文章编号:1671-9352(2007)12-0063-06

基于 Gaussian Copula 与 t-Copula 的沪深股指相关性分析

杨兴民[1,2],刘保东[1*],李娟[2]
(1. 山东大学　数学与系统科学学院,山东　济南　250100; 2. 鲁东大学　数学与信息学院,山东　烟台　264025)

摘要:针对沪深股指,讨论了 Gaussian Copula 与 t-Copula 的密度函数,并进行相关性建模,采用二步估计法对所建模型进行参数估计并给出了相关性指标。最后,通过 Monte Carlo 模拟的方法比较了 Copula 关联结构之间的差异。

关键词: Gaussian Copula; t-Copula; 相关性分析
中图分类号:O212　　**文献标志码:**A

Correlation analysis of the Shanghai-Shenzhen stock index based on Gaussian Copula and t-Copula

YANG Xing-min[1,2], LIU Bao-dong[1*], LI Juan[2]
(1. School of Mathematics and System Sciences, Shandong University, Jinan 250100, Shandong, China;
2. School of Mathematics and Information, Ludong University, Yantai 264025, Shandong, China)

Abstract: Gaussian Copula and t-Copula density functions were discussed according to the Shanghai-Shenzhen stock index, and a correlation model was provided. The parameter of this model was estimated by the two-step estimating method and the relative index was given. Finally, the difference between Copula correlation structures was compared by the Monte Carlo method.

Key words: Gaussian Copula; t-Copula; correlation analysis

0　前言

相关性分析在金融分析领域的应用非常广泛,如投资组合的分析、资产定价及违约风险等。我们常常用线性相关系数来分析随机变量间的相关关系,即使在多元分析中谈到典型相关系数,实际上也是基于二元的相关系数。随着相关研究的深入,用线性相关系数来度量相关性,已经不能完全适应某些金融现象的要求。而 Copula 函数为相关性分析提供了一条新的途径。因为 Copula 函数对于随机变量的严格单调增变换是不变的,所以由 $C(u_1,\cdots u_n)$导出的相关性指标,如 Kendall 的 τ,Spearman 的 ρ,都是单调不变的相关性度量,比常用的相关系数更加合乎实际要求。

对 Copula 函数的研究,可以追述到 1959 年, Sklar[1] 指出可以将一个联合分布分解为它的 k 个边缘分布和 1 个 Copula 函数, 其中 Copula 函数描述变量间的相关结构。1999 年,Nelsen[2] 提出了 Copula 理论在金融上的应用。此后 Copula 理论在统计上得到广泛应用,并开始应用于金融领域。把 Copula 理论引入到金融,目前已经取得了一些成果。Bouye E[3] 系统介绍了 Copula 理论在金融中的一些应用。Claudio Romano[4] 对意大利

收稿日期:2007-05-14
基金项目:山东省自然科学基金资助项目(Y2006A17);国家统计局资助项目(2006C08)
作者简介:杨兴民(1966-),男,副教授,主要研究方向:经济数学方法.Email:yxmyt@sina.com
* 通讯作者:刘保东(1964-),男,副教授,博士,主要研究方向:资源与环境数学、数学建模.Email: baodong@sdu.edu.cn

股市收益率进行了 Copula 分析，Embrechts[5] 等人用 Copula 进行了风险分析，Roberto[6] 对 Copula，特别是对 Archimedean Copula 做了较好的总结。Beatriz[7] 再次用 Copula 函数度量了金融风险，并比较了几种不同的 Copula。Stefano Demarta[8] 介绍了 t-Copula 函数的一些性质，提出了 t-Copula 函数的应用优势。

国内对 Copula 的研究还比较少，张尧庭[11] 从理论上探讨了 Copula 在金融上应用的可行性，韦艳华[13,14] 等人用 Copula 函数对相关性建模，得到了较好的结果，吴振翔[15] 等人用 Archimedean Copula 分析了外汇市场中最小风险的投资组合，陈守东[17]，何旭彪[16] 等用 Monte Carlo 方法对 Copula 函数度量风险价值进行了研究，对 Copula 函数分类研究以及相关的实证研究正处在深入研究阶段。

本文首先针对沪深股指，讨论了 Gaussian Copula 与 t-Copula 的密度函数；其次以上证综合指数的收益与深圳成分指数的收益作为边缘分布建模，研究沪深股指的相关性；然后采用二步估计对 Copula 函数进行参数估计，并在参数估计的基础上，给出了沪深股指相关性指标；最后利用 Monte Carlo 方法比较了两类 Copula 关联结构的差异。

