Digital Signal Processing 47 (2015) 5–16

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Dynamic network signal processing using latent threshold models

Jouchi Nakajima [1], Mike West *

Duke University, USA

ARTICLE INFO

Article history:
Available online 21 April 2015

Keywords:
Dynamic graphical models
EEG signal processing
Financial time series
Multivariate volatility
Sparse time-varying regression

ABSTRACT

We discuss multivariate time series signal processing that exploits a recently introduced approach to dynamic sparsity modelling based on latent thresholding. This methodology induces time-varying patterns of zeros in state parameters that define both directed and undirected associations between individual time series, so generating statistical representations of the dynamic network relationships among the series. Following an overview of model contexts and Bayesian analysis for dynamic latent thresholding, we exemplify the approach in two studies: one of foreign currency exchange rate (FX) signal processing, and one in evaluating dynamics in multiple electroencephalography (EEG) signals. These studies exemplify the utility of dynamic latent threshold modelling in revealing interpretable, data-driven dynamics in patterns of network relationships in multivariate time series.

1. Introduction

Dynamic network structures generating inter-related time series arise in many scientific fields—from neuroscience, to engineering signal processing, to financial econometrics, and others. For statistical definition and estimation of empirical networks, vector autoregressive (VAR) models have become standard tools (e.g. [37,38]). A typical approach examines the relevance of VAR coefficients, linking to the idea of Granger [10] causality to suggest and quantify *feed-forward* network connections. Time-varying VAR (TV-VAR) models refine this, in allowing for temporal changes in the strengths of such relationships (e.g. [14,30,9,26–28,15]). Linked to this, statistical graphical models are becoming increasingly popular in defining models of potentially sparse *contemporaneous* network associations induced by patterns of conditional independence [47,16,7,12,5,44].

This paper overviews a framework of dynamic sparsity in empirical network relationships using the latent threshold modelling (LTM) concept. Introduced in Nakajima and West [19], the LTM approach defines a class of parametrized models for thresholding state parameter processes in broad classes of multivariate time series models. Time- and data-adaptive thresholding can induce patterns of temporal sparsity in the resulting, practically effective state processes. We overlay the LTM method on TV-VAR models with Cholesky-style multivariate stochastic volatility (MSV) components (e.g. [18]). We denote the overall model class by TV-VAR-MSV. The time-varying VAR aspects flexibly represent and quantify dynamics in the *feed-forward* relationships among series, while the time-varying multivariate stochastic volatility elements represent and quantify dynamics in *contemporaneous* relationships.

The LTM concept allows for, and induces, dynamic patterns of sparsity in both feed-forward and contemporaneous networks linking the series. The model structure also allows decoupling of model fitting to a set of parallel but linked univariate dynamic models; computation scales only linearly with time series dimension as a result.

Following the introductory Section 1, we detail TV-VAR-MSV models and review the LTM ideas in Section 2. Section 3 outlines Bayesian analysis and computation for model fitting. Two applied studies follow: Section 4 presents an econometric finance example, exploring dynamic dependencies in international financial markets; Section 5 provides an example of dynamic network modelling of connectivities among multiple EEG signals from a neuropsychiatric study. Section 6 provides concluding comments.

Some notation: We denote vectors and matrices via bold font. We use the distributional notation $\boldsymbol{y} \sim N(\boldsymbol{m}, \boldsymbol{V})$, $d \sim U(a,b)$, $p \sim B(a,b)$, $v \sim G(a,b)$, for the multivariate normal, uniform, beta, and gamma distributions, respectively. We also use $s:t$ to denote $s, s+1, \ldots, t$ when $s < t$, for succinct subscripting; e.g., $\boldsymbol{y}_{1:T}$ denotes $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T\}$.

* Corresponding author at: The Arts & Sciences Professor of Statistics & Decision Sciences, Department of Statistical Science, Duke University, Durham 27708-0251, USA.

E-mail addresses: jouchi.nakajima@boj.or.jp (J. Nakajima), mw@stat.duke.edu (M. West).

