Parallel Bayesian inference for high dimensional dynamic factor copulas

Hoang Nguyen

joint work with **M. Concepcion Ausin, Pedro Galeano**
Universidad Carlos III de Madrid

Thursday 8th November, 2018

Motivation and research questions

- Investors usually face problems when measuring the return sensitivity for different groups of the financial assets.
    - But ... The traditional multivariate models can be very restrictive.
- Vine copula GARCH models are commonly used.
    - But ... Most of the vine copulas are static (except Almeida et al. (2012))
- Alternatively, Factor copulas proposed by Krupskii and Joe (2013) have several advantages. The observed variables are conditionally independent given one or more latent variables (see Oh and Patton (2017a),Oh and Patton (2017b), Creal and Tsay (2015)).

Motivation and research questions (cont.)

- We propose new models of time varying dependence in high dimensional time series and use the Bayesian approach to estimate the different factor copula models.
  - The dynamic factor loadings are modelled as generalized autoregressive score (GAS) processes.
  - The high dimensional problem is handled with a factor structure which also allows the inference strategy to run in a parallel setting.
  - Using group hyperbolic skew Student copula (GSt), we obtain different tail behaviour and asymmetric dependence among financial time series.
  - The model is extendible in which deleting or adding more variables does not change the form of model results.

Agenda

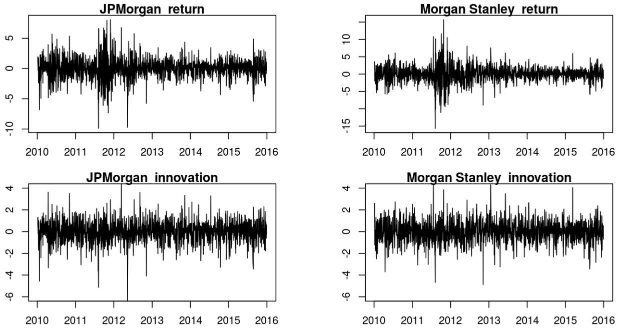
An example of GARCH copulas

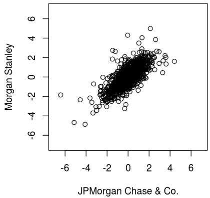
Let $r_t = (r_{1t}, \ldots, r_{dt})$ be a d-dimensional financial return time series at time t where $t = 1, \ldots, T$. Each marginal returns i are filtered out the conditional mean and conditional variance using $AR(k) - GARCH(p, q)$ model:

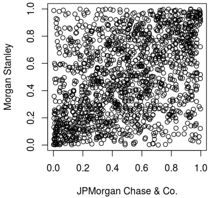
$$
\begin{aligned}
r_{it} &= c_i + \phi_{i1} r_{i,t-1} + \ldots + \phi_{ik} r_{i,t-k} + a_{it} \\
a_{it} &= \sigma_{it} \eta_{it} \text{ where } \eta_{it} \sim F_i(.|\vartheta_i) \\
\sigma^2_{it} &= \omega_i + \alpha_{i1} a^2_{i,t-1} + \ldots + \alpha_{ip} a^2_{i,t-p} + \beta_{i1} \sigma^2_{i,t-1} + \ldots + \beta_{iq} \sigma^2_{i,t-q}
\end{aligned}
$$

The standardized innovations η_{it} are taken from the filter.
Note that other models such as EGARCH, IGARCH, GJR-GARCH, Stochastic volatility could also be applied.

An example of GARCH copulas (cont.)







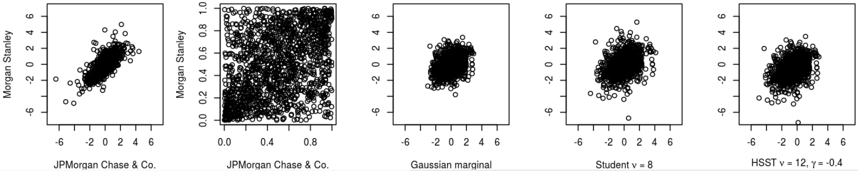
An example of GARCH copulas (cont.)

