第九章 优化基础

第 30 讲 凸函数和凸函数的保凸运算

黄定江

DaSE @ ECNU
djhuang@dase.ecnu.edu.cn

1. 30.1 凸函数
2. 30.2 凸函数的保凸运算
3. 30.3 共轭函数

1. 30.1 凸函数
2. 30.2 凸函数的保凸运算
3. 30.3 共轭函数

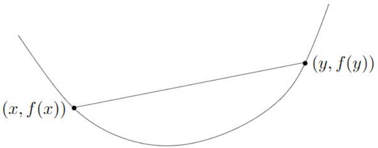
凸函数

定义 1

函数 $f: R^n \to R$ 是凸的，如果 $\mathbf{dom}\, f$ 是凸集，且对于任意 $x, y \in \mathbf{dom}\, f$ 和任意 $0 \le \theta \le 1$，有

$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)$$

- 从几何意义上看，上述不等式意味着点 $(x, f(x))$ 和 $(y, f(y))$ 之间的线段，即从 x 到 y 的弦，在函数 f 的图像上方（如下图所示）。称函数 f 是严格凸的，如果式子中的不等式当 $x \neq y$ 以及 $0 < \theta < 1$ 时严格成立。称函数 f 是凹的，如果函数 $-f$ 是凸的，称函数 f 是严格凹的，如果 $-f$ 严格凸。



凸函数

- 对于仿射函数，上面的不等式总成立，因此所有的仿射函数 (包括线性函数) 是既凸且凹的。反之，若某个函数是既凸又凹的，则其是仿射函数。
- 函数是凸的，当且仅当其在与其定义域相交的任何直线上都是凸的。换言之，函数 f 是凸的，当且仅当对于任意 $x \in \mathbf{dom}\, f$ 和任意向量 v，函数 $g(t) = f(x + tv)$ 是凸的 (其定义域为 $\{t | x + tv \in \mathbf{dom}\, f\}$)。这个性质非常有用，因为它容许我们通过将函数限制在直线上来判断其是否是凸函数。

保凸运算

确定函数凸性的方法

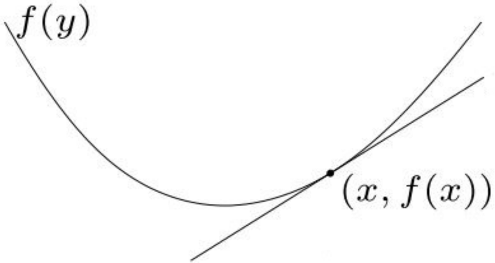
1. 定义
2. 求二阶导数，证明 $\nabla^2 f(x) \geq 0$
3. 说明 f 是由一些简单的凸函数通过一些保凸运算构造的
    - 非负加权求和
    - 复合仿射映射
    - 逐点最大和逐点上确界
    - 复合
    - 最小化
    - 透视函数

一阶条件

- 假设 f 可微 (即其梯度 $\triangledown f$ 在开集 $\mathbf{dom}\, f$ 内处处存在)，则函数 f 是凸函数的充要条件是 $\mathbf{dom}\, f$ 是凸集且对于任意 $x, y \in \mathbf{dom}\, f$, 下式成立

$$f(y) \geq f(x) + \triangledown f(x)^T (y - x)$$

下图描述了上述不等式:



一阶条件

- 由不等式可以知道，如果 $\triangledown f(x)=0$，那么对于所有的 $y \in \mathbf{dom}\, f$，存在 $f(y) \geq f(x)$，即 x 是函数 f 的全局极小点。
- 严格凸性同样可以由一阶条件刻画：函数 f 严格凸的充要条件是 $\mathbf{dom}\, f$ 是凸集且对于任意 $x, y \in \mathbf{dom}\, f, x \neq y$，有

$$f(y) > f(x) + \triangledown f(x)^T (y-x)$$

对于凹函数，亦存在与之对应的一阶条件：函数 f 是凹函数的充要条件是 $\mathbf{dom}\, f$ 是凸集且对于任意 $x, y \in \mathbf{dom}\, f$，下式成立

