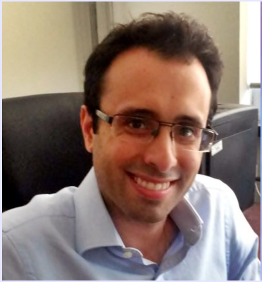
Forecasting with covariate-dependent copulas

Feng Li

School of Statistics and Mathematics
Central University of Finance and Economics

Collaborators



Anastasios Panagiotelis (Monash Uni.)

Yanfei Kang (Beihang Uni.)

Zhuojing He (CUFE)

Outline

- 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**.

- 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**

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

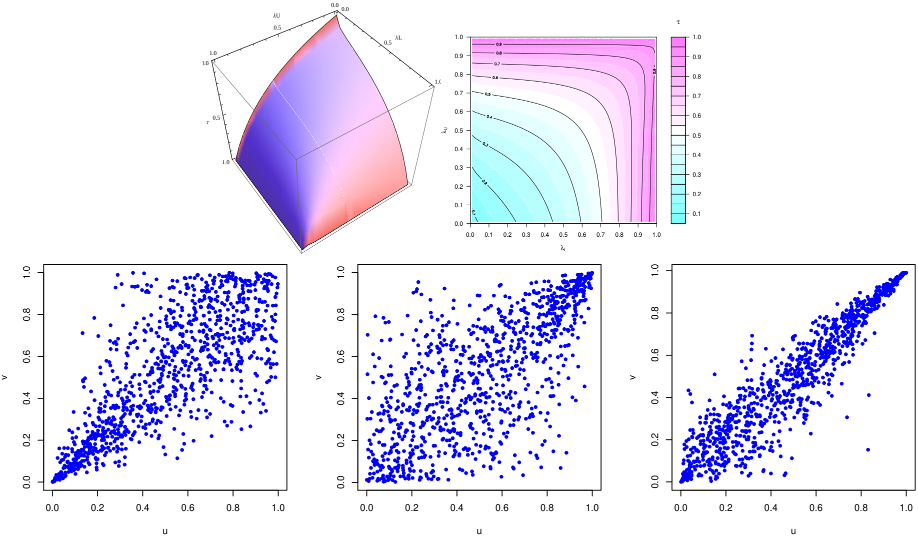
↳ The Joe-Clayton copula example

- 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:
  - $\lambda_L = 2^{-1/\delta}$ 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.



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

Covariate-dependent copula models

↳ The reparameterized copula model

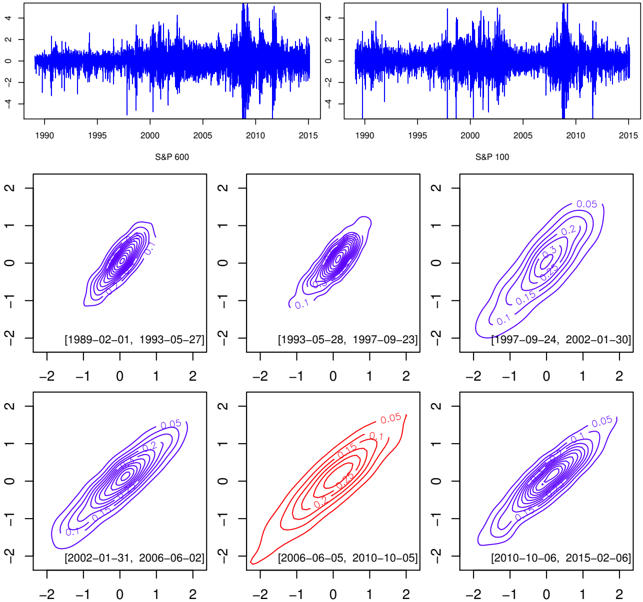
- **Reparametrization**: We reparameterize the copula as a function of tail-dependence and/or Kendall's tau $C(\mathbf{u}, \lambda_L, \tau)$.
- **Link with covariates**: All copula features in the k:th and l:th margins can be connected with covariates

