Lecture 25

當初 Taylor 的問題是, 如何將一個一般形式的多項式 $a_nx^n+\cdots+a_1x+a_0$, 改寫成 $x-x_0$ 的多項式 (此處 x_0 是任一個常數)? 也就是說, 如果

$$a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0 = p_n(x-x_0)^n+p_{n-1}(x-x_0)^{n-1}+\cdots+p_1(x-x_0)+p_0,$$

那麼如何求 p_k? 他發現, 如果把左式寫成 $p(x)$, 則 $p_0 = p(x_0)$. 如果求左右兩式在 x_0 的導數, 則得 $p_1 = p'(x_0)$. 一般而言, 如果寫出左右兩式的 k 階導函數 $(0 \le k \le n)$, 則得

$$\begin{aligned} p^{(k)}(x) = {} & p_n n\cdots(n-k+1)(x-x_0)^{n-k} + p_{n-1}(n-1)\cdots(n-k)(x-x_0)^{n-k-1} \\ & +\cdots+p_{k-1}(k+1)k\cdots 2(x-x_0)+p_k k!. \end{aligned}$$

所以 $p_k = p^{(k)}(x_0)/k!$.

推廣一點, 我們可以定義所謂的泰勒多項式如下.

> **若 $f(x)$ 在 x_0 處至少 n 次可微, 則所謂 $f(x)$ 在 x_0 附近的 n 階泰勒多項式就是一個符合以下條件的 n 階多項式 $p(x)$:**
>
> $$f^{(k)}(x_0) = p^{(k)}(x_0), \quad 0 \le k \le n. \tag{1}$$
>
> **因此, $p(x)$ 可以寫成 $x-x_0$ 的多項式:**
>
> $$f(x_0)+\frac{1}{1!}f'(x_0)(x-x_0)+\frac{1}{2!}f''(x_0)(x-x_0)^2+\cdots+\frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n.$$

因為 (1) 式, 我們想像 $p(x)$ 在 x_0 "附近" 應該很 "靠近" $f(x_0)$. 例如參考課本 **10.1** 節的圖 10.4 ($\sin x$ 在 0 附近的 7 階泰勒多項式, 在 $(-\pi,\pi)$ 之間很靠近 $\sin x$), 圖 10.5 ($\cos x$ 在 0 附近的 8 階泰勒多項式, 在 $(-\pi,\pi)$ 之間很靠近 $\cos x$), 圖 10.6 (e^x 在 0 附近的 4 階泰勒多項式, 在 $(-2,2)$ 之間很靠近 e^x), 圖 10.8 ($\ln x$ 在 1 附近的 4 階泰勒多項式, 在 $(-\frac{1}{2},\frac{1}{2})$ 之間很靠近 $\ln x$). 但是,

> **在 x_0 的多麼附近之內, 泰勒多項式 $p(x)$ 將會多麼靠近 $f(x)$ 呢?**

這個量化的問題, 是所謂**泰勒級數**或**泰勒展開**理論要討論的核心問題. 泰勒以及牛頓並沒有完全解決這個問題, 但是他們卻導出許多令人驚豔的結果. 而當時的許多結果都沒有證明, 純粹是"信不信由你."

首先, 我們可以觀察到 (泰勒當然也觀察到了), 如果 $f(x)$ 本身就是一個多項式的話, 那麼它可以在任意點 x_0 附近取任意高階的的泰勒多項式 (因為 $f(x)$ 在任意點處都有任意高階

導數, 大不了是 0). 我們何不把這泰勒"多項式"寫到無窮多階呢? 無窮多階的"多項式"就不再叫一個"多項式"了, 它叫做**冪級數** (power series). 其一般形式就是

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots. \tag{2}$$

基本上, 無窮高階的泰勒"多項式"就叫泰勒級數:

> 如果 $f(x)$ 在 x_0 處有無窮多次導數 (無窮多次可微), 則 $f(x)$ 在 x_0 附近的泰勒級數就是
>
> $$T(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n.$$

特別當 $x_0 = 0$ 的時候, 泰勒級數就是 $\sum \frac{f^{(n)}(0)}{n!} x^n$, 又稱做麥克勞林* 級數. 這或許是因為他在 1742 年出版了一本介紹微分的書 (其實是介紹牛頓的流數術), 而泰勒級數也被包含在他的書裡.