[1] *Present position and address*: Deputy Director, Monetary Affairs Department, Bank of Japan, 2-1-1 Nihonbashi-Hongokucho, Chuo-ku, Tokyo 103-0021, Japan.

http://dx.doi.org/10.1016/j.dsp.2015.04.008
1051-2004/

2. Latent threshold TV-VAR-MSV models

2.1. TV-VAR modelling

For the $m \times 1$-vector time series $\boldsymbol{y}_t$, $(t = 1, 2, \ldots)$, consider the TV-VAR(p) model

$$\boldsymbol{y}_t = \sum_{j=1}^{p} \boldsymbol{\Gamma}_{jt} \boldsymbol{y}_{t-j} + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_t), \tag{1}$$

where $\boldsymbol{\Gamma}_{jt}$ is the $m \times m$ matrix of time-varying coefficients at lag j, $(j = 1 : p)$, and $\boldsymbol{\Sigma}_t$ the variance matrix of the time t innovations vector $\boldsymbol{u}_t$. The model can easily be extended to include time-varying intercepts and dynamic regressions on other exogenous predictors known at time t, but that is not of main interest here.

Denote by $\boldsymbol{\Omega}_t$ the time-varying *precision* matrix $\boldsymbol{\Omega}_t = \boldsymbol{\Sigma}_t^{-1}$. Using Cholesky decomposition we can write

$$\boldsymbol{\Sigma}_t = \boldsymbol{A}_t^{-1} \boldsymbol{\Lambda}_t (\boldsymbol{A}_t')^{-1}, \quad \text{equivalently} \quad \boldsymbol{\Omega}_t = \boldsymbol{A}_t' \boldsymbol{\Lambda}_t^{-1} \boldsymbol{A}_t, \tag{2}$$

with

$$\boldsymbol{A}_t = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -a_{21,t} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ -a_{m1,t} & \cdots & -a_{m,m-1,t} & 1 \end{pmatrix} \quad \text{and}$$

$$\boldsymbol{\Lambda}_t = \begin{pmatrix} \lambda_{1t} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_{mt} \end{pmatrix}, \tag{3}$$

where the a_{ijt} elements are real-valued and the λ_{jt} are positive. Decompositions of this form are increasingly popular in multiple time series modelling (e.g. [23,33,30,18,19]). The decomposition allows for flexibility in modelling volatility matrices over time through models on the elements a_{ijt}, λ_{jt}, while also enabling efficient computation, as we detail further below.

Relating to graphical models [47,16,12], a zero in the i, j off-diagonal element of $\boldsymbol{\Omega}_t$ relates to conditional independence of the corresponding innovation elements u_{it}, u_{jt} in $\boldsymbol{u}_t$. The undirected graph on m nodes representing the m scalar elements of $\boldsymbol{u}_t$ has edges between only those pairs of nodes that are conditionally dependent, with the missing edges corresponding to node pairs with zero precision elements. Increased levels of conditional independence are key idea to reducing parameter dimension, inducing parsimonious structure in $\boldsymbol{\Omega}_t$, and hence $\boldsymbol{\Sigma}_t$, as a result. Any pattern of off-diagonal zeros then represents the network of contemporaneous dependencies at time t. Note that many zeros among the elements a_{ijt} of the lower-triangular matrix $\boldsymbol{A}_t$ can induce off-diagonal zeros in $\boldsymbol{\Omega}_t$, i.e. sparsity of $\boldsymbol{A}_t$ can lead to a sparse set of contemporaneous network connectivities.

2.2. Decoupled system of dynamic regressions

From Eqs. (2), (3) we see that

$$\boldsymbol{A}_t \boldsymbol{y}_t = \sum_{j=1}^{p} \boldsymbol{B}_{jt} \boldsymbol{y}_{t-j} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim N(\boldsymbol{0}, \boldsymbol{\Lambda}_t), \tag{4}$$

where: $\boldsymbol{B}_{jt} = \boldsymbol{A}_t \boldsymbol{\Gamma}_{jt}$, $(j = 1 : p)$, and $\boldsymbol{\varepsilon}_t = \boldsymbol{A}_t \boldsymbol{u}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{mt})'$ has independent elements $\varepsilon_{it} \sim N(0, \lambda_{it})$, $(i = 1 : m)$. The model can then be recast as the triangular system of univariate dynamic regressions

$$\begin{aligned} y_{1t} &= \boldsymbol{z}_{t-1}' \boldsymbol{a}_{1t} + \varepsilon_{1t}, \\ y_{2t} &= \boldsymbol{z}_{t-1}' \boldsymbol{a}_{2t} + a_{21,t} y_{1t} + \varepsilon_{2t}, \\ y_{3t} &= \boldsymbol{z}_{t-1}' \boldsymbol{a}_{3t} + a_{31,t} y_{1t} + a_{32,t} y_{2t} + \varepsilon_{3t}, \\ &\vdots \\ y_{mt} &= \boldsymbol{z}_{t-1}' \boldsymbol{a}_{mt} + a_{m1,t} y_{1t} + \cdots + a_{m,m-1,t} y_{m-1,t} + \varepsilon_{mt}, \end{aligned}$$

where $\boldsymbol{z}_{t-1} = (\boldsymbol{y}_{t-1}', \ldots, \boldsymbol{y}_{t-p}')'$ and, for each $i = 1 : m$, the $\boldsymbol{a}_{it}$ is a $mp \times 1$ vector formed by vertically stacking the transposed i-th rows of each of $\boldsymbol{B}_{1t}, \boldsymbol{B}_{2t}, \ldots, \boldsymbol{B}_{pt}$ in that order.

