

CPSC 303 Feb 7, 2024

- Recurrence Relations and Finite Precision (Section 5)
- Normal & Subnormal Numbers in Double Precision

Rem: [A&G] only covers "normal numbers"

Next Up! Chapter 10 [A&G]
or Interpolation (Polynomial)

Say

$$(\sigma - 1)(\sigma - 5)(x_n) = 0$$

$$(\sigma^2 - (1+5)\sigma + 5)(x_n) = 0$$

S likely; $0 < S < 1$

General solution

$$x_n = c_1 \cdot 1^n + c_2 S^n$$

$$= c_1 + c_2 S^n$$

Claim: This can have unexpected results

So for why c_1 , $0 < |s| < 1$

$$\lim_{n \rightarrow \infty} (c_1 + c_2 s^n)$$

$$= c_1 + c_2 \lim_{n \rightarrow \infty} s^n$$

$$= c_1$$

Strange behaviour:

$$c_1 = 0, \quad c_2 = 1, \quad x_n = s^n$$

$$x_0 = s^0 = 1, \quad x_1 = s^1 = s,$$

$$x_2 = s^2, \quad x_3 = s^3, \quad \dots$$

$$S = \frac{1}{2},$$

$$(\sigma - 1)\left(\sigma - \frac{1}{2}\right)$$

$$= \sigma^2 - \frac{3}{2}\sigma + \frac{1}{2}$$

$$\left(\sigma^2 - \frac{3}{2}\sigma + \frac{1}{2}\right) (\{x_n\}) = 0$$

$$x_{n+2} - \frac{3}{2}x_{n+1} + \frac{1}{2}x_n = 0$$

$$x_{n+2} = \frac{3}{2}x_{n+1} - \frac{1}{2}x_n$$

Rem: For MATLAB,

$x \leftarrow$ array $x \leftarrow$ cell

$$x_1 \text{ or } x\{1\} = 1$$

$$x_2 \text{ or } x\{2\} = 1/2$$

$$x_n = (1/2)^{n-1} \text{ in exact arithmetic}$$

Useful MATLAB

format long ← give more digits

format short ← gives fewer

Rem: Look at MATLAB's

x_n -- n large, cyclic, non-zero storage numbers

Scy $F_0 = \text{given}, F_1 = \text{given}$

MATLAB $F_1 = \text{given}, F_2 = \text{given}$

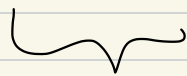
$$F_{n+2} = F_{n+1} + F_n$$

If $F_1 = 1, F_2 = \frac{1-\sqrt{5}}{2}$

sgen get

$$F_n = \left(\frac{1-\sqrt{5}}{2} \right)^{n-1}$$

similarly ~



- 0.6180... ~

$$f_n / (F_2)^{n-1} = 1$$

Double precision!

A real double precision number
is

$$\pm 1.b_1 \dots b_{52} \times 2^m$$

in binary
scientific: $_ . _ _ _ \times 2^{35}$
 \uparrow
has to be 1

$$-1022 \leq m \leq 1023$$

$$\text{So } m = -1022, -1021, \dots, 1023$$

possible values is

$$(1023) - (-1022) + 1$$

$$= 2046$$

There's 2 values left if

m is to have 11 bits

$$2^{11} = 2048$$

\pm } 1 bit

$b_1 - - b_{52}$ } 52 bits

m } at most 2048 values

11 bits

$$\text{Total} = 1 + 52 + 11 = 64$$

(One extra value of m is for
Inf, -Inf, NaN
↑ ↗
infinity Not a Number)

The other value of m

Note: in MATLAB:

$$2^{1023} \text{ is } 8.988 \times 10^{307}$$

$$2^{1024} \text{ is } \text{Inf}$$

but

$$(1/2)^{1074} = 4.94 \times 10^{-324}$$

$$(1/2)^{1075} = 0$$

In IEEE Double precision:

last value of m :

$$2^{-1022}$$

but no

$$2^{-1023}$$

It allow:

$$\pm 0.b_1b_2 \dots b_{52} 2^{-1022}$$

Friday Start Interpolation