Parallel Bayesian inference for high dimensional dynamic factor copulas

Hoang Nguyen, Concepcion Ausin and Pedro Galeano
Statistics Department, Universidad Carlos III de Madrid

Universidad Carlos III de Madrid
www.uc3m.es

Introduction

Investors usually face important problems when measuring the return sensitivity of financial assets. Practically, a multivariate volatility model for financial returns is employed where the standardized innovations are assumed to have either a multivariate Gaussian or a multivariate Student-t distribution. Despite that the number of parameters becomes explosive, such multivariate models still can be very restrictive. A recent approach, factor copulas as the truncated C-vines rooted at the latent variables are proposed for tackling the problem.

- Each marginal returns are filtered out its autocorrelation, for example, ARMA-GARCH process.
- The joint dependence of the standardized innovations are modelled flexibly by different copula models.
- It is assumed that each standardized innovation is affected by some common latent factors. Conditional on these factors, the returns are independent.

Main contribution

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 imposing a dynamic dependence structure in their densities.
- 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 (HSST), 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 and the number of parameters scales linearly with the dimension in parametric models.

Marginal model for each return series

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

$$
\begin{aligned}
r_{it} &= c_i + \phi_{i1} r_{i,t-1} + \dots + \phi_{ik} r_{i,t-k} + a_{it} \\
a_{it} &= \sigma_{it}\eta_{it} \\
\sigma_{it}^2 &= \omega_i + \alpha_{i1} a_{i,t-1}^2 + \dots + \alpha_{ip} a_{i,t-p}^2 + \beta_{i1}\sigma_{i,t-1}^2 + \dots + \beta_{iq}\sigma_{i,t-q}^2
\end{aligned}
$$

The standardized innovations η_{it} are taken from the filter. Note that other models such as EGARCH, IGARCH, GJR-GARCH, Stochastic volatility also could be applied. We make use of factor copulas to define the dependence structure over $u_{it} = F_i(\eta_{it}|\vartheta_i)$ which is known to be an uniform U(0,1) distribution for $i = 1, \dots, d$. Thus, the joint cdf of the vector $U_t = (u_{1t}, \dots, u_{dt})$, F, is given by a copula cdf

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

The computation for the marginal likelihood $C(u_{1t}, \dots, u_{dt}|\theta_t)$ is often expensive. One common approach is to transform the domain of copula data to the domain where the dependence structures are easier to obtain. Due to the fact that copulas are invariant under strictly increasing transformations of the random variables, It somehow coming back to standard factor models that had been widely discussed in the literature. see [1], [2] and [3].

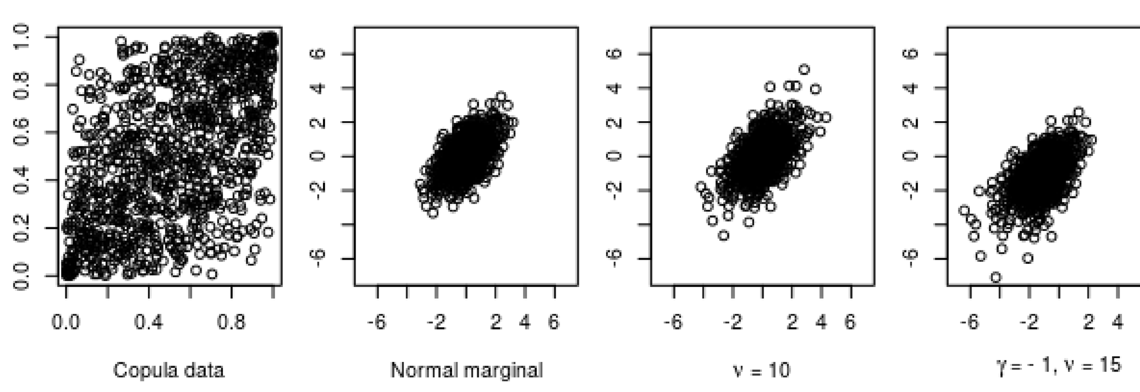

Figure 1 : Copula domain to real domain

Dynamic Hyperbolic skew Student copula model

The pseudo observable x_{it} are achieved from copula data $x_{it} = F_{HSST}^{-1}(u_{it}|\nu_g, \gamma_g)$. x_t follow multivariate HSST distribution then u_t are Hyperbolic Skew student copulas

$$x_{it} = \gamma_g \zeta_{gt} + \sqrt{\zeta_{gt}}(\rho_{it} z_t + \sqrt{1 - \rho_{it}^2}\epsilon_{it}) \tag{1}$$

