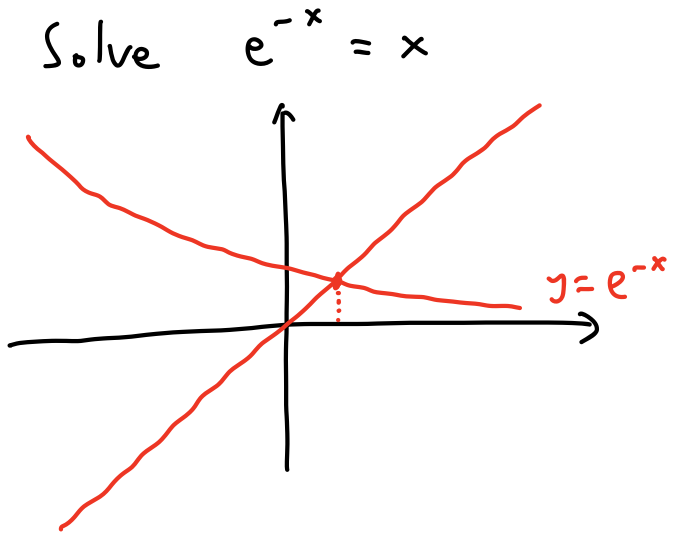
Math 8500 : Advanced numerical analysis I

- Why do we need numerical methods?

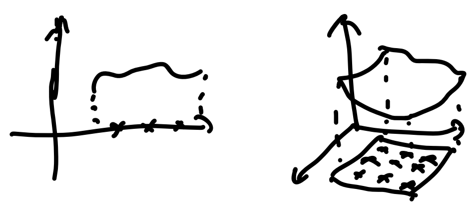
<u>Ex</u> Solve $e^{-x} = x$



no explicit sol'n

⇝ need to do approximation of the sol'n.

- How do we evaluate a numerical method?
  - Accuracy / cost
  - Stability / robustness
  - Regularity / information required.
  - High dimension efficiency.
  - good for parallel computing?
  - easy to implement



Round-off error

- normalized scientific notation (decimal system)

$$235.4127 = \underbrace{0.2354127}_{\text{mantissa}} \times 10^{\overset{\downarrow}{3}} \quad \text{exponent}$$

generally, for any $x \in \mathbb{R}$, $x \neq 0$,

$$x = \pm r \times 10^n \quad \text{where} \quad \frac{1}{10} \le r < 1,\ n \in \mathbb{Z}$$

- normalized scientific notation (binary system)

$$x = \pm q \times 2^m \quad \text{where} \quad \frac{1}{2} \le q < 1,\ m \in \mathbb{Z}$$

$$\underbrace{0.100101}_{\text{as binary}} \times 2^{-4} = (2^{-1} + 2^{-4} + 2^{-6}) \times 2^{-4}$$

In computer, we need to store $+$, 100101, -100

- 64-bit system

sign	exponent	mantissa
1	11	52

exponent → largest possible exp: $2^{10} - 1 = 1023$

$\Rightarrow$ largest possible number: $2^{1023} \approx 10^{308}$

trying to get a larger number leads to overflow

"NaN"

mantissa → 52 binary effective digits

$2^{52} = 10^n$

$\rightarrow$

16 decimal eff. digits.

"machine epsilon"

$\varepsilon \approx 10^{-16}$

- Why do we need round-off?

Say, using decimal system w/ 3 eff. digits

$0.563 + 0.611 = 1.174 \xrightarrow{\text{round-off}} 1.17$ (closest expressible number to 1.174)

$8.50 \times 0.850 = 7.225 \xrightarrow{\text{round-off}} 7.23$

- Be careful when subtracting two numbers which are close.

$$0.563 - 0.562 = 0.001$$

Say, given $f(x)$, want $f'(x_0) \approx \dfrac{f(x_0+h) - f(x_0)}{h}$ for small h.

For $f(x) = x^2$, $x_0 = 1$

$h = 0.1$ $\quad \dfrac{1.1^2 - 1^2}{0.1} = \dfrac{1.21 - 1}{0.1} = \dfrac{0.21}{0.1} = 2.1$

$h = 0.01$ $\quad \dfrac{1.01^2 - 1^2}{0.01} = \dfrac{1.0201 - 1}{0.01} \sim \dfrac{1.02 - 1}{0.01} = \dfrac{0.02}{0.01} = 2$

$h = 0.001$ $\quad \dfrac{1.001^2 - 1^2}{0.001} \sim \dfrac{1^2 - 1^2}{0.001} = 0$

choose $h \gg$ machine ε.

- Be careful when adding a small number to a large number

$$2.23 + 0.0123 = 2.2423 \to 2.24$$

$$2.23 + 0.0131 = 2.2431 \to 2.24$$