Bayesian Modeling of Dependence with Copulas

Feng Li

School of Statistics and Mathematics
Central University of Finance and Economics
If I should ever have a tattoo, that would be

\[ p(\theta|y) \propto p(y|\theta)p(\theta). \]
The stock market returns

SP100

SP600

Time
Our interests

- We would like to construct a multivariate model that some margins are continuous but some are discrete.
- We would like to dynamically model the rank correlations:
  \[ \tau = 4 \int \int F(x_1, x_2) dF(x_1, x_2) - 1. \]
- As well as the tail dependences
  \[ \lambda_L = \lim_{u \to 0^+} \Pr(X_1 < F_1^{-1}(u)|X_2 < F_2^{-1}(u)) \]
- We will see why we can not do it (easily) in the frequentist approach.
• The word “copula” means **linking**.

• **Sklar’s theorem**

Let $H$ be a multi-dimensional distribution function with marginal distribution functions $F_1(x_1), \ldots, F_m(x_m)$. Then there exists a function $C$ (**copula function**) such that

$$H(x_1, \ldots, x_m) = C(F_1(x_1), \ldots, F_m(x_m))$$

$$= C \left( \int_{-\infty}^{x_1} f(z_1) \, dz_1, \ldots, \int_{-\infty}^{x_m} f(z_m) \, dz_m \right) = C(u_1, \ldots, u_m).$$

Furthermore, if $F_i(x_i)$ are continuous, then $C$ is unique, and the derivative

$$c(u_1, \ldots, u_m) = \frac{\partial^m C(u_1, \ldots, u_m)}{\partial u_1 \cdots \partial u_m}$$

is the **copula density**.
The covariate-contingent copula model

The Joe-Clayton copula

- The Joe-Clayton copula function

\[ C(u, v, \theta, \delta) = 1 - \left[ 1 - \left\{ (1 - \bar{u}^\theta)^{-\delta} + (1 - \bar{v}^\theta)^{-\delta} - 1 \right\}^{-1/\delta} \right]^{1/\theta} \]

where \( \theta \geq 1, \delta > 0, \bar{u} = 1 - u, \bar{v} = 1 - v \).

- Our interests:
  - The rank correlation and tail dependence in the model.
  - The convenience for interpretation (knowing the underlying factors of dependences).

- We use the reparameterized copula \( C(u, v, \lambda_L, \tau) = C(u, v, \theta, \delta) \).

* **Note!** any other copulas can be equally well used.
The covariate-contingent copula model

The model

• The marginal models
  • In principle, any combination of univariate marginal models can be used.
  • In the continuous case, we use univariate model that each margin is from the student t distribution.

• The log likelihood

\[
\log p(\{\beta, I\}|y, x) = \text{constant} + \sum_{j=1}^{M} \log p(y_j|\{\beta, I\}_j, x_j) \\
+ \log \mathcal{L}_C(u|\{\beta, I\}_C, y, x) + \log p(\{\beta, I\})
\]

where all the parameters are connected with covariates via link function \(\varphi(\cdot)\), (identity, log, logit, probit,...)

Marginal features
\[
\mu = \varphi_{\beta_u}^{-1}(X_u \beta_u), \quad \sigma^2 = \varphi_{\beta_\sigma}^{-1}(X_\sigma \beta_\sigma), \ldots
\]

Copula features
\[
\lambda_L = \varphi_{\lambda}^{-1}((X_u, X_v) \beta_\lambda), \quad \tau = \varphi_{\tau}^{-1}((X_u, X_v) \beta_\tau).
\]
The covariate-contingent copula model

The Bayesian approach

- The priors
  - The priors for the copula functions are easy to specify due to our reparameterization.
  - The priors for the marginal distributions are specified in their usual ways.
  - When variable selection is used, we assume there are no covariates in the link functions \textit{a priori}.
- The posterior

\[
p(\beta | \mathbf{Y}) \propto \mathcal{L}(\mathbf{Y} | \beta) \times \prod_{i \in u,v,\tau,\lambda} p(\beta_i)
\]
- The posterior inference is straightforward although the model is very complicated.
The beauty of Bayesian approach is not because of its complexity, but because of its simplicity.
The covariate-contingent copula model

