Accepted Manuscript

Multivariate methods using mixtures: Correspondence analysis, scaling and pattern-detection

Shirley Pledger, Richard Arnold

PII: S0167-9473(13)00186-2
DOI: http://dx.doi.org/10.1016/j.csda.2013.05.013
Reference: COMSTA 5595

To appear in: *Computational Statistics and Data Analysis*

Received date: 21 June 2012
Revised date: 15 May 2013
Accepted date: 15 May 2013

Please cite this article as: Pledger, S., Arnold, R., Multivariate methods using mixtures: Correspondence analysis, scaling and pattern-detection. *Computational Statistics and Data Analysis* (2013), http://dx.doi.org/10.1016/j.csda.2013.05.013

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Multivariate methods using mixtures: correspondence analysis, scaling and pattern-detection

Shirley Pledger[a,1,*], Richard Arnold[a]

[a]*School of Mathematics, Statistics and Operations Research, Victoria University of Wellington, Wellington, New Zealand.*

Abstract

Matrices of binary or count data are modelled under a unified statistical framework using finite mixtures to group the rows and/or columns. This likelihood-based one-mode and two-mode fuzzy clustering provides maximum likelihood estimation of parameters and the options of using likelihood ratio tests or information criteria for model comparison. Geometric developments focused on pattern detection give likelihood-based analogues of various techniques in multivariate analysis, including multidimensional scaling, association analysis, ordination, correspondence analysis, and the construction of biplots. Illustrative examples demonstrate the effectiveness of these visualisations for identifying patterns of ecological significance (e.g. abrupt versus slow species turnover).

Keywords: Association analysis, Biclustering, Biplots, Cluster analysis, Correspondence analysis, Data visualisation, Dimension reduction, Finite mixture, Fuzzy clustering, Multidimensional scaling.

1. Introduction

There is widespread use of matrices of binary and count data across many disciplines, for example incidence and abundance matrices in ecological communities where the rows are species and the columns are samples, or binary item response analysis with respondents in the rows and questions in the columns. Major objectives include summarising the multivariate

*Corresponding author

Email address: shirley.pledger@vuw.ac.nz (Shirley Pledger)

[1]Tel +64 4 4636788, fax +64 4 4635045.

Download English Version:

https://daneshyari.com/en/article/6870354

Download Persian Version:

https://daneshyari.com/article/6870354

Daneshyari.com