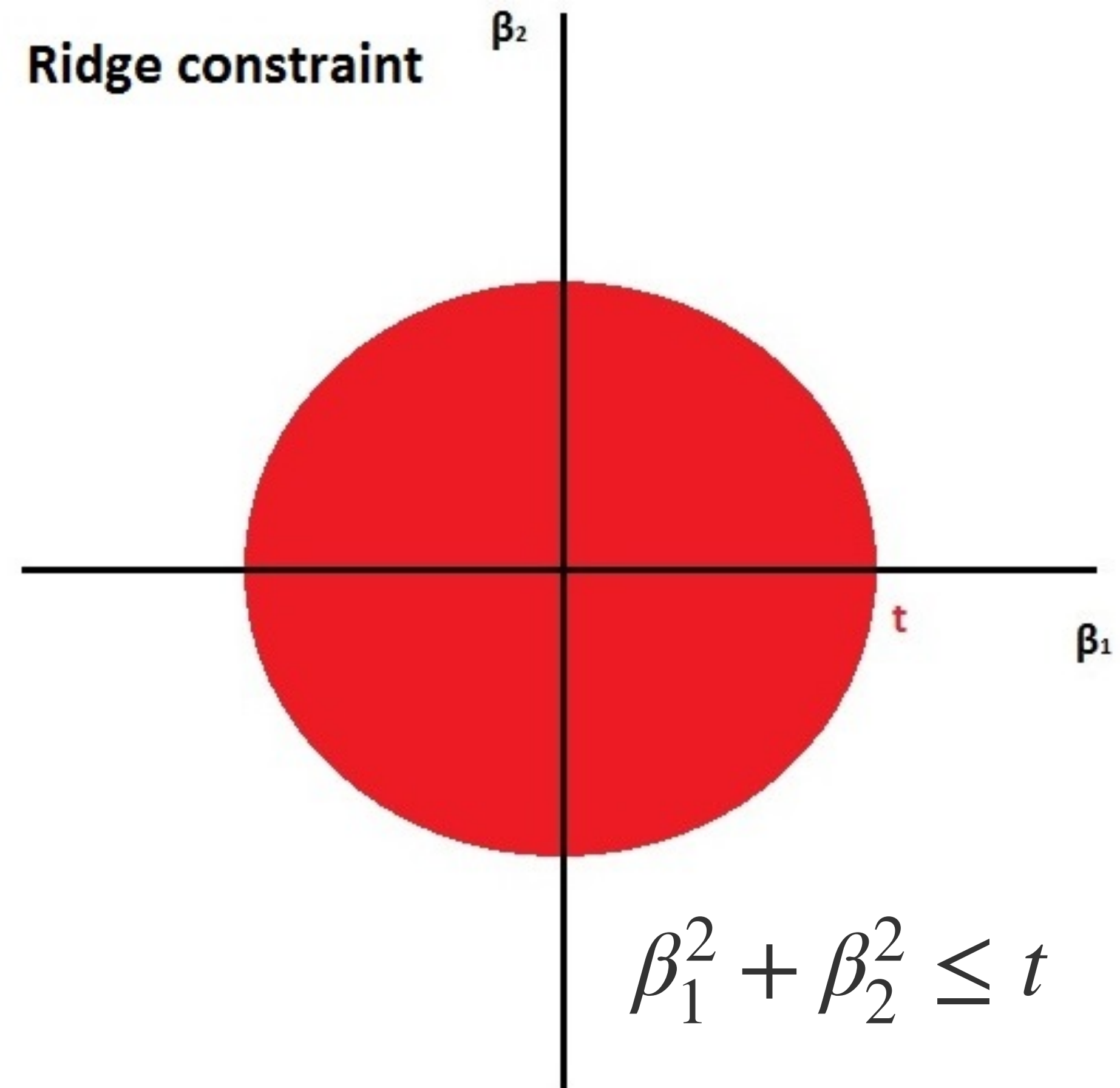
