

doi:10.11835/j.issn.1008-5831.2017.02.004

欢迎按以下格式引用:韩超,严太华.基于高维动态C-Vine Copula模型的中国股票市场行业组合风险研究[J].重庆大学学报(社会科学版),2017(2):40-50.

Citation Format: HAN Chao, YAN Taihua. A study on industry portfolio risk in China stock market based on high dimensional dynamic C-Vine Copula model [J]. Journal of Chongqing University(Social Science Edition), 2017(2):50-50.

# 中国股票市场行业组合风险研究

## ——基于高维动态C-Vine Copula模型

韩超,严太华

(重庆大学经济与工商管理学院,重庆 400044)

**摘要:**股市是经济的晴雨表,股市中不同行业风险的组合计量对于金融市场和实体经济投资意义重大。文章采用高维动态C-Vine Copula前沿技术计量多维行业组合风险,并且与静态C-Vine Copula作比较。结论显示:高维动态C-Vine Copula计量的VaR每次都能通过UC检验和稳定性测试,而静态C-Vine Copula方法每次都不能通过回溯检验,表明高维动态C-Vine Copula优于静态C-Vine Copula,可以作为股市行业风险组合计量的一种新方法。

**关键词:**高维动态C-Vine Copula;GPD;组合风险计量;VaR**中图分类号:**F830.91**文献标志码:**A**文章编号:**1008-5831(2017)02-0040-11

单一金融时间序列往往具有偏斜的尖峰厚尾特征;多维金融时间序列的联合分布不一定是椭圆分布。对于多维金融时间序列来说,其内部相关很可能不是线性的,传统的线性相关计量可能会导致错误结论。Sklar<sup>[1]</sup>引入的Copula函数能够较好地解决非线性、非对称相依计量问题,契合了多维金融时间序列的统计特征。自Embrechts<sup>[2]</sup>以来,Copula函数逐渐应用于金融研究领域,如Bouyé等<sup>[3]</sup>、Rockinger等<sup>[4]</sup>、Romano<sup>[5]</sup>、Embrechts等<sup>[6]</sup>、Rodriguez<sup>[7]</sup>、Ghorbel和Trabelsi<sup>[8]</sup>等。近期的文献表明:高维化和动态化是Copula理论研究的前沿动向。在Copula函数行之有效的多维化拓展中,Vine-Copula代表了典型研究方向,Vine-Copula借助于Pair-Copula-Construction(PCC)构造可以实现降维目的,使“维度诅咒”问题得以克服,而PCC构造使得降维后多元变量之间两两相依的Copula种类选择和参数拟合更为灵活。进一步赋予Vine-Copula参数以动态化,则产生高维动态Vine-Copula这一最新的研究方法,相关文献主要有So和Yeung<sup>[9]</sup>、Reboredo和Ugolini<sup>[10]</sup>、杜子平等<sup>[11]</sup>、曹洁和程希骏<sup>[12]</sup>等。在Vine Copula的研究文献中以C-Vine和D-Vine两种特殊的Regular Vine(R-Vine)类型最为常见,本文采用C-Vine Copula结构进行多元相依建模。

股市数据作为金融时间序列的重要组成部分,采用Copula函数对其相依性计量的文献也很多,如曹洁和程希骏<sup>[12]</sup>、Berger<sup>[13]</sup>、李强等<sup>[14]</sup>、Domino和Błachowicz<sup>[15]</sup>等。股市由各行业组成,各行业之间存在着复杂

修回日期:2016-11-26

基金项目:国家自然科学基金项目“交叉上市、投资者情绪与资产定价”(71373296)

作者简介:韩超,男,重庆大学经济与工商管理学院博士研究生,主要从事金融工程研究,E-mail:charlie.han@163.com;严太华,男,重庆大学经济与工商管理学院教授,博士研究生导师,主要从事数量经济学研究。

的作用机制,行业风险的组合计量方法可以扩展于测定股市整体风险,股市行业风险的研究为众多学者所关注,如刘琼芳和张宗益<sup>[15]</sup>、赵宁等<sup>[17]</sup>、张帮正和魏宇<sup>[18]</sup>、Righi 和 Ceretta<sup>[19]</sup>、Bartram 和 Wang<sup>[20]</sup> 等采用 Copula 函数对股市行业风险进行了相依性研究。以往的股市行业风险研究方法,多数止步于传统多元 Copula,或者仅用 Vine Copula 对数据进行拟合,运用仿真尤其是动态仿真进而进行风险组合计量的研究较少。针对这一研究现状,本文采用高维动态 C-Vine Copula 进行研究,边缘分布通过极值理论的 GPD 模型拟合,有效刻画金融时间序列的极值特征,C-Vine 结构能够突出高维变量中主导变量的作用,以 PCC 模块突破“维度诅咒”,实现灵活的高维动态建模目的。研究过程中采用二阶段建模方法:第一阶段利用 GJR-GARCH 模型过滤原始序列,以 GPD 模型对过滤所得新息序列进行概率积分变换(PIT);第二阶段对 PIT 序列分别进行动态和静态的 C-Vine Copula 建模比较。之后,通过(动态)Vine 结构的仿真技术计量四维行业组合风险,对于投资者组合风险控制与监管层多维风险监管具有一定的启示意义。举一反三,希望此方法还可以推广到更高维的甚至其他领域的研究之中。

## 一、理论基础

### (一) Copula、Vine Copula 与风险计量

相较于传统的建立在线性相关和正态分布基础上的研究方法而言,Copula 函数能够把具有非正态分布特征和非线性相依关系的多维变量“连接”起来,Copula 函数的这一优越特性满足了相互联系的多维世界的大数据数理分析要求,使其获得大范围的运用,尤其在多维风险组合计量方面颇有价值。Copula 函数的多维化具有广阔的现实意义,传统的多元 Copula 却极不灵活,Vine Copula 则可以实现降维目的,借助 Pair Copula 形式进行成对建模,赋予建模以模型种类和模型参数的灵活性,使多元 Copula 的计算大为简化、更为可行,具有多重优点。基于此,Vine Copula 用作多维风险建模就有了现实意义,其利用 Copula 的连接属性把若干风险变量以特殊的 Vine 结构粘连起来,能够用来进行高维风险组合计量,这一用途逐步获得认可,逐渐应用到金融风险研究领域。

