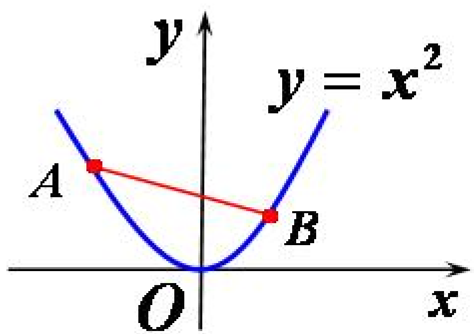
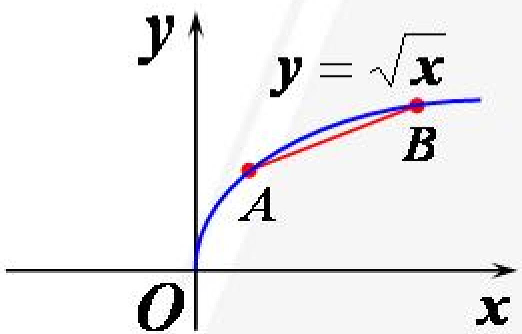
§5　函数的凸性与拐点

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





$y=x^2(y=\sqrt{x})$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

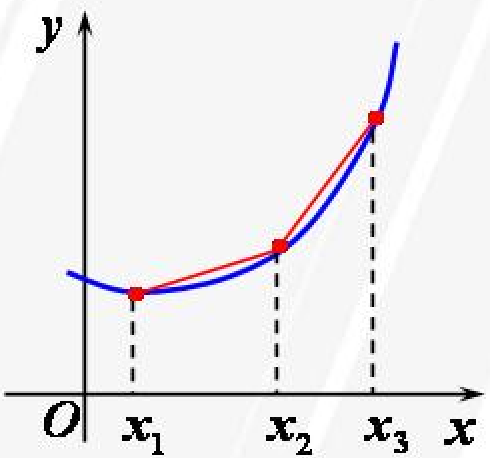
为(严格) 凹函数，反之亦然.

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

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

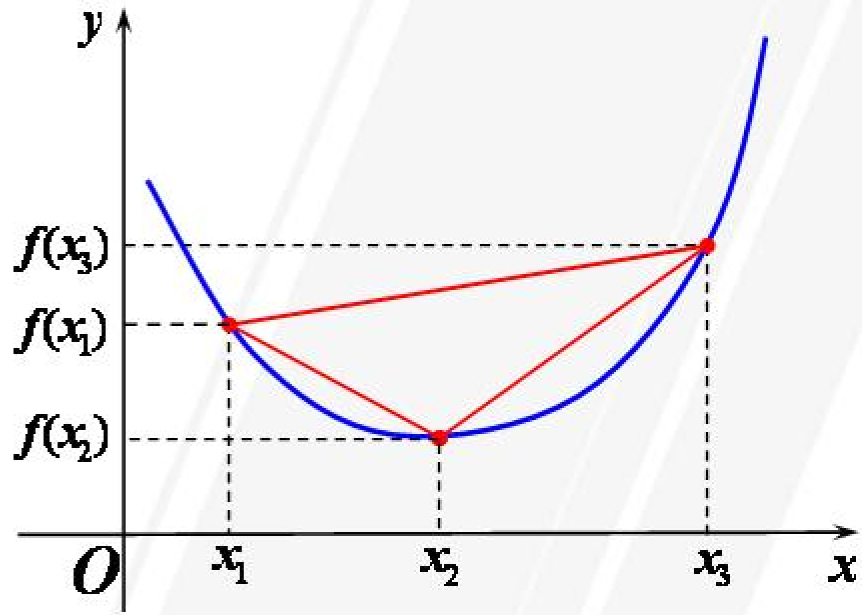
$$f(\lambda x_1+(1-\lambda)x_3)\le\lambda f(x_1)+(1-\lambda)f(x_3),$$

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，请读者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

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

即：

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

(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. 特别是在 (a,b) 上处处连续.

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

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

这就证明了$F(h)$有下界. 所以

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

同理可证 $f'_-(x_0)$ 存在.

注 开区间上的凸函数处处连续, 但不一定处处可导; 闭区间上的凸函数在端点不一定连续.

定理6.14

设 f 为区间 I 上的可导函数，则下述论断互相等价：

(i) $f(x)$ 为 I 上的凸函数；

(ii) $f'(x)$ 为 I 上的增函数；

(iii) 对于 I 上的任意两点 x_1, x_2，有

$$f(x_2) \geq f(x_1) + f'(x_1)(x_2 - x_1).$$

注 (iii) 中的不等式表示切线恒在凸曲线的下方.

