凸函数

徐海峰整理

October 27, 2024

数学科学学院, 扬州大学

此幻灯片的教学内容面向非数学专业的理工科本科生.

内容基于下面的教材:

梅加强编著《数学分析》, 高等教育出版社.

其他参考文献.

[1] 蒋国强、蔡蕃、张兴龙、汤进龙、孟国明、俞皓等编《高等数学》.

凸函数的引入

设 f 为区间 $I=[a,b]$ 上的二阶可导函数, 且 $f''>0$, 则 f 在区间内部取不到极大值.

设 f 为区间 $I = [a, b]$ 上的二阶可导函数, 且 $f'' > 0$, 则 f 在区间内部取不到极大值.进一步的, 可以证明 f 的图像位于过 $(a, f(a))$, $(b, f(b))$ 的直线下方.

设 f 为区间 $I=[a,b]$ 上的二阶可导函数, 且 $f''>0$, 则 f 在区间内部取不到极大值.进一步的, 可以证明 f 的图像位于过 $(a,f(a))$, $(b,f(b))$ 的直线下方.

凸函数

凸函数的定义

定义

设 f 为区间 I 上定义的函数. 如果对任意 $a < b \in I$, 均有

$$f(x) \leqslant \ell(x), \quad \forall\, x \in (a, b),$$

其中 $\ell(x)$ 是过点 $(a, f(a))$, $(b, f(b))$ 的直线方程,

$$\ell(x) = \frac{f(b) - f(a)}{b - a}(x - a) + f(a),$$

则称 f 为 I 上的**凸函数**.

导出凸函数的等价定义

上面直线方程 $\ell(x)$ 也可写为

$$\ell(x) = f(b) - \frac{f(b) - f(a)}{b - a}(b - x).$$

上面直线方程 $\ell(x)$ 也可写为

$$\ell(x) = f(b) - \frac{f(b) - f(a)}{b - a}(b - x).$$

从而 $f(x) \leqslant \ell(x)$ 可改写为

$$\frac{f(x) - f(a)}{x - a} \leqslant \frac{f(b) - f(a)}{b - a} \leqslant \frac{f(b) - f(x)}{b - x}, \quad \forall\, x \in (a, b).$$

上面直线方程 $\ell(x)$ 也可写为

$$\ell(x) = f(b) - \frac{f(b) - f(a)}{b - a}(b - x).$$

从而 $f(x) \leqslant \ell(x)$ 可改写为

$$\frac{f(x) - f(a)}{x - a} \leqslant \frac{f(b) - f(a)}{b - a} \leqslant \frac{f(b) - f(x)}{b - x}, \quad \forall\, x \in (a, b).$$

上式等价于

$$\frac{f(x) - f(a)}{x - a} \leqslant \frac{f(b) - f(x)}{b - x}, \quad \forall\, x \in (a, b).$$

上面直线方程 $\ell(x)$ 也可写为

$$\ell(x) = f(b) - \frac{f(b) - f(a)}{b - a}(b - x).$$

从而 $f(x) \leqslant \ell(x)$ 可改写为

$$\frac{f(x) - f(a)}{x - a} \leqslant \frac{f(b) - f(a)}{b - a} \leqslant \frac{f(b) - f(x)}{b - x}, \quad \forall\, x \in (a, b).$$

上式等价于

$$\frac{f(x) - f(a)}{x - a} \leqslant \frac{f(b) - f(x)}{b - x}, \quad \forall\, x \in (a, b).$$

几何意义就是, 若记 $A = (a, f(a))$, $B = (b, f(b))$, $P = (x, f(x))$, 则 $f(x)$ 是凸函数当且仅当对 A, B 之间曲线上任意点 P, 直线 PA 的斜率小于等于 PB 的斜率.

当 $x \in [a, b]$ 时, x 可以写为

$$x = ta + (1-t)b, \quad t = \frac{b-x}{b-a} \in [0, 1].$$

当 $x \in [a, b]$ 时, x 可以写为

$$x = ta + (1-t)b, \quad t = \frac{b-x}{b-a} \in [0,1].$$

于是 $x - a = ta + (1-t)b - a = (1-t)(b-a)$此时 $l(x)$ 可以写为

$$\begin{aligned} l(x) &= \frac{f(b)-f(a)}{b-a}(x-a) + f(a) = \frac{f(b)-f(a)}{b-a}(1-t)(b-a) + f(a) \\ &= (f(b)-f(a))(1-t) + f(a) = tf(a) + (1-t)f(b). \end{aligned}$$

当 $x \in [a, b]$ 时, x 可以写为

$$x = ta + (1-t)b, \quad t = \frac{b-x}{b-a} \in [0, 1].$$

于是 $x - a = ta + (1-t)b - a = (1-t)(b-a)$此时 $l(x)$ 可以写为

$$\begin{aligned} l(x) &= \frac{f(b)-f(a)}{b-a}(x-a) + f(a) = \frac{f(b)-f(a)}{b-a}(1-t)(b-a) + f(a) \\ &= (f(b)-f(a))(1-t) + f(a) = tf(a) + (1-t)f(b). \end{aligned}$$

从而 $f(x) \leqslant l(x), \forall x \in (a, b)$ 可以改写为

$$f(ta + (1-t)b) \leqslant tf(a) + (1-t)f(b), \quad \forall t \in (0, 1).$$

凸函数与 Young 不等式

Young 不等式

命题

设 $a, b > 0$, $p > 1$, $q = \frac{p}{p-1}$, 则有

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

Young 不等式

命题

设 $a, b > 0,\ p > 1,\ q = \frac{p}{p-1}$, 则有

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

注

由条件, p 和 q 满足 $\frac{1}{p} + \frac{1}{q} = 1$.

