라운드 오프 오차

계산기나 컴퓨터가 실수 계산을 수행할 때 발생하는 오차

이진 기계수

이진 부동 소수점 연산 표준

단정도, 배정도, 확장정도

64비트

64비트

정규화된 배정도 수 $(-1)^{s} 2^{c-1023} (1+f)$

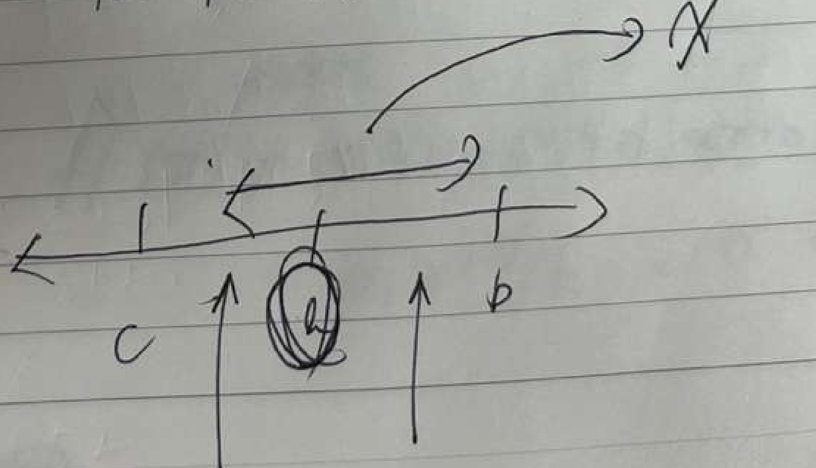
첫번째 비트 : s (1 bit)

2~12번째 비트 : c (11 bit)

나머지 비트 : f (52 bit)

소수부의 첫번째 비트는 항상 1

실제로 저장 X 한 비트의 추가적인 정확도 얻음



$\frac{a+c}{2}$ $\frac{a+b}{2}$

수렴률

$a_1 \quad a_2 \quad a_3 \quad \cdots \quad a_n \simeq x$

$$|\alpha - \alpha_n| \le \frac{K}{n^p}$$

$O(1/n^p)$ $\quad \alpha_n = \alpha + O(1/n^p)$

" $\alpha_n \to \alpha$ 이고, 수렴률 $1/n^p$ 이다"

예제 1.4.2 수열 $\{\alpha_n\}$ 과 $\{\hat{\alpha}_n\}$

$$\alpha_n = \frac{n+1}{n^2} \qquad \hat{\alpha}_n = \frac{n+3}{n^3}$$

$$|\alpha_n - 0| = \frac{n+1}{n^2} \le \frac{n+n}{n^2} = 2 \cdot \frac{1}{n}$$

$$|\hat{\alpha}_n - 0| = \frac{n+3n}{n^3} = 4 \cdot \frac{1}{n^2}$$

$$\alpha_n = 0 + O\left(\frac{1}{n}\right) \qquad \hat{\alpha}_n = 0 + O\left(\frac{1}{n^2}\right)$$

안정성

초기 데이터 값을 조금 변경시켰을 때,
조금 변경되는 방법 → 안정적

크게 변화 → 불안정적

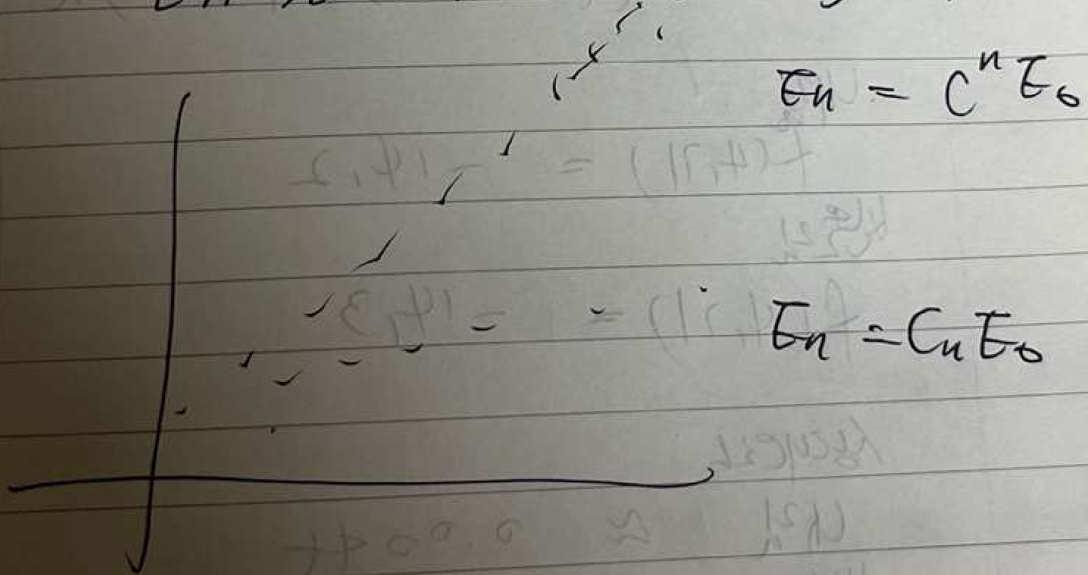
라운드 오류 오차가 증가하는 방식

$E_0 > 0$, n번의 연산이 더 적용 후

오차의 크기 E_n, n과 무관하게

$E_n \approx C_n E_0$ ⇒ 선형적 (linear)

$E_n \approx C^n E_0$ ⇒ 지수적 (exponential)



예제 1.4.1

예

$x = 4.71 \quad f(x) = x^3 - 6.1x^2 + 3.2x + 1.5$.

