Biplot videos

Michael Greenacre

Universitat Pompeu Fabra, Barcelona, Catalunya
email: michael.greenacre@upf.edu

Keywords: biplot, matrix product, regression, singular-value decomposition, triplot.

Most multivariate statistical methods that are used in practice have a common theory of matrix products – such methods include multiple regression, principal component analysis, correspondence analysis, log-ratio analysis, linear discriminant analysis, canonical correlation analysis, as well as several constrained variants of these methods which mix rank reduction with linear constraints, for example redundancy analysis and canonical correspondence analysis. Where there is a matrix product, there is a *biplot*, a type of multivariate scatterplot that graphically represents two sets of objects – usually cases and variables – in a common vector space. In the linearly constrained versions, the constraining variables can be added to the biplot to obtain what is often called a "triplot".

For a couple of years I have been experimenting with dynamic graphics in statistics, producing video animations of models, algorithms and results. The article by Greenacre and Hastie (2010) is a first product of this work, containing four videos embedded in the article where there would otherwise be static figures. The videos illustrate much more clearly the models and results of the complex statistical analyses presented in the article. Other articles with video content as supplementary material are by Greenacre (2010a, 2011).

Mainly as a complement to my book "Biplots in Practice" (Greenacre, 2010b) I have been developing a series of video animations, not only as an educational tool but also opening up new ways of understanding and interpreting multivariate statistical results. In this talk I will take you on a moving-picture journey from the simplest biplot, based on multiple regression, through several illustrations of other well-known multivariate methods, and finally the canonical correspondence analysis of a large ecological data set, including hundreds of cases and hundreds of dependent and independent variables.

References

[1] Greenacre, M. (2010a). Correspondence analysis of raw data. *Ecology* **91**, 958–963.

[2] Greenacre, M. (2010b). *Biplots in Practice*. BBVA Foundation, Madrid. Freely downloadable from http://www.multivariatestatistics.org.

[3] Greenacre, M. (2011). Contribution biplots. Working paper 1162, Department of Economics and Business, Pompeu Fabra University. Downloadable from http://www.econ.upf.edu/en/research/onepaper.php?id=1162.

[4] Greenacre, M., Hastie, T. (2010). Dynamic visualization of statistical learning algorithms in the context of high-dimensional textual data. *Journal of Web Semantics* **8**, 163–168.