Young 不等式

命题

设 $a, b > 0$, $p > 1$, $q = \frac{p}{p-1}$, 则有

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

注

由条件, p 和 q 满足 $\frac{1}{p} + \frac{1}{q} = 1$.

下面我们利用 e^x 是凸函数来证明 Young 不等式.

Proof.

考虑指数函数 $f(x) = e^x$,

Proof.

考虑指数函数 $f(x) = e^x$,由 $f''(x) = e^x > 0$, 可知 f 为(严格)凸函数.

Proof.

考虑指数函数 $f(x) = e^x$,由 $f''(x) = e^x > 0$, 可知 f 为(严格)凸函数.于是, 当 $a, b > 0$, p, q 满足条件 $p, q > 1$ 且 $\frac{1}{p} + \frac{1}{q} = 1$ 时, 有

$$
\begin{aligned}
ab = e^{\ln(ab)} = e^{\ln a + \ln b} &= e^{\frac{1}{p}\ln a^p + \frac{1}{q}\ln b^q} \\
&\leqslant \frac{1}{p}e^{\ln a^p} + \frac{1}{q}e^{\ln b^q} = \frac{a^p}{p} + \frac{b^q}{q}.
\end{aligned}
$$

□

Proof.

考虑指数函数 $f(x)=e^x$,由 $f''(x)=e^x>0$, 可知 f 为(严格)凸函数.于是, 当 $a,b>0$, p,q 满足条件 $p,q>1$ 且 $\frac{1}{p}+\frac{1}{q}=1$ 时, 有

$$\begin{aligned} ab &= e^{\ln(ab)} = e^{\ln a+\ln b} = e^{\frac{1}{p}\ln a^p+\frac{1}{q}\ln b^q} \\ &\leqslant \frac{1}{p}e^{\ln a^p}+\frac{1}{q}e^{\ln b^q} = \frac{a^p}{p}+\frac{b^q}{q}. \end{aligned}$$

□

更一般地, 有 Jensen 不等式.

Jensen 不等式

定理

设 f 是定义在区间 I 上的函数, 则 f 为凸函数当且仅当对任意的 $x_i \in I$, $\lambda_i \geqslant 0$ $(i = 1, 2, \ldots, n)$, $\sum_{i=1}^{n} \lambda_i = 1$, 有

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i).$$

续

Proof.

只需证明必要性.

Proof.

只需证明必要性.对 n 用数学归纳法.

续

Proof.

只需证明必要性.对 n 用数学归纳法.$n = 1$ 是显然的.

续

Proof.

只需证明必要性.对 n 用数学归纳法.$n=1$ 是显然的.$n=2$ 前面已经讨论过.

Proof.

只需证明必要性.对 n 用数学归纳法.$n=1$ 是显然的.$n=2$ 前面已经讨论过.假设不等式对 $n=k$ 成立, 则当 $n=k+1$ 时, 不妨假设 $0<\lambda_{k+1}<1$, 此时

$$\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} = 1.$$

续

Proof.

只需证明必要性.对 n 用数学归纳法.$n = 1$ 是显然的.$n = 2$ 前面已经讨论过.假设不等式对 $n = k$ 成立, 则当 $n = k + 1$ 时, 不妨假设 $0 < \lambda_{k+1} < 1$, 此时

$$\sum_{i=1}^{k} \frac{\lambda_i}{1 - \lambda_{k+1}} = 1.$$

□

续

Proof.

由归纳假设, 有

$$
\begin{aligned}
f\Big(\sum_{i=1}^{k+1}\lambda_i x_i\Big) &= f\Big((1-\lambda_{k+1})\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}x_i+\lambda_{k+1}x_{k+1}\Big)\\
&\leqslant (1-\lambda_{k+1})f\Big(\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}x_i\Big)+\lambda_{k+1}f(x_{k+1})\\
&\leqslant (1-\lambda_{k+1})\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}f(x_i)+\lambda_{k+1}f(x_{k+1})\\
&= \sum_{i=1}^{k+1}\lambda_i f(x_i).
\end{aligned}
$$

续

Proof.

由归纳假设, 有

$$
\begin{aligned}
f\Big(\sum_{i=1}^{k+1}\lambda_i x_i\Big) &= f\Big((1-\lambda_{k+1})\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}x_i+\lambda_{k+1}x_{k+1}\Big)\\
&\leqslant (1-\lambda_{k+1})f\Big(\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}x_i\Big)+\lambda_{k+1}f(x_{k+1})\\
&\leqslant (1-\lambda_{k+1})\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}f(x_i)+\lambda_{k+1}f(x_{k+1})\\
&= \sum_{i=1}^{k+1}\lambda_i f(x_i).
\end{aligned}
$$

这说明不等式对 $n=k+1$ 也成立, 从而定理得证. □

凸函数的性质

命题 (连续性质)

设 f 为区间 I 上的凸函数, 如果 $[a, b]$ 为 I 中不含 I 的端点的闭区间, 则 f 为 $[a, b]$ 上的 *Lipschitz* 函数. 特别地, 凸函数在区间内部总是连续的.

命题 (连续性质)

设 f 为区间 I 上的凸函数, 如果 $[a, b]$ 为 I 中不含 I 的端点的闭区间, 则 f 为 $[a, b]$ 上的 *Lipschitz* 函数. 特别地, 凸函数在区间内部总是连续的.

Proof.

由条件, 可取 $a', b' \in I$, 使得 $a' < a$, $b < b'$.

命题 (连续性质)

设 f 为区间 I 上的凸函数, 如果 $[a, b]$ 为 I 中不含 I 的端点的闭区间, 则 f 为 $[a, b]$ 上的 *Lipschitz* 函数. 特别地, 凸函数在区间内部总是连续的.

Proof.

