

The PLS approach to Generalised Linear Models and Causal Path Modeling: Algorithms and Applications

Vincenzo Esposito Vinzi
(Department of Mathematics and Statistics, University of Napoli)
vincenzo.espositovinzi@unina.it

Abstract

PLS (Partial Least Squares or Projection to Latent Structures) methods represent a new generation of statistical procedures. They cover a very broad area of statistical methods, from regression to generalised linear models, from data analysis to causal and path modeling, with several theoretical and statistical properties being demonstrated.

The success of PLS methods is widely recognized in chemistry, oil industry, food industry, medicine, biology. These methods are now getting the same level of success in business and industry, especially in the areas of marketing and strategic management. PLS regression (PLS1 for the univariate case and PLS2 for the multivariate one) fits particularly well to situations where classical OLS regression is unstable or not feasible at all (high degree of multicollinearity, small number of observations compared to the number of variables, missing data).

There exist several versions of the PLS algorithm within the regression framework. At first, the paper will deal with some extensions of these algorithms to the case of ordinal response variables (PLS logistic regression) as well as to the wider framework of generalised linear models. Then, it will be shown how the PLS principle allows to develop a distribution-free approach to path modeling as an alternative to LISREL for the study of causal relationships. Comparisons between LISREL and PLS path modeling will be shown with respect to their objectives, statistical properties and performances. In particular, the paper will focus on the algorithmic/computational aspects of the presented methods as well as on the related problems of validation and variables selection within a non parametric context. An outline of the open problems and of the most interesting research perspectives will be also sketched. Examples on real data will be finally shown in order to provide insights on the interpretation rules of the proposed methods.