### (二) Vine 的类型与排序

在 Vine Copula 的研究文献中以 C-Vine 和 D-Vine 两种特殊的 Regular Vine(R-Vine)类型最为常见,其中 C-Vine 可以用来组织存在关键变量风险因子之间的多维 Copula 函数关系,而 D-Vine 则主要用来对不存在关键变量的多元 Copula 函数进行描述,而介于 C-Vine 和 D-Vine 之间也有很多其他的 Vine 类型。在 Vine 的排列中,第一棵树的排列尤为重要,它决定了之后树的结点和边的构造。对于第一棵树,C-Vine 需要将关键变量置于根部,其余变量按照与根部的紧密关系依次排列;D-Vine 也面临着变量之间相关强度大小的问题,依据相关度大小排列成对变量,据以构成树形图,需要注意的是 D-Vine 中当一个变量被选择两次的时候,它不能在第一棵树上形成新的 Pair Copula 对。本文的着重点在 C-Vine 类型的排序上。

## 二、高维动态 Vine Copula 的构建

赋予 Vine Copula 的参数动态化,即可得到动态 Vine Copula。动态 Vine Copula 在赋予 Copula 函数特殊结构的同时,也赋予了 Vine 的结点、边与 Copula 参数以动态化特征,是目前该学术领域的前沿研究方法。高维动态 Vine Copula 的构造根据 Vine 的类型不同而变化,高维动态 Vine Copula 的 h 函数的求解与数据仿真是一个难点,决定了该方法能否得以顺利应用。h 函数提法引自 Aas 等<sup>[21]</sup>。本文采用 t-Copula 函数形式开展研究,其动态 h 函数如式(4)。

本文采用两阶段建模方法拟合 Vine 结构,第一阶段根据边际分布模型参数进行概率积分变换,得到 PIT 序列,第二阶段进行 Vine 结构的拟合。模型构建步骤如下。

(1)用 GJR-GARCH-SKewT 模型对多元收益率序列进行过滤,模型如式(1)所示:

$$r_{it} = \mu_i + z_{it}, z_{it} = \sigma_{it} \cdot \varepsilon_{it}$$

$$\sigma_{it}^2 = \omega_i + \alpha_i z_{i,t-1}^2 + \gamma_i I_{i,t-1} z_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2 \quad \varepsilon_{it} \sim i. i. dST(\lambda_i, \eta_i) \quad (1)$$

其中  $\lambda_i, \eta_i$  分别表示 SkewT 分布的偏斜度和形状参数。

(2)对上一步中的标准残差序列进行 GPD 极值建模,模型如式(2)所示:

$$F_{u_i}^{[t]}(y_{it}) \approx G_{\xi_i, \beta_i}^{[t]}(y_{it}) = \begin{cases} 1 - (1 + \xi_i y_{it} / \widehat{\beta}_i)^{-1/\xi_i}, \xi_i \neq 0 \\ 1 - \exp(-y_{it} / \widehat{\beta}_i), \xi_i = 0 \end{cases} \quad i = 1, 2, 3, 4 \quad (2)$$

其中,  $y_{it} = \hat{x}_{it} - \hat{u}_i$  表示超阈值,  $\hat{x}_{it}$  是独立同分布的随机变量序列,  $\hat{u}_i$  是  $\hat{x}_{it}$  尾部极值的阈值,  $\widehat{\beta}_i > 0$ 。继而以相应阈值和参数的 GPD 模型去刻画上下尾分布形态, 以 Epanechnikov 核函数对数据中间区域进行平滑处理, 得到较为满意的 PIT(概率积分变换)序列。

(3) 对 PIT 序列进行动态 Pair Copula 建模, 求出相应的时变相关系数序列, 对其求均值。这里采用的是 Pair-t-Copula 函数, 模型表示为式(3), 其中,  $u_{it}, u_{jt}$  均表示 PIT 序列元素, 令  $x_t = (x_{it}, x_{jt})'$ , 则:

$$G_{ij}^{[1]}(u_{it}, u_{jt}; \rho_{ij,t}, v_{ij}) = T_{v_{ij}}(T_{v_{ij}}^{-1}(u_{it}), T_{v_{ij}}^{-1}(u_{jt})) = \int_{-\infty}^{T_{v_{ij}}^{-1}(u_{it})} \int_{-\infty}^{T_{v_{ij}}^{-1}(u_{jt})} \frac{\Gamma\left(\frac{v_{ij}+2}{2}\right) |\rho_{ij,t}|^{-\frac{1}{2}}}{\Gamma\left(\frac{v_{ij}}{2}\right) (v_{ij}\pi)} \left(1 + \frac{1}{v_{ij}} x_t' \rho_{ij,t}^{-1} x_t\right)^{-\frac{v_{ij}+2}{2}} dx_{it} dx_{jt} \quad (3)$$

(4) 比较高维 PIT 序列的两两相关系数均值, 分析是否存在主导变量, 以确定 Vine 的类型和关键变量, 如果存在主导变量, 则采用 C-Vine。

(5) 本文选择的是 C-Vine 结构, 进而拟合边和节点的参数。其中, 一棵树的边是下一棵树的节点, 下一棵树的节点由上一棵树连接对应边的两节点拟合的(条件)Copula 函数所对应的 h 函数求出。由此, 不断迭代, 直到求出所有树上的边和节点的对应参数。当然这里的 Vine 结构参数和 h 函数值都是动态数组, 其中动态化的 h 函数表示为式(4)所示:

$$h^{[1]}(u_{it} | u_{jt}) = t_{v_{ij}+1} \left\{ \frac{t_{v_{ij}}^{-1}(u_{it}) - \rho_{ij,t} t_{v_{ij}}^{-1}(u_{jt})}{\sqrt{\frac{(v_{ij} + (t_{v_{ij}}^{-1}(u_{jt}))^2)(1 - \rho_{ij,t}^2)}{v_{ij} + 1}}} \right\} \quad (4)$$

这里带有下标的  $\rho_{ij,t}$  是时变的, 其动态化演进服从式(5)所示的 AR(1)过程:

$$\rho_{ij,t} = (1 - \alpha_{ij} - \beta_{ij}) \cdot R_{ij} + \alpha_{ij} \cdot r_{ij,t-1} + \beta_{ij} \cdot \rho_{ij,t-1}, 0 < \alpha_{ij} + \beta_{ij} < 1, \alpha_{ij} > 0, \beta_{ij} \geq 0 \quad (5)$$

其中,  $\rho_{ij,t}$ 、 $R_{ij}$ 、 $r_{ij,t-1}$  分别表示时变相关参数、样本相关参数、残差相关系数,  $\alpha_{ij}$  和  $\beta_{ij}$  表示其后变量对应的参数系数。

### 三、高维动态 Vine Copula 的仿真

理论上, C-Vine Copula 和 D-Vine Copula 的仿真算法形式是一样的, 如式(6)所示:

$$\begin{aligned} x_{1t} &= w_{1t} \\ x_{2t} &= h_{21}^{[1]-1}(w_{2t} | x_{1t}) \\ x_{3t} &= h_{31,2}^{[1]-1}(w_{3t} | x_{1t}, x_{2t}) \\ x_{4t} &= h_{41,2,3}^{[1]-1}(w_{4t} | x_{1t}, x_{2t}, x_{3t}) \end{aligned} \quad (6)$$

