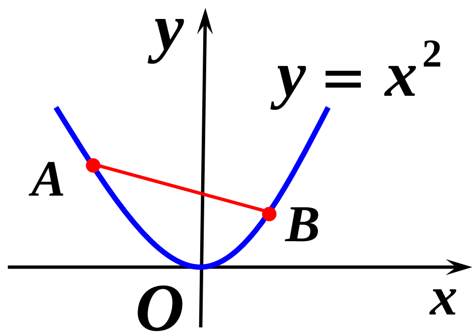
数学分析　第六章
微分中值定理及其应用

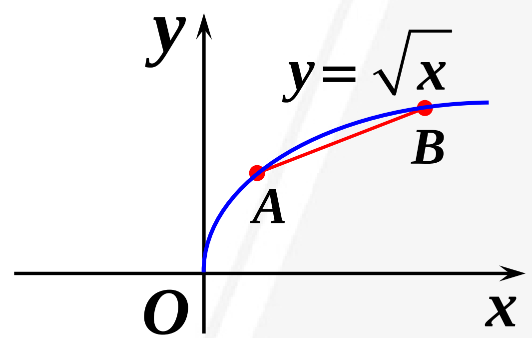
§5 函数的凸性与拐点

第二十讲

函数的凸性，詹森不等式

从两个熟悉的函数 $y=x^2$ 与 $y=\sqrt{x}$ 的图像来看凸性的不同：





$y=x^2(y=\sqrt{x})$上任取两点A,B，弦 $\overline{AB}$ 恒在曲线段 $\overset{\frown}{AB}$ 的上方(下方)．

定义1

设 f 为区间 I 上的函数. 若对于 I 上的任意两点 x_1, x_2 和任意实数 $\lambda\in(0,1)$, 总有

$$f(\lambda x_1+(1-\lambda)x_2)\le\lambda f(x_1)+(1-\lambda)f(x_2),\quad (1)$$

则称 f 为 I 上的一个凸函数. 反之如果总有

$$f(\lambda x_1+(1-\lambda)x_2)\ge\lambda f(x_1)+(1-\lambda)f(x_2),\quad (2)$$

则称 f 为 I 上的一个凹函数.

如(1)和(2)式中的不等号改为严格不等号, 则相应的函数称为严格凸函数和严格凹函数.

由此可得 $y = x^2$ 在 $(-\infty, +\infty)$ 上为严格的凸函数，

$y = \sqrt{x}$ 为 $[0, +\infty)$ 上的严格凹函数 .

很明显，若 $f(x)$ 为(严格)的凸函数, 那么 $-f(x)$ 就

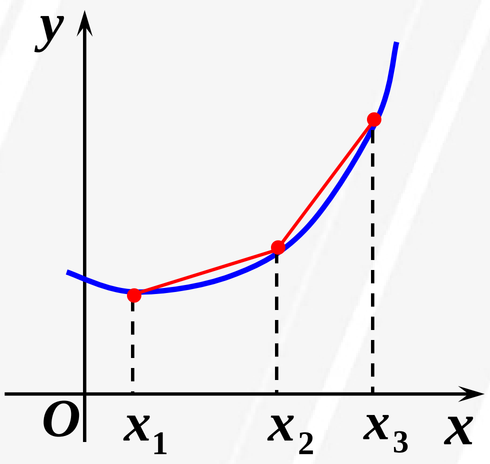
为(严格) 凹函数，反之亦然.

引理 $f(x)$为区间 I上的凸函数的充要条件是：对于I中的任意三点$x_1 < x_2 < x_3$，有

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \le \frac{f(x_3)-f(x_2)}{x_3-x_2}. \qquad (3)$$

证（必要性） 设 $\lambda = \dfrac{x_3-x_2}{x_3-x_1}$，于是

$$x_2 = \lambda x_1 + (1-\lambda)x_3.$$



因为$f(x)$为 I 上的凸函数，所以

$$f(x_2) = f(\lambda x_1 + (1-\lambda)x_3)$$

$$\le \lambda f(x_1) + (1-\lambda) f(x_3)$$

$$= \frac{x_3-x_2}{x_3-x_1} f(x_1) + \frac{x_2-x_1}{x_3-x_1} f(x_3).$$

从而有

$$(x_3-x_1)f(x_2)\le(x_3-x_2)f(x_1)+(x_2-x_1)f(x_3),$$

即

$$\boxed{\begin{aligned}&(x_3-x_2)f(x_2)+(x_2-x_1)f(x_2)\\&\quad\le(x_3-x_2)f(x_1)+(x_2-x_1)f(x_3),\end{aligned}}$$

整理后即为 (3) 式.