由条件, 可取 $a', b' \in I$, 使得 $a' < a$, $b < b'$.任取 $x < y \in [a, b]$, 因为 f 为凸函数, 故有下列不等式

$$\frac{f(a') - f(a)}{a' - a} \leqslant \frac{f(a') - f(x)}{a' - x} \leqslant \frac{f(y) - f(x)}{y - x} \leqslant \frac{f(b') - f(y)}{b' - y} \leqslant \frac{f(b') - f(b)}{b' - b}.$$

命题 (连续性质)

设 f 为区间 I 上的凸函数, 如果 $[a, b]$ 为 I 中不含 I 的端点的闭区间, 则 f 为 $[a, b]$ 上的 *Lipschitz* 函数. 特别地, 凸函数在区间内部总是连续的.

Proof.

由条件, 可取 $a', b' \in I$, 使得 $a' < a$, $b < b'$.任取 $x < y \in [a, b]$, 因为 f 为凸函数, 故有下列不等式

$$\frac{f(a') - f(a)}{a' - a} \leqslant \frac{f(a') - f(x)}{a' - x} \leqslant \frac{f(y) - f(x)}{y - x} \leqslant \frac{f(b') - f(y)}{b' - y} \leqslant \frac{f(b') - f(b)}{b' - b}.$$

(注意这些不等式成立的前提是 a, b 不是区间 I 的端点.)

命题 (连续性质)

设 f 为区间 I 上的凸函数, 如果 $[a, b]$ 为 I 中不含 I 的端点的闭区间, 则 f 为 $[a, b]$ 上的 *Lipschitz* 函数. 特别地, 凸函数在区间内部总是连续的.

Proof.

由条件, 可取 $a', b' \in I$, 使得 $a' < a$, $b < b'$.任取 $x < y \in [a, b]$, 因为 f 为凸函数, 故有下列不等式

$$\frac{f(a') - f(a)}{a' - a} \leqslant \frac{f(a') - f(x)}{a' - x} \leqslant \frac{f(y) - f(x)}{y - x} \leqslant \frac{f(b') - f(y)}{b' - y} \leqslant \frac{f(b') - f(b)}{b' - b}.$$

(注意这些不等式成立的前提是 a, b 不是区间 I 的端点.)令

$$M = \max\left\{\left|\frac{f(a') - f(a)}{a' - a}\right|, \left|\frac{f(b') - f(b)}{b' - b}\right|\right\}$$

□

Proof.

则不等式意味着

$$|f(x) - f(y)| \leqslant M|x - y|, \quad \forall\, x, y \in [a, b].$$

续

Proof.

则不等式意味着

$$|f(x) - f(y)| \leqslant M|x - y|, \quad \forall\, x, y \in [a, b].$$

这说明 f 在 $[a, b]$ 上为 Lipschitz 函数. $\square$

连续函数的凸性判定

命题

设 f 为区间 I 上的连续函数，则 f 为凸函数当且仅当对任意 $x_1 < x_2 \in I$，有

$$f\left(\frac{x_1 + x_2}{2}\right) \leqslant \frac{1}{2}[f(x_1) + f(x_2)]. \quad \text{（平均值不等式）}$$

连续函数的凸性判定

命题

设 f 为区间 I 上的连续函数, 则 f 为凸函数当且仅当对任意 $x_1 < x_2 \in I$, 有

$$f\left(\frac{x_1 + x_2}{2}\right) \leqslant \frac{1}{2}[f(x_1) + f(x_2)]. \quad \text{（平均值不等式）}$$

Proof.

只要证明充分性即可.

连续函数的凸性判定

命题

设 f 为区间 I 上的连续函数, 则 f 为凸函数当且仅当对任意 $x_1 < x_2 \in I$, 有

$$f\left(\frac{x_1+x_2}{2}\right) \leqslant \frac{1}{2}[f(x_1)+f(x_2)]. \quad \text{(平均值不等式)}$$

Proof.

只要证明充分性即可.设 $a < b \in I$, $\ell(x)$ 是过 $(a, f(a))$ 和 $(b, f(b))$ 的线性函数.

命题

设 f 为区间 I 上的连续函数, 则 f 为凸函数当且仅当对任意 $x_1 < x_2 \in I$, 有

$$f\left(\frac{x_1+x_2}{2}\right) \leqslant \frac{1}{2}[f(x_1)+f(x_2)]. \quad \text{（平均值不等式）}$$

Proof.

只要证明充分性即可.设 $a < b \in I$, $\ell(x)$ 是过 $(a, f(a))$ 和 $(b, f(b))$ 的线性函数.显然 $g(x) := f(x) - \ell(x)$ 仍然满足平均值不等式.

命题

设 f 为区间 I 上的连续函数, 则 f 为凸函数当且仅当对任意 $x_1 < x_2 \in I$, 有

$$f\left(\frac{x_1+x_2}{2}\right) \leqslant \frac{1}{2}[f(x_1)+f(x_2)]. \quad \text{（平均值不等式）}$$

Proof.

只要证明充分性即可.设 $a < b \in I$, $\ell(x)$ 是过 $(a, f(a))$ 和 $(b, f(b))$ 的线性函数.显然 $g(x) := f(x) - \ell(x)$ 仍然满足平均值不等式.

因为 g 是 $[a, b]$ 上的连续函数, 故在某点 $x_0 \in [a, b]$ 处取到最大值 M.

命题

设 f 为区间 I 上的连续函数, 则 f 为凸函数当且仅当对任意 $x_1 < x_2 \in I$, 有

$$f\left(\frac{x_1 + x_2}{2}\right) \leqslant \frac{1}{2}[f(x_1) + f(x_2)]. \quad \text{（平均值不等式）}$$

Proof.

只要证明充分性即可.设 $a < b \in I$, $\ell(x)$ 是过 $(a, f(a))$ 和 $(b, f(b))$ 的线性函数.显然 $g(x) := f(x) - \ell(x)$ 仍然满足平均值不等式.