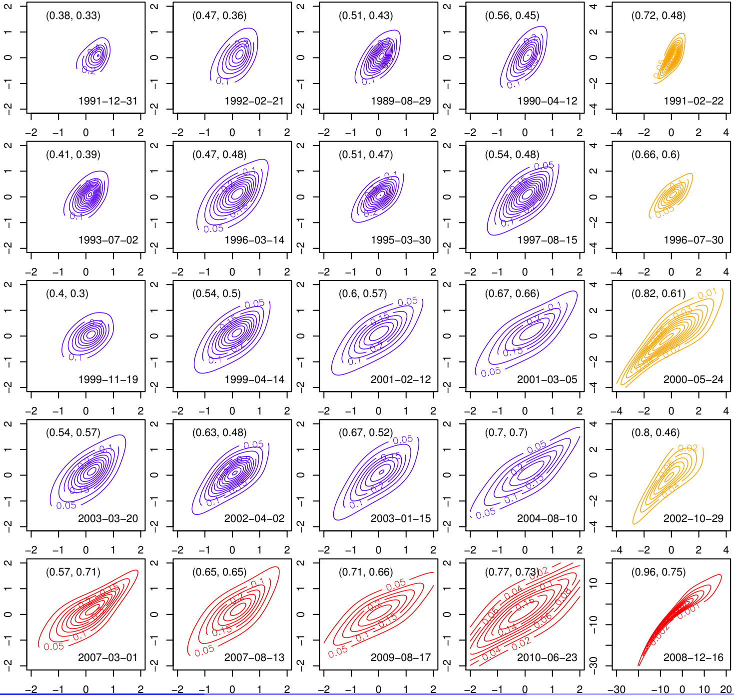
$$\tau_{kl} = l_{\tau}^{-1}(\mathbf{X}_{kl}\boldsymbol{\beta}_{\tau}),$$
$$\lambda_{kl} = l_{\lambda}^{-1}(\mathbf{X}_{kl}\boldsymbol{\beta}_{\lambda})$$

- **Applicable Copulas**: Any copula can be equally well used with such reparameterization when there is closed form of tail-dependence and Kendall's τ.
  - **Archimedean copulas**: Joe-Clayton, Clayton, Gumbel,...
  - **Elliptical copulas**: Bivariate or multivariate Gaussian and *t* copulas
- Marginal models we have used
  - Mixture of asymmetric student's-*t* distributions (Li, Villani, and Kohn, 2011).
  - Dynamic conditional correlation GARCH models (Engle, 2002)
  - stochastic volatility (SV) models (Kastner and Frühwirth-Schnatter, 2014).
  - Poisson regression models.

Covariate-dependent copula models

↳ Why introducing complications? Evidence from SP100 & SP600





(0.38, 0.33)
1991–12–31
(0.47, 0.36)
1992–02–21
(0.51, 0.43)
1989–08–29
(0.56, 0.45)
1990–04–12
(0.72, 0.48)
1991–02–22
(0.41, 0.39)
1993–07–02
(0.47, 0.48)
1996–03–14
(0.51, 0.47)
1995–03–30
(0.54, 0.48)
1997–08–15
(0.66, 0.6)
1996–07–30
(0.4, 0.3)
1999–11–19
(0.54, 0.5)
1999–04–14
(0.6, 0.57)
2001–02–12
(0.67, 0.66)
2001–03–05
(0.82, 0.61)
2000–05–24
(0.54, 0.57)
2003–03–20
(0.63, 0.48)
2002–04–02
(0.67, 0.52)
2003–01–15
(0.7, 0.7)
2004–08–10
(0.8, 0.46)
2002–10–29
(0.57, 0.71)
2007–03–01
(0.65, 0.65)
2007–08–13
(0.71, 0.66)
2009–08–17
(0.77, 0.73)
2010–06–23
(0.96, 0.75)
2008–12–16

- Any multivariate density can be decomposed into a product of densities. For example,

$$f(\mathbf{u}) = f(u_1|u_2, \ldots, u_d)f(u_2|u_3, \ldots, u_d) \ldots f(u_{d-1}|u_d)f(u_d)\,,$$

- For example the so called canonical vine (C-vine) decomposition of a trivariate density would be:

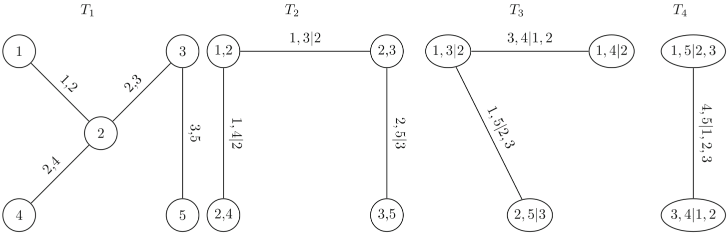
$$f_{123}(u_1, u_2, u_3) = c_{23|1}(F_{2|1}(u_2|u_1), F_{3|1}(u_3|u_1))c_{12}(u_1, u_2)c_{13}(u_1, u_3)\,,$$

- Not all vines can be used to obtain a pair copula construction, and a vine structure that is consistent with a valid pair copula construction is known as a regular vine or R-vine.
- We extend our work to R-vine that M variables is a vine in which two edges in tree j are joined by an edge in tree $j + 1$ only if these edges share a common node.

