Comparative study of some methods of regularization in the linear model.

Ibtissem Redjam[1], Dalel Zerdazi[2]

[1]Abdelhafid Boussouf University Center of Mila, Mila, Algeria

[2]University of Constantine 1, Constantine, Algeria

ibtissem.redjam@centre-univ-mila.dz, dalel.zerdazi@hotmail.com

In statistics, multiple linear regression is a mathematical regression method extending simple linear regression to describe the variations of a endogenous (dependent) variable, associated with variations in several exogenous variables (independent).

One of the most recurrent issues in the statistical learning field is the quality of generalization. Methods for solving general problems are mainly depicted in the context of the linear model.

In this study, we will focus mainly on the issues related to regression and modeling. The purpose is to present regularization methods which allow to produce parsimonious solutions and improve the quality of generalization within our statistical model. Simulation studies are proposed to guarantee the effectiveness of the studied methods in order to test the different cases which are mentioned in the theorectical part.

The purpose of regression, whether linear or nonlinear, is to assess the quality generalization of the method that generates the estimators. We address the problem of how to assess quality of generalization within the following variatons:

- Each method provides its own estimator, and different samples produce different estimators for the same method.

- The quality of generalization of a method is the quality of generalization in view of all the possible samples.

- A method is better than another if its quality of generalization is better on all possible samples.