1 密度函数与相关性指标计算

1.1 定义

定义 1 $(X,Y)^{\mathrm{T}}$，$(\bar{X},\bar{Y})^{\mathrm{T}}$ 是独立同分布的向量。则 Kendall 的 τ 可以定义为

$$\tau(X,Y)=P\{(X-\bar{X})(Y-\bar{Y})>0\}-P\{(X-\bar{X}(Y-\bar{Y}))<0\}。$$

定义 2 令 $(X,Y)^{\mathrm{T}}$ 是连续随机变量的向量，边缘分布函数分别为 F,G，则

(1) $(X,Y)^{\mathrm{T}}$ 的上尾相关系数为

$$\lim_{u\to1^-}P\{Y>G^{-1}(u)\mid X>F^{-1}(u)\}=\lambda_U。$$

如果 $\lambda_U\in(0,1]$，X,Y 称为上尾相关；如果 $\lambda_U=0$，X,Y 称为上尾独立。

(2) $(X,Y)^{\mathrm{T}}$ 的下尾相关系数为

$$\lim_{u\to0^+}P\{Y<G^{-1}(u)\mid X<F^{-1}(u)\}=\lambda_l。$$

如果 $\lambda_l\in(0,1]$，X,Y 称为下尾相关；如果 $\lambda_l=0$，X,Y 称为下尾独立。

根据定义 2，因为

$$P\{Y>G^{-1}(u)\mid X>F^{-1}(u)\}=\frac{P\{Y>G^{-1}(u),\ X>F^{-1}(u)\}}{P\{X>F^{-1}(u)\}}=$$

$$\frac{1-P\{X\leqslant F^{-1}(u)\}-P\{Y\leqslant G^{-1}(u)\}+P\{X\leqslant F^{-1}(u),\ Y\leqslant G^{-1}(u)\}}{1-P\{X\leqslant F^{-1}(u)\}}=$$

$$\frac{1-2u+C(u,u)}{1-u}。$$

同理，$P\{Y<G^{-1}(u)\mid X<F^{-1}(u)\}=\dfrac{P\{Y<G^{-1}(u),X<F^{-1}(u)\}}{P\{X<F^{-1}(u)\}}=\dfrac{C(u,u)}{u}$。

则有下述等价定义：

定义 3 有二元 Copula 函数 C 使得 $\lim\limits_{u\to1^-}\dfrac{1-2u+C(u,u)}{1-u}=\lambda_U$ 存在，则如果 $\lambda_U\in(0,1]$，X,Y 称为上尾相关；如果 $\lambda_U=0$，X,Y 称为上尾独立。二元 Copula 函数 C，存在 $\lim\limits_{u\to0^+}\dfrac{C(u,u)}{u}=\lambda_L$，如果 $\lambda_L\in(0,1]$，X,Y 称为下尾相关；如果 $\lambda_L=0$，X,Y 称为下尾独立。

1.2 密度函数的计算

对于 Gaussian Copula：

$C_{Ga}^{\rho}(-1<\rho<1)$ 的分布函数为

$$C_{Ga}^{\rho}(u_1,u_2)=\int_{-\infty}^{\Phi^{-1}(u_1)}\int_{-\infty}^{\Phi^{-1}(u_2)}\frac{1}{2\pi\sqrt{1-\rho^2}}\mathrm{e}^{\frac{-(s^2-2\rho st+t^2)}{2(1-\rho^2)}}\mathrm{d}s\mathrm{d}t。\tag{1}$$

因为

$$F(x_1,\cdots x_n,\cdots x_N) = C(F_1(x_1),\cdots F_n(x_n),\cdots F_N(x_N)),$$

所以

$$f(x_1,x_2,\cdots x_n) = c(F_1(x_1),F_2(x_2),\cdots F_n(x_n))\prod_{i=1}^{n} f_i(x_i)。\tag{2}$$

而标准多元正态分布的密度函数为 $f(x) = \dfrac{1}{\sqrt{(2\pi)^N}\sqrt{|\boldsymbol{\rho}|}}\mathrm{e}^{-\frac{1}{2}x^{\mathrm{T}}\boldsymbol{\rho}^{-1}x}$,

一元标准正态的密度函数为 $f(x_n) = \dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}x_n^2}$,

则由式(2),有

$$\frac{1}{\sqrt{(2\pi)^N}\sqrt{|\boldsymbol{\rho}|}}\mathrm{e}^{-\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\rho}^{-1}x} = c(\Phi(x_1),\cdots\Phi_n(x_n),\cdots\Phi_N(x_N))\left(\prod_{n=1}^{N}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}x_n^2}\right)。$$

