Modeling covariate-contingent correlation and tail dependence with copulas

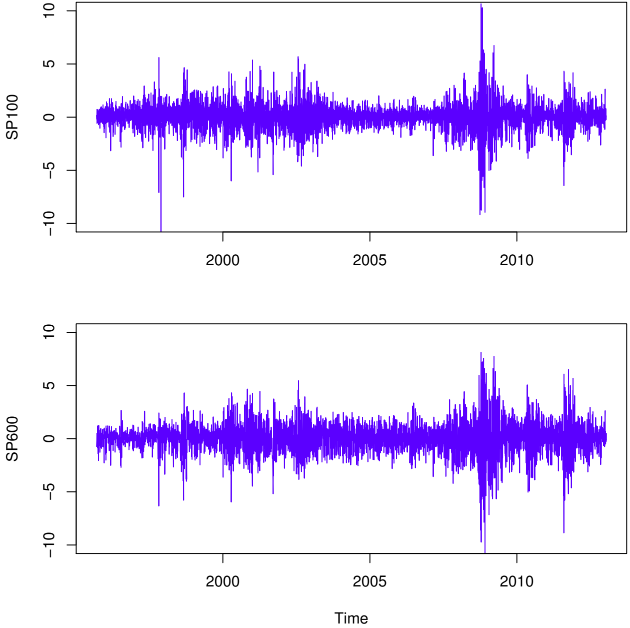
Feng Li

School of Statistics and Mathematics
Central University of Finance and Economics

Outline of the talk

1. A financial story
2. Measuring correlation and tail dependence with copulas
3. The covariate-contingent copula model
4. Extensions

The stock market returns



Our interests

- We would like to construct a multivariate model that some margins are continuous but some are discrete.
  - One margin: A company's stock credited as A, A^+ over time by Standard & Poor's.
  - The other margin: the stock returns over time
- In the *big data* world: we would like to estimate a very heavy multivariate model in such way that
  1. Independently build each marginal model. Parallel them!
  2. Build the multivariate dependences on top of the margins.

- The word “copula” means **linking**.
- **Sklar’s theorem**
  Let H be a multi-dimensional distribution function with marginal distribution functions $F_1(x_1), ..., F_m(x_m)$. Then there exists a function C (**copula function**) such that

$$\begin{aligned} H(x_1, ..., x_m) =& C(F_1(x_1), ..., F_m(x_m)) \\ =& C\left(\int_{-\infty}^{x_1} f(z_1)dz_1, ..., \int_{-\infty}^{x_m} f(z_m)dz_m\right) = C(u_1, ..., u_m). \end{aligned}$$

  Furthermore, if $F_i(x_i)$ are continuous, then C is unique, and the derivative $c(u_1, ..., u_m) = \partial^m C(u_1, ..., u_m)/(\partial u_1 ... \partial u_m)$ is the **copula density**.

Introduction to copulas

↪ Some arbitrary examples

- If $X_1, \ldots, X_m$ are independent, and iff C is a product copula, then

$$C(F_1(x_1), \ldots, F_m(x_m)) = \prod_{i=1}^{m} F_i(x_i)$$

- The bivariate Gaussian copula

$$C(u_1, u_2, \rho) = \Phi_2(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \rho)$$
$$= \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{z_1^2 - 2\rho z_1 z_2 + z_2^2}{2(1-\rho^2)}\right\} dz_1 dz_2$$

- The multivariate probit model is a simple example of a Gaussian copula, with univariate probit regressions as the marginals.
- There are many ways to construct a copula youself.

