第二章 基础知识

修贤超

https://xianchaoxiu.github.io

目录

- 2.1 范数
- 2.2 导数
- 2.3 广义实值函数
- 2.4 凸集
- 2.5 凸函数
- 2.6 共轭函数
- 2.7 次梯度

凸函数的定义

- 设 $f:\mathbb{R}^n \to \mathbb{R}$ 为适当函数，如果 $\mathrm{dom}\ f$ 是凸集，且

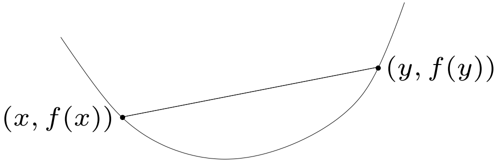
$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)$$

对所有 $x, y \in \mathrm{dom}\ f$, $0 \le \theta \le 1$ 都成立，则称 f 是**凸函数**

- 若对所有 $x, y \in \mathrm{dom}\ f$，$x \neq y$，$0 < \theta < 1$，有

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

则称 f 是**严格凸函数**



一元凸函数的例子

- **仿射函数**　对任意 $a, b \in$，$ax + b$ 是 $\mathbb{R}$ 上的凸函数
- **指数函数**　对任意 $a \in$，e^{ax} 是 $\mathbb{R}$ 上的凸函数
- **幂函数**　对 $\alpha \geq 1$ 或 $\alpha \leq 0$，x^α 是 $\mathbb{R}_{++}$ 上的凸函数
- **绝对值的幂**　对 $p \geq 1$，$|x|^p$ 是 $\mathbb{R}$ 上的凸函数
- **负熵**　$x \log x$ 是 $\mathbb{R}_{++}$ 上的凸函数
- **仿射函数**　对任意 $a, b \in \mathbb{R}$，$ax + b$ 是 $\mathbb{R}$ 上的凹函数
- **幂函数**　对 $0 \leq \alpha \leq 1$，x^α 是 $\mathbb{R}_{++}$ 上的凹函数
- **对数函数**　$\log x$ 是 $\mathbb{R}_{++}$ 上的凹函数

多元凸函数的例子

- 所有的仿射函数既是凸函数，又是凹函数

$$f(x) = a^\top x + b$$

$$f(X) = \mathrm{Tr}(A^\top X) + b = \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}X_{ij} + b$$

- 所有的范数都是凸函数

$$f(x) = \|x\|_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p} \ (p \geq 1)$$

$$f(X) = \|X\|_2 = \sigma_{\max}(X) = (\lambda_{\max}(X^\top X))^{1/2}$$

强凸函数

- 若存在常数 $m > 0$，使得

$$g(x) = f(x) - \frac{m}{2}\|x\|^2$$

为凸函数，则称 $f(x)$ 为**强凸函数**，其中 m 为强凸参数

- 若存在常数 $m > 0$，使得对任意 $x, y \in \operatorname{dom} f$ 以及 $\theta \in (0,1)$，有

$$f(\theta x + (1-\theta y)) \le \theta f(x) + (1-\theta) f(y) - \frac{m}{2}\theta(1-\theta)\|x-y\|^2,$$

则称 $f(x)$ 为**强凸函数**，其中 m 为强凸参数

- 为了方便也称 $f(x)$ 为 m-强凸函数
- 设 f 为强凸函数且存在最小值，则 f 的最小值点唯一

凸函数判定定理

- 设$f:\mathbb{R}^n\to\mathbb{R}$ 是凸函数，当且仅当对每个 $x\in\mathrm{dom}\ f v\in\mathbb{R}^n$，函数 $g:\mathbb{R}\to\mathbb{R}$ 是关于 t 的凸函数

$$g(t)=f(x+tv),\quad \mathrm{dom}\ g=\{t\mid x+tv\in\mathrm{dom}\ f\}$$

- $f(X)=-\log\det X$ 是凸函数，其中 $\mathrm{dom}\ f=\mathcal{S}^n_{++}$

  证明　任取 $X\succ 0$ 以及方向 $V\in\mathcal{S}^n$，将 f 限制在直线 $X+tV$（t 满足 $X+tV\succ 0$）上，那么

$$\begin{aligned}g(t)=-\log\det(X+tV)&=-\log\det X-\log\det(I+tX^{-1/2}VX^{-1/2})\\&=-\log\det X-\sum_{i=1}^{n}\log(1+t\lambda_i)\end{aligned}$$

