Bayesian covariate-dependent copula modeling
— with applications in stocks, text sentiments and firm credit risks

Feng Li

feng.li@cufe.edu.cn
http://feng.li/

School of Statistics and Mathematics
Central University of Finance and Economics

Outline

1. Covariate-dependent copula models
2. The Bayesian scheme
3. Improving forecasting performance with covariate-dependent tail-dependence (Li & Kang, 2016)
4. Detecting credit risk clustering (Li & He, 2017)
5. Multivariate covariate-dependence with mixed margins (Li, Panagiotelis & Kang, ongoing research)

Collaborators on the series of papers

- Yanfei Kang, Associate Professor, Beihang University, Beijing China.
- Zhuojing He, PhD student, Central University of Finance and Economics, Beijing China.
- Anastasios Panagiotelis, Associate Professor, Monash University, Australia.

* Feng Li's research is supported by National Natural Science Foundation of China.

Covariate-dependent copula models
$\looparrowright$ 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.

Covariate-dependent copula models

$\looparrowright$ The reparameterized copula model

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

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

- **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**: Gaussian and t copulas
- Marginal models we have used
  - Mixture of asymmetric student's-t distributions.
  - GARCH models
  - stochastic volatility (SV) models.
  - Poisson regression models.

The Bayesian approach

- **The log Posterior**

$$\log p(\{\beta, \mathfrak{I}\}|\boldsymbol{y}, \boldsymbol{x}) = \mathrm{c} + \sum_{j=1}^{M} \{\log p(\boldsymbol{y}_{.j}|\{\beta, \mathfrak{I}\}_j, \boldsymbol{x}_j) + \log p(\{\beta, \mathfrak{I}_j\})\}$$
$$+ \log \mathcal{L}_C(\boldsymbol{u}_{1:M}|\{\beta, \mathfrak{I}\}_C, \boldsymbol{y}, \boldsymbol{x}) + \log p_C(\{\beta, \mathfrak{I}\})$$

where

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

The Bayesian approach

- **The priors** for the copula model are easy to specify due to our reparameterization.
  - It it **not easy** to specify priors directly on $\{\beta, \mathfrak{I}\}$
  - But it is **easy** to puts prior information on the model parameters features $(\tau,\ \mu,\ \sigma^2)$ and then derive the implied prior on the intercepts and variable selection indicators.
  - When variable selection is used, we assume there are no covariates in the link functions *a priori*.
- **The posterior** inference is straightforward although the model is very complicated.

The Bayesian approach

↪ Sampling the posterior with an efficient MCMC scheme

- 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) $\{\beta_\tau, \mathcal{I}_\tau\}_C \vert \{\beta_\tau, \mathcal{I}_\tau\}_{-C}$
(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) $\{\beta_\kappa, \mathcal{I}_\kappa\}_M \vert \{\beta_\kappa, \mathcal{I}_\kappa\}_{-M}$	

Model Comparison

- 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)

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

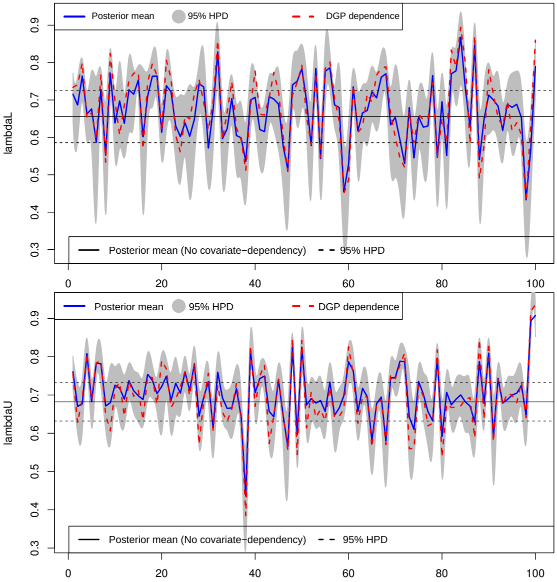
where $D_{a:b}$ is the dataset from time a to b and θ are the model parameters.

Simulation

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

Simulation



Posterior mean
95% HPD
DGP dependence
Posterior mean (No covariate−dependency)
95% HPD
lambdaL
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0
20
40
60
80
100
Posterior mean
95% HPD
DGP dependence
Posterior mean (No covariate−dependency)
95% HPD
lambdaU
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0
20
40
60
80
100

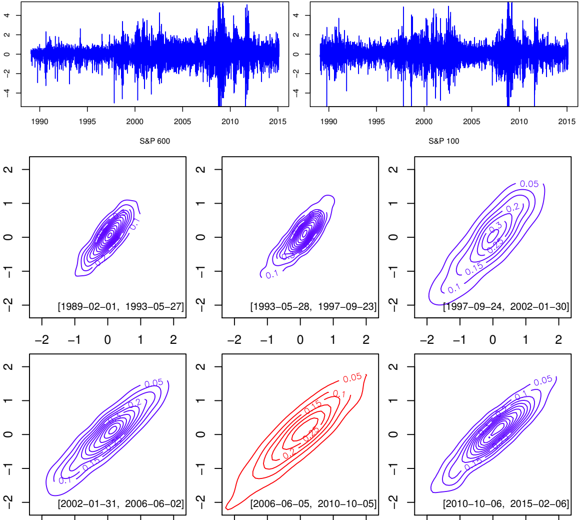
Improving forecasting performance with covariate-dependent tail-dependence (Li & Kang, 2016)