其中,  $w_{it}, x_{it}, i = 1, \dots, 4$  分别表示独立的均匀分布元素和待仿真变量。由式(6), h 逆函数对于 Vine 结构的倒推和数据仿真至关重要。对于动态 Vine Copula, h 逆函数是动态的。动态 h 逆函数的条件信息集存在多个变量, 求解较为复杂, 我们可以通过层层推导把条件变量一一剔除, 求出最终解。本文采用 C-Vine Copula 结构开展研究, 着重介绍 C-Vine Copula 的数据仿真方法。对于 C-Vine 来说, 求解  $x_{it}$  的推导过程如式(7), 其中,  $x_{it}$  表示要仿真的第  $i$  维变量, 令  $w_{i1, \dots, i-1, t} = h^{[1]}(x_{it} | x_{1t}, \dots, x_{i-1, t}) \quad i = 2, 3, 4$ 。

$$\begin{aligned} x_{1t} &= w_{1t} \\ x_{2t} &= h_{21}^{[1]-1}(w_{21,t}, w_{1t}) \\ x_{3t} &= h_{31}^{[1]-1}(h_{312}^{[1]-1}(w_{312,t}, w_{21,t}), w_{1t}) \\ x_{4t} &= h_{41}^{[1]-1}(h_{412}^{[1]-1}(h_{4123}^{[1]-1}(w_{4123,t}, w_{312,t}), w_{21,t}), w_{1t}) \end{aligned} \quad (7)$$

但是,式(7)的嵌套表述是隐晦的、不直白的。接下来,为了表述方便,以静态的第四维  $x_4$  求解为例,展示这一求解过程,如式(8),动态的思路是一致的,只不过参数是动态的。

$$x_4 = h_{41}^{-1}{}_{123}(w_4 | x_1, x_2, x_3) \Rightarrow w_4 = h_{41}{}_{123}(x_4 | x_1, x_2, x_3) = F(x_4 | x_1, x_2, x_3) =$$

$$\frac{\partial C_{43112} \{F(x_4 | x_1, x_2), F(x_3 | x_1, x_2)\}}{\partial F(x_3 | x_1, x_2)} = \frac{\partial C_{43112} \{h_{41}{}_{12}(x_4 | x_1, x_2), h_{31}{}_{12}(x_3 | x_1, x_2)\}}{\partial h_{31}{}_{12}(x_3 | x_1, x_2)} =$$

$$h_{431}{}_{12} \{h_{41}{}_{12}(x_4 | x_1, x_2), h_{31}{}_{12}(x_3 | x_1, x_2)\} \Rightarrow h_{41}{}_{12}(x_4 | x_1, x_2) = h_{431}{}_{12}^{-1}(w_4, h_{31}{}_{12}(x_3 | x_1, x_2)),$$

$$\bar{w} h_{31}{}_{12}(x_3 | x_1, x_2) = w_3, \therefore h_{41}{}_{12}(x_4 | x_1, x_2) = h_{431}{}_{12}^{-1}(w_4, w_3), \text{令 } z = h_{431}{}_{12}^{-1}(w_4, w_3) = h_{41}{}_{12}(x_4 | x_1, x_2) =$$

$$h_{421}{}_{1}(h_{41}{}_{1}(x_4, x_1), h_{21}{}_{1}(x_2, x_1)) \Rightarrow h_{41}{}_{1}(x_4, x_1) = h_{421}{}_{1}^{-1}(z, h_{21}{}_{1}(x_2, x_1)), z' = h_{41}{}_{1}(x_4, x_1) \Rightarrow x_4 = h_{41}{}_{1}^{-1}(z', x_1) \quad (8)$$

对于高维动态 C-Vine 的数据仿真来说,其计算过程是动态化的,动态参数由相应的动态型 h 函数和时变 Copula 函数求得,动态参数服从式(5)中的动态演进式。为了更为直观地表达仿真计算过程,以  $x_{4t}$  为例,画出 C-Vine 的动态仿真路径图 1,其中,  $W_4$  为  $w_4$  的集合向量,  $X_4$  为  $x_{4t}$  的集合向量。

当通过式(7)、式(8)和图 1 仿真出四维  $x_{it}$  变量以后,对其按照两阶段建模法第一阶段的模型和参数进行返回倒推,即可仿真出高维动态相关收益率序列,本文的仿真次数以 10 000 次为 1 组。

#### 四、VaR 与回溯检验

VaR 表示资产收益率序列在一定显著性水平下的分位数。求解 VaR 有历史模拟法、方差-协方差方法和蒙特卡罗模拟法等。本文基于对高维数据动态 C-Vine-Copula 结构的 Pair Copula 分解,如果通过复杂的多重积分表达式采用方差-协方差方法,理论上是可以求出显著性水平  $\alpha$  下  $t$  时序的风险价值  $\text{VaR}_t(\alpha)$  的,然而这个方法在实际计算上不太可行。因此,本文借助 Matlab 编程实现蒙特卡罗模拟,按照拟合的高维动态 C-Vine 结构仿真多维数据,倒推出模拟收益率序列,每 10 000 个数据模拟实现之后,通过计算机按照相应的显著性水平自动选择位于数据排序左端的对应数据作为此 10 000 个数据组的 1 个 VaR 值,如此得出分别包含 318 个元素的 VaR 随机序列组,将其与样本外 318 个数据分别进行返回检验,观察比较高维动态与高维静态 C-VineCopula 的风险效果。VaR 的回溯检验通过 R 语言编程实现,理论基础是 Kupiec<sup>[22]</sup> 的 UC 检验理论,进而验证模型的有效性。

