Homework 4

For a given set of training sample $(y_i, \mathbf{x}_i) \in \mathbb{R} \times \mathbb{R}^p$, the (linear) kernel quantile regression (KQR) to estimate the τth conditional quantile of y given $\mathbf{x}$, $f(\mathbf{x}) = \beta_0 + \boldsymbol{\beta}^T\mathbf{x}$ solves

$$(\hat{\beta}_0, \hat{\boldsymbol{\beta}}) = \underset{\beta_0, \boldsymbol{\beta}}{\operatorname{argmin}} \sum_{i=1}^n \rho_\tau\{y_i - f(\mathbf{x}_i)\} + \frac{\lambda}{2}\boldsymbol{\beta}^T\boldsymbol{\beta} \tag{1}$$

where $\rho_\tau(u) = u\tau - u\mathbb{1}\{u < 0\}$.

1. Justify that (1) is equivalent to solve

$$\min_{\beta_0, \boldsymbol{\beta}, \boldsymbol{\xi}, \boldsymbol{\zeta}} \quad \tau\sum_{i=1}^n \xi_i + (1-\tau)\sum_{i=1}^n \zeta_i + \frac{\lambda}{2}\boldsymbol{\beta}^T\boldsymbol{\beta} \tag{2}$$

   subject to

$$-\zeta_i \le y_i - f(\mathbf{x}_i) \le \xi_i \quad \text{and} \quad \xi_i, \zeta_i \ge 0.$$

2. Introducing Lagrange multipliers $\alpha_i, \gamma_i, \kappa_i, \rho_i \ge 0, i = 1, \cdots, n$, the primal function of (2) is given by

$$\begin{aligned} L_p : \tau\sum_{i=1}^n \xi_i + (1-\tau)\sum_{i=1}^n \zeta_i + \frac{\lambda}{2}\boldsymbol{\beta}^T\boldsymbol{\beta} + \sum_{i=1}^n \alpha_i\{y_i - f(\mathbf{x}_i) - \xi_i\} \\ - \sum_{i=1}^n \gamma_i\{y_i - f(\mathbf{x}_i) + \zeta_i\} - \sum_{i=1}^n \kappa_i\xi_i - \sum_{i=1}^n \rho_i\zeta_i. \end{aligned}$$

   (a) Derive KKT (stationary / complementary slackness) conditions.

   (b) Show that corresponding dual problem is given by

$$\begin{aligned} \max_{\theta_1, \cdots, \theta_n} \quad & \sum_{i=1}^n \theta_i y_i - \frac{1}{2\lambda}\sum_{i=1}^n\sum_{j=1}^n \theta_i\theta_j\mathbf{x}_i^T\mathbf{x}_j \\ \text{subject to} \quad & -(1-\tau) \le \theta_i \le \tau, \quad i = 1, \cdots, n \\ & \sum_{i=1}^n \theta_i = 0. \end{aligned} \tag{3}$$

   where $\theta_i = \alpha_i - \gamma_i$.

3. Write your own `R` code to solve linear KQR that returns β_0 and $\boldsymbol{\beta}$ using either `ipop{}` function in `kernlab` package or `solve.QP()` in `quadprog`. You can get $\boldsymbol{\beta}$ by solving QP problem in (3) and β_0 from the fact that $y_i - (\beta_0 + \boldsymbol{\beta}^T\mathbf{x}_i) = 0$ for all i such that $-(1-\tau) < \alpha_i < \tau$.