1 FISTA

Our function:

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}[\hat{\hat{p}}_t^{(k)}(1|S_t)(1-\hat{\hat{p}}_t^{(k)}(1|S_t))(\tilde{Y}_{i,t+1}^{(DR)}-f_t(S_t)^T\beta)^2]+\lambda\|\beta\|_1-w^t\beta \tag{1}$$

We can let

$$f(\beta)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}[\hat{\hat{p}}_t^{(k)}(1|S_t)(1-\hat{\hat{p}}_t^{(k)}(1|S_t))(\tilde{Y}_{i,t+1}^{(DR)}-f_t(S_t)^T\beta)^2]-\omega^t\beta$$

$$g(\beta)=\lambda\|\beta\|_1$$

Then

$$\nabla f(\beta)=\frac{-2}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\hat{p}}_t^{(k)}(1|S_t)(1-\hat{\hat{p}}_t^{(k)}(1|S_t))(\tilde{Y}_{i,t+1}^{(DR)}-f_t(S_t)^T\beta)f_t(S_t)-\omega$$

$$\nabla^2 f(\beta)=\frac{2}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\hat{p}}_t^{(k)}(1|S_t)(1-\hat{\hat{p}}_t^{(k)}(1|S_t))f_t(St)f_t(S_t)^T$$

Then L the Lipschitz constant of ∇f is the $L=sup\|\nabla^2 f(\beta)\|$ the largest eigenvalue.

In the algorithm we also need to get $p_L(y_k)$ and

$$p_L(y_k)=\underset{\beta}{\operatorname{argmin}}\left\{\lambda\|\beta\|_1+\frac{L}{2}(\beta-[y_k-\frac{1}{L}\nabla f(y_k)])^2\right\}$$

Then $x_k=(|y_k-\frac{1}{L}\nabla f(y_k)|-\frac{\lambda}{L})_+sgn(y_k-\frac{1}{L}\nabla f(y_k))$

The rest is just following the FISTA algorithm in page 12. Better to use FISTA with backtracking, and you even don't need L.