

# 基于 Copula 的沪深股市相依结构与相关模式研究

卢俊香, 武宇, 杜艳丽

(西安工程大学理学院, 西安 710048)

**摘要:**针对传统 Pearson 线性相关系数与 Granger 因果分析法的不足,采用一种特殊的相关性分析方法—Copula 函数方法对沪深股市相关结构与相关模式进行研究。首先用核密度估计方法对 Copula 函数的边缘分布进行估计,再结合秩相关系数对数据拟合较好的 Copula 函数进行选择,最后用离散  $L_2$  范数评价方法对其拟合程度进行检验。研究发现, $t$ -Copula 可以较好地拟合沪深股市的日收益率序列,沪深股市日收益率序列呈现出较强的相关性以及对称的尾部相关性,当沪深两市出现大幅震荡时,两市收益率的协同作用将大幅增强。

**关键词:**Copula;相关性;收益率;模型选择;拟合度检验

**中图分类号:**F831

**文献标志码:**A

## 引言

Copula 函数作为相关分析和多变量建模的一种工具,处理非线性非对称的数据有其独特的优势,因此国内外对 Copula 方法及其应用有广泛关注,其应用范围也不断扩展。

1959 年,Sklar A 提出将一个  $k$  元联合分布函数分解成其  $k$  个边缘分布函数和一个 Copula 函数,而这个 Copula 函数恰能描述变量间的相关性<sup>[1]</sup>。1999 年,Nelson R B 对函数的含义和性质做了全面总结<sup>[2]</sup>。Dias, Emberches 等将 Copula 函数与 GARCH 模型相结合使得能准确的刻画时间序列间的关系<sup>[3-4]</sup>。Ding 和 Engle 等研究了资产间的相关性及模型的估计和检验<sup>[5]</sup>。随后 Copula 被应用到投资组合、风险管理等重要领域<sup>[6-10]</sup>。国内的张尧庭教授最先将 Copula 函数应用到相关分析中<sup>[11]</sup>。韦艳华和张世英把时间序列与 Copula 函数结合在一起,对变量间的条件相关性进行分析<sup>[12-13]</sup>。周艳菊等提出基于 Kendall $\tau$ 、尾相关系数的 Copula 的非参数估

计,对道琼斯工业指数和恒生指数做实证分析<sup>[14]</sup>。赵丽琴和张连增等对 Copula 函数的参数估计做了分析,提出对 Copula 函数的参数采用参数估计、半参数估计以及非参数估计的方法<sup>[15-16]</sup>。由于 Copula 函数的特点突出,近年来在诸多领域都有其研究应用<sup>[17-21]</sup>。

本文利用 Copula 函数,以深证 300 指的日收益率和上证 B 指的日收益率为边缘分布建模,研究沪深股指的相关性,首先运用核密度估计法对 Copula 函数的边缘分布进行估计,在对 Copula 函数参数估计的基础上,结合秩相关系数对沪深收益率拟合较好 Copula 函数进行选择,最后通过离散  $L_2$  范数进行拟合优度检验。结果表明, $t$ -Copula 可以较好的拟合沪深股市的日收益率序列,沪深股市日收益率序列呈现出较强的相关性以及对称的尾部相关性,当沪深两市出现大幅震荡时,其协同作用将大幅增强。

## 1 Copula 函数的定义<sup>[2]</sup>

$N$  元 Copula 函数是指具有以下性质的函数  $C$ :

收稿日期:2016-02-23

基金项目:陕西省教育厅专项科研基金(14JK1299)

作者简介:卢俊香(1980-),女,山东沂水人,副教授,博士,主要从事金融与统计方面的研究,(E-mail)jun-xianglu@163.com;

武宇(1991-),女,陕西渭南人,硕士生,主要从事金融与统计方面的研究,(E-mail)1069432692@qq.com

- (1)  $C = I^N = [0, 1]^N$ ;
- (2)  $C$  对它的每一个变量都是递增的;
- (3)  $C$  的边缘分布  $C_n(\cdot)$  满足:  $C_n(u_n) = C(1, \dots, 1, u_n, 1, \dots, 1) = u_n$ , 其中,  $u \in [0, 1], n \in [0, N]$ 。

显然,若  $F_1(\cdot), \dots, F_N(\cdot)$  是一元分布函数,令  $u_n = F_n(x_n)$  是随机变量,则  $C(F_1(\cdot), \dots, F_N(\cdot))$  是一个具有边缘分布  $F_1(\cdot), \dots, F_N(\cdot)$  的多元分布函数。

