1.

EECS 545 Homework 3 Yuan Yin

2) Robust Regression

(a) According to Majorize-Minimize Algorithm, we find a majorizing function $\bar{J}(\vec{\theta}, \vec{\theta}_t)$ and then iterate $\vec{\theta}_t$ with solution for $\vec{\theta}_{t+1} = \arg\min \bar{J}(\vec{\theta}, \vec{\theta}_t)$

According to the lemma in notes, we have such $\bar{J}$ s.t.

$$\bar{J}(\vec{\theta}, \vec{\theta}_t) = \sum_{i=1}^{n} \bar{\rho}(y_i - \vec{w}^T\vec{x}_i - b)$$

where $\bar{\rho}(r) = \rho(r_{t,i}) - \frac{1}{2} r_{t,i}\psi(r_{t,i}) + \frac{1}{2}\frac{\psi(r_{t,i})}{r_{t,i}} r^2$

$$\psi(r) = \rho'(r), \quad r_{t,i} = y_i - \vec{w}_t^T\vec{x}_i - b_t$$

It's nothing but

$$\vec{\theta}_{t+1} = \arg\min \left(\sum_{i=1}^{n} \left(\rho(r_{t,i}) - \frac{1}{2} r_{t,i}\psi(r_{t,i}) + \frac{1}{2}\frac{\psi(r_{t,i})}{r_{t,i}} r^2 \right)\right)$$

since for time t, $r_{t,i}$ is known (can be computed)

Then $\vec{\theta}_{t+1} = \arg\min \sum_{i=1}^{n} \frac{1}{2}\frac{\psi(r_{t,i})}{r_{t,i}} r^2$

Suppose $c_{t,i} = \frac{1}{2}\frac{\psi(r_{t,i})}{r_{t,i}}$, then

$$\vec{\theta}_{t+1} = \arg\min_{\vec{\theta}} \sum_{i=1}^{n} c_{t,i}(y_i - \vec{\theta}^T\vec{x}_i)^2$$

Since that $c_{t,i}$ is changing with iteration time t changes, we can think MM algorithm as "iteratively reweighted least squares"

Explain of robustness: From the lemma we know that $c_{t,i} = \frac{1}{2}\frac{\psi(r_{t,i})}{r_{t,i}}$ is nonincreasing for $r_{t,i} > 0$. That is to say, for outliers who have larger error $r_{t,i}$, the weight $c_{t,i}$ will be smaller. (Noting that since $\rho(r)$ is symmetric