Covariate-dependent copula models
↳ Extending to Vine structure

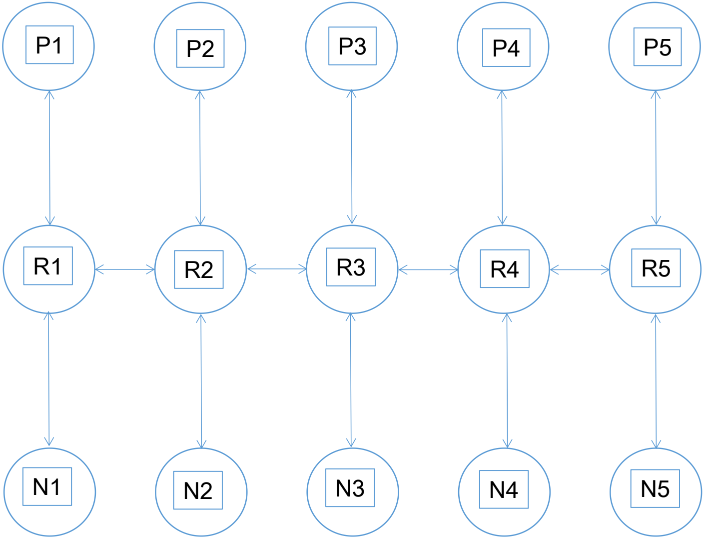
- To extend the bivariate covariate-dependent copula model of Li and Kang (2018) to vine copulas we simple allow all pair copula parameters including those that describe conditional dependence to be functions of linear predictors.
- For the simplicity purpose, we keep the vine structure to be fixed in our application. But algorithms that can quickly select a vine copula model to accurately captures features in the dependence structure are possible to adapt (Panagiotelis, Czado, Joe, and Stöber, 2017).
- Our framework allows for margins to be combined from discrete and continuous variables.

↳ An R-vine tree sequence in 5 dimensions



Covariate-dependent copula models

↳ An R-vine for stock market returns R_i and texts sentiments (P_i, N_i)



- The log likelihood for the vine copula in **tree** i part is

$$\sum_{e \in E_i} \log \left(c_{j(e),k(e)|D(e)}(F_{j(e)|D(e)}(x_{j(e)}|\boldsymbol{x}_{D(e)}), F_{k(e)|D(e)}(x_{k(e)}|\boldsymbol{x}_{D(e)})|\theta_e) \right)$$

where

$$\begin{aligned} F(x_{j(e)}|\boldsymbol{x}_{D(e)}) &= \frac{\partial C_{j(e),j'(e)|D(e)\setminus j'(e)}(F(x_{j(e)}|\boldsymbol{x}_{D(e)\setminus j'(e)}), F(x_{j'(e)}|\boldsymbol{x}_{D(e)\setminus j'(e)}))}{\partial F(x_{j'(e)}|\boldsymbol{x}_{D(e)\setminus j'(e)})} \\ &=: h_{j(e),j'(e)|D(e)\setminus j'(e)}(F(x_{j(e)}|\boldsymbol{x}_{D(e)\setminus j'(e)}), F(x_{j'(e)}|\boldsymbol{x}_{D(e)\setminus j'(e)})), \end{aligned}$$

- We also want gradients

Covariate-dependent copula models II

↳ The formulas under the Vine structure

$$
\begin{aligned}
\frac{\partial}{\partial \theta} \ln \left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}),\, F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta\right)\right) &= \frac{\frac{\partial}{\partial \theta}\left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}),\, F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta\right)\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}),\, F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta\right)} \\
&= \frac{\partial_\theta c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}),\, F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}),\, F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta\right)}.
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial \theta} \ln \left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),\, F_{V|\mathbf{Z}}(v|\mathbf{z})\right)\right) &= \frac{\frac{\partial c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),\, F_{V|\mathbf{Z}}(v|\mathbf{z})\right)}{\partial F_{U|\mathbf{Z}}(u|\mathbf{z},\theta)}}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),\, F_{V|\mathbf{Z}}(v|\mathbf{z})\right)} \times \frac{\partial}{\partial \theta} F_{U|\mathbf{Z}}(u|\mathbf{z},\theta) \\
&= \frac{\partial_1 c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),\, F_{V|\mathbf{Z}}(v|\mathbf{z})\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),\, F_{V|\mathbf{Z}}(v|\mathbf{z})\right)} \times \frac{\partial}{\partial \theta} F_{U|\mathbf{Z}}(u|\mathbf{z},\theta).
\end{aligned}
$$