$$f(y) \leq f(x) + \triangledown f(x)^T (y-x)$$

二阶条件

- 现在假设函数 f 二阶可微，即对于开集 $\mathbf{dom}\, f$ 内的任意一点，它的 Hessian 矩阵或者二阶导数 $\bigtriangledown^2 f$ 存在，则函数 f 是凸函数的充要条件是，其 Hessian 矩阵是半正定阵：即对于所有的 $x \in \mathbf{dom}\, f$ 有
- 对于 R 上的函数，上式可以简化为一个简单的条件 $f''0(x) \geq 0(\mathbf{dom}\, f$ 是凸的，即一个区间)，此条件说明函数 f 的导数是非减的。条件

二阶条件

- 类似地，函数 f 是凹函数的充要条件是，$\textbf{dom}\, f$ 是凸集且对于任意 $x \in \textbf{dom}\, f$, $\nabla^2 f(x) \preceq 0$。严格凸的条件可以部分由二阶条件刻画。如果对于任意的 $x \in \textbf{dom}\, f$ 有 $\nabla^2 f(x) \succ 0$, 则函数 f 是严格凸。反过来则不一定成立：例如，函数 $f: R \to R$ 其表达式为 $f(x) = x^4$，它是严格凸的，但是在 $x = 0$ 处，二阶导数为零。

例子

前文已经提到所有的线性函数和仿射函数均为凸函数 (同时也是凹函数)，并描述了凸和凹的二次函数。下面给出更多的凸函数和凹函数的例子。首先考虑 R 上的一些函数，其自变量为 x。

- 指数函数。对任意 $a \in R$, 函数 e^{ax} 在 R 上是凸的。
- 幂函数。当 $a \geq 1$ 或 $a \leq 0$ 时，x^a 是在 R_{++} 上的凸函数，当 $0 \leq a \leq 1$ 时 x^a 是在 R_{++} 上的凹函数。
- 绝对值幂函数。当 $p \leq 1$ 时，函数 $|x|^p$ 在 R 上是凸函数。
- 对数函数。函数 $logx$ 在 R_{++} 上的凹函数。
- 负熵。函数 $xlogx$ 在其定义域上是凸函数。

例子

下面我们给出 R^n 上的一些例子。

- 范数。R^n 上的任意范数均为凸函数。
- 最大值函数。函数 $f(x) = max\{x_1, \cdots, x_n\}$ 在 R^n 上是凸的。
- 二次-线性分式函数。如 $f(x, y) = x^2/y$，在其定义域是凸的。
- 指数和对数。$\log(\sum_i \exp(x_i))$，在其定义域是凸的。
- 几何平均。几何平均是其定义域上面的凹函数。
- 对数-行列式。$f(\boldsymbol{X}) = \log\det\boldsymbol{X}$，在定义域内函数时凹的。

1 30.1 凸函数

2 30.2 凸函数的保凸运算

3 30.3 共轭函数

非负加权求和 & 复合仿射映射

非负乘积：如果 f 是凸的，$\alpha > 0$，那么 αf 是凸的
和：如果 f_1, f_2 都是凸的，那么 $f_1 + f_2$ 是凸的
复合仿射映射：如果 f 是凸的，那么 $f(Ax + b)$ 是凸的

非负加权求和 & 复合仿射映射

例 1

- 线性不等式对应的对数障碍

$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x), \qquad \mathbf{dom}\, f = \{x | a_i^T x \leq, i = 1, \cdots, m\}$$

- 仿射函数的 (任何) 范数:$f(x) = \|Ax + b\|$

逐点最大

如果 $f_1, \cdots, f_m$ 是凸的，那么 $f(x) = \max\{f_1(x), \cdots, f_m(x)\}$ 是凸的

例 2

- 分段线性函数：$f(x) = \max_{i=1,\cdots,m}\{a_i^T x + b_i\}$ 是凸集
- $x \in R^n$ 最大的 r 个分量之和：