Measuring correlation and tail dependence

↪ Kendall's τ and tail-dependences

- The **Kendall's** τ can be written in terms of copula function:

$$\tau = 4\int\int F(x_1, x_2) dF(x_1, x_2) - 1 = 4\int\int C(u_1, u_2) dC(u_1, u_2) - 1.$$

- As well as the bivariate lower and upper **tail dependences**

$$\lambda_L = \lim_{u \to 0^+} \Pr(X_1 < F_1^{-1}(u) | X_2 < F_2^{-1}(u)) = \lim_{u \to 0^+} \frac{C(u, u)}{u},$$
$$\lambda_U = \lim_{u \to 1^-} \Pr(X_1 > F_1^{-1}(u) | X_2 > F_2^{-1}(u)) = \lim_{u \to 1^-} \frac{1 - C(u, u)}{1 - u}.$$

- Some facts:
  - The Kendall's τ is invariant w.r.t. **strictly** increasing transformations.
  - For all copulas in the elliptical class (Gaussian, *t*,...), $\tau = \frac{2}{\pi} \arcsin(\rho)$.
  - The Gaussian copula has zero tail dependence.
  - The student t copula has asmptotic upper tail dependence even for negative and zero correlations. The tail dependence decreases when degrees of freedom increases.

The covariate-contingent copula model

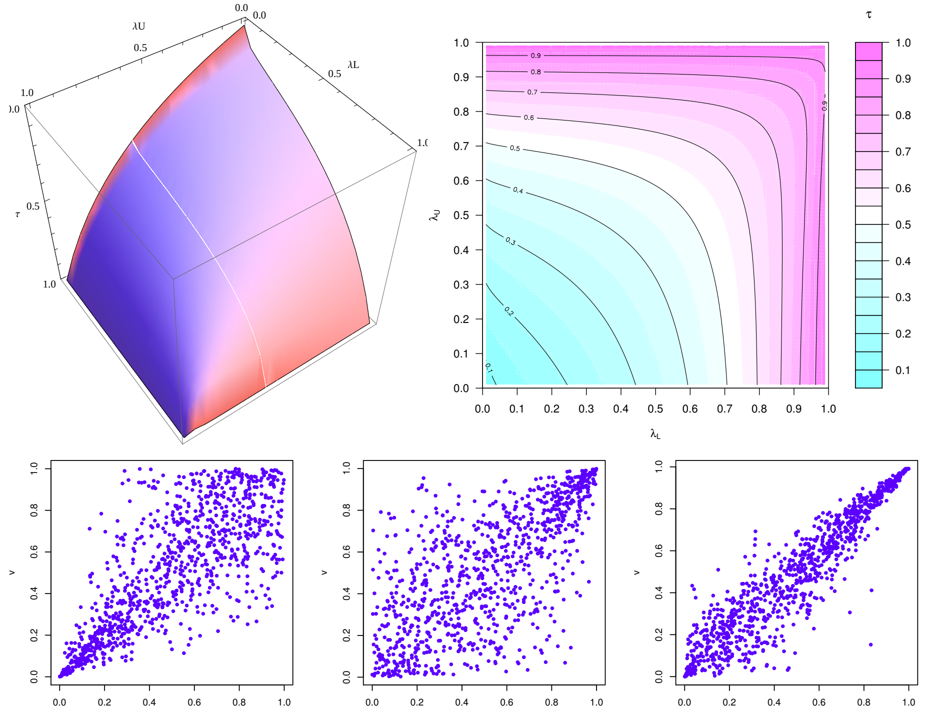
The Joe-Clayton copula

- The Joe-Clayton copula function

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

  where $\theta \geqslant 1$, $\delta > 0$, $\bar{u} = 1 - u$, $\bar{v} = 1 - v$.
- Some properties:
  - One type of Archimedean copula.
  - $\lambda_L = 2^{-1/\delta}$ does not depend on $\lambda_U = 2 - 2^{-1/\theta}$.
  - $\tau = 1 - 4\int_0^\infty s \times (\varphi'(s))^2 ds$ is calculated via Laplace transform.
- 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.


