Assignment - 3 (RMJ916)

(1) When l_1 regularization is used & Hessian matrix is diagonal with positive entries so, objective function

$$\hat{L}_R(\theta) := \sum_{i=1}^{d} [\tfrac{1}{2} H_{ii} (\theta_i - \theta_i^*)^2 + \alpha |\theta_i|]$$

note: $\alpha / H_{ii} > 0$

→ gradient of regularized objective

$$\nabla \hat{L}_R (\theta^*_{R_i}) = H_{ii} ((\theta^*_{R_i}) - \theta_i^*) + \alpha \, sign(\theta^*_R)_i$$

$$\nabla \hat{L}_R(\theta^*_R)_i = 0$$

so,

$$H_{ii} ((\theta^*_{R})_i - \theta_i^*) + \alpha \, sign(\theta^*_R)_i = 0$$

$$\therefore H_{ii} (\theta^*_R)_i - H_{ii} \theta^*_i + \alpha \, sign(\theta^*_R)_i = 0$$

$$\therefore H_{ii} \theta^*_i = H_{ii}(\theta^*_R)_i + \alpha \, sign(\theta^*_R)_i$$

now divide by H_{ii} (Hessian matrix)

$$\therefore \cancel{H_{ii}} \left(\theta_i^* - \frac{\alpha \, sign(\theta^*_R)_i}{H_{ii}}\right) = \cancel{H_{ii}} (\theta^*_R)_i$$

so, after applying Relu activation function

→ $(\theta^*_R)_i = \max \{\theta_i^* - \frac{\alpha}{H_{ii}}, 0\}$ where $\theta^*_{R_i} \geq 0$

& here $sign(\theta_{R_i})$ is positive; ——— (1)

$(\theta^*_R)_i = \min \{\theta_i^* + \frac{\alpha}{H_{ii}}, 0\}$ where $\theta^*_{R_i} < 0$

here $sign(\theta_{R_i})$ is negative; ——— (2)

So,

$$\therefore \theta^*_{R_i} = \begin{cases} \max(\theta_i^* - \frac{\alpha}{H_{ii}}, 0), & \theta_i^* \geq 0 \\ \min(\theta_i + \frac{\alpha}{H_{ii}}, 0), & \theta_i^* < 0 \end{cases}$$