Computer Arithmetics

β, N, L, U

$$x = \pm \left(d_1 + d_2/\beta + \dots + d_N/\beta^{N-1}\right) \cdot \beta^E$$

E, d_i is integer

$$0 \le d_i \le \beta - 1$$

$$L \le E \le U$$

Finite # of digits

$$\left| \frac{x - Fl(x)}{x} \right| \le \frac{\epsilon_{mach}}{2}$$ normalized #

$$\epsilon_{mach} = \beta^{1-N}$$

$$f(x) = \frac{1-(1-x)}{x}$$

$f(\epsilon_{mach} \cdot 0.99)$

$f(\epsilon_{mach} \cdot 1.01)$

+

- loose digits

*

/ -11 a lot of 1-1-1-1

integers

vpa

sym

C

are represented

uniquely

$13 = 1.3 \cdot 10 = 0.13 \cdot 10\not{0}$

agreement $d_1 \neq \not{0}$

Binaries $d_1 = 1$

Subnormals

d_1 may be $\not{0}$ if $E = L$

$N = 3, \beta = ?, L = -1, U = 1$

$\pm$ 1.00
1.01
1.10
1.11 $* \not{1}0^{-1, \not{0}, 1}$

$U - L = 2$

D

Subnormals

d_1 may be $\emptyset$

IF $E = L$

0.01

$0.10 * 10^{-1}$

0.11

Subnormals

$1 \cdot 10^{-1}$ - smallest normalized

$0.01 \cdot 10^{-1}$

What is smallest #

normalized #

realmin

$X = 1.00 \cdot 10^{L}$

$X = \beta^{L}$

smallest subnormal

$y = \beta^{L} \cdot \beta^{1-N}$

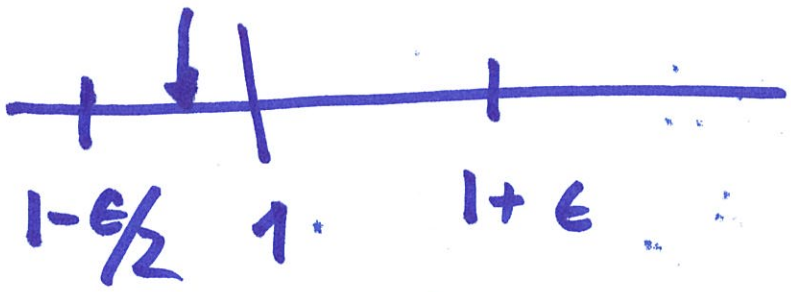
~~F++~~

Smallest possible that can not be representa

$\underbrace{1111}_{54} \approx \underbrace{1000}_{53}$

$$f(x) = \frac{1 - (1 - x)}{x}$$

- $x_1 = \epsilon_{mach} \cdot 0.99$



$$1 - x \approx 1; \quad f(x_1) = 0$$

- $x_2 = \epsilon_{mach} \cdot 1.01$

$$1 - x_2 \approx 1 - \epsilon_{mach}/2$$

$$f(x_2) = \frac{1 - (1 - \epsilon_{mach}/2)}{\epsilon_{mach} \cdot 1.01} \approx 2$$

G

double precision

8 Bytes = 64 Bits

N = 53 ← normalized $d_1 = 1, 52$

E = 11 Bits

1 Bit for sign

52 + 11 + 1

$\emptyset .. 2047$, (?)

