Combining Random and Specific Directions for Robust Estimation of High-Dimensional Multivariate Data

D. Peña and J. Prieto[1]

[1] University Carlos III de Madrid , Madrid 126, Getafe 28906, Spain

Keywords: Projection pursuit, Stahel-Donoho estimator, Kurtosis coefficient.

1 Abstract

A powerful procedure for robust estimation of shape and location with multivariate data is proposed. The procedure searches for outliers in univariate projections on directions that are obtained both randomly, using a modification of the Stahel-Donoho procedure, and by maximizing and minimizing the kurtosis coefficient of the projected data, as proposed by Peña and Prieto (2001). It is shown that the resulting procedure inherits the good properties of both methods and provides a powerful tool for Data Mining. The performance of the procedure is illustrated with a Monte Carlo experiment.

References

D. Peña and F.J. Prieto (2001,a). Robust Covariance Matrix Estimation and Multivariate Outlier Detection, *Technometrics,* 3, 286-310.

D. Peña and F.J. Prieto (2001,b). Cluster Identification Using Projections. *Journal of the American Statistical Association*, 96, 1433–1445.