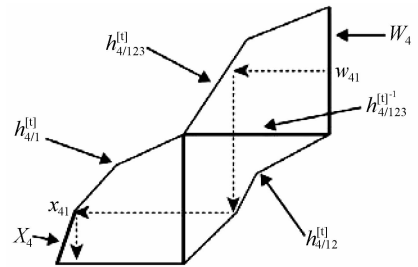


图 1 动态 C-Vine-Copula 的仿真路径图

#### 五、数理分析

##### (一) 数据预处理

本文对沪深 300 指数中的基建、银行、运输和地产四大行业指数从 2006 年 1 月 4 日到 2014 年 4 月 30 日的开展研究,研究对象为日间对数收益率  $r_t = \ln(P_t/P_{t-1}) * 100\%$ ,共 2 018 组数据。其中前 1 700 组四维数据用于模型拟合,后 318 组数据用来进行 VaR 的样本外返回检验。研究过程中借助 R 和 Matlab 语言编程实现。本文选择的四大行业具有典型特征:基建和地产是典型的资本密集型行业,是 GDP 投资乘数效应的主要因素,包含了政府直接投入和撬动的民间资本,运输是实体经济繁荣程度的典型代表,银行是实现货币乘数效应的典型中介,其造血和输血功能对经济繁荣起到催化作用。这四大行业之间不是孤立的,基建的风险和地产风险本身具有交叉重叠部分,基建的增长显然会为地产打造增值基础,地产的繁荣会创造税收,为基建形成资金来源,地产与基建的共同发展会创造就业、提高收入,借助于 GDP 的加速度效应为实体经济产生投资动能;同时地产与基建的投资增长也会对实体经济产生挤占效应,遏制实体经济投资增长;实体经济的兴衰会体现到运输行业上来,运输行业又会对基础设施建设产生需求变化,而银行以信贷收缩与扩张的形式为三大行业的兴衰产生顺周期的或者熨平周期的货币行为,在此过程中实现息差利润。如此,四大行业之间一定存在着某种复杂的关系,这四大典型行业的风险研究对于整个国家的经济风险具有指示意义。本文的研究数据取自国泰安数据库。表 1 为四维序列的统计量表,显见四维数据的尖峰厚尾性, Jarque-Bera 统计量强烈拒绝正态性假定,数据分布具有左偏特性。

表1 四维数据统计量表

指标	地产	基建	银行	运输
Mean	0.000 9	0.000 2	0.000 6	0.000 0
Skewness	-0.279 8	-0.587 5	-0.034 9	-0.508 8
Kurtosis	4.429 9	6.193 7	5.487 9	5.639 1
Jarque - Bera	167.013 8	820.245 7	438.782 6	566.681 8
Probability	0	0	0	0

## (二) 平稳性检验

本文运用模型对数据进行过滤之前必须对数据进行平稳性检验,平稳性是金融时间序列建模和预测的前提,相应的单位根 ADF 检验结果如表 2 所示,可见四种金融时间序列均强烈拒绝了单位根存在的原假设,没有理由认为数据不是平稳时间序列。

## (三) 自相关检验与波动性聚类判别

为了更好地进行数据建模,进行数据特征分析很有必要。以基建和地产为例,如图 1 和图 2 所示,序列显示出轻微的自相关性和波动性聚类持久记忆特征。本文采用 GJR - GARCH 模型对数据进行处理,可以看到模型实现了预定效果。

## (四) 单个时间序列的模型提炼与数据过滤

由 Glostern 等<sup>[23]</sup>、Peiró<sup>[24]</sup>等文献可以知道,GARCH 过滤金融时间序列后的标准化残差分布显著偏斜,再加上股指收益率序列的波动性聚类和对正负新息冲击的非对称性特征,因此,本文采用 GJR - SkewT 模型对序列进行拟合。结合本文研究数据,四种对数收益率序列均为尖峰厚尾的偏态分布,具有平稳性,有轻微的序列相关性和偏斜特征,存在 GARCH 效应,同时对数据进行 ARMA 识别,发现均为 0 阶。

故而,本文采用 ARMA(0,0) - GJR - SkewT 模型对序列进行拟合,相应的数学表达式见式(1),其中  $\lambda_i, \eta_i$  分别表示 SkewT 分布的偏斜度和形状参数,拟合参数见表 3。

表2 日对数收益率的平稳性检验表

指标 1	指标 2	值
ADF 统计量	基建	-41.243 8
	地产	-40.172 0
	银行	-41.222 2
	运输	-39.961 2
$t$ -统计量	$\alpha = 1\%$	-3.434
显著性水平	$\alpha = 5\%$	-2.863
$\alpha$	$\alpha = 10\%$	-2.567 6
	USD	0
	EUR	0
	JPY	0
	HKD	0
$p$ 值	EUR	0
	JPY	0
	HKD	0
	USD	0

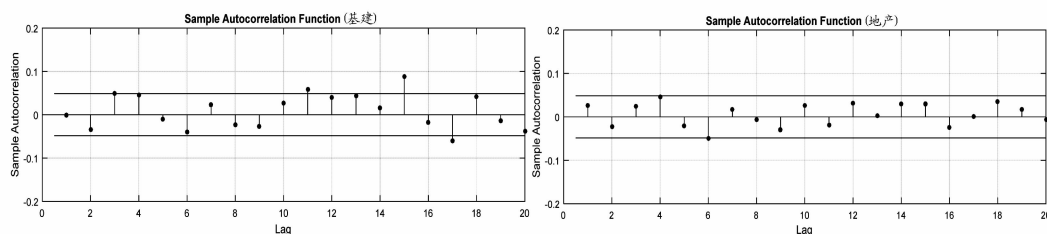


图2 收益率的自相关图

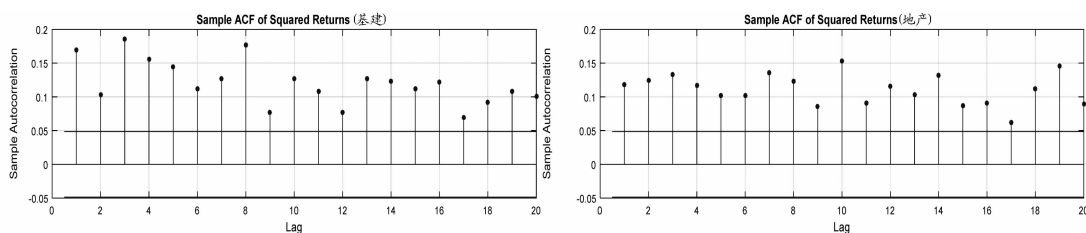


图3 收益率平方的自相关图

表3 ARMA(0,0) - GJR - SkewT 拟合参数表

参数	地产	基建	银行	运输
$\mu_i$	0.000 7	-0.000 2	0.000 3	-0.000 3
$\omega_i$	0.000 0	0.000 0	0.000 0	0.000 0
$\alpha_i$	0.031 7	0.068 2	0.038 1	0.064 9
$\beta_i$	0.967 4	0.938 0	0.962 1	0.938 6
$\gamma_i$	-0.001 5	-0.016 2	-0.002 3	-0.009 3
$\lambda_i$	0.953 1	0.877 1	1.040 4	0.896 3
$\eta_i$	7.099 3	5.637 1	6.951 0	5.742 4
LogLike	3 899.021	4 648.031	4 305.757	4 416.353

为了验证 ARMA(0,0) - GJR - SkewT 模型的实际效果,本文对过滤所得的标准残差序列分别绘制自相关图 4 和波动性聚类图 5,仍以基建和地产为例,从图形中我们可以看到自相关性有进一步的削弱,而波动性聚类则基本消除,可以认为模型设定初步达到了预定效果。