-1023 ;

I

Total # of #

$$2 \cdot 3^{N-1} \times (U - L + 1) + 1$$

2 ← sign

3 ← mantissa

$(U - L + 1)$ ← exponent

1 ← zero

Finding root of

$f(x) = 0$

- Bisection $r = 1$
- Newton $r = 2$
- Secant $1 < r < 2$

"golden ratio"

$$\epsilon_{k+1} = C\,\epsilon_k^r$$

ϵ_{k+1}: error at step $k+1$

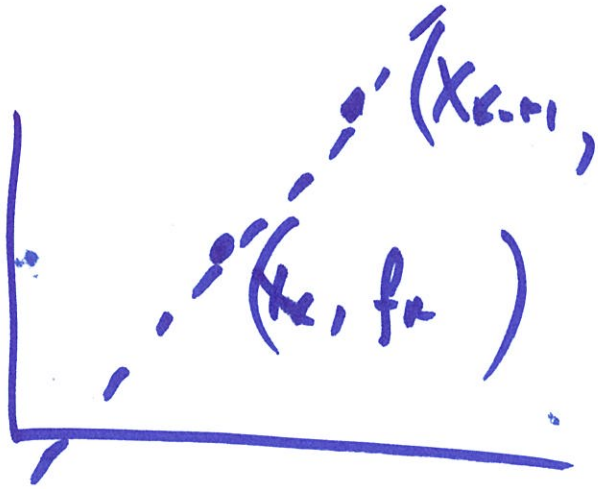
ϵ_k: error at step k

$m = \frac{a+b}{2}$; IF $f(a)f(m) < 0$, $b = m$ else $a = m$

END IF

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

$$x_{k+1} = x_k - \frac{f(x_k)\,(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}$$



1.12 $\varepsilon_1 = 0.12$ Newton

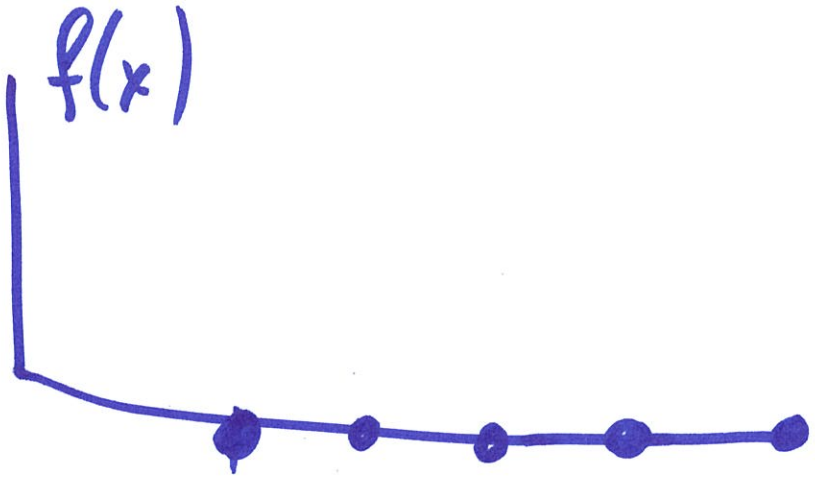
1.01 $\varepsilon_2 = 0.01$ $f'(x^*) \neq \emptyset$

1.00 $\varepsilon_3 = \emptyset$

1.00 ←

1.16
1.08
1.04
1.02
1.01
1.005

$\epsilon_k = \epsilon_{k-1}/2$



$f(x) = x^2$

$f'(x) = 2x$

$\epsilon_{k+1} = \epsilon_k/2$

$x_{k+1} = x_k - \frac{x_k^2}{2x_k} = \frac{x_k}{2}$

$f(x)$

$f'(x) = \emptyset$

M

LU decomposition

↳ upper

↳ lower

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

M_1 eliminate Below a_{11}

$$M_1 A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cancel{0} & a'_{22} & \cdots & a'_{1n} \\ \cancel{0} & \vdots & & \\ \vdots & \vdots & & \\ \cancel{0} & a'_{n2} & \cdots & a'_{nn} \end{pmatrix}$$

M_2

↳ "eliminates" a'_{22}

$$M_2 M_1 A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cancel{0} & a'_{22} & \cdots & a'_{2n} \\ \cancel{0} & \cancel{0} & a''_{33} & \cdots & a''_{3n} \\ & & & & \end{pmatrix}$$

