函數的凸凹性

法蘭克

1　函數的凸凹性

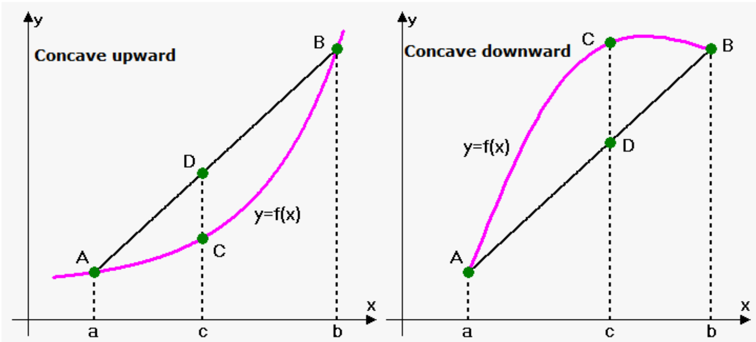
假設$y = f(x)$是定義在實數線上某個區間$[a, b]$的連續函數，並且f', f''在開區間(a, b)上可微分・如果任給$x_1, x_2 \in [a, b]$且$t \in [0, 1]$，恆有

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2),$$

我們稱此函數在$[a, b]$區間上凹向上(concave up)・如果不等式相反

$$f(tx_1 + (1-t)x_2) \geq tf(x_1) + (1-t)f(x_2),$$

我們稱此函數在$[a, b]$區間內凹向下(concave down)・



定理 1.1 如果$f''(x) > 0$, $(f''(x) < 0)$ $x \in (a, b)$則$y = f(x)$在$[a, b]$區間內凹向上(凹向下)・

令$x_1 < x_2$是$[a, b]$區間內任意給定的兩個數・我們希望證明函數是凹向上， 我們必須證明上面的第一個不等式・於是我們定義一個新的函數

$$g(t) = tf(x_1) + (1-t)f(x_2) - f(tx_1 + (1-t)x_2).$$

要證明此不等式，我們就是要證明此函數g在$[0, 1]$區間上非負$(g(t) \geq 0,\ t \in [0, 1])$・ 觀察一下$g(0) = g(1) = 0.$・利用Rolle定理，存在$c \in (0, 1)$使得$g'(c) = 0$・計算一下函數的一次微分

$$g'(t) = f(x_1) - f(x_2) - f'(tx_1 + (1-t)x_2)(x_1 - x_2),$$

與函數的二次微分：

$$g''(t) = -f''(tx_1 + (1-t)x_2)(x_1 - x_2)^2.$$

因為$x_1 \neq x_2$所以$(x_1 - x_2) > 0$.由於$f'' > 0$,所以馬上知道

$$g''(t) < 0, \quad t \in (0,1).$$

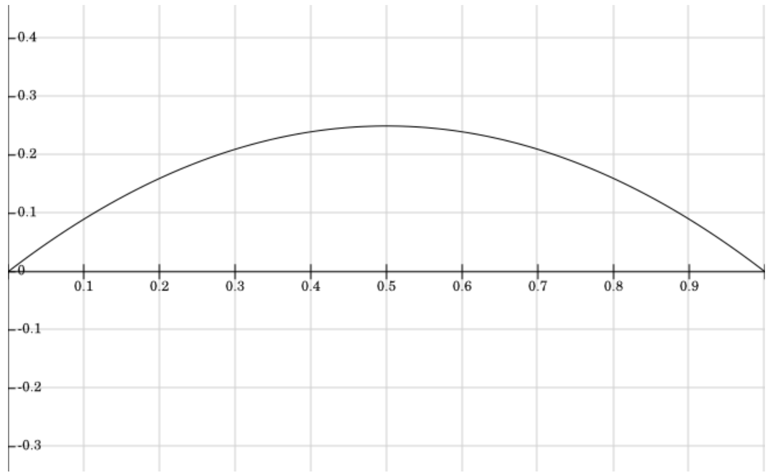
因為$g''(t) < 0$所以$g'(t)$在$[0,1]$區間上是遞減函數．所以在$0 < t < c$,

