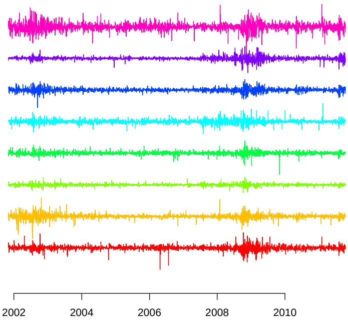
Hierarchical Gaussian Graphical Models: Beyond Reversible Jump

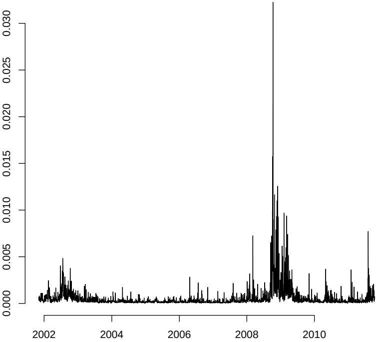
Alex Lenkoski

Institut für Angewandte Mathematik
Universität Heidelberg

A Troubling Dataset

Daily returns on 20 assets from 2002 to 2012





The G-Wishart Distribution

Suppose $\mathcal{D} = \{\mathbf{Z}^{(1)}, \ldots, \mathbf{Z}^{(n)}\}$ with

$$\mathbf{Z}^{(j)} \sim \mathcal{N}_p(0, \mathbf{K}^{-1}).$$

We consider the G-Wishart prior $\mathcal{W}_G(\delta, \mathbf{D})$

$$pr\left(\mathbf{K}|G, \delta, \mathbf{D}\right) = \frac{1}{I_G(\delta, \mathbf{D})}|\mathbf{K}|^{(\delta-2)/2} \exp\left(-\frac{1}{2}\langle \mathbf{K}, \mathbf{D}\rangle\right) \mathbf{1}_{\mathbf{K}\in \mathrm{P}_G},$$

where

$$\mathbf{K} \in \mathrm{P}_G \Rightarrow K_{ij} = 0, \quad (i,j) \notin G.$$

Then

$$\mathbf{K}|\mathcal{D} \sim \mathcal{W}_G(\delta + n, \mathbf{D}^*)$$

where $\mathbf{D}^* = \mathbf{D} + \sum \mathbf{Z}^{(j)}\mathbf{Z}^{(j)\prime}$.

Cholesky Basics

Let $\mathbf{\Phi}$ be the Cholesky decomposition of $\mathbf{K} \sim \mathcal{W}_G(\delta, \mathbf{D})$. Then

$$pr(\mathbf{\Phi}|G) = \prod_{i=1}^{p} \Phi_{ii}^{(\delta+\nu_i^G-1)} \exp\left(-\frac{1}{2}\langle \mathbf{\Phi}'\mathbf{\Phi}, \mathbf{D}\rangle\right)$$

Let F be the fill-in graph associated with G, thus

$$\Phi_{ij} = 0 \quad (i,j) \notin F$$

and

$$\Phi_{ij} = -\frac{1}{\Phi_{ii}} \sum_{l=1}^{i} \Phi_{li}\Phi_{lj} \quad (i,j) \in F \setminus G.$$

Finally, if $\mathbf{K} \sim \mathcal{W}_G(\delta, \mathbb{I}_p)$, then

$$pr(\mathbf{\Phi}|G) = \prod_{i=1}^{p} \Phi_{ii}^{\delta+\nu_i^G-1} \exp\left(-\frac{1}{2} \sum_{(i,j)\in F} \Phi_{ij}^2\right).$$

Reversible Jump

Key result of Dobra, Lenkoski and Rodriguez (2011):

$$R_g^+ = \frac{pr\left(\mathcal{D}|\mathbf{K}'\right)}{pr\left(\mathcal{D}|\mathbf{K}^{[s]}\right)} \frac{pr\left(\mathbf{K}'|G'\right)}{pr\left(\mathbf{K}^{[s]}|G^{[s]}\right)} \times$$
$$\times \frac{J\left(\mathbf{K}' \to (\boldsymbol{\Phi}')^{\nu(G')}\right)}{J\left(\mathbf{K}^{[s]} \to (\boldsymbol{\Phi}^{[s]})^{\nu(G^{[s]})}\right)} \frac{J\left(\left((\boldsymbol{\Phi}^{[s]})^{\nu(G^{[s]})}, \gamma\right) \to (\boldsymbol{\Phi}')^{\nu(G')}\right)}{\frac{1}{\sigma_g\sqrt{2\pi}} \exp\left(-\frac{\left(\Phi'_{i_0j_0} - \Phi^{[s]}_{i_0j_0}\right)^2}{2\sigma_g^2}\right)}.$$

