Bayesian Modeling Strategies for Multivariate Non-Gaussian Time Series Data

Refik Soyer

The George Washington University
Department of Decision Sciences

UARK, Fayetteville
April 20, 2019

Joint work with:
Tevfik Aktekin, UNH, and Nicholas Polson, UChicago

Outline

1. Background: Modeling temporal and (contemporary) component correlations.
2. A family of multivariate state space models: Dynamic multivariate distributions.
3. Modeling multivariate durations and multivariate counts.
4. Bayesian inference: MCMC and PL.
5. Numerical illustrations.
6. Concluding remarks.

Earlier Bayesian work goes back to early 1990s.

- Grunwald, Raftery and Guttorp (1993, JRSSB): Dirichlet time series of proportions.
- Ord, Fernandes and Harvey (1993): Multivariate counts.
- Cargnoni, Muller and West (1997, JASA): Multinomial time series.

Recent interest in discrete valued time series and count data:

Handbook of Discrete Valued Time Series by Davis et al. (2015).

Modeling Correlations in Multivariate Time Series

- Modeling temporal correlations.

  Cox (1981) classifies time series models into observation and parameter driven processes.

- Modeling contemporaneous correlations.

  Common strategies: Marshall and Olkin (1988, JASA)

  1. Modeling via conditionals Arnold et al. (1992, 2001 StatSci).
  2. Common environment models: Generation of dependence via mixtures [Arbous and Kerrich (1951, Biometrics) and Lindley and Singpurwalla (1986, JAP)].

- Observation driven models:

  Multivariate INAR models: Pedeli and Karlis (2011, StatMod), Pedeli and Karlis (2012, JTSA).

  Multivariate Poisson Series: Ravishanker et al. (2014, StatInt), Serhiyenko et al. (2017, ASMB)

- Parameter driven (state-space) models:

  Ord et al. (1993) and Jorgensen et al. (1999, Biometrika)

  Chen et al. (2016, JASA), Berry and West (2018): "Decouple/recouple"

  Aktekin, Polson and Soyer (2018, BA): "random environment"

A Family of Multivariate State Space Models

- Consider J component multivariate time series Y_{jt}'s, $j = 1, \ldots, J$, subject to a common environment such that [Gamerman et al. (2013)]

$$p(Y_{jt}|\theta_t, \lambda_j, \boldsymbol{\nu}) = f(Y_{jt}, \lambda_j, \boldsymbol{\nu})\theta_t^{g(Y_{jt}, \boldsymbol{\nu})} \exp\{-\theta_t h(Y_{jt}, \lambda_j, \boldsymbol{\nu})\},$$

  where θ_t's, λ_j's, and $\boldsymbol{\nu}$ are three sets of model parameters and the functions $f(.), g(.)$, and $h(.)$ are specified so that we have a proper density.

- θ_t represents the effect of the common random environment on each component at time t.

- Both λ_j's and $\boldsymbol{\nu}$ represent static effects where λ_j's are component specific and $\boldsymbol{\nu}$ may include common as well as specific effects.

- We assume that, conditional on θ_t's, λ_j's, and $\boldsymbol{\nu}$, Y_{jt} are independent over time.
- Also, assume that, conditional on θ_t's, λ_j's, and $\boldsymbol{\nu}$, components Y_{jt} are independent of each other at time t.
- Thus, for $\boldsymbol{Y}_t = \{Y_{1t}, \ldots, Y_{Jt}\}$ we can obtain

$$p(\boldsymbol{Y}_t|\theta_t, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \prod_{j=1}^{J} p(Y_{jt}|\theta_t, \lambda_j, \boldsymbol{\nu})$$

and the general form can be written as

$$p(\boldsymbol{Y}_t|\theta_t, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f(\boldsymbol{Y}_t, \boldsymbol{\lambda}, \boldsymbol{\nu})\theta_t^{g(\boldsymbol{Y}_t, \boldsymbol{\nu})} \exp\{-\theta_t h(\boldsymbol{Y}_t, \boldsymbol{\lambda}, \boldsymbol{\nu})\},$$

where $\boldsymbol{\lambda} = \{\lambda_1, \ldots, \lambda_J\}$.

Evolution of θ_t's

- Environmental process $\{\theta_t\}$ follows a Markovian evolution: Bather (1965), and Smith and Miller (1986).

$$\theta_t = \frac{\theta_{t-1}}{\gamma}\epsilon_t,$$

where

$$(\epsilon_t | D^{t-1}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \sim Beta[\gamma\alpha_{t-1}, (1-\gamma)\alpha_{t-1}]$$

$0 < \gamma < 1$, and $D^t = (D^{t-1}, \boldsymbol{Y}_t)$.