因为 g 是 $[a, b]$ 上的连续函数, 故在某点 $x_0 \in [a, b]$ 处取到最大值 M.要证 f 为凸函数, 即证 $g(x) \leqslant 0$.

命题

设 f 为区间 I 上的连续函数, 则 f 为凸函数当且仅当对任意 $x_1 < x_2 \in I$, 有

$$f\left(\frac{x_1 + x_2}{2}\right) \leqslant \frac{1}{2}[f(x_1) + f(x_2)]. \quad \text{（平均值不等式）}$$

Proof.

只要证明充分性即可.设 $a < b \in I$, $\ell(x)$ 是过 $(a, f(a))$ 和 $(b, f(b))$ 的线性函数.显然 $g(x) := f(x) - \ell(x)$ 仍然满足平均值不等式.

因为 g 是 $[a, b]$ 上的连续函数, 故在某点 $x_0 \in [a, b]$ 处取到最大值 M.要证 f 为凸函数, 即证 $g(x) \leqslant 0$.若能说明 $M = 0$ 即可.

连续函数的凸性判定

命题

设 f 为区间 I 上的连续函数, 则 f 为凸函数当且仅当对任意 $x_1 < x_2 \in I$, 有

$$f\left(\frac{x_1+x_2}{2}\right) \leqslant \frac{1}{2}[f(x_1)+f(x_2)]. \quad \text{（平均值不等式）}$$

Proof.

只要证明充分性即可.设 $a < b \in I$, $\ell(x)$ 是过 $(a, f(a))$ 和 $(b, f(b))$ 的线性函数.显然 $g(x) := f(x) - \ell(x)$ 仍然满足平均值不等式.

因为 g 是 $[a, b]$ 上的连续函数, 故在某点 $x_0 \in [a, b]$ 处取到最大值 M.要证 f 为凸函数, 即证 $g(x) \leqslant 0$.若能说明 $M = 0$ 即可.

如果 $x_0 = a$ 或 $x_0 = b$, 则 $M = 0$. □

续

Proof.

如果 $x_0 \in (a, b)$,

续

Proof.

如果 $x_0 \in (a, b)$,不妨设 $x_0 \leqslant \frac{a+b}{2}$.

续

Proof.

如果 $x_0 \in (a, b)$,不妨设 $x_0 \leqslant \frac{a+b}{2}$.考虑 a 关于 x_0 的对称点 $x_1 = 2x_0 - a$,

Proof.

如果 $x_0 \in (a, b)$,不妨设 $x_0 \leqslant \frac{a+b}{2}$.考虑 a 关于 x_0 的对称点 $x_1 = 2x_0 - a$,则 $x_1 \in [a, b]$, 且

$$M = g(x_0) = g\left(\frac{a + x_1}{2}\right) \leqslant \frac{1}{2}[g(a) + g(x_1)] \leqslant M,$$

Proof.

如果 $x_0 \in (a,b)$,不妨设 $x_0 \leqslant \frac{a+b}{2}$.考虑 a 关于 x_0 的对称点 $x_1 = 2x_0 - a$,则 $x_1 \in [a,b]$, 且

$$M = g(x_0) = g\left(\frac{a+x_1}{2}\right) \leqslant \frac{1}{2}[g(a)+g(x_1)] \leqslant M,$$

这说明 $M = g(a) = 0$. □

Proof.

如果 $x_0 \in (a,b)$,不妨设 $x_0 \leqslant \frac{a+b}{2}$.考虑 a 关于 x_0 的对称点 $x_1 = 2x_0 - a$,则 $x_1 \in [a,b]$, 且

$$M = g(x_0) = g\left(\frac{a+x_1}{2}\right) \leqslant \frac{1}{2}[g(a)+g(x_1)] \leqslant M,$$

这说明 $M = g(a) = 0$. □

注

从上面的证明过程可以看出, 如果函数满足平均值不等式且在 x_0 处达到最大值, 则在关于 x_0 对称的任意一点处也达到最大值.

续

Proof.

如果 $x_0 \in (a, b)$,不妨设 $x_0 \leqslant \frac{a+b}{2}$.考虑 a 关于 x_0 的对称点 $x_1 = 2x_0 - a$,则 $x_1 \in [a, b]$, 且

$$M = g(x_0) = g\left(\frac{a + x_1}{2}\right) \leqslant \frac{1}{2}[g(a) + g(x_1)] \leqslant M,$$

这说明 $M = g(a) = 0$. □

注

从上面的证明过程可以看出, 如果函数满足平均值不等式且在 x_0 处达到最大值, 则在关于 x_0 对称的任意一点处也达到最大值.（上面 $g(x_0) = M$ 推出 $g(a) = g(x_1) = M$.）

续

Proof.

如果 $x_0 \in (a,b)$,不妨设 $x_0 \leqslant \frac{a+b}{2}$.考虑 a 关于 x_0 的对称点 $x_1 = 2x_0 - a$,则 $x_1 \in [a,b]$, 且

$$M = g(x_0) = g\left(\frac{a+x_1}{2}\right) \leqslant \frac{1}{2}[g(a)+g(x_1)] \leqslant M,$$

这说明 $M = g(a) = 0$. □

注

从上面的证明过程可以看出, 如果函数满足平均值不等式且在 x_0 处达到最大值, 则在关于 x_0 对称的任意一点处也达到最大值.（上面 $g(x_0) = M$ 推出 $g(a) = g(x_1) = M$.）特别地, 有下面的推论.

推论

设 f 为区间 I 上的凸函数, 如果 f 在 I 的内部达到最大值, 则 f 是常值函数.

推论

设 f 为区间 I 上的凸函数, 如果 f 在 I 的内部达到最大值, 则 f 是常值函数.

Proof.

设 f 在 x_0 处达到最大值.