↪ Log predictive density score comparison

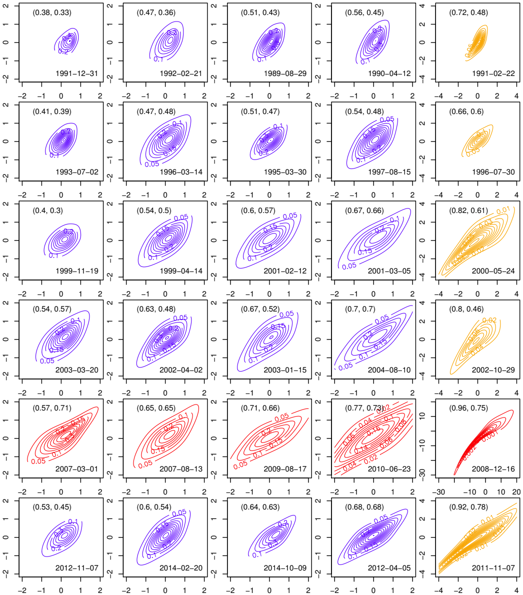
Margins	LPS decomposition	Reparameterized Copulas			
		Joe-Clayton	Clayton	Gumbel	t-Copula
		(*Joint modeling approach*)			
SPLIT-t	M_1	−1743.12	−1741.04	−1754.36	−1741.47
	M_2	−1435.98	−1468.25	−1485.68	−1430.07
	$C(CD.)$	837.50	690.22	797.78	792.14
	Global	−2344.12	−2523.75	−2448.14	−2380.12
SPLIT-t	M_1	−1747.99	−1747.15	−1754.61	−1782.37
	M_2	−1434.22	−1449.95	−1446.84	−1658.09
	$C(Const.)$	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
	$C(CD.)$	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
	$C(CD.)$	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
	$C(CD.)$	964.37	698.30	1012.10	1053.19
	Global	−3013.90	−3300.46	−2964.68	−2934.40
		(*Bivariate volatility models*)			
Bivariate DCC-GARCH		−2730.78			
Bivariate SV		−2999.63			

Improving forecasting performance with covariate-dependent tail-dependence (Li & Kang, 2016)

↪ The S&P 100 and S&P 600 and their empirical copulas



Contour plots of the posterior densities



Detecting credit risk clustering with distance-to-default index (Li & He, 2017)

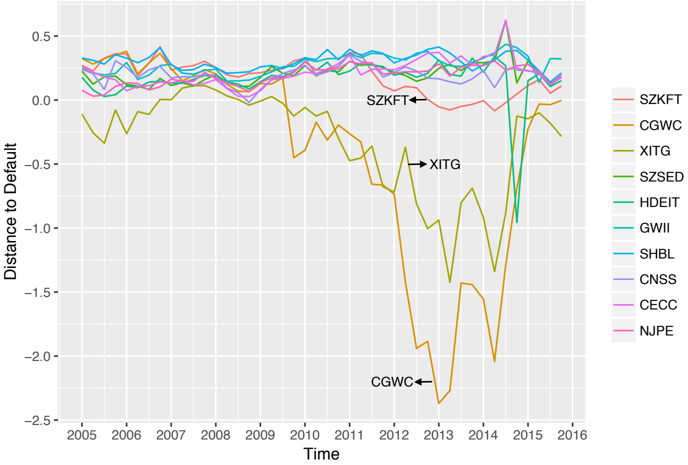


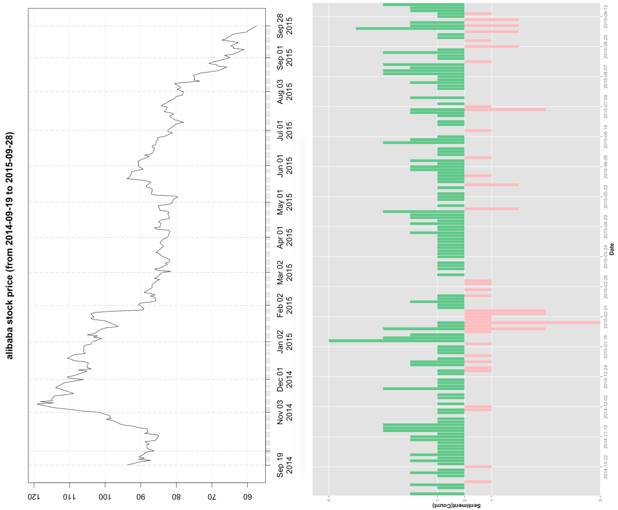
Figure: Distance-to-default (DTD) for 10 firms. In risk management, the probability of default is high if the value of DTD is small.

