MATH5015P 最优化算法 **2025 春**

作业 5

提交日期: 6 月 8 日

问题 1 (group lasso) 对于 *group lasso* 问题，请使用交替线性极小化方法求解（见 *lecture 17, 14-16* 页），写出迭代公式. *Group lasso* 问题如下：

$$\min_{\beta} \frac{1}{2}\|y - X\beta\|_2^2 + \lambda \sum_{g=1}^{m} \sqrt{n_g}\|\beta_g\|_2,$$

where:

- $y \in \mathbb{R}^n$ 和 $X \in \mathbb{R}^{n\times p}$ 是给定的数据,
- 变量 $\beta \in \mathbb{R}^p$，$\lambda > 0$ 是给定的参数,
- 下标集 $\{1, 2, \ldots, p\}$ 被分为不相交的 m 组, g 是分组下标，β_g 是对应分组 g 的分量，n_g 是分组 g 的分量维度.

问题 2 给定原问题

$$\min_{w\in\mathbb{R}^d} P(w) = \frac{1}{n}\sum_{i=1}^{n} \phi_i(w^T x_i) + \frac{\lambda}{2}\|w\|^2,$$

其中 ϕ_i 为闭凸函数。

证明：对偶问题表述为*:*

$$\max_{\alpha\in\mathbb{R}^n} D(\alpha) = \frac{1}{n}\sum_{i=1}^{n} -\phi_i^*(-\alpha_i) - \frac{\lambda}{2}\left\|\frac{1}{\lambda n}\sum_{i=1}^{n}\alpha_i x_i\right\|^2$$

其中 $\phi_i^*(u) = \max_z(zu - \phi_i(z))$ 是 ϕ_i 的共轭函数。

问题 3 (随机梯度法) 给定矩阵 $A \in \mathbb{R}^{m\times n}$, 使用随机梯度法求解如下问题，写出迭代具体公式

$$\min_{x\in\mathbb{R}^{n\times r}} \mathrm{Tr}(x^T A^T A x). \tag{5.1}$$

问题 4 在最小二乘问题中出现的平方误差损失函数非常适合使用随机梯度方法进行最小化。我们的问题是：

$$\min_{\theta\in\mathbb{R}^n} \frac{1}{2}\|X\theta - y\|_2^2 = \frac{1}{2}\sum_{i=1}^{m}(x_i^T\theta - y_i)^2,$$

其中 x_i^T 是 $X \in \mathbb{R}^{m\times n}$ 的第 i 行，$y \in \mathbb{R}^m$ 。我们可以将这个目标函数写为：

$$f(\theta) = \sum_{i=1}^{m} f_i(\theta),$$

其中

$$f_i(\theta) := \frac{1}{2}(x_i^T\theta - y_i)^2, \quad \text{对于} i = 1, \dots, m.$$

然后随机梯度方法给出更新规则：

$$\theta_{k+1} = \theta_k - \eta_k \nabla f_{s[k]}(\theta_k),$$

其中 η_k 是步长（也称为学习率），$s[k] \in \{1, \dots, m\}$, 通常是通过从集合 $\{1, \dots, m\}$ 中随机抽取一个数字。

1. 假设 $\{x_i\}_{i=1}^m$ 是一组相互正交的向量。找到一个固定的步长 η，使得随机梯度方法收敛到最小二乘问题的解。
2. 如果没有条件上述条件，即 $\{x_i\}_{i=1}^m$ 并不相互正交，那么随机梯度法的收敛需要什么条件？