where $\epsilon_{it} \sim iidN(0, 1)$. In this case, $z_t \sim iidN(0, 1)$ is one state latent variable which affects each individual pseudo-observable innovation x_{it}. $\zeta_{gt} \sim Inv - Gamma(\frac{\nu_g}{2}, \frac{\nu_g}{2})$ is the mixing variable who shares the same value for asset i belong to the same group ($i \in A_g$, $g = 1, \dots, G$). γ_g accounts for the asymmetric distribution. Then, the correlation matrix $R_t = \rho_t \rho_t' + diag(1 - \rho_t^2)$ where $\rho_t = (\rho_{1t}, \dots, \rho_{dt})$. The dynamic process of ρ_t are modelled as an observation driven process of a modified logistic transformation of $f_t = (f_{1t}, \dots, f_{dt})$ which guarantees $\rho_{it} \in (-1, 1)$.

$$
\begin{aligned}
\rho_{it} &= \frac{1 - \exp(-f_{it})}{1 + \exp(-f_{it})} \\
f_{i,t+1} &= (1 - b_g) f_{i0} + a_g s_{it} + b_g f_{it} \\
s_{it} &= \frac{\partial \mathbf{log} p(u_{it}|z_t, \zeta_t, \gamma_g, f_t, \mathcal{F}_t, \theta)}{\partial f_{it}}
\end{aligned} \tag{2}
$$

f_{it} are proposed based on the GAS Model (see [4]) in which the score s_{it} depends on the complete density. Also let $\tilde{x}_{it} = \frac{x_{it} - \gamma_g \zeta_{gt}}{\sqrt{\zeta_{gt}}}$, the equation (2) becomes

$$s_{it} = \frac{1}{2}\tilde{x}_{it} z_t + \frac{1}{2}\rho_{it} - \rho_{it}\frac{\tilde{x}_{it}^2 + z_t^2 - 2\rho_{it}\tilde{x}_{it} z_t}{2(1 - \rho_{it}^2)} \tag{3}$$

Bayesian Estimation

For d = 100, T = 1000, iterations = 10.000, it took 13 minutes, 35 minutes, 45 minutes for Gaussian, Student, HSST copulas on Intel i7-4770 PC (4 cores - 8 threads - 3.4GHz).