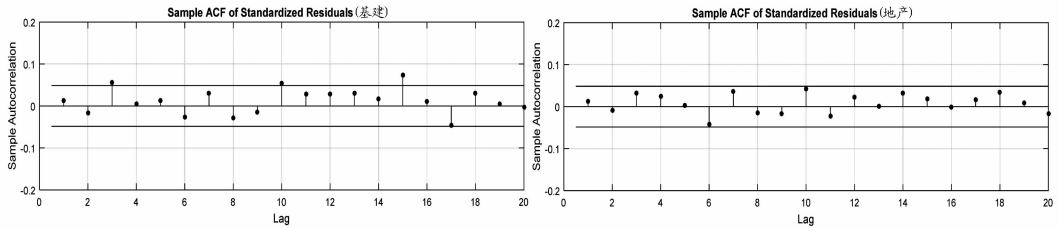


图4 标准残差的自相关图

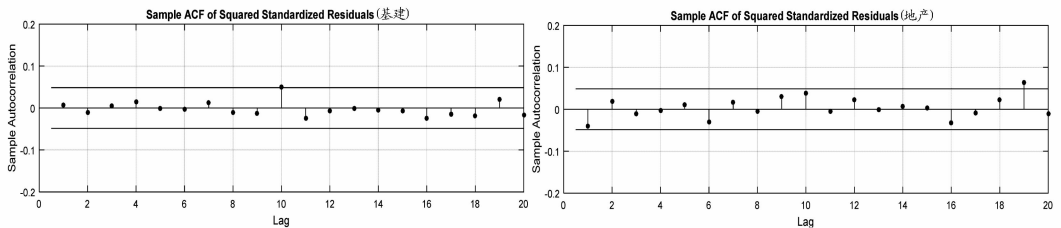


图5 标准残差平方的自相关图

(五) 广义帕累托分布(GPD)参数估计与拟合检验

金融序列尾分布的刻画对于 Vine Copula 求 VaR 的准确性极其重要,合适的尾分布拟合对于研究结论很关键。为了进一步准确刻画数据特征,本文对过滤的标准残差再进行 GPD 尾部拟合,相应的 GPD 拟合参数如表 4 所示,其中括号内的阈值为上尾阈值,其余 GPD 参数均为下尾参数。

表4 GPD 拟合参数表及 PIT 检验表

行业	GPD 拟合参数				PIT 检验	
	阈值	位置	$\xi_i$	$\hat{\beta}_i$	K-S 的 Z 值	双尾渐近显著性
基建	-1.08(1.11)	11.41%	0.021 4	0.670 2	0.871	0.434
地产	-1.27(1.31)	8.94%	-0.026 3	0.620 6	0.776	0.584
银行	-1.26(1.20)	9.47%	-0.091 4	0.627 7	0.831	0.494
运输	-1.09(1.10)	11.71%	0.086 5	0.599 9	0.792	0.558

然后,本文对过滤得到的标准化残差序列进行概率积分变换(PIT)。在 PIT 过程中,以相应阈值和参数的 GPD 模型去刻画上下尾分布状态,以 Epanechnikov 核函数对数据中间区域进行平滑处理,得到一个整体的序列特征分布组合(分布组合的拟合效果如图 6 所示,仍以基建和地产为例),进而以此对残差序列进行 PIT,并进行 K-S 检验,结果显示四组 PIT 序列均不能拒绝服从 U(0,1)分布的原假设,如表 4 右半部分所

示。这也就为高维动态 Vine - Copula 建模做好了准备工作。

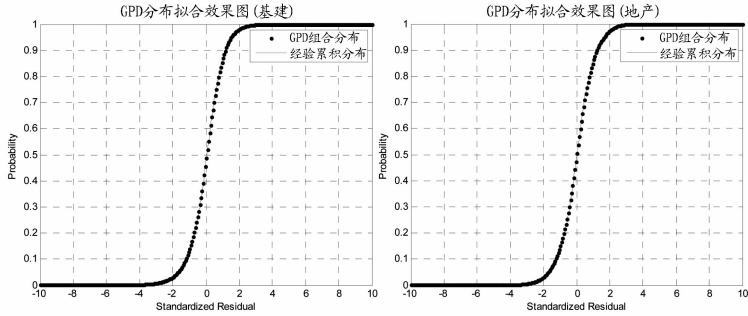


图 6 GPD 分布组合的拟合效果诊断图

(六)高维动态 C - Vine Copula 结构的 Pair - Copula 分解

在接下来的研究中,为简化起见,均选用 t Copula 函数进行研究分析。t Copula 属于椭圆类 Copula 的一种,能够捕捉序列间的对称性和厚尾性特征。本文选用 t Copula 只是一个次优化选择,感兴趣的研究者完全可以把选择的视野放宽,追求更优拟合的 Copula 类型。尽管如此,我们仍然画出四维行业变量两两 t - copula 的概率图(7),显见两两关系的对称的厚尾形态,说明选择 t Copula 开展研究是较为合理的。

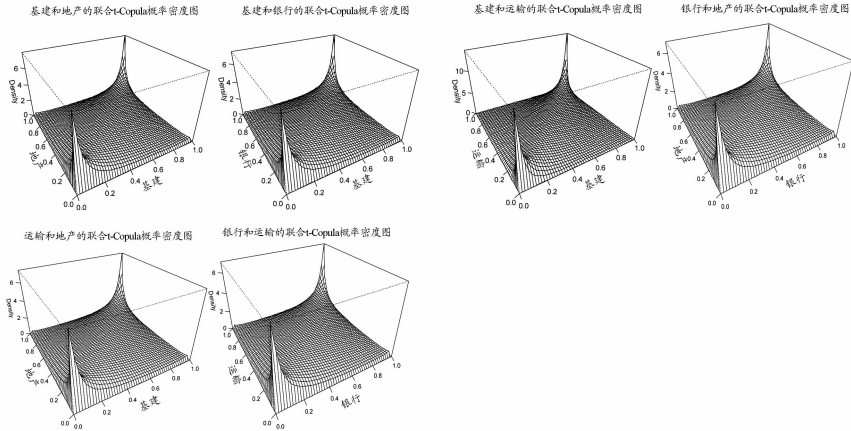


图 7 四维 Pair - Copula 概率密度图

在进行动态化建模之前需要确定 Vine 的类型,将 t Copula 参数的动态过程设为式(5),定义  $\bar{\rho} = \frac{1}{n} \cdot \sum_{i=1}^n |R_{ij,i}|$  作为归一化相关关系的测度指标,表 5 为四维变量动态归一化相关表。由表 5,可以比较出基建在与其他变量的相关关系中处于强势地位,进而根据两两动态相关强度的大小可以依次排序为基建、运输、地产和银行。这一排序与我们的先验知识相符,基建的投资规模最为宏大,运输处于行业发展的纽带与指示地位,超过了地产排在第二位,而地产和银行则紧紧绑在了一起,说明地产的资本运作属性,银行为地产的信贷扩张所牵制。因此可以把基建确定为关键变量,这也就说明了 Vine Copula 的类型是 C - Vine,根据强弱关系可以绘制出图 8 的 C - Vine 结构图,其中,数字 1, 2, 3, 4 分别表示基建、地产、银行和运输四种行业指数收益率。

