

中央研究院統計科學研究所

學術演講

講題：Automated learning of mixtures of factor analyzers with missing values

演講人：Prof. Tsung-I Lin (林宗儀 教授)

Institute of Statistics, National Chung Hsing University,
Taichung, Taiwan (中興大學統計學研究所)

時間：2020年1月6日 (星期一) 上午10:30-12:00

地點：中央研究院統計科學研究所6005會議室(環境變遷研究大樓A棟)

※茶會：上午10:10開始

Abstract

The mixture of factor analyzers (MFA) model has emerged as a useful tool to perform dimensionality reduction and model-based clustering of heterogeneous data. In seeking the most appropriate number of factors (q) of a MFA model with the number of components (g) fixed a priori, a two-stage procedure is commonly implemented by firstly carrying out parameter estimation over a set of prespecified numbers of factors, and then selecting the best q according to certain penalized likelihood criteria. When the dimensionality of data grows higher, such a procedure can be computationally prohibitive. To overcome this obstacle, we develop an automated learning scheme, called the automated MFA (AMFA) algorithm, to effectively merge parameter estimation and selection of q into a one-stage algorithm. The proposed AMFA procedure that allows for much lower computational cost is also extended to accommodate missing values. Moreover, we explicitly derive the score vector and the empirical information matrix for calculating standard errors associated with the estimated parameters. The potential and applicability of the proposed method are demonstrated through a number of real datasets with genuine and synthetic missing values.

Keywords: automated learning; factor analysis; maximum likelihood estimation; missing values; model selection; one-stage algorithm

中央研究院
統計科學研究所