

Correspondence analysis: Jack of all trades, Master of one

Michael Greenacre

Abstract Correspondence analysis (CA) as a method of multivariate data visualization has an almost 60-year history and I have been a part of it for 46 years since my doctoral studies in Prof. Jean-Paul Benzécri's laboratory in Paris in the years 1973-1975. In this talk, I reflect on (i) how this method has developed theoretically and matured over the last decades, (ii) what distinguishes it from other approaches to dimension reduction, and (iii) why this method is an important and essential addition to the applied multivariate toolbox.

These days CA has found use in almost every field of multivariate research. Published applications as well as methodological papers on CA have been rising exponentially, especially in the fields of biology and ecology. CA also has intimate connections with other multivariate areas, notably compositional data analysis, discriminant analysis, analysis of variance and principal component analysis. Because CA is applicable to the most basic of data types, namely categorical data, it can be applied to almost any multivariate data set, thanks to ingenious ways of recoding data to categorical scales. But the prime example, where all its properties make perfect sense, is that of bivariate categorical data in the form of a contingency table, and its generalization to multivariate categorical data.

This talk will include several emblematic applications of CA as well as personal musings on the future of what can be called "the joy of CA".

Keywords Multivariate analysis; dimension reduction; categorical data; contingency tables;

Michael Greenacre

Universitat Pompeu Fabra, Spain, e-mail: michael.greenacre@upf.edu