2.2 Composite quantile regression

$$\underset{b_{\theta_k},\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n}\sum_{k=1}^{K}\omega_k\rho_{\theta_k}\{y_i - b_{\theta_k} - \mathbf{x}_i^T\boldsymbol{\beta}\},$$

$$\Updownarrow$$

$$L(y|\alpha,\boldsymbol{\beta}) = \prod_{i=1}^{n}\sum_{k=1}^{K}\omega_k\theta_k(1-\theta_k)exp\{-\rho_{\theta_k}(y_i - b_{\theta_k} - \mathbf{x}_i^T\boldsymbol{\beta})\} \tag{4}$$

To solve (4), Huang and Chen [2015] introduce a cluster matrix $\mathbf{C}$, whose i,k th element C_{ik} is equal to 1 if the ith subject belongs to the kth cluster, otherwise $C_{ik} = 0$. It is treated as a missing value, and they consider following complete likelihood

$$\begin{aligned}
&\prod_{i=1}^{n}\prod_{k=1}^{K}\Big[\omega_k\theta_k(1-\theta_k)exp\{-\rho_{\theta_k}(y_i - b_{\theta_k} - \mathbf{x}_i^T\boldsymbol{\beta})\}\Big]^{C_{ik}} \qquad (5)\\
&= \prod_{i=1}^{n}\prod_{k=1}^{K}\Big[\theta_k(1-\theta_k)exp\{-\rho_{\theta_k}(y_i - b_{\theta_k} - \mathbf{x}_i^T\boldsymbol{\beta})\}\Big]^{C_{ik}} \times \prod_{i=1}^{n}\prod_{k=1}^{K}\omega_k^{C_{ik}}\\
&= P(y|C) \times P(C),
\end{aligned}$$

where $P(C)$ is marginal distribution which is multinomial distribution.

Then, the posterior distribution of $\boldsymbol{\beta}$ is

$$\begin{aligned}
f(\boldsymbol{\beta}|\mathbf{y},\mathbf{X},\tau,\lambda,\mathbf{b},\boldsymbol{\omega},\mathbf{C}) &\propto \text{ Complete likelihood} \times \text{ Priors}\\
&\propto \prod_{i=1}^{n}\prod_{k=1}^{K}\Big[\omega_k\theta_k(1-\theta_k)exp\{-\rho_{\theta_k}(y_i - b_{\theta_k} - \mathbf{x}_i^T\boldsymbol{\beta})\}\Big]^{C_{ik}} \times \pi(\boldsymbol{\beta}|\tau,\lambda)\pi(\boldsymbol{\omega}) \qquad (6)\\
&\propto exp\{-\sum_{i=1}^{n}\sum_{k=1}^{K}\omega_k\rho_{\theta_k}\{y_i - b_{\theta_k} - \mathbf{x}_i^T\boldsymbol{\beta}\} - \tau\lambda\sum_{j=1}^{p}|\beta_j|\} \text{ : (Not sure, but no need to derive)}
\end{aligned}$$

: We are draw samples from the posterior distribution using MCMC.

First, consider hierarchical model as follows:

$$
\begin{aligned}
y_i =& b_{\theta_k} + \mathbf{x}_i^T \boldsymbol{\beta} + \xi_{1k}\tilde{\nu}_i + \tau^{-1/2}\xi_{2k}\sqrt{\tilde{\nu}_i} z_i, \quad \text{if } y_i \in \text{ cluster } k. \\
\tilde{\boldsymbol{\nu}}|\tau \sim& \prod_{i=1}^{n} \tau exp(-\tau\tilde{\nu}_i) \\
\mathbf{z} \sim& \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} exp\Big(-\frac{1}{2} z_i^2 \Big) \\
\boldsymbol{\beta}, \mathbf{s}|\eta^2 \sim& \prod_{j=1}^{p} \frac{1}{\sqrt{2\pi s_j}} exp\Big(-\frac{\beta_j^2}{2s_j} \Big) \prod_{j=1}^{p} \frac{\eta^2}{2} exp\Big(-\frac{\eta^2}{2} s_j \Big) \\
(\tau, \eta^2) \sim& Gamma(b, d) \\
\omega \sim& \text{Dirichlet}(\alpha_1, \cdots, \alpha_K)
\end{aligned}
$$

Then, the complete likelihood based on ALD form is

$$
f(\mathbf{y}|\mathbf{X}, \tau, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{s}, \eta^2, \mathbf{b}, \boldsymbol{\omega}, \mathbf{C}) = \prod_{i=1}^{n}\prod_{k=1}^{K} \Big(\frac{1}{\sqrt{2\pi\tau^{-1}\xi_{2k}^2\tilde{\nu}_i}} \Big)^{C_{ik}} exp\Big\{ -\frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{K} \frac{C_{ik}(y_i - b_{\theta_k} - \mathbf{x}_i^T\boldsymbol{\beta} - \xi_{1k}\tilde{\nu}_i)^2}{\tau^{-1}\xi_{2k}^2\tilde{\nu}_i} \Big\}, \tag{7}
$$

and compute conditional distributions of the parameters as follows.