推论

设 f 为区间 I 上的凸函数, 如果 f 在 I 的内部达到最大值, 则 f 是常值函数.

Proof.

设 f 在 x_0 处达到最大值.任取 $a, b \in I$, $a < x_0 < b$.

推论

设 f 为区间 I 上的凸函数, 如果 f 在 I 的内部达到最大值, 则 f 是常值函数.

Proof.

设 f 在 x_0 处达到最大值.任取 $a, b \in I$, $a < x_0 < b$.则 $x_0 = ta + (1-t)b$, $t \in (0, 1)$.

推论

设 f 为区间 I 上的凸函数, 如果 f 在 I 的内部达到最大值, 则 f 是常值函数.

Proof.

设 f 在 x_0 处达到最大值.任取 $a, b \in I$, $a < x_0 < b$.则 $x_0 = ta + (1-t)b$, $t \in (0, 1)$.此时

$$\max f = f(x_0) \leqslant tf(a) + (1-t)f(b) \leqslant t \max f + (1-t) \max f = \max f,$$

推论

设 f 为区间 I 上的凸函数, 如果 f 在 I 的内部达到最大值, 则 f 是常值函数.

Proof.

设 f 在 x_0 处达到最大值.任取 $a, b \in I$, $a < x_0 < b$.则 $x_0 = ta + (1-t)b$, $t \in (0,1)$.此时

$$\max f = f(x_0) \leqslant tf(a) + (1-t)f(b) \leqslant t\max f + (1-t)\max f = \max f,$$

这说明 $f(a) = f(b) = \max f$,

推论

设 f 为区间 I 上的凸函数, 如果 f 在 I 的内部达到最大值, 则 f 是常值函数.

Proof.

设 f 在 x_0 处达到最大值.任取 $a, b \in I$, $a < x_0 < b$.则 $x_0 = ta + (1-t)b$, $t \in (0,1)$.此时

$$\max f = f(x_0) \leqslant tf(a) + (1-t)f(b) \leqslant t\max f + (1-t)\max f = \max f,$$

这说明 $f(a) = f(b) = \max f$,因此 f 恒等于 $f(x_0)$. $\square$

例

证明函数 $f(x) = -\ln x \ (x > 0)$ 是凸函数.

例

证明函数 $f(x) = -\ln x \ (x > 0)$ 是凸函数.

Proof.

函数 $f(x) = -\ln x$ 在 $(0, +\infty)$ 上连续, 故可用平均值不等式验证其凸性.

例

证明函数 $f(x) = -\ln x\ (x > 0)$ 是凸函数.

Proof.

函数 $f(x) = -\ln x$ 在 $(0, +\infty)$ 上连续, 故可用平均值不等式验证其凸性.任取 $x, y > 0$, 要证 $f(\frac{x+y}{2}) \leqslant \frac{1}{2}[f(x) + f(y)]$.

例

证明函数 $f(x) = -\ln x\ (x > 0)$ 是凸函数.

Proof.

函数 $f(x) = -\ln x$ 在 $(0, +\infty)$ 上连续, 故可用平均值不等式验证其凸性.任取 $x, y > 0$, 要证 $f(\frac{x+y}{2}) \leqslant \frac{1}{2}[f(x) + f(y)]$.这等价于

$$
\begin{aligned}
& -\ln\left(\frac{x+y}{2}\right) \leqslant \frac{1}{2}[-\ln x - \ln y] \\
\Leftrightarrow\ \ & \ln\left(\frac{x+y}{2}\right) \geqslant \ln\sqrt{xy} \\
\Leftrightarrow\ \ & \frac{x+y}{2} \geqslant \sqrt{xy}.
\end{aligned}
$$

例

证明函数 $f(x) = -\ln x\ (x > 0)$ 是凸函数.

Proof.

函数 $f(x) = -\ln x$ 在 $(0, +\infty)$ 上连续, 故可用平均值不等式验证其凸性.任取 $x, y > 0$, 要证 $f(\frac{x+y}{2}) \leqslant \frac{1}{2}[f(x) + f(y)]$.这等价于

$$\begin{aligned} & -\ln\left(\frac{x+y}{2}\right) \leqslant \frac{1}{2}[-\ln x - \ln y] \\ \Leftrightarrow\ & \ln\left(\frac{x+y}{2}\right) \geqslant \ln\sqrt{xy} \\ \Leftrightarrow\ & \frac{x+y}{2} \geqslant \sqrt{xy}. \end{aligned}$$

而这是显然成立的, 故 $f(x) = -\ln x$ 是 $(0, +\infty)$ 上的凸函数. □

命题 (导数性质)

设 f 为区间 I 上的凸函数, x 为 I 的内点, 则 f 在 x 处的左导数和右导数均存在, 且 $f'_-(x) \leqslant f'_+(x)$.

命题 (导数性质)

设 f 为区间 I 上的凸函数, x 为 I 的内点, 则 f 在 x 处的左导数和右导数均存在, 且 $f'_-(x) \leqslant f'_+(x)$.

Proof.

任取 $x_1, x_2 \in I$, 使得 $x_1 < x < x_2$.

命题 (导数性质)

设 f 为区间 I 上的凸函数, x 为 I 的内点, 则 f 在 x 处的左导数和右导数均存在, 且 $f'_-(x) \leqslant f'_+(x)$.

Proof.

任取 $x_1, x_2 \in I$, 使得 $x_1 < x < x_2$.因为 f 为凸函数, 我们有不等式

$$\frac{f(x) - f(x_1)}{x - x_1} \leqslant \frac{f(x_2) - f(x)}{x_2 - x},$$

命题 (导数性质)

设 f 为区间 I 上的凸函数, x 为 I 的内点, 则 f 在 x 处的左导数和右导数均存在, 且 $f'_-(x) \leqslant f'_+(x)$.