  其中 λ_i 是 $X^{-1/2}VX^{-1/2}$ 第 i 个特征值. 对每个 $X\succ 0$ 以及方向 V，g 关于 t 是凸的，因此 f 是凸的

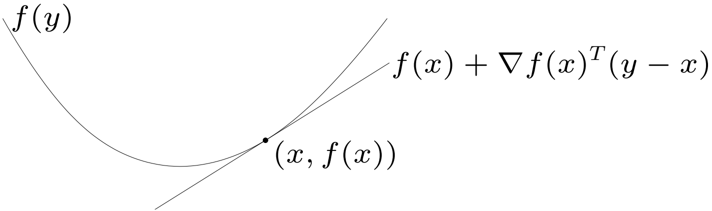
一阶条件

- 凸集上的可微函数 f 是凸函数当且仅当

$$f(y) \geq f(x) + \nabla f(x)^\top (y - x) \quad \forall x, y \in \mathrm{dom}\ f$$

- 设 f 为可微函数，则 f 为凸函数当且仅当 $\mathrm{dom}\ f$ 为凸集且 ∇f 为单调映射

$$(\nabla f(x) - \nabla f(y))^\top (x - y) \geq 0, \quad \forall\ x, y \in\ \mathrm{dom}\ f$$



二阶条件

- 设f 为定义在凸集上的二阶连续可微函数, f 是凸函数当且仅当

  $$\nabla^2 f(x) \succeq 0 \quad \forall x \in \mathrm{dom}\ f$$

  如果 $\nabla^2 f(x) \succ 0\ \forall x \in \mathrm{dom}\ f$，则 f 是**严格凸函数**

- 最小二乘函数 $f(x) = \|Ax - b\|_2^2$

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

  对任意 A，函数 f 都是凸函数

- 二次函数 $f(x) = (1/2)x^\top Px + q^\top x + r$ (其中 $P \in \mathcal{S}^n$)

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

  f 是凸函数当且仅当 $P \succeq 0$

上方图

- **函数 $f(x)$ 为凸函数当且仅当其上方图 $\mathrm{epi} f$ 是凸集**

必要性 若 f 为凸函数，则对任意 $(x_1, y_1), (x_2, y_2) \in \mathrm{epi} f, t \in [0, 1]$,

$$ty_1 + (1-t)y_2 \geq tf(x_1) + (1-t)f(x_2) \geq f(tx_1 + (1-t)x_2),$$

故 $(tx_1 + (1-t)x_2, ty_1 + (1-t)y_2) \in \mathrm{epi} f, t \in [0, 1]$

充分性 若 $\mathrm{epi} f$ 是凸集，则对任意 $x_1, x_2 \in \mathrm{dom}\ f, t \in [0, 1]$,

$$\begin{aligned} &(tx_1 + (1-t)x_2, tf(x_1) + (1-t)f(x_2)) \in \mathrm{epi} f \\ \Rightarrow\ & f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2). \end{aligned}$$

凸函数的判断方法

- 用定义验证（通常将函数限制在一条直线上）
- 利用一阶条件、二阶条件
- 直接研究 f 的上方图 $\mathrm{epi}\ f$
- 说明 f 可由简单的凸函数通过一些保凸的运算得到
  - 非负加权和
  - 与仿射函数的复合
  - 逐点取最大值
  - 与标量、向量函数的复合

非负加权和与仿射函数的复合

- 若 f 是凸函数，则 αf 是凸函数，其中 $\alpha \geq 0$
- 若 f_1, f_2 是凸函数，则 $f_1 + f_2$ 是凸函数
- 若 f 是凸函数，则 $f(Ax + b)$ 是凸函数
- 线性不等式的对数障碍函数

$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^\top x), \quad \mathrm{dom}\ f = \{x \mid a_i^\top x < b_i, i = 1, ..., m\}$$

