APPM 4600: Floating Point and Algorithms & Stability

Friday, August 30, 2024 10:22 AM

Class 03: August 30, 2024

Recall: We defined important concepts having to do with floating point number arithmetic and systems. These include **double precision** (Float 64 bit) and **single precision** (Float 32 bit) numbers, how to find the largest and smallest number that can be represented, the machine epsilon and its relationship to how many decimal digits I can trust. We began a discussion on **loss of precision**, and what operations may incur in serious **loss of significant digits of relative accuracy**.

WARM UP: Doing 4 digit arithmetic

1. $128.5 + 24.43 = ?$

$$\begin{array}{r} {\scriptstyle 1}\\ 128.5\\ +\ 24.43\\ \hline 152.93 \end{array}$$

"True": 152.93

4 digit: 152.9

$err_{abs} = 0.03$

$err_{rel} = 0.03/152.93$

$1.96e\text{-}4$

2. $128.5 - 127.2 = ?$

1.3

True 1.3

4 digit 1.300...0

$err_{abs} \leq 0.1$

$err_{rel} \leq 0.1/13 \approx \_ \times 10^{-1}$

We concluded that:

$+, \times, \div$ → do not introduce loss of precision

$-$ → loss of prec when subtracting two numbers close to each other!

QUESTION: Does this show up in actual / relevant math problems?

/ relevant math problems?

A: Yes. We will see it throughout the course. Here are some examples:

E1: Quadratic Equation

PROBLEM: Find the roots of

$$0.2x^2 - 47.91x + 6 = 0$$

PROPOSED METHOD - quadratic formula

w/ $a = 0.2$, $b = -47.91$, $c = 6.0$

$$r = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

(4 digits)

$$r = \frac{47.91 \pm \sqrt{2295.3681 - 4.8}}{0.4}$$

$$r = \frac{47.91 \pm 47.85}{0.4}$$

$r_1 = 239.4$ $r_2 = 0.15$

TRUE: $r_1 = 239.4$ ✓ $r_2 = 0.1253$ ✗

RELERR $O(10^{-4})$ 0.1971 ✗

$$r_{bad} = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \left[\frac{-b + \sqrt{b^2 - 4ac}}{-b + \sqrt{b^2 - 4ac}} \right]$$

$$= \frac{b^2 - (b^2 - 4ac)}{2a(-b + \sqrt{b^2 - 4ac})} = \frac{4ac}{2a(-b + \sqrt{\quad})}$$

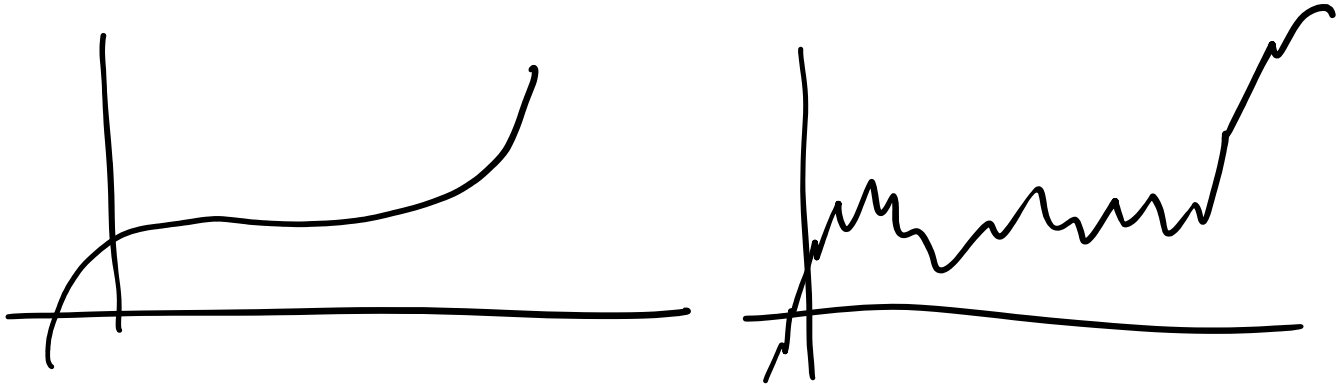
$$= \frac{2c}{-b + \sqrt{b^2 - 4ac}}$$

$$r = \frac{2(6)}{47.91 + 47.85} = 0.1253 \checkmark$$

Polynomial eval:

$$p(x) = (x-2)^9 \quad \text{for } x = 2.01$$

$$p(x) = x^9 - 18x^8 + 144x^7 - \ldots$$



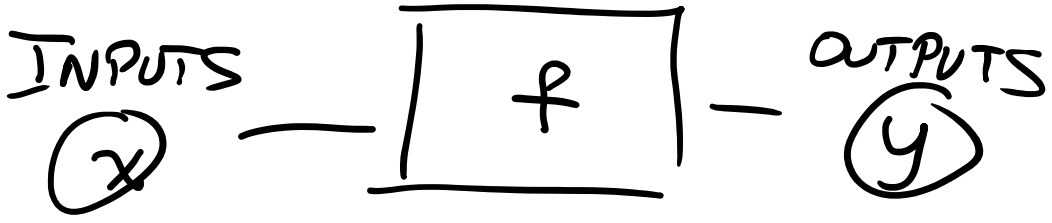
Algorithms & Stability

Sequence of instructions to be carried out in order.

"Pseudocode" → Inputs, Outputs, Detailed list of instructions.

Solve system of lin. eqs $A\vec{x} = \vec{b}$

Solve system of nonlin. eqs $F(\vec{x}) = \vec{0}$



MATH PROBLEM