$$g'(t) > g(c) = 0.$$

所以在$0 < t < c$區間內，函數g是遞增函數．在$c < t < 1$區間內，

$$g'(t) < g'(c) = 0.$$

所以在$c < t < 1$區間內，函數g是遞減函數．如圖所示(此圖給出這類型函數一個範例)：



所以我們馬上知道在$0 < t < c$區間內，因為函數遞增，所以$g(t) > g(0) = 0$ 在$c < t < 1$區間內，因為函數遞減，所以$g(t) > g(1) = 0.$．於是我們證明了，對任意的$0 \leq t \leq 1$函數恆有 $g(t) \geq 0$，於是我們證明了不等式．

2 Jensen不等式

事實上，如果我們令$\lambda_1 = t$且$\lambda_2 = 1 - t$,那麼凹向上的函數，就可以改寫為

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2),$$

其中$\lambda_1, \lambda_2 \geq 0$且$\lambda_1 + \lambda_2 = 1$.

定理 2.1 假設$y = f(x)$在區間$[a,b]$上是凹向上的函數．則任給$x_1, \cdots, x_n \in [a,b]$且 $\lambda_1, \cdots, \lambda_n \geq 0$滿足$\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$, 恆有

$$f(\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \cdots + \lambda_n f(x_n).$$

附註：如果函數是凹向下的函數，不等號方向改變．

證明：我們假設此不等式對$n=2$成立，我們希望能推得$n=3$也成立．觀察發現

$$\lambda_1 x_1+\lambda_2 x_2+\lambda_3 x_3=(\lambda_1+\lambda_2)\left(\frac{\lambda_1}{\lambda_1+\lambda_2}x_1+\frac{\lambda_2}{\lambda_1+\lambda_2}x_2\right)+\lambda_3 x_3.$$

假設$\lambda_1,\lambda_2,\lambda_3\geq 0$且$\lambda_1+\lambda_2+\lambda_3=1$, $x_1,x_2,x_3\in[a,b]$．我們令$t=\lambda_1+\lambda_2$,則$\lambda_3=1-t$．再令$y=\dfrac{\lambda_1}{\lambda_1+\lambda_2}x_1+\dfrac{\lambda_2}{\lambda_1+\lambda_2}x_2$.則

$$\lambda_1 x_1+\lambda_2 x_2+\lambda_3 x_3=ty+(1-t)x_3.$$

如果$y=f(x)$在$[a,b]$上是凹向上的函數，則

$$\begin{aligned} f(\lambda_1 x_1+\lambda_2 x_2+\lambda_3 x_3)&=f(ty+(1-t)x_3)\\ &\leq tf(y)+(1-t)f(x_3)\\ &=tf(y)+\lambda_3 f(x_3). \end{aligned} \tag{2.1}$$

如果我們令$s=\dfrac{\lambda_1}{\lambda_1+\lambda_2}$則$\dfrac{\lambda_2}{\lambda_1+\lambda_2}=1-s$,於是

$$y=\frac{\lambda_1}{\lambda_1+\lambda_2}x_1+\frac{\lambda_2}{\lambda_1+\lambda_2}x_2=sx_1+(1-s)x_2.$$

再利用函數的性質：

$$f(y)=f(sx_1+(1-s)x_2)\leq sf(x_1)+(1-s)f(x_2).$$

所以利用上面這個不等式我們可以推得

$$\begin{aligned} tf(y)&=(\lambda_1+\lambda_2)f(y)\\ &\leq(\lambda_1+\lambda_2)\cdot(sf(x_1)+(1-s)f(x_2))\\ &=(\lambda_1+\lambda_2)\cdot\left(\frac{\lambda_1}{\lambda_1+\lambda_2}f(x_1)+\frac{\lambda_2}{\lambda_1+\lambda_2}f(x_2)\right)\\ &=\lambda_1 f(x_1)+\lambda_2 f(x_2). \end{aligned}$$

