ARTICLE IN PRESS
JID: ECOSTA [m3Gsc;March 13, 2018;5:36]

Econometrics and Statistics 000 (2018) 1–19

Contents lists available at ScienceDirect

Econometrics and Statistics

journal homepage: www.elsevier.com/locate/ecosta

Tail dependence of recursive max-linear models with regularly varying noise variables

Nadine Gissibl, Claudia Klüppelberg*, Moritz Otto

Center for Mathematical Sciences, Technical University of Munich, Boltzmannstrasse 3, Garching 85748, Germany

A R T I C L E I N F O

Article history:
Received 22 January 2017
Revised 14 February 2018
Accepted 14 February 2018
Available online xxx

Keywords:
Causal inference
Directed acyclic graph
Extreme value theory
Graphical model
Max-linear model
Max-stable model
Regular variation
Structural equation model
Tail dependence coefficient

A B S T R A C T

Recursive max-linear structural equation models with regularly varying noise variables are considered. Their causal structure is represented by a directed acyclic graph (DAG). The problem of identifying a recursive max-linear model and its associated DAG from its matrix of pairwise tail dependence coefficients is discussed. For example, it is shown that if a causal ordering of the associated DAG is additionally known, then the minimum DAG representing the recursive structural equations can be recovered from the tail dependence matrix. For a relevant subclass of recursive max-linear models, identifiability of the associated minimum DAG from the tail dependence matrix and the initial nodes is shown. Algorithms find the associated minimum DAG for the different situations. Furthermore, given a tail dependence matrix, an algorithm outputs all compatible recursive max-linear models and their associated minimum DAGs.

© 2018 EcoSta Econometrics and Statistics. Published by Elsevier B.V. All rights reserved.

1. Introduction

Causal inference is fundamental in virtually all areas of science. Examples for concepts established over the last years to understand causal inference include structural equation modeling (see e.g. Bollen, 1989; Pearl, 2009) and graphical modeling (see e.g. Lauritzen, 1996; Spirtes et al., 2000; Koller and Friedman, 2009).

In extreme risk analysis it is especially important to understand causal dependencies. We consider *recursive max-linear models* (RMLMs), which are max-linear structural equation models whose causal structure is represented by a *directed acyclic graph* (DAG). Such models are directed graphical models (Pearl, 2009, Theorem 1.4.1); i.e., the DAG encodes conditional independence relations in the distribution via the (directed global) Markov property. RMLMs were introduced and studied in Gissibl and Klüppelberg (2018). They may find their application in situations when extreme risks play an essential role and may propagate through a network, for example, when modeling water levels or pollution concentrations in a river or when modeling risks in a large industrial structure. In Einmahl et al. (2016), a RMLM was fitted to data from the EURO STOXX 50 Index, where the DAG structure was assumed to be known.

In this paper, we assume *regularly varying* noise variables. This leads to models treated in classical multivariate extreme value theory. The books by Beirlant et al. (2004), de Haan and Ferreira (2006), and Resnick (1987, 2007) provide a detailed introduction into this field. A RMLM with regularly varying noise variables is in the *maximum domain of attraction* of an *extreme value* (*max-stable*) *distribution*. The spectral measure of the limit distribution, which describes the dependence

* Corresponding author.
E-mail addresses: n.gissibl@tum.de (N. Gissibl), cklu@tum.de (C. Klüppelberg), moritz.otto@kit.edu (M. Otto).

https://doi.org/10.1016/j.ecosta.2018.02.003
2452-3062/© 2018 EcoSta Econometrics and Statistics. Published by Elsevier B.V. All rights reserved.

Please cite this article as: N. Gissibl et al., Tail dependence of recursive max-linear models with regularly varying noise variables, Econometrics and Statistics (2018), https://doi.org/10.1016/j.ecosta.2018.02.003

structure given by the DAG, is discrete. Every max-stable random vector with discrete spectral measure is *max-linear* (ML), and every multivariate max-stable distribution can be approximated arbitrarily well via a ML model (e.g. Yuen and Stoev, 2014, Section 2.2). This demonstrates the important role of ML models in extreme value theory. They have been investigated, generalized, and applied to real world problems by many researchers; see e.g. Schlather and Tawn (2002), Wang and Stoev (2011), Falk et al. (2015), Strokorb and Schlather (2015), Einmahl et al. (2012), Cui and Zhang (2018), and Kiriliouk (2017).

One main research problem that is addressed for restricted recursive structural equation models, where the functions are required to belong to a specified function class, is the *identifiability* of the coefficients and the DAG from the observational distribution. Recently, particular attention in this context has been given to recursive structural equation models with additive Gaussian noise; see e.g. Peters et al. (2014), Ernest et al. (2018), and references therein. For RMLMs this problem is investigated in Gissibl et al. (2018). In the present paper, we discuss the identifiability of RMLMs from their *(upper) tail dependence coefficients* (TDCs).