- 仿射函数的（任意）范数 $f(x) = \|Ax + b\|$

逐点取最大值

- 若 $f_1, \cdots, f_m$ 是凸函数，则 $f(x) = \max\{f_1(x), \cdots, f_m(x)\}$ 是凸函数
- 分段线性函数
$$f(x) = \max_{i=1,\cdots,m}(a_i^\top x + b_i)$$
- $x \in \mathbb{R}^n$ 的前 r 个最大分量之和
$$f(x) = x_{[1]} + x_{[2]} + \cdots + x_{[r]}$$
事实上，$f(x)$ 可以写成如下多个线性函数取最大值的形式
$$f(x) = \max\{x_{i_1} + x_{i_2} + \cdots + x_{i_r} \mid 1 \leq i_1 < i_2 < \cdots < i_r \leq n\}$$

逐点取上界

- 若对每个 $y \in \mathcal{A}$，$f(x,y)$ 是关于 x 的凸函数，则
$$g(x) = \sup_{y\in\mathcal{A}} f(x,y)$$
是凸函数
- 集合 C 的支撑函数
$$S_C(x) = \sup_{y\in C} y^\top x$$
- 集合 C 点到给定点 x 的最远距离
$$f(x) = \sup_{y\in C} \|x-y\|$$
- 对称矩阵 $X \in \mathcal{S}^n$ 的最大特征值
$$\lambda_{\max}(X) = \sup_{\|y\|_2=1} y^\top X y$$

与标量函数的复合

- 给定函数 $g:\mathbb{R}^n \to \mathbb{R}$ 和 $h:\mathbb{R}\to\mathbb{R}$,

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

若 $\begin{array}{l} g \text{ 是凸函数}, h \text{ 是凸函数}, \tilde{h} \text{ 单调不减} \\ g \text{ 是凹函数}, h \text{ 是凸函数}, \tilde{h} \text{ 单调不增} \end{array}$，那么 f 是凸函数

证明 对 $n=1$，g,h 均可微的情形

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

- 如果 g 是凸函数，则 $\exp g(x)$ 是凸函数
- 如果 g 是正值凹函数，则 $1/g(x)$ 是凸函数

与向量函数的复合

- 给定函数 $g:\mathbb{R}^n\to\mathbb{R}^k$ 和 $h:\mathbb{R}^k\to\mathbb{R}$:

$$f(x)=h(g(x))=h(g_1(x),g_2(x),...,g_k(x))$$

若 $\begin{array}{l} g_i\ \text{是凸函数}, h\ \text{是凸函数}, \tilde{h}\ \text{关于每个分量单调不减} \\ g_i\ \text{是凹函数}, h\ \text{是凸函数}, \tilde{h}\ \text{关于每个分量单调不增} \end{array}$，那么 f 是凸函数

证明 对 $n=1$，g,h 均可微的情形

$$f''(x)=g'(x)^\top\nabla^2h(g(x))g'(x)+\nabla h(g(x))^Tg''(x)$$

- 如果 g_i 是正值凹函数，则 $\sum_{i=1}^m\log g_i(x)$ 是凹函数
- 如果 g_i 是凸函数，则 $\log\sum_{i=1}^m\exp g_i(x)$ 是凸函数

取下确界

- 若 $f(x,y)$ 关于 (x,y) 整体是凸函数，C 是凸集，则

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

  是凸函数

- 考虑函数 $f(x,y) = x^\top Ax + 2x^\top By + y^\top Cy$，海瑟矩阵满足

$$\begin{bmatrix} A & B \\ B^\top & C \end{bmatrix} \succeq 0, \quad C \succ 0,$$

  则 $f(x,y)$ 为凸函数. 对 y 求最小值得

$$g(x) = \inf_y f(x,y) = x^\top (A - BC^{-1}B^\top)x,$$

  因此 g 是凸函数. 进一步地，A 的 Schur 补 $A - BC^{-1}B^\top \succeq 0$

- 点 x 到凸集 S 的距离 $\mathrm{dist}(x,S) = \inf_{y \in S} \|x - y\|$ 是凸函数

透视函数

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

$$g(x,t) = tf(x/t), \quad \mathrm{dom}\ g = \{(x,t) | x/t \in \mathrm{dom}\ f, t > 0\}$$