证 (i) $\Rightarrow$ (ii) 任取 $x_1, x_2 \in I$ 和正数 h，使

$$x_1 < x_2,\ 且\ x_1 - h \in I,\ x_2 + h \in I.$$

已知 f 是凸函数，由(4)式

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

$$\le \frac{f(x_2 + h) - f(x_2)}{h}.$$

令 $h \to 0^+$，因为

$$\lim_{h \to 0^+} \frac{f(x_1) - f(x_1 - h)}{h} = f'_-(x_1) = f'(x_1),$$

$$\lim_{h \to 0^+} \frac{f(x_2 + h) - f(x_2)}{h} = f'_+(x_2) = f'(x_2),$$

所以 $$f'(x_1) \le \frac{f(x_2) - f(x_1)}{x_2 - x_1} \le f'(x_2),$$

故 $f'(x)$ 递增.

(ii) $\Rightarrow$ (iii) 对于任意 $x_1, x_2 \in I$, 不妨设 $x_1 < x_2$.

则 $f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1) \quad x_1 < \xi < x_2$.

因为 $f'(x)$ 递增, 所以

$$f(x_2) \geq f(x_1) + f'(x_1)(x_2 - x_1).$$

(ii) $f'(x)$ 为 I 上的增函数;

(iii) 对于 I 上的任意两点 x_1, x_2, 有

$$f(x_2) \geq f(x_1) + f'(x_1)(x_2 - x_1).$$

(iii) $\Rightarrow$ (i) 仍设$x_1 < x_2$，$x_0 = \lambda x_1 + (1-\lambda)x_2$ $(0 < \lambda < 1)$, 则

$$f(x_1) \geq f(x_0) + f'(x_0)(x_1 - x_0), \qquad (6)$$

$$f(x_2) \geq f(x_0) + f'(x_0)(x_2 - x_0). \qquad (7)$$

将(6)式乘以λ, (7) 式乘以$(1-\lambda)$作和，并注意到

$\lambda x_1 + (1-\lambda)x_2 - x_0 = 0$, 得

$$\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(x_0) = f(\lambda x_1 + (1-\lambda)x_2)$$

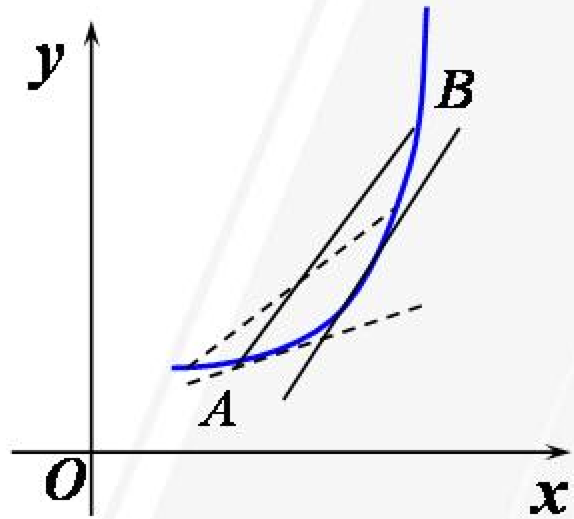
(iii) 对于 I 上的任意两点 x_1, x_2, 有

$$f(x_2) \geq f(x_1) + f'(x_1)(x_2 - x_1).$$

(i) $f(x)$ 为 I 上的凸函数；

我们在这里再一次强调，

函数 f 是凸函数的几何意义是：曲线 $y = f(x)$ 的弦位于相应曲线段的上方；而它的切线位于曲线的下方.



我们在定理中列出了凸函数的三个等价性质. 对于凹函数也有类似的性质，请大家写出相应的定理.

定理6.15

设 $f(x)$ 在区间 I 上二阶可导, 则 $f(x)$在区间I上是凸(凹)函数的充要条件为:

$$f''(x) \geq 0 \quad (f''(x) \leq 0).$$

证 由定理 6.14 立即可得.

例 2 讨论函数 $f(x)=\arctan x$ 的凹凸性区间.

解 因为

$$f'(x)=\frac{1}{1+x^2},\quad x\in(-\infty,+\infty),$$

$$f''(x)=\frac{-2x}{(1+x^2)^2},\ x\in(-\infty,+\infty).$$

所以当 $x\in(-\infty,0)$时, $f''(x)>0$, $f(x)$ 为凸函数;

当 $x\in(0,+\infty)$时, $f''(x)<0$, $f(x)$ 为凹函数 .

例 3 设函数 $f(x)$为 (a, b) 上的可导凸(凹)函数. 那么$f'(x_0)=0$的充要条件是 x_0 为$f(x)$的极值点. (本例说明：在凸(凹)函数的条件下，可微函数的极值点与稳定点是等价的.)