Famous Last Words:

"Basically, at this point, I feel like this problem has been solved"

– A.L., Zürich, 2011

Conditional Bayes Factors (CBFs)

Let $f = (p-1, p)$ and $\mathbf{\Phi}^{-f} = \mathbf{\Phi} \setminus \{\Phi_{p-1,p}, \Phi_{p,p}\}$. Suppose that $G \subset G'$ with $G' \setminus G = f$.
Instead of considering

$$\frac{pr(\mathcal{D}|G')}{pr(\mathcal{D}|G)}$$

we consider the CBF

$$\frac{pr(\mathcal{D}, \mathbf{\Phi}^{-f}|G')}{pr(\mathcal{D}, \mathbf{\Phi}^{-f}|G)}$$

and note

$$pr(\mathcal{D}, \mathbf{\Phi}^{-f}|G) = \int_0^{+\infty} pr(\mathcal{D}|\mathbf{\Phi}) pr(\mathbf{\Phi}|G) d\Phi_{pp}$$

$$pr(\mathcal{D}, \mathbf{\Phi}^{-f}|G') = \int_0^{+\infty} \int_{-\infty}^{+\infty} pr(\mathcal{D}|\mathbf{\Phi}) pr(\mathbf{\Phi}|G') d\Phi_f d\Phi_{pp}.$$

Really? (A Humbling Calculation)

Let $a = p - 1$, we have that

$$pr(\mathcal{D}, \mathbf{\Phi}^{-f} | G) \propto \frac{1}{I_G} \exp\left(-\frac{1}{2}(2D^*_{ap}\phi_{aa}\phi_0 + D^*_{pp}\phi_0^2)\right)$$

where ϕ_0 is the fill-in from $\mathbf{\Phi}^{-f}$ of f. Likewise,

$$pr(\mathcal{D}, \mathbf{\Phi}^{-f} | G') \propto \frac{\Phi_{aa}}{I_{G'}} \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}(2D^*_f \Phi_{aa}\Phi_f + D^*_{pp}\Phi_f^2)\right) d\Phi_f.$$

Thus,

$$\frac{pr(\mathcal{D}, \mathbf{\Phi}^{-f} | G')}{pr(\mathcal{D}, \mathbf{\Phi}^{-f} | G)} = \Phi_{aa} \left(\frac{2\pi}{D^*_{pp}}\right)^{1/2} \exp\left(\frac{1}{2} D^*_{pp}(\phi_0 - \mu)^2\right) \frac{I_G}{I_{G'}}$$

where $\mu = -\Phi_{aa} D^*_f / D^*_{pp}$.

Avoiding Normalizing Constants

We thus have

$$\frac{pr(G|\mathcal{D}, \mathbf{\Phi}^{-f})}{pr(G'|\mathcal{D}, \mathbf{\Phi}^{-f})} = H(\mathbf{\Phi}^{-f}, \mathbf{D}^*) \frac{I_{G'}}{I_G} \frac{pr(G')}{pr(G)}.$$

Suppose that $\tilde{\mathbf{K}} \sim \mathcal{W}_{G'}(\delta, \mathbf{D})$. The exchange algorithm (Murray et al. 2006) allows us to replace this with

$$\frac{H(\mathbf{\Phi}^{-f}, \mathbf{D}^*)}{H(\tilde{\mathbf{\Phi}}^{-f}, \mathbf{D})} \frac{pr(G')}{pr(G)}$$

The double Metropolis Hastings (Liang 2010) algorithm allows $\tilde{\mathbf{K}}$ to be sampled via an MCMC step from $\mathbf{K}$.