## 2 常见 Copula 函数

### 2.1 椭球 Copula 函数

椭球 Copula 函数主要有正态 Copula 和  $t$ -Copula。

#### 2.1.1 正态 Copula(Gaussian Copula)

多元正态 Copula:

$$C_{Ga}(u_1, \dots, u_N; \rho) = \frac{1}{(2\pi)^{\frac{N}{2}} \sqrt{|\rho|}} \int_{-\infty}^{\varphi^{-1}(u_1)} \dots \int_{-\infty}^{\varphi^{-1}(u_N)} \exp\left(-\frac{1}{2}x^{-1}\rho^{-1}x\right) dx_1 \dots dx_N \quad (1)$$

其中,  $\rho$  为对角线上元素为 1 的对称正定矩阵,  $|\rho|$  表示矩阵  $\rho$  相对应的行列式的值,  $\varphi_p$  是相关矩阵为  $\rho$  的  $N$  元高斯分布函数,  $\varphi^{-1}$  是一元标准正态分布函数的分位数函数,  $x = (x_1, x_2, \dots, x_N)$ 。

#### 2.1.2 $t$ -Copula

自由度为  $v$  的多元  $t$ -Copula:

$$C_t(u_1, \dots, u_N; \rho, v) = \frac{\Gamma\left(\frac{v+N}{2}\right) |\rho|^{-1/2} T_v^{-1}(u_1) \dots T_v^{-1}(u_N)}{\Gamma\left(\frac{v}{2}\right) (v\pi)^{N/2}} \int_{-\infty}^{\frac{v+1}{v} T_v^{-1}(u_1)} \dots \int_{-\infty}^{\frac{v+1}{v} T_v^{-1}(u_N)} \left(1 + \frac{1}{v} x' \rho^{-1} x\right)^{-\frac{v+N}{2}} dx_1 dx_2 \dots dx_N \quad (2)$$

其中,  $\Gamma$  是 Gamma 函数,  $T_v^{-1}$  是自由度为  $v$  的一元标准  $t$  分布函数的反函数。图 1 和图 2 为 1500 个  $\rho = 0.6$  时二元正态 Copula 与二元  $t$ -Copula 的概率密度函数图。

由图 1 和图 2 可知:取相同的相关系数时,二元 Gaussian Copula 函数和二元  $t$ -Copula 函数具有对称的尾部,其中二元  $t$ -Copula 函数具有较厚的尾部,对随机变量间的尾部相关变化较为敏感,能更好的捕捉到随机变量之间的对称尾部相关关系。而二元正态 Copula 函数在分布的尾部变量间是渐进独立的,所以适合于具有对称尾部,但上、下尾部渐进独立的情形。这两种 Copu-

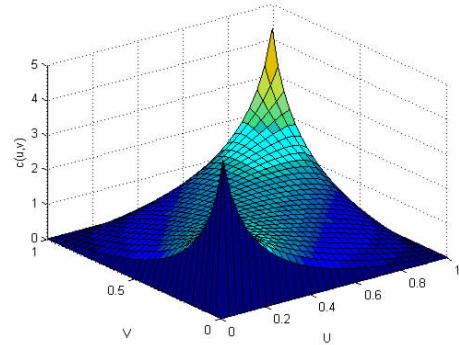


图 1 Gaussian Copula 函数概率密度函数图

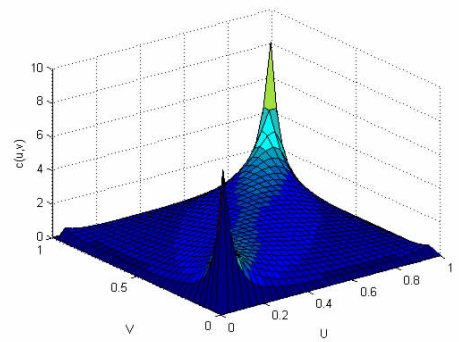


图 2  $t$ -Copula 函数概率密度函数图

la 函数都无法捕捉到随机变量间非对称的相关关系。

### 2.2 经验 Copula 函数

若  $X, Y$  的经验分布函数分别为  $\hat{F}_n(x), \hat{G}_n(y)$ , 则样本的经验 Copula  $\hat{C}_n(u, v)$  定义为:

$$\hat{C}_n(u, v) = \frac{1}{n} \sum_{i=1}^n I\{\hat{F}_n(x_i) \leq u, \hat{G}_n(y_i) \leq v\} \quad u, v \in [0, 1] \quad (3)$$

