Differential 和 Derivatives

Differential 和 Derivative 的詮釋

u 是某變量。習慣上,

$\triangle u$ 是指 變量 u 的變化量;

du 是指 變量 u 的微變化量, 稱爲 *differential of u* ;

f 是以 x 爲變量的函數。我們可以求 自變量 x 的變化量 和對應的 因變量 f 的變化量 之間的比值 :

$\frac{\triangle f}{\triangle x}$ 指的是**函數的**(平均)**變化率**;

$\frac{df}{dx}$ 指的是**函數的微變化率**, 稱爲 derivative of f , 即, 兩個微變化量 df 和 dx 的比值。

所以很自然地,

$\frac{df}{dx}(a)$, 也記作 $\frac{d}{dx}\Big[f(x)\Big]_{x=a}$ 、$f'(a)$ 、$\boldsymbol{D}_x f\,(x)\Big|_{x=a}$ 、$\boldsymbol{D}f\,(a)$, 定義爲

$$\frac{df}{dx}(a) \stackrel{\text{def}}{=} \lim_{\triangle x\to 0}\frac{f(a+\triangle x)-f(a)}{(a+\triangle x)-a} = \lim_{\triangle x\to 0}\frac{f(a+\triangle x)-f(a)}{\triangle x},$$

稱爲 derivative of f at a 或 (instantaneous) rate of change of f at a / f 在 a 處的瞬間變化率 ; 給 limit 的變量換個寫法, 就變成

$$\frac{df}{dx}(a) \stackrel{\text{def}}{=} \lim_{h\to 0}\frac{f(a+h)-f(a)}{h} = \lim_{y\to a}\frac{f(y)-f(a)}{y-a}$$

$\frac{df}{dx}(x)$, 也記作 $\frac{d}{dx}\Big[f(x)\Big]$ 、$\frac{df}{dx}$ 、$f'(x)$ 、f' 、$\boldsymbol{D}_x f\,(x)$ 、$\boldsymbol{D}f\,(x)$ 、$\boldsymbol{D}f$, 定義爲

$$\frac{df}{dx}(x) \stackrel{\text{def}}{=} \lim_{\triangle x\to 0}\frac{f(x+\triangle x)-f(x)}{(x+\triangle x)-x} = \lim_{\triangle x\to 0}\frac{f(x+\triangle x)-f(x)}{\triangle x}\left(= \text{“}\lim_{\triangle x\to 0}\frac{\triangle f}{\triangle x}\text{”}\right),$$

稱爲 derivative of f (at x) 或 (instantaneous) rate of change of f / f (在 x 處) 的瞬間變化率 。給 limit 的變量換個寫法, 就變成

$$\frac{df}{dx} \stackrel{\text{def}}{=} \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = \lim_{y\to x}\frac{f(y)-f(x)}{y-x}$$

當然, 所有的定義一定要在極限值存在時才有意義: If $\frac{df}{dx}(a)$ is well-defined, then we say f is differentiable at a; If $\frac{df}{dx}(x)$ is well-defined $\forall\, x$ in the domain of f, then we say f is differentiable(可微). 很自然地, 如果函數 f 在 a 處不連續, 則 $\frac{df}{dx}(a)$ 不存在, 即 f 在 a 處不可微。逆命題即 若 f 在 a 處可微, 則 f 在 a 處連續。

既然 derivative 是用 limit 定義的, 而 limit 又分 left limit(左極限) 和 right limit(右極限), 所以 derivative 也分成

left derivative: $\boldsymbol{D}^{-}f\,(a) = f'(a^{-}) \stackrel{\text{def}}{=} \lim_{y\to a^{-}}\frac{f(y)-f(a)}{y-a} = \lim_{h\to 0^{-}}\frac{f(x+h)-f(x)}{h},$

right derivative: $\boldsymbol{D}^{+}f\,(a) = f'(a^{+}) \stackrel{\text{def}}{=} \lim_{y\to a^{+}}\frac{f(y)-f(a)}{y-a} = \lim_{h\to 0^{+}}\frac{f(x+h)-f(x)}{h}.$

於是: $\boldsymbol{D}f\,(a)$ 存在 (即 f 在 a 處可微) $\iff$ $\boldsymbol{D}^{-}f\,(a)$ 和 $\boldsymbol{D}^{+}f\,(a)$ 皆存在並相等。
所以, 我們又可以得到: f (在 a 處) 可微 $\Rightarrow$ $\boldsymbol{D}f$ (在 a 處) 連續 (因爲 左極限=右極限)。

