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

\[ x, \sum_{i=1}^{m} \| A_i x_i - b_i \|^2 + \lambda \sum_{1 \leq i < j \leq m} \| x_i - x_j \|_1 \]  

(1)

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

3 Theory

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

\[ \langle x, v \rangle \geq c \quad \text{and} \quad \langle y, v \rangle \leq c \]

for all \( x \in A \) and \( y \in B \); i.e., the hyperplane \( \langle \cdot, v \rangle = c \), \( v \) the normal vector, separates \( A \) and \( B \).

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

\[ \min_x f(x) := g(x) + h(x) \]

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

\[ \| \nabla g(x) - \nabla g(y) \| \leq M \| x - y \|, \quad \forall x, y. \]

\( h \) is also convex. Given \( x^0 \) and let \( z^1 = x^0 \), and iterate

\[ x^k = \text{prox}_{\alpha,h}(z^k - \alpha \nabla g(z^k)) \]

\[ \beta_{k+1} = \frac{1}{2} \left( 1 + \sqrt{1 + 4\beta_k^2} \right) \]

\[ z^{k+1} = x^k + \frac{\beta_k - 1}{\beta_{k+1}} (x^k - x^{k-1}) \]  

(2)

[1] https://www.cs.utexas.edu/~huangqx/hw4_data.mat
Derive a upper bound on $\alpha$ so that

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

where $x^*$ is one optimal solution.

**Problem 5.** Consider the following constrained optimization problem

$$\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad g_i(x) \leq 0, \quad 1 \leq i \leq m
\end{align*}$$

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

$$\begin{align*}
& f(x) - f(u) \geq \eta(x, u)^T \nabla f(u) \\
& -g_i(u) \geq \eta(x, u)^T \nabla g_i(u)
\end{align*}$$

$\forall x, u, i = 1, \ldots, m.$

Then KKT conditions are become sufficient conditions for optimality.