- $f(\tilde{\nu}_i|\mathbf{X}, \mathbf{y}, \tilde{\boldsymbol{\nu}}_{-i}, \boldsymbol{\beta}, \mathbf{b}, \mathbf{s}, \tau, \eta^2, \boldsymbol{\omega}, \mathbf{C}) \propto f(\mathbf{y}|\mathbf{X}, \tilde{\boldsymbol{\nu}}, \boldsymbol{\beta}, \mathbf{b}, \mathbf{s}, \tau, \eta^2, \boldsymbol{\omega}, \mathbf{C}) \times \pi(\tilde{\nu}_i|\tau)$: Inverse Gaussian distribution $\left(\lambda = \Big(\sum_{k=1}^{K} \frac{C_{ik}\xi_{1k}^2}{\xi_{2k}^2} + 2 \Big)\tau, \mu = \sqrt{\sum_{k=1}^{K} \frac{C_{ik}(\xi_{1k}^2 + 2\xi_{2k}^2)}{(y_i - b_{\theta_k} - \mathbf{x}_i^T\boldsymbol{\beta})^2}} \right)$
- $f(s_j|\mathbf{X}, \mathbf{y}, \tilde{\boldsymbol{\nu}}, \boldsymbol{\beta}, \mathbf{b}, \mathbf{s}_{-j}, \tau, \eta^2, \boldsymbol{\omega}, \mathbf{C}) \propto \pi(\beta_j|s_j) \times \pi(s_j|\eta^2)$: Inverse Gaussian distribution$(\lambda = \eta^2, \mu = \sqrt{\eta^2/\beta_j^2})$
- $f(\beta_j|\mathbf{X}, \mathbf{y}, \tilde{\boldsymbol{\nu}}, \boldsymbol{\beta}_{-j}, \mathbf{b}, \mathbf{s}, \tau, \eta^2, \boldsymbol{\omega}, \mathbf{C}) \propto f(\mathbf{y}|\mathbf{X}, \tilde{\boldsymbol{\nu}}, \boldsymbol{\beta}, \mathbf{b}, \mathbf{s}, \tau, \eta^2, \boldsymbol{\omega}, \mathbf{C}) \times \pi(\beta_j|s_j)$: $N\left(\frac{\sum_{i=1}^{n} x_{ij}\tilde{y}_i/\tilde{\sigma}_i^2}{\sum_{i=1}^{n} x_{ij}^2/\tilde{\sigma}_i^2 + 1/s_j}, \frac{1}{\sum_{i=1}^{n} x_{ij}^2/\tilde{\sigma}_i^2 + 1/s_j} \right)$, where
- $f(\tau|\mathbf{X}, \mathbf{y}, \tilde{\boldsymbol{\nu}}, \boldsymbol{\beta}, \mathbf{b}, \mathbf{s}, \eta^2, \boldsymbol{\omega}, \mathbf{C}) \propto f(\mathbf{y}|\mathbf{X}, \tilde{\boldsymbol{\nu}}, \boldsymbol{\beta}, \mathbf{b}, \mathbf{s}, \tau, \eta^2, \boldsymbol{\omega}, \mathbf{C}) \times \pi(\tilde{\nu}|\tau) \times \pi(\tau)$: Gamma distribution
- $f(\eta^2|\mathbf{X}, \mathbf{y}, \tilde{\boldsymbol{\nu}}, \boldsymbol{\beta}, \mathbf{b}, \mathbf{s}, \tau, \boldsymbol{\omega}, \mathbf{C}) \propto \pi(\mathbf{s}|\eta^2) \times \pi(\eta^2)$: Gamma distribution
- $f(\boldsymbol{\omega}|\mathbf{X}, \mathbf{y}, \tilde{\boldsymbol{\nu}}, \boldsymbol{\beta}, \mathbf{b}, \mathbf{s}, \tau, \eta^2, \mathbf{C})$: Dirichlet distribution
- $f(\mathbf{C}_i|\mathbf{X}, \mathbf{y}, \tilde{\boldsymbol{\nu}}, \boldsymbol{\beta}, \mathbf{b}, \mathbf{s}, \tau, \eta^2, \boldsymbol{\omega}, \mathbf{C}_{-i})$: Multinomial distribution
- $f(b_{\theta_k}|\mathbf{X}, \mathbf{y}, \tilde{\boldsymbol{\nu}}, \boldsymbol{\beta}, \mathbf{s}, \mathbf{b}_{-\theta_k}, \tau, \eta^2, \boldsymbol{\omega}, \mathbf{C})$: Gaussian distribution

They are all known distributions [Huang and Chen, 2015].