Linear Algebra and its Applications 473 (2015) 93–113

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

On the causal interpretation of acyclic mixed graphs under multivariate normality

Christopher J. Fox [a], Andreas Käufl [b], Mathias Drton [c,*]

[a] Department of Statistics, The University of Chicago, Chicago, IL, USA
[b] Institute for Mathematics, University of Augsburg, Augsburg, Germany
[c] Department of Statistics, University of Washington, Seattle, WA, USA

ARTICLE INFO

Article history:
Received 23 September 2013
Accepted 17 February 2014
Available online 11 March 2014
Submitted by L.-H. Lim

MSC:
62H05

Keywords:
Covariance matrix
Graphical model
Hidden variable
Multivariate normal distribution
Latent variable
Structural equation model

ABSTRACT

In multivariate statistics, acyclic mixed graphs with directed and bidirected edges are widely used for compact representation of dependence structures that can arise in the presence of hidden (i.e., latent or unobserved) variables. Indeed, under multivariate normality, every mixed graph corresponds to a set of covariance matrices that contains as a full-dimensional subset the covariance matrices associated with a causally interpretable acyclic digraph. This digraph generally has some of its nodes corresponding to hidden variables. We seek to clarify for which mixed graphs there exists an acyclic digraph whose hidden variable model coincides with the mixed graph model. Restricting to the tractable setting of chain graphs and multivariate normality, we show that decomposability of the bidirected part of the chain graph is necessary and sufficient for equality between the mixed graph model and some hidden variable model given by an acyclic digraph.

* Corresponding author.
E-mail addresses: chrisfox@uchicago.edu (C.J. Fox), andreas.kaeufl@googlemail.com (A. Käufl), md5@uw.edu (M. Drton).

http://dx.doi.org/10.1016/j.laa.2014.02.032

1. Introduction

Acyclic digraphs are standard representations of causally interpretable statistical models in which the involved random variables are noisy functions of each other, with all noise terms being independent. In generalization, acyclic mixed graphs with directed and bidirected edges are widely used for compact representation of causal structure when important variables are hidden (that is, unobserved) [12,15,18,20,23]. Such mixed graphs are also known as path diagrams in the field of structural equation modeling [2,11]. The graphs provide, in particular, a framework for statistical model selection in the presence of hidden variables [3,19,20].

Under joint multivariate normality, it is well-known that for every acyclic mixed graph there exists an acyclic digraph, generally with additional vertices, such that the statistical model associated with the digraph is a full-dimensional subset of the model determined by the mixed graph. Here, nodes that appear in the digraph but not in the mixed graph are treated as hidden variables and marginalized over. In this paper we ask which mixed graphs induce a statistical model that is not only a superset but equal to a hidden variable model given by some acyclic digraph. We focus on the particularly tractable class of chain graphs, that is, mixed graphs without semi-directed cycles. Our main result characterizes the chain graphs for which there exists an acyclic digraph with hidden variable model equal to the chain graph model. We begin by formally introducing the concerned graphical models and stating the precise form of the problem and main result.

Let $\mathcal{D} = (V, E)$ be an acyclic digraph with finite vertex set V and edge set $E \subseteq V \times V$. We denote possible edges (u, v) by $u \to v$. Let $\mathbb{R}^E$ be the set of matrices $\Lambda = (\lambda_{uv}) \in \mathbb{R}^{V\times V}$ that are supported on E, that is, $\lambda_{uv} = 0$ if $u \to v \notin E$ or $u = v$. For any $\Lambda \in \mathbb{R}^E$, the matrix $I - \Lambda$ is invertible because $\det(I - \Lambda) = 1$. Throughout, I denotes the identity matrix with size determined by the context.

Definition 1.1. The *Gaussian graphical model* $\mathbf{N}(\mathcal{D})$ is the family of all multivariate normal distributions $\mathcal{N}(\mu, \Sigma)$ on $\mathbb{R}^V$ that have covariance matrix

$$\Sigma = (I - \Lambda)^{-T} \Omega (I - \Lambda)^{-1} \tag{1.1}$$

with $\Lambda \in \mathbb{R}^E$ and $\Omega \in \mathbb{R}^{V\times V}$ diagonal with positive diagonal entries.

The motivation for consideration of the model $\mathbf{N}(\mathcal{D})$ becomes clearer through the following construction. Let $\epsilon = (\epsilon_v)_{v\in V}$ be a multivariate normal random vector with covariance matrix $\Omega = (\omega_{uv})$, and let $\mathrm{pa}(v) = \{u : u \to v \in E\}$ be the set of *parents* of vertex v. Define the random vector $X = (X_v)_{v\in V}$ to be the solution of the linear equation system

$$X_v = \lambda_{0v} + \sum_{u\in \mathrm{pa}(v)} \lambda_{uv} X_u + \epsilon_v, \quad v \in V. \tag{1.2}$$