凸函数（Convex Function）

张伯雷
南京邮电大学 计算机学院、通达学院
https://bolei-zhang.github.io/course/opt.html
bolei.zhang@njupt.edu.cn

仿射集、凸集、凸锥

- 仿射集
  - 一个集合C是仿射集，则连接集合内任意两点的直线也在仿射集内

$$\forall x_1, x_2 \in \mathrm{C} \Rightarrow x = \theta x_1 + (1-\theta) x_2 \in \mathrm{C}$$

- 凸集
  - 一个集合C是凸集，则集合中任意两点所组成的线段仍然在该集合中

$$\forall x_1, x_2 \in \mathrm{C}, 0 \leq \theta \leq 1. \Rightarrow \ \theta x_1 + (1-\theta) x_2 \in \mathrm{C}$$

- 凸锥
  - 一个集合C是凸锥，则集合中任意两点的非负组合仍然在集合中

$$\forall x_1, x_2 \in \mathrm{C} \Rightarrow \theta_1 \geq 0, \theta_2 \geq, x = \theta_1 x_1 + \theta_2 x_2 \in \mathrm{C}$$

几种重要的凸集

- 超平面、半空间
- 多面体
- 对称半正定矩阵集合
- 球、椭球

凸集的证明

- 1. 根据定义：集合C中任意两点所组成的线段仍然在该集合中

$$x_1, x_2 \in C, 0 \le \theta \le 1. \Rightarrow \theta x_1 + (1-\theta) x_2 \in C$$

- 2.集合 C 可由简单的凸集(超平面、 半空间、 范数球等)经过**保凸运算**后得到
  - 交集
  - 仿射变换

目录

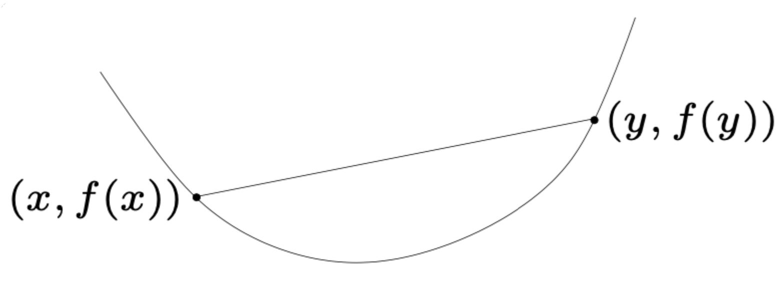
- 凸函数的定义
- 常见的凸函数
- 保持凸性的操作
- 共轭函数

定义1

- 对于一个函数f: $\mathbb{R}^n \to \mathbb{R}$，如果$\mathrm{dom} f$为凸集，且

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

对所有的$x, y \in \mathrm{dom} f$，$0 \le \theta \le 1$成立，则f为（下）凸函数



连接凸函数的图像上任意两点的线段都在函数图像上方

- f是凸函数，则$-f$是凹函数
- 如果$\mathrm{dom} f$ 是凸集，且

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

对所有的$x, y \in \mathrm{dom} f$，$x \neq y$，$0 < \theta < 1$成立，则f是严格凸函数

定义2：一阶条件

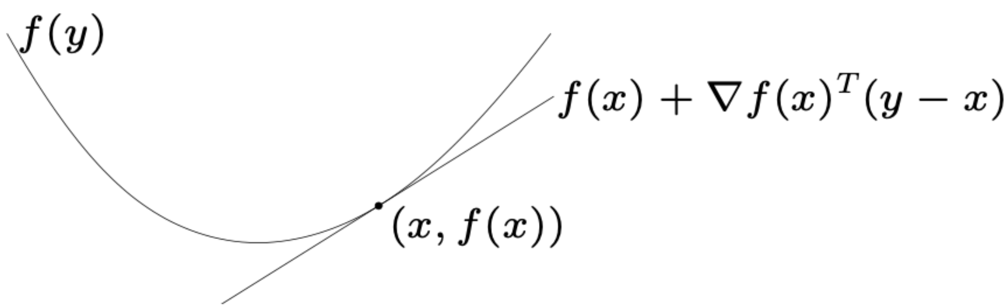
- 可微：如果dom$f$为开集，且函数的导数

$$\nabla f(x) = (\frac{\partial f(x)}{\partial x_1}; \frac{\partial f(x)}{\partial x_2}; \dots; \frac{\partial f(x)}{\partial x_n})$$