  若 f 是凸函数，则 g 是凸函数
- $f(x) = x^\top x$ 是凸函数，则 $g(x,t) = x^\top x/t$ 是区域 $\{(x,t) \mid t > 0\}$ 上的凸函数
- $f(x) = -\log x$ 是凸函数，则 $g(x,t) = t\log t - t\log x$ 是 $\mathbb{R}^2_{++}$ 上的凸函数
- 若 f 是凸函数，则

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

  是区域 $\{x \mid c^\top x + d > 0, (Ax+b)/(c^\top x + d) \in \mathrm{dom}\ f\}$ 上的凸函数

目录

- 2.1 范数
- 2.2 导数
- 2.3 广义实值函数
- 2.4 凸集
- 2.5 凸函数
- 2.6 共轭函数
- 2.7 次梯度

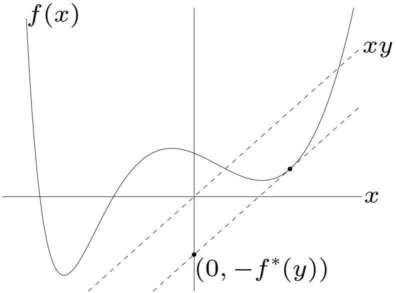
共轭函数

- **定义 2.19** 适当函数 f 的**共轭函数**定义为

$$f^*(y) = \sup_{x \in \mathrm{dom}\ f} \left(y^\top x - f(x) \right)$$

- f^* 恒为凸函数，无论 f 是否是凸函数 **命题 2.5** Fenchel 不等式

$$f(x) + f^*(y) \geq x^\top y \quad \forall x, y$$



例 2.6

- 考察二次函数

$$f(x) = \frac{1}{2}x^\top Ax + b^\top x + c$$

- 强凸情形（$A \succ 0$）

$$f^*(y) = \frac{1}{2}(y-b)^\top A^{-1}(y-b) - c$$

- 一般凸情形（$A \succeq 0$）

$$f^*(y) = \frac{1}{2}(y-b)^\top A^\dagger(y-b) - c, \quad \mathrm{dom}\ f^* = \mathcal{R}(A) + b$$

这里 $\mathcal{R}(A)$ 为 A 的像空间

例 2.7-2.8

- 考凸集 C 的示性函数

$$f(x) = \begin{cases} 0, & x \in C \\ +\infty, & x \notin C \end{cases}$$

共轭函数

$$f^*(y) = \sup_{x \in C} y^\top x$$

- 范数

$$f(x) = \|x\|$$

共轭函数

$$f^*(y) = \begin{cases} 0, & \|y\|_* \leq 1 \\ +\infty, & \|y\|_* > 1 \end{cases}$$

二次共轭函数

- **定义 2.20** 任一函数 f 的二次共轭函数定义为

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

- **定理 2.12** 若 f 为闭凸函数，则

$$f^{**}(x) = f(x)$$

- 如果 f 是闭凸函数，则

$$y \in \partial f(x) \quad \Leftrightarrow \quad x \in \partial f^*(y) \quad \Leftrightarrow \quad x^\top y = f(x) + f^*(y)$$

目录

- 2.1 范数
- 2.2 导数
- 2.3 广义实值函数
- 2.4 凸集
- 2.5 凸函数
- 2.6 共轭函数
- 2.7 次梯度

次梯度

- **可微凸函数 f 的一阶条件**

$$f(y) \geq f(x) + \nabla f(x)^\top (y - x)$$

- **设 f 为适当凸函数，$x \in \mathrm{dom}\ f$，若向量 $g \in \mathbb{R}^n$ 满足**

$$f(y) \geq f(x) + g^\top (y - x), \quad \forall y \in \mathrm{dom}\ f$$

则称 g 为函数 f 在点 x 处的一个次梯度

- **进一步地，称集合**

$$\partial f(x) = \{g \mid g \in \mathbb{R}^n, f(y) \geq f(x) + g^\top (y - x), \forall y \in \mathrm{dom}\ f\}$$