Proof.

任取 $x_1, x_2 \in I$, 使得 $x_1 < x < x_2$.因为 f 为凸函数, 我们有不等式

$$\frac{f(x)-f(x_1)}{x-x_1} \leqslant \frac{f(x_2)-f(x)}{x_2-x},$$

f 为凸函数也意味着 $\frac{f(x)-f(x_1)}{x-x_1}$ 是关于 x_1 的单调递增函数, 上式则表明这个单调递增函数有上界, 从而极限

$$\lim_{x_1 \to x^-} \frac{f(x)-f(x_1)}{x-x_1} = f'_-(x)$$

存在,

命题 (导数性质)

设 f 为区间 I 上的凸函数, x 为 I 的内点, 则 f 在 x 处的左导数和右导数均存在, 且 $f'_-(x) \leqslant f'_+(x)$.

Proof.

任取 $x_1, x_2 \in I$, 使得 $x_1 < x < x_2$.因为 f 为凸函数, 我们有不等式

$$\frac{f(x)-f(x_1)}{x-x_1} \leqslant \frac{f(x_2)-f(x)}{x_2-x},$$

f 为凸函数也意味着 $\frac{f(x)-f(x_1)}{x-x_1}$ 是关于 x_1 的单调递增函数, 上式则表明这个单调递增函数有上界, 从而极限

$$\lim_{x_1 \to x^-} \frac{f(x)-f(x_1)}{x-x_1} = f'_-(x)$$

存在,且

$$f'_-(x) \leqslant \frac{f(x_2)-f(x)}{x_2-x}.$$

Proof.

同理, 当 $x_2 \to x^+$ 时, $\frac{f(x_2)-f(x)}{x_2-x}$ 单调递减且有下界, 从而极限存在且满足

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

Proof.

同理, 当 $x_2 \to x^+$ 时, $\frac{f(x_2)-f(x)}{x_2-x}$ 单调递减且有下界, 从而极限存在且满足

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

这就证明了命题. $\square$

Proof.

同理, 当 $x_2 \to x^+$ 时, $\frac{f(x_2)-f(x)}{x_2-x}$ 单调递减且有下界, 从而极限存在且满足

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

这就证明了命题. $\square$

注

(1) 凸函数不一定可微.

Proof.

同理, 当 $x_2 \to x^+$ 时, $\frac{f(x_2)-f(x)}{x_2-x}$ 单调递减且有下界, 从而极限存在且满足

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

这就证明了命题. $\square$

注

(1) 凸函数不一定可微.例如 $f(x)=|x|$ 是凸函数, 但它在 $x=0$ 处不可微.

Proof.

同理, 当 $x_2 \to x^+$ 时, $\frac{f(x_2)-f(x)}{x_2-x}$ 单调递减且有下界, 从而极限存在且满足

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

这就证明了命题. □

注

(1) 凸函数不一定可微.例如 $f(x) = |x|$ 是凸函数, 但它在 $x = 0$ 处不可微.

(2) 凸函数的不可微点只有至多可数个.

Proof.

同理, 当 $x_2 \to x^+$ 时, $\frac{f(x_2)-f(x)}{x_2-x}$ 单调递减且有下界, 从而极限存在且满足

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

这就证明了命题. □

注

(1) 凸函数不一定可微.例如 $f(x) = |x|$ 是凸函数, 但它在 $x = 0$ 处不可微.

(2) 凸函数的不可微点只有至多可数个.（凸函数 f 的导函数 f' 在其定义域内是单调递增函数, 因此其间断点至多为可数多个.）

命题

设 f 是区间 I 上的可微函数, 则

(1) f 为凸函数当且仅当 f' 为单调递增函数*;*

(2) f 为凸函数当且仅当对任意 $x_0, x \in I$, 有 $f(x) \geqslant f'(x_0)(x - x_0) + f(x_0)$*.*

命题

设 f 是区间 I 上的可微函数, 则

(1) f 为凸函数当且仅当 f' 为单调递增函数;

(2) f 为凸函数当且仅当对任意 $x_0, x \in I$, 有 $f(x) \geqslant f'(x_0)(x - x_0) + f(x_0)$.

Proof.

(1) 设 f 为可微函数, $x_1 < x_2$.

命题

设 f 是区间 I 上的可微函数, 则

(1) f 为凸函数当且仅当 f' 为单调递增函数;

(2) f 为凸函数当且仅当对任意 $x_0, x \in I$, 有 $f(x) \geqslant f'(x_0)(x - x_0) + f(x_0)$.

Proof.

(1) 设 f 为可微函数, $x_1 < x_2$.则当 $x \in (x_1, x_2)$ 时,

$$\frac{f(x) - f(x_1)}{x - x_1} \leqslant \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leqslant \frac{f(x_2) - f(x)}{x_2 - x},$$

命题

设 f 是区间 I 上的可微函数, 则

(1) f *为凸函数当且仅当* f' *为单调递增函数;*

(2) f *为凸函数当且仅当对任意* $x_0, x \in I$, *有* $f(x) \geqslant f'(x_0)(x - x_0) + f(x_0)$.

Proof.

(1) 设 f 为可微函数, $x_1 < x_2$.则当 $x \in (x_1, x_2)$ 时,

$$\frac{f(x) - f(x_1)}{x - x_1} \leqslant \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leqslant \frac{f(x_2) - f(x)}{x_2 - x},$$

分别在这个不等式的左边和右边令 $x \to x_1$ 以及 $x \to x_2$, 得

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

命题

设 f 是区间 I 上的可微函数, 则

(1) f 为凸函数当且仅当 f' 为单调递增函数;

(2) f 为凸函数当且仅当对任意 $x_0, x \in I$, 有 $f(x) \geqslant f'(x_0)(x - x_0) + f(x_0)$.

