凸函数（Convex Function）(1)

张伯雷
南京邮电大学 计算机学院
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

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

- 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

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

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

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

- 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

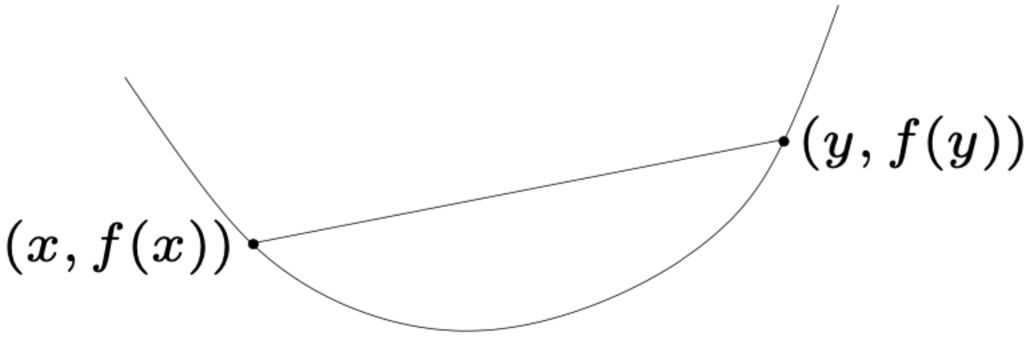
- 凸函数的定义
- 常见的凸函数

定义1

- 对于一个函数$f\colon \mathbb{R}^n \to \mathbb{R}$，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$x, y \in \mathrm{dom}f$，$x \ne y$，$0 < \theta < 1$成立，则f是严格凸函数

定义2

- f是凸函数的充分必要条件为：$\forall x, v \in dom\, f, g(t) = f(x + tv)$是凸函数
- 证明

定义3：一阶条件

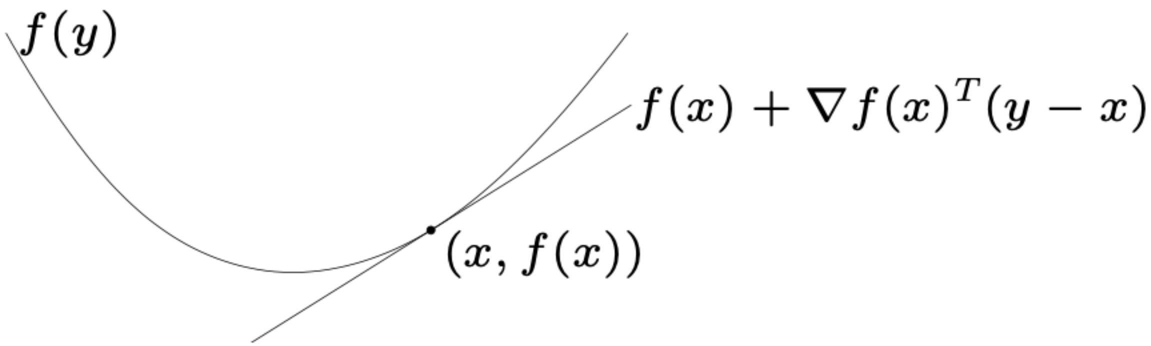
- 可微：如果dom$f$为开集，且函数的导数

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n}\right)$$

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

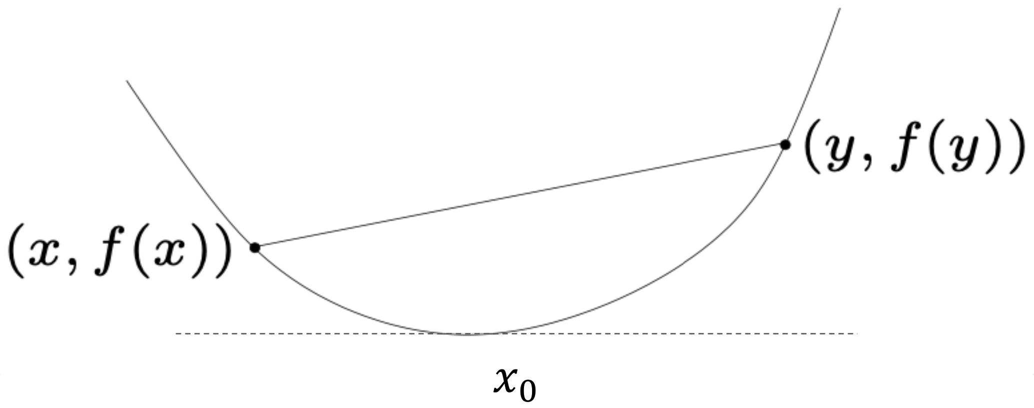
一阶条件的证明

- =>

- <=

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4：二阶条件

- 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

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

  - 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

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： $p(x) = 0 \; iff \; x = 0$
  - 三角不等式：$p(x+y) \leq p(x) + p(y)$
- 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_p范数：$||x||_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$
  - 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：$f(x) = \max\{\mathrm{x}_1, \ldots, \mathrm{x}_\mathrm{n}\}, \mathrm{x} \in R^m$
    - 证明
  - L0范数（作业）：x中非零元素的个数

例-二次函数

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

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

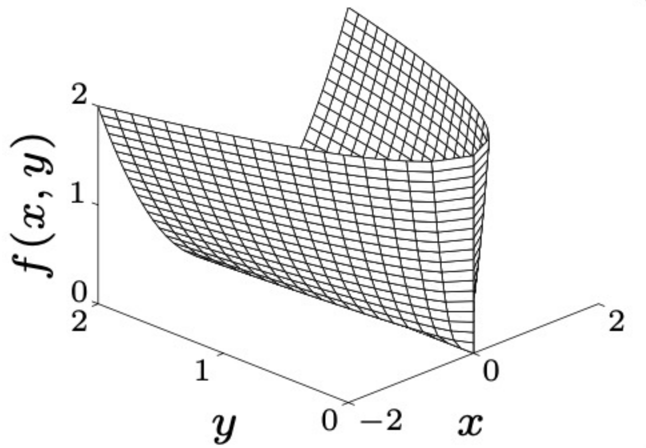
当P为半正定矩阵（$P \succcurlyeq 0$）时为凸函数

- 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$y > 0$都为凸函数



例

- log-sum-up（极大值函数的解析逼近）

$$f(x) = \log(e^{x_1} + \cdots + e^{x_n}), \mathrm{x} \in R^m$$

  - 性质： $\max\{x_1, \dots, x_n\} \le f(x) \le \max\{x_1, \dots, x_n\} + \log n$

- 证明

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