Copula data $u_{it} = F_i(\eta_{it}|\vartheta_i)$ is known to be an uniform U(0,1) distribution for $i = 1, ..., d$. Thus, the joint cdf of the vector $U_t = (u_{1t}, ..., u_{dt})$, is given by a copula cdf

$$F(u_{1t}, ..., u_{dt}|\theta_t) = C(u_{1t}, ..., u_{dt}|\theta_t)$$

The computation for the likelihood $C(u_{1t}, ..., u_{dt}|\theta_t)$ is often computational expensive (see Krupskii and Joe (2013)).

Alternatively, Creal and Tsay (2015), Oh and Patton (2017b) choose another approach based on the transformation of copula data into real domain where it is trivial to model the joint dependence.



Gaussian one factor copulas

Assume that the variables x_{it} are conditionally independent given one standard Normal latent variable z_t and they are bivariate Gaussian with correlation ρ_{it},

$$\begin{aligned} x_{it} &= \rho_{it} z_t + \sqrt{1-\rho_{it}^2}\epsilon_{it}, \text{ for } t = 1,\ldots,T; i = 1,\ldots,d; \\ x_t &= \{x_{1t},\ldots,x_{dt}\} \sim \Phi_d(\cdot|R_t) \text{ where } R_{ijt} = \rho_{it}\rho_{jt}; i \neq j; \\ u_t &= \{u_{1t},\ldots,u_{dt}\} = \{\Phi(x_{1t}),\ldots,\Phi(x_{dt})\} \sim C_{Gauss}(\cdot|R_t) \end{aligned} \tag{1}$$

where $\epsilon_{it} \sim iid\Phi(0,1)$ is the individual shock and $z_t \sim iid\Phi(0,1)$ is one state latent variable. x_t follow a multivariate normal one factor.

Gaussian one factor copulas (Cont.)

The dynamic Gaussian one factor copulas are constructed by allowing $\rho_t = (\rho_{1t}, \ldots, \rho_{dt})$ to vary across time as a modified logistic transformation of $f_t = (f_{1t}, \ldots, f_{dt})$ which follows the generalized autoregressive score process(GAS, see Creal et al. (2013)), Oh and Patton (2017b)

$$
\begin{aligned}
\rho_{it} &= \frac{1 - \exp(-f_{it})}{1 + \exp(-f_{it})} \in [-1, 1] \\
f_{it} &= (1 - b)\, f_{ic} + a s_{i,t-1} + b f_{i,t-1} \\
s_{it} &= \frac{\partial \log p(u_t | z_t, f_t, \mathcal{F}_t, \theta)}{\partial f_{it}} \\
s_{it} &= \frac{1}{2} x_{it} z_t + \frac{1}{2} \rho_{it} - \rho_{it} \frac{x_{it}^2 + z_t^2 - 2\rho_{it} x_{it} z_t}{2(1 - \rho_{it}^2)}
\end{aligned}
\tag{2}
$$

Gaussian one factor copulas (Cont.)

In comparison to Creal et al. (2011), Oh and Patton (2017b), the score of GAS process,

$$\begin{aligned} s_{it}^{org} &= \frac{\partial \log p\,(u_t|f_t, \mathcal{F}_t, \theta)}{\partial f_{it}} \\ &= \mathbf{E}_{z_t}\left[\frac{\partial \log p(u_t|z_t, f_t, \theta)}{\partial f_t}\middle|\, u_t, f_t, \theta\right] = \mathbf{E}_{z_t}\left[s_{it}|\, u_t, f_t, \theta\right] \end{aligned} \quad (3)$$

where z_t is sampled from its posterior $p(z_t|x_t, f_t, \theta)$ (see Lucas et al. (2018)).

- The score s_{it} is close to the s_{it}^{org}
- Reducing the computational burden (contrasts with Oh and Patton (2017b) where the expressions for s_{it} are obtained by numerical differentiation of the joint copula density).