对所有定义域中的值都存在，则函数f可微

- 一阶条件：对于可微函数f，如果dom$f$为凸集，则f为凸函数当且仅当

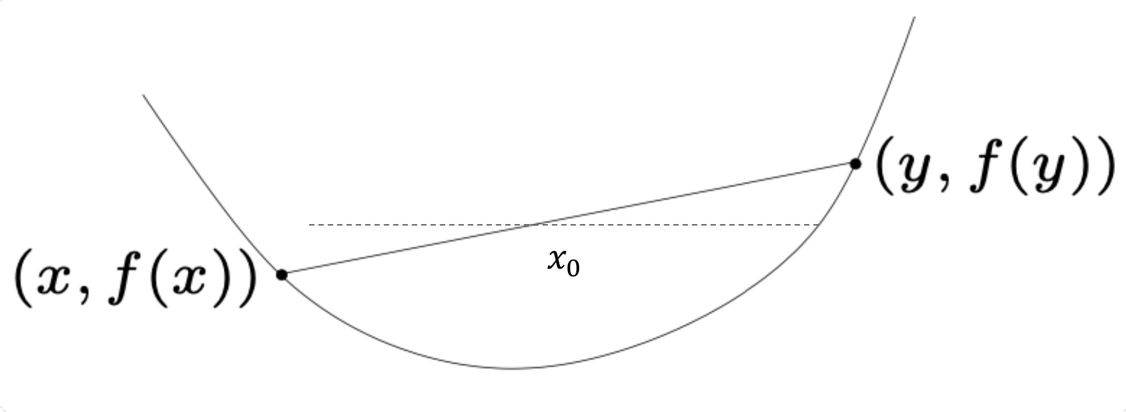
$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \quad \text{for all } x, y \in \mathbf{dom}\, f$$



意义：凸函数在局部的一阶泰勒展开是对函数值的低估

凸函数的最小值

- 假设对于凸函数f，存在某个x_0，使得$\nabla f(x_0) = 0$



则对任意的y， $f(y) \geq f(x_0)$

定义3：二阶条件

- 二阶可微：如果dom$f$为开集，且函数的Hessian矩阵满足$\nabla^2 f(x) \in S^n$

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, i, j = 1, \dots, n$$

对所有定义域中的值都存在，则函数f二阶可微

- 二阶条件：对于二阶可微函数f，如果定义域为凸集
  - f为凸函数当且仅当

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \mathbf{dom}\, f$$

  - 如果对所有定义域内的x满足$\nabla^2 f(x) \succ 0$，则f为严格凸函数
- 几何解释

目录

- 凸函数的定义
- **常见的凸函数**
- 保持凸性的操作
- 共轭函数

一维空间举例

- 凸函数
  - 仿射函数：$ax + b, \forall a, b \in \mathrm{R}$
  - 指数函数：$e^{ax}, \forall a, b \in \mathrm{R}$
  - 幂函数：x^{a} on $\mathrm{R}_{++}, \forall a \geq 1$ or $a \leq 0$
  - 负熵：$x\log x$ on R_{++}
- 凹函数
  - 仿射函数：$ax + b, \forall a, b \in \mathrm{R}$
  - 幂函数：x^{α} on $\mathrm{R}_{++}, \forall 0 \leq \alpha \leq 1$
  - 对数函数：$\log x$ on R_{++}

向量范数(norm)

- 范数是一个函数p，满足以下性质：
  - 非负性：$\forall x \in X, p(x) \geq 0$
  - 齐次性：$p(sx) = |s|p(x)$
  - 正定性： $p(x) = 0 \text{ iff } x = 0$
  - 三角不等式：$p(x+y) \leq p(x) + p(y)$

l_p范数：$||x||_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$

- 常见的范数函数
  - L1范数：$||x||_1 = |x_1| + |x_2| + \cdots + |x_n|$
  - L2范数：$||x||_2 = \sqrt{x_1^2 + \cdots + x_n^2}$
  - l_∞范数：$||x||_\infty = \max(|x_1|, \ldots, |x_n|)$

多维空间举例

- 多维向量R^n
  - 仿射函数：$f(x) = a^T x + b$
  - 范数函数：$||x||_p = (\sum_{i=1}^{n}|x_i|^p)^{\frac{1}{p}}$ for $p \geq 1$;
    - 证明
  - 极大值函数：$f(x) = \max\{\mathrm{x_1}, \ldots, \mathrm{x_n}\}, \mathrm{x} \in R^m$
    - 证明
  - L0范数（x中非零元素的个数）是否为凸函数，请证明(作业)