表 5 两两动态 t copula 归一化相关表

$\bar{\rho}$	基建	地产	银行	运输
基建	1	0.712 4	0.682 5	0.928 8
地产	0.712 4	1	0.667 3	0.686 5
银行	0.682 5	0.667 3	1	0.672 7
运输	0.928 8	0.686 5	0.672 7	1

根据 C - Vine 的结构图 8 来估计相应的树、边和节点的参数,相应的动态 C - Vine 结构参数见表 6,其中,  $\alpha_{ij}$  和  $\beta_{ij}$  分别对应动态方程式(5)中的参数系数  $\alpha_{ij}$  和  $\beta_{ij}$ 。

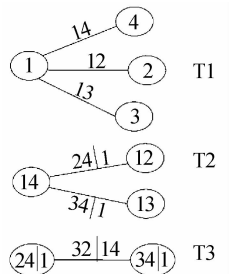


图 8 四维 C - Vine 结构图

表 6 动态 C-Vine 结构参数表

参数	(基建, 运输)	(基建, 地产)	(基建, 银行)	(地产, 运输   基建)	(银行, 运输   基建)	(地产, 银行   基建, 运输)
d. f.	5.679 9	9.417 5	7.178 8	34.597 7	48.182	16.484 2
St. E	(0.850 7)	(2.267 6)	(1.105 1)	(19.932 6)	(20.236 6)	(6.890 7)
$\alpha_{ij}$	0.036 7	0.034 9	0.034 1	0.032 6	0.026 9	0.020 4
St. E	(0.007 6)	(0.007 5)	(0.009 1)	(0.011)	(0.038)	(0.007)
$\beta_{ij}$	0.942 6	0.965 1	0.965 9	0.925 7	0.901 9	0.974 9
St. E	(0.011 5)	(0.098 1)	(0.011 7)	(0.032)	(0.279)	(0.010 6)
LogL	1 752.158	664.024	597.649	18.598	23.875	120.409

从参数表 6 可以看出,动态模型拟合效果较为理想,系数比较显著,在考虑到条件 Copula 的因素以后,自由度明显增大,对数似然绝对值明显减少,参数显著性有所下降,进一步说明了在以基建为条件变量时,地产和运输,银行和运输,地产和银行的相关性显著减弱,进而表明基建在四维变量联合分布中的主导作用。值得注意的是第一棵树上的参数拟合值  $\alpha_{ij}$  与  $\beta_{ij}$  的和都很接近于 1,说明动态因素在动态相关系数的决定中相当活跃,  $\beta_{ij}$  远大于  $\alpha_{ij}$ ,且参数估计很显著,说明动态相关的一阶滞后在动态相关的求解中起到了决定性作用;而第二、三棵树上的系数显著性较第一棵树明显降低,说明拟合参数值的波动域更宽,动态因素更强,也说明在第二、三棵树上以基建为条件变量时,动态参数的影响增大。根据四维 C-Vine 动态结构的具体参数,画出相应的两两动态相关图 9,可以看出考虑到条件因素后的相关图波动更为剧烈,从某个方面说明基建变量在动态趋势中的支配作用。以上分析与我们的先验知识也是一致的,佐证了拟合结果的合理性。

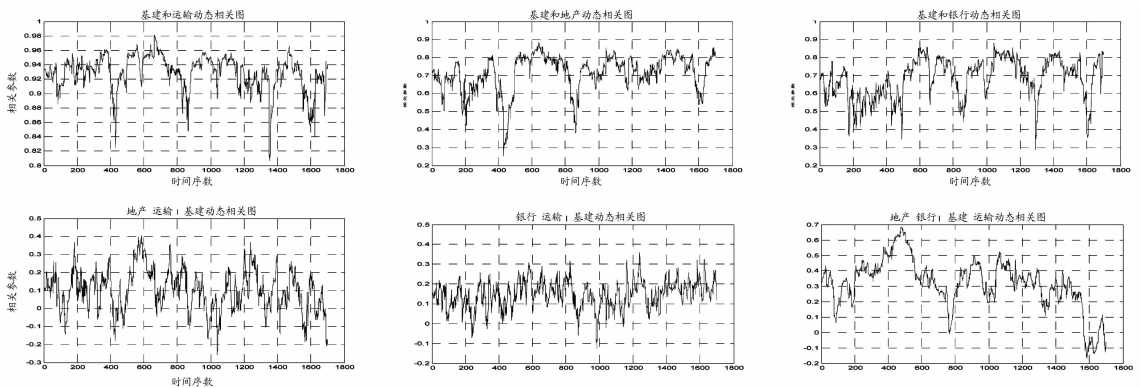


图 9 四维 C-Vine Pair-Copula 动态相关图

(七) VaR 的仿真与回溯检验

在确定了四维动态 C-Vine Copula 结构之后,按照该结构进行  $318 \times 10000$  次的动态仿真,并且按照第二节中步骤(2)(1)的对应参数进行还原,模拟出包含  $318 \times 10000$  个收益率仿真数据的数据集合,对于 318 次中每 10 000 次的仿真数据,通过计算机程序选择排序左端对应显著性水平的数值作为每 10 000 次仿真产生的一个 VaR 值,如此得到包含 318 个元素的 VaR 值序列组,然后运用 318 组样本外数据进行返回检验,并且与静态 C-Vine 的结果进行比较。

为了使研究更具有一般意义,随机选取两种组合比例分别为 (25%, 25%, 25%, 25%) 与 (10%, 30%, 20%, 40%),所产生的对应的组合风险序列分别为 VaR1 和 VaR2。

按照如上的方法进行仿真,求得 1%, 5%, 10% 三种显著性水平两种比例的 VaR 序列组 VaR1 和 VaR2,与相应的样本外数据比较。以 5% 显著性水平 VaR1 的超越情形为例分别绘制动态和静态模型的超越图 10-图 11,比较发现,动态模型较之静态模型更能灵活捕捉变化的风险。最后运用 Kupiec Test 方法进行回溯检验,得到无条件覆盖(UC)检验结论,如表 7 所示。而表 8 则证明了动态 C-Vine 模型风险计算结果的稳定性。表 7 中可以看出,动态 C-Vine 模型每次都能通过回溯检验,而静态 C-Vine 每次都不能通过检验,尽管其稳定性较好,仍然认为动态 C-Vine 可以用于股市行业风险组合计量,静态 C-Vine 模型则不符