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

是凸的（$x_{[i]}$ 是 x 的第 i 个最大元）

证明：

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

逐点上确界

如果对于任意 $y \in \mathcal{A}$，函数 $f(x,y)$ 关于 x 都是凸的，则

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

关于 x 亦是凸的。

例 3

- 集合 C 的支撑函数:$S_C(x) = sup_{y \in C} y^T x$ 是凸的
- 集合 C 的最远点的距离：

$$f(x) = \sup_{y \in C} \|x - y\|$$

- n 阶对称矩阵的最大特征值:$X \in S^n$

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

标量函数复合

$g: R^n \to R, h: R \to R$。

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

$$f\text{ 是凸函数，如果} \begin{cases} g\text{是凸的，}h\text{是凸的，非减} \\ g\text{是凹的，}h\text{是凸的，非增} \end{cases}$$

- 证明（对 $n=1$，g, h 可微的情况）

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

- 扩展值延伸 $\tilde{h}$ 必须保持单调性

例 4

- 如果 g 是凸函数，则 $\exp g(x)$ 是凸函数
- 如果 g 是凹函数且大于零，则 $l/g(x)$ 是凸函数.

向量函数复合

$g: R^n \to R^k, h: R^k \to R$。

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

f 是凸函数，如果 $\begin{cases} g_i\text{是凸的，对每个分量}h\text{是凸的，非减} \\ g_i\text{是凹的，对每个分量}h\text{是凸的，非增} \end{cases}$

证明（对 $n=1$，g, h 可微的情况）

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

例 5

例

- 如果 g_i 是凸函数并且是正的，则 $\sum_{i=1}^{m} \log g_i(x)$ 是凸函数
- 函数 $h(z) = \log(\sum_{i=1}^{k} e^{z_i})$ 是凸函数且在每一维分量上非减，因此只要 g_i 是凸函数，$\log(\sum_{i=1}^{k} e^{g_i})$ 就是凸函数。

1. 30.1 凸函数
2. 30.2 凸函数的保凸运算
3. 30.3 共轭函数

共轭函数

定义 2

共轭函数f 的定义为：

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

- f^* 是凸的 *(*即使 f 不是*)*

例 6

- 非负对数函数 $f(x) = -\log x$

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

- 严格凸二次型 $f(x) = (1/2)x^TQx$, 其中 $Q \in S^n_{++}$

$$f^*(y) = \sup_x (y^Tx - (1/2)x^TQx) = \frac{1}{2}y^TQ^{-1}y$$

性质

- Fenchel 不等式：从共轭函数的定义可知，对于任意 x 和 y 有

$$f(x)+f^*(y)\le y^Tx$$

- 共轭的共轭：若 $f(x)$ 为凸函数，且 $f(x)$ 的上半图是闭集，则有

$$f^{**}=f$$

- 可微函数：设 $f(x)$ 为凸函数，且可微，对于 $z\in R^n$，若

$$y=\nabla f(z)$$

$$\text{则}f^*(y)=z^T\nabla f(z)-f(z)$$

性质

- 伸缩变换和符合仿射变换：设 $A \in R^{n\times n}$ 非奇异，$b \in R^n$，则函数 $g(x) = f(Ax + b)$ 的共轭函数为

$$g^*(y) = f^*(A^{-T}y) - b^T A^{-T} y,$$

其定义域为 $\mathbf{dom}\, g^* = A^T \mathbf{dom}\, f^*$。

- 独立函数的和：如果函数 $f(u, v) = f_1(u) + f_2(v)$，其中 f_1 和 f_2 是凸函数，且共轭函数分别为 f_1^* 和 f_2^*，则

$$f^*(w, z) = f_1^*(w) + f_2^*(z).$$

换言之，独立函数的和的共轭函数是各个凸函数的共轭函数的和。(“独立” 的含义是各个函数具有不同的变量)