- γ acts as a discount factor such that $\theta_t < \frac{\theta_{t-1}}{\gamma}$.

- It can be shown that

$$(\theta_t | \theta_{t-1}, D^{t-1}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \sim Beta[\gamma\alpha_{t-1}, (1-\gamma)\alpha_{t-1}; (0, \frac{\theta_{t-1}}{\gamma})],$$

is a scaled Beta density over $(0, \theta_{t-1}/\gamma)$

Conditional Filtering Density

- With $(\theta_{t-1}|D^{t-1}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \sim Gamma(\alpha_{t-1}, \beta_{t-1})$, the forecast distribution of θ_t can be obtained as

$$(\theta_t|D^{t-1}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \sim Gamma(\gamma\alpha_{t-1}, \gamma\beta_{t-1}).$$

- Starting at time 0 with $(\theta_0|D_0) \sim Gamma(\alpha_0, \beta_0)$, the posterior density of θ_t at time t can be obtained as

$$(\theta_t|D^t, \boldsymbol{\lambda}, \boldsymbol{\nu}) \sim Gamma(\alpha_t, \beta_t),$$

where

$$\alpha_t = \gamma\alpha_{t-1} + g(\boldsymbol{Y}_t, \boldsymbol{\nu})$$

$$\beta_t = \gamma\beta_{t-1} + h(\boldsymbol{Y}_t, \boldsymbol{\lambda}, \boldsymbol{\nu})$$

- By integrating out the environment we can obtain the distribution of $\boldsymbol{Y}_t$ given the past data and the static parameters

$$p(\boldsymbol{Y}_t|D^{t-1}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \frac{\Gamma[\gamma\alpha_{t-1} + g(\boldsymbol{Y}_t, \boldsymbol{\nu})] f(\boldsymbol{Y}_t, \boldsymbol{\lambda}, \boldsymbol{\nu})(\gamma\beta_{t-1})^{\gamma\alpha_{t-1}}}{\Gamma(\gamma\alpha_{t-1})[\gamma\beta_{t-1} + h(\boldsymbol{Y}_t, \boldsymbol{\lambda}, \boldsymbol{\nu})]^{\gamma\alpha_{t-1}+g(\boldsymbol{Y}_t,\boldsymbol{\nu})}}.$$

  which is a dynamic multivariate distribution.

- The above provides us with dynamic multivariate generalizations of known distributions such as multivariate negative binomial, multivariate Lomax, multivariate beta prime (multivariate generalized Lomax), multivariate Burr (compound Weibull).

- Consider J different Poisson time series operating in a common environment such as

$$Y_{jt} \sim Pois(\lambda_j \theta_t), \text{ for } j = 1, \dots, J$$

- $f(\boldsymbol{Y}_t, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f(\boldsymbol{Y}_t, \boldsymbol{\lambda}) = (\prod_j \frac{\lambda_j^{Y_{jt}}}{Y_{jt}!})$

  $g(\boldsymbol{Y}_t, \boldsymbol{\nu}) = g(\boldsymbol{Y}_t) = \sum_j Y_{jt}$

  $h(\boldsymbol{Y}_t, \boldsymbol{\lambda}, \boldsymbol{\nu}) = h(\boldsymbol{Y}_t, \boldsymbol{\lambda}) = \sum_j \lambda_j$.

- For $(\theta_t | D^t, \boldsymbol{\lambda}) \sim Gamma(\alpha_t, \beta_t)$, we have

$$\alpha_t = \gamma \alpha_{t-1} + \sum_j Y_{jt}.$$

$$\beta_t = \gamma \beta_{t-1} + \sum_j \lambda_j.$$

Distribution of Multivariate Counts

- The marginal distributions of Y_{jt} for any j can be obtained as a negative binomial model.

$$p(Y_{jt}|\lambda_j, D^{t-1}) = \binom{\gamma\alpha_{t-1} + Y_{jt} - 1}{Y_{jt}} \left(1 - \frac{\lambda_j}{\gamma\beta_{t-1} + \lambda_j}\right)^{\gamma\alpha_{t-1}} \left(\frac{\lambda_j}{\gamma\beta_{t-1} + \lambda_j}\right)$$

- The multivariate distribution $p(\boldsymbol{Y}_t|\boldsymbol{\lambda}, D^{t-1})$ is given by

$$\frac{\Gamma(\gamma\alpha_{t-1} + \sum_j Y_{jt})}{\Gamma(\gamma\alpha_{t-1}) \prod_j \Gamma(Y_{jt} + 1)} \prod_j \left(\frac{\lambda_j}{\gamma\beta_{t-1} + \sum_j \lambda_j}\right)^{Y_{jt}} \left(\frac{\gamma\beta_{t-1}}{\gamma\beta_{t-1} + \sum_j \lambda_j}\right)^{\gamma\alpha_{t-1}}$$

  which is a dynamic multivariate negative binomial.

