Model Comoparison

1 LR model

$$\sum_{k=1}^{q}\sum_{\ell=1}^{b}\sum_{i=1}^{n}\frac{1}{nb}\rho_{\tau_\ell}(y_i^{(k)}-Z_i^{(k)})+\sum_j P_\lambda(\sigma_j(Z))$$

$$\begin{aligned}L_\delta(Z,E,\Gamma)=&\sum_{k=1}^{q}\sum_{\ell=1}^{b_n}\sum_{i=1}^{n}\frac{1}{nb_n}\rho_{\tau_\ell}(e_{ik}^{(\ell)})+\sum_j P_\lambda(\sigma_j(Z))\\&+\sum_{\ell=1}^{b}\langle\Gamma^{(\ell)},Y-Z-E^{(\ell)}\rangle+\frac{\delta}{2}\sum_{\ell=1}^{b}\|Y-Z-E^{(\ell)}\|_F^2\end{aligned}$$

1.1 update Z

$$\begin{aligned}&\frac{\delta}{2}\sum_{\ell=1}^{b_n}\|Y-Z-E^{(\ell)}\|_F^2-\sum_{\ell=1}^{b_n}\mathrm{tr}([\Gamma^{(\ell)}]^TZ)\\
&\quad=\sum_\ell\Big\{\mathrm{tr}\Big(\frac{\delta}{2}(Y-Z-E^{(\ell)})^T(Y-Z-E^{(\ell)})-[\Gamma^{(\ell)}]^TZ\Big)\Big\}\\
&\quad=\frac{\delta}{2}\sum_\ell\Big\{\mathrm{tr}\Big((Y-Z-E^{(\ell)})^T(Y-Z-E^{(\ell)})-\frac{2}{\delta}[\Gamma^{(\ell)}]^TZ\Big)\Big\}\\
&\quad\propto\frac{\delta}{2}\sum_\ell\Big\{\mathrm{tr}\Big(Z^TZ-2(Y-E^{(\ell)}+\frac{1}{\delta}\Gamma^{(\ell)})^TZ\Big)\Big\}\\
&\quad=\frac{\delta b}{2}\Big\{\mathrm{tr}\Big(Z^TZ-2[\frac{1}{b}\sum_\ell(Y-E^{(\ell)}+\frac{1}{\delta}\Gamma^{(\ell)})]^TZ\Big)\Big\}\\
&\quad\propto\frac{\delta b}{2}\Big\{\mathrm{tr}\Big([Z-\frac{1}{b}\sum_\ell(Y-E^{(\ell)}+\frac{1}{\delta}\Gamma^{(\ell)})]^T[Z-\frac{1}{b}\sum_\ell(Y-E^{(\ell)}+\frac{1}{\delta}\Gamma^{(\ell)})]\Big)\Big\}\\
&\quad=\frac{\delta b}{2}\|\frac{1}{b}\sum_\ell(Y-E^{(\ell)}+\frac{1}{\delta}\Gamma^{(\ell)})-Z\|_F^2\end{aligned}$$

Z is updated by optimizing

$$\min_Z\frac{1}{2}\|\frac{1}{b}\sum_\ell(Y-E^{(\ell)}+\frac{1}{\delta}\Gamma^{(\ell)})-Z\|_F^2+\frac{1}{\delta b}\sum_j P_\lambda(\sigma_j(Z))\tag{1}$$

Using taylor approximation, equation 1 can be rewritten as follows:

$$\min_Z \frac{1}{2}\|\frac{1}{b}\sum_\ell (Y - E^{(\ell)} + \frac{1}{\delta}\Gamma^{(\ell)}) - Z\|_F^2 + \frac{1}{\delta b}\sum_j \nu_j \sigma_j(Z),$$

where $\nu_j = P'_\lambda\{\sigma_j(\tilde{Z})\}$.

$$\therefore Z^{k+1} = PD_{\frac{1}{\delta b},\nu}(\Sigma)Q^T$$

where $P\Sigma Q^T$ is the SVD of $\frac{1}{b}\sum_\ell(Y - E^{(\ell)} + \frac{1}{\delta}\Gamma^{(\ell)})$ and $D_{\frac{1}{\delta b},\nu}(\Sigma) = \mathrm{diag}(\{\sigma_i - \frac{\nu_i}{\delta b}\}_+)$.