合风险建模要求。

表7 回溯检验统计量比较表

10% 显著性水平下 VaR1 的回溯检验			5% 显著性水平下 VaR1 的回溯检验			1% 显著性水平下 VaR1 的回溯检验		
模型种类			模型种类			模型种类		
指标	静态 C - Vine	动态 C - Vine	指标	静态 C - Vine	动态 C - Vine	指标	静态 C - Vine	动态 C - Vine
期望超越	31	31	期望超越	15	15	期望超越	3	3
实际超越	12	41	实际超越	3	16	实际超越	0	1
UC 概率	0	0.10	UC 概率	0	0.98	UC 概率	0	0.15
结果	拒绝	不拒绝	结果	拒绝	不拒绝	结果	拒绝	不拒绝

10% 显著性水平下 VaR2 的回溯检验			5% 显著性水平下 VaR2 的回溯检验			1% 显著性水平下 VaR2 的回溯检验		
模型种类			模型种类			模型种类		
指标	静态 C - Vine	动态 C - Vine	指标	静态 C - Vine	动态 C - Vine	指标	静态 C - Vine	动态 C - Vine
期望超越	31	31	期望超越	15	15	期望超越	3	3
实际超越	12	29	实际超越	3	13	实际超越	0	1
UC 概率	0	0.60	UC 概率	0	0.44	UC 概率	0	0.15
结果	拒绝	不拒绝	结果	拒绝	不拒绝	结果	拒绝	不拒绝

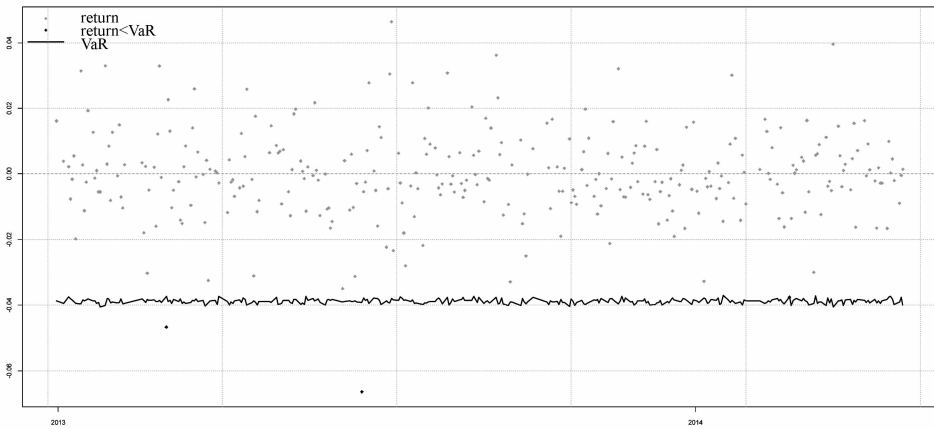


图10 静态 C - Vine VaR1 的超越图

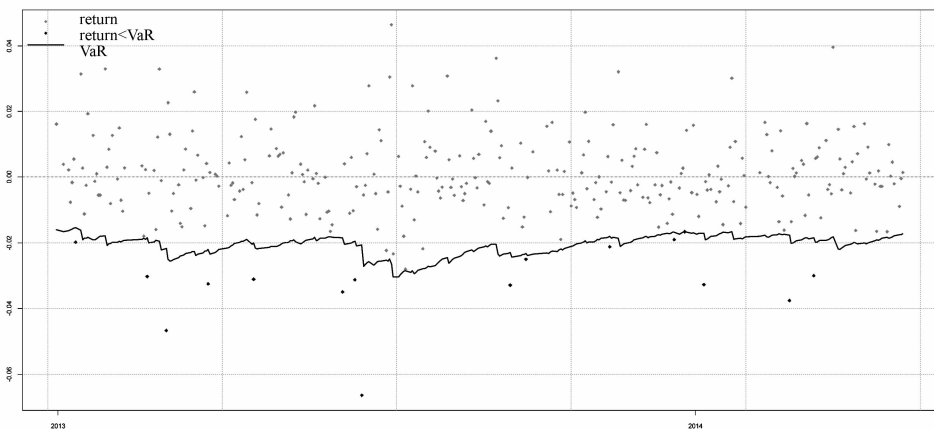


图11 动态 C - Vine VaR1 的超越图

### 六、动态 C - Vine Copula 的风险效果分析

根据投资学理论,我们知道马科维茨投资组合有效集的最优边界即为有效前沿。结合上文,基于动态 C - Vine Copula 模型的 VaR 能够在相应的显著性水平下准确地捕捉到左尾极端风险事件,据此可以绘制出原始投资组合和动态 C - Vine Copula 的资产组合有效前沿图 12。图中与静态 C - Vine Copula 作比较,发现动态和静态 C - Vine 的资产有效前沿相较于原始投资组合均发生了形状改变。静态 C - Vine 由于模型设定问题,对于原始投资组合的风险倾向没有准确刻画,而动态 C - Vine 则完全摆脱了原始投资组合的轨迹,显示理性投资者因为更准确地捕捉到风险而表现出风险厌恶,资产组合有效前沿完全左移,投资者因此采取保守型投资策略,规避风险带来的损失。也进一步说明基于动态 C - Vine Copula 模型的风险计量对于投资者来说具有重要的现实意义。

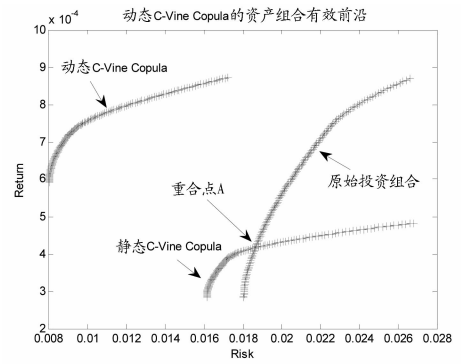


图 12 资产组合有效前沿图

表 8 两种组合动态 C - Vine VaR 的统计比较

指标	数量	极小值	极大值	均值	离差
10% VaR1	318	-0.030	-0.008	-0.015	0.004
10% VaR2	318	-0.031	-0.009	-0.017	0.004
5% VaR1	318	-0.043	-0.013	-0.023	0.006
5% VaR2	318	-0.044	-0.013	-0.025	0.005
1% VaR1	318	-0.075	-0.026	-0.043	0.010
1% VaR2	318	-0.078	-0.029	-0.047	0.009

### 七、研究结论与意义

本文运用两阶段建模方法开展研究:第一阶段采用 ARMA - GJR - GARCH 模型和 GPD 模型拟合实现 PIT 序列;第二阶段进行四维变量的动态 C - Vine Copula 建模。之后进行风险变量的仿真逆运算,并且与静态 C - Vine Copula 的风险效果作比较,进一步分析资产组合前沿曲线的移动情形。

