# Diagnostic biplots for linear models

### Jan Graffelman

Universitat Politècnica de Catalunya, Spain

#### Abstract

In observational studies multicollinearity is a common problem in regression models (Belsley et. al, 1980). Several statistics are available that help to detect collinearity such as variance inflation factors (vif's), tolerances, correlations between regression coefficients, condition indices and others. However, it often remains a complicated task to identify the offending predictors and their relationships. Graphics are extensively used in the analysis of the residuals of a regression model, but a rarely used for an analysis of multicollinearity. Recently, Friendly and Kwan (2009) have suggested the use of biplots (Gabriel, 1971; Gower and Hand, 1996) as a graphical tool for uncovering collinearity in multiple regression. They proposed biplots involving the last principal component of the matrix of predictors in particular. Biplots have mostly been used to produce pictures of data matrices (raw quantitative data matrices, contingency tables), but can also be constructed for several matrices involved in a linear model that are of diagnostic interest. Several examples of such diagnostic biplots will be discussed.

### **Keywords**

Biplot, Multicollinearity, Variance inflation factor.

## References

Belsley, D.A., Kuh, E., and Welsch, R.E. (1980). Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York: Wiley.

Friendly, M. and Kwan, E. (2009). Where's Waldo? Visualizing Collinearity Diagnostics. *Amer. Statist.* 63, 56–65.

Gabriel, K.R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika* 58(3), 453–467.

Gower, J.C. and Hand, D.J. (1996). Biplots. London: Chapman & Hall.