λU
λL
τ
0.0
0.5
1.0
λ_U
λ_L
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
v

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**

$$
\begin{aligned}
\log p(\{\beta, \mathcal{I}\}|y, x) =& \text{constant} + \sum_{j=1}^{M} \log p(y_{.j}|\{\beta, \mathcal{I}\}_j, x_j) \\
& + \log \mathcal{L}_C(u|\{\beta, \mathcal{I}\}_C, y, x) + \log p(\{\beta, \mathcal{I}\})
\end{aligned}
$$

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

$$
\begin{array}{lll}
\text{Marginal features} & \mu = \varphi_{\beta_u}^{-1}(X_u \beta_u), & \sigma^2 = \varphi_{\beta_\sigma}^{-1}(X_\sigma \beta_\sigma), \ldots \\
\text{Copula features} & \lambda_L = \varphi_{\lambda}^{-1}((X_u, X_v)\beta_\lambda), & \tau = \varphi_{\tau}^{-1}((X_u, X_v)\beta_\tau).
\end{array}
$$

- 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 *a priori*.
- The posterior

$$p(\boldsymbol{\beta}|\mathbf{Y}) \propto \mathcal{L}(\mathbf{Y}|\beta) \times \prod_{i \in u,v,\tau,\lambda} p(\beta_i)$$

- The posterior inference is straightforward although the model is very complicated.

The covariate-contingent copula model

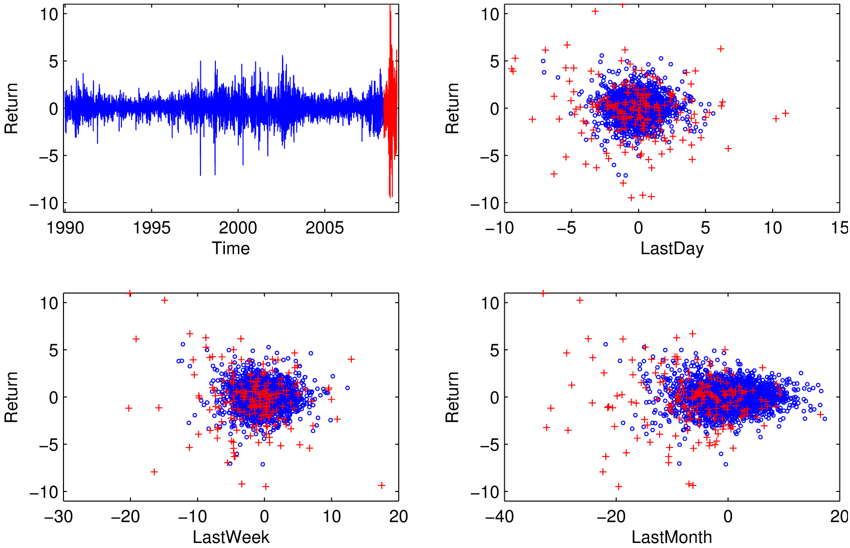
↪ Sampling the posterior with an efficient MCMC method

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

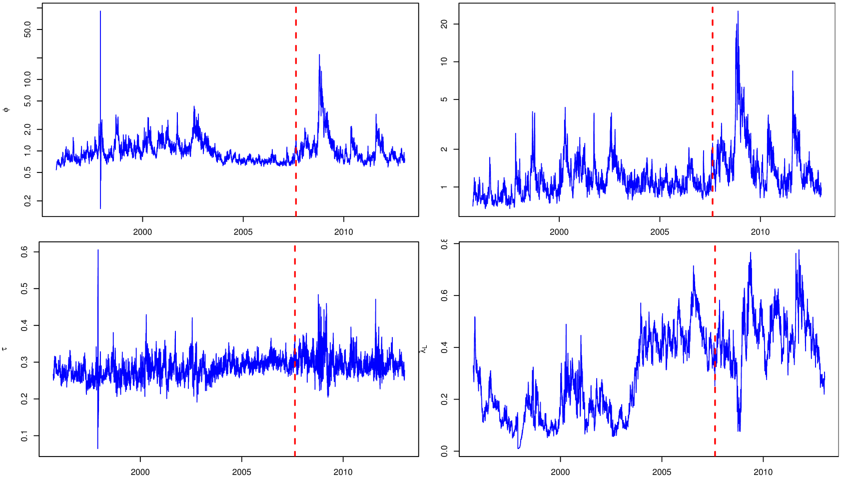
$$\boldsymbol{\beta}_p | \boldsymbol{\beta}_c \sim \mathbf{MVT}\left[\hat{\boldsymbol{\beta}},\ -\left(\frac{\partial^2 \ln p(\boldsymbol{\beta}|\mathbf{Y})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'}\right)^{-1}\Bigg|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}},\ \mathrm{df}\right],$$

  where $\hat{\boldsymbol{\beta}}$ is obtained by R steps ($R \leqslant 3$) 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