例-二次函数

- 二次函数：$f(x) = \frac{1}{2}x^T P x + q^T x + r, P \in S^n$

$$\nabla f(x) = Px + q, \qquad \nabla^2 f(x) = P$$

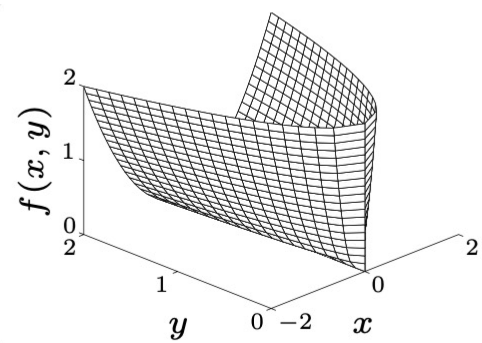
当P为半正定矩阵（$P \succcurlyeq 0$）时为凸函数

- 最小二乘目标函数：$f(x) = ||Ax - b||_2^2$

$$\nabla f(x) = 2A^T(Ax - b), \quad \nabla^2 f(x) = 2A^T A$$

对任意的A都为凸函数

- (作业)二次函数比线性函数：$f(x,y) = \mathrm{x}^2/y$对所有的$y > 0$



半正定矩阵的性质

- 矩阵A为行列式，则存在以下性质：
  - A的特征值均为非负的。
  - 存在n阶实矩阵C，使$A = C^T C$.
  - A的行列式是非负的

凸函数（Convex Function）(2)

张伯雷
南京邮电大学 计算机学院、通达学院
https://bolei-zhang.github.io/course/opt.html
bolei.zhang@njupt.edu.cn

目录

- 凸函数的定义
- 常见的凸函数
- **保持凸性的操作**
- 共轭函数

保持凸性的操作

- 验证一个函数是否是凸函数的方法
  - 1. 根据3种定义
  - 2. 验证函数是由多个相对简单的凸函数通过一些保持凸性的操作来得到

保持凸性的操作1：非负加权和

- 命题：若$f_1, \ldots, f_m$为凸函数，则$f = \sum_{i=1}^{m} w_i f_i$ 为凸函数，$\mathrm{w}_i \geq 0$
- 证明:
  - 1. 定义域是否为凸集
  - 2. 根据定义1

保持凸性的操作2：仿射组合

- 命题：若f为凸函数，则$f(Ax+b)$为凸函数
- 证明：

保持凸性的操作3：最大化/极大化

- 最大化与极大化
  - max：最大值
  - sup：上确界
- 命题：若$f_1, \ldots, f_m$为凸函数，则$f(x) = \max\{f_1(x), \ldots, f_m(x)\}$为凸函数
  - 证明:
- 命题：若$f(x, y)$对所有$y \in A$为关于x的凸函数，则$f(x) = \sup\limits_{y \in A} f(x, y)$为凸函数

保持凸性的操作4：标量函数的组合

- 对于函数$g: R^n \to R^k$与函数$h: R^k \to R$，函数组合为$f(x) = h(g(x))$
- 考虑n=k=1的情况
- f为凸函数如果：$\begin{cases} \text{g是凸函数，h是凸函数，h递增} \\ \text{g是凹函数，h是凸函数，h递减} \end{cases}$

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

- 例：
  - $\exp g(x)$ 当g是凸函数时为凸函数
  - $1/g(x)$当g是凹函数且为正时为凸函数

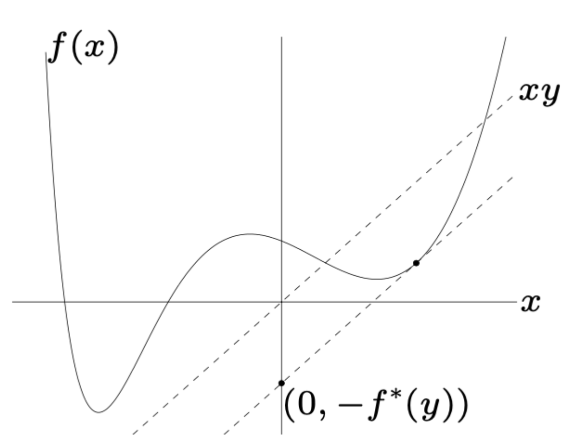
目录

- 凸函数的定义
- 常见的凸函数
- 保持凸性的操作
- **共轭函数**

共轭函数（conjugate function）

- 函数f的共轭定义为

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



- 性质：$f^*(y)$一定为凸函数

例-计算函数的共轭

- 1. $f(x) = ax + b, \mathrm{domf} = \mathrm{R}$

- 2. $f(x) = -\log x\,, \mathrm{domf} = \mathrm{R}_{++}$

作业

- 1：$f(x) = x^{-2}, \mathrm{x} \neq 0$，是否为凸函数，请说明理由

- 2：0范数：$||\mathrm{x}||_0 = \mathrm{x}$中非零元素的个数 ， 是否为范数，是否为凸函数，写出证明过程

- 3. 求证：$f(x, y) = \mathrm{x}^2/y$对所有的$y > 0$都为凸函数

- 4*. 已知$\forall x, v \in dom\, f, g(t) = f(x + tv)$是凸函数，求证f是凸函数

谢谢！！