1 ADMM

Our optimization problem is:

$$\begin{aligned}
\min_{\mathbf{z},\boldsymbol{\lambda}} \quad & \frac{1}{2}\mathbf{z}^T K\mathbf{z} - \sum_{n=1}^{m} \lambda_n + \delta(\boldsymbol{\lambda}) \\
\text{s.t.} \quad & Y\boldsymbol{\lambda} - \mathbf{z} = 0 \\
& \mathbf{y}^T\boldsymbol{\lambda} = 0
\end{aligned}$$

with $\delta(\boldsymbol{\lambda}) = \begin{cases} 0 & 0 \le \boldsymbol{\lambda} \le C \\ \infty & \text{otherwise} \end{cases}$, $(K)_{i,j} = K(\mathbf{x}_i, \mathbf{x}_j)$, $Y = diag(\mathbf{y})$
Let $\Gamma(\boldsymbol{\lambda}, \mathbf{z}) = \begin{bmatrix} Y \\ \mathbf{y}^T \end{bmatrix} \boldsymbol{\lambda} - \begin{bmatrix} I \\ 0 \end{bmatrix} \mathbf{z}$, then $\frac{\partial \Gamma}{\partial \boldsymbol{\lambda}} = \begin{bmatrix} Y & \mathbf{y} \end{bmatrix}$ $\quad \frac{\partial \Gamma}{\partial \mathbf{z}} = \begin{bmatrix} -I & 0 \end{bmatrix}$
The augumented Lagrange function is:

$$L_\rho(\boldsymbol{\lambda}, \mathbf{z}, \boldsymbol{\mu}) = \frac{1}{2}\mathbf{z}^T K\mathbf{z} - \sum_{n=1}^{m} \lambda_n + \delta(\boldsymbol{\lambda}) + \frac{\rho}{2}\|\Gamma(\boldsymbol{\lambda}, \mathbf{z}) + \boldsymbol{\mu}\|^2$$

The ADMM algorithm is:

step 1:

$$\begin{aligned}
\mathbf{z}^{t+1} &= \arg\min_{\mathbf{z}} \quad L_\rho(\boldsymbol{\lambda}^t, \mathbf{z}, \boldsymbol{\mu}^t) \\
\nabla_{\mathbf{z}} L_\rho &= K\mathbf{z} + \rho\frac{\partial \Gamma}{\partial \mathbf{z}}(\Gamma(\boldsymbol{\lambda}^t, \mathbf{z}) + \boldsymbol{\mu}^t) \\
&= K\mathbf{z} + \rho(\mathbf{z} - Y\boldsymbol{\lambda}^t - \begin{bmatrix} I & 0 \end{bmatrix} \boldsymbol{\mu}^t) \\
\mathbf{z}^{t+1} &= \rho(K + \rho I)^{-1}(Y\boldsymbol{\lambda}^t + \begin{bmatrix} I & 0 \end{bmatrix} \boldsymbol{\mu}^t)
\end{aligned}$$

step 2:

$$\begin{aligned}
\boldsymbol{\lambda}^{t+1} &= \arg\min_{\boldsymbol{\lambda}} \quad L_\rho(\boldsymbol{\lambda}, \mathbf{z}^{t+1}, \boldsymbol{\mu}^t) \\
&= \arg\min_{\boldsymbol{\lambda}}\{-\sum_{n=1}^{m} \lambda_n + \delta(\boldsymbol{\lambda}) + \frac{\rho}{2}\|\Gamma(\boldsymbol{\lambda}, \mathbf{z}^{t+1}) + \boldsymbol{\mu}^2\|^2\} \\
&= \pi(\tilde{\boldsymbol{\lambda}}^{t+1}) \\
\tilde{\boldsymbol{\lambda}}^{t+1} &= (Y^2 + \mathbf{y}\mathbf{y}^T)^{-1}(\frac{1}{\rho}\mathbf{1} + Y\mathbf{z}^{t+1} - \begin{bmatrix} Y & \mathbf{y} \end{bmatrix} \boldsymbol{\mu}^t) \\
\lambda_i^{t+1} &= \begin{cases} 0 & \tilde{\lambda}_i^{t+1} < 0 \\ C & \tilde{\lambda}_i^{t+1} > C \\ \tilde{\lambda}_i^{t+1} & \text{otherwise} \end{cases}
\end{aligned}$$

step 3:

$$\boldsymbol{\mu}^{t+1} = \boldsymbol{\mu}^t + \Gamma(\boldsymbol{\lambda}^{t+1}, \mathbf{z}^{t+1})$$