

## 作业 3

提交日期: 4 月 20 日

**问题 1 (Lipschitz 常数)** • 若  $f(x)$  是二阶连续可微, 证明  $\nabla f(x)$  是  $L$ -Lipschitz 连续等价于  $LI \succeq \nabla^2 f(x) \succeq -LI$

- 估计逻辑回归函数的梯度的 Lipschitz 常数:

$$\min_x \ell(x) \stackrel{\text{def}}{=} \frac{1}{m} \sum_{i=1}^m \ln(1 + \exp(-b_i a_i^\top x)).$$

其中,  $a_i \in \mathbb{R}^n, b_i \in \{-1, +1\}, i = 1, \dots, m$  是给定的数据。

**问题 2** 对于次梯度算法, 请构造一个非光滑函数例子, 说明常数步长不收敛。

**问题 3 (邻近点映射计算)** 计算下面函数的邻近点映射, 即  $\text{prox}_h(x) = \arg \min_y h(y) + \frac{1}{2} \|y - x\|^2$ .

- $h(x) = \|x\|_\infty$  (需要求解一个一维子问题)。
- $h(x) = \max(0, \|x\|_2 - 1)$ .

**问题 4 (条件梯度法应用: 最优实验设计问题)** 考虑  $D$ -最优实验设计 ( $D$ -optimal experimental design), 其目标是最大化估计量的信息内容, 通过差分香农熵测量, 具体到最大化  $\det V(m_1, \dots, m_n)$ , 具体背景参考《convex optimization: 7.5 节》。

该问题需要求解下述约束问题:

$$\min_{x \in \Delta_n} -\ln \det V(x) \quad (3.1)$$

其中  $V(x) = \sum_{i=1}^n x_i a_i a_i^\top, a_i \in \mathbb{R}^d, i = 1, 2, \dots, n$  是给定的数据,  $\Delta_n = \{x | \sum_{i=1}^n x_i = 1, x \geq 0\}$ .

请使用条件梯度法求解该问题, 写出迭代公式, 并且给出子问题的解。

**问题 5 (镜像梯度法子问题)** 求解问题

$$\min f(x) \quad \text{s.t.} \quad x \in \Delta,$$

其中,  $\Delta = \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1, x_i \geq 0\}$ . 使用镜像梯度法, 迭代公式为

$$x^{k+1} = \arg \min_{x \in \Delta} \nabla f(x^k)^\top (x - x^k) + \frac{1}{\alpha_k} \sum_{i=1}^n x_i \log \frac{x_i}{x_i^k}.$$

证明:

$$x_i^{k+1} = \frac{x_i^k \exp(-\alpha_k \nabla f(x^k)_i)}{\sum_{j=1}^n x_j^k \exp(-\alpha_k \nabla f(x^k)_j)}.$$