Covariate-dependent copula models III

↳ The formulas under the Vine structure

$$\frac{\partial}{\partial\theta}\ln\left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),\, F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)\right)\right) = \frac{\partial_1 c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),\, F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),\, F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)\right)} \times \frac{\partial}{\partial\theta} F_{U|\mathbf{Z}}(u|\mathbf{z},\theta) + \frac{\partial_2 c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),\, F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),\, F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)\right)} \times \frac{\partial}{\partial\theta} F_{V|\mathbf{Z}}(v|\mathbf{z},\theta).$$

$$\begin{aligned}
\frac{\partial}{\partial\theta} F_{U|V,\mathbf{Z}}(u|v,\mathbf{z},\theta) &= \frac{\partial}{\partial\theta}\left(h_{U|V,\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),\, F_{V|\mathbf{Z}}(v|\mathbf{z})\right)\right) \\
&= \partial_1 h_{U|V,\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),\, F_{V|\mathbf{Z}}(v|\mathbf{z})\right) \times \frac{\partial}{\partial\theta} F_{U|\mathbf{Z}}(u|\mathbf{z},\theta) \\
&= c_{U|V,\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),\, F_{V|\mathbf{Z}}(v|\mathbf{z})\right) \times \frac{\partial}{\partial\theta} F_{U|\mathbf{Z}}(u|\mathbf{z},\theta)
\end{aligned}$$

and

$$\frac{\partial}{\partial\theta} h_{U|V,\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}),\, F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)\right) = \partial_2 h_{U|V,\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}),\, F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)\right) \times \frac{\partial}{\partial\theta} F_{V|\mathbf{Z}}(v|\mathbf{z},\theta).$$

Covariate-dependent copula models IV
↳ The formulas under the Vine structure

- More equations available at Stöber and Schepsmeier (2013) and our working paper (Li, Panagiotelis & Kang, 2019).
- A nice C++ library with R/Python interface handling these equations well (but without covariate-dependent structure) written by Nagler et al., (2019) https://github.com/vinecopulib/
- Our `cdcopula` package (*in preparation*) allows for covariate-dependent structures.

- **The log posterior**

$$\log p(\{\beta, \mathcal{I}\}|y, x) = c + \sum_{j=1}^{M} \{\log p(y_{.j}|\{\beta, \mathcal{I}\}_j, x_j) + \log p(\{\beta, \mathcal{I}_j\})\}$$
$$+ \log \mathcal{L}_C(u_{1:M}|\{\beta, \mathcal{I}\}_C, y, x) + \log p_C(\{\beta, \mathcal{I}\})$$

where

- $\{\beta\}$ are the coefficient in the linking function,
- $\{\mathcal{I}\}$ are the corresponding variable selection indicators.
- $\{\beta, \mathcal{I}\}$ can be estimated jointly via Bayesian approach.
- $u_j = F_j(y_j)$ is the CDF of the j:th marginal model.

The Bayesian framework
↳ The prior specification

- **The priors** for the copula model are easy to specify due to our reparameterization.
  - It it **not easy** to specify priors directly on $\{\boldsymbol{\beta}, \mathcal{J}\}$
  - But it is **easy** to puts prior information on the model parameters features (τ, μ, σ^2) and then derive the implied prior on the intercepts and variable selection indicators.
- **Spike-and-slab** prior for variable selection.
  - The prior for each variable selection indicator is identically distributed as $\mathrm{Bern}(p)$ where p is the hyperparameter.
  - Note that when the number of covariates is large, the independent Bernoulli prior can be very informative (Yau, Kohn, and Wood, 2003).
  - We also implement the recommended Beta prior (Scott, Berger, et al., 2010) where the uniform prior is a special case.
  - Our comparison in simulation studies show that the difference between independent Bernoulli prior and uniform prior is negligible with less than twenty covariates for each feature.

The MCMC scheme
↳ Metropolis-Hastings within Gibbs

- We update all the parameters **jointly** by using tailored Metropolis-Hastings within Gibbs. This is more efficient compared to the two-stage inference according to our study.
- **Taming the Beast:** the analytical gradients require the derivative for the copula density and marginal densities which can be conveniently decomposed via the chain rule that greatly reduces the complexity of the the gradient calculation.
- **Bayesian variable selection** is carried out simultaneously.
- The Gibbs sampler for covariate-dependent copula.

