Ridge and Lasso Regression: The Constraints

Lasso constraint

$$|\beta_1| + |\beta_2| \le t$$
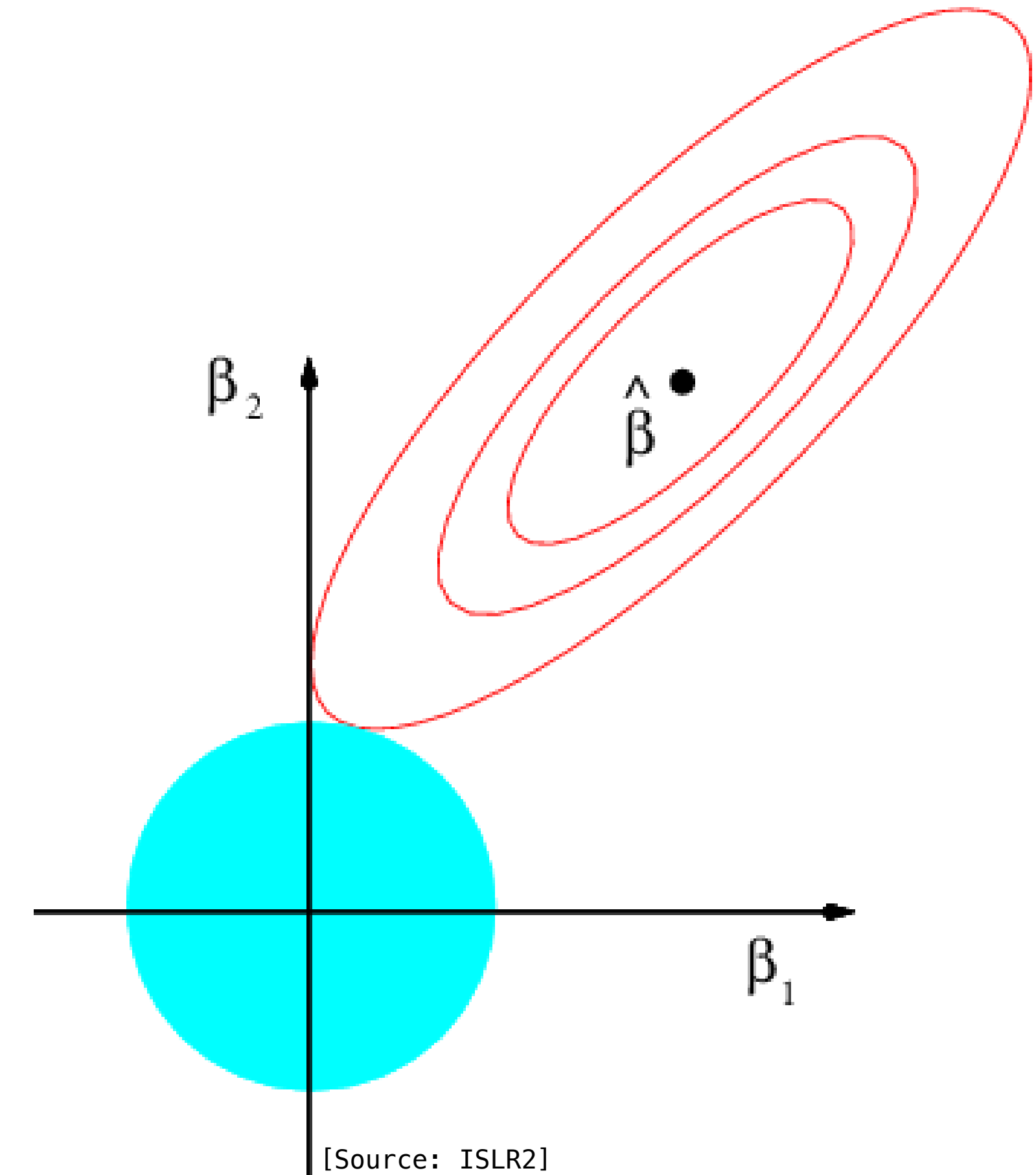
Ridge constraint