- It is a dynamic version of the (bivariate) negative binomial distribution proposed by Arbous and Kerrich (1951) for modeling number of accidents.

Conditional Distributions of Y_{jt}'s

- The conditionals of Y_{jt}s will also be negative binomial type distributions with the dynamic conditional mean (or regression) of Y_{jt} given Y_{it} for $i \neq j$ is given by

$$E[Y_{jt} | Y_{it}, \lambda_i, \lambda_j, D^{t-1}] = \frac{\lambda_j(\gamma\alpha_{t-1} + Y_{it})}{(\lambda_i + \gamma\beta_{t-1})},$$

  which is linear in Y_{it}.

- The bivariate counts are positively correlated with the correlation is given by

$$Cor(Y_{it}, Y_{jt} | \boldsymbol{\lambda}, D^{t-1}) = \sqrt{\frac{\lambda_i \lambda_j}{(\lambda_i + \gamma\beta_{t-1})(\lambda_j + \gamma\beta_{t-1})}}.$$

- Consider J different Gamma time series operating in a common environment such as

$$Y_{jt} \sim Gamma(\phi_j,\ \lambda_j\theta_t),\ \text{for } j = 1, \ldots, J.$$

- $f(\boldsymbol{Y}_t, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \prod_{j=1}^{J} \frac{\lambda_j^{\phi_j} Y_{jt}^{\phi_j - 1}}{\Gamma(\phi_j)}$

  $g(\boldsymbol{Y}_t, \boldsymbol{\nu}) = \sum_j \phi_j$

  $h(\boldsymbol{Y}_t, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \sum_j \lambda_j Y_{jt}.$

- For $(\theta_t | D^t, \boldsymbol{\lambda}, \boldsymbol{\nu}) \sim Gamma(\alpha_t, \beta_t)$, we have

$$\alpha_t = \gamma\alpha_{t-1} + \sum_j \phi_j.$$

$$\beta_t = \gamma\beta_{t-1} + \sum_j \lambda_j Y_{jt}.$$

Distribution of Multivariate Durations

- The marginal distribution of Y_{jt}'s can be obtained as a scaled beta prime density.

$$p(Y_{jt}|D^{t-1},\boldsymbol{\lambda},\boldsymbol{\nu}) = \frac{\Gamma(\gamma\alpha_{t-1}+\phi_j)}{\Gamma(\phi_j)\Gamma(\gamma\alpha_{t-1})} \frac{Y_{jt}^{\phi_j-1}(\lambda_j/\gamma\beta_{t-1})^{\phi_j}}{\left(1+(\lambda_j/\gamma\beta_{t-1})Y_{jt}\right)^{\gamma\alpha_{t-1}+\sum_j \phi_j}}$$

- The multivariate distribution $p(\boldsymbol{Y}_t|\boldsymbol{\lambda},\boldsymbol{\nu},D^{t-1})$ is given by

$$\frac{\Gamma(\gamma\alpha_{t-1}+\sum_j \phi_j)}{\prod_{j=1}\Gamma(\phi_j)\Gamma(\gamma\alpha_{t-1})} \frac{\prod_j Y_{jt}^{\phi_j-1}\prod_j(\lambda_j/\gamma\beta_{t-1})^{\phi_j}}{\left(1+\sum_j(\lambda_j/\gamma\beta_{t-1})Y_{jt}\right)^{\gamma\alpha_{t-1}+\sum_j \phi_j}}$$

  is the dynamic version of the generalized multivariate Lomax (beta prime) distribution of Nayak (1987).

- Case $\phi_j = 1$ for all j, provides us with dynamic version of multivariate Lomax distributions of Lindley and Singpurwalla (1986).

- The dynamic conditional mean (or regression) of Y_{jt} given Y_{it} for $i \neq j$ is given by

$$E[Y_{jt}|Y_{it}, \boldsymbol{\lambda}, \boldsymbol{\nu}, D^{t-1}] = \frac{\phi_j(\gamma\beta_{t-1} + \lambda_i Y_{it})}{\lambda_j(\gamma\alpha_{t-1} + \phi_i - 1)},$$

which is linear in Y_{it}.