Generalized hyperbolic skew Student-t one factor copulas

Assuming that for all assets i,

$$x_{it} = \gamma\zeta_t + \sqrt{\zeta_t}(\rho_{it} z_t + \sqrt{1 - \rho_{it}^2}\epsilon_{it})$$

where $\zeta_t \sim Inv - Gamma(\frac{\nu}{2}, \frac{\nu}{2})$ is the mixing variable who shares the same value for asset i and γ accounts for the asymmetric distribution. $x_t = \{x_{1t}, \dots, x_{dt}\} \sim F_{MGSt}(\cdot \mid \mu = 0, R_t, \nu, \gamma)$. And further, if we define $\tilde{x}_{it} = \frac{x_{it} - \gamma\zeta_t}{\sqrt{\zeta_t}}$, then $\tilde{x}_t = \{\tilde{x}_{1t}, \dots, \tilde{x}_{dt}\} \sim \Phi_d(\cdot \mid \mu = 0, R_t)$, the dynamic equation (2) remains

$$s_{it} = \frac{\partial \log p(u_{it} | z_t, \zeta_t, \gamma, f_t, \mathcal{F}_t, \theta)}{\partial f_{it}} = \frac{1}{2}\tilde{x}_{it} z_t + \frac{1}{2}\rho_{it} - \rho_{it}\frac{\tilde{x}_{it}^2 + z_t^2 - 2\rho_{it}\tilde{x}_{it} z_t}{2(1 - \rho_{it}^2)}$$

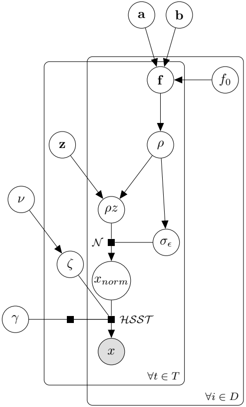
Group dynamic hyperbolic skew Student-t one factor copulas

The fact that using a few parameters to control for the comovement in extreme events is really restrictive. The strategy (suggested by Creal and Tsay (2015)) is to split the assets into G smaller groups that can share the same characteristics.

$$\begin{aligned} x_{igt} &= \gamma_g \zeta_{gt} + \sqrt{\zeta_{gt}} \left(\rho_{igt} z_t + \sqrt{1 - \rho_{igt}^2} \epsilon_{git} \right) \\ \rho_{igt} &= \frac{1 - \exp\left(-f_{igt}\right)}{1 + \exp\left(-f_{igt}\right)} \\ f_{igt} &= \left(1 - b_g\right) f_{igc} + a_g s_{i,g,t-1} + b_g f_{i,g,t-1} \end{aligned} \tag{4}$$

where $\left(a_g, b_g, \nu_g, \gamma_g\right)$ accounts for the dynamic and asymmetric dependence between individual assets belong to the same group.

Bayesian approach



Algorithm 1: BUG illustration - MGSt copulas

Prior: $a \sim U(-0.5, 0.5)$;
$b \sim U(0, 1)$;
$f_c \sim U(-5, 5)$;
$\nu \sim 4 + G(2, 2.5)$;
$\gamma \sim N(0, 1)$;
$z_t \sim N(0, 1)$;
$\zeta_g \sim IG(\nu_g/2, \nu_g/2)$;

Data transformation

$$x_{igt} = F^{-1}(u_{it}|\nu_g, \gamma_g)$$
$$\tilde{x}_{igt} = \frac{x_{igt} - \gamma_g \zeta_{gt}}{\sqrt{\zeta_{gt}}}$$

Parameters transformation

$$s_{igt} = \frac{1}{2}\tilde{x}_{igt} z_t + \frac{1}{2}\rho_{igt} - \rho_{igt}\frac{\tilde{x}_{igt}^2 + z_t^2 - 2\rho_{igt}\tilde{x}_{igt} z_t}{2(1-\rho_{igt}^2)}$$
$$f_{i,g,t+1} = f_{0c} + a_g s_{igt} + b_g f_{igt}$$
$$\rho_{i,g,t+1} = \frac{1-\exp(-f_{i,g,t+1})}{1+\exp(-f_{i,g,t+1})};$$