所以再結合不等式(2.1)，我們推得

$$f(\lambda_1 x_1+\lambda_2 x_2+\lambda_3 x_3)\leq\lambda_1 f(x_1)+\lambda_2 f(x_2)+\lambda_3 f(x_3).$$

所以，我們利用歸納法，假設不等式對$n=k-1$時成立．給定非負實數$\lambda_1,\cdots,\lambda_k$滿足$\sum_{i=1}^{k}\lambda_i=1$． 任取$[a,b]$上的$k$個點$x_1,\cdots,x_k$．令$t=\sum_{i=1}^{k-1}\lambda_i$則 $\lambda_k=1-t$．於是

$$\sum_{i=1}^{k}\lambda_i x_i=t\left(\sum_{i=1}^{k-1}\frac{\lambda_i}{t}x_i\right)+(1-t)x_k.$$

利用函數的凹向上性質：

$$\begin{aligned} f\left(\sum_{i=1}^{k}\lambda_i x_i\right)&=f\left(t\left(\sum_{i=1}^{k-1}\frac{\lambda_i}{t}x_i\right)+(1-t)x_k\right)\\ &\leq tf\left(\sum_{i=1}^{k-1}\frac{\lambda_i}{t}x_i\right)+(1-t)f(x_k)\\ &=tf\left(\sum_{i=1}^{k-1}\frac{\lambda_i}{t}x_i\right)+\lambda_k f(x_k). \end{aligned}$$

因為$\displaystyle\sum_{i=1}^{k-1}\frac{\lambda_i}{t}=1$,所以我們可以再次利用函數凹向上的性質：

$$f\left(\sum_{i=1}^{k-1}\frac{\lambda_i}{t}x_i\right)\le\sum_{i=1}^{k-1}\frac{\lambda_i}{t}f(x_i).$$

帶入上面的不等式後，發現

$$\begin{aligned}f\left(\sum_{i=1}^{k}\lambda_i x_i\right)&\le tf\left(\sum_{i=1}^{k-1}\frac{\lambda_i}{t}x_i\right)+\lambda_k f(x_k)\\&\le t\left(\sum_{i=1}^{k-1}\frac{\lambda_i}{t}f(x_i)\right)+\lambda_k f(x_k)=\left(\sum_{i=1}^{k-1}\lambda_i f(x_i)\right)+\lambda_k f(x_k)=\sum_{i=1}^{k}\lambda_i f(x_i).\end{aligned}$$

於是我們證明了不等式對$n=k$也成立，利用數學歸納法可知，不等式對任意的$n\ge 2$恆成立．

3 應用

我們可以利用函數的凸凹性來證明許多重要的不等式．

範例 3.1 假設A,B,C是三角形的三內角，證明

$$\sin A+\sin B+\sin C\le\frac{3\sqrt{3}}{2}.$$

由於A,B,C是三角形三內角，$A,B,C\in(0,\pi)$且$A+B+C=\pi$. 考慮函數$f(x)=\sin x$, $x\in[0,\pi]$.則$f'(x)=\cos x$, $f''(x)=-\sin x$. 所以在$(0,\pi)$上$f''(x)<0$．於是$f(x)$是凹向下的函數，因此利用Jensen不等式：

$$\begin{aligned}\frac{\sin A+\sin B+\sin C}{3}&=\frac{1}{3}(f(A)+f(B)+f(C))\\&\le f\left(\frac{A+B+C}{3}\right)\\&=\sin\frac{A+B+C}{3}=\sin\frac{\pi}{3}=\frac{\sqrt{3}}{2}.\end{aligned}$$

