$\boxed{\text{MIO:}}$

$\boxed{\text{Computational cost}}$ $\xRightarrow{\text{solution}}$ $\min F(\beta) \quad s.t. \ \beta_j = 0, \ j \notin I(\tilde{\beta})$

$f(\beta) + \lambda \|\beta\|_q$

$\uparrow I(\tilde{\beta})$

$\boxed{\text{DFO}}$: $\beta^{(m+1)} \in S\left(\beta^{(m)} - \frac{1}{L}\nabla f(\beta^{(m)}); k; \frac{\lambda}{L} \ell q\right) \Rightarrow I(\tilde{\beta})$

$= \{i, \tilde{\beta}_j \neq 0, j \in [p]\}$

Discrete First Order

thresholding operator

※ sensitive to the initialization $\beta^{(1)}$

$\boxed{\beta^{(1)}}, \lambda, L, \tau$

$(L \geq L_0)$

k - sparse

Δ

$\boxed{\text{Neighborhood Continuation:}}$ ① initialize $\hat{\beta}(\lambda_i; k_j) \leftarrow 0, \quad i, j \in [N] \times [r]$

$\Downarrow$

swap

$\beta^{(1)}$

② for $i \in [N], j \in [r]$

a. $\hat{\beta}(\lambda_a; k_b) \longrightarrow \text{DFO} \longrightarrow \{\hat{\beta}_{a,b}, F_{a,b}\}$

4 neighborhood initializations

$\boxed{\hat{\beta}(\lambda_a, k_b)}$ all OR one ?

b. $\hat{\beta}(\lambda_i; k_j) \leftarrow \hat{\beta}_{a,b}$ with minimum $\{F_{a,b}\}$

③ repeat ② until $\{F(\lambda_i; k_j)\}_{i,j}$ stop changing between successive sweeps.

$\boxed{\text{A (Randomized) Local Search Heuristic:}}$

for every nonzero initialization $\hat{\beta}(\lambda_a, k_b)$, randomly swap roughly 50% of the nonzero coefficients with an equal number of zero coefficients before passing it as an initialization to the DFO.

How to ensure 50% of nonzero $\xleftrightarrow{\text{swap}}$ number of zero ?