如果 $f(x)$ 是個多項式, 則其泰勒級數 $T(x)$ 其實只有有限項 (因為 $f^{(k)}(x) = 0, \forall k \geq \deg f$), 所以沒有 $T(x)$ 收斂或發散的問題. 而且 $f(x)$ 和 $T(x)$ 不僅是"靠近"而已, 它們其實"相等" ($T(x)$ 不過是 $f(x)$ 換一個形式的寫法而已). 但是, 如果 $f(x)$ 是個一般的函數 (例如基本函數中的有理函數, 指數對數函數, 三角函數), 事情可就沒有這麼簡單了. 我們首先應該搞清楚, 冪級數是什麼意思? 模倣級數的部分和定義, 我們做如下的定義:

> 考慮冪級數如 (2), 定義其部分和函數 $f_n(x)$ 爲
>
> $$f_n(x) = \sum_{k=0}^{n} a_k x^k.$$
>
> 若函數數列 $f_n(x)$ 在 $[a, b]$ 內收斂, 且極限函數爲 $f(x)$, 則稱此冪級數收斂, 且寫作
>
> $$\sum_{n=0}^{\infty} a_n x^n = f(x), \quad x \in [a, b].$$
>
> 否則稱此冪級數發散.

* Maclaurin (1698–1746), 史稱此君是牛頓最成功的門徒. 不過, 此人出生的時候牛頓已經五十六歲了, 他可能不是牛頓親炙的學生.

接著, 我們得要定義什麼是函數數列的收斂了. 其實, 就是固定一個點 x 來看, 其函數值所形成的實數數列個自收斂; 把所有這些極限值聚在一起定義一個新的映射, 由極限的唯一性得知這個映射還是個函數. 這就是極限函數了.

> 若每個 $f_n(x)$ 均在 $[a,b]$ 中有定義, 且存在某個函數 $f(x)$, 使得任給一個 $\epsilon > 0$, 對每個 (選定的) $x \in [a,b]$, 都找得到一個 $N > 0$ $(N = N(\epsilon, x))$, 以至於 $n \geq N \Longrightarrow |f_n(x) - f(x)| < \epsilon$. 則稱函數數列 $\{f_n(x)\}$ 在 $[a,b]$ 內收斂, 且其極限函數是 $f(x)$.

注意, 此處的 $[a,b]$ 可以視情況換成開區間, 或是無限長的區間. 上面這個定義, 是一個點一個點來討論是否收斂, 而必須對每個點都有極限值, 才可以定義極限函數. 這種收斂定義又細稱做**逐點收斂**. 以後, 各位同學或許會學到其他方式的函數數列收斂的定義.

至此, 我們有了第二個問題:

> 如果 $f(x)$ 在 x_0 附近有泰勒級數, 那麼在 x_0 的多麼附近之內該級數收斂? 收斂後的極限函數是否等於 $f(x)$?

其實, 我們將發現, 這個問題和前一個問題, 有共同的答案.

由於時間的限制, 我們將割捨所有理論推導的部分, 而只能綜合介紹結論.

課本 10.5 綜合理論

> 如果 $f(x)$ 在 x_0 處至少 $n+1$ 次可微, 且其第 $n+1$ 次之導函數連續. 令 $p_n(x)$ 爲 $f(x)$ 在 x_0 附近的 n 階泰勒多項式, 則
>
> $$f(x) = p_n(x) + R_n(x) := p_n(x) + \frac{1}{(n+1)!} f^{(n+1)}(\xi)(x - x_0)^{n+1}.$$
>
> 其中 ξ 是 x 和 x_0 之間的某一個數.