將不等式兩邊同乘3我們就得到想證明的式子．

定理 3.1 算幾不等式:假設$a_1,\cdots,a_n\ge 0$則

$$\frac{a_1+\cdots+a_n}{n}\ge\sqrt[n]{a_1\cdots a_n}.$$

如果a_i其中一個為零，則不等式顯然成立．於是我們假設$a_1, \cdots, a_n \neq 0$． 對$\sqrt[n]{a_1 \cdots a_n}$取ln後，我們發現

$$\ln(\sqrt[n]{a_1 \cdots a_n}) = \frac{1}{n}\left(\ln a_1 + \ln a_2 + \cdots + \ln a_n\right).$$

我們令函數

$$f(x) = \ln x, \quad x > 0.$$

則我們得到：

$$\sqrt[n]{a_1 \cdots a_n} = \frac{1}{n} f(a_1) + \cdots + \frac{1}{n} f(a_n).$$

計算函數的微分後，我們得到$f'(x) = 1/x$且$f''(x) = -1/x^2 < 0$．於是我們推得函數凹向下．利用上面的Jensen不等式，我們可以推得

$$f\left(\frac{1}{n}a_1 + \cdots + \frac{1}{n}a_n\right) \geq \frac{1}{n} f(a_1) + \cdots + \frac{1}{n} f(a_n).$$

這個不等式就是：

$$\ln\left(\frac{a_1 + \cdots + a_n}{n}\right) \geq \ln(\sqrt[n]{a_1 \cdots a_n}).$$

兩邊同取exp,我們就證明了算幾不等式．

定理 3.2 (Young不等式)假設$p, q > 0$且$\frac{1}{p} + \frac{1}{q} = 1$．則任給$a, b > 0$恆有

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

我們知道$a^p = e^{\ln a^p}$.．所以考慮函數$f(x) = e^x$, $x \in \mathbb{R}$.則 $f(\ln a^p) = a^p$且$f(\ln b^q) = b^q$.計算函數的二次微分可得$f''(x) = e^x > 0$，可知函數為凹向上．因此

$$f\left(\frac{1}{p}\ln a^p + \frac{1}{q}\ln b^q\right) \leq \frac{1}{p} f(\ln a^p) + \frac{1}{q} f(\ln b^p). \tag{3.1}$$

由於

$$\frac{1}{p}\ln a^p + \frac{1}{q}\ln b^q = \ln a + \ln b = \ln(ab).$$

所以

$$f\left(\frac{1}{p}\ln a^p + \frac{1}{q}\ln b^q\right) = f(\ln ab) = ab.$$

我們發現不等式(3.1)就是要證明的不等式．

利用Young不等式，我們可以推論下列不等式：

推論 3.1 (Holder不等式) 假設$a_1, \cdots, a_n$與$b_1, \cdots, b_n$是正實數，且p, q同上．則

$$(a_1 b_1 + \cdots + a_n b_n) \leq \left(a_1^p + \cdots + a_n^p\right)^{1/p} \left(b_1^q + \cdots + b_n^q\right)^{1/q}.$$

當$p = q = 2$時，就是算幾不等式．

證明：我們令$A_i = a_i / \left(a_1^p + \cdots + a_n^p\right)^{1/p}$, $1 \le i \le n$, 且$B_i = b_i / \left(b_1^q + \cdots + b_n^q\right)^{1/q}$.要證明上述不等式，我們只需證明

$$A_1 B_1 + \cdots + A_n B_n \le 1.$$

觀察一下發現：$A_1^p + A_2^p + \cdots + A_n^p = B_1^q + B_2^q + \cdots + B_n^q = 1$. 利用Young不等式，我們可以推得$A_i B_i \le \dfrac{A_i^p}{p} + \dfrac{B_i^q}{q}$,其中$1 \le i \le n$.我們把這個不等式兩邊相加後得到：

$$A_1 B_1 + \cdots + A_n B_n \le \frac{A_1^p + \cdots + A_n^p}{p} + \frac{B_1^q + \cdots + B_n^q}{q} = \frac{1}{p} + \frac{1}{q} = 1.$$

於是我們就證明了Holder不等式．