1.2 update $E^{(\ell)}$

$e_{ik}^{(\ell)}$ is updated by optimizing

$$\min_{e_{ik}^{(\ell)}} \frac{1}{nb}\rho_{\tau_\ell}(e_{ik}^{(\ell)}) - [\gamma_{ik}^{(\ell)}]^t e_{ik}^{(\ell)} + \frac{\delta}{2}(y_i^{(k)} - [Z_i^{(k)}]^{t+1} - e_{ik}^{(\ell)})^2.$$

By KKT condition, $e_{ik}^{(\ell)}$ satisfies the following equation for each $i \in \{1, \ldots, n\}$, $k \in \{1, \ldots, q\}$, and $\ell \in \{1, \ldots, b\}$:

$$\frac{1}{nb}\partial\rho_{\tau_\ell}(e_{ik}^{(\ell)}) = [\gamma_{ik}^{(\ell)}]^t + \delta(y_i^{(k)} - [Z_i^{(k)}]^{t+1} - e_{ik}^{(\ell)}),$$

where $\partial\rho_{\tau_\ell}(e_{ik}^{(\ell)})$ is subgradient of $\rho_{\tau_\ell}(e_{ik}^{(\ell)})$:

$$\partial\rho_{\tau_\ell}(e_{ik}^{(\ell)}) = \begin{cases} \tau_\ell - 1 & \text{if } e_{ik}^{(\ell)} < 0, \\ \{c \in \mathbb{R} : \tau_\ell - 1 \le c \le \tau_\ell\} & \text{if } e_{ik}^{(\ell)} = 0, \\ \tau_\ell & \text{if } e_{ik}^{(\ell)} > 0. \end{cases}$$

Let $\epsilon_{ik}^{(\ell)} = y_i^{(k)} - [Z_i^{(k)}]^{t+1}$. Then, $e_{ik}^{(\ell)}$ is updated as follows.

$$\therefore e_{ik}^{(\ell)} = \begin{cases} \epsilon_{ik}^{(\ell)} + [\gamma_{ik}^{(\ell)}]^t/\delta - (\tau_\ell - 1)/(nb\delta) & \text{if } \epsilon_{ik}^{(\ell)} + [\gamma_{ik}^{(\ell)}]^t/\delta < (\tau_\ell - 1)/(nb\delta), \\ 0 & \text{if } (\tau_\ell - 1)/(nb\delta) \le \epsilon_{ik}^{(\ell)} + [\gamma_{ik}^{(g)}]^t/\delta \le \tau_\ell/(nb\delta), \\ \epsilon_{ik}^{(\ell)} + [\gamma_{ik}^{(\ell)}]^t/\delta - \tau_\ell/(nb\delta) & \text{if } \epsilon_{ik}^{(\ell)} + [\gamma_{ik}^{(\ell)}]^t/\delta > \tau_\ell/(nb\delta). \end{cases}$$

1.3 update multiplier : $\gamma_{ik}^{(\ell)}$

$$[\gamma_{ik}^{(\ell)}]^{t+1} = [\gamma_{ik}^{(\ell)}]^t + \delta(y_i^{(k)} - [Z_i^{(k)}]^{t+1} - [e_{ik}^{(\ell)}]^{t+1})$$