Covariates effects on the tail-dependency

	No covariates	Macroeconomic covariates	Specific covariates	Macroeconomic and specific covariates
Constant	−4.931	2.014	−2.174	10.515
	(1.000)	(1.000)	(1.000)	(1.000)
CPI		−0.431		−71.814
		(0.593)		(0.507)
M2 growth rate		−0.122		2.169
		(0.586)		(0.636)
Short-term interest rate		−0.012		11.998
		(0.988)		(0.254)
RMB/USD spot rate		−0.526		−0.650
		(0.605)		(0.309)
CGWC's solvency capacity			−0.017	4.498
			(0.866)	(0.590)
CGWC's developing capacity			0.012	−1.680
			(0.637)	(0.624)
CGWC's profitability			0.004	−12.948
			(0.751)	(0.597)
CGWC's operating capacity			−0.039	4.819
			(0.716)	(0.615)
XITG's solvency capacity			0.089	58.419
			(0.813)	(0.537)
XITG's developing capacity			0.030	104.257
			(0.652)	(0.578)
XITG's profitability			−0.409	294.978
			(0.531)	(0.389)
XITG's operating capacity			−0.057	−0.305
			(0.857)	(0.463)
LPS	−308.732	−200.606	−106.542	−52.831

Multivariate covariate-dependence with mixed margins (Li, Panagiotelis & Kang, ongoing research)

↬ Modeling stock returns and text sentiments



Covariates in texts data

财新网 Caixin.com

阿里巴巴 搜索 帮助？

您搜索 **阿里巴巴** 获得大约 **647** 条查询结果，以下是第 **1 – 20** 条。 （搜索用时 0.18117423 秒）

- **所有结果**
- 财新网
- 杂志
- 图片
- 博客
- 视听
- 数字说
- 专题

- 时间不限
- 一天内
- 一周内
- 一月内
- **一年内**
- 自定日期范围

- **全文**
- 正文
- 作者
- 标题

联姻东风日产不到两年 恒大三次提出回购要求（要闻）[2015-11-27]

2月1日至2016年1月31日。 2014年6月，恒大俱乐部增资扩股引入**阿里巴巴**。**阿里巴巴**（中国）网络技术有限公司注资12亿，持有恒大足球俱乐部40%股权。广州恒大队更名为广州恒大淘宝队。 2014年

阿里影业积攒超级IP 年产电影三部左右（TMT）[2015-11-26]

的整条电影产业链。此前阿里影业还收购了粤科软件，**阿里巴巴**还对优酷土豆发起私有化邀约，也有意与阿里影业产生协同，如何贯通业务是阿里影业首先要解决问题。 C2B影视娱乐内容投融资平台娱乐宝就是一个众筹平台

互联网跨界融合中的"共生金融"新模式（王永利）[2015-11-26]

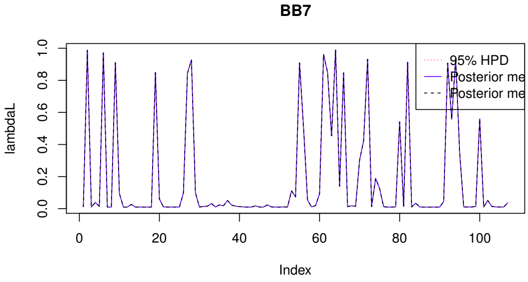
，不仅新兴的互联网企业（如**阿里巴巴**、腾讯、百度等）依托其核心竞争力不断拓展经营范围，形成跨业经营的共生经济模式，而且一些商业物流企业、房地产企业等也运用互联网技术进行改造，加快共生经济模式的发展（如京

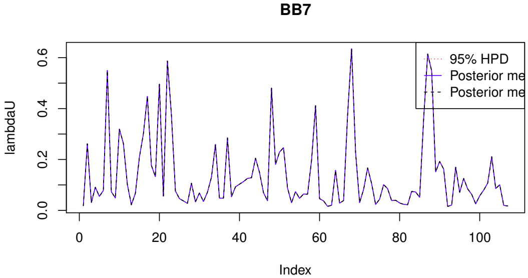
中国优步瞄准企业服务和二三线市场（TMT）[2015-11-26]

。"2015年加满油、蓄势待发，2016年才是真正全速前进的一年。"柳甄说。就在发布会前一天，她刚与**阿里巴巴**旗下钉钉签署合作协议，钉钉的员工将通过优步实现上下班通勤等，并正式宣布针对企业客户的U4B（即

Date	P	N	U	上涨	下跌	打击	合作	增加	影响	违法
2014-10-15	2	0	1	0	0	1	3	0	1	0
2014-11-19	3	0	1	0	0	1	3	0	1	0
2015-01-28	3	3	3	0	0	2	5	1	2	4
2015-01-29	1	1	2	1	2	1	1	1	2	1
2015-01-30	1	7	2	0	1	3	1	1	5	5
2015-07-08	2	3	2	0	2	1	2	2	2	0

The dependence between positive/negative sentiments and stocks





Working in progress

- Modeling multivariate covariate-dependent structure via the vine copula.
- Looking into more efficient inference techniques, VB?
- Including probabilistic topic model to model the texts margins.

Thank you!

feng.li@cufe.edu.cn

http://feng.li/