第八讲：凸函数的运算

杨 林

大 纲

1.保持凸性的运算

2.共轭函数

***3.对数凹函数和对数凸函数**

大 纲

1.保持凸性的运算

2.共轭函数

*3.对数凹函数和对数凸函数

1 保持凸性的运算

■ 确定函数凸性的实用方法：

1. 验证定义(通常通过限制在一条线上简)
2. 对于两次可微函数，证明$\nabla^2 f(x) \geq 0$
3. 证明 f 是通过保持凸性的运算从简单的凸函数得到的

- 非负加权求和
- 与仿射函数的复合
- 点态最大值与上确界
- 复合
- 最小化
- 透视

1 保持凸性的运算

- **和函数:** $f_1 + f_2$ 是凸函数若 f_1, f_2是凸函数(可推广到无穷和、积分)
- **非负倍数:** αf 是凸函数若 f 是凸函数, $\alpha \geq 0$
- **正加权求和**
- **与仿射函数复合:** $f(Ax + b)$ 是凸函数若 f 是凸函数

例 1:

1. 线性不等式的对数障碍

$$f(x) = -\log(b_i - a_i^T x), \mathbf{dom}\, f = \{x | a_i^T x < b_i, i = 1, \ldots, m\}$$

2. 仿射函数的任意范数: $f(x) = \| Ax + b \|$

1 保持凸性的运算

- **逐点最大值:** $f_1, \cdots, f_m$ 是凸函数,那么$\max\{f_1(x), \cdots, f_m(x)\}$ 是凸函数

例 2:

1. 分段线性函数: $f(x) = \max_{i=1,\ldots,m} a_i^T x + b_i$是凸函数
2. $x \in \mathbb{R}^n$ 中 r 个最大分量的和:

$$f(x) = x_{[1]} + x_{[2]} + \cdots + x_{[r]}$$

是凸函数 ($x_{[1]}$ 是 x 的第 i 个最大分量)

- **证明:**

$$f(x) = \max\{x_{i_1} + x_{i_2} + \cdots + x_{i_r} | 1 \leq i_1 < i_2 < \cdots < i_r \leq n$$

1 保持凸性的运算

- **逐点上确界:** 如果 $f(x,y)$ 在 x 上是凸的，对于每个 $y\in\mathcal{A}$，那么

$$g(x)=\sup_{y\in\mathcal{A}} f(x,y)$$

是凸函数

- **例** 3:

1. 集合 C 的支撑函数: $S_C(x)=\sup\limits_{y\in C} y^T x$ 是凸函数
2. 集合C中到最远点的距离: $f(x)=\sup\limits_{y\in C}\|x-y\|$
3. 对称矩阵的最大特征值:对于 $X\in S^n$

$$\lambda_{max}(X)=\sup_{\|y\|_2=1} y^T X y$$

1 保持凸性的运算

■ 标量函数的复合:

$g: \mathbb{R}^n \to \mathbb{R}$ 和 $h: \mathbb{R} \to \mathbb{R}$ 的复合函数:

$$f(x) = h\big(g(x)\big)$$

是凸函数，若: (1) g 是凸函数, h 是凸函数, $\tilde{h}$ 非递减; (2) g 是凹函数, h 是凸函数, $\tilde{h}$ 非递增.

■ **证明:**(对于 $n = 1$, 可微函数g, h)

$$f''(x) = h''(g(x))\mathrm{g}'(\mathrm{x})^2 + h'(g(x))g''(x)$$

注意:扩展值扩展必须保持单调性

□ **例** 4:

1. $\exp g(x)$ 是凸函数若 g 是凸函数

2. $1/g(x)$ 是凸函数若 g 是凹函数且为正

1 保持凸性的运算

■ 向量函数的复合:

$g: \mathbb{R}^n \to \mathbb{R}^k$ 和 $h: \mathbb{R}^k \to \mathbb{R}$ 的复合函数:

$$f(x) = h\big(g(x)\big) = h\big(g_1(x), \cdots, g_k(x)\big)$$

是凸函数，若: (1) g_i 是凸函数, h 是凸函数, $\tilde{h}$ 在每一个参数上是非递减的; (2) g_i是凹函数, h 是凸函数, $\tilde{h}$ 在每一个参数上是非递增的.

■ **证明:**(对于 $n = 1$, 可微函数g, h)

$$f''(x) = g'(x)^T \nabla^2 h(g(x)) g'(x) + \nabla h(g(x))^T g''(x)$$

注意:扩展值扩展必须保持单调性

1 保持凸性的运算

- **例** 5**:**

1. $\sum_{i=1}^{m} \log g_i(x)$ 是凹函数若 g_i 是凹函数且为正
2. $\log \sum_{i=1}^{m} \exp g_i(x)$ 是凸函数若 g_i 是凸函数

1 保持凸性的运算

- **最小化:** 如果 $f(x,y)$ 在 (x,y) 上是凸的，且 C 是一个凸集，那么

$$g(x)=\sup_{y\in C} f(x,y)$$

  是凸函数

例 6:

1. 到集合的距离: $\mathbf{dist}(x,S)=\inf_{y\in S}\|x-y\|$ 是凸函数若 S 是凸集

1 保持凸性的运算

- **证明:**

$$(x,t) \in \textbf{epi}\ g \Rightarrow \begin{cases} (1)\ t \ \geq \ f(x,y)\ \text{或}\ (x,y,t) \in \textbf{epi}\ f \\ (2) y \ \in \ C\ \text{其中}\ C\ \text{是凸集} \end{cases}$$

