成对融合惩罚与 Huber 损失结合方法的求解过程：

考虑目标函数

$$S_n(\boldsymbol{\alpha},\boldsymbol{\beta})=\frac{1}{n}\sum_{i=1}^{n}H_\tau\left(y_i-\alpha_i-\boldsymbol{x}_{\boldsymbol{i}}^{\top\boldsymbol{\beta}}\right)+\frac{1}{n}\sum_{i<i'}P_{\lambda_1}\left(\left|\alpha_i-\alpha_{i'}\right|\right)+\sum_{j=1}^{p}P_{\lambda_2}\left(\left|\beta_j\right|\right), \qquad (1)$$

为求解上式，考虑利用 ADMM 算法将其转变为容易求解的形式,

$$L_n(\boldsymbol{\alpha},\boldsymbol{\beta})=\frac{1}{n}\sum_{i=1}^{n}H_\tau\left(s_{1,i}\right)+\frac{1}{n}\sum_{i<i'}P_{\lambda_1}\left(\left|s_{2,ii'}\right|\right)+\sum_{j=1}^{p}P_{\lambda_2}\left(\left|\beta_j\right|\right),$$

$$\text{s.t.}\quad s_{1,i}=y_i-\alpha_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}, s_{2,ii'}=\alpha_i-\alpha_{i'}$$

其对应的增广拉格朗日函数为

$$\mathcal{L}_n(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{s}_1,\boldsymbol{v}_1,\boldsymbol{s}_2,\boldsymbol{v}_2)=L_n(\boldsymbol{\alpha},\boldsymbol{\beta})+\frac{\rho_1}{2}\|\boldsymbol{y}-\boldsymbol{\alpha}-\boldsymbol{X\beta}-\boldsymbol{s}_1\|_2^2+\langle\boldsymbol{y}-\boldsymbol{\alpha}-\boldsymbol{X\beta}-\boldsymbol{s}_1,\boldsymbol{v}_1\rangle \quad (2)$$

$$+\frac{\rho_2}{2}\|\boldsymbol{D\alpha}-\boldsymbol{s}_2\|_2^2+\langle\boldsymbol{D\alpha}-\boldsymbol{s}_2,\boldsymbol{v}_2\rangle$$

其中$\boldsymbol{D}=\{\boldsymbol{e}_i-\boldsymbol{e}_{i'},i<i'\}$，$\rho_1,\rho_2$是惩罚参数，$\boldsymbol{v}_1,\boldsymbol{v}_2$是拉格朗日乘子向量。在该框架下，上式的算法过程如下：对于迭代次数$l=0,1,2,\cdots$，不断重复如下步骤直至收敛：

$$\text{step 1: }\boldsymbol{\alpha}^{(l+1)}=\underset{\boldsymbol{\alpha}}{\operatorname{argmin}}\,\mathcal{L}_n\left(\boldsymbol{\alpha},\boldsymbol{\beta}^{(l)},\boldsymbol{s}_1^{(l)},\boldsymbol{v}_1^{(l)},\boldsymbol{s}_2^{(l)},\boldsymbol{v}_2^{(l)}\right),$$

$$\text{step 2: }\boldsymbol{\beta}^{(l+1)}=\underset{\boldsymbol{\beta}}{\operatorname{argmin}}\,\mathcal{L}_n\left(\boldsymbol{\alpha}^{(l+1)},\boldsymbol{\beta},\boldsymbol{s}_1^{(l)},\boldsymbol{v}_1^{(l)},\boldsymbol{s}_2^{(l)},\boldsymbol{v}_2^{(l)}\right),$$

$$\text{step 3: }\boldsymbol{s}_1^{(l+1)}=\underset{\boldsymbol{s}_1}{\operatorname{argmin}}\,\mathcal{L}_n\left(\boldsymbol{\alpha}^{(l+1)},\boldsymbol{\beta}^{(l+1)},\boldsymbol{s}_1,\boldsymbol{v}_1^{(l)},\boldsymbol{s}_2^{(l)},\boldsymbol{v}_2^{(l)}\right),$$

$$\text{step 4: }\boldsymbol{s}_2^{(l+1)}=\underset{\boldsymbol{s}_2}{\operatorname{argmin}}\,\mathcal{L}_n\left(\boldsymbol{\alpha}^{(l+1)},\boldsymbol{\beta}^{(l+1)},\boldsymbol{s}_1^{(l+1)},\boldsymbol{v}_1^{(l+1)},\boldsymbol{s}_2,\boldsymbol{v}_2^{(l)}\right),$$

$$\text{step 5: }\boldsymbol{v}_1^{(l+1)}=\boldsymbol{v}_1^{(l)}+\rho_1\left(\boldsymbol{y}-\boldsymbol{\alpha}^{(l+1)}-\boldsymbol{x\beta}^{(l+1)}-\boldsymbol{s}_1^{(l+1)}\right),$$

$$\text{step 6: }\boldsymbol{v}_2^{(l+1)}=\boldsymbol{v}_2^{(l)}+\rho_2\left(\boldsymbol{D\alpha}^{(l+1)}-\boldsymbol{s}_2^{(l+1)}\right).$$