- The bivariate durations are positively correlated with the correlation is given by

$$Cor(Y_{it}, Y_{jt}|\boldsymbol{\lambda}, \boldsymbol{\nu}, D^{t-1}) = \sqrt{\frac{\phi_i\phi_j}{(\gamma\alpha_{t-1} + \phi_i - 1)(\gamma\alpha_{t-1} + \phi_j - 1)}}.$$

- For the Lomax case of $\phi_i = \phi_j = 1$ this reduces to

$$Cor(Y_{it}, Y_{jt}|\boldsymbol{\lambda}, \boldsymbol{\nu}, D^{t-1}) = \frac{1}{\gamma\alpha_{t-1}},$$

where $\gamma\alpha_{t-1} > 1$.

- Estimation can be done using MCMC (Gibbs sampler) or Particle Filtering.
- If we assume independent gamma priors for λ_j's as

$$\lambda_j \sim Gamma(a_j, b_j); j = 1, \dots, J,$$

then we can obtain

$$p(\lambda_j | \theta_1, \dots, \theta_t, \gamma, \boldsymbol{\nu}, D^t) \sim Gamma(a_{jt}, b_{jt}),$$

where $a_{jt} = a_{j,t-1} + Y_{jt}$ and $b_{jt} = b_{j,t-1} + \theta_t$ in the negative binomial and $a_{jt} = a_{j,t-1} + \phi_j$ and $b_{jt} = b_{j,t-1} + \theta_t Y_{jt}$ in the generalized Lomax cases.

- Updating of the discount parameter γ and $\boldsymbol{\nu}$ requires a Metropolis step.

- Given T multivariate observations, we can draw from the full conditional

$$p(\theta_1, \cdots, \theta_T | \boldsymbol{\lambda}, \boldsymbol{\nu}, \gamma, D^T)$$

via forward filtering backward sampling (FFBS) of Fruhwirth-Schnatter (1994).

$$p(\theta_T | \boldsymbol{\lambda}, \boldsymbol{\nu}, \gamma, D^T) \cdots p(\theta_1 | \boldsymbol{\lambda}, \boldsymbol{\nu}, \gamma, D^1)$$

- This is feasible since

$$(\theta_{t-1} | \theta_t, \boldsymbol{\lambda}, \boldsymbol{\nu}, \gamma, D^{t-1}) \sim Gamma[(1-\gamma)\alpha_{t-1}, \beta_{t-1}]$$

where $\gamma\theta_t < \theta_{t-1} < \infty$, a shifted gamma density.

Particle Filtering

- MCMC is not very attractive for on-line updating of θ_t's since it needs to be rerun for every new observation.

- Due to the availability of conditional distributions of dynamic θ_t's and the static λ_j's, we have conditional sufficient statistics which enables us to use particle learning (PL) approach of Carvalho et al. (2010, Stat. Sci.)

- Since the predictive distribution $p(\boldsymbol{Y}_{t+1}|\theta_t, \boldsymbol{\lambda}, \boldsymbol{\nu}, D^t)$ and the propogation density $p(\theta_{t+1}|\theta_t, \boldsymbol{\lambda}, \boldsymbol{\nu}, D^{t+1})$ are available, we can use the PL approach instead of APF.

- The marginal likelihood of γ and $\boldsymbol{\nu}$ is available conditional on $\boldsymbol{\lambda}$ and thus we can use discrete priors for these in the PF updating.

Particle Learning Algorithm

Assume that γ and $\boldsymbol{\nu}$ are known and define the conditional sufficient statistic $s_t = f(s_{t-1}, \theta_t, \boldsymbol{Y}_t)$ where $\boldsymbol{Y}_t = (Y_{1t}, \ldots, Y_{Jt})$ and $z_t = \{\theta_t, s_t, \boldsymbol{\lambda}\}$.

The algorithm can be summarized as:

1. (Resample) $\{z_t\}_{i=1}^{K}$ from $z_t^{(i)} = \{\theta_t, s_t, \boldsymbol{\lambda}\}^{(i)}$ using weights $w_t^{(i)} \propto p(\boldsymbol{Y}_{t+1} | z_t^{(i)})$
2. (Propagate) $\{\theta_t^{(i)}\}$ to $\{\theta_{t+1}^{(i)}\}$ via $p(\theta_{t+1} | z_t^{(i)}, \boldsymbol{Y}_{t+1})$
3. (Update) $s_{t+1}^{(i)} = f(s_t^{(i)}, \theta_{t+1}^{(i)}, \boldsymbol{Y}_{t+1})$
4. (Sample) $(\boldsymbol{\lambda})^{(i)}$ from $p(\boldsymbol{\lambda} | s_{t+1}^{(i)})$

Some Remarks on PL

1. In step 1, z_t will be stored at each point in time and it only includes one state parameter (θ_t), hence eliminating the need to update all state parameters.
2. In step 3, $f(.)$ represents the deterministic updating of the conditional sufficient statistic based on the a_{jt} and b_{jt} recursions.
3. For PL to work, we need $p(\boldsymbol{Y}_{t+1}|z_t^{(i)})$, the predictive likelihood, for computing the weights in step 1 and $p(\theta_{t+1}|z_t^{(i)}, \mathbf{Y}_{t+1})$, the propagation density, for step 3.

