

VARIABLE SELECTION WITH GENERALIZED SIMSEL METHODS

Silvelyn Zwanzig

Uppsala University , Sweden

The principle of SimSel is to study the influence of stepwise perturbation of the variable under consideration on the quality of a model fit. The main idea is that the disturbance of an unimportant variable will have no effect. Furthermore an independent pseudo variable is implemented to the data set which serves as an untreated control group and is explored for estimating the null distribution in the test step. The SimSel method introduced in Eklund,Zwanzig (2012) is based on an approximative quadratic model and the least squares criterion. Now generalizations in two extensions are considered.

- Several variables are disturbed simultaneously.
- Instead of least squares other criteria for the model fit are applied, as Ridge or Lasso.

The goal is to detect unimportant variables as well as variables, which can be replaced by others.

Keywords: Variable selection, Ridge, Lasso, Perturbation, Simex.

References:

- M. Eklund and S. Zwanzig (2012). SimSel: a new simulation method for variable selection. *Journal of Statistical Computation & Simulation*. 82, 515-527
- M. Eklund and S. Zwanzig (2009). Ridge-SimSel - a generalization of the variable selection method SimSel to multicollinear data sets. *U.U.D.M. 2009:12*, ISSN 1101-3591.