Conditional distribution

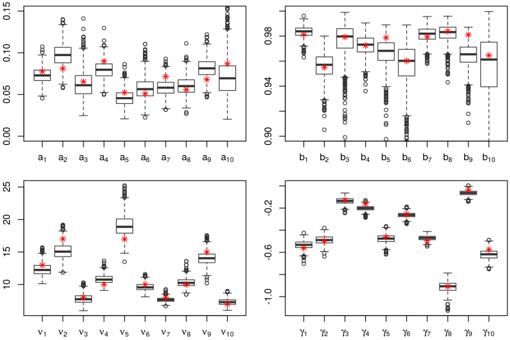
$$\tilde{x}_{igt} \sim N(\rho_{igt} z_t, \sqrt{1-\rho_{igt}^2});$$

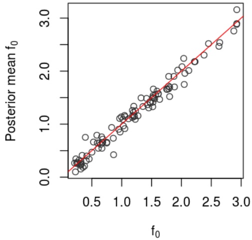
https://github.com/hoanguc3m/FactorCopula

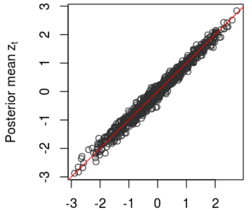
Data simulation

We simulate $d = 100$, $G = 10$, $T = 1000$, iterations $= 10,000$

It took 13 minutes, 35 minutes, 45 minutes for Gaussian, Student, Generalized hyperbolic student copulas on Intel i7-4770 PC (4 cores - 8 threads - 3.4GHz).







Monte Carlo study

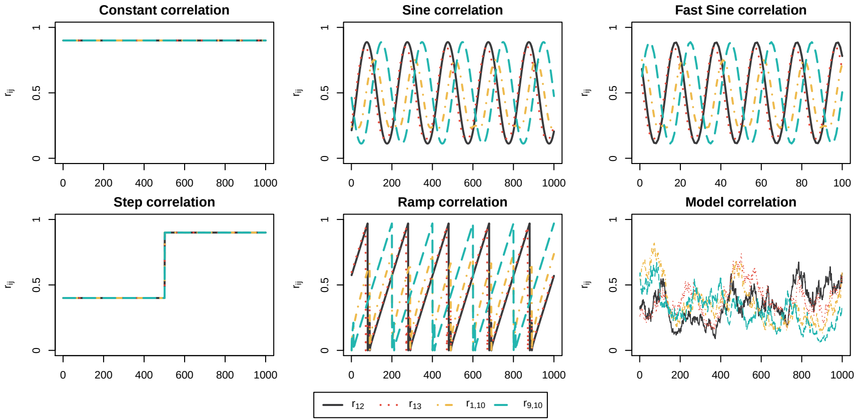
Next, we compare several dynamic correlation models in different stress scenarios based on Engle (2002). We estimate R_t accuracy (based on the mean absolute error - MAE) using EWMA, DCC, GAS models.

$$\begin{aligned} \text{Multivariate Gaussian:}\ & y_t = R_t^{1/2}\epsilon_t \\ \text{Multivariate Student-t :}\ & y_t = \sqrt{\zeta_t}R_t^{1/2}\epsilon_t \\ \text{Multivariate GSt:}\ & y_t = \gamma\zeta_t + \sqrt{\zeta_t}R_t^{1/2}\epsilon_t \end{aligned}$$

where $\epsilon_t \sim \Phi_d(0, I_d)$, $\zeta_t \sim IG(\nu/2, \nu/2)$, we let $\nu = 10$ and $\gamma = -0.1$ and the time varying scale matrix

$$R_t = \begin{bmatrix} 1 & \dots & r_{1dt} \\ \dots & 1 & \dots \\ r_{d1t} & \dots & 1 \end{bmatrix} = \rho_t\rho_t' + diag\left(1 - \rho_t^2\right) \tag{5}$$

Monte Carlo study (next)



Monte Carlo study (next)

Table: MAE results: in-sample

	Constant	Sine	Fast sine	Step	Ramp	Model
MAE - Multivariate Gaussian						
GAS	0.562	0.639	0.794	0.765	0.738	0.861
DCC	1.000	1.000	1.000	1.000	1.000	1.000
EWMA	5.153	1.048	1.068	1.068	0.921	1.199
MAE - Multivariate Student-t						
GAS	0.506	0.792	0.883	0.871	0.805	0.947
DCC	1.000	1.000	1.000	1.000	1.000	1.000
EWMA	4.636	1.160	1.193	1.043	1.059	1.251
MAE - MGSt						
GAS	0.867	0.878	0.981	0.953	0.892	0.916
DCC	1.000	1.000	1.000	1.000	1.000	1.000
EWMA	5.135	1.107	1.194	1.137	1.045	1.181

The table shows the relative MAE for the estimated dynamic correlation of EWMA, DCC, GAS models.