Margin component (1)	...	Margin component (M)	Copula component (C)
(1.1) $\{\beta_\mu, \mathcal{I}_\mu\}_1 \vert \{\beta_\mu, \mathcal{I}_\mu\}_{-1}$	...	(M.1) $\{\beta_\mu, \mathcal{I}_\mu\}_M \vert \{\beta_\mu, \mathcal{I}_\mu\}_{-M}$	(C.1) $\{\beta_\lambda, \mathcal{I}_\lambda\}_C \vert \{\beta_\lambda, \mathcal{I}_\lambda\}_{-C}$
(1.2) $\{\beta_\phi, \mathcal{I}_\phi\}_1 \vert \{\beta_\phi, \mathcal{I}_\phi\}_{-1}$	...	(M.2) $\{\beta_\phi, \mathcal{I}_\phi\}_M \vert \{\beta_\phi, \mathcal{I}_\phi\}_{-M}$	(C.2) $\{\beta_\lambda, \mathcal{I}_\lambda\}_C \vert \{\beta_\lambda, \mathcal{I}_\lambda\}_{-C}$
(1.3) $\{\beta_\nu, \mathcal{I}_\nu\}_1 \vert \{\beta_\nu, \mathcal{I}_\nu\}_{-1}$	...	(M.3) $\{\beta_\nu, \mathcal{I}_\nu\}_M \vert \{\beta_\nu, \mathcal{I}_\nu\}_{-M}$	...
(1.4) $\{\beta_\kappa, \mathcal{I}_\kappa\}_1 \vert \{\beta_\kappa, \mathcal{I}_\kappa\}_{-1}$	...	(M.4) $\{\beta_\kappa, \mathcal{I}_\kappa\}_M \vert \{\beta_\kappa, \mathcal{I}_\kappa\}_{-M}$	(C.ν) $\{\beta_\lambda, \mathcal{I}_\lambda\}_C \vert \{\beta_\lambda, \mathcal{I}_\lambda\}_{-C}$

where $\nu = M(M-1)/2$ is the number of paired conditional copulas.

The MCMC scheme

↳ Stochastic Gradient with Langevin Dynamics

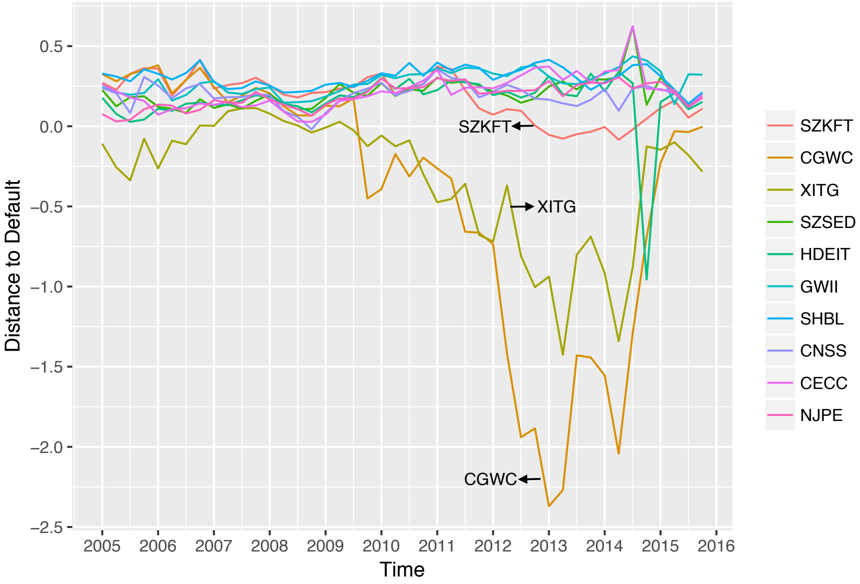
- In addition to a standard Gibbs sampler we also consider Stochastic Gradient MCMC with Langevin Dynamics algorithm (Welling and Teh, 2011).
- First denote the full set of parameters by $\theta := \left(\boldsymbol{\beta}^{(p)}, \boldsymbol{\beta}^{(c)}\right)$. The algorithm simulates values of the parameter vector as a discretization of a stochastic differential equation. Specifically, θ is updated as

$$\theta_{s+1} = \theta_s + \frac{\epsilon_s}{2}\left(\frac{\partial \log p(\theta_s)}{\partial \theta} + \frac{T}{|\mathcal{T}|}\sum_{t\in\mathcal{T}} \frac{\partial \log p(y_t|\theta)}{\partial \theta}\right) + \eta_s\,, \qquad (1)$$

- The algorithm combines the features of stochastic gradient ascent and Langevin dynamics.
- The SGLD algorithm can be generalised by multiplying both ϵ_s and η_s in equation by a matrix $M^{-1/2}$ that is chosen to account for the geometry of the posterior surface. For instance, one potential choice is to choose M as the Fisher Information matrix (Ma, Chen, and Fox, 2015).
- The SGLD could also be calibrated (or monitored) via Metropolis-Hastings rejections.
- As such, we cannot conduct variable selection via the use of indicator variables when using the SGLD algorithm.

