A Tractable State-Space Model for Symmetric Positive-Definite Matrices

Jesse Windle\textsuperscript{1}
Carlos Carvalho\textsuperscript{2}

August 9, 2015

\textsuperscript{1} Hi Fidelity Genetics
\textsuperscript{2} The University of Texas at Austin
The Basic Story

1. The Bayesian analysis of covariance-matrix-valued state-space models can be difficult.
2. The subsequent model is computationally tractable, but it comes at a cost.
State-Space Models

Latent States: $x_{t-1}$ $x_t$ $x_{t+1}$

Observations: $y_{t-1}$ $y_t$ $y_{t+1}$

System’s parameters, $\theta$
State-Space Models

Latent States:
- $x_{t-1}$
- $x_t$
- $x_{t+1}$

Observations:
- $y_{t-1}$
- $y_t$
- $y_{t+1}$

System’s parameters, $\theta$

$$\prod_{i=1}^{T} p(y_t | x_t, \theta) \prod_{i=2}^{T} p(x_t | x_{t-1}, \theta) p(x_1 | \theta)$$
State-Space Models

Latent States:

\[ x_{t-1} \rightarrow x_t \rightarrow x_{t+1} \]

Observations:

\[ y_{t-1} \rightarrow y_t \rightarrow y_{t+1} \]

System’s parameters, \( \theta \)

Filter: \( p(x_t|y_{1:t}) \).
State-Space Models

Latent States: $x_{t-1} \rightarrow x_t \rightarrow x_{t+1}$

Observations: $y_{t-1} \rightarrow y_t \rightarrow y_{t+1}$

System's parameters, $\theta$

Smooth: $p(x_{1:T}|y_{1:T})$. 
State-Space Models

Latent States:

Observations:

Infer: $p(\theta|y_{1:T})$. 
State-Space Models in Finance

\[ R_t \sim N(0, V_t), \]
\[ V_t \sim P(V_{t-1}). \]
State-Space Models in Finance

\[ R_t \sim N(0, V_t), \]
\[ V_t \sim \mathbb{P}(V_{t-1}). \]
$R_t \sim N(0, V_t),$

$V_t \sim \mathcal{P}(V_{t-1}).$
State-Space Models in Finance

\[ R_{t,i} \sim N(0, V_t/k), \ i = 1, \ldots, k, \]
\[ V_t \sim \mathbb{P}(V_{t-1}). \]
State-Space Models in Finance

\[ Y_t \sim W_m(k, V_t/k), \quad Y_t = \sum_{i=1}^{k} R_{t,i} R'_{t,i} \]

\[ V_t \sim P(V_{t-1}). \]
State-Space Models in Finance

\[ Y_t \sim W_m(k, X_t^{-1}/k), \quad Y_t = \sum_{i=1}^{k} R_{t,i} R'_{t,i} \]

\[ X_t \sim \mathbb{P}(X_{t-1}). \]
Our hands are now tied

\[
\prod_{i=1}^{T} p(Y_t|X_t, \theta) \prod_{i=2}^{T} p(X_t|X_{t-1}, \theta) p(X_1|\theta)
\]

Wishart

Problem: Moving around the state-space.

\[ x_t = \text{Lower}(X_t) \sim \text{GP} \ ? \]
Pick a new set of coordinates?

Matrix logarithm [Bauer and Vorkink, 2011]:

\[ X_t = U_t \exp(D_t) U_t', \]

\[ \log X_t = U_t D_t U_t', \]

\[ Z_t = \text{Lower}(\log X_t). \]

\[ \prod_{i=1}^{T} p(Y_t|X_t, \theta) \prod_{i=1}^{T} p(X_t|X_{t-1}, \theta) p(X_0|\theta) \]

\[ p(X_1:T|Y_{1:T}, \theta) \rightarrow \text{Gibbs + Metropolis-Hastings}. \]

\[ p(X_t|X_{-t}, Y_{1:T}, \theta). \]
Pick a new set of coordinates?