Estimation results I

We illustrate an empirical example using $d = 140$ stock returns divided into $G = 12$ groups from 01/01/2007 to 01/09/2014 of the companies listed in S&P 500 index. The daily data contain $T = 2000$ observation days. We use AR(1)-GARCH(1,1) - skew Student-t innovation to marginalize each stock returns.

Table: Estimation results for alternative copula models

	Gaussian multi.group (3)	Student-t multi.group (6)	MGSt multi.group (9)
AIC	-164658	-192197	-196901
BIC	-163739	-191211	-195848
DIC	-167622	-212414	-213090
# params	164	176	188
a	[0.056,0.146]	[0.025,0.067]	[0.026,0.069]
b	[0.855,0.984]	[0.945,0.995]	[0.956,0.995]
ν		[6.816,11.870]	[7.885,23.086]
γ			[-1.215,-0.184]
f_c	[0.965,2.369]	[1.030,2.442]	[1.015,2.432]

Posterior estimations for nine one factor copula models and model selection criteria.

Empirical estimations

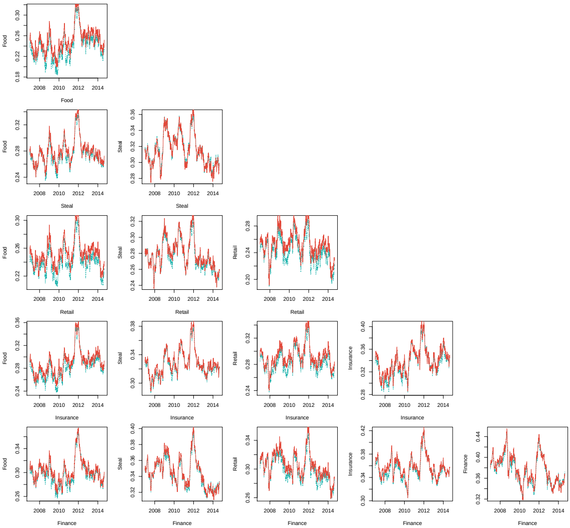


Figure: The Kendall-τ correlation among group sectors

Empirical estimations

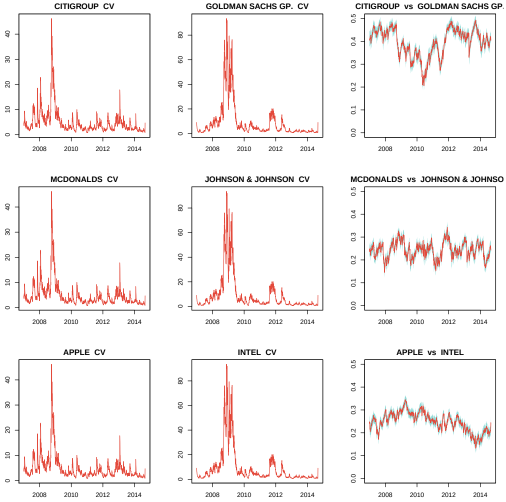


Figure: Posterior Kendall-τ correlation among time series

Main contribution

We employ the Bayesian approach to estimate the different specifications of the factor copula models.

- Condition on the latent factors, time series become independent which allows the algorithm to run in a parallel setting.
- Factor loadings could be modelled as generalized autoregressive score (GAS) processes imposing a dynamic dependence structure in their densities.
- Using hyperbolic skew Student copulas, we obtain different types of tail and asymmetric dependence.
- Furthermore, the model is extendible. The number of parameters scales linearly with the dimension in parametric models.

Discussion and Extension

- The more complex copula functions are build based on the distribution of ζ_{gt}.
- Computational expensive on $x_{it} = F^{-1}(u_{it}|\theta)$.
- Dynamic multi-skewness factor copulas.
- Understanding about the tail behaviour of the GSt copula.
- Taking into account more factors.