条件 (1) 和 (2) 分别导致两个集合

$$S_1 = \{(x,y,t) | y \in C\}$$

$$S_2 = \{(x,y,t) | (x,y,t) \in \textbf{epi}\ f\}$$

S_1 和 S_2 都是凸集

通过分析条件， **epi** g 是两个凸集交互的投影

因此， **epi** g 是凸的

1 保持凸性的运算

- **函数 $f: \mathbb{R}^n \to \mathbb{R}$ 的透视函数 $g: \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$：**

$$g(x,t) = tf\left(\frac{x}{t}\right), \mathbf{dom}\ g = \left\{(x,t) \middle| \frac{x}{t} \in \mathbf{dom}\ f, t > 0\right\}$$

g 是凸函数若 f 是凸函数

- **证明:**

$$\begin{aligned}(x,t,s) \in \mathbf{epi}\ g \ &\Leftrightarrow g(x,t) \le s \\ &\Leftrightarrow tf\left(\frac{x}{t}\right) \le s \\ &\Leftrightarrow f\left(\frac{x}{t}\right) \le \frac{s}{t} \\ &\Leftrightarrow (x/t, s/t) \in \mathbf{epi}\ f\end{aligned}$$

这意味着 $\mathbf{epi}\ f$ 是 $\mathbf{epi}\ g$ 的透视.因此, $\mathbf{epi}\ f$ 凸当且仅当 $\mathbf{epi}\ g$ 凸

1 保持凸性的运算

- **例** 7:

1. $f(x) = x^T x$ 是凸函数; 因此, 对于$t > 0$, $g(x,t) = x^T x/t$ 是凸函数
2. 负对数函数 $f(x) = -\log x$ 是凸函数;因此相对熵 $\mathrm{g}(x,t) = t\log t - t\log x$ 在 $\mathbb{R}^2_{++}$ 上是凸函数
3. 如果 f 是凸函数，那么

$$g(x) = (c^T x + d) f((Ax + b)/(c^T x + d))$$

在$\{x|c^T x + d > 0, (Ax + b)/(c^T x + d) \in \mathbf{dom}\, f\}$上是凸函数

大 纲

1.保持凸性的运算

2.共轭函数

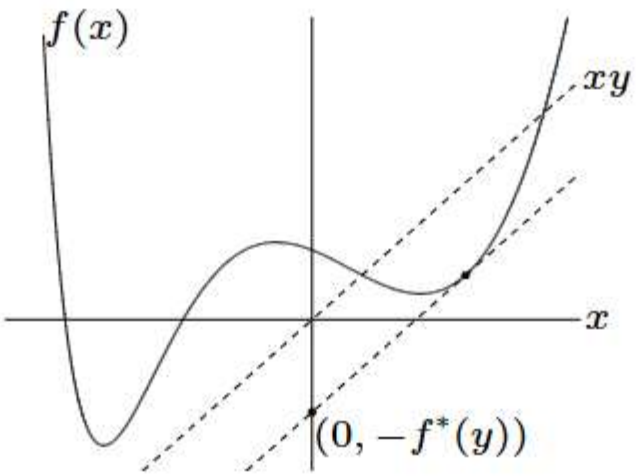
***3.对数凹函数和对数凸函数**

2 共轭函数

■ **定义1（共轭函数）：**

函数 f 的共轭函数是

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



f^* 是凸函数（即使 f 不是），为什么？

这在第 5 章将很有用

2 共轭函数

- **例** 8:

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 $Q \in S_{++}^n$

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

预备知识：矩阵的导数

$$\frac{\partial(\mathbf{a}^T\mathbf{x})}{\partial\mathbf{x}} = \frac{\partial(\mathbf{x}^T\mathbf{a})}{\partial\mathbf{x}} = \mathbf{a}$$

$$\frac{\partial(\mathbf{a}^T\mathbf{X}\mathbf{b})}{\partial\mathbf{X}} = \mathbf{a}\mathbf{b}^T = \mathbf{a}\otimes\mathbf{b} \in \mathbb{R}^{n\times n}$$

$$\frac{\partial(\mathbf{a}^T\mathbf{X}^T\mathbf{b})}{\partial\mathbf{X}} = \mathbf{b}\mathbf{a}^T = \mathbf{b}\otimes\mathbf{a} \in \mathbb{R}^{n\times n}$$

$$\frac{\partial(\mathbf{x}^T\mathbf{A}\mathbf{x})}{\partial\mathbf{x}} = (\mathbf{A}+\mathbf{A}^T)\mathbf{x}$$

$$\frac{\partial[(\mathbf{X}\mathbf{b}+\mathbf{c})^T\mathbf{A}(\mathbf{X}\mathbf{b}+\mathbf{c})]}{\partial\mathbf{X}} = (\mathbf{A}+\mathbf{A}^T)(\mathbf{X}\mathbf{b}+\mathbf{c})\mathbf{b}^T$$

$$\frac{\partial(\mathbf{a}^T\mathbf{X}^T\mathbf{X}\mathbf{b})}{\partial\mathbf{X}} = \mathbf{X}(\mathbf{a}\otimes\mathbf{b}+\mathbf{b}\otimes\mathbf{a})$$

大 纲

1.保持凸性的运算

2.共轭函数

***3.对数凹函数和对数凸函数**

3 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

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

函数 f 是对数凸的若 $\log f$ 是凸的

3 对数凹函数和对数凸函数

- **例** 9:

1. 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
2. 许多常见的概率密度函数是对数凹的，例如正态分布：

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

3. 累积高斯分布函数 Φ 是对数凹函数

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du$$

谢 谢 ！