15.097 - Homework 1

Colin Pawlowski

March 2, 2016

1 Compressed Sensing

2 Algorithmic Framework for Regression using MIO

3 First Order Method

Here, we derive a first order method following the notes from Lecture 2 - Best Subset Selection. Consider the problem

$$\begin{aligned} \min_{\boldsymbol{\beta}} \quad & g(\boldsymbol{\beta}) := \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \Gamma\|\boldsymbol{\beta}\|_1 \\ \text{s.t.} \quad & \|\boldsymbol{\beta}\|_0 \leq k. \end{aligned} \tag{1}$$

Since $g(\boldsymbol{\beta})$ is convex and $\|\nabla g(\boldsymbol{\beta}) - \nabla g(\boldsymbol{\beta}_0)\| \leq \ell\|\boldsymbol{\beta} - \boldsymbol{\beta}_0\|$, it follows that for all $L \geq \ell$

$$g(\boldsymbol{\beta}) \leq Q(\boldsymbol{\beta}) := g(\boldsymbol{\beta}_0) + \nabla g(\boldsymbol{\beta}_0)^T(\boldsymbol{\beta} - \boldsymbol{\beta}_0) + \frac{L}{2}\|\boldsymbol{\beta} - \boldsymbol{\beta}_0\|_2^2 + \Gamma\|\boldsymbol{\beta}\|_1. \tag{2}$$

To find feasible solutions, we solve the following problem

$$\begin{aligned} \min_{\boldsymbol{\beta}} \quad & Q(\boldsymbol{\beta}) \\ \text{s.t.} \quad & \|\boldsymbol{\beta}\|_0 \leq k. \end{aligned} \tag{3}$$

This is equivalent to

$$\begin{aligned}
\min_{\boldsymbol{\beta}} \quad & \frac{L}{2}\Big\|\boldsymbol{\beta} - \Big(\boldsymbol{\beta}_0 - \frac{1}{L}\nabla g(\boldsymbol{\beta}_0)\Big)\Big\|_2^2 - \frac{1}{2L}\|\nabla g(\boldsymbol{\beta}_0)\|_2^2 + \Gamma\|\boldsymbol{\beta}\|_1 \\
\text{s.t.} \quad & \|\boldsymbol{\beta}\|_0 \leq k,
\end{aligned} \tag{4}$$

which reduces to the following plus a constant term:

$$\begin{aligned}
\min_{\boldsymbol{\beta}} \quad & \frac{L}{2}\|\boldsymbol{\beta} - \mathbf{u}\|_2^2 + \Gamma\|\boldsymbol{\beta}\|_1 \\
\text{s.t.} \quad & \|\boldsymbol{\beta}\|_0 \leq k.
\end{aligned} \tag{5}$$