1 = 61/11$. In exact arithmetic the x_k form a monotonically increasing sequence that converges to 6. Implement the recurrence in MATLAB and compare the computed x_{34} with the true value 5.998 (correct to four digits). Try to explain what you see.

10. Find the smallest value of p for which the expression calculated in double precision arithmetic at $x = 10^{-p}$ has no correct significant digits (no correct digits in the mantissa). (Hint: First find the limit of the expression as $x \to 0$.)

    (a) $\dfrac{\tan x - x}{x^3}$

    (b) $\dfrac{e^x + \cos x - \sin x - 2}{x^3}$.

11. This exercise illustrates the difficulty in handling polynomials in finite precision computation. Consider the polynomial

$$\begin{aligned} p(x) = (x-2)^9 &= x^9 - 18x^8 + 144x^7 - 672x^6 + 2016x^5 \\ &\quad -4032x^4 + 5376x^3 - 4608x^2 + 2304x - 512 \end{aligned}$$

Write a MATLAB script to evaluate p at 151 equidistant points (use `linspace` command) in the interval [1.95, 2.05] using two methods:

(a) Apply Horner's method, or call MATLAB function `polyval` (Type `help polyval` for more info).

(b) Calculate $p(x) = (x-2)^9$ directly.

Plot these results in two separate figures. For example, if x is a row vector of points in the given interval then the commands

```
>> y = p(x); plot(x, y)
```

will do the job. Do the plots differ from one another? If yes, can you think of possible reasons?

12. **Assignment:** A banker creates a program that takes fractions of cents that are truncated in a bank's transactions and deposits them to his own account. This is not a new idea, and hackers who have actually attempted it have been arrested. In this exercise we will simulate the program to determine how long it would take to become a millionaire this way.

Assume that we have access to 50,000 bank accounts. Initially we can take the account balances to be uniformly distributed between, say, \$100 and \$100,000. The annual interest rate on the accounts is 5%, and interest is compounded daily and added to the accounts, except that fractions of a cent are truncated. These will be deposited to an illegal account that initially has balance \$0.

Write a MATLAB program that simulates the Office Space scenario. You can set up the initial accounts with the commands

```
accounts = 100 + (100000-100) * rand(50000,1);
% Sets up 50,000 accounts with balances between $100 and $100000.
accounts = floor(100 * accounts)/100;
% Deletes fractions of a cent from initial balances.
```

(a) Write a MATLAB program that increases the accounts by (5/365)% interest each day, truncating each account to the nearest penny and placing the truncated amount into an account, which we will call the illegal account. Assume that the illegal account can hold fractional amounts (i.e., do not truncate this account's values) and let the illegal account also accrue daily interest. Run your code to determine how many days it would take to become a millionaire assuming the illegal account begins with a balance of zero.

(b) Without running your MATLAB code, answer the following questions: On average, about how much money would you expect to be added to the illegal account each day due to the embezzlement? Suppose you had access to 100,000 accounts, each initially with a balance of, say, \$5000. About how much money would be added to the illegal account each day in this case? Explain your answers.

Note that this type of rounding corresponds to fixed-point truncation rather than floating-point, since only two places are allowed to the right of the decimal point, regardless of how many or few decimal digits appear to the left of the decimal point.

10 marks

*** End ***