Module 4:

Linear regression with multiple features

$$\hat{y}_i = \hat{\beta}_0 + \sum_{j=1}^{p} \hat{\beta}_j x_{i,j}\,, \quad \forall i = 1,\ 2,\ \ldots N$$

Least squares estimates: obtained by minimizing the residual sum of squares (RSS)

$$RSS(\hat{\beta}_0, \ldots, \hat{\beta}_p) = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

$$= \sum_{i=1}^{N} \left(y_i - (\hat{\beta}_0 + \sum_{j=1}^{p} \hat{\beta}_j x_{i,j})\right)^2$$

Module 7:

Penalized (or **regularized**) RSS

$$RSS^{regularized}(\hat{\beta}_0, \ldots, \hat{\beta}_p, \lambda) = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 + \lambda R(\hat{\beta}_1, \ldots, \hat{\beta}_p)$$

$$= \sum_{i=1}^{N} \left(y_i - (\hat{\beta}_0 + \sum_{j=1}^{p} \hat{\beta}_j x_{i,j})\right)^2 + \lambda R(\hat{\beta}_1, \ldots, \hat{\beta}_p)$$

Ridge Regression

$$R(\hat{\beta}_1, \ldots, \hat{\beta}_p) = \sum_{j=1}^{p} |\hat{\beta}_j|^2$$

$$RSS^{ridge}(\hat{\beta}_0, \ldots, \hat{\beta}_p, \lambda) = \sum_{i=1}^{N} \left(y_i - (\hat{\beta}_0 + \sum_{j=1}^{p} \hat{\beta}_j x_{i,j})\right)^2 + \lambda \sum_{j=1}^{p} |\hat{\beta}_j|^2$$

The Lasso

$$R(\hat{\beta}_1, \ldots, \hat{\beta}_p) = \sum_{j=1}^{p} |\hat{\beta}_j|$$

$$RSS^{lasso}(\hat{\beta}_0, \ldots, \hat{\beta}_p, \lambda) = \sum_{i=1}^{N} \left(y_i - (\hat{\beta}_0 + \sum_{j=1}^{p} \hat{\beta}_j x_{i,j})\right)^2 + \lambda \sum_{j=1}^{p} |\hat{\beta}_j|$$