**LDL decomposition** [Chiriac and Voev, 2010, Loddo et al., 2011]:

\[
X_t = L_t \exp(D_t)L'_t,
\]

\[
Z_t = (\text{StrictLower}(L_t), \text{Diag}(D_t)).
\]

\[
\prod_{i=1}^{T} p(Y_t|X_t, \theta) \prod_{i=1}^{T} p(X_t|X_{t-1}, \theta) p(X_0|\theta)
\]

\[
p(X_{1:T}|Y_{1:T}, \theta) \rightarrow \text{Gibbs + Metropolis-Hastings.}
\]

\[
p(X_t|X_{-t}, Y_{1:T}, \theta).
\]
Use the original coordinates?

\[ X_t = S_t \Psi_t S_t', \quad S_t S_t' = f(X_{t-1}) \]

| Source | \( f(X_{t-1}) \) | \( \Psi_t \) | \( p(X_1:T | Y_1:T, \theta) \) |
|--------|------------------|---------------|-----------------------------|
| (1)    | \( \lambda^{-1} X_{t-1} \) | \( \mathcal{W}_m(\rho, \mathbf{I}_m/\rho) \) |                                           |
| (2)    | \( \lambda^{-1} X_{t-1} \) | \( \beta_m\left(\frac{n}{2}, \frac{1}{2}\right) \) |                                           |

(1) Philipov and Glickman [2006], Asai and McAleer [2009]
(2) Uhlig [1997], Rank \( m=1 \) Case Only

Other relevant work: Gourieroux et al. [2009], Fox and West [2011]; Prado and West [2010], Jin and Maheu [2013], Shirota et al. [2015]. GARCH literature... Bauwens et al. [2006].
Use the original coordinates?

\[ X_t = S_t \Psi_t S_t', \quad S_t S_t' = f(X_{t-1}) \]

| Source | \( f(X_{t-1}) \) | \( \Psi_t \) | \( p(X_{1:T} | Y_{1:T}, \theta) \) |
|--------|------------------|----------------|---------------------|
| (1)    | \( \lambda^{-1}X_{t-1} \) | \( W_m(\rho, 1_m/\rho) \) | Gibbs + MH          |
| (2)    | \( \lambda^{-1}X_{t-1} \) | \( \beta_m(\frac{n}{2}, \frac{1}{2}) \) | \( p(X_t | Y_{1:t}, \theta) \) |

(1) Philipov and Glickman [2006], Asai and McAleer [2009]
(2) Uhlig [1997], Rank \( m=1 \) Case Only

Other relevant work: Gourieroux et al. [2009], Fox and West [2011]; Prado and West [2010], Jin and Maheu [2013], Shirota et al. [2015]. GARCH literature... Bauwens et al. [2006].
Use the original coordinates?

\[ X_t = S_t \psi_t S'_t, \quad S_t S'_t = f(X_{t-1}) \]

| Source | \( f(X_{t-1}) \) | \( \psi_t \) | \( p(X_{1:T}|Y_{1:T}, \theta) \) |
|--------|-----------------|----------|--------------------------|
| (1)    | \( \lambda^{-1}X_{t-1} \) | \( W_m(\rho, I_m/\rho) \) | Gibbs + MH |
| (2)    | \( \lambda^{-1}X_{t-1} \) | \( \beta_m\left(\frac{n}{2}, \frac{1}{2}\right) \) | \( p(X_t|Y_{1:t}, \theta) \) |

(1) Philipov and Glickman [2006], Asai and McAleer [2009]
(2) Uhlig [1997], Rank \( m=1 \) Case Only

(Other relevant work: Gourieroux et al. [2009], Fox and West [2011]; Prado and West [2010], Jin and Maheu [2013], Shirota et al. [2015]. GARCH literature... Bauwens et al. [2006].)
Uhlig Extension

\[ X_t = S_t \Psi_t S_t', \quad S_t S_t' = \lambda^{-1} X_{t-1} \]

\[ \Psi_t \sim \beta_m \left( \frac{n}{2}, \frac{k}{2} \right), \quad k \in \mathbb{N}; \]