具体地，step 1,2 有显式解：

$$\boldsymbol{\alpha}^{(l+1)}=\underset{\boldsymbol{\alpha}}{\operatorname{argmin}}\,\frac{\rho_1}{2}\left\|\boldsymbol{y}-\boldsymbol{X\beta}^{(l)}-\boldsymbol{s}_1^{(l)}+\boldsymbol{v}_1^{(l)}/\rho_1-\boldsymbol{\alpha}\right\|_2^2+\frac{\rho_2}{2}\left\|\boldsymbol{D\alpha}-\boldsymbol{s}_2^{(l)}+\boldsymbol{v}_2^{(l)}/\rho_2\right\|_2^2$$

$$=\left(\rho_1\boldsymbol{I}_n+\rho_2\boldsymbol{D}^\top\boldsymbol{D}\right)^{-1}\left(\rho_1\left(\boldsymbol{y}-\boldsymbol{X\beta}^{(l)}-\boldsymbol{s}_1^{(l)}+\boldsymbol{v}_1^{(l)}/\rho_1\right)+\rho_2\boldsymbol{D}^\top\left(\boldsymbol{s}_2^{(l)}-\boldsymbol{v}_2^{(l)}/\rho_2\right)\right);$$

step 2, step 3 要求解

$$\boldsymbol{\beta}^{(l+1)}=\underset{\boldsymbol{\beta}}{\operatorname{argmin}}\,\frac{\rho_1}{2}\left\|\boldsymbol{y}-\boldsymbol{\alpha}^{(l+1)}-\boldsymbol{s}_1^{(l)}+\boldsymbol{v}_1^{(l)}/\rho_1-\boldsymbol{X\beta}\right\|_2^2+\sum_{j=1}^{p}P_{\lambda_2}\left(\left|\beta_j\right|\right).$$

$$\boldsymbol{s}_1^{(l+1)}=\underset{\boldsymbol{s}_1}{\operatorname{argmin}}\,\frac{1}{n}\sum_{i=1}^{n}H_\tau\left(s_{1,i}\right)+\frac{\rho_1}{2}\left\|\boldsymbol{y}-\boldsymbol{\alpha}^{(l+1)}-\boldsymbol{X\beta}^{(l+1)}-\boldsymbol{s}_1\right\|_2^2\langle\boldsymbol{y}-\boldsymbol{\alpha}^{(l+1)}-\boldsymbol{X\beta}^{(l+1)}-\boldsymbol{s}_1,\boldsymbol{v}_1^{(l)}\rangle.$$

此与本文方法 $\boldsymbol{\beta},\boldsymbol{s}$ 解相同；

step 4 要求解

$$\boldsymbol{s}_2^{(l+1)}=\underset{\boldsymbol{s}_2}{\operatorname{argmin}}\,\frac{\rho_2}{2}\left\|\boldsymbol{D\alpha}^{(l+1)}-\boldsymbol{s}_2+\boldsymbol{v}_2^{(l)}/\rho_2\right\|_2^2+\frac{1}{n}\sum_{i<i'}P_{\lambda_1}\left(\left|s_{2,ii'}\right|\right)=\underset{\boldsymbol{s}_2}{\operatorname{argmin}}\,\frac{\rho_2}{2}\sum_{i<i'}\left(\alpha_i^{(l+1)}-\alpha_{i'}^{(l+1)}+v_{2,ii'}^{(l)}/\rho_2-s_{2,ii'}\right)^2+\frac{1}{n}\sum_{i<i'}P_{\lambda_1}\left(\left|s_{2,ii'}\right|\right).$$

当惩罚为 SCAD 惩罚时，其具体解可根据本文式(7)给出：

$$s_{2,ii'}^{(l+1)}=\begin{cases}\mathrm{ST}\left(c_{ii'},\lambda_1/a_{ii'}\right), & \left|c_{ii'}\right|\le\lambda_1+\lambda_1/a_{ii'}\\ \dfrac{\mathrm{ST}\left(c_{ii'},\gamma_1\lambda_1/\left((\gamma_1-1)a_{ii'}\right)\right)}{1-1/\left((\gamma_1-1)a_{ii'}\right)}, & \lambda_1+\lambda_1/a_{ii'}<\left|c_{ii'}\right|\le\gamma_1\lambda_1\\ c_{ii'}, & \left|c_{ii'}\right|>\gamma_1\lambda_1\,.\end{cases}$$

其中$a_{ii'}=n\rho_2, c_{ii'}=\alpha_i^{(l+1)}-\alpha_{i'}^{(l+1)}+v_{2,ii'}^{(l)}\ /\rho_2$.

给定初值$\boldsymbol{\alpha}^{(0)},\boldsymbol{\beta}^{(0)}$，参考以往文献，我们令$\boldsymbol{s}_1^{(0)}=\boldsymbol{y}-\boldsymbol{\alpha}^{(0)}-\boldsymbol{X}\boldsymbol{\beta}^{(0)},\boldsymbol{s}_2^{(0)}=\boldsymbol{D}\boldsymbol{\alpha}^{(0)},\boldsymbol{v}_1^{(0)}=\boldsymbol{0},\boldsymbol{v}_2^{(0)}=\boldsymbol{0}$.