Closed form

$$
\begin{aligned}
&L_\delta(H_0, H, \Theta_0, \Theta, Z, E, \Gamma, A) \\
&\quad = \sum_{k=1}^{q} \sum_{\ell=1}^{b_n} \sum_{i=1}^{n} \frac{1}{n b_n} \rho_{\tau_\ell}(e_{ik}^{(\ell)}) + P_{\lambda_1}(\|\theta_j^{(k)}\|_2) + \sum_j P_{\lambda_2}(\sigma_j(Z)) \\
&\quad + \sum_{\ell=1}^{b_n} \langle \Gamma^{(\ell)}, Y - Z - V^{(\ell)*} H^* - E^{(\ell)} \rangle \\
&\quad + \frac{\delta}{2} \sum_{\ell=1}^{b_n} \|Y - Z - V^{(\ell)*} H^* - E^{(\ell)}\|_F^2 + \langle A, \Theta^* - H^* \rangle + \frac{\delta}{2} \|\Theta^* - H^*\|_F^2,
\end{aligned}
\tag{1}
$$

where $P_\lambda(\cdot)$'s are SCAD functions with regularization parameters $\lambda_1, \lambda_2, a > 0:$

$$
P_\lambda(x) = \begin{cases} \lambda|x| & \text{if } |x| \le \lambda, \\ (2a\lambda|x| - x^2 - \lambda^2)/\{2(a-1)\} & \text{if } \lambda < |x| \le a\lambda, \\ \lambda^2(a+1)/2 & \text{otherwise.} \end{cases}
$$

Let (1) be Q.

1 Updating H_0 and H

Updating H_0 and H is equivalent to solving

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

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

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

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

2 Updating Θ_0 and Θ

Updating Θ_0 and Θ is equivalent to solving

$$\min_{\theta_j^{(k)}} \langle [\alpha_j^{(k)}]^t, \theta_j^{(k)} \rangle + \frac{\delta}{2} \| \theta_j^{(k)} - [\eta_j^{(k)}]^{t+1} \|_2^2 + P_{\lambda_1}(\|\theta_j^{(k)}\|_2), \tag{2}$$

Using taylor approximation, 2 can be rewritten as follows:

$$\min_{\theta_j^{(k)}} \langle [\alpha_j^{(k)}]^t, \theta_j^{(k)} \rangle + \frac{\delta}{2} \| \theta_j^{(k)} - [\eta_j^{(k)}]^{t+1} \|_2^2 + \lambda_1 w_j^{(k)} \|\theta_j^{(k)}\|_2.$$

By KKT condition, Θ^{t+1} satisfies the following equation for each $j \in \{0, \ldots, p_n\}$ and $k \in \{1, \ldots, q\}$:

$$[\alpha_j^{(k)}]^t + \delta(\theta_j^{(k)} - [\eta_j^{(k)}]^{t+1}) + \lambda_1 w_j^{(k)} \partial \|\theta_j^{(k)}\|_2 = 0, \tag{3}$$

where weight $w_j^{(k)} = 0$ for $j = 0$, $w_j^{(k)} = P'_{\lambda_1}(\|\tilde{\theta}_j^{(k)}\|_2)$ for $j \neq 0$, and $P'_\lambda(|\cdot|)$ is SCAD derivative function.

$$P'_\lambda(|x|) = \begin{cases} \lambda & \text{if } |x| \le \lambda, \\ \frac{a\lambda - |x|}{a-1} & \text{if } \lambda < |x| \le a\lambda, \\ 0 & \text{if } a\lambda < |x|. \end{cases}$$

First, if $\theta_j^{(k)} \neq 0$, (3) can be rewritten as $\theta_j^{(k)} = r_j^{(k)}(1 + \frac{\lambda_1 w_j^{(k)}}{(\delta \cdot \|\theta_j^{(k)}\|_2)})^{-1}$ where $r_j^{(k)} = [\eta_j^{(k)}]^{t+1} - [\alpha_j^{(k)}]^t/\delta$. In this case, we can easily check that $\|r_j^{(k)}\|_2 > \lambda_1 w_j^{(k)}/\delta$. Second, if $\theta_j^{(k)} = 0$, (3) can be rewritten as $\lambda_1 w_j^{(k)} c = \delta r_j^{(k)}$ and we can easily check that $\|r_j^{(k)}\|_2 \le \lambda_1 w_j^{(k)}/\delta$. To sum up, $\theta_j^{(k)}$ which is a subvector of kth column of Θ with a K_n length is updated as follows:

$$\therefore \theta_j^{(k)} = \begin{cases} r_j^{(k)}(1 - \frac{\lambda_1 w_j^{(k)}}{(\delta \cdot \|r_j^{(k)}\|_2)}) & \text{if } \|r_j^{(k)}\|_2 > \lambda_1 w_j^{(k)}/\delta, \\ 0 & \text{if } \|r_j^{(k)}\|_2 \le \lambda_1 w_j^{(k)}/\delta. \end{cases}$$

3 Updating Z

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

Z is updated by optimizing

$$\min_{Z} \frac{1}{2}\|\frac{1}{b}\sum_{\ell=1}^{b}(Y - V^{(\ell)*}[H^*]^{t+1} - [E^{(\ell)}]^t + \frac{1}{\delta}[\Gamma^{(\ell)}]^t) - Z\|_F^2 + \sum_j P_{\lambda_2/(\delta b)}\{\sigma_j(Z)\}, \tag{4}$$

Using taylor approximation, 4 can be rewritten as follows:

$$\min_{Z} \frac{1}{2}\|\frac{1}{b}\sum_{\ell=1}^{b}(Y - V^{(\ell)*}[H^*]^{t+1} - [E^{(\ell)}]^t + \frac{1}{\delta}[\Gamma^{(\ell)}]^t) - Z\|_F^2 + \sum_j \nu_j\sigma_j(Z),$$

where $\nu_j = P'_{\lambda_2/(\delta b)}\{\sigma_j(\tilde{Z})\}$. The solution is obtained by an adaptively soft-thresholded singular value decomposition. Thus, Z_{t+1} is updated as follows:

$$\therefore Z_{t+1} = PD_{\lambda_2/(\delta\cdot b),\nu}(\Sigma)Q,$$

where $P\Sigma Q^T$ is singular value decomposition of $\frac{1}{b}\sum_{\ell=1}^{b}(Y - V^{(\ell)*}[H^*]^{t+1} - [E^{(\ell)}]^t + \frac{1}{\delta}[\Gamma^{(\ell)}]^t)$, $D_{\lambda_2/(\delta\cdot b),\nu}(\Sigma) = \mathrm{diag}(\{\sigma_i - \frac{\lambda_2\nu_i}{\delta\cdot b}\}_+)$, and $\{x\}_+ = \max(0, x)$. Furthermore, if X is full rank, $\hat{L}$ can be recovered by $\hat{L} = (X^TX)^{-1}X^T\hat{Z}$, and if X is not full rank, $\hat{L} = X^+Z$ where X^+ is pseudo-inverse of X.

4 Updating $E^{(\ell)}$

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

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

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

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

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

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

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

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

5 Updating multipliers : $\Gamma^{(\ell)}$ and A

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

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