A perfectly terrible update

Under the prior $\mathcal{W}_G(\delta, \mathbb{I}_p)$ and current state $\mathbf{K}$, consider the following algorithm

1. Determine $\mathbf{\Phi}$ from $\mathbf{K}$
2. Propose $\mathbf{\Psi}$:
   a. Sample $c \sim \chi^2_{\delta+\nu_i^G}$ and set $\Psi_{ii} = c^{1/2}$
   b. Sample $\Psi_{ij} \sim \mathcal{N}(0, 1)$ for $(i, j) \in G$
   c. Complete $\mathbf{\Psi}$ for all $(i, j) \in F \setminus G$, $i < j$
3. Compute

$$\alpha = \exp\left(-\frac{1}{2} \sum_{(i,j)\in F\setminus G} (\Psi_{ij}^2 - \Phi_{ij}^2)\right)$$

4. With probability $\min\{\alpha, 1\}$ set $\mathbf{K} = \mathbf{\Psi}'\mathbf{\Psi}$

A terrible posterior sampler, but perfect here!

Simulation Results

	Time (sec)		MSE	
	Mean	SD	Mean	SD
CL	182.5	(4.1)	0.0088	(6e-04)
RJ	1518.4	(30.2)	0.0349	(0.0025)

All of this, and in six dimensions!

A Multivariate Graphical Stochastic Volatility Model

We can now address our "troubling" dataset.

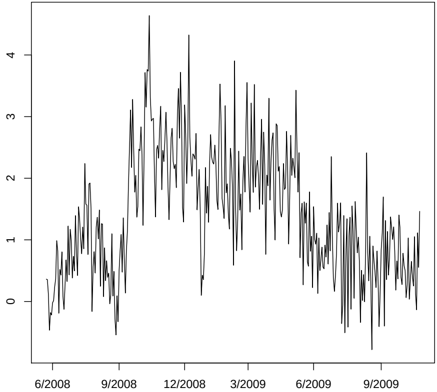
$$\begin{aligned} \mathbf{Y}_t | \mathbf{K}, X_t &\sim \mathcal{N}_p(\mathbf{0}, \exp(X_t)\mathbf{K}^{-1}) \\ X_t | \phi, X_{t-1}, \beta, \tau &\sim \mathcal{N}(\phi X_{t-1} + \beta VIX_{t-1}, \tau^{-1}) \\ \mathbf{K} | G &\sim \mathcal{W}_G(\delta, \mathbb{I}_p) \\ G &\sim pr(G) \end{aligned}$$

And note that

$$\mathbf{K} | \{\mathbf{Y}^{(1)}, \ldots, \mathbf{Y}^{(T)}\}, \mathbf{X}, G \sim \mathcal{W}_G \left(\delta + T, \mathbf{D} + \sum_{t=1}^{T} \frac{\mathbf{Y}^{(t)} \mathbf{Y}^{(t)\prime}}{\exp(X_t)} \right)$$

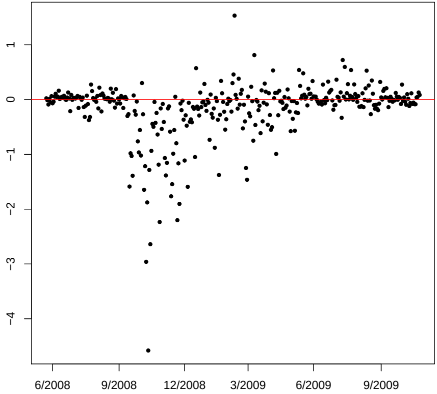
Volatility Components

Posterior mean of $X_{t+1}|\mathbf{Y}_1, \ldots, \mathbf{Y}_t$



Predictive Performance

We now forecast one step ahead $pr(\mathbf{Y}^{(t+1)}|\mathbf{Y}^{(1:t)})$, comparing the volatility model to one without heteroskedasticity



Above, the difference in Energy Scores.

Conclusions

Paper on the arXiv

- Yuan Cheng and Alex Lenkoski. *A Multivariate Graphical Stochastic Volatility Model*, in submission

More work on CBFs

- Anna Karl and Alex Lenkoski. *Instrumental Variable Bayesian Model Averaging via Conditional Bayes Factors*, in submission
- Alexander Jordan and Alex Lenkoski. *Tobit Bayesian Model Averaging and Foreign Direct Investment*, in submission
- Tobias Kaiser and Alex Lenkoski. *Bayesian Verification of Instrument Validity via Conditional Bayes Factors*, in submission