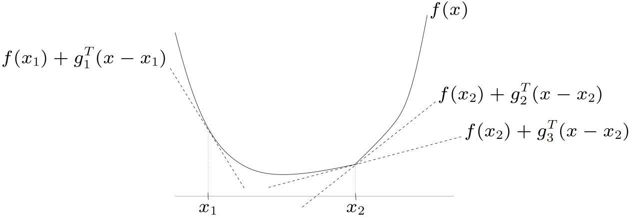
为 f 在点 x 处的次微分

次梯度

- 知$f(x) + g^\top(y - x)$ 是 $f(y)$ 的一个全局下界
- g 可以诱导出上方图 $\mathrm{epi}\ f$ 在点 $(x, f(x))$ 处的一个支撑超平面

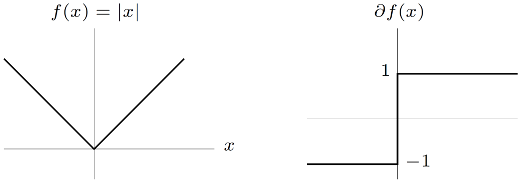
$$\begin{bmatrix} g \\ -1 \end{bmatrix}\left(\begin{bmatrix} y \\ t \end{bmatrix} - \begin{bmatrix} x \\ f(x) \end{bmatrix}\right) \le 0 \quad \forall\ (y, t) \in \mathrm{epi}\ f$$

- 如果 f 是可微凸函数, 那么 $\nabla f(x)$ 是 f 在点 x 处的一个次梯度
- g_2, g_3 是点 x_2 处的次梯度, g_1 是点 x_1 处的次梯度



次梯度存在性

- 设f 为凸函数，$\mathrm{dom}\ f$ 为其定义域. 如果 $x \in \mathrm{int\ dom}\ f$，则 $\partial f(x)$ 是非空的，其中 $\mathrm{int\ dom}\ f$ 的含义是集合 $\mathrm{dom}\ f$ 的所有内点.
- 绝对值函数 $f(x) = |x|$



- 欧几里得范数 $f(x) = \|x\|_2$

$$\text{如果 } x \neq 0, \partial f(x) = \frac{1}{\|x\|_2} x, \quad \text{如果 } x = 0, \partial f(x) = \{g \mid \|g\|_2 \leq 1\}$$

次梯度的性质

- 对任何 $x \in \mathrm{dom}\ f$，$\partial f(x)$ 是一个闭凸集（可能为空集）
- 如果 $x \in \mathrm{int\ dom}\ f$，则 $\partial f(x)$ 非空有界集
- 设凸函数 $f(x)$ 在 $x_0 \in \mathrm{int\ dom}\ f$ 处可微，则 $\partial f(x_0) = \{\nabla f(x_0)\}$
- 设 $f: \mathbb{R}^n \to \mathbb{R}$ 为凸函数，$x, y \in \mathrm{dom}\ f$，则 $(u-v)^\top (x-y) \geq 0$, 其中 $u \in \partial f(x)$，$v \in \partial f(y)$
- 设 $f(x)$ 是闭凸函数且 ∂f 在点 $\bar{x}$ 附近存在且非空．若序列 $x^k \to \bar{x}$，$g^k \in \partial f(x^k)$ 为 $f(x)$ 在点 x^k 处的次梯度，且 $g^k \to \bar{g}$，则 $\bar{g} \in \partial f(\bar{x})$

方向导数

- 设f 为适当函数，给定点 x_0 以及方向 $d \in \mathbb{R}^n$，方向导数（若存在）定义为

$$\lim_{t\downarrow 0} \phi(t) = \lim_{t\downarrow 0} \frac{f(x_0+td)-f(x_0)}{t}$$

  其中 $t\downarrow 0$ 表示 t 单调下降趋于 0
- 若 f 是凸函数，则 $\phi(t)$ 在 $(0,+\infty)$ 上是单调不减的，$\lim$ 可替换为 $\inf$
- 对于凸函数 f，给定点 $x_0 \in \mathrm{dom}\ f$ 以及方向 $d \in \mathbb{R}^n$，其方向导数定义为

$$\partial f(x_0; d) = \inf_{t>0} \frac{f(x_0+td)-f(x_0)}{t}.$$

方向导数有限

- 设$f(x)$ 为凸函数，$x_0 \in \mathrm{int\ dom}\ f$，则对任意 $d \in \mathbb{R}^n$，$\partial f(x_0; d)$ 有限