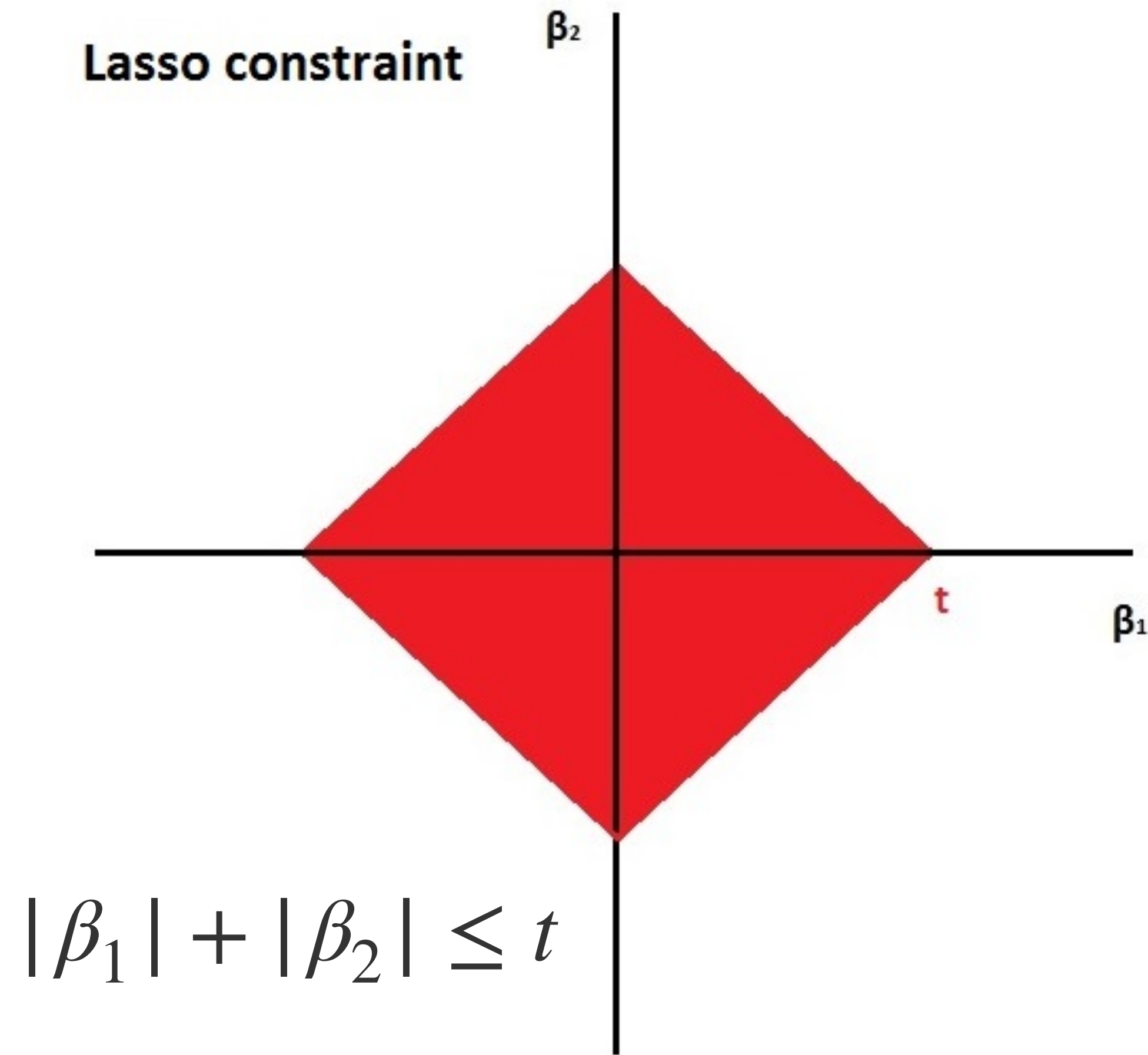