참값 = -14.263899

세자리 (내림) = -13.5

세자리 (반올림) = -13.4

상대오차

내림시 ≈ 0.05

반올림시 ≈ 0.06

중첩법에 쓸수 없다.

$f(x) = ((x - 6.1)\cancel{x} + 3.2)x + 1.5$

내림

$f(4.71) = -14.2$

반올림

$f(4.71) = -14.3$

상대오차

내림 ≈ 0.0045

반올림 ≈ 0.0025

VS.

1.4 과학 계산에서의 오차.

$ax^2+bx+c=0$ 꼴은 $a\neq 0$ 일때

$$x_1 = \frac{-b + \sqrt{b^2-4ac}}{2a}, \quad x_2 = \frac{-b-\sqrt{b^2-4ac}}{2a}$$

$x^2 + 62.10x + 1 = 0$ 같은

b^2가 $4ac$보다 훨씬 크기 때문에

$\sqrt{b^2-4ac} \approx b$.

상대오차 증가

$$x_1 = \left(\frac{-b+\sqrt{b^2-4ac}}{2a}\right)\left(\frac{-b-\sqrt{b^2-4ac}}{-b-\sqrt{b^2-4ac}}\right)$$

$$x_1 = \frac{-2c}{b+\sqrt{b^2-4ac}}$$ 로 구하면

상대오차가 작아져서 더 정밀한 계산 가능.

이식은 b가 양수일때만 사용, 아니면 상대오차 커

유한 자릿수 연산. 비슷한 값 빼기는 연산 가능하면 X
만약 작다면

예제 1.3.3 자리수 늘려야 한다.
네자리 십진 내림.

$$fl(\pi) = 0.3141 \times 10^1, \quad fl\left(\frac{22}{7}\right) = 0.3142 \times 10^1$$

$$fl(\pi) - fl\left(\frac{22}{7}\right) = -0.0001 \times 10$$

부동 소수점 표현

$$p^* = fl\left(fl(\pi) - fl\left(\frac{22}{7}\right)\right) = -0.1000 \times 10^{-2}$$

$$\left|\frac{\pi - fl(\pi)}{\pi}\right| \leq 0.0002,$$

$$\left|\frac{\frac{22}{7} - fl\left(\frac{22}{7}\right)}{\frac{22}{7}}\right| \leq 0.0003$$

$$\left|\frac{\pi - \frac{22}{7} - p^*}{\pi - \frac{22}{7}}\right| \approx 0.2092$$

거의 70%

절대오차 $|p-p^*|$

상대오차 $|p-p^*| / |p|$

[예제 1.3.2]

a. $p=3.000$, $p^*=3.100$:

절대오차 0.1 상대오차 $0.333\overline{3} \times 10^{-1}$

b. $p=0.003000$, $p^*=0.003100$

절대오차 0.0001 상대오차 $0.333\overline{3} \times 10^{-1}$

c. $p=3000$, $p^*=3100$

절대오차 : 100 상대오차 $0.333\overline{3} \times$

정확도 측면에서 절대오차보다 상대오차를 사용하는 것이 의미가 있다

십진 기계수

$\pm\ 0.d_1 d_2 \cdots d_k \times 10^n,$

$1 \le d_1 \le 9, \quad 0 \le d_i \le 9$

k자리 십진 기계수

$y = 0.d_1 d_2 \cdots d_k / d_{k+1} d_{k+2} \cdots \times 10^n$

내림, 반올림 두 가지 방법.

$\pi = 0.314159 \,26\, 25 \cdots \times 10^1$

a. 다섯자리 이후 내림

b. 다섯자리 이후 반올림

반올림.

$fl(\pi) = (0.31415 + 0.00001) \times 10^1$

$= 0.31416 \times 10^1 = 3.1416$

채원

└ 정규화된

표현 가능한 가장 작은 양수 $s=0,\ e=1,\ f=0$

$2^{-1022} \cdot (1+0) \approx 0.225 \times 10^{-307}$

〃 가장 큰 수 $s=0,\ e=2046,$

$f=1-2^{-52}.$

$2^{1023} \cdot (1+(1-2^{-52})) \approx 0.17977 \times 10^{309}$?

양의 0 $s=0,\ e=0,\ f=0$

음의 0 $s=1,\ e=0,\ f=0$