其中,当  $\{\hat{F}_n(x_i) \leq u, \hat{G}_n(y_i) \leq v\}$  时,  $I\{\hat{F}_n(x_i) \leq u, \hat{G}_n(y_i) \leq v\} = 1$ , 否则  $I\{\hat{F}_n(x_i) \leq u, \hat{G}_n(y_i) \leq v\} = 0$ 。

## 3 实证分析

### 3.1 描述性统计

选取 2009 年 12 月 3 日至 2015 年 11 月 6 日上证 B 指与深证 300 指的日收盘价共 1417 组有效数据,数据来源同花顺。采取对原序列进行一阶对数差分并乘以 100:

$$R_t = 100 \times (\ln P_t - \ln P_{t-1}) \quad (4)$$

得到对数收益率将对数收益率序列样本分别记为  $\{x_i\}$  和  $\{Y_i\}$ 。经计算可知沪深两市的每日收益率数据的均值都比较小,标准差比较大,峰度都大于 5,偏度都大于

0.5, 并且 JB 统计量 p 值均小于 0.001, 说明两收益率序列均不服从正态分布, 由于上证 B 指与深证 300 指的日收益率分布具有尖峰厚尾的特性且拒绝正态性假设, 故很难确定总体的分布, 所以采用核光滑方法估计总体的分布。

### 3.2 边缘分布估计

用核密度法估计 X, Y 的联合分布密度函数  $f(x, y)$ , 首先选取正态核  $K(x, y) = \frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right)$  作为核密度函数, 取窗宽  $h_1 = 0.3488, h_2 = 0.2421$ , 用样本经验分布作为边缘分布的近似, 用核密度方法估计边缘分布, 由图 3 与图 4 中知, 经验分布函数图和核密度估计所得的分布图几乎重合, 可见核密度估计效果良好。

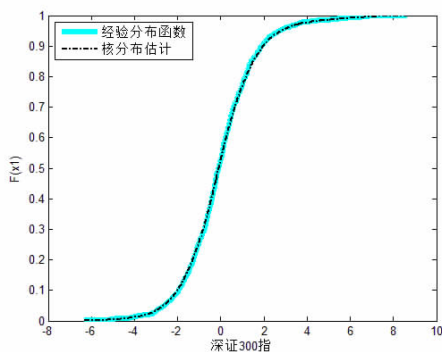


图 3 深证 300 指收益率的经验分布图和核分布估计图

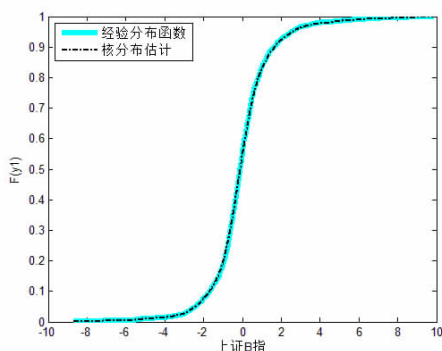


图 4 上证 B 指收益率的经验分布图和核分布估计图

### 3.3 Copula 函数参数估计及选取

用核密度估计方法确定深证 300 指收益率和上证 B 指收益率的边缘分布, 分别记为  $U = F(x)$  和  $V = G(y)$ , 就可以根据  $(U_i, V_i) (i = 1, 2, \dots, n)$  的二元频数直方图得到直观的相依关系图, 在频数直方图的基础上得到频率直方图, 而频率直方图可以作为 Copula 密度函数的估计, 频率直方图如图 5 所示。

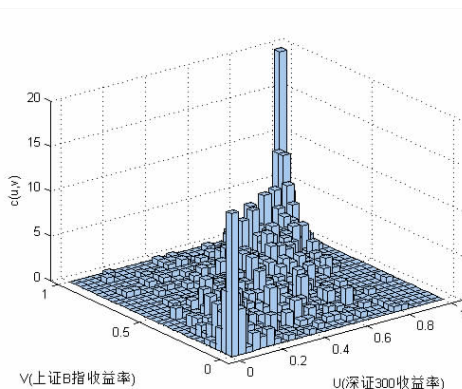


图 5 深证 300 指和上证 B 指收益率的频率直方图

由图 5 可知, 频率直方图具有基本对称的尾部, 据此初步选取能够刻画对称的尾部相关性的  $t$ -Copula、Gaussian Copula 函数描述和刻画深证 300 指和上证 B 指收益率的相依结构。

#### 3.3.1 $t$ -Copula 与 Gaussian Copula 的参数估计

经计算求得, 二元 Gaussian Copula 中的线性相关系数  $\rho$  的估计值为:

$$\hat{\rho}_{Ga} = \begin{pmatrix} 1.0000 & 0.8363 \\ 0.8363 & 1.0000 \end{pmatrix}$$