Thank you
and have a good time!

C. Almeida, C. Czado, and H. Manner. Modeling high dimensional time-varying dependence using d-vine scar models. *arXiv preprint arXiv:1202.2008*, 2012.

D. Creal and R. Tsay. High dimensional dynamic stochastic copula models. *Journal of Econometrics*, 189(2):335 – 345, 2015.

D. Creal, S. Koopman, and A. Lucas. A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business & Economic Statistics*, 29(4):552–563, 2011.

D. Creal, S. J. Koopman, and A. Lucas. Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5): 777–795, 2013.

R. Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.

P. Krupskii and H. Joe. Factor copula models for multivariate data. *Journal of Multivariate Analysis*, 120:85–101, 2013.

A. Lucas, J. Schaumburg, and B. Schwaab. Bank business models at zero interest rates. *Journal of Business & Economic Statistics*, (just-accepted), 2018.

H. Manner and J. Segers. Tails of correlation mixtures of elliptical copulas. *Insurance: Mathematics and Economics*, 48(1):153–160

Tail dependence for GSt copula

Consider the bivariate GSt copula, we derive the tail dependence of pseudo observable x_1 and x_2 from the Equation (4) as

$$
\begin{aligned}
x_1 &= \gamma_g \zeta_g + \sqrt{\zeta_g}(\rho_1 z + \sqrt{1-\rho_1^2}\epsilon_1) \\
x_2 &= \gamma_g \zeta_g + \sqrt{\zeta_g}(\rho_2 z + \sqrt{1-\rho_2^2}\epsilon_2)
\end{aligned} \tag{6}
$$

$$
\begin{aligned}
&C_{GSt}(u,u|\nu_g,\gamma_g) = P(x_1 \leq F_{HS}^{-1}(u), x_2 \leq F_{HS}^{-1}(u)) \\
&= \mathbb{E}\left(P\left(\epsilon_1 \leq \frac{F_{HS}^{-1}(u) - \gamma_g\zeta_g}{\sqrt{\zeta_g}\sqrt{1-\rho_1^2}} - \frac{\rho_1 z}{\sqrt{1-\rho_1^2}}, \epsilon_2 \leq \frac{F_{HS}^{-1}(u) - \gamma_g\zeta_g}{\sqrt{\zeta_g}\sqrt{1-\rho_2^2}} - \frac{\rho_2 z}{\sqrt{1-\rho_2^2}}\right)\right) \\
&= \mathbb{E}(\Phi(p_1).\Phi(p_2))
\end{aligned} \tag{7}
$$

Then, we obtain $C_{GSt}(u,u|\nu_g,\gamma_g)$ as the numerical integral over $z \sim N(0,1)$ and $\zeta_g \sim IG(\nu_g/2, \nu_g/2)$ As taking the equation to the limit, the tail dependence are

$$\lambda_L = \lim_{u\to 0} \frac{C(u,u)}{u} \; ; \; \lambda_U = 2 + \lim_{u\to 0} \frac{C(1-u,1-u)-1}{u}$$

Model selection

The statistics of model selection are calculated based on the average of log likelihood. We take the average of the log likelihood after MCMC iterations at the posterior mean of the interested parameters as the integral over the nuisance parameter space. The set parameters of interest is $\theta_{int} = \{a, b, f_0, \nu, \gamma\}$ and the nuisance parameters are $\theta_{nui} = \{z, \zeta\}$.

$$
\begin{aligned}
AIC &= -2\mathbb{E}_{\theta_{nui}}(\log(p(u|\bar{\theta}_{int}, f, \mathcal{F}))) + 2p \\
BIC &= -2\mathbb{E}_{\theta_{nui}}(\log(p(u|\bar{\theta}_{int}, f, \mathcal{F}))) + p\log T \\
DIC &= -4\mathbb{E}_{\theta}[\log p(u|\theta, f, \mathcal{F})|u] + 2E_{\theta_{nui}}(\log(p(u|\bar{\theta}_{int}, f, \mathcal{F})))
\end{aligned}
$$

where p is the number of parameters of interest.