Derivative 的特性 & 基本函數的 Derivative

形式上 (不考慮 limit 存在性的問題), derivative 具有下列性質 (以下的 a, b, c 皆為 constant):

特性: $\frac{d}{dx}[cf] = c\frac{df}{dx}$, $\frac{d}{dx}[f+g] = \frac{df}{dx} + \frac{dg}{dx}$ (derivative 具有 *linearity*, 即 $\frac{d}{dx}$、$\boldsymbol{D}$ 是 "線性算子")

$\frac{d}{dx}[fg] = \frac{df}{dx}g + f\frac{dg}{dx}$ (*product rule*)

$\frac{d}{dx}[\frac{f}{g}] = \frac{\frac{df}{dx}g - f\frac{dg}{dx}}{g^2}$ (*quotient rule*)

$\frac{d}{dx}[f(g(x))] = \frac{df}{du}\Big|_{u:=g(x)} \frac{du}{dx}$ (*chain rule*) "$\frac{df}{dx} = \frac{df}{dg}\frac{dg}{dx}$" 比較直觀、好記

基本式: $\frac{d}{dx}[x^a] = ax^{a-1}$,

$\frac{d}{dx}[\ln x] = \frac{1}{x}$,

$\frac{d}{dx}[e^x] = e^x$,

$$\left.\begin{array}{l} \frac{d}{dx}[\sin x] = \cos x, \\ \frac{d}{dx}[\cos x] = -\sin x, \end{array}\right\} \Rightarrow \left\{\begin{array}{l} \frac{d}{dx}[\tan x] = \sec^2 x, \\ \frac{d}{dx}[\sec x] = \sec x \tan x, \\ \frac{d}{dx}[\cot x] = -\csc^2 x, \\ \frac{d}{dx}[\csc x] = -\csc x \cot x. \end{array}\right.$$

例如, 由基本式可以導出

$$\frac{d}{dx}[\ln(ax)] = \frac{d}{dx}[\ln x + \ln a] \underbrace{= \frac{d}{dx}[\ln x] + \frac{d}{dx}[\ln a]}_{\text{by linearity}} = \frac{1}{x} + 0 = \frac{1}{x},$$

$$\frac{d}{dx}[\log_b x] \underbrace{= \frac{d}{dx}[\frac{\ln x}{\ln b}]}_{\text{change of base}} \underbrace{= \frac{1}{\ln b} \cdot \frac{d}{dx}[\ln x]}_{\text{by linearity}} = \frac{1}{\ln b}\frac{1}{x},$$

$$\begin{aligned} \frac{d}{dx}[e^{ax}] &\left(= \tfrac{d}{du}[e^u]\tfrac{du}{dx} = e^u a\right) \quad (u := ax \text{ 並根據 chain rule}) \\ &= e^{ax}a, \end{aligned}$$

$$\begin{aligned} \frac{d}{dx}[b^x] &= \frac{d}{dx}[e^{x\ln b}] \\ &\left(= \tfrac{d}{du}[e^u]\tfrac{du}{dx} = e^u \ln b\right) \quad (u := x\ln b \text{ 並根據 chain rule}) \\ &= b^x \ln b. \end{aligned}$$

這些都還算是很基本的。再看複雜一點的:

$$\begin{aligned} \frac{d}{dx}[\sin(\cos x)] &= \frac{d}{du}[\sin u] \cdot \frac{du}{dx} \quad (u := \cos x \text{ 並根據 chain rule}) \\ &= \cos u \cdot (-\sin x) = -\cos(\cos x) \cdot \sin x. \end{aligned}$$

$$\begin{aligned} \frac{d}{dx}[x^x] &= \frac{d}{dx}[e^{x\ln x}] = \frac{d}{du}[e^u] \cdot \frac{d}{dx}[x\ln x] \quad (u := x\ln x \text{ 並根據 chain rule}) \\ &= e^u \cdot \underbrace{(\ln x + x \cdot \tfrac{1}{x})}_{\text{by product rule}} = x^x(\ln x + 1). \end{aligned}$$

