第九讲：凸函数扩展

凸函数相关的其他定义

杨 林

大 纲

1. 共轭函数

2.对数凹函数和对数凸函数

3.拟凸函数

大 纲

1. 共轭函数

2.对数凹函数和对数凸函数

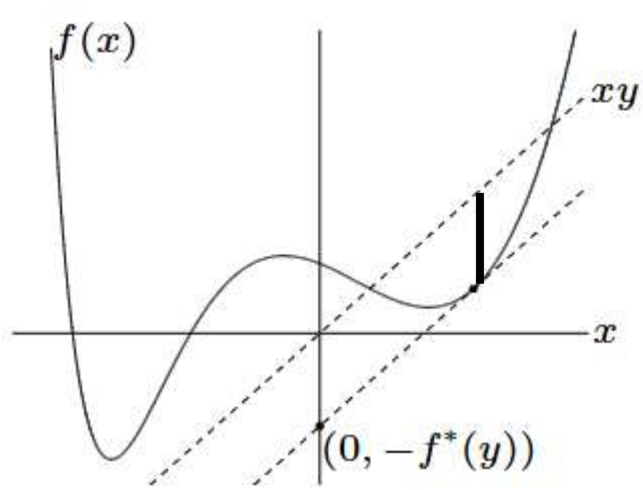
3.拟凸函数

1 共轭函数

■ **定义1（共轭函数）：**

函数 f 的共轭函数是

$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} (y^T x - f(x))$$



f 是凸函数（即使 f 不是），为什么？

1 共轭函数

- **例** 1:

1. 负对数 $f(x) = -\log x$

$$f^*(y) = \sup_{x>0}(xy + \log x) = \begin{cases} -1 - \log(-y) & y < 0 \\ \infty & \text{其他情况} \end{cases}$$

2. 严格凸的二次函数 $f(x) = (1/2)x^T Qx$，其中 $Q \in S_{++}^n$

$$f^*(y) = \sup_x(y^T x - (1/2)x^T Qx) =) = (1/2)y^T Q^{-1} y$$

分析: 关于x求导：$y - \left(\frac{1}{2}\right)(Q + Q^T)x = y - Qx$，极值在$x = Q^{-1}y$时取得，从而有$f^*(y) = y^T Q^{-1} y - \left(\frac{1}{2}\right) y^T Q^{-1} y = (1/2)y^T Q^{-1} y$

大 纲

1. 共轭函数

2.对数凹函数和对数凸函数

3.拟凸函数

2 对数凹函数和对数凸函数

■ **定义2（对数凹函数和对数凸函数）：**

一个正函数 f 是对数凹的若 $\log f$ 是凹的

$$f(\theta x+(1-\theta)y)\geq f(x)^{\theta}f(y)^{1-\theta}\quad 对于\ 0\leq\theta\leq 1$$

或者

$$\log f\ (\theta x+(1-\theta)y)\geq\theta\log f\ (x)+(1-\theta)\log f\ (y)$$

函数 f 是对数凸的若 $\log f$ 是凸的

注： log函数不加说明默认底数为e

2 对数凹函数和对数凸函数

- **例** 2 (幂函数): $\mathbb{R}_{++}$ 上的 x^a 在 $a \leq 0$ 时对数凸，在 $a \geq 0$ 时对数凹
- **证明:**

对于任意$0 \leq \theta \leq 1$, 当 $a \leq 0$ 时

$$f(\theta x + (1-\theta)y) = (\theta x + (1-\theta)y)^a = e^{a\log(\theta x+(1-\theta)y)}$$

由于 $f(x) = -\log x$ 是 $\mathbb{R}_{++}$ 上的凸函数,所以有

$$\log(\theta x + (1-\theta)y) \geq \theta \log x + (1-\theta)\log y$$

故

$$e^{a\log(\theta x+(1-\theta)y)} \leq e^{a(\theta\log x+(1-\theta)\log y)}$$

$$= x^{a\theta}y^{a(1-\theta)} = f(x)^\theta f(y)^{1-\theta}$$

$\mathbb{R}_{++}$ 上的 x^a 在 $a \leq 0$ 时对数凸, 同理可证在 $a \geq 0$ 时对数凹

2 对数凹函数和对数凸函数

- **例** 3: 许多常见的概率密度函数是对数凹的，例如正态分布：

$$f(x)=\frac{1}{\sqrt{(2\pi)^n\det\Sigma}}e^{-\frac{1}{2}(x-\bar{x})^T\Sigma^{-1}(x-\bar{x})}$$

- **证明:**

$$\log f(x)=\log\frac{1}{\sqrt{(2\pi)^n\det\Sigma}}-\frac{1}{2}(x-\bar{x})^T\Sigma^{-1}(x-\bar{x})$$

因为 $\Sigma^{-1}\succeq 0$, 所以常值函数 $g(x)=\log\frac{1}{\sqrt{(2\pi)^n\det\Sigma}}$ 与二次函数 $h(x)=-\frac{1}{2}(x-\bar{x})^T\Sigma^{-1}(x-\bar{x})$ 都是凹函数