证 充分性是显然的(费马定理). 下面证明必要性. 设 $f(x)$是凸函数, x_0 是 $f(x)$ 的稳定点, 即$f'(x_0)=0$. 由定理 6.14 的 (ii), $f'(x)$是递增的. 所以

(i) 当$x\in(a, x_0)$时，$f'(x)\le 0$, $f(x)$ 是递减的, 故

$$f(x)\ge f(x_0),\ x\in(a, x_0);$$

(ii) 当 $x \in (x_0, b)$ 时, $f'(x) \geq 0$, $f(x)$ 是递增的, 故

$$f(x) \geq f(x_0),\ x \in (x_0, b).$$

综上, $f(x) \geq f(x_0),\ x \in (a, b)$.

即 $f(x_0)$ 是 $f(x)$ 的极小值.

注 我们实际上已经证明, 对于可微凸函数，其极值总是极小值，可微凹函数的极值总是极大值.

因此下面这个例题自然就产生了.

例 4 设 $f(x)$ 是区间 (a, b) 上的一个严格凸函数，若 $f(x_0)$是 $f(x)$ 的一个极值，则 $f(x)$ 仅有惟一的极值，并且是极小值.

证 应当注意，这里并没有假设函数 $f(x)$ 的可微性，所以例 3 的方法就失效了.

因为 $f(x)$ 严格凸，所以当 $x_1 < x_0 < x_2$ 时，

$$\frac{f(x_0)-f(x_1)}{x_0-x_1} < \frac{f(x_2)-f(x_0)}{x_2-x_0}.$$

由于 $f(x_0)$ 是极值，因此当 x_1, x_2充分接近 x_0时，有

$$[f(x_0)-f(x_1)]\cdot[f(x_2)-f(x_0)] \le 0.$$

又因 $f(x)$ 是严格凸的，所以

$$f(x_0)-f(x_1)\le 0,\ f(x_2)-f(x_0)\ge 0,$$

即 $f(x_0)$是极小值 .

对于任意 $x\in(x_0,b)$, 因为 $f(x_0)$ 是极小值，　所以

存在 $x_1\in(x_0,x)$, 使得

$$f(x_1)\ge f(x_0).$$

又因为 $f(x)$ 是严格凸函数，所以

$$0\le\frac{f(x_1)-f(x_0)}{x_1-x_0}<\frac{f(x)-f(x_0)}{x-x_0},$$

即 $f(x)>f(x_0)$.

同理可证：对于任意 $x\in(a,x_0)$, 仍有 $f(x_0)>f(x)$.

设 $f(x)$ 有另一极小值 $f(x^*)$. 根据以上讨论，把 x^* 和 x_0 分别看作极值点时, 有

$$f(x_0) > f(x^*) \text{ 和 } f(x^*) > f(x_0)$$

同时成立, 矛盾. 所以极值点惟一.

例 5 证明不等式 $(abc)^{\frac{a+b+c}{3}} \le a^a b^b c^c$，其中 a,b,c 均为正数.

证 设 $f(x) = x\ln x$，则

$$f'(x) = \ln x + 1,\ f''(x) = \frac{1}{x} > 0,$$

所以 $f(x)$ 在 $x > 0$ 时为严格凸的. 由詹森不等式

$$f\left(\frac{a+b+c}{3}\right) \le \frac{1}{3}(f(a) + f(b) + f(c)),$$

即

$$\frac{a+b+c}{3}\ln\frac{a+b+c}{3} \le \frac{1}{3}\ln a^a b^b c^c.$$

又因

$$\sqrt[3]{abc} \le \frac{a+b+c}{3},$$

故有

$$\frac{a+b+c}{3}\ln\sqrt[3]{abc} \le \frac{1}{3}\ln(a^a b^b c^c).$$

再由对数函数是严格增的，就证得

$$(abc)^{\frac{a+b+c}{3}} \le a^a b^b c^c.$$

*例 6 设$a>0,b>0,p>0,q>0,\frac{1}{p}+\frac{1}{q}=1$. 求证

$$ab<\frac{1}{p}a^p+\frac{1}{q}b^q.$$

证 设$f(x)=\ln x$. 因$f''(x)<0$, 故 $f(x)$是$x>0$上的严格凹函数，所以有

$$f\left(\frac{1}{p}a^p+\frac{1}{q}b^q\right)>\frac{1}{p}f(a^p)+\frac{1}{q}f(b^q)$$