$$\begin{aligned} \frac{d}{dx}[(x^2+1)^{(x^2-1)}] &= \frac{d}{dx}[e^{(x^2-1)\ln(x^2+1)}] = \frac{d}{du}[e^u] \cdot \frac{du}{dx} \quad (u := (x^2-1)\ln(x^2+1) \text{ 並根據 chain rule}) \\ &= e^u \cdot \underbrace{\left(2x\ln(x^2+1) + (x^2-1) \cdot \frac{d}{dx}[\ln(x^2+1)]\right)}_{\text{by product rule}} \\ &= e^u \cdot \left(2x\ln(x^2+1) + (x^2-1) \cdot \underbrace{\tfrac{2x}{x^2+1}}_{\text{by chain rule}}\right) = (x^2+1)^{(x^2-1)}\left(2x\ln(x^2+1) + \frac{2x^3-2x}{x^2+1}\right). \end{aligned}$$

從以上得知, 求函數的 derivative 多半須要用到 chain rule, 特別是 合成函數的 derivative。可見 函數合成 和 chain rule 的關連性和重要性。

再次強調, df、dx、$\cdots$ 是從 $\triangle f$、$\triangle x$、$\cdots$ 來的, d 是 differential 算子, 不是數字, 不可以分開。雖然 $\boldsymbol{D}=\frac{d}{dx}$ 是 derivative 算子, 但因爲 $\frac{df}{dx}$ 是 $\triangle f$ 除以 $\triangle x$ 取極限, 所以 $\frac{df}{dx}$ 可以看作是兩個數相除: df 除以 dx, 即 $df=\frac{df}{dx}\cdot dx=f'\cdot dx$, 於是:

- Chain rule 是非常自然的: $\frac{df}{dx}=\frac{df}{du}\frac{du}{dx}=\frac{df}{du}\frac{du}{dv}\frac{dv}{dx}=\frac{df}{du}\frac{du}{dv}\frac{dv}{dw}\frac{dw}{dx}=\cdots$, 再多層的函數合成也不怕。用普通話解釋 $\frac{d}{dx}f(u(x))$: 原來 在不知 u 是 x 的函數的情況下,f 對其變量的變化率爲 $\frac{d}{d\spadesuit}f(\spadesuit)=\frac{d}{du}f(u)$, 後來知道 u 是 x 的函數, u 對其變量 x 的變化率爲 $\frac{du}{dx}$ — 原來 f 隨 u 而變, 後來知道 u 隨 x 而變 $\Rightarrow$ f 隨 x 而變 — f 對 x 的變化率 等於 f 對 u 的變化率 乘以 u 對 x 的變化率。就像是:A 是 B 的 r_1 倍, B 是 C 的 r_2 倍, 所以 A 是 C 的 $r_1\cdot r_2$ 倍 一樣簡單。(df 是 du 的 $\frac{df}{du}$ 倍, du 是 dx 的 $\frac{du}{dx}$ 倍, 所以 df 是 dx 的 $\frac{df}{du}\frac{du}{dx}$ 倍)

- $$\begin{aligned} \because df &= \tfrac{df}{dx}\cdot dx, && \text{所以 , 當 } \triangle x\approx 0, && \triangle f\approx f'\cdot\triangle x, \\ &= f'\cdot dx, && \text{即 當 } x\approx a, && f(x)-f(a)\approx f'(a)\cdot(x-a), \quad \text{— 函數值變化量的估計,} \\ & && \text{即 當 } x\approx a, && f(x)\approx f(a)+f'(a)\cdot(x-a), \quad \text{— \textit{linear approximation}。} \end{aligned}$$

  再仔細看, 這個 $y-f(a)=f'(a)\cdot(x-a)$ by point-slope formula (點斜式),
  即 $y=f(a)+f'(a)\cdot(x-a)$, 就是在 a 處的切線方程式。

- 令 $y=f(x)$, 則 $\frac{dx}{dy}=\frac{1}{\frac{dy}{dx}}=\frac{1}{f'(x)}$ — 除了 $\frac{dy}{dx}=0$ 的地方以外都可以求 $\frac{dx}{dy}$。我們以前學過, 假使 x 可以用 y 唯一表示則反函數 $x=f^{-1}(y)$ 存在, 那麼 x 就可以對 y 求導。所以, 前面的意義, 就是: $\boxed{\text{在 } f'\neq 0 \text{ 的地方, \textit{locally}, 都有反函數 } f^{-1}}$。雖然如此, 但我們是否須要確實知道 x 如何用 y 表示、y 如何用 x 表示, 才能夠求 $\frac{dx}{dy}$、$\frac{dy}{dx}$ 嗎? 當然不必如此麻煩。