其和函数 $\log f(x)$ 也是凹函数

注： $\Sigma=\mathrm{E}[(x-\bar{x})(x-\bar{x})^T]$为协方差矩阵，为半正定对称矩阵

大 纲

1. 共轭函数

2.对数凹函数和对数凸函数

3.拟凸函数

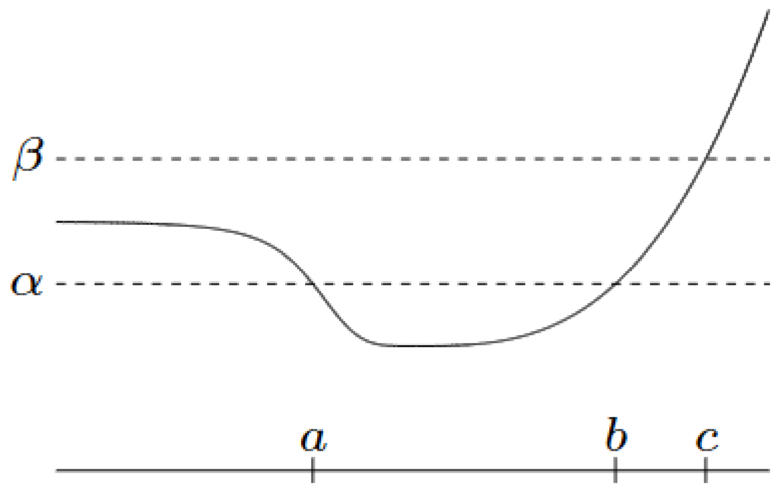
3 拟凸函数

■ **定义3(拟凸函数)：**

$f : \mathbb{R}^n \to \mathbb{R}$ 是拟凸的，如果 $\mathbf{dom}\, f$ 是凸的，并且下水平集

$$S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$$

对所有 α 都是凸的



如果 $-f$ 是拟凸的，则 f 是拟凹的

如果 f 是拟凸且拟凹的，则它是拟线性的

3 拟凸函数

■ 性质

修正的 Jensen 不等式：对于拟凸函数 f

$$0 \leq \theta \leq 1 \Longrightarrow f(\theta x + (1-\theta)y) \leq \max\{f(x), f(y)\}$$

■ 等价充要条件

■ 证明（必要性）：

显然，x 和 y 分别在 $f(x)$ 和 $f(y)$ 的下水平集内

x 和 y 在 $f(x)$ 和 $f(y)$ 中较大者对应的下水平集内

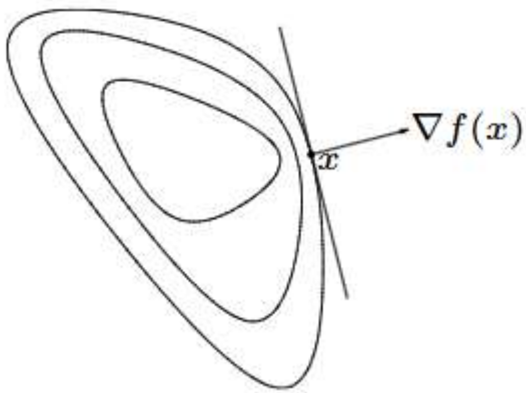
由于下水平集是凸的，$\theta x + (1-\theta)y$ 必须在 $f(x)$ 或 $f(y)$ 的下水平集内，即 $f(\theta x + (1-\theta)y) \leq \max\{f(x), f(y)\}$

3 拟凸函数*

■ 性质

一阶条件： 可微函数 f 在凸域上为拟凸当且仅当

$$f(y) \leq f(x) \Rightarrow \nabla f(x)^T(y-x) \leq 0$$



■ 证明：

（提示）$f(y) \leq f(x)$ 意味着 y 位于凸子水平集 $C_{f(x)}$ 的内部，而 x 位于边界上，$\nabla f(x)$ 和 x 决定了包含子水平集的半空间

3 拟凸函数*

■ 性质

二阶条件： f二次可微，如果函数f在凸域上为拟凸，则有

$$y^T\nabla f(x) = 0 \implies y^T\nabla^2 f(x)y \geq 0$$

反之，如果函数f满足

$$y^T\nabla f(x) = 0 \implies y^T\nabla^2 f(x)y > 0$$

则函数为拟凸函数

充分性解释：在斜率为0的点二阶导数非负

3 拟凸函数

性质

3. 拟凸函数的和不一定为拟凸，例如$f_1(x) = -1/x$, $f_2(x) = -1/(1-x), 0 \leq x \leq 1$

3 拟凸函数

□ 例 4:

1. $\text{ceil}(x) = \inf\{z \in \mathbb{Z} \mid z \geq x\}$ 是拟线性函数
2. $\log x$ 在 $\mathbb{R}_{++}$ 上是拟线性函数
3. $f(x_1, x_2) = x_1 x_2$在 $\mathbb{R}_{++}^2$ 上是拟凹函数
4. 线性分数函数

$$f(x) = \frac{a^T x + b}{c^T x + d}, \textbf{dom } f = \{x | c^T x + d > 0\}$$

是拟线性的

5. 距离比函数

$$f(x) = \frac{\|x - a\|_2}{\|x - b\|_2}, \textbf{dom } f = \{x \mid \|x - a\|_2 \leq \|x - b\|_2\}$$

是拟凸函数

3 拟凸函数

- **证明** 4.1 $\text{ceil}(x) = \inf\{z \in \mathbb{Z} \mid z \geq x\}$：

$$\text{ceil}(\theta x + (1-\theta)y) = \lceil \theta x + (1-\theta)y \rceil$$
$$\leq \lceil \max\{x, y\} \rceil = \max\{\text{ceil}(x), \text{ceil}(y)\}$$

所以,$\text{ceil}(x)$ 是一个拟凸函数, 同理可证 $\text{ceil}(x)$ 是一个拟凹函数. 故$\text{ceil}(x)$ 是一个拟线性函数

3 拟凸函数

■ **证明** 4.2 ($\log x$) :

因为$\log x$ 是 $\mathbb{R}_{++}$ 上的单调递增函数, 所以有

$$\min\{\log x, \log y\} \leq \log[\theta x + (1-\theta)y] \leq \max\{\log x, \log y\}$$

故 $\log x$ 是一个拟线性函数

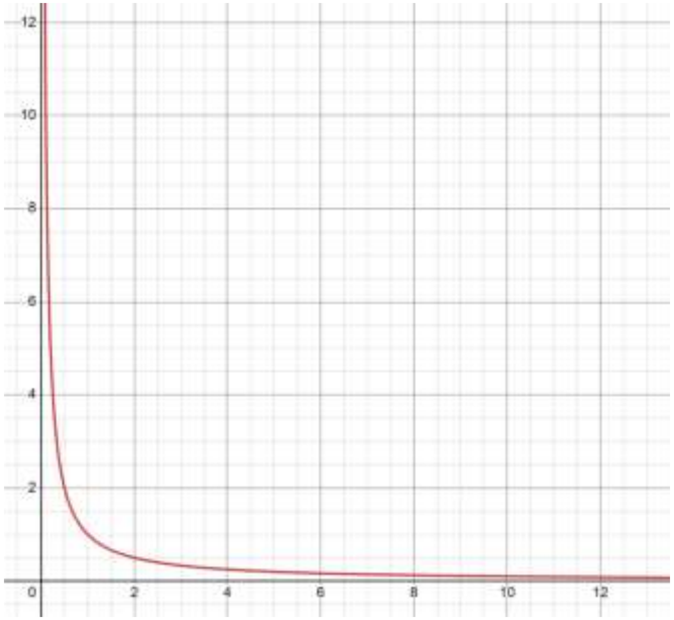
3 拟凸函数

■ **证明** 4.3 $f(x_1, x_2) = x_1 x_2$：

因为 $f(x_1, x_2)$ 的上水平集

$$\{x \in \mathbb{R}^2_{++} \mid x_1 x_2 \geq \alpha\}$$

都是凸集



3 拟凸函数

- **证明** 4.4 $f(x)=\frac{a^Tx+b}{c^Tx+d}$**:**

$f(x)=\frac{a^Tx+b}{c^Tx+d}$, $\mathbf{dom}\ f=\{x|c^Tx+d>0\}$ 的α-下水平集为

$$\begin{aligned}S_\alpha&=\{x\ |c^Tx+d>0,(a^Tx+b)/(c^Tx+d)\le\alpha\}\\&=\{x\ |c^Tx+d>0,(a^Tx+b)\le\alpha(c^Tx+d)\}\end{aligned}$$

因为它是一个开的半平面和闭的半平面的交集,所以它是一个凸集

3 拟凸函数

- **证明** 4.5 $f(x)=\frac{\|x-a\|_2}{\|x-b\|_2}$:

$f(x)=\frac{\|x-a\|_2}{\|x-b\|_2}$, $\mathbf{dom}\, f=\{x \mid \|x-a\|_2 \le \|x-b\|_2\}$ 的α-下水平集为

$$S_\alpha=\{x \mid \|x-a\|_2 \le \alpha\|x-b\|_2\}$$

由于在半平面 $\{x \mid \|x-a\|_2 \le \|x-b\|_2\}$ 上 $f(x) \le 1$,所以我们选取 $\alpha \le 1$

$\|x-a\|_2 \le \alpha\|x-b\|_2$ 两端平方, 并重新排列各项得到

$$(1-\alpha^2)x^Tx-2(a-\alpha^2 b)^T x+a^Ta-\alpha^2 b^Tb \le 0$$

当 $\alpha \le 1$ 时是一个凸集(实际上是一个欧几里得球)

谢 谢 ！