Propogation Density and Resampling Weights

- The propagation density of PL in step 2 is given by

$$p(\theta_{t+1}|\theta_t, D^{t+1}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \propto \theta_{t+1}^{\gamma\alpha_t + g(\boldsymbol{Y}_t, \boldsymbol{\nu}) - 1} \left(1 - \frac{\gamma}{\theta_t}\theta_{t+1}\right)^{(1-\gamma)\alpha_t} \exp\{-\theta_{t+1} h(\boldsymbol{Y}_t, \boldsymbol{\lambda}, \boldsymbol{\nu})\}$$

which is the density form of the scaled hyper-geometric beta distribution.

- The predictive likelihood $p(\boldsymbol{Y}_{t+1}|z_t^{(i)}) = p(\boldsymbol{Y}_{t+1}|\theta_t, \boldsymbol{\lambda}, \boldsymbol{\nu})$ is

$$f(\boldsymbol{Y}_t, \boldsymbol{\lambda}, \boldsymbol{\nu}) \left(\frac{\theta_t}{\gamma}\right)^{g(\boldsymbol{Y}_t, \boldsymbol{\nu})} \frac{B[\gamma\alpha_t + g(\boldsymbol{Y}_t, \boldsymbol{\nu}), (1-\gamma)\alpha_t]}{B[\gamma\alpha_t, (1-\gamma)\alpha_t]} {}_1F_1(a^\star, b^\star, c^\star),$$

where ${}_1F_1(a^\star, b^\star, c^\star)$ represents confluent hyper-geometric function (CHF).

Updating of Discount Factor γ

- For the sequential updating of the γ posterior at each point in time, we can use the marginal likelihood conditional on λ_j's.
- The conditional posterior is given by

$$p(\gamma = k|\boldsymbol{\lambda}, D^{t+1}) \propto \prod_{i=1}^{t+1} p(\boldsymbol{Y}_i|\boldsymbol{\lambda}, D^{i-1}, \gamma = k)p(\gamma = k),$$

where $p(\gamma = k)$ is a discrete uniform prior.

- To incorporate the learning of γ to PL, we first estimate the posterior of γ using the Monte Carlo average based on the updated samples of $\boldsymbol{\lambda}$ and then, we resample particles from this distribution to update $f(.)$ in step 3 of the algorithm.

Bayesian Analysis (2018) **13**, Number 2, pp. 385–409

Sequential Bayesian Analysis of Multivariate Count Data

Tevfik Aktekin*, Nick Polson†, and Refik Soyer‡

Abstract. We develop a new class of dynamic multivariate Poisson count models that allow for fast online updating. We refer to this class as multivariate Poisson-scaled beta (MPSB) models. The MPSB model allows for serial dependence in count data as well as dependence with a random common environment across time series. Notable features of our model are analytic forms for state propagation, predictive likelihood densities, and sequential updating via sufficient statistics for the static model parameters. Our approach leads to a fully adapted particle learning algorithm and a new class of predictive likelihoods and marginal distributions which we refer to as the (dynamic) multivariate confluent hyper-geometric negative binomial distribution (MCHG-NB) and the dynamic multivariate negative binomial (DMNB) distribution, respectively. To illustrate our methodology, we use a simulation study and empirical data on weekly consumer non-durable goods demand.

Keywords: state space, count time series, multivariate poisson, scaled beta prior, particle learning.

Illustration: Weekly Grocery Visits of Households

- Data: The weekly grocery store visits of 540 Chicago based households accumulated over 104 weeks.
- Only 2 households considered in the illustration.
- There is dependence over time and over households (correlation is about 0.4).
- Households are affected by the same random common environment.
- We use independent flat but proper priors for θ_0 and λ_j's.
- For discount parameter γ we define a discrete uniform prior defined over $(0, 1)$.

