Homework

December 6, 2019

1 Lecture 8

1. Propose an adaptive to Lipschitz constant of the gradient version of the Dual Gradient Method. Prove the convergence rate theorem.

2. Propose a generalization of the Dual Gradient Method for the composite optimization problem

$$\min_{x \in Q} f(x) + h(x),$$

where f us L-smooth and h is simple convex.

3. Propose an adaptive to Lipschitz constant of the gradient version of the Accelerated Gradient Method. Prove the convergence rate theorem.

4. Consider the LASSO problem

$$\min_x \|Ax - b\|_2^2 + \lambda \|x\|_1$$

with the Euclidean proximal setup. Write the problem

$$\min_{x \in Q} \{\alpha_{k+1}(f(y_{k+1}) + \langle \nabla f(y_{k+1}), x - y_{k+1} \rangle + h(x)) + V(x, u_k)\}$$

in the AGM step in this case and find its solution.

5. Implement Primal Gradient Method, Dual Gradient Method and Accelerated Gradient Method.

1. Apply all three methods to "Bad" functions which give the lower complexity bounds for convex case. (See [Nesterov, 2004], Section 2.1.2).

2. Apply DGM and AGM to the LASSO problem on some real dataset. See. e.g. Boston Housing Dataset `https://towardsdatascience.com/linear-regression-on-boston-housing-dataset-f409b7e4a155`. Given feature vectors $a_i \in \mathbb{R}^n$, $i = 1, ..., m$ (data) the goal is to predict the target b, based on the observable values $b_i \in \mathbb{R}$, $i = 1, ..., m$. This leads to the linear regression problem

$$\min_x \sum_{i=1}^m (a_i^T x - b_i)^2 = \|Ax - b\|_2^2,$$

which is then regularized by $\lambda \|x\|_1$.