  $f(x)$ 是 以 x 爲變量的函數。令 變量 y 等於函數值 $f(x)$, 則 $y=f(x)$ 就可視爲 x,y 的方程式, 才有所謂的自變量、因變量 (因爲 x 是函數 f 的變量、加上那個"令"的動作把原本不相干的變量 y 說成函數值$\cdots$)。

- 方程式 是用來說明變量之間的關係, 一般來說, 沒有一定誰是 主、從、自變量、因變量的說法。例如說, 我們把任何 兩個變量 x,y 的方程式寫成 $E(x,y)=0$, 則 因爲 在此限制下 左邊等於右邊 $\Rightarrow$ 左邊 E 的微變化量 等於 右邊的微變化量 (=0), 而左邊 E 的變化源自於 x,y 的變化, 所以 E 的微變化量由 x,y 的微變化量構成, 所以由原本 $E=0$ 是 x、y 的關係式, 最後 $d(E)=d(0)$ 推導下去 就變成 dx、dy 的關係式了。這就是求 所謂的 *implicit derivative* 的觀念。

  例如, $u^v=v^u$, $\Rightarrow v\ln u=u\ln v$, $\Rightarrow d(v\ln u)=d(u\ln v)$, $\Rightarrow dv\cdot\ln u+v\cdot\frac{du}{u}=du\cdot\ln v+u\cdot\frac{dv}{v}$, $\Rightarrow(\frac{v}{u}-\ln v)\cdot du=(\frac{u}{v}-\ln u)\cdot dv$, 如此, $\frac{du}{dv}$ 或 $\frac{dv}{du}$ 就輕易求得了。當然, 這只是 了解 differential 的好處之一。

  我們一再強調 不要拘泥在名字上, 要 摒棄自以爲是的 主從、因果 關係。在這兒就可以看出好處。如果了解這個道理, 以後凡是求 derivative, 都要用 "這個的 differential 和 那個的 differential 的關係" 來想。(求 differential 的方法 在前一點已經說了 : $df=\frac{df}{dx}\cdot dx=\frac{df}{du}\cdot\frac{du}{dx}\cdot dx=\frac{df}{du}\cdot\frac{du}{dv}\cdot\frac{dv}{dx}\cdot dx=\cdots$)

- 姑且不論兩變量 u,v 的關係。若知 u,v 的參數式 $\begin{cases} u=u(t) \\ v=v(t) \end{cases}$, 則 一樣可求 n 階導數 $\frac{d^n}{dv^n}[u]\overset{\text{def}}{=}\frac{d}{dv}[\cdots\frac{d}{dv}[\frac{d}{dv}[u]]\cdots]$:

  $\begin{cases} du=\frac{du}{dt}\,dt=\dot{u}\,dt \\ dv=\frac{dv}{dt}\,dt=\dot{v}\,dt \end{cases}$, $\frac{du}{dv}=\frac{\dot{u}}{\dot{v}}$; $\frac{d^2u}{dv^2}\overset{\text{def}}{=}\frac{d}{dv}(\frac{du}{dv})=\frac{d(\frac{\dot{u}}{\dot{v}})}{dv}=\frac{\frac{d}{dt}(\frac{\dot{u}}{\dot{v}})}{\dot{v}}$; $\cdots\cdots$ 只要分母的 $\dot{v}$ 不爲 0 。

- 了解 一般反函數的 derivative 後, 現在具體來求**反三角函數的 derivatives**

sin 的反函數 記作 **arcsin** (或 $\sin^{-1}$)。令 $r=\sin\theta$ 並限制 $\theta\in[\frac{-\pi}{2},\frac{\pi}{2}]$ 致使 一對一 而有反函數 $\theta=\arcsin(r)$, $r\in[-1,1]$, 則 $\frac{d}{dr}[\arcsin(r)]=\frac{d\theta}{dr}=\frac{1}{\frac{dr}{d\theta}}=\frac{1}{\frac{d}{d\theta}[\sin\theta]}=\frac{1}{\cos\theta}=\frac{1}{+\sqrt{1-\sin^2\theta}}=\frac{1}{\sqrt{1-r^2}}$, 最後要用自己原來的變量 r 來表示。($\cos\theta=+\sqrt{1-\sin^2\theta}$ 是因爲在此 $\theta\in[\frac{-\pi}{2},\frac{\pi}{2}]\Rightarrow\cos\theta\geq 0$)

