數値分析

Chapter 1
Mathematical Preliminaries
and
Error Analysis

- Ex. 證明 $0.\overline{9} = 1$

  Proof: Let $x = 0.\overline{9}$

  $10x = 9.\overline{9}$

  $9x = 9$

  $\Rightarrow x = 1$

- This book examines problems that can be solved by methods of a approximation, which we call numerical methods. For example, numerical differential.

- Ex. $f(x) = x^5 + 5x^4 + 3x^3 - 2x^2 + x - 1$, 求 $f(\pi)$.

  Sol: $((((x+5)x+3)x-2)x+1-1$

  (因 trumcate or round off 之次數少, 所以誤差小)

- Continuity and Differentiability (P.2∼P.3)

1.2 Review of Calculus

- <u>Mean Value Thorem</u>

  If $f \in C[a,b]$ and f is differentiable on (a,b), then a number c in (a,b) exists such that (see P.4 Figure 1.3)

  $f'(c) = \frac{f(b)-f(a)}{b-a}$

- <u>Extreme Value Thorem</u>

  If $f \in C[a,b]$, then c_1 and c_2 in $[a,b]$ exist with $f(c_1) \leq f(x) \leq f(c_2)$ for all x in $[a,b]$. If, in addition ,f is differentiable on (a,b), then the number c_1 and c_2 occur either at endpoints of $[a,b]$ or where f' is zero.

  (極值不是在邊界, 就是在 $f' = 0$)

- Mean Value Thorem for Integrals and Intermediate Value Theorem

  If $f \in C[a,b]$, g is integrable on $[a,b]$ and $g(x)$ does not change sign on $[a,b]$, then there exists a number c in (a,b) with

  $\int_a^b f(x)g(x)dx = f(c)\int_a^b g(x)dx$

  (可找到一個等面積之長方形)

- Taylor's Thorem

  Suppose $f \in C^n[a,b]$ and f^{n+1} exists on $[a,b]$. Let x_0 be a number in $[a,b]$. For every x in $[a,b]$, there exists a number $\xi(x)$ between x_0 and x with

  $f(x) = P_n(x) + R_n(x)$

  where

  $$\begin{aligned} P_n(x) &= f(x_0) + f'(x_0)(x-x_0) + \tfrac{f''(x_0)}{2!}(x-x_0)^2 \\ &\quad + ... + \tfrac{f^n(x_0)}{n!}(x-x_0)^n \\ &= \sum_{k=0}^{n} \tfrac{f^k(x_0)}{k!}(x-x_0)^k \end{aligned}$$

  and

  $R_n(x) = \frac{f^{n+1}(\xi(x))}{(n+1)!}(x-x_0)^{n+1}$

  (用 x_0 這個點上的各階動差來找出整個函數)

  (P.9 Figure 1.7)

- $f(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x-x_0) + \frac{\partial f}{\partial y}(y-y_0) + \frac{1}{2!}\frac{\partial^2 f}{\partial x^2}(x-x_0)^2 + \frac{1}{2!}\frac{\partial^2 f}{\partial y^2}(y-y_0)^2 + \frac{\partial^2 f}{\partial x \partial y}(x-x_0)(y-y_0)$ (二維)

  (It$\hat{o}'$s Lemma)

1.3 Round-off Error and Computer Arithmetic

- To save storage and provide a unique representation for each floating-point number, a normalization is imposed. Using this system gines a floating-point number of the form

  $(-1)^s \times 2^{c-1023} \times (1+f)$

  Consider for example, the machine number

  (s c f) = (sign characteristic mantissa)

  0 10000000011 1011100100010000000000000000000000000000

  The leftmost bit is zero, which indicates that the number is positive. The next 11 bits, 10000000011, giving the characteristic, are equivalent to the decimal number

  $c = 1 \times 2^{10} + 0 \times 2^9 + ... + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 1024 + 2 + 1 = 1027$

  The exponential part of the number is, therefore, $2^{1027-1023} = 2^4$. The final 52 bits specify that the mantissa is

  $f = 1 \times (\frac{1}{2})^1 + 1 \times (\frac{1}{2})^3 + 1 \times (\frac{1}{2})^4 + 1 \times (\frac{1}{2})^5 + 1 \times (\frac{1}{2})^8 + 1 \times (\frac{1}{2})^{12}$

  As a consequence, this machine number precisely represents the decimal number

  $(-1)^s \times 2^{c-1023} \times (1+f) = (-1)^0 \times 2^{1027-1023}(1 + (\frac{1}{2} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{256} + \frac{1}{4096})) = 27.56640625$

  However, the next smaller machine number is

  0 10000000011 1011100100001111111111111111111111111111

  and the next larger machine number is

  0 10000000011 1011100100010000000000000000000000000001

  This means that our original machine number represents not only 27.56640625, but also half of the real numbers that are between 27.56640625 and its two nearest machine-number neighbors.

- The smallest normalized positive number that can be represented has all zeros except for the rightmost bit of 1 and os equivalent to

  $2^{-1023}(1+2^{-52}) \approx 10^{-308}$

  and the largest has a leading 0 followed by all 1s and is equivalent to

  $2^{1024}(2-2^{-52}) \approx 10^{308}$

  Numbers occuring in calculations that have a magnitude less than $2^{-1023}(1+2^{-52})$ result in underflow and are generally set to zero . Numbers greater than $2^{1024}(2-2^{-52})$ result in overflow and typically cause the computation to halt.

- Rounding and Chopping

  (The error that results from replacing with its floating-point form is called round-off error)

1.4 Errors in Scientific Computation

- (P.21)

  Ex 1. $x^2 + 62.10x + 1 = 0$, whose roots are approximately

  $x_1 = -0.0160723$ and $x_2 = -62.08390$

  In this equation, b^2 is much larger than $4ac$, so the numerator in the calculation for x_1 involves the subtration of nearly equal numbers. Since

  $\sqrt{b^2-4ac} = \sqrt{(62.10)^2-(4)(1)(1)} = 62.06$ (62.06778552)

  we have

  $fl(x_1) = \frac{-62.10+62.06}{2.000} = \frac{-0.04000}{2.000} = -0.02000$ (因 62.10 與 62.06 很接近, 使得 trumcate 掉的 0.00778552 顯得很重要),

  a poor approximation to $x_1 = -0.01611$ with the larger relative error

  $\frac{|-0.01611+0.002000|}{|-0.01611|} = 2.4 \times 10^{-1}$ (誤差很大)

On the other hand, the calculation for x_2 involves the addition of nearly equal numbers $-b$ and $-\sqrt{b^2-4ac}$. The presents no problem since

$fl(x_2) = \frac{-62.10-62.06}{2.000} = \frac{-124.2}{2.000} = -62.10$,

has the small relative error

$\frac{|-62.08+62.10|}{|-62.08|} = 3.2 \times 10^{-4}$ (誤差很小)

要獲得更精確的 x_1 逼近值, 使用反有理化.

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

$\Rightarrow fl(x_1) = \frac{-2.000}{62.10+62.06} = -0.01610$

which has the small relative error 6.2×10^{-4}

(反有理化後, 誤差變小)

但同時, $fl(x_2)$ 若也做反有理化, 會從誤差小 $\longrightarrow$ 誤差大.