$$\beta_1^2 + \beta_2^2 \le t$$
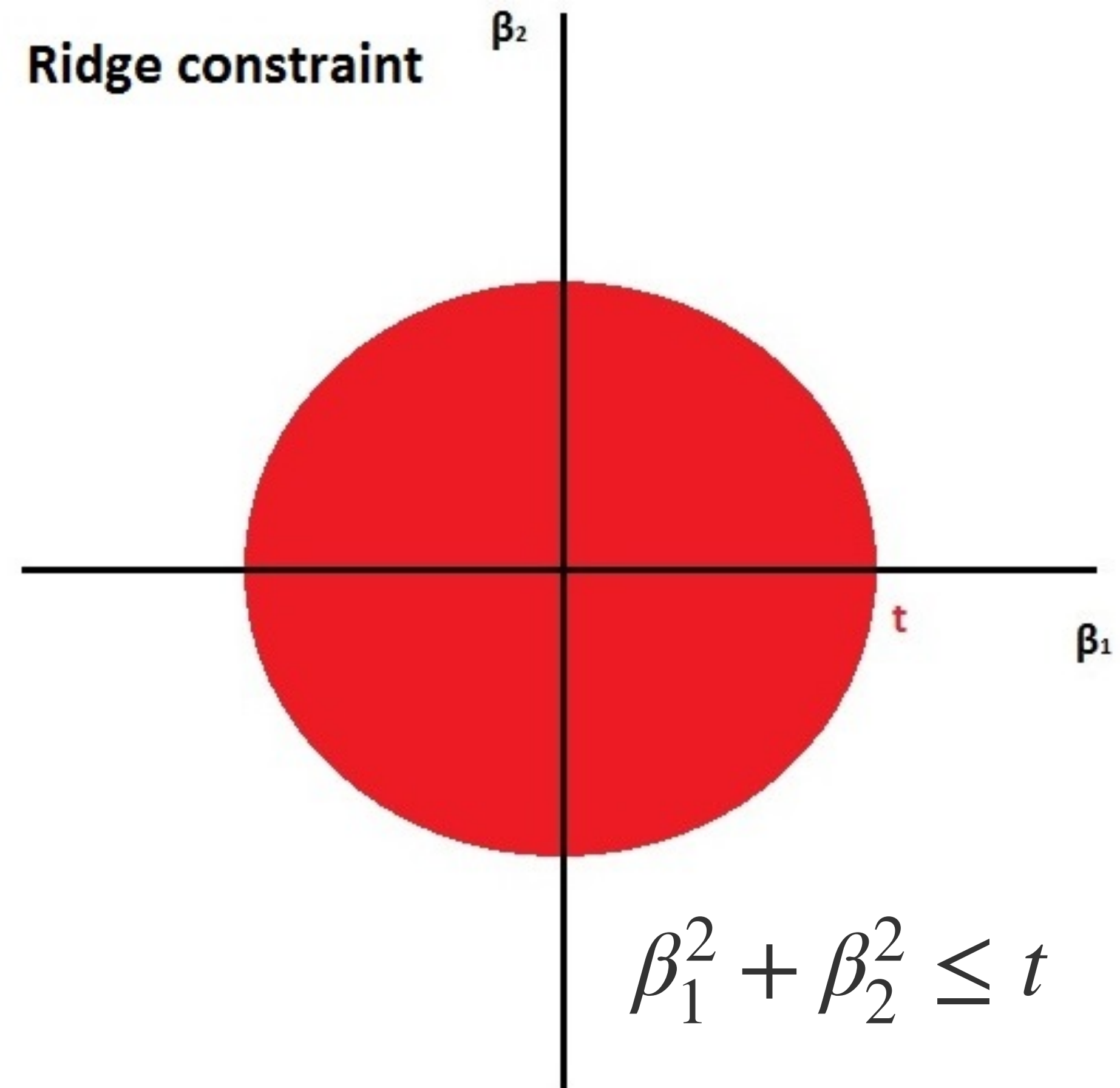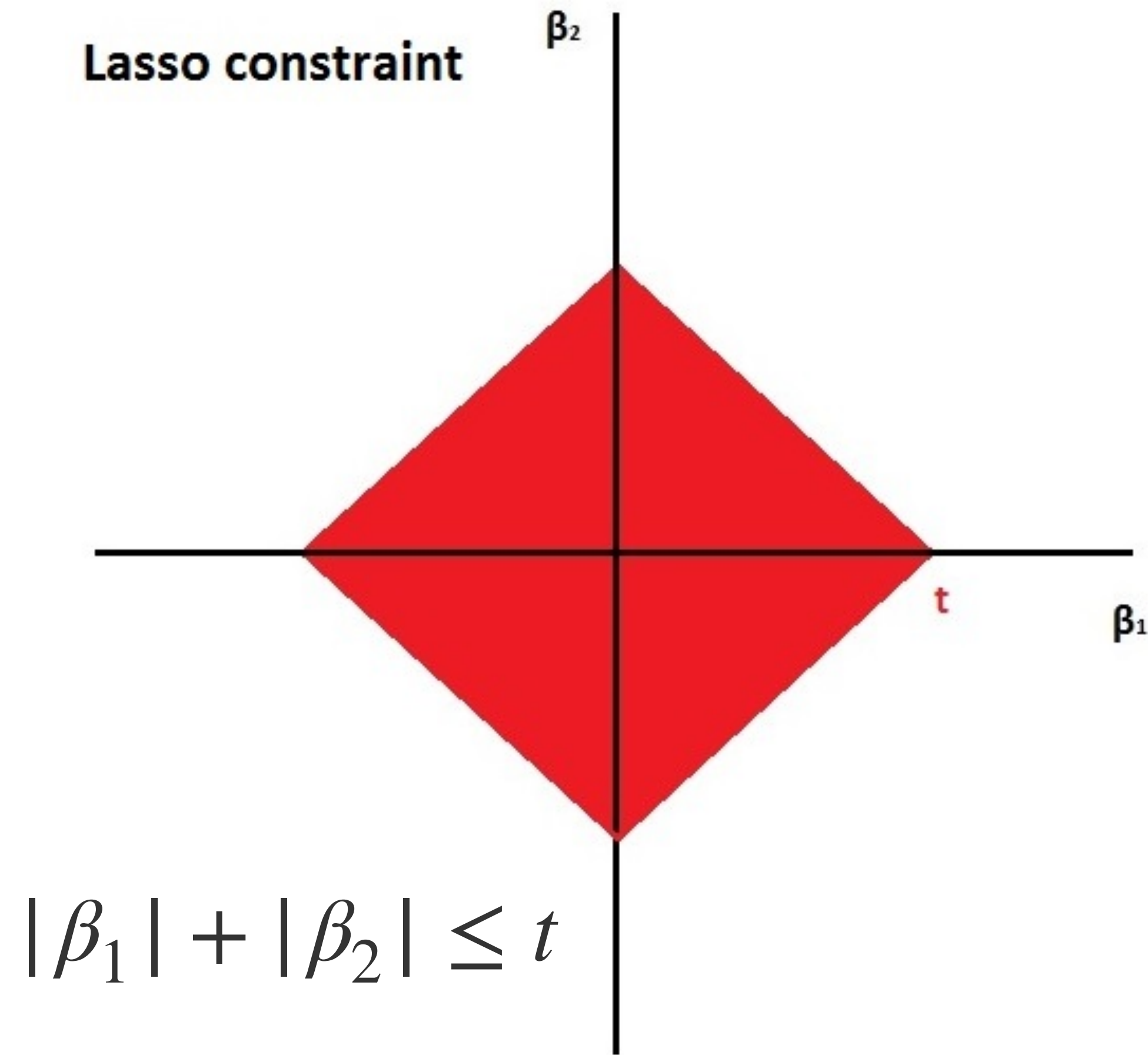
Ridge Regression

Least Squares produces estimates by minimizing

$$\text{RSS} = \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \sum_{j-1}^{p}\hat{\beta}_{j_1}x_{ij})^2$$

Ridge regression instead minimizes

$$\underbrace{\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \sum_{j-1}^{p}\hat{\beta}_{j_1}x_{ij})^2}_{\text{model fit}} + \underbrace{\lambda\sum_{j=1}^{p}\beta_j^2}_{\text{penalty}} = \text{RSS} + \lambda\sum_{j=1}^{p}\beta_j^2$$

[Source: ISLR2]

where $\lambda \geq 0$ is the tuning parameter controlling trade off

between model fit and size of coefficients ($\lambda \to \infty$, $\hat{\beta}_j \to 0$)