

penalty	function	optimizer	reference
ridge	$p(x_j) = \lambda x_j^2$	glmnet, ista	(Hoerl & Kennard, 1970)
lasso	$p(x_j) = \lambda x_j $	glmnet, ista	(Tibshirani, 1996)
adaptiveLasso	$p(x_j) = \frac{1}{w_j} \lambda x_j $	glmnet, ista	(Zou, 2006)
elasticNet	$p(x_j) = \alpha \lambda x_j + (1 - \alpha) \lambda x_j^2$	glmnet, ista	(Zou & Hastie, 2005)
cappedL1	$p(x_j) = \lambda \min(x_j , \theta); \theta > 0$	glmnet, ista	(Zhang, 2010)
lsp	$p(x_j) = \lambda \log(1 + x_j /\theta); \theta > 0$	glmnet, ista	(Candès et al., 2008)
scad	$p(x_j) = \begin{cases} \lambda x_j & \text{if } x_j \leq \lambda \\ \frac{-x_j^2 + 2\theta \lambda x_j - \lambda^2}{2(\theta - 1)} & \text{if } \lambda < x_j \leq \lambda \theta ; \theta > 2 \\ (\theta + 1) \lambda^2 / 2 & \text{if } x_j \geq \theta \lambda \end{cases}$	glmnet, ista	(Fan & Li, 2001)
mcp	$p(x_j) = \begin{cases} \lambda x_j - x_j^2 / (2\theta) & \text{if } x_j \leq \theta \lambda \\ \theta \lambda^2 / 2 & \text{if } x_j > \lambda \theta \end{cases}; \theta > 0$	glmnet, ista	(Zhang, 2010)