

Closed form

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

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

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

Let (1) be Q .

1 Update 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

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

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

2 Update Θ_0 and Θ

Updating Θ_0 and Θ is equivalent to solving

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

Using Taylor approximation, $P_{\lambda_1}(\|\theta_j^{(k)}\|_2) \approx P_{\lambda_1}(\|\tilde{\theta}_j^{(k)}\|_2) + P'_{\lambda_1}(\|\tilde{\theta}_j^{(k)}\|_2)(\|\theta_j^{(k)}\|_2 - \|\tilde{\theta}_j^{(k)}\|_2)$. Thus, 2 can be rewritten as follows:

$$\min_{\theta_j^{(k)}} \langle [\alpha_j^{(k)}]^t, \theta_j^{(k)} \rangle + \frac{\delta}{2} \|\theta_j^{(k)} - [\eta_j^{(k)}]^{t+1}\|_2^2 + P'_{\lambda_1}(\|\tilde{\theta}_j^{(k)}\|_2) \|\theta_j^{(k)}\|_2.$$

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

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

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.

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

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

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

2.1 Update Θ for initial value

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

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

where $r_j^{(k)} = [\eta_j^{(k)}]^{t+1} - [\alpha_j^{(k)}]^t / \delta$.

3 Update Z

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

Z is updated by optimizing

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

Using taylor approximation, $P_{\lambda_2} \{\sigma_j(Z)\} \approx P_{\lambda_2} \{\sigma_j(\tilde{Z})\} + P'_{\lambda_2} \{\sigma_j(\tilde{Z})\} (\sigma_j(Z) - \sigma_j(\tilde{Z}))$. Thus, 4 can be rewritten as follows:

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

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

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

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

3.1 Update Z for initial value

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

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

where $P\Sigma Q^T$ is singular value decomposition of $\frac{1}{b} \sum_{\ell=1}^b (Y - V^{(\ell)*} [H^*]^{t+1} - [E^{(\ell)}]^t + \frac{1}{\delta} [\Gamma^{(\ell)}]^t)$, $D_{\lambda_2/\delta b}(\Sigma) = \text{diag}(\{\sigma_i - \frac{\lambda_2}{\delta b}\}_+)$, and $\{x\}_+ = \max(0, x)$.

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

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

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

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

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

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

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

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