tan 的反函數 記作 **arctan** (或 $\tan^{-1}$)。令 $r=\tan\theta$ 並限制 $\theta\in(\frac{-\pi}{2},\frac{\pi}{2})$ 致使 一對一 而有反函數 $\theta=\arctan(r)$, $r\in(-\infty,\infty)$, 則 $\frac{d}{dr}[\arctan(r)]=\frac{d\theta}{dr}=\frac{1}{\frac{dr}{d\theta}}=\frac{1}{\frac{d}{d\theta}[\tan\theta]}=\frac{1}{\sec^2\theta}=\frac{1}{1+\tan^2\theta}=\frac{1}{1+r^2}$, 最後要用自己原來的變量 r 來表示。

sec 的反函數 記作 **arcsec** (或 $\sec^{-1}$)。令 $r=\sec\theta$ 並限制 $\theta\in[0,\frac{\pi}{2})\cup(\frac{\pi}{2},\pi]$ 致使 一對一 而有反函數 $\theta=\operatorname{arcsec}(r)$, $|r|\geq 1$, 則 $\frac{d}{dr}[\operatorname{arcsec}(r)]=\frac{d\theta}{dr}=\frac{1}{\frac{dr}{d\theta}}=\frac{1}{\frac{d}{d\theta}[\sec\theta]}=\frac{1}{\sec\theta\tan\theta}=\frac{1}{+\sec\theta\sqrt{\sec^2\theta-1}}=\frac{1}{|r|\sqrt{r^2-1}}$, 最後要用自己原來的變量 r 來表示。(若 $\theta\in[0,\frac{\pi}{2})$, 則 $\sec\theta\geq 1,\tan\theta\geq 0$; 若 $\theta\in(\frac{\pi}{2},\pi]$, 則 $\sec\theta\leq -1,\tan\theta\leq 0$。所以都是$+\sec\theta\sqrt{\sec^2\theta-1}$)

記得我們上課時說的, $\begin{array}{l}\arcsin(r)+\arccos(r)=\frac{\pi}{2}\\ \arctan(r)+\operatorname{arccot}(r)=\frac{\pi}{2}\\ \operatorname{arcsec}(r)+\operatorname{arccsc}(r)=\frac{\pi}{2}\end{array}$ 的餘角關係嗎? 所以一個的 derivative 是另一個的 derivative 的反號。但是剩下的 derivatives 你還是應該要會推導:

$\theta=\arcsin(r)$,	$r\in[-1,1]$,	$\theta\in[\frac{-\pi}{2},\frac{\pi}{2}]$,	$\frac{d\theta}{dr}=\frac{1}{\sqrt{1-r^2}}$
$\theta=\arccos(r)$,	$r\in[-1,1]$,	$\theta\in[0,\pi]$,	$\frac{d\theta}{dr}=\frac{-1}{\sqrt{1-r^2}}$
$\theta=\arctan(r)$,	$r\in(-\infty,\infty)$,	$\theta\in(\frac{-\pi}{2},\frac{\pi}{2})$,	$\frac{d\theta}{dr}=\frac{1}{1+r^2}$
$\theta=\operatorname{arccot}(r)$,	$r\in(-\infty,\infty)$,	$\theta\in(0,\pi)$,	$\frac{d\theta}{dr}=\frac{-1}{1+r^2}$
$\theta=\operatorname{arcsec}(r)$,	$|r|\geq 1$,	$\theta\in[0,\frac{\pi}{2})\cup(\frac{\pi}{2},\pi]$,	$\frac{d\theta}{dr}=\frac{1}{|r|\sqrt{r^2-1}}$
$\theta=\operatorname{arccsc}(r)$,	$|r|\geq 1$,	$\theta\in[\frac{-\pi}{2},0)\cup(0,\frac{\pi}{2}]$,	$\frac{d\theta}{dr}=\frac{-1}{|r|\sqrt{r^2-1}}$