其中 $x_n = \Phi^{-1}(u_n), n = 1,2,\cdots N$; $\boldsymbol{x} = (x_1,x_2,\cdots x_n)^{\mathrm{T}}$,$\boldsymbol{\rho}$ 是一个对称矩阵,主对角线元素为 1。

二维情形下,令 $\Phi(x_1) = u_1, \Phi(x_2) = u_2$ 有:

$$c(u_1,u_2;\boldsymbol{\rho}) = \frac{1}{\sqrt{|\boldsymbol{\rho}|}}\frac{\mathrm{e}^{-\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\rho}^{-1}\boldsymbol{x}}}{\mathrm{e}^{-\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}},\tag{3}$$

其中 $x_n = \Phi^{-1}(u_n), (n = 1,2)$; $\boldsymbol{x} = (x_1,x_2)^{\mathrm{T}}$,$\boldsymbol{\rho}$ 是一个对称矩阵,主对角线元素为 1。式(3) 即为 Gaussian Copula 的密度函数。

对于 t-Student Copula: $C_t^{\rho,\nu}(-1 < \rho < 1)$ 的分布函数为

$$C_t^{\rho,v}(u_1,u_2) = \int_{-\infty}^{t_\nu^1}\int_{-\infty}^{t_\nu^{-1}(u_2)}\frac{1}{2\pi\sqrt{1-\rho^2}}\left[1-\frac{s^2-2\rho st+t^2}{\nu(1-\rho^2)}\right]^{\frac{-(\nu+2)}{2}}\mathrm{d}s\mathrm{d}t。$$

因为标准多元 t 分布的密度函数为:

$$f(x) = \frac{1}{\sqrt{|\boldsymbol{\rho}|}}\frac{\Gamma\left(\frac{\nu+N}{2}\right)\left(1+\frac{1}{\nu}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\rho}^{-1}\boldsymbol{x}\right)^{-\frac{\nu+N}{2}}}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu\pi)^N}},$$

一元标准 t 分布的密度函数为

$$f(x_n) = \frac{\prod_{n=1}^{N}\left(1+\frac{x_n^2}{\nu}\right)^{-\frac{\nu+1}{2}}\Gamma\left(\frac{\nu+1}{2}\right)^N}{\Gamma\left(\frac{\nu}{2}\right)^N\sqrt{(\nu\pi)^N}},$$

由式(2),有

$$\frac{1}{\sqrt{|\boldsymbol{\rho}|}}\frac{\Gamma\left(\frac{\nu+N}{2}\right)\left(1+\frac{1}{\nu}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\rho}^{-1}\boldsymbol{x}\right)^{-\frac{\nu+N}{2}}}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu\pi)^N}} = C(t_1(x_1),\cdots t_n(x_n),\cdots t_N(x_N))\frac{\prod_{n=1}^{N}\left(1+\frac{x_n^2}{\nu}\right)^{-\frac{\nu+1}{2}}\Gamma\left(\frac{\nu+1}{2}\right)^N}{\Gamma\left(\frac{\nu}{2}\right)^N\sqrt{(\nu\pi)^N}}。$$

其中 $x_n = t^{-1}(u_n), (n = 1,2,\cdots N)$, $\boldsymbol{x} = (x_1,x_2,\cdots x_n)^{\mathrm{T}}$,$\boldsymbol{\rho}$ 是一个对称矩阵,主对角线元素为 1。

二维情形下,令 $t(x_1) = u_1, t(x_2) = u_2$,有

$$c(u_1,u_2;\boldsymbol{\rho}) = \frac{1}{\sqrt{|\boldsymbol{\rho}|}}\frac{\Gamma(\nu+2)\left[\Gamma\left(\frac{\nu}{2}\right)\right]^2\left(1+\frac{1}{\nu}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\rho}^{-1}\boldsymbol{x}\right)^{-\frac{\nu+2}{2}}}{\left[\Gamma\left(\frac{\nu+1}{2}\right)\right]^N\Gamma\left(\frac{\nu}{2}\right)\prod_{n=1}^{N}\left(1+\frac{x_n^2}{\nu}\right)^{-\frac{\nu+1}{2}}},\tag{4}$$

其中 $x_n = t^{-1}(u_n), (n = 1,2)$, $\boldsymbol{x} = (x_1,x_2)^{\mathrm{T}}$,$\boldsymbol{\rho}$ 是一个对称矩阵,主对角线元素为 1。式(4) 即为 t-Copula 的密度函数。

1.3　相关性指标的计算

对于 Gaussian Copula:记

$$P\{V \leqslant v \mid U = u\} = \frac{\partial C(u,v)}{\partial u},\ P\{V > v \mid U = u \mid\} = 1 - \frac{\partial C(u,v)}{\partial u}。$$