2 SP model

$$\sum_{k=1}^{q}\sum_{\ell=1}^{b}\sum_{i=1}^{n}\frac{1}{nb}\rho_{\tau_\ell}(y_i^{(k)}-v_i^{(\ell)*}\Theta^{(k)*})+\sum_j\sum_k P_\lambda(\|\theta_j^{(k)}\|_2)$$

$$\begin{aligned} & L_\delta(H_0,H,\Theta_0,\Theta,E,\Gamma,A)\\ & \quad =\sum_{k=1}^{q}\sum_{\ell=1}^{b_n}\sum_{i=1}^{n}\frac{1}{nb_n}\rho_{\tau_\ell}(e_{ik}^{(\ell)})+\sum_j\sum_k P_{\lambda_1}(\|\theta_j^{(k)}\|_2)\\ & \quad +\sum_{\ell=1}^{b_n}\langle\Gamma^{(\ell)},Y-V^{(\ell)*}H^*-E^{(\ell)}\rangle+\frac{\delta}{2}\sum_{\ell=1}^{b_n}\|Y-V^{(\ell)*}H^*-E^{(\ell)}\|_F^2\\ & \quad +\langle A,\Theta^*-H^*\rangle+\frac{\delta}{2}\|\Theta^*-H^*\|_F^2, \end{aligned} \tag{2}$$

2.1 update η

Updating H_0 and H is equivalent to solving

$$\begin{aligned} \min_{H_0,H} & -\sum_{\ell=1}^{b_n}\langle[\Gamma^{(\ell)}]^t,V^{(\ell)*}H^*\rangle+\frac{\delta}{2}\sum_{\ell=1}^{b_n}\|Y-V^{(\ell)*}H^*-[E^{(\ell)}]^t\|_F^2\\ & -\langle A^t,H^*\rangle+\frac{\delta}{2}\|[\Theta^*]^t-H^*\|_F^2. \end{aligned}$$

By KKT condition, $[H^*]^{t+1}$ satisfies

$$\delta(\sum_{\ell=1}^{b}[V^{(\ell)*}]^TV^{(\ell)*}+I_{(p+1)K})[H^*]^{t+1}=\sum_{\ell=1}^{b}[V^{(\ell)*}]^T[\Gamma^{(\ell)}]^t+\delta\sum_{\ell=1}^{b}[V^{(\ell)*}]^T(Y-[E^{(\ell)}]^t)+A+\delta[\Theta^*]^t,$$

$$\therefore [H^*]^{t+1}=\frac{1}{\delta}\Big(\sum_{\ell=1}^{b}[V^{(\ell)*}]^TV^{(\ell)*}+I_{(p+1)K}\Big)^{-1}\Big(\sum_{\ell=1}^{b}[V^{(\ell)*}]^T[\Gamma^{(\ell)}]^t+\delta\sum_{\ell=1}^{b}[V^{(\ell)*}]^T(Y-[E^{(\ell)}]^t)+A+\delta[\Theta^*]^t\Big)$$

2.2 update θ

Updating Θ_0 and Θ is equivalent to solving

$$\min_{\theta_j^{(k)}}\langle[\alpha_j^{(k)}]^t,\theta_j^{(k)}\rangle+\frac{\delta}{2}\|\theta_j^{(k)}-[\eta_j^{(k)}]^{t+1}\|_2^2+P_\lambda(\|\theta_j^{(k)}\|_2), \tag{3}$$

Using taylor approximation, equation 3 can be rewritten as follows:

$$\min_{\theta_j^{(k)}}\langle[\alpha_j^{(k)}]^t,\theta_j^{(k)}\rangle+\frac{\delta}{2}\|\theta_j^{(k)}-[\eta_j^{(k)}]^{t+1}\|_2^2+w_j^{(k)}\|\theta_j^{(k)}\|_2.$$

where $w_j^{(k)}=P'_\lambda(\|\tilde{\theta}_j^{(k)}\|_2)$ for $j\neq 0$. First, if $\theta_j^{(k)}\neq 0$, $\theta_j^{(k)}=r_j^{(k)}(1+\frac{w_j^{(k)}}{(\delta\cdot\|\theta_j^{(k)}\|_2)})^{-1}$ where $r_j^{(k)}=[\eta_j^{(k)}]^{t+1}-[\alpha_j^{(k)}]^t/\delta$. In this case, we can easily check that $\|r_j^{(k)}\|_2>w_j^{(k)}/\delta$. Second, if $\theta_j^{(k)}=0$, $w_j^{(k)}c=\delta r_j^{(k)}$

