MATH 697 FALL 2017

Problem Set 2: Due Thursday Week 6

(1) Let $J : \mathbb{R}^p \mapsto \mathbb{R}$ be a convex (not necessarily differentiable) function. In class, we defined the **subdifferential** of J at a point $x \in \mathbb{R}^p$, denoted $\partial J(x)$.

 (a) Show that $\partial J(x)$ is a closed and convex set.

 (b) Let x_k and p_k be sequences such that $p_k \in \partial J(x_k)$ for every k. Show that if $x_k \to x$ and $p_k \to p$, then $p \in \partial J(p)$.

(2) Consider the (constrained) minimization problem

$$|\mathbf{X}\beta - \mathbf{y}|_2^2 \text{ with constraint } |\beta|_2 \leq t,$$

where $\mathbf{X}^t\mathbf{X}$ is assumed to be invertible. Show that if $|\hat{\beta}|_2 > t$, with $\hat{\beta}$ being the least squares solution, then the constrained solution, denoted $\hat{\beta}_0$, must be such that

$$|\hat{\beta}_0|_2 = t \text{ and } \mathbf{X}^t(\mathbf{X}\hat{\beta}_0 - \mathbf{y}) \perp \hat{\beta}_0.$$

(3) This exercise is intended to generate more interesting simulated data using a (non-linear) additive model. Fix coefficients $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)$ of your choice, and then generate a set of $n = 100$ data points (x_i, y_i) of the form

$$y_i = f_\beta(x_i) + \varepsilon_i$$

the x_i are always given by i.i.d. standard normals, the ε_i are i.i.d. normals with mean zero and variance 0.5, and $f_\beta(x) = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$.

Then, save the generated data of pairs (x_i, y_i) in a file (in csv format, for instance) and produce a plot of this data (when you submit the plot, indicate the β you used).

(4) To the data generated in the previous problem, apply the standard least squares method and indicate the $\hat{\beta}$ you obtained. Produce a plot containing the $\hat{y}$ and training data.

(5) Do the same as before but instead of least squares, applying ridge regression with $\lambda = 10, 20$, and 50. Produce a plot as in the previous case for each of the values of λ.