$(AB)^{-1} = B^{-1}A^{-1}$ ✓

$U = (M_{n-1} \dots M_3 M_2 M_1) A$

$M = M_{n-1} M_{n-2} \dots M_3 M_2 M_1$

$M^{-1}(U = MA)$

$A = M^{-1}U$

$M^{-1} = (M_{n-1} M_{n-2} \dots M_3 M_2 M_1)^{-1}$

$= \underbrace{M_1^{-1} M_2^{-1} M_3^{-1}} = L$

$\boxed{A = LU}$

$$\underline{\underline{A}}\,\underline{x} = \underline{b}$$

$$A = L\,U$$

$$\underline{\underline{L}}\,(\underline{\underline{U}}\,\underline{x}) = \underline{b}$$

$$\underline{z}$$

$$\underline{\underline{L}}\,\underline{z} = \underline{b} \Leftarrow \text{Forward substitution}$$

$$\underline{\underline{U}}\,\underline{x} = \underline{z} \Leftarrow \text{Backward substitution}$$

Example

$$\begin{pmatrix} 3 & 4 \\ 6 & 9 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 11 \\ 24 \end{pmatrix}$$

A

$$\begin{pmatrix} 3 & 4 \\ 6 & 9 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 11 \\ 24 \end{pmatrix}$$

$$M=\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 3 & 4 \\ 6 & 9 \end{pmatrix}=\begin{pmatrix} 3 & 4 \\ 0 & 1 \end{pmatrix}$$

M

$$MA = U$$

$$\begin{pmatrix} 3 & 4 \\ 6 & 9 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 3 & 4 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 4 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 11 \\ 2 \end{pmatrix}$$

$$x_2 = 2$$

$$3x_1 + 4x_2 = 11$$

$$x_1 = \frac{1}{3}(11 - 4x_2) = \frac{1}{3}(11 - 8) = 1$$

$$x_1 = 1$$
$$x_2 = 2 \qquad \underline{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

9

$$\begin{pmatrix}1 & 0\\ 2 & 1\end{pmatrix}\underbrace{\begin{pmatrix}3 & 4\\ 0 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}}_{z}=\begin{pmatrix}11\\ 24\end{pmatrix}$$

$$\begin{pmatrix}1 & 0\\ 2 & 1\end{pmatrix}\begin{pmatrix}z_1\\ z_2\end{pmatrix}=\begin{pmatrix}11\\ 24\end{pmatrix}$$

$z_1 = 11$

$z_2 = 24 - 2z_1 =$

$= 24 - 22 = 2$

$$z=\begin{pmatrix}11\\ 2\end{pmatrix}$$

$$\begin{pmatrix}1 & 0\\ 2 & 1\end{pmatrix}\begin{pmatrix}3 & 4\\ 0 & 1\end{pmatrix}\begin{pmatrix}z_1\\ z_2\end{pmatrix}=\begin{pmatrix}11\\ 2\end{pmatrix}$$

$$\begin{pmatrix}3 & 4\\ 0 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}11\\ 2\end{pmatrix}$$

$x_2 = 2;\quad x_1 = \frac{1}{3}(11 - 4x_2) =$

$= \frac{1}{3}(11 - 8) = 1$

Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 8 \\ 3 & 10 & 20 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 10 \\ 24 \\ 49 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 8 \\ 3 & 10 & 20 \end{pmatrix}$$

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}$$

$-3/1$

$-3/1$

$$M_1 A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 8 \\ 3 & 10 & 20 \end{pmatrix}$$

$$= \left(\begin{array}{c|cc} 1 & 2 & 3 \\ \hline 0 & 1 & 2 \\ \hline 0 & 4 & 11 \end{array}\right) ; \quad M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{pmatrix}$$

$$M_2 M_1 A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 4 & 11 \end{pmatrix} S$$

$$= \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix} = U$$

$$A = (M_2 M_1)^{-1} U = M_1^{-1} M_2^{-1} U$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 8 \\ 3 & 10 & 20 \end{pmatrix}$$