and we can easily check that $\|r_j^{(k)}\|_2 \leq w_j^{(k)}/\delta$. To sum up, $\theta_j^{(k)}$ which is a subvector of kth column of Θ with a K_n length is updated as follows:

$$\therefore \theta_j^{(k)} = \begin{cases} r_j^{(k)}(1 - \frac{w_j^{(k)}}{(\delta \cdot \|r_j^{(k)}\|_2)}) & \text{if } \|r_j^{(k)}\|_2 > w_j^{(k)}/\delta, \\ 0 & \text{if } \|r_j^{(k)}\|_2 \leq w_j^{(k)}/\delta. \end{cases}$$

2.3 update e

$e_{ik}^{(\ell)}$ is updated by optimizing

$$\min_{e_{ik}^{(\ell)}} \frac{1}{nb}\rho_{\tau_\ell}(e_{ik}^{(\ell)}) - [\gamma_{ik}^{(\ell)}]^t e_{ik}^{(\ell)} + \frac{\delta}{2}(y_i^{(k)} - V_i^{(\ell)*}[H^{*(k)}]^{t+1} - e_{ik}^{(\ell)})^2.$$

By KKT condition, $e_{ik}^{(\ell)}$ satisfies the following equation for each $i \in \{1,\ldots,n\}$, $k \in \{1,\ldots,q\}$, and $\ell \in \{1,\ldots,b\}$:

$$\frac{1}{nb}\partial\rho_{\tau_\ell}(e_{ik}^{(\ell)}) = [\gamma_{ik}^{(\ell)}]^t + \delta(y_i^{(k)} - V_i^{(\ell)*}[H^{*(k)}]^{t+1} - e_{ik}^{(\ell)}),$$

where $\partial\rho_{\tau_\ell}(e_{ik}^{(\ell)})$ is subgradient of $\rho_{\tau_\ell}(e_{ik}^{(\ell)})$. Let $\epsilon_{ik}^{(\ell)} = y_i^{(k)} - V_i^{(\ell)*}[H^{*(k)}]^{t+1}$. Then, $e_{ik}^{(\ell)}$ is updated as follows.

$$\therefore e_{ik}^{(\ell)} = \begin{cases} \epsilon_{ik}^{(\ell)} + [\gamma_{ik}^{(\ell)}]^t/\delta - (\tau_\ell - 1)/(nb\delta) & \text{if } \epsilon_{ik}^{(\ell)} + [\gamma_{ik}^{(\ell)}]^t/\delta < (\tau_\ell - 1)/(nb\delta), \\ 0 & \text{if } (\tau_\ell - 1)/(nb\delta) \leq \epsilon_{ik}^{(\ell)} + [\gamma_{ik}^{(g)}]^t/\delta \leq \tau_\ell/(nb\delta), \\ \epsilon_{ik}^{(\ell)} + [\gamma_{ik}^{(\ell)}]^t/\delta - \tau_\ell/(nb\delta) & \text{if } \epsilon_{ik}^{(\ell)} + [\gamma_{ik}^{(\ell)}]^t/\delta > \tau_\ell/(nb\delta). \end{cases}$$

2.4 update multipliers Γ and $\mathbf{A}$

$$[\Gamma^{(\ell)}]^{t+1} = [\Gamma^{(\ell)}]^t + \delta(Y - V_0^{(\ell)}H_0^{t+1} - V^{(\ell)}H^{t+1} - [E^{(\ell)}]^{t+1}),$$

$$A^{t+1} = A^t + \delta(\Theta^{t+1} - H^{t+1}).$$