Weekly Grocery Visits

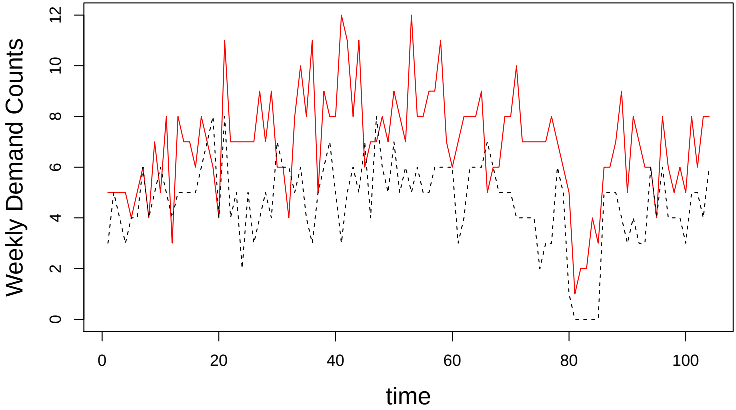
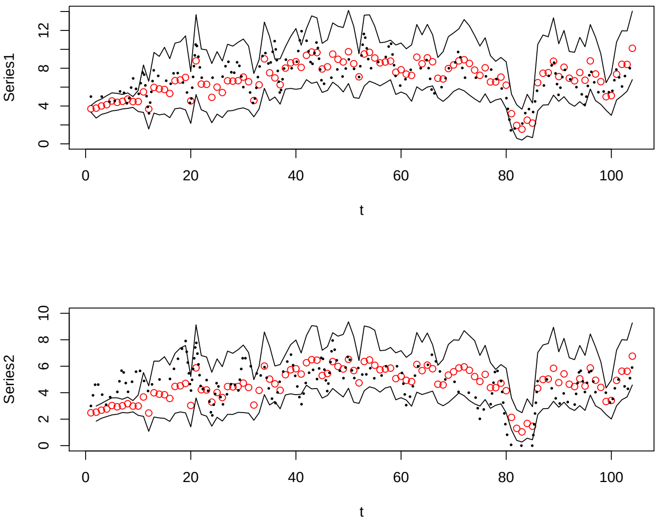
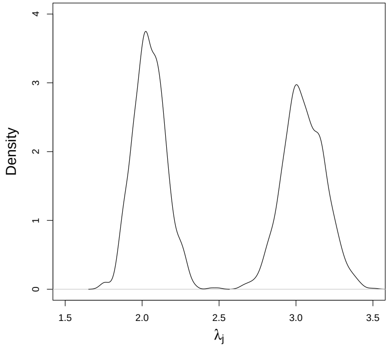


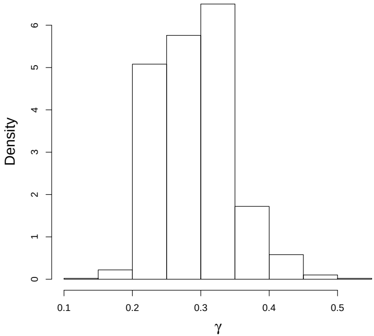
Figure: Weekly demand for households 1 (solid red line) and 2 (dashed line).