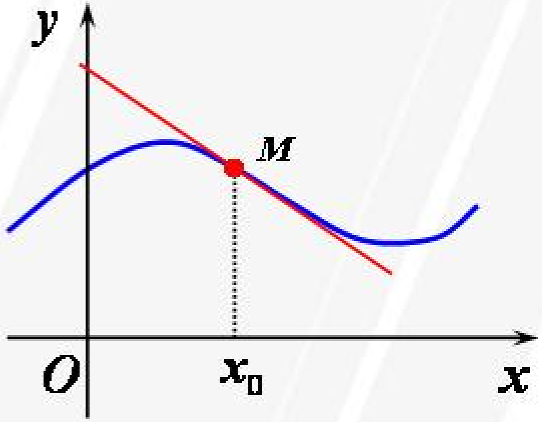
$$=\frac{1}{p}\ln a^p+\frac{1}{q}\ln b^q=\ln ab.$$

即 $ab<\dfrac{1}{p}a^p+\dfrac{1}{q}b^q.$

定义2

设曲线 $y = f(x)$ 在点 $M(x_0, f(x_0))$ 处有穿过曲线的切线，并且切线的两侧分别是严格凸和严格凹的，这时称点 M 为曲线 $y = f(x)$ 的拐点.

图中所示的 M 是一个拐点.



又如点 $(0, 0)$ 是曲线 $y = \arctan x$ 的一个拐点；而余弦曲线 $y = \cos x$ 的所有拐点为 $\left(k\pi + \dfrac{\pi}{2},\ 0\right)$，其中 $k \in \mathbf{Z}$.

下面两个定理是显然的.

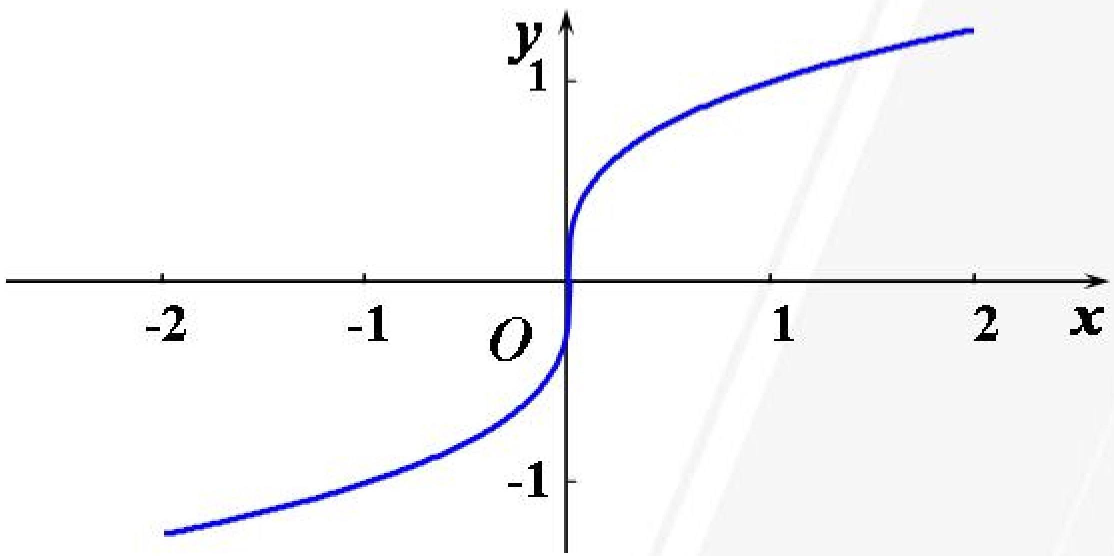
定理6.16

若 $f(x)$在点 x_0二阶可导,则$(x_0, f(x_0))$为曲线 $y = f(x)$的拐点的必要条件是 $f''(x_0) = 0$.

定理6.17

设 $f(x)$在点 x_0 可导，在$U^\circ(x_0)$ 二阶可导.
若 $f''(x)$在$U_+^\circ(x_0)$，$U_-^\circ(x_0)$的符号相反,则 $(x_0, f(x_0))$ 是 $f(x)$的一个拐点 .

应当注意：若$(x_0, f(x_0))$是曲线$y = f(x)$的一个拐点，那么f在点x_0的导数不一定存在. 比如函数 $y = \sqrt[3]{x}$ 在 $x = 0$ 的导数不存在，但根据定义2，点(0, 0) 却是曲线 $y = \sqrt[3]{x}$ 的一个拐点 .



复习思考题

1. 两个凸函数的乘积是否是凸函数 ?

2. 两个凸函数的复合是否是凸函数 ?

3. 任选一个凸函数, 利用詹森不等式构造出新的不等式.