The TDC, which goes back to Sibuya (1960), measures the extremal dependence between two random variables and is a simple and popular dependence measure in extreme value theory. Methods to construct multivariate max-stable distributions with given TDCs have been proposed, for example, by Schlather and Tawn (2002), Falk (2005), Falk et al. (2015), and Strokorb and Schlather (2015). Somehow related we identify all RMLMs with the same given TDCs.

1.1. *Problem description and important concepts*

First we briefly review RMLMs and introduce the TDC formally. We then describe the idea of this work in more detail and state the main results.

Max-linear models on DAGs

Consider a *RMLM* $\boldsymbol{X} = (X_1, \dots, X_d)$ *on a DAG* $\mathcal{D} = (V, E)$ with nodes $V = \{1, \dots, d\}$ and edges $E = \{(k, i) : i \in V \text{ and } k \in \text{pa}(i)\}$:

$$X_i = \bigvee_{k \in \text{pa}(i)} c_{ki} X_k \vee c_{ii} Z_i, \quad i = 1, \dots, d, \tag{1}$$

where pa(i) denotes the parents of node i in $\mathcal{D}$ and $c_{ki} > 0$ for $k \in \text{pa}(i) \cup \{i\}$; the noise variables $Z_1, \dots, Z_d$, represented by a generic random variable Z, are assumed to be independent and identically distributed with support $\mathbb{R}_+ := (0, \infty)$ and *regularly varying* with index $\alpha \in \mathbb{R}_+$, abbreviated by $Z \in \text{RV}(\alpha)$. Denoting the distribution function of Z by F_Z, the latter means that

$$\lim_{t \to \infty} \frac{1 - F_Z(xt)}{1 - F_Z(t)} = x^{-\alpha}$$

for every $x \in \mathbb{R}_+$. Examples for F_Z include Cauchy, Pareto, and log-gamma distributions. For details and background on regular variation, see e.g. Resnick (1987, 2007).

The properties of the noise variables imply the existence of a normalizing sequence $a_n \in \mathbb{R}_+$ such that for independent copies $\boldsymbol{X}^{(1)}, \dots, \boldsymbol{X}^{(n)}$ of $\boldsymbol{X}$,

$$a_n^{-1} \bigvee_{\nu=1}^{n} \boldsymbol{X}^{(\nu)} \xrightarrow{d} \boldsymbol{M}, \quad n \to \infty, \tag{2}$$

where $\boldsymbol{M}$ is a non-degenerate random vector with distribution function denoted by G and all operations are taken componentwise. Thus $\boldsymbol{X}$ is in the maximum domain of attraction of G; we write $\boldsymbol{X} \in \text{MDA}(G)$. The limit vector $\boldsymbol{M}$ (its distribution function G) is necessarily max-stable: in the present situation we have for all $n \in \mathbb{N}$ and independent copies $\boldsymbol{M}^{(1)}, \dots, \boldsymbol{M}^{(n)}$ of $\boldsymbol{M}$, the distributional equality $n^{1/\alpha} \boldsymbol{M} \stackrel{d}{=} \bigvee_{\nu=1}^{n} \boldsymbol{M}^{(\nu)}$. Furthermore, $\boldsymbol{M}$ is again a RMLM on $\mathcal{D}$, with the same weights in (1) as $\boldsymbol{X}$ and standard α-Fréchet distributed noise variables, i.e.,

$$F_Z(x) = \Phi_\alpha(x) = \exp\{-x^{-\alpha}\}, \quad x \in \mathbb{R}_+.$$

A proof of (2) as well as an explicit formula for G and its univariate and bivariate marginal distributions can be found in Appendix A.2, Proposition A.2.

In what follows we summarize the most important properties of $\boldsymbol{X}$ presented in Gissibl and Klüppelberg (2018) which are needed throughout the paper. Every component of $\boldsymbol{X}$ can be written as a max-linear function of its ancestral noise variables:

$$X_i = \bigvee_{j \in \text{An}(i)} b_{ji} Z_j, \quad i = 1, \dots, d, \tag{3}$$

where An(i) = an$(i) \cup \{i\}$ and an(i) are the ancestors of i in $\mathcal{D}$ (Gissibl and Klüppelberg, 2018, Theorem 2.2). For $i \in V$, $b_{ii} = c_{ii}$. For $j \in \text{an}(i)$, b_{ji} can be determined by a path analysis of $\mathcal{D}$ as explained in the following. Throughout we write $k \to i$ whenever $\mathcal{D}$ has an edge from k to i. With every path $p = [j = k_0 \to k_1 \to \cdots \to k_n = i]$ we associate a weight, which we define to be

Please cite this article as: N. Gissibl et al., Tail dependence of recursive max-linear models with regularly varying noise variables, Econometrics and Statistics (2018), https://doi.org/10.1016/j.ecosta.2018.02.003

Download English Version:

https://daneshyari.com/en/article/8919479

Download Persian Version:

https://daneshyari.com/article/8919479

Daneshyari.com