  证明 首先 $\partial f(x_0; d)$ 不为正无穷是显然的. 由于 $x_0 \in \mathrm{int\ dom}\ f$，根据次梯度的存在性定理可知 $f(x)$ 在点 x_0 处存在次梯度 g. 根据方向导数的定义，有

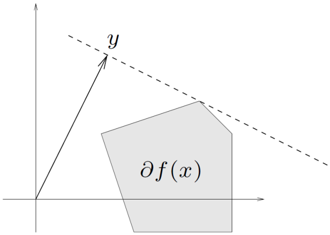
$$
\begin{aligned}
\partial f(x_0; d) &= \inf_{t>0} \frac{f(x_0 + td) - f(x_0)}{t} \\
&\geq \inf_{t>0} \frac{tg^\top d}{t} = g^\top d
\end{aligned}
$$

  其中的不等式利用了次梯度的定义. 这说明 $\partial f(x_0; d)$ 不为负无穷

方向导数和次梯度

- 设 $f:\mathbb{R}^n \to (-\infty,+\infty]$ 为凸函数，$x_0 \in \mathrm{int\ dom}\ f$，$d$ 为 $\mathbb{R}^n$ 中任一方向，则

$$\partial f(x_0;d) = \max_{g\in\partial f(x_0)} g^\top d$$



- $\partial f(x;y)$ 是 $\partial f(x)$ 的支撑函数
- $\partial f(x_0;d) = \nabla f(x_0)^\top d$, 切对所有的 $x_0 \in \mathrm{int\ dom}\ f$ 以及所有的 d 都存在

次梯度的计算规则

- 若凸函数 f 在点 x 处可微，则 $\partial f(x)=\{\nabla f(x)\}$
- 设凸函数 f_1,f_2 满足 $\mathrm{int\ dom}\ f_1\cap\mathrm{dom}\ f_2\neq\emptyset$，而 $x\in\mathrm{dom}\ f_1\cap\mathrm{dom}\ f_2$. 若

$$f(x)=\alpha_1 f_1(x)+\alpha_2 f_2(x),\quad \alpha_1,\alpha_2\geq 0$$

  则 $f(x)$ 的次微分

$$\partial f(x)=\alpha_1\partial f_1(x)+\alpha_2\partial f_2(x)$$

- 设 h 为适当凸函数，f 满足 $f(x)=h(Ax+b)$. 若存在 $x^\sharp\in\mathbb{R}^m$，使得 $Ax^\sharp+b\in\mathrm{int\ dom}\ h$，则

$$\partial f(x)=A^\top\partial h(Ax+b),\quad \forall\ x\in\mathrm{int\ dom}\ f$$

两个函数之和的次梯度

- 设 $f_1, f_2 : \mathbb{R}^n \to (-\infty, +\infty]$ 是两个凸函数，则对任意的 $x_0 \in \mathbb{R}^n$，

$$\partial f_1(x_0) + \partial f_2(x_0) \subseteq \partial(f_1 + f_2)(x_0).$$

进一步地，若 $\mathrm{int\ dom}\ f_1 \cap \mathrm{dom}\ f_2 \neq \emptyset$，则对任意的 $x_0 \in \mathbb{R}^n$，

$$\partial(f_1 + f_2)(x_0) = \partial f_1(x_0) + \partial f_2(x_0).$$

证明 对于任意给定的 x_0，设 $g \in \partial(f_1 + f_2)(x_0)$。如果 $f_1(x_0) = +\infty$，则 $(f_1 + f_2)(x_0) = +\infty$。由次梯度的定义，我们有

$$(f_1 + f_2)(x) \geq (f_1 + f_2)(x_0) + g^\top(x - x_0)$$

对任意 $x \in \mathbb{R}^n$ 成立，故 $f_1 + f_2 \equiv +\infty$。这与 $\mathrm{int\ dom}\ f_1 \cap \mathrm{dom}\ f_2 \neq \emptyset$ 矛盾，因此以下我们假设 $f_1(x_0), f_2(x_0) < +\infty$