本文研究得出如下结论:(1) 动态 C - Vine Copula 模型可以用于股市行业组合风险计量,而静态 C - Vine Copula 模型则不满足行业组合风险计量要求;(2) 动态 C - Vine Copula 模型的有效前沿摆脱了原始投资组合的轨迹完全左移,显示理性投资者因为更准确地捕捉到风险而表现出风险厌恶;而静态 C - Vine 由于模型设定问题,无法准确刻画资产组合有效前沿,相较于原始投资组合没有固定的风险偏好或者厌恶倾向。

本文的研究意义在于:(1) 研究过程中首次将极值理论 GPD 模型与高维动态 C - Vine Copula 结合起来开展研究,并与相应的静态 C - Vine Copula 模型结论作比较;(2) 研究结果为揭示多维行业风险之间的相依性提供了一种高效可行的方法,具有广阔的经济含义,可以借助高维动态 C - Vine Copula 方法进行有主导变量的多维风险组合计量,以此计量结果为依据进一步估算经济资本。结合当前的巴塞尔新资本协议的全面风险管理要求,该研究结论对于金融监管部门的多维行业风险监测与调控具有一定的借鉴价值。

#### 参考文献:

[1] SKLAR A. Fonctions de répartition à n dimension set leursmarges[J]. Publication de l' Institut de Statistique de L' Universit - é de Paris, 1959(8) :229 - 231

[2] EMBRECHTS P, McNEIL A, STRAUMANN D. Correlation: Pitfalls and alternatives[J]. Risk, 1999, 12(5) :69 - 71.

[3] BOUYÉ E, DURRLEMAN V, NIKEGHBALI A. Copulas for finance: A reading guide and some applications[R]. Working Paper, Financial Econometrics Research Centre, City University Business School, 2000.

[4] ROCKINGER M, JONDEAU E. Conditional dependency of financial series; An application of copulas[R]. Working paper, HEC School of Management, France, 2001.

[5] ROMANO C. Applying copula function to risk management[R]. Rome: University of Rome, 2002.

- [6] EMBRECHTS P, HÖING A, JURI A. Using copulae to bound the value-at-risk for functions of dependent risks[J]. *Finance and Stochastics*, 2003(7): 145 – 167.
- [7] RODRIGUEZ J C. Measuring financial contagion: A copula approach[J]. *Journal of Empirical Finance*, 2007, 14(3): 401 – 423.
- [8] GHORBEL A, TRABELSI A. Optimal dynamic hedging strategy with futures oil markets via FIEGARCH – EVT copula models[J]. *International Journal of Managerial and Financial Accounting*, 2012, 4(1): 1 – 28.
- [9] SO K P, YEUNG Y T. Vine-copula GARCH model with dynamic conditional dependence[J]. *Computational Statistics and Data Analysis*, 2014, 76: 655 – 671.
- [10] REBOREDO J C, UGOLINI A. Downside/upside price spillovers between precious metals: A vine copula approach[J]. *North American Journal of Economics and Finance*, 2015, 34: 84 – 102.
- [11] 杜子平, 闫鹏, 张勇. 基于“藤”结构的高维动态 copula 的构建[J]. *数学的实践与认识*, 2009, 39(10): 96 – 102.
- [12] 曹洁, 程希骏. 基于时变 Pair-copula 的多资产投资组合 VaR 分析[J]. *中国科学技术大学学报*, 2011, 41(12): 1047 – 1051.
- [13] BERGER T. Forecasting value-at-risk using time varying copulas and EVT return distributions[J]. *International Economics*, 2013, 133: 93 – 106.
- [14] 李强, 周孝华, 张保帅. 基于 Copula 的我国台湾和韩国股票市场相关性研究[J]. *管理工程学报*, 2014, 28(2): 100 – 107.
- [15] DOMINO K, BŁACHOWICZ T. The use of copula functions for modeling the risk of investment in shares traded on the Warsaw stock exchanges[J]. *Physica A: Statistical Mechanics and its Applications*, 2014, 413: 77 – 85.
- [16] 刘琼芳, 张宗益. 基于 Copula 房地产与金融行业的股票相关性研究[J]. *管理工程学报*, 2011, 25(1): 165 – 169.
- [17] 赵宁, Winston T Lin, 孙雪卿. 基于 Copula 贝叶斯估计的行业风险差异分析[J]. *数学的实践与认识*, 2015, 45(10): 17 – 27.
- [18] 张帮正, 魏宇. 基于 R\_Vine Copula 方法的上证行业指数相关性研究[J]. *北京理工大学学报(社会科学版)*, 2015, 17(3): 100 – 108.
- [19] RIGHI M B, CERETTA P S. Analyzing the dependence structure of various sectors in the Brazilian market: A Pair Copula Construction approach[J]. *Economic Modelling*, 2013, 35: 199 – 206.
- [20] BARTRAM S M, WANG Y H. European financial market dependence: An industry analysis[J]. *Journal of Banking and Finance*, 2015, 59: 146 – 163.
- [21] AAS K, CZADO C, FRIGESSI A, BAKKEN H. Pair-copula constructions of multiple dependence[J]. *Insurance: Mathematics and Economics*, 2009, 44(2): 182 – 198.
- [22] KUPIEC P. Techniques for verifying the accuracy of risk measurement models[J]. *Journal of Derivatives*, 1995(2): 73 – 84.
- [23] GLOSTENL R, JAGANNATHAN R, RUNKLE D E. On the relation between the expected value and the volatility of the nominal excess return on stocks[J]. *The Journal of Finance*, 1993, 48(5): 1779 – 1801.
- [24] PEIRÓ A. Skewness in financial returns[J]. *Journal of Banking and Finance*, 1999, 23(6): 847 – 862.

## A study on industry portfolio risk in China stock market based on high dimensional dynamic C – Vine Copula model

HAN Chao, YAN Taihua

(School of Economics and Business Administration, Chong Qing University, Chongqing 400044, P. R. China)

**Abstract:** Stock market is a “weatherglass” of macro-economy, portfolio risk measurement of different industries in stock market is very important for investors whether in financial markets or in real economy. This article measures multi-dimensional industries portfolio risk by the frontier method of high dimensional dynamic C – Vine Copula, and compares the results to static model. The result shows that VaR series got from high dimensional dynamic C – Vine Copula can get through UC and stability tests, while the static model can not. Thus we can get the conclusion that high dimensional dynamic C – Vine Copula performs better than static model and can be used as a new method for portfolio risk measurement of different industries in stock market.

**Key words:** high dimensional dynamic C – Vine Copula; GPD; portfolio risk measurement; VaR

(责任编辑 傅旭东)