$R_n(x)$ 稱為 $p_n(x)$ (n 階泰勒多項式) 與 $f(x)$ 之間的**誤差項**, 或是**餘式**. 有幾種不同的表達 $R_n(x)$ 的方法, 上述是所謂 Lagrange* 表達式. 我們無法確知 ξ 是何物, 但是, 上述理論至少有兩個推論. 第一, 令 $h > 0$, 如果在 $[x_0 - h, x_0 + h]$ 的範圍之內 (x_0 的"附近"),

$$L \leq |f^{(n+1)}(x)| \leq M, \quad (\forall x \in [x_0 - h, x_0 + h])$$

(連續函數在閉區間中有最大值和最小值), 則可以估計

$$L \frac{h^{n+1}}{(n+1)!} \leq |f(x) - p_n(x)| = |R_n(x)| \leq M \frac{h^{n+1}}{(n+1)!}.$$

* Lagrange (1736–1813), 他是在泰勒身後五年才出世的.

如果固定了 n 和 h, 那麼 L 和 M 也跟著固定了. 所以可以估計 $p_n(x)$ 有多麼"靠近" $f(x)$. 如果有泰勒級數, 那就要問, 當 h 在什麼範圍內 (x 有多麼"靠近" x_0),

$$\lim_{n\to\infty} R_n(x) = 0?$$

注意, L 和 M 也許會跟著 n, x_0 和 h 變化. 例如當 $f(x) = 1/x$ 的時候.

如果 $\lim R_n(x) = 0$, 那麼也就是

$$\lim_{n\to\infty} |f(x) - p_n(x)| = 0.$$

所以泰勒級數若收斂, 其極限函數必定是 $f(x)$. 而且, 以泰勒多項式逼近函數 $f(x)$ 的情況果真就像以十進位的小數逼近實數一樣, 越後面的項越小; 也就是說, 當 n 越大, $|\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n|$ 就越小. 隨著 n 的增加, 泰勒多項式也就越來越"靠近" $f(x)$, 而其"靠近的程度"就是 $|R_n(x)|$.

第二, 若取 $n = 1$, 則得

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(\xi), \quad \xi \text{ 在 } x \text{ 和 } x_0 \text{ 之間}.$$

這意味著, 如果 $f(x)$ 在 $[a,b]$ 區間內的割線斜率 (平均變化率) 是 c, 則在 $[a,b]$ 之中必有一點 ξ, 使得 $f(x)$ 在 ξ 的切線斜率 (瞬間變化率) 也是 c. 這叫做**微分平均值定理**.

課本 10.2 重要的泰勒級數

泰勒多項式的發現, 想必使得牛頓回憶起自己早年 (二十三歲) 發現的二項級數 (binomial series)*

e^x, $\cos x$, $\sin x$ 在 0 附近的泰勒級數, 其實在整個 $\mathbb{R}$ 上都會收斂. 以 e^x 為例,

$$R_n(x) = \frac{1}{(n+1)!} e^{\xi} x^{n+1}, \quad \text{某個介於 } 0 \text{ 和 } x \text{ 之間的 } \xi.$$

任給定一個 $x \in \mathbb{R}$, 令 $M = \lceil 2|x| \rceil$. 那麼, 隨便指定一個 $\epsilon > 0$, 則取 $N = \lceil \log_2 \frac{|2x|^M}{\epsilon M!} \rceil$, 每當 $n > N$,

$$\begin{aligned} |\frac{x^{n+1}}{(n+1)!}| &= \frac{|x|^M}{M!} \cdot \frac{|x|}{M+1} \cdots \frac{|x|}{n+1} \\ &\le \frac{|x|^M}{M!} \cdot \frac{1}{2^{n-M+1}} \le \frac{|2x|^M}{M!} \cdot \frac{1}{2^N} < \epsilon. \end{aligned}$$

* 史稱此為牛頓在數學上的第一個貢獻. 時值 1665 年, 牛頓剛從劍橋畢業, 回到鄉下躲避瘟疫. 此後的十二年, 他獨力寫下了微積分基本定理和萬有引力原理, 發現了光的三原色, 還發明了反射望遠鏡.