Finally, define $\boldsymbol{x}_{1t} = \boldsymbol{z}_{t-1}$, $\boldsymbol{b}_{1t} = \boldsymbol{a}_{1t}$, and, for $i = 2 : m$, $\boldsymbol{x}_{it} = (\boldsymbol{z}_{t-1}', y_{1t}, \ldots, y_{i-1,t})'$ and $\boldsymbol{b}_{it} = (\boldsymbol{a}_{it}', a_{i1,t}, \ldots, a_{i,i-1,t})'$. Then the model notation is simplified as

$$y_{it} = \boldsymbol{x}_{it}' \boldsymbol{b}_{it} + \varepsilon_{it} \quad \text{with} \quad \varepsilon_{it} \sim N(0, \lambda_{it}), \qquad (i = 1 : m), \tag{5}$$

with noise terms ε_{it} that are independent across $i = 1 : m$.

The system of Eqs. (5) is the model representation adopted. For each i, note that: (i) the leading mp elements of $\boldsymbol{b}_{it}$ relate the time t univariate signal y_{it} to the previous p lagged signals on all series, so defining feed-forward network structure; while (ii) for signals $i = 2 : m$, the last $i - 1$ elements of $\boldsymbol{b}_{it}$ relate the y_{it} to the contemporaneous values of other series y_{ht} for $h < i$; together these define the elements of $\boldsymbol{A}_t$ and hence, when coupled with $\boldsymbol{\Lambda}_t$, we recover the full multivariate volatility matrix $\boldsymbol{\Sigma}_t$. That is, this formulation coverts the goal of modelling multivariate volatility structure $\boldsymbol{\Sigma}_t$ over time to that of a collection of dynamic regression vectors $\boldsymbol{b}_{it}$ and scalar volatilities λ_{it}. This transformed representation of MSV structure is also then key to decoupling the analyses. Under dynamic model structures and priors for each $(\boldsymbol{b}_{it}, \lambda_{it})$ that are independent over $i = 1 : m$, Bayesian analysis reduces to computations that are conditionally independent across series i and can be implemented in parallel.

The Cholesky-style representation of MSV structures that the general formulation here represents has become increasingly popular as a general approach due to (i) the decoupling that yields computational tractability, and (ii) the opportunity to model the resulting components $(\boldsymbol{b}_{it}, \lambda_{it})$ in various ways. On the latter point, the entire focus of our work here is to apply latent threshold models to the $\boldsymbol{b}_{it}$ and one of two standard univariate volatility models to the λ_{it}, as detailed below. Connections with prior work on multivariate volatility include well-known latent factor models that structure $\boldsymbol{\Sigma}_t$ via various forms of latent factor representation (e.g. [1,24,2,6,17,48]). There are similarities with our framework, in that Eq. (3) is a form of (full-rank) factor decomposition in which $\boldsymbol{A}_t$ is the inverse of a time-varying factor loadings matrix. More directly and practically, we note that we can explicitly include latent factor structure in more elaborate LTMs by adding latent factor components to the right-hand side of Eq. (4) directly [20].

2.3. Latent threshold structure

The concept of using Bayesian variable selection methods to induce zeros in TV-VAR state parameters has become of interest in a number of literatures, especially signal processing in time series econometrics (e.g. [15]). Coupled with this is the now traditional use of Bayesian graphical modelling to induce zeros in precision matrices of innovations in TV-VAR and other dynamic models (e.g. [5,41,40]). The concept and resulting methodology of latent threshold modelling is a very general, and widely applicable strategy that permits the *existence* of relationships to vary over time, i.e., allowing time- and data-adaptive dynamics in sparsity patterns in all model components within an overall model structure.

We follow Nakajima and West [19] in applying dynamic latent thresholding to the model state vectors $\boldsymbol{b}_{1:m,t}$ in Eq. (5). We use the simplest, practicable state models, based on underlying latent autoregressive models of order one, or AR(1). For each