Proof.

(1) 设 f 为可微函数, $x_1 < x_2$.则当 $x \in (x_1, x_2)$ 时,

$$\frac{f(x) - f(x_1)}{x - x_1} \leqslant \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leqslant \frac{f(x_2) - f(x)}{x_2 - x},$$

分别在这个不等式的左边和右边令 $x \to x_1$ 以及 $x \to x_2$, 得

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

这说明 f' 是单调递增函数. □

续

Proof.

反之，

Proof.

反之,如果 f 可微, 任取 $a < b \in I$, 设 $a < x < b$,

Proof.

反之,如果 f 可微, 任取 $a < b \in I$, 设 $a < x < b$,由微分中值定理, 存在 $\xi_1 \in (a, x)$, $\xi_2 \in (x, b)$, 使得

$$f(x) - f(a) = f'(\xi_1)(x - a), \quad f(b) - f(x) = f'(\xi_2)(b - x).$$

Proof.

反之,如果 f 可微, 任取 $a < b \in I$, 设 $a < x < b$,由微分中值定理, 存在 $\xi_1 \in (a, x)$, $\xi_2 \in (x, b)$, 使得

$$f(x) - f(a) = f'(\xi_1)(x - a), \quad f(b) - f(x) = f'(\xi_2)(b - x).$$

如果 f' 为单调递增函数, 则 $f'(\xi_1) \leqslant f'(\xi_2)$,

Proof.
反之,如果 f 可微, 任取 $a < b \in I$, 设 $a < x < b$,由微分中值定理, 存在 $\xi_1 \in (a, x)$, $\xi_2 \in (x, b)$, 使得

$$f(x) - f(a) = f'(\xi_1)(x - a), \quad f(b) - f(x) = f'(\xi_2)(b - x).$$

如果 f' 为单调递增函数, 则 $f'(\xi_1) \leqslant f'(\xi_2)$,从而

$$\frac{f(x) - f(a)}{x - a} = f'(\xi_1) \leqslant f'(\xi_2) = \frac{f(b) - f(x)}{b - x}.$$

Proof.
反之,如果 f 可微, 任取 $a < b \in I$, 设 $a < x < b$,由微分中值定理, 存在 $\xi_1 \in (a, x)$, $\xi_2 \in (x, b)$, 使得

$$f(x) - f(a) = f'(\xi_1)(x - a), \quad f(b) - f(x) = f'(\xi_2)(b - x).$$

如果 f' 为单调递增函数, 则 $f'(\xi_1) \leqslant f'(\xi_2)$,从而

$$\frac{f(x) - f(a)}{x - a} = f'(\xi_1) \leqslant f'(\xi_2) = \frac{f(b) - f(x)}{b - x}.$$

由前面的讨论, 知 f 为凸函数. $\square$

Proof.

(2) 设 f 为可微凸函数, $x_0 \in I$.

Proof.

(2) 设 f 为可微凸函数, $x_0 \in I$.如果 $x > x_0$, 则由 (1) 的证明, 有

$$f'(x_0) \leqslant \frac{f(x) - f(x_0)}{x - x_0};$$

Proof.

(2) 设 f 为可微凸函数, $x_0 \in I$.如果 $x > x_0$, 则由 (1) 的证明, 有

$$f'(x_0) \leqslant \frac{f(x) - f(x_0)}{x - x_0};$$

如果 $x < x_0$, 仍由 (1) 的证明, 有

$$\frac{f(x_0) - f(x)}{x_0 - x} \leqslant f'(x_0);$$

续

Proof.

(2) 设 f 为可微凸函数, $x_0 \in I$.如果 $x > x_0$, 则由 (1) 的证明, 有

$$f'(x_0) \leqslant \frac{f(x) - f(x_0)}{x - x_0};$$

如果 $x < x_0$, 仍由 (1) 的证明, 有

$$\frac{f(x_0) - f(x)}{x_0 - x} \leqslant f'(x_0);$$

总之, 对任何 $x, x_0 \in I$, 有

$$f(x) \geqslant f'(x_0)(x - x_0) + f(x_0). \tag{*}$$

□

Proof.

反之, 如果不等式(∗)成立, 则任取 $a < x < b$, 有

$$\frac{f(x) - f(a)}{x - a} \leqslant f'(x) \leqslant \frac{f(b) - f(x)}{b - x},$$

续

Proof.

反之, 如果不等式$(*)$成立, 则任取 $a < x < b$, 有

$$\frac{f(x)-f(a)}{x-a} \leqslant f'(x) \leqslant \frac{f(b)-f(x)}{b-x},$$

由前面的讨论知 f 为凸函数. $\square$

命题

设 f 为区间 I 上的连续函数. 如果 f 的左导数 f'_- 存在且单调递增, 则 f 为凸函数. 对右导数有类似结论.

命题

设 f 为区间 I 上的连续函数. 如果 f 的左导数 f'_- 存在且单调递增, 则 f 为凸函数. 对右导数有类似结论.

Proof.

我们来证明 f 满足 Jensen 不等式.

命题

设 f 为区间 I 上的连续函数. 如果 f 的左导数 f'_- 存在且单调递增, 则 f 为凸函数. 对右导数有类似结论.

Proof.

我们来证明 f 满足 Jensen 不等式.首先, 设 $x_0 \in I$, 记

$$L(x) = f'_-(x_0)(x - x_0) + f(x_0), \quad g(x) = f(x) - L(x).$$

命题

设 f 为区间 I 上的连续函数. 如果 f 的左导数 f'_- 存在且单调递增, 则 f 为凸函数. 对右导数有类似结论.

Proof.

我们来证明 f 满足 Jensen 不等式.首先, 设 $x_0 \in I$, 记

$$L(x) = f'_-(x_0)(x - x_0) + f(x_0), \quad g(x) = f(x) - L(x).$$

于是 $g'_-(x) = f'_-(x) - f'_-(x_0)$.