则二元 Gaussian Copula 的表达式为:

$$\hat{C}^{Ga}(u, v) = \int_{-\infty}^{\varphi^{-1}(u)} \int_{-\infty}^{\varphi^{-1}(v)} \frac{1}{2\pi \sqrt{1 - 0.8363^2}} \exp \left[ -\frac{s^2 - 2 \times 0.8363st + t^2}{2 \times (1 - 0.8363^2)} \right] ds dt \quad (5)$$

二元  $t$ -Copula 函数中的线性相关系数  $\rho$  和自由度  $k$  的估计值分别为:

$$\hat{\rho}_t = \begin{pmatrix} 1.0000 & 0.8540 \\ 0.8540 & 1.0000 \end{pmatrix}$$

$$\hat{k} = 3.2715 \approx 3$$

二元  $t$ -Copula 的表达式为:

$$\hat{C}^t(u, v) = \int_{-\infty}^{\hat{k}^{-1}(u)} \int_{-\infty}^{\hat{k}^{-1}(v)} \frac{1}{2\pi \sqrt{1 - 0.8540^2}} \left[ 1 + \frac{s^2 - 2 \times 0.8540st + t^2}{3 \times (1 - 0.8540^2)} \right]^{-(3+2)/2} ds dt \quad (6)$$

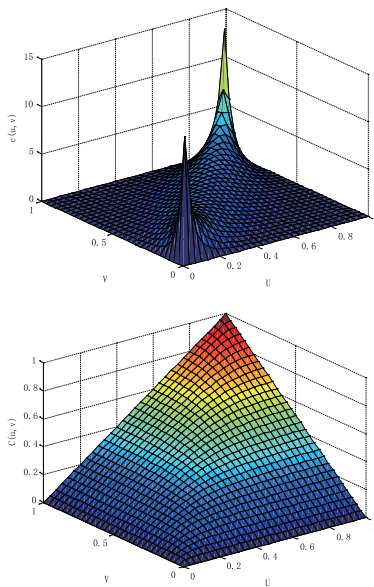
#### 3.3.2 二元 Gaussian Copula 与 $t$ -Copula 函数的选取分析

首先从秩相关方面考虑对 Copula 函数进行选择。通过估计出的参数可以求出 Gaussian Copula、 $t$ -Copula 以及原始序列所对应的 Kendall 秩相关系数与 Spearman 秩相关系数(表 1)。

表 1 Kendall 秩相关系数与 Spearman 秩相关系数

名称	Kendall	Spearman
原始预测	0.6459	0.8196
Gaussian Copula	0.6305	0.8239
$t$ -Copula	0.6517	0.8274

由表 1 可知沪深两市收益率之间有很强的相关性,将 Gaussian Copula、 $t$ -Copula 秩相关系数与由原始序列预测的秩相关系数进行对比,其中  $t$ -Copula 的秩相关系数更接近于原始序列预测的秩相关系数,由此可知从秩相关性角度出发, $t$ -Copula 比 Gaussian Copula 更好的拟合了深证 300 指和上证 B 指收益率之间的秩相关性。根据估计参数可得到能描述沪深收益率相关性的  $t$ -Copula 密度函数图与分布函数图(图 6)。

图 6 二元  $t$ -Copula 密度函数图和分布函数图

### 3.4 拟合度检验

通过距离法对选出较优的 Copula 函数进行拟合程度检验,通过计算 Copula 函数  $C$  与经验 Copula 函数  $\hat{C}$  之间的距离(多取  $L_2$  范数距离),该距离越小,对应 Copula 函数的拟合程度越好。

Copula 函数  $C$  与经验 Copula 函数  $\hat{C}$  的  $L_2$  范数距离定义为:

$$d^2 = \sum_{i=1}^n |\hat{C}_n(u_i, v_i) - C(u_i, v_i)|^2 \quad (7)$$

经计算, Gaussian Copula、 $t$ -Copula 函数与经验 Copula 分布函数之间的  $L_2$  范数距离分别为 0.0387 和 0.0188。所以,在  $L_2$  范数距离的评价标准上  $t$ -Copula 比 Gaussian Copula 能更好的拟合深证 300 指和上证 B 指日收益率序列。

## 4 结束语

通过 Gaussian Copula 与  $t$ -Copula 参数的估计,秩相关系数的分析、离散  $L_2$  范数的拟合程度检验,对沪深股市的相关性进行了定量与定性分析,研究结果表明: $t$ -Copula 可以较好的拟合沪深股市的日收益率序列,沪深股市日收益率序列呈现出很强相关性及对称的尾部相关性,当沪深两市出现大幅震荡时,两市收益率的协同作用将大幅增强。