（充分性）对于任意 $x_1<x_3,\lambda\in(0,1)$. 设

$$x_2=\lambda x_1+(1-\lambda)x_3,$$

则

$$\frac{f(x_2)-f(x_1)}{x_2-x_1}\le\frac{f(x_3)-f(x_2)}{x_3-x_2}.$$

由于必要性的证明是可逆的，从而得到

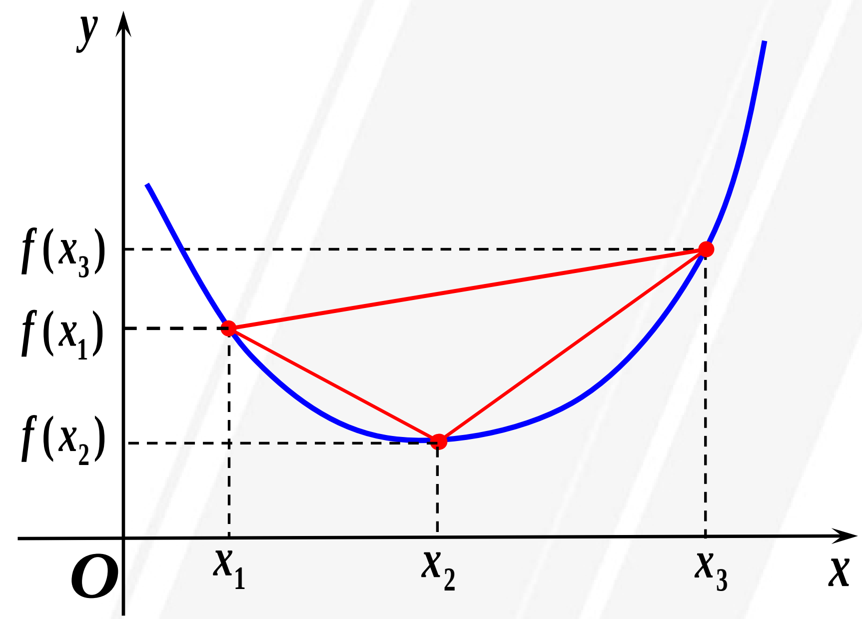
$$f(\lambda x_1+(1-\lambda)x_3)\le\lambda f(x_1)+(1-\lambda)f(x_3),$$

所以 f 为 I 上的凸函数.

同理可证 f 为 I 上的凸函数的充要条件是：对于 I 中的任意三点 $x_1 < x_2 < x_3$ ，有

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \le \frac{f(x_3)-f(x_1)}{x_3-x_1} \le \frac{f(x_3)-f(x_2)}{x_3-x_2}. \quad (4)$$

注 (4) 式与 (1) 式是等价的. 所以有些教材将 (4) 式作为凸函数的定义.



对于凹函数，请大家自行写出相应的定理.

由数学归纳法不难证明：f 为 I 上的凸函数充要条件是：任给 $x_1,\cdots,x_n \in I, 0<\lambda_i<1,\ i=1,2,\cdots,n,$ $\lambda_1+\lambda_2+\cdots+\lambda_n=1$, 必有

$$f(\lambda_1 x_1+\cdots+\lambda_n x_n) \le \lambda_1 f(x_1)+\cdots+\lambda_n f(x_n).$$

这是著名的**詹森不等式**.

詹森(Jensen,J.L. 1859-1925,丹麦)

特别取 $\lambda_i = \dfrac{1}{n}$, 有

$$f\left(\frac{x_1+x_2+\cdots+x_n}{n}\right) \le \frac{1}{n}\left[f(x_1)+f(x_2)+\cdots+f(x_n)\right],$$

即：

$$f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \le \frac{1}{n}\sum_{i=1}^{n} f(x_i) \qquad (5)$$

(5) 式是凸函数最常用的不等式 .

下面举例说明凸函数的内在性质.

例 1 设 f 为开区间 $(a,\ b)$ 上的凸函数, 那么它在 (a, b) 中每一点的左、右导数存在.

证 对于任意的 $x_0 \in (a, b), 0 < h_1 < h_2$,使

$$x_0 < x_0 + h_1 < x_0 + h_2 < b,$$

由引理的(4)式得到

$$\frac{f(x_0+h_1)-f(x_0)}{h_1} \le \frac{f(x_0+h_2)-f(x_0)}{h_2}.$$

令$F(h)=\dfrac{f(x_0+h)-f(x_0)}{h}$, 则$F(h)$在$(0,b-x_0)$上递增.

取$x' \in (a,b), x' < x_0$, 由引理又得

$$\frac{f(x_0)-f(x')}{x_0-x'} \le \frac{f(x_0+h)-f(x_0)}{h},\ h\in(0,b-x_0).$$

因为左端是一个定值，所以$F(h)$有下界.

根据第三章的相关定理，下面极限存在：

$$\lim_{h\to 0^+} F(h)=\lim_{h\to 0^+}\frac{f(x_0+h)-f(x_0)}{h}=f'_+(x_0).$$

$f'_-(x_0)$的存在性留作练习.

注 开区间上的凸函数左右导数处处存在，但不一定处处可导.

例子同时证明了：开区间上凸函数是连续函数.