Ridge and Lasso Regression: The Constraints



Ridge Regression

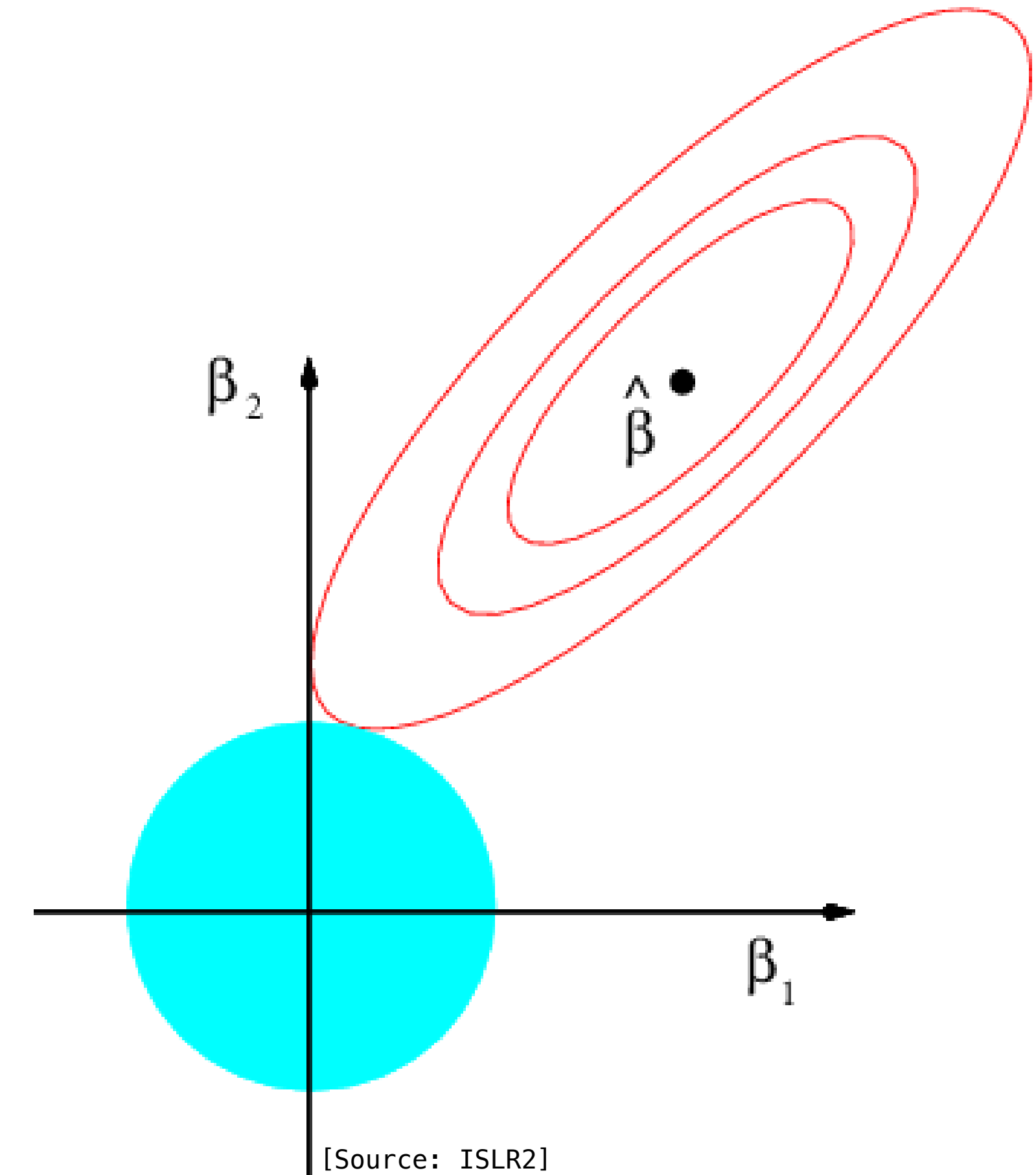
ridge uses ℓ_2 penalty

Least Squares produces estimates by minimizing

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

Ridge regression instead minimizes

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



where $\lambda \geq 0$ is **the tuning parameter** controlling trade off between model fit and size of coefficients ($\lambda \rightarrow \infty, \hat{\beta}_j \rightarrow 0$)