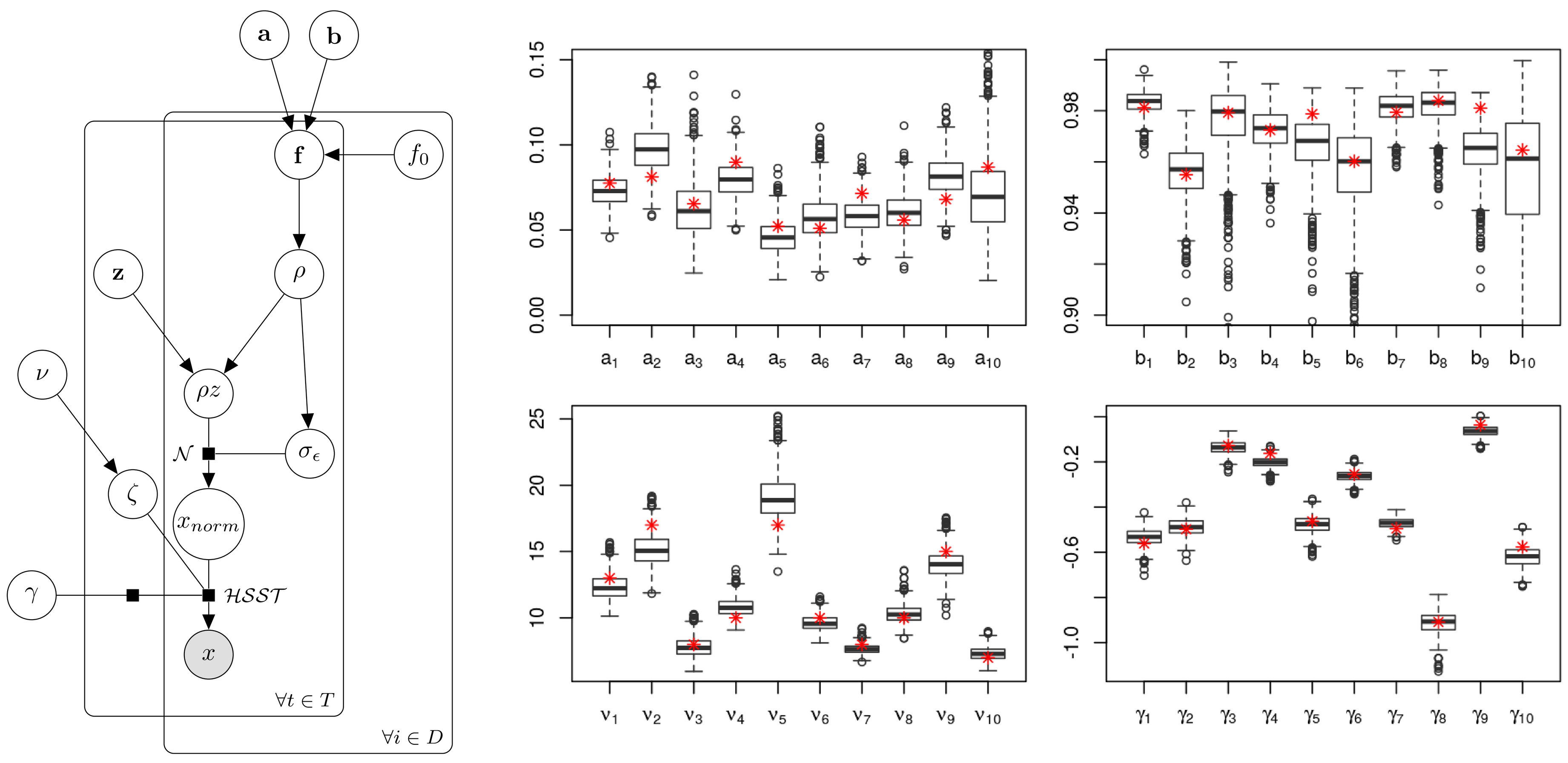

Figure 2 : Box plots for Posterior estimation of a_g, b_g, ν_g, γ_g with simulated data

Empirical Results

We illustrate an empirical example using $d = 150$ stock returns from 01/01/2010 to 31/12/2015 of the companies listed in S&P 500 index. The daily data contain $T = 1509$ observation days. We use $AR(1) - GARCH(1, 1) - HSST$ innovation to marginalize each stock returns.

Table 1 : Estimation results for alternative copula models

	Gaussian block equi (1)	Gaussian 1G (2)	Gaussian dynamic (3)	Student-t block equi (4)	Student-t 1G (5)	Student-t dynamic (6)	Skew Student block equi (7)	Skew Student 1G (8)	Skew Student dynamic (9)
AIC	-139990	-140848	-141211	-168123	-153480	-168957	-168709	-153690	**-169362**
BIC	-139766	-140039	-140264	-167825	-152666	-167935	-168336	-152871	**-168266**
DIC	-139968	-140851	-141214	-168145	-153500	-169037	-169092	-153772	**-170008**
# params	42	152	178	56	153	192	70	154	206
a	[0.022,0.144]	0.039	[0.023,0.140]	[0.017,0.078]	0.026	[0.023,0.135]	[0.017,0.062]	0.025	[0.023,0.132]
b	[0.793,0.998]	0.982	[0.611,0.996]	[0.901,0.999]	0.993	[0.585,0.997]	[0.924,0.999]	0.993	[0.597,0.997]
ν				[7.734,20.802]	35.606	[7.766,21.253]	[8.242,35.468]	34.234	[8.121,35.857]
γ							[-1.678,-0.058]	-0.378	[-1.689,-0.036]
f_0	[1.303,1.767]	[0.935,2.439]	[0.913,2.476]	[1.366,1.834]	[0.953,2.521]	[0.933,2.520]	[1.361,1.825]	[0.949,2.519]	[0.907,2.518]

Discussion and Extension

- The more complex copula functions 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 hyperbolic skew Student-t copula
- Taking into account more factors
- Using GPU to fasten the inference process.

References

[1] D. D. Creal and R. S. Tsay, "High dimensional dynamic stochastic copula models," *Journal of Econometrics*, 2015.
[2] P. Krupskii and H. Joe, "Factor copula models for multivariate data," *Journal of Multivariate Analysis*, 2013.
[3] D. H. Oh and A. J. Patton, "Time-varying systemic risk: Evidence from a dynamic copula model of cds spreads," *J. Bus Econ Stat*, 2016.
[4] D. Creal, S. J. Koopman, and A. Lucas, "Generalized autoregressive score models with applications," *J. of Applied Econometrics*, 2013.