

Program lasso regularization (lasso2):

$$\begin{aligned} & \text{minimize} && \frac{1}{2} \xi^\top \xi + \tau \mathbf{1}_n^\top \epsilon \\ & \text{subject to} && \\ & && y - Xw = \xi \\ & && w \leq \epsilon \\ & && -w \leq \epsilon. \end{aligned}$$

minimizing over ξ, w and ϵ , with $y, \xi \in \mathbb{R}^m$, and $w, \epsilon, \mathbf{1}_n \in \mathbb{R}^n$.

The constraints $w \leq \epsilon$ and $-w \leq \epsilon$ are equivalent to $|w_i| \leq \epsilon_i$ for $i = 1, \dots, n$, so for an optimal solution we must have $\epsilon \geq 0$ and $|w_i| = \epsilon_i$, that is, $\|w\|_1 = \epsilon_1 + \dots + \epsilon_n$.

The Lagrangian $L(\xi, w, \epsilon, \lambda, \alpha_+, \alpha_-)$ is given by

$$\begin{aligned} L(\xi, w, \epsilon, \lambda, \alpha_+, \alpha_-) &= \frac{1}{2} \xi^\top \xi + \tau \mathbf{1}_n^\top \epsilon + \lambda^\top (y - Xw - \xi) \\ &\quad + \alpha_+^\top (w - \epsilon) + \alpha_-^\top (-w - \epsilon) \\ &= \frac{1}{2} \xi^\top \xi - \xi^\top \lambda + \lambda^\top y \\ &\quad + \epsilon^\top (\tau \mathbf{1}_n - \alpha_+ - \alpha_-) + w^\top (\alpha_+ - \alpha_- - X^\top \lambda), \end{aligned}$$

with $\lambda \in \mathbb{R}^m$ and $\alpha_+, \alpha_- \in \mathbb{R}_+^n$. Since the objective function is convex and the constraints are affine (and thus qualified), the Lagrangian L has a minimum with respect to the primal variables, ξ, w, ϵ iff $\nabla L_{\xi, w, \epsilon} = 0$. Since the gradient $\nabla L_{\xi, w, \epsilon}$ is given by

$$\nabla L_{\xi, w, \epsilon} = \begin{pmatrix} \xi - \lambda \\ \alpha_+ - \alpha_- - X^\top \lambda \\ \tau \mathbf{1}_n - \alpha_+ - \alpha_- \end{pmatrix},$$

we obtain the equations

$$\begin{aligned} \xi &= \lambda \\ \alpha_+ - \alpha_- &= X^\top \lambda \\ \alpha_+ + \alpha_- &= \tau \mathbf{1}_n. \end{aligned}$$

Using these equations, the dual function $G(\lambda, \alpha_+, \alpha_-) = \min_{\xi, w, \epsilon} L$ is given by

$$\begin{aligned} G(\lambda, \alpha_+, \alpha_-) &= \frac{1}{2} \xi^\top \xi - \xi^\top \lambda + \lambda^\top y = \frac{1}{2} \lambda^\top \lambda - \lambda^\top \lambda + \lambda^\top y \\ &= -\frac{1}{2} \lambda^\top \lambda + \lambda^\top y = -\frac{1}{2} (\|y - \lambda\|_2^2 - \|y\|_2^2), \end{aligned}$$

so

$$G(\lambda, \alpha_+, \alpha_-) = -\frac{1}{2} (\|y - \lambda\|_2^2 - \|y\|_2^2).$$