根据定义 3

$$\lambda_U = \lim_{u\to 1}\frac{1-2u+C(u,u)}{1-u} = \lim_{u\to 1}\frac{2-2u+C(u,u)-1}{1-u} = -\lim_{u\to 1}\left(-2+\frac{1-C(u,u)}{1-u}\right) =$$

$$-\lim_{u\to 1}\left(-2+\frac{\partial}{\partial s}C(s,t)\mid_{s=t=u}\mid + \frac{\partial}{\partial t}C(s,t)\mid_{s=t=u}\mid\right) =$$

$$\lim_{u\to 1}(P\{V > u \mid U = u \mid\} + P\{V > u \mid v = u \mid\})。$$

因为 C 是可交换的,即 $$C(u,v) = C(v,u),$$

所以 $$\lambda_u = 2\lim_{u\to 1}P\{V > u \mid U = u\}。$$

又因为

$$\lim_{u\to 1}P\{V > u \mid U = u\} = \lim_{x\to\infty}P\{\Phi^{-1}(V) > x \mid \Phi^{-1}(U) = x\} = \lim_{x\to\infty}P\{X > x \mid Y = x\},$$

所以,在正态分布情形下

$$f(x\mid y) = \frac{f(x,y)}{f_Y(y)} = \frac{\dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-u_1^2)}{\sigma_1^2}-\frac{2\rho(x-u_1)(y-u_2)}{\sigma_1\sigma_2}+\frac{(y-u_2)^2}{\sigma_2^2}\right]}}{\dfrac{1}{\sqrt{2\pi}\sigma_2}\mathrm{e}^{-\frac{1}{2\sigma_2^2}(y-u_2)^2}} =$$

$$\frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}}\mathrm{e}^{\frac{-1}{2\sigma_1^2(1-\rho^2)}\left[x-u_1-\rho\frac{\sigma_1}{\sigma_2}(y-u_2)\right]^2}。$$

即: $$X\mid Y = y \sim N\left(u_1+\rho\frac{\sigma_1}{\sigma_2}(y-u_2),\sigma_1^2(1-\rho^2)\right)。$$

标准正态下: $$X\mid Y = y \sim N(\rho y, 1-\rho^2),$$

所以 $$\lambda_u = 2\lim_{u\to 1}P\{V > u \mid U = u\} = 2\lim_{x\to\infty}P\{X > x \mid Y = x\} =$$

$$2\lim_{x\to\infty}\Phi\left((x-\rho x)\mid\sqrt{1-\rho^2}\right) = 2\lim_{x\to\infty}\Phi(x\sqrt{1-\rho}\mid\sqrt{1+\rho})。$$

则对于 Gaussian Copula,当 $-1 < \rho < 1$ 时,$\lambda_u = 0$;

由椭圆 Copula 的对称性,当 $-1 < \rho < 1$ 时,$\lambda_l = 0$。

对于 t-Student Copula:$\lambda_u = 2\lim_{x\to\infty}P(Y > x \mid X = x)$。

$$E(Y\mid X = x) = \rho x,\ D(Y\mid X = x) = \left(\frac{\nu+x^2}{\nu+1}\right)(1-\rho^2),$$

则 $$\lambda_u = 2\lim_{x\to\infty}P(Y > x\mid X = x) = 2\lim_{x\to\infty}t_{\nu+1}\left(\left(\frac{\nu+1}{\nu+x^2}\right)^{\frac{1}{2}}\frac{x-\rho x}{\sqrt{1-\rho^2}}\right) = 2t_{\nu+1}\frac{\sqrt{\nu+1}\sqrt{1-\rho}}{\sqrt{1+\rho}}。$$

由椭圆 Copula 的对称性,得

$$\lambda_u = \lambda_l = 2t_{\nu+1}\frac{\sqrt{\nu+1}\sqrt{1-\rho}}{\sqrt{1+\rho}}。\tag{5}$$

可见,用 Gaussian Copula 建模方法,无法捕捉到尾部变化,而且尾部相关系数不存在。用 t-Copula 建模方法可以克服上述不足。

2 对沪深股指的相关性建模问题实证研究

本节以上证综合指数的收益与深圳成分指数的收益作为边缘分布建模,旨在进一步研究沪深股市的相关性。

我们选取 2000.01.04 ~ 2005.11.04 日上证综合指数和深圳成分指数的每日收盘价为样本,共 1289 组有效数据。将价格 p_t 定义为市场第 t 日指数收盘价,第 t 日收益 R_t 定义为 $R_t = 100(\ln p_t - \ln p_{t-1})$。