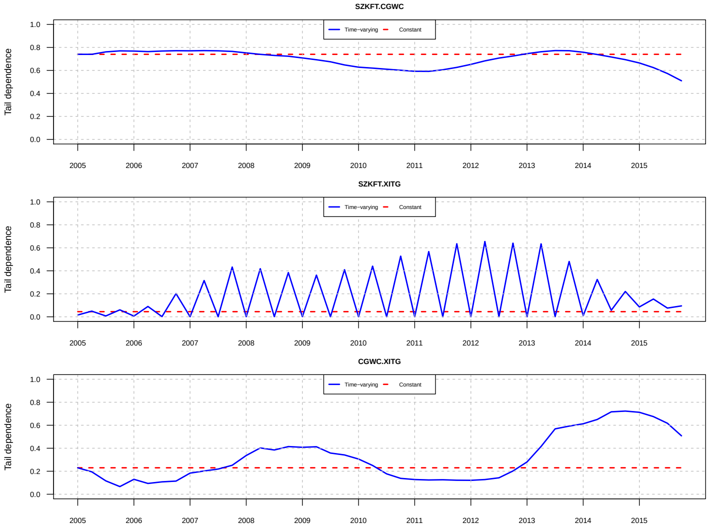
Modeling tail-dependence

↳ Determining default risk clustering factors (Li and He, 2019)

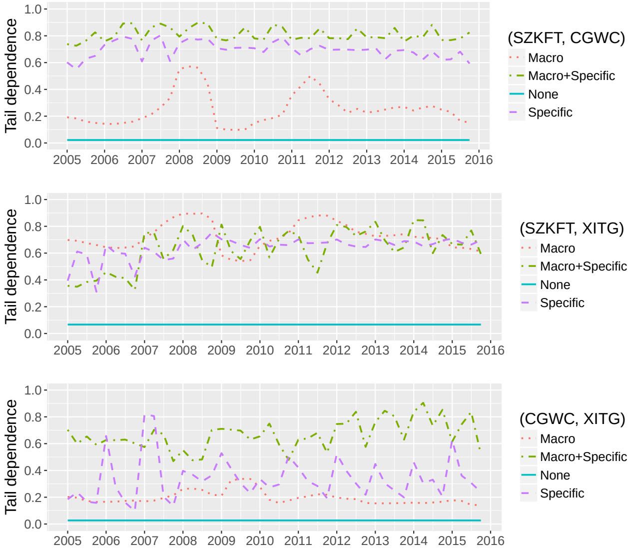


Modeling tail-dependence

↳ with time-varying copulas (Patton, 2006)



Modeling tail-dependence
↳ with covariate-dependent copulas



Forecasting evaluation

↳ Log predictive score

- We evaluating the model performance based on **out-of-sample prediction**.
- In our time series application, we estimate the model based on the 80% of historical data and then predict the last 20% data.
- We evaluate the quality of the one-step-ahead predictions using the **log predictive score** (LPS)

$$\begin{aligned} \mathrm{LPS} &= \log p(\mathbf{y}_{(T+1):(T+p)} | \mathbf{y}_{1:T}) \\ &= \sum_{i=1}^{p} \log \int p(\mathbf{y}_{T+i} | \theta, \mathbf{y}_{1:(T+i-1)}) p(\theta | \mathbf{y}_{1:(T+i-1)}) d\theta \end{aligned}$$

  where $\mathbf{y}_{a:b}$ is the dataset from time a to b and θ are the model parameters.
- The global LPS in a copula model equals to the sum of the LPS values in each marginal model and the copula component

$$\mathrm{LPS} = \sum_{i=1}^{M} \mathrm{LPS}_i + \sum_{j,k|D} \mathrm{LPS}_{j,k}$$

  which allows us to parallel compute the LPS values and compare the (conditional) contributions of different paired copulas.

Forecasting evaluation

↳ Improving forecasting performance with variable selection

- LPS comparison based on four-fold cross-validation for Joe-Clayton copula with 16 DGP settings where the **true model is covariate-independent** in all Joe-Clayton copula and split-t marginal features.
- Each dataset consists of $1,000$ observations with given mean values $\bar{\lambda}_L$ and $\bar{\lambda}_U$, and standard deviation 0.1. We estimate the Joe-Clayton copula model jointly with three modeling strategies (*CD.+VS.*, *CD.* and *Const.*).
- Efficiency improved for a misspecified model:
  - Average LPS improved due to covariate-dependent structure is **14**.
  - Overall LPS improved with covariate-dependent structure and variable selection: **19**
  - Average efficiency improvement due to variable selection: **32.5%**

