NM 02

- Real Number (Floating-point form)

$$x = \pm(.b_1 b_2 \cdots b_k) \times \beta^{e}$$

$\rightarrow$ Exponent

Mantissa

base

Sign	Exponent	Mantissa
1bit	7bits	24bits
0 양수		
1 음수		

- example 2.1

1 0000010 1101000……

$\downarrow$ $\downarrow$

$-$ 2 $\longrightarrow 11.01_2 \rightarrow 3.25_{10}$

- example 2.2

42.78125_{10}

$\downarrow$

101010 $0.78125 = 0.5 + 0.25 + 0.03125$

1 1 00 1

$101010.11001_2 = 0.10101011001 \times 2^6$

$+$ 2^6 4278125

0 0000110 10101011001 00000……

- 표준 $(-1)^S \times (1.f)_{2진수} \times (2^{c-127})_{10진수}$

S	C	f
1bit	8bit	23bit

$+\infty$	0	11111111	000 · · ·
$-\infty$	1	11111111	000 · · ·
$+0$	0	00000000	000 · · ·
-0	1	00000000	000 · · ·

→ 뒤에가 non-zero 면 NaN

- example 2.3

-45.8125

↓ ↘

-45 0.8125

-101101_2 0.1101 → $-101101.1101 = -1.011011101 \times 2^5$

- $127+5=132$

1 10000100 011011101 · · · ·

- IEEE 에서 표현할 수 있는

가장 큰 수 0 11111110 1111 · · · · · $1.1111 \cdots \times 2^{127} = (2-2^{23}) \times 2^{127}$

양수중 가장 작은 수 0 00000001 0000 · · · · 1.0×2^{-126}

half precision	1 sign	5 exponent	10 fraction
single precision	1 sign	8 exponent	23 fraction
double precision	1 sign	11 exponent	52 fraction

- round off error

$fl(x)$ ← round off 한 결과

chopping 버림

rounding 반올림

- absolute error $|x - fl(x)|$

relative error $\dfrac{|x - fl(x)|}{|x|}$

chopping의 relative error : $\dfrac{|0.b_1b_2\cdots b_k\cdots \times 10^n - 0.b_1b_2\cdots b_k \times 10^n|}{|0.b_1b_2\cdots b_k\cdots \times 10^n|}$

$= \dfrac{|0.b_{k+1}b_{k+2}\cdots|}{|0.b_1b_2\cdots b_k\cdots|} \times 10^{-k} \leq \dfrac{1}{0.1} \times 10^{-k}$ 약…

$\therefore \left|\dfrac{x - fl(x)}{x}\right| \leq 10^{-k+1}$

rounding의 relative error : $\left|\dfrac{x - fl(x)}{x}\right| \leq 5 \times 10^{-k}$

따라서 약 2^{-k}의 rounding, 2^{-k+1}의 chopping relative error가 있다.

- 유효숫자 Significant digits n 까지 정확하다고 할 수 있다.
  leading zero → 안 셈
  Tailing zero → 셈

rounding 0.0045678 to 3 sigifcant digits → 0.00457

- x^* is approximation x with k significant digits.

$$\left|\frac{x-x^*}{x}\right| \le 5\times10^{-k}$$

- 테일러 급수

$$f(x) \sim f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots\cdots$$

$$f(x) \sim \sum_{k=0}^{\infty} \frac{f^k(c)}{k!}(x-c)^k$$ (c=0 이면 매클로린 급수)

x와 c는 충분히 가까워야 한다.

- example $\ln(x)$

$\ln(1.1)$을 테일러급수 5번으로 구해보자 기준점 c를 0으로두고 . . .

$f'(0) = \frac{1}{1+x} = 1$, $f''(0) = -\frac{1}{(1+x)^2} = -1$

$f'''(0) = \frac{2}{(1+x)^3} = 2$, $f''''(0) = -\frac{6}{(1+x)^4} = -6$

$\therefore \ln(1+x) \sim x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5}$ x에 0.1을 대입하면 끝!

$(-1 < x \leq 1)$

시험

- example $3x^5 - 2x^4 + 15x^3 + 13x^2 - 12x - 5$ $c=2$

$\downarrow$

$f(2) = 207$

$f'(2) = 396$

$\vdots$

$f^{6}(2) = 0$

$\vdots$

$f(x) = 207 + 396(x-2) + 295(x-2)^2 \cdots + 3(x-2)^5$

- example e^x

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad (|x| < \infty)$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \quad (|x| < \infty)$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \quad (|x| < \infty)$$

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 \cdots = \sum_{k=0}^{\infty} x^k \quad (|x| < 1)$$

- $f(x) = \sum_{k=0}^{n} \frac{f^{k}(x_0)}{k!}(x-x_0)^k + R_{n+1}(x)$ 로 나타낼 수 있다.

$$R_{n+1}(x) = \frac{f^{n+1}(\xi)}{(n+1)!}(x-x_0)^{n+1} \quad \text{for } x_0 < \xi < x$$

: remainder 또는 truncation error

매클로린 : $f(x) = \sum_{k=0}^{n} \frac{f^{k}(0)}{k!} x^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1}$ for $0 < \xi < x$

- example 테일러 급수의 error e^x near $x_0 = 0$

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \underbrace{\frac{e^\xi}{(n+1)!} x^{n+1}}_{R_{n+1}(x)}$$

$$= \sum_{k=0}^{n} \frac{x^k}{k!} + \frac{e^\xi}{(n+1)!} x^{n+1}$$

e^1을 $n=3$으로 계산했을 때 error

$R_4 = \frac{e^\xi}{4!}$ $(0 < \xi < 1)$

$|e^\xi| \le e \le 2.8$ 로 가정하면 R_4는 약 $\frac{1}{4!} \times 2.8 = 0.11667$

시험

- example e를 소숫점 3자리까지 근사하라 : 2.718

↓

유효숫자는 4자리 이어야 함.

$\left|\frac{x - x^*}{x}\right| \le 5 \times 10^{-k}$, $|e^\xi| \le e \le 2.8$ 활용

$$\underbrace{\frac{e^{\xi}}{(n+1)!}}_{\text{오차}} < \frac{2.8}{(n+1)!} < 5\times10^{-4}$$

$n \geq 7$ 이상이면 만족

$x - x_0$를 h (거리)로 놓으면

$$f(x+h) = \sum_{k=0}^{n} \frac{f^{k}(x)}{k!} h^{k} + \underbrace{\frac{f^{n+1}(\xi)}{(n+1)!} h^{n+1}} \quad \text{for } x < \xi < x+h$$

error $F(h) = O(h^{n})$ ($C \cdot h^{n}$ 보다 작거나 같다)

$n+1$일 때, $F(h) = O(h^{n+1})$

$$\therefore f(x+h) = f(x) + O(h)$$

$$= f(x) + f'(x)h + O(h^{2})$$

$$= f(x) + f'(x)h + \frac{1}{2!} f''(x) h^{2} + O(h^{3})$$