Multiple and Joint Correspondence Analysis: Testing the True Dimension of a Study

Sergio Camiz
Dipartimento di Matematica, Sapienza Università di Roma
E-mail: sergio.camiz@uniroma1.it,
and Gastão Coelho Gomes
Departamento de Métodos Estatísticos, Universidade Federal do Rio de Janeiro
E-mail: gastao@im.ufrj.br

Abstract

The problem of the proper dimension of a Multiple Correspondence Analysis (*MCA*) is discussed, based on both the re-evaluation of the explained inertia sensu Benzécri (1979) and Greenacre (2006) and a test proposed by Ben Ammou and Saporta (1998). This leads to the consideration of a better reconstruction of the off-diagonal sub-tables of the Burt's table crossing the nominal characters taken into the account. Thus, Greenacre (1988) Joint Correspondence Analysis (*JCA*) is introduced and the results obtained on two applications are shown. The quality of reconstruction of both *MCA* and *JCA* solutions are compared to the Simple Correspondence Analysis results of the two-way tables. It results that *JCA*'s reduced-dimensional reconstruction is much better than the *MCA*'s one, that reveals highly biased and non-monotonous. Keywords: Correspondence Analysis, Multiple

Correspondence Analysis, Joint Correspondence Analysis.

Résumé

On discute le problème de la dimension d'une Analyse des Correspondances Multiples, basé soit sur la ré-evaluation de l'inertie expliquée sensu Benzécri (1979) et Greenacre (2006) et le test proposé par Ben Ammou and Saporta (1998). Ceci conduit à la considération d'une meilleure reconstruction des sous-tableaux hors-diagonale du tableau de Burt qui croise les charactères nominaux considérés. On introduit donc l'Analyse des Correspondances Conjointe (*JCA*, Greenacre, 1988) et on montre les résultats obtenu dans deux applications. On compare aussi la qualité de la reconstruction obtenue par les solutions *MCA* et *JCA* avec les résultats de l'Analyse des Correspondances Simples sur les tableaux à deux voies. l résulte que la reconstruction de dimensione réduite de la *JCA* est fort meilleure de celle de la *MCA*, qui s'avère fort biaisée et non-monotone. Mots-clés:

Analyse des Correspondances Simple, Analyse des Correspondances Multiples, Analyse des Correspondances Conjointe.

1 Introduction

The identification of the dimension of a data table under study is a crucial issue of most multidimensional scaling techniques. A distinction should be done between linear scaling, in which the encapsulated solutions allows an *a posteriori* choice of the user, and non-linear one, in which usually the solution dimension is an *a priori* choice that conditions the results. As the latter may be only hypothesized, e.g. according to the results of a previous linear scaling that may be used as a starting configuration, the identification in the linear case has an importance that goes beyond the simple linear case, to involve most of the analysis that follow the scaling