DGP settings	$\bar{\lambda}_U^{(DGP)} = 0.3$			$\bar{\lambda}_U = 0.5$			$\bar{\lambda}_U = 0.7$			$\bar{\lambda}_U = 0.9$		
	CD.+VS.	*CD.*	*Const.*	*CD.+VS.*	*CD.*	*Const.*	*CD.+VS.*	*CD.*	*Const.*	*CD.+VS.*	*CD.*	*Const.*
$\bar{\lambda}_L = 0.3$	−980.97	−988.00	−960.95	−984.03	−989.06	−967.98	−973.88	−978.02	−965.53	−991.62	−996.06	−957.56
$\bar{\lambda}_L = 0.5$	−959.95	−964.87	−958.32	−994.70	−998.60	−975.69	−968.10	−977.24	−967.35	−992.07	−998.23	−962.59
$\bar{\lambda}_L = 0.7$	−975.20	−975.58	−968.03	−974.75	−981.30	−968.03	−970.17	−973.98	−958.12	−987.26	−991.37	−970.30
$\bar{\lambda}_L = 0.9$	−985.44	−989.74	−964.16	−980.20	−982.21	−967.38	−985.02	−992.05	−965.61	−971.10	−975.07	−969.30

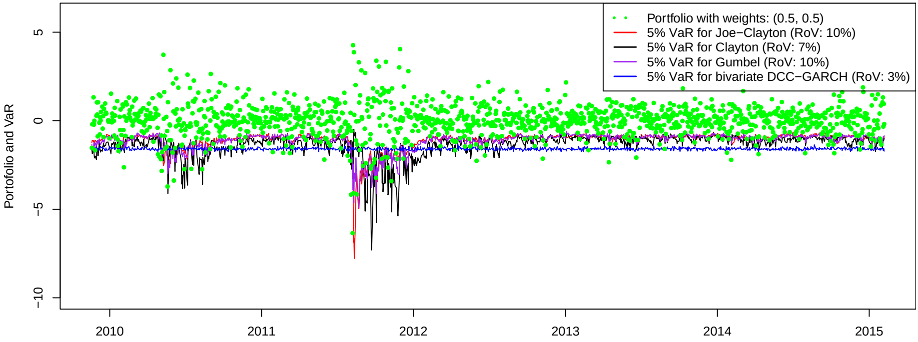
Table 4: LPS of four-fold cross-validation for Joe-Clayton copula with 16 DGP settings and 64 simulations based on different combination of lower tail-dependence and upper tail-dependence, respectively. Each dataset consists of 1,000 observations with given mean ($\bar{\lambda}_L$ and $\bar{\lambda}_U$) and standard deviation (0.1) for lower and upper tail-dependences. Each dataset is estimated with four models (*J. + CD.*, *J. + Const.*, *T.+ CD.* and *T. + Const.*) and the LPS for the best model is marked in bold.

DGP settings		$\bar{\lambda}_U^{(DGP)} = 0.3$		$\bar{\lambda}_U = 0.5$		$\bar{\lambda}_U = 0.7$		$\bar{\lambda}_U = 0.9$	
	MCMC	*CD.*	*Const.*	*CD.*	*Const.*	*CD.*	*Const.*	*CD.*	*Const.*
$\bar{\lambda}_L^{(DGP)} = 0.3$	*J.*	**−519.56**	−520.91	**−506.90**	−508.95	**−427.72**	−432.35	**−273.93**	−306.99
	T.	−523.25	−522.00	−510.60	−511.75	−444.32	−439.68	−310.67	−321.38
$\bar{\lambda}_L = 0.5$	*J.*	**−501.33**	−502.57	**−468.30**	−471.97	**−424.30**	−436.54	**−244.02**	−268.56
	T.	−510.51	−507.29	−476.68	−474.30	−446.38	−451.83	−299.08	−314.36
$\bar{\lambda}_L = 0.7$	*J.*	**−440.81**	−454.16	**−424.20**	−439.24	**−380.30**	−390.38	**−243.16**	−244.78
	T.	−457.76	−460.83	−440.01	−440.70	−397.72	−402.37	−283.96	−295.11
$\bar{\lambda}_L = 0.9$	*J.*	**−228.83**	−256.11	**−218.61**	−294.52	**−241.21**	−255.13	**−210.11**	−269.86
	T.	−244.01	-294.00	−292.74	−317.60	−280.67	−289.88	−259.15	−297.25

- We also evaluate our model's capacity in capturing tail-dependences based on the out-of-sample performance of the Value-at-Risk (VaR) estimation.
- We consider our portfolio return R_t composed by asset returns denoted as $y_{1t}, \ldots, y_{Mt}$. The portfolio return can be approximately written as:

$$R_t = \omega_1 y_{1t} + \ldots + \omega_M y_{Mt}, \text{ with } \sum_{m=1}^{M} \omega_m = 1,$$

- We approximate the out-of-sample VaR by firstly simulating the assets vector $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{Mt})'$ from the joint predictive density $p(\boldsymbol{y}_t | \boldsymbol{y}_{1:(t-1)}, \boldsymbol{x})$ and then calculating the portfolio return R_t and its $\alpha\%$ empirical quantile.