Prediction Intervals

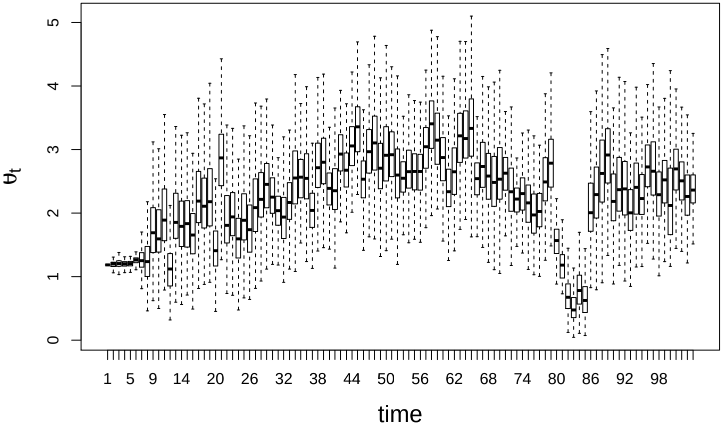


Posterior density plots of λ_1, λ_2 and γ





Posterior Box Plots for the Random Environment



Family of Multivariate Non-Gaussian State Space Models

Tevfik Aktekin
Decision Sciences
*University of New Hampshire**

Nicholas G. Polson
Booth School of Business
University of Chicago

Refik Soyer
Decision Sciences
George Washington University

Abstract

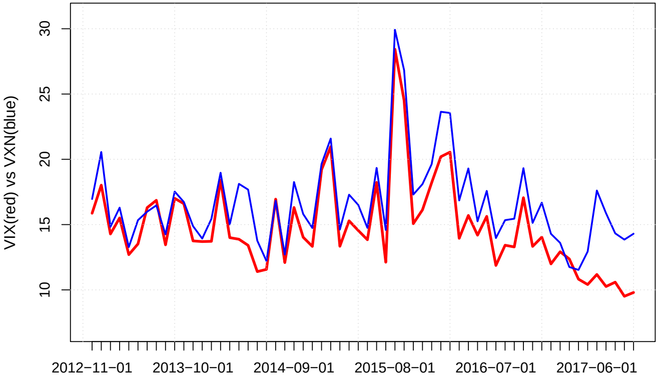
In this paper, we consider the Bayesian analysis of dynamic multivariate non-Gaussian time series models which include many well-known distributions. A key feature of our proposed model is its ability to account for correlations across time as well as across series (contemporary). The proposed modeling approach yields analytically tractable dynamic marginal likelihoods, a property not typically found outside of linear Gaussian time series models. These dynamic marginal likelihoods can be tied back to known static multivariate distributions such as the Lomax, generalized Lomax, and the multivariate Burr distributions. The availability of the marginal likelihoods allows us to develop efficient estimation methods for various settings using Markov chain Monte Carlo as well as particle based methods. To illustrate our methodology, we use simulated data examples and a real application of multivariate time series for modeling the joint dynamics of stochastic volatility in financial indexes, the VIX and VXN.

Keywords: state space, non-Gaussian, dynamic time series, particle learning, stochastic volatility

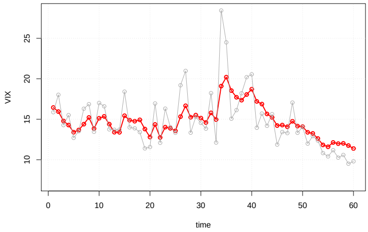
Illustration: Modeling Volatility Market Indices

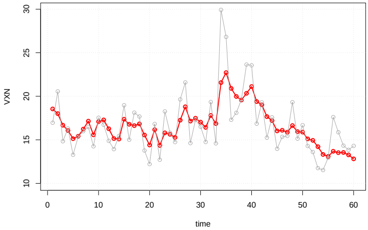
- Monthly time-series of VIX and VXN (October 2012-October 2017)
- Highly correlated series.
- Consider the bivariate generalized Lomax model.
- Parameters $\phi_1 = 1.23$ and $\phi_2 = 1.44$ are estimated and treated as fixed.
- A 100-point discrete prior used for γ over $(0, 1)$.
- We use independent flat but proper priors for θ_0 and λ_j's.