Sampling the posterior with an efficient MCMC method

- We update all the parameters jointly by using Metropolis-Hastings within Gibbs.
- The proposal density for each parameter vector $\beta$ is a multivariate $t$-density with $\text{df} > 2$,

$$
\beta_p | \beta_c \sim \text{MVT} \left[ \hat{\beta}, - \left( \frac{\partial^2 \ln p(\beta|Y)}{\partial \beta \partial \beta'} \right)^{-1} \right|_{\beta=\hat{\beta}}, \text{df} \right],
$$

where $\hat{\beta}$ is obtained by $R$ steps ($R \leq 3$) Newton’s iterations during the proposal with analytical gradients.
- Variable selections are carried out simultaneously.
- **The key:** The analytical gradients require the derivative for the copula density and marginal densities.
The stock returns, a revisit
The tail-dependence and Kendall’s $\tau$ over time (posterior mode)
**Table:** Posterior summary of copula model with S&P100 and S&P600 data. In the copula component part, the first row and second row for $\beta$ and $I$ are the results for the combined covariates that are used in the first and second marginal model, respectively. The intercept are always included in the model.

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>RM1</th>
<th>RM5</th>
<th>RM20</th>
<th>CloseAbs95</th>
<th>CloseAbs80</th>
<th>MaxMin95</th>
<th>MaxMin80</th>
<th>CloseSqr95</th>
<th>CloseSqr80</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Copula component (C)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{\lambda_L}$</td>
<td>$-8.165$</td>
<td>$-0.555$</td>
<td>$1.793$</td>
<td>$0.005$</td>
<td>$-0.170$</td>
<td>$0.110$</td>
<td>$-0.667$</td>
<td>$-1.448$</td>
<td>$-0.636$</td>
<td>$0.050$</td>
</tr>
<tr>
<td></td>
<td>$1.463$</td>
<td>$0.405$</td>
<td>$0.934$</td>
<td>$-2.138$</td>
<td>$-1.288$</td>
<td>$-1.954$</td>
<td>$-1.577$</td>
<td>$-1.873$</td>
<td>$-1.805$</td>
<td></td>
</tr>
<tr>
<td>$I_{\lambda_L}$</td>
<td>$1.00$</td>
<td>$0.98$</td>
<td>$0.37$</td>
<td>$0.63$</td>
<td>$0.02$</td>
<td>$0.61$</td>
<td>$0.36$</td>
<td>$0.35$</td>
<td>$0.39$</td>
<td>$0.29$</td>
</tr>
<tr>
<td></td>
<td>$1.00$</td>
<td>$1.00$</td>
<td>$0.00$</td>
<td>$0.30$</td>
<td>$0.35$</td>
<td>$0.40$</td>
<td>$0.00$</td>
<td>$0.61$</td>
<td>$0.34$</td>
<td></td>
</tr>
<tr>
<td>$\beta_{\tau}$</td>
<td>$-1.726$</td>
<td>$0.181$</td>
<td>$-0.217$</td>
<td>$-0.304$</td>
<td>$-0.107$</td>
<td>$0.115$</td>
<td>$0.005$</td>
<td>$-0.257$</td>
<td>$1.068$</td>
<td>$0.037$</td>
</tr>
<tr>
<td></td>
<td>$-0.191$</td>
<td>$0.170$</td>
<td>$0.274$</td>
<td>$0.144$</td>
<td>$-0.051$</td>
<td>$-0.671$</td>
<td>$0.059$</td>
<td>$-0.209$</td>
<td>$-0.181$</td>
<td></td>
</tr>
<tr>
<td>$I_{\tau}$</td>
<td>$1.00$</td>
<td>$0.00$</td>
<td>$0.00$</td>
<td>$0.00$</td>
<td>$0.00$</td>
<td>$0.90$</td>
<td>$0.99$</td>
<td>$1.00$</td>
<td>$0.85$</td>
<td>$0.00$</td>
</tr>
<tr>
<td></td>
<td>$1.00$</td>
<td>$1.00$</td>
<td>$0.00$</td>
<td>$0.00$</td>
<td>$1.00$</td>
<td>$1.00$</td>
<td>$0.00$</td>
<td>$0.00$</td>
<td>$0.00$</td>
<td>$0.00$</td>
</tr>
</tbody>
</table>

The inefficiency factors for the parameters are all below 25.
Extensions and future work

- Our bivariate tail-dependence method can be other higher-order multivariate models.
- Mixtures of copulas.
- Modeling multivariate volatility surface with copulas.
- Our copula method is general and can also be applied to other areas, e.g. optimal design for multivariate data.
Can we have a model that is big like an elephant?

by John Godfrey Saxe (1816-1887)
Knowing the elephant

• Sophisticated models are essential for such situations.
• In principle, the complicated model should be able to capture more complicated data features.
• Estimating such model is not easy.
• There is huge space to explore.
  • Techniques like parallel MCMC should be explored to speed up the computation.
  • Statistics with big data is the new challenge.
Thank you!

feng.li@cufe.edu.cn

http://feng.li/