Easy to compute:

- \( p(X_t | Y_{1:t}, \theta) \) Wishart
- \( p(X_t | Y_{1:t}, X_{t+1}, \theta) \) Shifted Wishart
- \( p(X_{1:T} | Y_{1:T}, \theta) \)
Uhlig Extension

\[ X_t = S_t \Psi_t S'_t, \quad S_t S'_t = \lambda^{-1} X_{t-1} \]

\[ \Psi_t \sim \beta_m \left( \frac{n}{2}, \frac{k}{2} \right), \quad k \in \mathbb{N}; \]

Easy to compute:

- \( p(X_t|Y_{1:t}, \theta) \) Wishart
- \( p(X_t|Y_{1:t}, X_{t+1}, \theta) \) Shifted Wishart
- \( p(X_{1:T}|Y_{1:T}, \theta) \)
- \( p(Y_t|Y_{t-1}, \theta) \) Multivariate compound gamma

\[ \implies p(Y_{1:T}|\theta). \]
Uhlig Extension

\[ X_t = S_t \psi_t S_t', \quad S_t S_t' = \lambda^{-1} X_{t-1} \]

\[ \psi_t \sim \beta_m \left( \frac{n}{2}, \frac{k}{2} \right), \quad k \in \mathbb{N}; \]

Easy to compute:

- \( p(X_t | Y_{1:t}, \theta) \) Wishart
- \( p(X_t | Y_{1:t}, X_{t+1}, \theta) \) Shifted Wishart
- \( p(X_{1:T} | Y_{1:T}, \theta) \)
- \( p(Y_t | Y_{t-1}, \theta) \) Multivariate compound gamma

\[ \Rightarrow p(Y_{1:T} | \theta). \]

Only need to record:

\[ \Sigma_t = \lambda \Sigma_{t-1} + Y_t. \]
How does this work? Key Transformation

Muirhead [1982], Uhlig [1997], Díaz-García and Jáimez [1997]:
How does this work? Key Transformation

Muirhead [1982], Uhlig [1997], Díaz-García and Jáimez [1997]:

Density of rank-deficient Wishart

\[
\frac{\pi^{-(mk-k^2)/2} |L|(k-m-1)/2}{2^{mk/2} \Gamma_k \left( \frac{k}{2} \right) |V|^{k/2}} \exp \left( \text{tr} - \frac{1}{2} V^{-1} Y \right) \exp \left( \text{tr} - \frac{1}{2} V^{-1} Y \right)
\]

\[
(dY) = 2^{-k} \prod_{i=1}^{k} I_i^{m-k} \prod_{i<j} (l_i - l_j)(H_1' d H_1) \wedge \bigwedge_{i=1}^{k} dl_i.
\]

(Introductory text: Mikusiński and Taylor [2002])
Example

- 30 stocks from DJIA as of Oct. 2010.
- $Y_t$: Realized kernels (e.g. Barndorff-Nielsen et al. [2009])
Prediction Exercise

- Predictive portfolios:

\[ \pi_t^* = \arg\min_{\pi' \mathbf{1} = 1} \pi' \hat{V}_t \pi \]

\[ \hat{V}_t = \mathbb{E}[V_t|Y_{1:t-1}] \]

- Performance:

portfolio variation = \( \text{var}(\pi_t^*'r_t) \).

<table>
<thead>
<tr>
<th></th>
<th>root mean variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSV Extension</td>
<td>0.00977</td>
</tr>
<tr>
<td>Uhlig Extension</td>
<td>0.00936.</td>
</tr>
</tbody>
</table>
Prediction Exercise

Accumulated Portfolio Variation vs. Time

Squared returns

Time

UE
FSVE
JPM
Drawbacks

Discussion:

- Roberto Casarin
- Catherine Scipione Forbes
- Enrique ter Horst, German Molina
Drawback: $X_t$ is not stationary (realism)
Drawback: $X_t$ is not stationary (predictions)
Drawback: $X_t$ is not stationary (predictions)
Drawback: $X_t$ is not stationary (predictions)

