計算機による数値表現と誤差

1. 計算機による数値表現

浮動小数点数 (floating point number)

$$-247.72 = -0.24732 \times 10^{-3}$$

$$\text{Avogadro数} \fallingdotseq \underbrace{0.60221409}_{\text{有限桁}} \times 10^{24}$$

→ コンピュータでは 2進数(16進数)で表現

一般形: $\pm f \times \beta^{m}$ (β 進 n 桁浮動小数点数)

β : 基数 (base)　コンピュータでは 2 or 16

f : 仮数部 (mantissa)

$$f = \frac{c_1}{\beta} + \frac{c_2}{\beta^2} + \cdots + \frac{c_n}{\beta^n}$$

$$c_1 = 1, \cdots, \beta-1, \quad c_2, \cdots, c_n = 0, 1, \cdots, \beta-1$$

m : 指数部 (exponent)

$$-m_L \leq m \leq m_U$$

$\beta = 2$, プリントの例では 7桁使うから (ただし先頭1桁は符号±1)

$$m_L = 2^6 = 64, \quad m_U = 2^6 - 1 = 63$$

(* 実際には m の代わりに $m+64$ をコンピュータ内に記録することが多い)

プリント例では, 全体で 1語32ビット … 単精度浮動小数点数

1語64ビット … 倍精度浮動小数点数

浮動小数点数の絶対値の

最大値: $\left(\frac{\beta-1}{\beta} + \frac{\beta-1}{\beta^2} + \cdots + \frac{\beta-1}{\beta^n}\right) \times \beta^{m_U} = (\beta-1) \times \beta^{-1} \frac{(1-\beta^{-n})}{1-\beta^{-1}} \times \beta^{m_U}$

$= (1-\beta^{-n})\beta^{m_U}$ —— ①

最小値: $\left(\frac{1}{\beta} + \frac{0}{\beta^2} + \cdots + \frac{0}{\beta^n}\right) \times \beta^{-m_L} = \beta^{-(m_L+1)}$ —— ②

オーバーフロー (overflow) … 絶対値が①を超えること, アンダーフロー (underflow) … 絶対値が②より小さくなること

これが起こらないよう, プログラミングで注意する.

2. 誤差

丸め … 実数 → 有限桁の浮動小数点で近似.
↓
丸め誤差 (round-off error)

β進n桁の浮動小数点数に切捨てを行ったときの相対誤差の上限

$$\varepsilon_M = \frac{\beta^{-n}}{\beta^{-1}} = \beta^{1-n}$$ … 計算機イプシロン (machine epsilon)

i.e. $1 \oplus \varepsilon_M = 1$ となる最大の正の浮動小数点数
($\oplus$ … 計算機における足し算)

* 手元のPCでは $\varepsilon_M = 1.084 \times 10^{-19}$ (倍精度)

ex. … 台形則による数値積分における丸め誤差の累積

$$\int_0^1 f(x)\,dx \simeq \frac{h}{2}f(0) + h\sum_{k=1}^{N-1} f(kh) + \frac{h}{2}f(1) \quad \left(h = \frac{1}{N}\right)$$

h をやみくもに小さくとればいいものではない.

情報落ち　大きい数に小さい数を足す
→ 小さい数の下位何桁かの情報が落ち,
足し算の結果に誤差が生じる.

$1.00000 + 1.23456 \times 10^{-3}$
$= 1.00000 + 0.00123\cancel{456}$　桁ずらし
$= 1.00123$

ex.: $\sum_{n=1}^{\infty} \frac{1}{n^2} \left(= \frac{\pi^2}{6}\right) \fallingdotseq \sum_{n=1}^{N} \frac{1}{n^2}$　単精度計算, $N = 1000000$

$1 + \frac{1}{2^2} + \cdots + \frac{1}{N^2}$ ⟶ 相対誤差 1.3×10^{-4}

$\frac{1}{N^2} + \cdots + \frac{1}{2^2} + 1$ ⟶ 〃 6.5×10^{-7}

桁落ち (cancellation of significant digits)

値の近い数どうしの引き算 → 有効桁数減少

$$1.2345678 - 1.2345555 = 0.0000123$$

よくある例： 数値微分 $f'(x) \simeq \dfrac{f(x+h)-f(x)}{h}$

例： 二次方程式の解 $ax^2+bx+c=0$

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

$a, b, c > 0$, $c \approx 0$ の場合

$$x = \frac{-b + \sqrt{b^2-4ac}}{2a} \rightarrow \text{桁落ち}$$

$$\downarrow$$

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

桁落ち解消