所以

$$\lim_{n\to\infty} R_n(x) = e^{\xi} \lim_{n\to\infty} \frac{x^{n+1}}{(n+1)!} = 0, \quad \forall x \in \mathbb{R}.$$

注意 $\sin x$ 和 $\cos x$ 的泰勒 (麥克勞林) 級數, 它們對應了 $\sin x$ 和 $\cos x$ 分別為奇函數和偶函數的性質.

歐拉公式 $e^{i\theta} = \cos\theta + i\sin\theta$, 複數平面的單位圓.

課本 10.3 泰勒級數的變化和應用

如果 $f(x)$ 和 $g(x)$ 都在 x_0 處有一次以上的導數, 而且 $f(x_0) = g(x_0) = 0$. 則

$$\lim_{x\to x_0} \frac{f(x)}{g(x)}$$

稱為不定形 (indetermined form) 之極限問題. 因為 $f(x) = f'(x_0)(x-x_0) + f''(x_0)(x-x_0)^2 + \cdots$, $g(x)$ 亦類似, 所以

$$\frac{f(x)}{g(x)} = \frac{f'(x_0)(x-x_0) + f''(x_0)(x-x_0)^2 + \cdots}{g'(x_0)(x-x_0) + g''(x_0)(x-x_0)^2 + \cdots}.$$

上下各除一個 $x - x_0$ 項, 而且如果 $f'(x_0)/g'(x_0)$ 不再是不定形, 則

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{f'(x_0)}{g'(x_0)}.$$

這就是所謂的羅必達 (L'Hôpital, 1661–1704) 法則. 這雖然是個方便求極限的工具, 但是不能告訴我們收斂的速度. 直接使用泰勒多項式, 不但可以求得極限, 還可以告訴我們收斂的速度.

習題 25

25.1 初學者常常以為一個函數如果可微就可以一直微分下去. 其實不然. 例如考慮 $f(x) = x^2 \sin\frac{1}{x}$. 注意 $x^2 \sin\frac{1}{x}$ 本身在 0 點沒有定義, 但是因為

$$\lim_{x\to 0} x^2 \sin\frac{1}{x} = 0,$$

所以 $f(x)$ 屬於可以補救的不連續. 只要定義 $f(0) = 0$ 就連續了. 現在, 請您 (1) 以微分公式求當 $x \neq 0$ 的時候 $f(x)$ 的導函數. (2) 以微分定義證明 $f(x)$ 在 $x = 0$ 處是可微的, 而其導數值是多少? 所以 $f(x)$ 的一次導函數 $f'(x)$ 是存在的. 但是, 請證明: (3) $f'(x)$ 在 $x = 0$ 處不連續. 因此, $f'(x)$ 在 $x = 0$ 處不可微; 換句話說, $f(x)$ 在 $x = 0$ 處恰僅有一次導數.

25.2 初學者常常以為如果每個 $f_n(x)$ 都是連續函數, 則它們的極限函數也連續. 其實不然. 例如在 $[0,1]$ 區間中考慮函數數列 $f_n(x) = x^n$. 每個 $f_n(x)$ 都在 $[0,1]$ 中連續. 請證明 (1) 若 $0 \le x < 1$, 則 $\lim f_n(x) = 0$; (2) 若 $x = 1$, 則 $\lim f_n(x) = 1$. 因為對每個 $x \in [0,1]$, $\lim f_n(x)$ 的極限值都存在, 所以可以定義函數極限

$$f(x) = \begin{cases} 0 & \text{當 } x \in [0,1), \\ 1 & \text{否則,} \end{cases} \qquad x \in [0,1].$$

可見極限函數 $f(x)$ 是不連續的.

25.3 課本 **10.5** 習題 5, 7.

25.4 從大園到中壢的距離是十二公里, 全線的速限是時速六十公里. 某人花了十分鐘開車走完全程, 而他宣稱自己絕對沒有超速. 請證明他說謊.

25.5 課本 **10.2** 習題 19–24.

25.6 課本 **10.2** 習題 31.

25.7 課本 **10.3** 習題 1, 5, 8, 10, 11.

25.8 課本 **10.3** 習題 18.

25.9 課本 **10.3** 習題 20.

25.10 課本 **10.3** 習題 23.