Monthly VIX and VXN

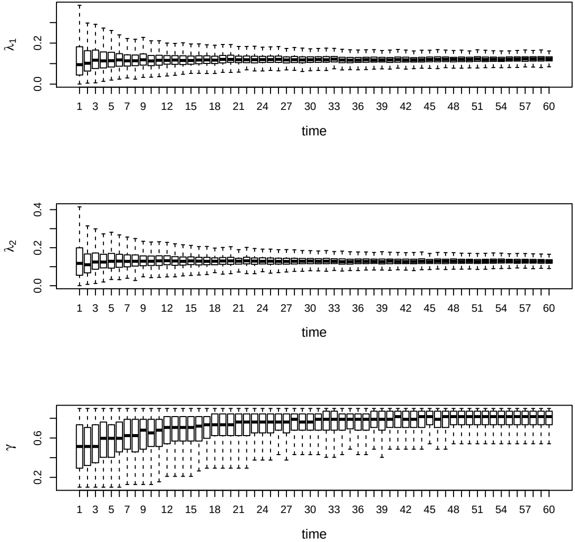


Posterior Predictive Means versus Actual VIX and VXN





Behavior of λ_1, λ_2 and γ



Posterior Distribution of θ_t

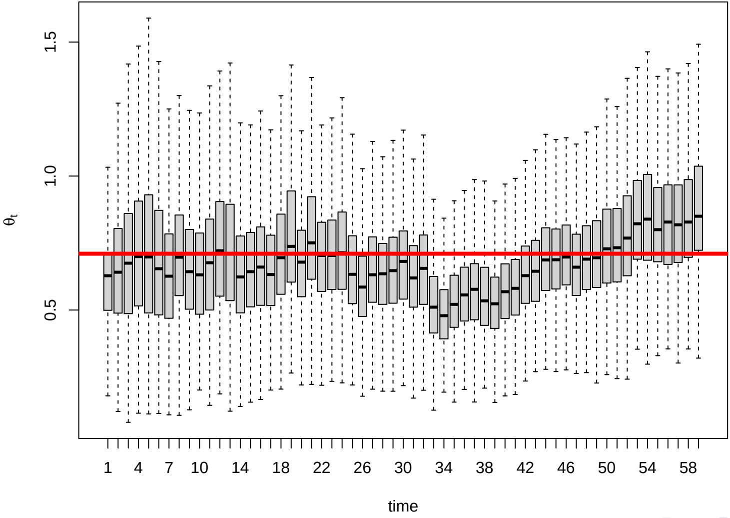


Illustration: Multivariate Burr

- Consider J conditionally independent Weibull series with density

$$p(Y_{jt}|\theta_t, \lambda_j, \phi_j) = \theta_t \lambda_j \phi_j Y_{jt}^{\phi_j - 1} \exp\{-\theta_t \lambda_j Y_{jt}^{\phi_j}\}.$$

- The marginal distribution of Y_{jt} given D_{t-1} can be obtained as a Burr density

$$p(Y_{jt}|D^{t-1}, \lambda_j, \phi_j) = \frac{\frac{\lambda_j}{\gamma\beta_{t-1}} \phi_j Y_{jt}^{\phi_j - 1}}{(1 + \sum_{j=1}^{J} \frac{\lambda_j}{\gamma\beta_{t-1}} Y_{jt}^{\phi_j})^{\gamma\alpha_{t-1}+J}}.$$

- The multivariate distribution $p(\boldsymbol{Y}_t|\boldsymbol{\lambda}, \boldsymbol{\nu}, D^{t-1})$ is given by

$$\frac{\Gamma(\gamma\alpha_{t-1} + J) \prod_{j=1}^{J} \frac{\lambda_j}{\gamma\beta_{t-1}} \phi_j Y_{jt}^{\phi_j - 1}}{\Gamma(\gamma\alpha_{t-1})(1 + \sum_{j=1}^{J} \frac{\lambda_j}{\gamma\beta_{t-1}} Y_{jt}^{\phi_j})^{\gamma\alpha_{t-1}+J}}.$$

- The above distribution is the dynamic version of multivariate generalized Burr distribution; see Tadikamalla (1980).

- When $\phi_j = 1$, for $j = 1, \ldots, J$ it reduces to the dynamic multivariate Lomax distribution.

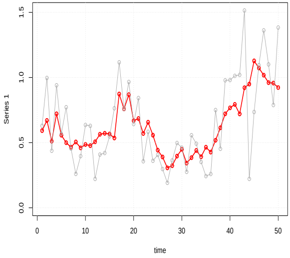
- An interesting property of the multivariate Burr is the nonlinearity of the regressions.

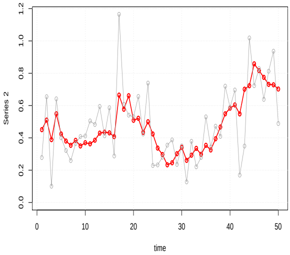
  For example, we can show that the dynamic conditional mean $E[Y_{it}|Y_{jt}, \boldsymbol{\lambda}, \boldsymbol{\nu}, D^{t-1}]$ is given by

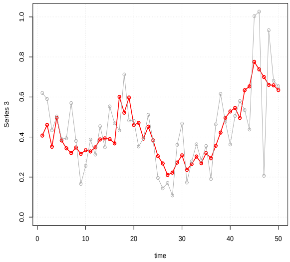
$$\frac{\Gamma(1 + 1/\phi_i)\Gamma(\gamma\alpha_{t-1} + 1 - 1/\phi_i)(\gamma\beta_{t-1} + \lambda_j Y_{jt}^{\phi_j})^{1/\phi_i}}{\lambda_i^{1/\phi_i}\Gamma(\gamma\alpha_{t-1} + 1)},$$

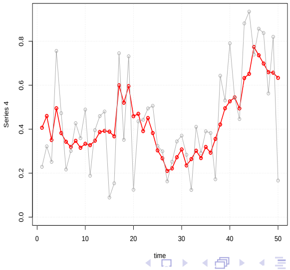
  which is not linear unless $\phi_i = \phi_j = 1$.

Illustration: Simulated Multivariate Burr Data









Concluding Remarks

- Multivariate time series models based on a random environment were developed.
- The multivariate time series family includes members of generalized-gamma family and the members of time transformed exponential family.
- MCMC and Particle filtering methods with PL can be developed.
- Availability of the propogation density still as a scaled hyper-geometric beta density and the resampling weights being in the form of some multivariate confluent hyper-geometric distribution.
- Experience with the Poisson (Aktekin et al. 2018, BA), gamma and Weibull models (Aktekin et al. 2019).