Partial Robust M-regression

S. Serneels[1], C. Croux[2], P. Filzmoser[3] and P.J. Van Espen[1]

[1] Department of Chemistry, University of Antwerp, Belgium
[2] Department of Applied Economics, K.U. Leuven, Belgium
[3] Department of Statistics and Probability Theory, Vienna University of Technology, Austria

Keywords: Calibration, Partial least squares, M-estimator.

Abstract

Partial Least Squares (PLS) is a standard statistical method in chemometrics. It can be considered as an incomplete, or "partial", version of the Least Squares estimator of regression, applicable when high or perfect multicollinearity is present in the predictor variables. The Least Squares estimator is well-known to be an optimal estimator for regression, but only when the error terms are normally distributed. In absence of normality, and in particular when outliers are in the data set, other more robust regression estimators have better properties.

We will present a "partial" version of M-regression estimators (Serneels, Croux, Filzmoser, and Van Espen, 2005). If an appropriate weighting scheme is chosen, partial M-estimators become entirely robust to any type of outlying points. It will be shown that robust M-regression outperforms existing methods for robust PLS regression in terms of statistical precision and computational speed, while keeping the robustness properties. The method is applied to a data set consisting of EPXMA spectra of archæological glass vessels. This data set contains several outliers, and the advantages of Partial Robust M-regression are illustrated. Applying Partial Robust M-regression yields much smaller prediction errors for noisy calibration samples than PLS. On the other hand, if the data follow perfectly well a normal model, the loss in efficiency to be paid for is very small.

References

S. Serneels, C. Croux, P. Filzmoser and P.J. Van Espen (2005). Partial robust M-regression. *Journal of Chemometrics and Intelligent Laboratory Systems.* Conditionally accepted, 2005.

M. Tenenhaus (1998). *L'Approche PLS.* Avec le concours de la Fondation HEC. Groupe HEC, 78351 Jouy-en-Josas Cedex, France.

H. Wold (1973). In: P.R. Krishnaiah (ed.), *Multivariate Analysis III*, Academic Press, New York, pp. 383–407.