- The ratio of violations (RoV) is the percentage of sample observations lying out of the VaR critical values.
- The benchmark DCC-GARCH model has the smallest variation in both 5% VaR and 1% VaR compared to other three models, and also has the smallest RoV, which is not suitable to reflect the risk of portfolio returns with time-varying effects.
- Joe-Clayton copula and the Gumbel copula have the same RoV but the VaR for Joe-Clayton copula is more robust than the Gumbel copula.

Improving forecasting performance

↳ Joint modeling vs two-stage modeling

		Reparameterized Copulas			
		Joe-Clayton	Clayton	Gumbel	t-Copula
Margins	LPS decomposition				
		(*Joint modeling approach*)			
SPLIT-t	M_1	−1743.12	−1741.04	−1754.36	−1741.47
	M_2	−1435.98	−1468.25	−1485.68	−1430.07
	$\mathrm{C(CD.)}$	837.50	690.22	797.78	792.14
	Global	−2344.12	−2523.75	−2448.14	−2380.12
SPLIT-t	M_1	−1747.99	−1747.15	−1754.61	−1782.37
	M_2	−1434.22	−1449.95	−1446.84	−1658.09
	$\mathrm{C(Const.)}$	779.14	654.46	780.33	703.96
	Global	−2411.06	−2547.14	−2421.15	−2736.49
		(*Two-stage modeling approach*)			
SPLIT-t	M_1	−1740.10	−1741.05	−1737.73	−1741.47
	M_2	−1428.39	−1436.63	−1427.83	−1433.41
	$\mathrm{C(CD.)}$	819.63	694.84	781.39	788.22
	Global	−2346.61	−2483.93	−2392.13	−2389.41
GARCH(1,1)	M_1	−1948.07	−1948.07	−1948.07	−1948.07
	M_2	−1673.85	−1673.85	−1673.85	−1673.85
	$\mathrm{C(CD.)}$	702.35	530.48	810.39	791.55
	global	−2919.57	-3091.44	−2811.53	−2830.37
SV	M_1	−2166.90	−2154.18	−2168.17	−2179.36
	M_2	−1811.36	−1844.57	−1808.61	−1808.24
	$\mathrm{C(CD.)}$	964.37	698.30	1012.10	1053.19
	Global	−3013.90	−3300.46	−2964.68	−2934.40

Working in progress

- Improve the computational efficiency of the covariate-dependent vine model.
- Better variable selection tool.
- High-dimensional value at risk forecasting.

References I

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

Kastner, G. and Frühwirth-Schnatter, S. (2014), "Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models," *Computational Statistics & Data Analysis*, 76, 408–423.

Li, F. and He, Z. (2019), "Credit risk clustering in a business group: which matters more, systematic or idiosyncratic risk?" *Cogent Economics & Finance*, 1632528.

Li, F. and Kang, Y. (2018), "Improving forecasting performance using covariate-dependent copula models," *International Journal of Forecasting*, 34, 456–476.

Li, F., Villani, M., and Kohn, R. (2011), "Modeling Conditional Densities Using Finite Smooth Mixtures," in *Mixtures: estimation and applications*, eds. Mengersen, K., Robert, C., and Titterington, M., John Wiley & Sons Inc, Chichester, pp. 123–144.

Ma, Y.-A., Chen, T., and Fox, E. (2015), "A complete recipe for stochastic gradient MCMC," in *Advances in Neural Information Processing Systems*, pp. 2917–2925.

References II

Panagiotelis, A., Czado, C., Joe, H., and Stöber, J. (2017), "Model selection for discrete regular vine copulas," *Computational Statistics & Data Analysis*, 106, 138–152.

Patton, A. (2006), "Modelling Asymmetric Exchange Rate Dependence," *International economic review*, 47, 527–556.

Scott, J. G., Berger, J. O., et al. (2010), "Bayes and empirical-Bayes multiplicity adjustment in the variable-selection problem," *The Annals of Statistics*, 38, 2587–2619.

Stöber, J. and Schepsmeier, U. (2013), "Estimating standard errors in regular vine copula models," *Computational Statistics*, 28, 2679–2707.

Welling, M. and Teh, Y. W. (2011), "Bayesian learning via stochastic gradient Langevin dynamics," in *Proceedings of the 28th international conference on machine learning (ICML-11)*, pp. 681–688.

Yau, P., Kohn, R., and Wood, S. (2003), "Bayesian variable selection and model averaging in high-dimensional multinomial nonparametric regression," *Journal of Computational and Graphical Statistics*, 12, 23–54.

Thank you!

feng.li@cufe.edu.cn

http://feng.li/