Annals of the Institute of Statistical Mathematics (2024) 76:1–33
https://doi.org/10.1007/s10463-023-00884-4

INVITED ARTICLE: FOURTH AKAIKE MEMORIAL LECTURE

# Identifiability of latent-variable and structural-equation models: from linear to nonlinear

**Aapo Hyvärinen**[1] · **Ilyes Khemakhem**[2] · **Ricardo Monti**[2]

Received: 27 January 2023 / Revised: 4 April 2023 / Accepted: 24 April 2023 /
Published online: 4 November 2023


## Abstract

An old problem in multivariate statistics is that linear Gaussian models are often unidentifiable. In factor analysis, an orthogonal rotation of the factors is unidentifiable, while in linear regression, the direction of effect cannot be identified. For such linear models, non-Gaussianity of the (latent) variables has been shown to provide identifiability. In the case of factor analysis, this leads to independent component analysis, while in the case of the direction of effect, non-Gaussian versions of structural equation modeling solve the problem. More recently, we have shown how even general nonparametric nonlinear versions of such models can be estimated. Non-Gaussianity is not enough in this case, but assuming we have time series, or that the distributions are suitably modulated by observed auxiliary variables, the models are identifiable. This paper reviews the identifiability theory for the linear and nonlinear cases, considering both factor analytic and structural equation models.

**Keywords** Identifiability · Independent component analysis · Structural equation model · Factor analysis · Disentanglement · Non-Gaussianity

The Related Articles are https://doi.org/10.1007/s10463-023-00885-3; https://doi.org/10.1007/s10463-023-00886-2; https://doi.org/10.1007/s10463-023-00887-1.

✉ Aapo Hyvärinen
aapo.hyvarinen@helsinki.fi

[1] Department of Computer Science, University of Helsinki, Pietari Kalmin katu 5, 00560 Helsinki, Finland

[2] Gatsby Computational Neuroscience Unit, UCL, 25 Howland Street, London, W1T 4JG, UK