CS395T: Numerical Optimization for Graphics and AI: Homework IV

1 Guideline

- Please complete **3** problems out of **5** problems, and please complete at least one problem in the theory session.
- You are welcome to complete more problems.

2 Programming

Each problem in this section counts as two.
Problem 1 and Problem 2. We are interested in solving

$$\min_{\boldsymbol{x}_i, 1 \le i \le m} \sum_{i=1}^{m} \|A_i \boldsymbol{x}_i - \boldsymbol{b}_i\|^2 + \lambda \sum_{1 \le i < j \le m} \|\boldsymbol{x}_i - \boldsymbol{x}_j\|_1 \tag{1}$$

where $\|\boldsymbol{a}\|_1 = \sum_{i=1}^{n} |a_i|$ stands for the L1-norm of a vector. Here A_i and $\boldsymbol{b}_i$ constant matrices. Please apply proximal gradient method to solve (1). An example dataset can be downloaded from [1].

3 Theory

Problem 3. Let $\mathcal{A}$ and $\mathcal{B}$ be two disjoint nonempty convex subsets of $\mathbb{R}^n$. Then there exists a nonzero vector $\boldsymbol{v}$ and a real number c such that

$$\langle \boldsymbol{x}, \boldsymbol{v} \rangle \ge c \quad \text{and} \quad \langle \boldsymbol{y}, \boldsymbol{v} \rangle \le c$$

for all $\boldsymbol{x} \in \mathcal{A}$ and $\boldsymbol{y} \in \mathcal{B}$; i.e., the hyperplane $\langle \cdot, \boldsymbol{v} \rangle = c$, $\boldsymbol{v}$ the normal vector, separates $\mathcal{A}$ and $\mathcal{B}$.

Problem 4. This problem studies the convergence rate of a variant of proximal gradient method. Consider the problem of solving

$$\underset{\boldsymbol{x}}{\text{minimize}}\ f(\boldsymbol{x}) := g(\boldsymbol{x}) + h(\boldsymbol{x})$$

where g is a smooth function whose gradient is Lipschitz continuous, i.e.,

$$\|\nabla g(\boldsymbol{x}) - \nabla g(\boldsymbol{y})\| \le M \|\boldsymbol{x} - \boldsymbol{y}\|, \quad \forall \boldsymbol{x}, \boldsymbol{y}.$$

h is also convex. Given $\boldsymbol{x}^0$ and let $\boldsymbol{z}^1 = \boldsymbol{x}^0$, and iterate

$$\begin{aligned} \boldsymbol{x}^k &= \text{prox}_{\alpha,h}\big(\boldsymbol{z}^k - \alpha \nabla g(\boldsymbol{z}^k)\big) \\ \beta_{k+1} &= \frac{1}{2}(1 + \sqrt{1 + 4\beta_k^2}) \\ \boldsymbol{z}^{k+1} &= \boldsymbol{x}^k + \frac{\beta_k - 1}{\beta_{k+1}}(\boldsymbol{x}^k - \boldsymbol{x}^{k-1}) \end{aligned} \tag{2}$$

[1] `https://www.cs.utexas.edu/~huangqx/hw4_data.mat`

Derive a upper bound on α so that

$$f(\boldsymbol{x}^k) - f(\boldsymbol{x}^\star) \leq \frac{2M}{(k+1)^2}\|\boldsymbol{x}^0 - \boldsymbol{x}^\star\|^2,$$

where $\boldsymbol{x}^\star$ is one optimal solution.

Problem 5. Consider the following constrained optimization problem

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & g_i(\boldsymbol{x}) \leq 0, \quad 1 \leq i \leq m \end{aligned} \tag{3}$$

Suppose there exists a common function $\eta(\boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}^n$, so that

$$\begin{aligned} f(\boldsymbol{x}) - f(\boldsymbol{u}) &\geq \eta(\boldsymbol{x}, \boldsymbol{u})^T \nabla f(\boldsymbol{u}) && \forall \boldsymbol{x}, \boldsymbol{u} \\ -g_i(\boldsymbol{u}) &\geq \eta(\boldsymbol{x}, \boldsymbol{u})^T \nabla g_i(\boldsymbol{u}) && \forall \boldsymbol{x}, \boldsymbol{u}, i = 1, \cdots, m. \end{aligned}$$

Then KKT conditions are become sufficient conditions for optimality.