Predictions of future variance:

$$M_h = \mathbb{E}[X_{t+h}^{-1} | X_t^{-1}], \ h > 0.$$  

Konno [1988]:

$$M_h = \frac{n + k - m - 1}{n - m - 1} \lambda M_{h-1}$$

where $M_0 = X_t^{-1}$. 
What does this work at all?
What does this work at all?
Volatility models: think in terms of forecasts

- Uhlig extension:

\[
\mathbb{E}[X_{t+1}^{-1} \mid Y_{1:t}, \theta] = \frac{\lambda k}{n - m - 1} \left( \sum_{i=0}^{t-1} \lambda^i Y_{t-i} + \lambda^t \Sigma_0 \right).
\]
Volatility models: think in terms of forecasts

- Uhlig extension (EWMA):

\[
\mathbb{E}[X_{t+1}^{-1} | Y_{1:t}, \theta] = \frac{\lambda k}{n - m - 1} \left( \sum_{i=0}^{t-1} \lambda^i Y_{t-i} + \lambda^t \Sigma_0 \right).
\]

\[
\frac{n + k - m - 1}{n - m - 1} \lambda = 1 \quad \implies \quad \mathbb{E}[X_{t+1}^{-1} | Y_{1:t}, \theta] = (1 - \lambda) \left( \sum_{i=0}^{t-1} \lambda^i Y_{t-i} + \lambda^t \Sigma_0 \right).
\]
Volatility models: think in terms of forecasts (continued)

- "GARCH" (EWMA-MR):

\[ \mathbb{E}[X_{t+1}^{-1} | Y_{1:t}, \theta] \simeq (1 - \gamma)C + \gamma (1 - \lambda) \left( \sum_{i=0}^{t} \lambda^i Y_{t-i} \right). \]

- Univariate stochastic volatility:

  EWMA-MR of the log squared returns

- Leverage effects:

  asymmetrically weight past observations depending on market movements.
Estimating $\theta = (n, k, \lambda, \Sigma_0)$

The model:

\[ Y_t = W_m(k, (kX_t)^{-1}), \]
\[ X_t = S_t \psi_t S_t', \quad S_t S_t' = \lambda^{-1} X_{t-1}, \]
\[ \psi_t \sim \beta_m\left(\frac{n}{2}, \frac{k}{2}\right), \quad k \in \mathbb{N}. \]

Conjugate prior:

\[ X_1 \sim W_m(n, (\lambda k \Sigma_0)^{-1}). \]

\[ Y_{-\tau}, \ldots, Y_0, Y_1, \ldots, Y_T. \]

\[ \Sigma_t = \sum_{i=0}^{t-1} \lambda^i Y_{t-i} + \lambda^t \Sigma_0 \]
Estimating $\theta = (n, k, \lambda, \Sigma_0)$

The model:

\[
Y_t = W_m(k, (kX_t)^{-1}), \\
X_t = S_t \psi_t S_t', \quad S_t S_t' = \lambda^{-1} X_{t-1}, \\
\psi_t \sim \beta_m\left(\frac{n}{2}, \frac{k}{2}\right), \quad k \in \mathbb{N}.
\]

Conjugate prior:

\[
X_1 \sim W_m(n, (\lambda k \Sigma_0)^{-1}).
\]

\[
Y_{-\tau}, \ldots, Y_0, Y_1, \ldots, Y_T.
\]

\[
\Sigma_t = \sum_{i=0}^{t-1} \lambda^i Y_{t-i} + \lambda^t \Sigma_0 \implies \Sigma_0(\lambda) = \sum_{i=0}^{-\tau} \lambda^i Y_{-i} + 0.
\]
Recapitulation

1. Given our specific observation distribution, it isn't easy to construct tractable matrix-valued state-space models.

2. Uhlig essentially provides a way to do this, but it comes with a cost.

Slides with references:

http://www.jessewindle.com/
Thank you for your attention.

http://www.jessewindle.com/