Return
10
5
0
−5
−10
1990
1995
2000
2005
Time
Return
−10
−5
0
5
10
15
LastDay
Return
−30
−20
−10
0
10
20
LastWeek
Return
−40
−20
0
20
LastMonth

The tail-dependence and Kendall's τ over time



The posterior copula plot

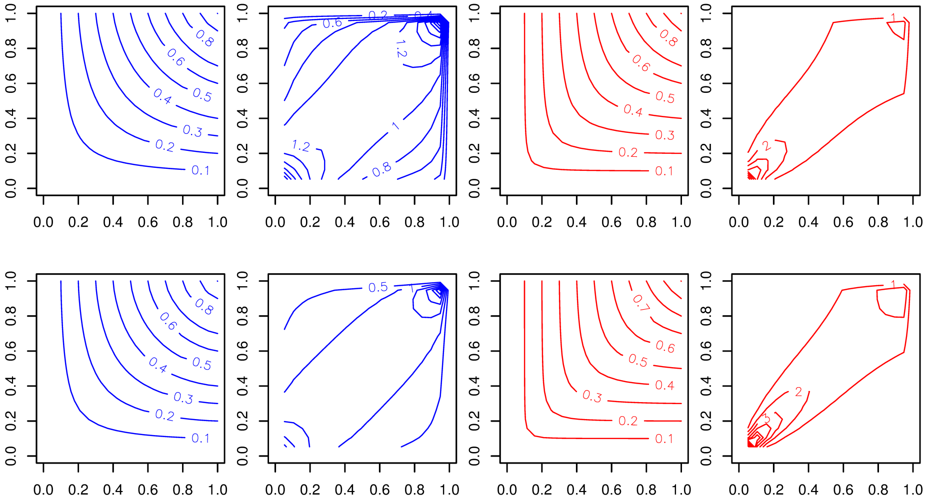


Table: Posterior summary of copula model with S&P100 and S&P600 data. In the copula component part, the first row and second row for β and $\mathcal{I}$ are the results for the combined covariates that are used in the first and second marginal model, respectively. The intercept are always included in the model.

	Intercept	RM1	RM5	RM20	CloseAbs95	CloseAbs80	MaxMin95	MaxMin80	CloseSqr95	CloseSqr80
					Copula component ($\mathcal{C}$)					
β_{λ_L}	−8.165	**−0.555**	1.793	**0.005**	−0.170	**0.110**	−0.667	−1.448	−0.636	0.050
		1.463	**0.405**	0.934	−2.138	−1.288	−1.954	−1.577	**−1.873**	−1.805
$\mathcal{I}_{\lambda_L}$	1.00	**0.98**	0.37	**0.63**	0.02	**0.61**	0.36	0.35	0.39	0.29
		1.00	**1.00**	0.00	0.30	0.35	0.40	0.00	**0.61**	0.34
β_{τ}	−1.726	0.181	−0.217	−0.304	−0.107	**0.115**	**0.005**	**−0.257**	**1.068**	0.037
		−0.191	**0.170**	0.274	0.144	**−0.051**	**−0.671**	0.059	−0.209	−0.181
$\mathcal{I}_{\tau}$	1.00	0.00	0.00	0.00	0.00	**0.90**	**0.99**	**1.00**	**0.85**	0.00
		1.00	**1.00**	0.00	0.00	**1.00**	**1.00**	0.00	0.00	0.00

The inefficiency factors for the parameters are all bellow 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.

Thank you!

feng.li@cufe.edu.cn

http://feng.li/