2022秋季 学院路校区(凸优化与近似)
最优化方法 第 2 次作业
提交日期：2022 年10 月11日周二课前

2022 年 9 月 26 日

1. 请依次完成以下问题(上次做过的同学请忽略)：

   (a) 设$\{f_i\}_{i\in I}$是$\mathbb{R}^n$上的凸函数族(可能不可数无穷多). 证明

   $$f(x) := \sup_{i\in I} f_i(x)$$

   在$\mathbb{R}^n$上是凸函数. (提示：可用定义，或者第2 题的结论和凸集的交还是凸集证明)

   (b) 设$x \in \mathbb{R}^n$使得$\mathcal{A}(x) := \{i \in I : f(x) = f_i(x)\} \neq \emptyset$, 此即$f$的定义中，取到上确界的函数指标集. 证明

   $$\left\{\sum_{i\in I_k} \theta_i g_i : \forall k \in \mathbb{Z}_{++}, g_i \in \partial f_i(x), \theta_i \geq 0, \forall i \in I_k \subseteq \mathcal{A}(x), |I_k| = k, \sum_{i\in I_k} \theta_i = 1\right\} \subseteq \partial f(x).$$

   并分别对函数$|x| = \max\{x, -x\}$和$\max\{0, x\}$ 应用该结论.

   (c) 证明凸函数的和是凸函数；凸函数的非负倍数是凸函数.

   (d) (Nesterov构造的世界最差函数) 已知$t < n$. 对于某$\gamma, \alpha \in \mathbb{R}_{++}$，考虑函数[1]

   $$f(x) = \gamma \max_{1\leq i\leq t} x[i] + \tfrac{\alpha}{2}\|x\|^2.$$

   这里$\alpha > 0, \gamma > 0$是待定参数.

   - 由(a)和(c)说明函数f是凸函数.
   - 由(b)给出$\partial f(x)$的计算公式.

   $$\partial f(x) = \alpha x + \gamma \mathrm{conv}\Big\{e_i \colon i \in \operatorname*{argmax}_{1\leq j\leq t} x[j]\Big\}.$$

[1]为了避免与迭代指标混淆，这里用$x[i]$ 表示x 的第i 个分量.

- 已知$x_* \in \mathbb{R}^n$ 满足

$$x_*[i] = \begin{cases} -\frac{\gamma}{\alpha t} & \text{如果} 1 \le i \le t \\ 0 & \text{否则}. \end{cases}$$

  对于$t = 2$，请给出$\partial f(x_*)$的显式表示. 请问$0 \in \partial f(x_*)$吗？为什么？给出函数的极小点和极小值.

2. 判断下面的函数在指定区域上是否是凸函数，并给出理由:

   (a) $f(x) \equiv 1$ 在$\mathbb{R}$上.

   (b) $f(x) = x$ 在$\mathbb{R}$上.

   (c) $f(x) = |x|$ 在$\mathbb{R}$上.

   (d) $f(x) = -|x|$ 在$\mathbb{R}$上.

   (e) $f(x) = -|x|$ 在$\mathbb{R}_+ = \{x \in \mathbb{R} : x \ge 0\}$ 上.

   (f) $\exp\{x\}$在$\mathbb{R}$上.

   (g) $\exp\{x^2\}$在$\mathbb{R}$上.

   (h) $\exp\{-x^2\}$在$\mathbb{R}$上.

   (i) $\exp\{-x^2\}$在$\{x \in \mathbb{R} : x \ge 100\}$ 上.

3. 证明下面的函数在指定区域上是凸函数:

   (a) $\frac{x^2}{y}$ 在$Q = \{(x, y) \in \mathbb{R}^2 : y > 0\}$ 上.

   (b) $\ln(\exp\{x\} + \exp\{y\})$在 2D 平面上.

4. 对于n维向量$\boldsymbol{x}$, 设$\hat{\boldsymbol{x}} = (\hat{x}^1, \cdots, \hat{x}^n)$ 是对向量$\boldsymbol{x}$ 的分量以非增方式进行排序后得到的新向量. 例如$\boldsymbol{x} = (2, 1, 3, 1), \hat{\boldsymbol{x}} = (3, 2, 1, 1)$. 固定$k, 1 \le k \le n$.

   (a) 函数$\hat{x}^k$($\boldsymbol{x}$的第k大元素)是$\boldsymbol{x}$ 的凸函数吗?

   (b) 函数$s_k(\boldsymbol{x}) = \hat{x}^1 + \cdots + \hat{x}^k$($\boldsymbol{x}$中前$k$ 大元素之和)是凸函数吗？

5. $f(x) = x^4$是强凸函数吗？为什么？

6. 已知$y \in \mathbb{R}^n$. 写出求$y \in \mathbb{R}^n$到以下集合的欧氏投影对应的问题，并求解:

   (a) $\Omega = \mathbb{R}^n_+$.

   (b) 集合(ℓ_∞空间中的单位球)

$$\Omega = \left\{ x \in \mathbb{R}^d : \max_{1 \le i \le d} |x_i| \le 1 \right\}.$$

(c) 考虑超平面$H=\{x\in\mathbb{R}^n: s^Tx=c\}$，其中$s\in\mathbb{R}^n\backslash\{\mathbf{0}\}$ 并且$c\in\mathbb{R}$是已知的. 设$x\in\mathbb{R}^n$ 是任意的. 找到用s 和c表示的$\Pi_H(x)$ 的公式，并证明它的正确性.

7. 设$\Omega\in\mathbb{R}^n$是闭凸集，并且对任意$x\in\mathbb{R}^n$, 用$\pi(x)$ 表示它在Ω 上的投影. 证明对任意$x,y\in\mathbb{R}^n$,

$$\|\pi(x)-\pi(y)\|\le\|x-y\|$$

成立，其中$\|\cdot\|$表示欧氏范数.

证明对任意$y\in\Omega$有

$$\|\pi(x)-y\|\le\|x-y\|.$$

8. [关于梯度下降] 想求解约束优化问题：

$$\min_{x\in\Omega} f(x),$$

其中f是L-Lipschitz凸函数，$\Omega\subset\mathbb{R}^n$为一个直径(在欧氏范数下)最多为$R$ 的有界闭凸集. 记x^*为f 在Ω 上的一个极小点. 假设我们将投影梯度下降算法中的更新替换为

$$\begin{aligned} &y_{s+1}=x_s-\eta\tfrac{g_s}{\|g_s\|},\ g_s\in\partial f(x_s).\\ &x_{s+1}=\pi_\Omega(y_{s+1}), \end{aligned} \tag{1}$$

其中$\pi_\Omega(\cdot)$为在Ω上的投影. 在相同的假设下，可以为该算法证明哪些保证？

9. 考虑以下关于负熵函数$\Phi(x)=\sum_{i=1}^n x_i\ln x_i$的相关问题.

(a) 证明$\Phi(x)$是凸函数.

(b) 写出由$\Phi(x)$定义的Bregman散度$D_\Phi(x,y)$.

(c) 利用条件极值(等式约束优化问题)的Lagrange乘子法，分别求$\Phi(x)$在$n-1$维标准单纯形$\Delta_n=\{x\in\mathbb{R}^n:\sum_{i=1}^n x_i=1,x\ge 0\}$上的极小点和极小值，及已知向量$y\in\mathbb{R}^n_{++}$的Bregaman投影$\Pi^\Phi_{\Delta_n}(y)$.