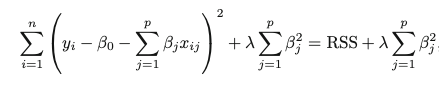
Ridge Regression

Ridge regression is a fitting procedure which allows to estimate the regression coefficients β’s by shrinking the impact of least important explanatory variables in association with a given response.

In particular, coefficient estimates with Ridge are the ones obtained by minimizing the following quantity



Leading to the minimization problem



under the constraint



Where t is a finite positive number.

The parameter λ ≥ 0, called tuning or regularization parameter, allows to tune the shrinkage penalty term. In particular,

* λ = 0 implies that the estimates obtained using Ridge and the ones obtained using OLS happen to be the same, since there is not a shrinkage penalty term.
* λ → ∞ implies a growth in the penalty and coefficient estimates approaching to zero.

As a different set of coefficient estimates is produced for each λ considered, the choice of the right tuning parameter is fundamental.

By solving the minimization problem of the cost function



It is possible to obtain



Where the inversion of the matrix is possible since it does not have singularity problems for finite values of the parameter λ.

In the expression above, ***𝐼*** is a pxp identity matrix and there is, again, the constraint  where t is a finite number larger than zero.

Once having all the coefficient estimates computed, predicted values of the response variable can be computed as

***𝑋X***𝜷Ridge ***= Y***predicted

Analytical expression for the expected value and variance of 𝜷Ridge’s can be obtained as





By observing those formulas, it is clear that the ridge estimator is biased and that, for λ → ∞, the variance of ridge estimates goes to zero.