函数族的上确界

- 设 $f_1, f_2, \cdots, f_m : \mathbb{R}^n \to (-\infty, +\infty]$ 均为凸函数，令

$$f(x) = \max\{f_1(x), f_2(x), \cdots, f_m(x)\}, \quad \forall x \in \mathbb{R}^n$$

对 $x_0 \in \bigcap_{i=1}^{m} \text{int dom } f_i$，定义 $I(x_0) = \{i \mid f_i(x_0) = f(x_0)\}$，则

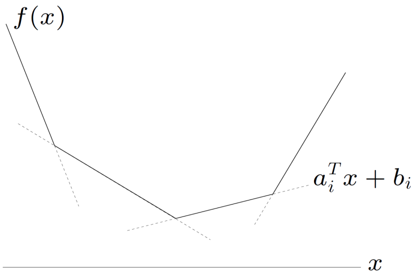
$$\partial f(x_0) = \text{conv} \bigcup_{i \in I(x_0)} \partial f_i(x_0)$$

- $I(x_0)$ 表示点 x_0 处"有效"函数的指标
- $\partial f(x_0)$ 是点 x_0 处"有效"函数的次微分并集的凸包
- 如果 f_i 可微, $\partial f(x_0) = \text{conv}\{\nabla f_i(x_0) \mid i \in I(x_0)\}$

例子

- 分段线性函数

$$f(x) = \max_{i=1,2,\cdots,m} \{a_i^\top x + b_i\}$$



- 点 x 处的次微分是一个多面体

$$\partial f(x) = \operatorname{conv}\{a_i \mid i \in I(x)\}$$

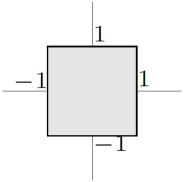
其中 $I(x) = \{i \mid a_i^\top x + b_i = f(x)\}$

例子

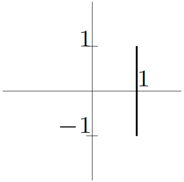
- ℓ_1-范数

$$f(x) = \|x\|_1 = \max_{s \in \{-1,1\}^n} s^\top x$$

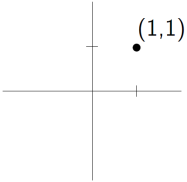
$$\partial f(x) = J_1 \times \cdots \times J_n, \quad J_k = \begin{cases} [-1,1], & x_k = 0 \\ \{1\}, & x_k > 0 \\ \{-1\}, & x_k < 0 \end{cases}.$$



$\partial f(0,0) = [-1,1] \times [-1,1]$



$\partial f(1,0) = \{1\} \times [-1,1]$



$\partial f(1,1) = \{(1,1)\}$

复合函数

- 设 $f_1, f_2, \cdots, f_m : \mathbb{R}^n \to (-\infty, +\infty]$ 为 m 个凸函数，$h : \mathbb{R}^m \to (-\infty, +\infty]$ 为关于各分量单调递增的凸函数，令

$$f(x) = h(f_1(x), f_2(x), \cdots, f_m(x))$$

- $z = (z_1, z_2, \cdots, z_m) \in \partial h(f_1(\hat{x}), f_2(\hat{x}), \cdots, f_m(\hat{x}))$ 以及 $g_i \in \partial f_i(\hat{x})$
- $g z_1 g_1 + z_2 g_2 + \cdots + z_m g_m \in \partial f(\hat{x})$

证明

$$\begin{aligned} f(x) &\geq h(f_1(\hat{x}) + g_1^\top (x - \hat{x}), f_2(\hat{x}) + g_2^\top (x - \hat{x}), \cdots, f_m(\hat{x}) + g_m^\top (x - \hat{x})) \\ &\geq h(f_1(\hat{x}), f_2(\hat{x}), \cdots, f_m(\hat{x})) + \sum_{i=1}^{m} z_i g_i^\top (x - \hat{x}) \\ &= f(\hat{x}) + g^\top (x - \hat{x}) \end{aligned}$$

Q&A

Thank you!

感谢您的聆听和反馈