命题

设 f 为区间 I 上的连续函数. 如果 f 的左导数 f'_- 存在且单调递增, 则 f 为凸函数. 对右导数有类似结论.

Proof.

我们来证明 f 满足 Jensen 不等式.首先, 设 $x_0 \in I$, 记

$$L(x) = f'_-(x_0)(x - x_0) + f(x_0), \quad g(x) = f(x) - L(x).$$

于是 $g'_-(x) = f'_-(x) - f'_-(x_0)$.由 f'_- 单调递增可知, 当 $x \leqslant x_0$ 时, $g'_-(x) \leqslant 0$;

命题

设 f 为区间 I 上的连续函数. 如果 f 的左导数 f'_- 存在且单调递增, 则 f 为凸函数. 对右导数有类似结论.

Proof.

我们来证明 f 满足 Jensen 不等式.首先, 设 $x_0 \in I$, 记

$$L(x) = f'_-(x_0)(x - x_0) + f(x_0), \quad g(x) = f(x) - L(x).$$

于是 $g'_-(x) = f'_-(x) - f'_-(x_0)$.由 f'_- 单调递增可知, 当 $x \leqslant x_0$ 时, $g'_-(x) \leqslant 0$;当 $x \geqslant x_0$ 时, $g'_-(x) \geqslant 0$.

命题

设 f 为区间 I 上的连续函数. 如果 f 的左导数 f'_- 存在且单调递增, 则 f 为凸函数. 对右导数有类似结论.

Proof.

我们来证明 f 满足 Jensen 不等式.首先, 设 $x_0 \in I$, 记

$$L(x) = f'_-(x_0)(x - x_0) + f(x_0), \quad g(x) = f(x) - L(x).$$

于是 $g'_-(x) = f'_-(x) - f'_-(x_0)$.由 f'_- 单调递增可知, 当 $x \leqslant x_0$ 时, $g'_-(x) \leqslant 0$;当 $x \geqslant x_0$ 时, $g'_-(x) \geqslant 0$.尽管是左导数, 也可以说明 g 函数在 x_0 的左边单调递减, 在 x_0 的右边单调递增.(需要用到[梅]《数学分析》第四章命题4.3.1的证明方法及其注记.)

命题

设 f 为区间 I 上的连续函数. 如果 f 的左导数 f'_- 存在且单调递增, 则 f 为凸函数. 对右导数有类似结论.

Proof.

我们来证明 f 满足 Jensen 不等式.首先, 设 $x_0 \in I$, 记

$$L(x) = f'_-(x_0)(x - x_0) + f(x_0), \quad g(x) = f(x) - L(x).$$

于是 $g'_-(x) = f'_-(x) - f'_-(x_0)$.由 f'_- 单调递增可知, 当 $x \leqslant x_0$ 时, $g'_-(x) \leqslant 0$;当 $x \geqslant x_0$ 时, $g'_-(x) \geqslant 0$.尽管是左导数, 也可以说明 g 函数在 x_0 的左边单调递减, 在 x_0 的右边单调递增.(需要用到[梅]《数学分析》第四章命题4.3.1的证明方法及其注记.)这说明 g 的最小值为 $g(x_0) = 0$, 即

$$f(x) \geqslant L(x) = f'_-(x_0)(x - x_0) + f(x_0), \quad x \in I. \tag{**}$$

□

Proof.

其次, 设 $x_i \in I$, $\lambda_i \geqslant 0$, $\sum_{i=1}^{n} \lambda_i = 1$.

Proof.

其次, 设 $x_i \in I$, $\lambda_i \geqslant 0$, $\sum\limits_{i=1}^{n} \lambda_i = 1$.记 $x_0 = \sum\limits_{i=1}^{n} \lambda_i x_i$, 利用 $(**)$ 可得

$$
\begin{aligned}
\sum_{i=1}^{n} \lambda_i f(x_i) \geqslant f'_-(x_0) \sum_{i=1}^{n} \lambda_i (x_i - x_0) + f(x_0) \sum_{i=1}^{n} \lambda_i \\
= f(x_0) = f\Big(\sum_{i=1}^{n} \lambda_i x_i\Big),
\end{aligned}
$$

Proof.

其次, 设 $x_i \in I$, $\lambda_i \geqslant 0$, $\sum\limits_{i=1}^{n} \lambda_i = 1$.记 $x_0 = \sum\limits_{i=1}^{n} \lambda_i x_i$, 利用 $(**)$ 可得

$$\begin{aligned}\sum_{i=1}^{n} \lambda_i f(x_i) &\geqslant f'_-(x_0) \sum_{i=1}^{n} \lambda_i (x_i - x_0) + f(x_0) \sum_{i=1}^{n} \lambda_i \\ &= f(x_0) = f\Big(\sum_{i=1}^{n} \lambda_i x_i\Big),\end{aligned}$$

这说明 f 满足 Jensen 不等式. □

二阶可导函数的凸性

命题

如果 f 在 I 上二阶可导, 则 f 为凸函数当且仅当 $f'' \geqslant 0$.

命题

如果 f 在 I 上二阶可导, 则 f 为凸函数当且仅当 $f'' \geqslant 0$.

Proof.

如果 f 二阶可导, 则 f 为凸函数当且仅当 f' 为单调递增函数,

命题

如果 f 在 I 上二阶可导, 则 f 为凸函数当且仅当 $f'' \geqslant 0$.

Proof.

如果 f 二阶可导, 则 f 为凸函数当且仅当 f' 为单调递增函数,而可微函数 f' 是单调递增函数当且仅当其导数 f'' 非负. $\square$

习题

习题

利用函数 $f(x) = -\ln x$ 是凸函数这个事实证明算术–几何平均不等式.