### 参考文献:

- [1] SKLAR A. Functions de repartition(a) n dimensions et leurs marges[J]. Publication de l'Institut de Statistique de l'Universite de Paris, 1959, 8(12): 229-231.
- [2] NELSON R B. An introduction to copulas[M]. New York: Springer, 1999.
- [3] DIAS A, EMBRECHTS P. Dynamic Copula models for multivariate high-frequency data in finance[J]. ETH Zurich, 2004, 7(5): 24-37.
- [4] LI M, YANG L. Modeling the volatility of futures return in rubber and oil-A Copula-Based GARCH approach[J]. Economic Modeling, 2013, 7(16): 576-581.
- [5] DING Z X, ENGLE R F. Large scale conditional covariance matrix modeling, estimation and testing[J]. Academia Economic Papers, 2001, 29(2): 157-184.
- [6] ROMANO C. Applying Copula function to risk management[J]. Rome University, 2002, 26(3): 73-77.
- [7] UMBERTO Cherebin, ELISA Luciano, WALTER V. Copula Methods in Finance[J]. England: John Wiley & Sons Ltd, 1988.
- [8] LESNEVSK V, NELSON L B, STAUM J. Simulation of Coherent Risk Measures Based on Generalized Scenarios[J]. Management Science, 2007, 53(11): 1756-1769.
- [9] SONGSAK S, HUNG T N. Modeling and Dependency of Agricultural Price and Production Indices of Thailand: Static Versus Time-Varying Copulas[J]. Journal of Approximate Reasoning, 2013, 54(7): 793-808.
- [10] MARTIN H K, VOLATILITY S. Comovements and Contagion in Securitized Real Estate Markets[J]. Real Estate Financial Economic, 2013, 47(6): 1-35.
- [11] 张尧庭. 我们应该选用什么样的相关性指标[J]. 统

- 计研究,2002(9):41-44.
- [12] 韦艳华,张世英.多元 Copula-GARCH 模型及其在金融风险分析上的应用[J].数理统计与管理,2007,26(3):432-439.
- [13] 韦艳华,张世英.Copula 理论及其在金融分析上的应用[M].北京:清华大学出版社,2008.
- [14] 周艳菊,王亚滨,王宗润.基于 Copula 函数的股指尾部相关性研究—以道琼斯工业指数与恒生指数为例[J].数学的实践与认识,2012,12(5):19-26.
- [15] 赵丽琴,籍艳丽.Copula 函数的非参数核密度估计[J].统计与决策,2009,9(2):29-33.
- [16] 张连增,胡祥.Copula 的参数与半参数估计方法的比较[J].统计研究,2014,31(2):91-95.
- [17] 杨湘豫,周再立.基于 Copula-TARCH 的开放式基金投资组合风险的实证分析[J].系统工程,2011,6(7):65-70.
- [18] 高杰,付冀.基于 Copula 模型的汇率与股市波动溢出效应分析[J].求索,2011,9(3):38-39.
- [19] 李强,周孝华.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.
- [20] 邱树萍,吴薇薇,侯美丽.基于 Copula 函数的航班延误相关性分析[J].武汉理工大学学报,2015,39(1):117-119.
- [21] 陈晓雷,李自胜,武鑫,等.基于 Copula 函数的中美大豆期货波动溢出效应研究[J].科技通报,2015,31(3):2-3.

## Research on The Dependence Structure and Correlation Pattern of Shanghai and Shenzhen Stock Market Based on Copula

LU Junxiang, WU Yu, DU Yanli

(School of Science, Xi'an Polytechnic University, Xi'an 710048, China)

**Abstract:** For the shortage of traditional Pearson linear correlation coefficient and Granger method of Cause and effect analysis, a special correlation analysis method — the method of Copula function was used to study the dependence structure and correlation pattern of Shanghai and Shenzhen stock market. First, the method of kernel density estimation was used to estimate marginal distributions of Copula function, then the rank correlation coefficient was combined with to choose a better copula function that can fit the data well, at last discrete  $L_2$  norm was used to fit degree test. The study found that  $t$ -Copula can fit the Shanghai and Shenzhen stock market's daily return rate sequence better. The return rate sequence shows a strong correlation and symmetrical tail dependence between Shanghai and Shenzhen stock market. When the Shanghai or Shenzhen stock appear sharp situation, the synergistic effect will be greatly enhanced.

**Key words:** Copula; correlation; return rate; model selection; fitting degree test