第 8 卷 第 1 期
2006 年 3 月

长安大学学报(社会科学版)
Journal of Chang'an University (Social Science Edition)

Vol 8 No. 1
Mar. 2006

【交通运输与经济】

金融市场相关性分析及其度量方法改进

战雪丽, 张世英, 张瑞锋
(天津大学 管理学院, 天津 300072)

摘 要: 介绍了目前测量金融市场相关性的几种分析方法, 尤其是定性方法和对线性相关性的刻画。引入连接函数(Copula)分析技术, 将随机变量的边缘分布借助于一个恰当的 Copula 函数得到其联合分布, 用以研究随机变量间的非线性相关结构。国内外实证研究结果表明, 与常见的相关性度量相比, 基于 Copula 相关性的分析方法适用范围更广, 对变量相关性的刻画更充分。介绍了目前常用的几种基于 Copula 相关性的度量方法, 并指出了其应用的领域和进一步研究的方向。

关键词: 应用经济学; 金融学; 金融市场; 随机变量; 连接函数; 条件概率
中图分类号: F830.9 文献标识码: A 文章编号: 1671-6248(2006)01-0051-04

Analysis on Financial Market Dependence and Improvement for Measuring Methods

ZHAN Xue-li, ZHANG Shi-ying, ZHANG Rui-feng
(School of Management, Tianjin University, Tianjin 300072, China)

Abstract: This paper introduces several methods for the financial markets, especially qualitative and linear methods. The Copula is introduced to establish the relationship between a multidimensional probability function and its lower dimensional margins, and from it the dependent structure can be discovered. Copula contains more dependent information of the random variables than the common dependent methods. The overseas and domestic demonstrations show that the Copula method can be used more widely and efficiently in practice. This paper also presents some dependent quantitative analysis methods based on Copula theory and shows some further-applied fields and researching areas.

Key words: applied economics; financial science; financial market; random varlable; Copula; conditional probability

0 引 言

相关性分析在金融活动中具有十分广泛的应用, 如投资组合分析、资产定价及风险分析等问题都用到相关性分析。目前在研究变量之间的相关性方法中, 常用的有线性相关系数和 Granger 因果分析方法。线性相关系数分析方法首先要求变量间的关系是线性的, 否则无法捕捉变量间非线性的关系; 其次要求方差有限, 而现实中很多分布的方差有时并不存在; 实际上相关性很强的变量, 用该方法分析得到的相关系数是 0。Granger 因果分析法通常只能给出定性的结论, 不能加以定量的描述。

实际上, 一组样本的相关系数总是可以计算的, 但它可能不反映什么关系, 即是说相关系数是 0, 但随机变量之间可能是有相关关系的。而用概率来反映随机变量之间的相关性, 对于任何分布的变量都是适用的。Copula 连接函数对于随机变量的严格单调增变换是不变的, 因此由它导出的度量相关性

收稿日期: 2005-06-15
基金项目: 国家自然科学基金资助项目(70471050); 河北省科技厅研发计划(05457242)
作者简介: 战雪丽(1979-), 女, 山东烟台人, 经济学博士研究生。

的指标是严格单调增变换下的相关性,比线性相关的范围要宽,并且由它导出的相关性度量并不是唯一的。首先,由 Copula 函数导出的一致性和相关性测度可以捕捉变量间非线性、非对称的相关关系,而且更容易捕捉分布尾部的相关关系。其次,Copula 模型不限制边缘分布的选择,而且 Copula 函数有很多分布族,当二元或多元正态分布假设被拒绝的时候,可以选用不同的边缘分布和 Copula 函数,不仅可以得到度量相关程度的相关参数,还可以得到描述相关模式的 Copula 函数,因此可以更全面地刻画随机变量间的相关关系。此外,Copula 函数中的相关参数与几种重要的一致性和相关性测度如 Kendall 的 τ、Spearman 的 ρ 和 Gini 系数常常有一一对应的关系,这种对应关系使不同的 Copula 函数之间具有了可比性。

1 相关性及其度量方法介绍

“相关”是指 2 个测量点数据或随机变量的波动方式是否一致[1]。如果数据或随机变量呈现相同的波动方式,即同时上升或者同时下降,则表示相关性强;反之,则表示相关性弱。在金融活动中,相关性分析往往需要有一个定量的指标来度量变量间的相关性强弱。

1.1 相关系数的描述

设随机变量 X 与 Y,则协方差为

$$\mathrm{Cov}(X,Y)=E(X-EX)(Y-EY) \tag{1}$$

相关系数为

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{DX}\,\sqrt{DY}} \tag{2}$$

由于任何 2 个随机变量 X、Y 的分布都是客观存在的,其方差 DX、DY 也是确定的,所以 ρ_{XY} 与 $\mathrm{Cov}(X,Y)$ 总是成正比的。显然,协方差 $\mathrm{Cov}(X,Y)$ 也是 1 个用来描述随机变量 X 与 Y 的相关程度的 1 个量。对于任何 2 个随机变量 X 与 Y 来说,均有以下结论:$\mathrm{Cov}(X,Y)$ 越大(小),X 与 Y 的相关程度就越高(低)。

但在实际中,人们总是用相关系数 ρ_{XY},而不用协方差 $\mathrm{Cov}(X,Y)$ 来判断 2 个随机变量 X 与 Y 的线性相关程度[2]。因为 ρ_{XY} 是 1 个无量纲的量,用它来描述 X 与 Y 的线性相关关系不受单位影响;而 $\mathrm{Cov}(X,Y)$ 则不同,它是 1 个有量纲的量,必须依赖于 X 与 Y 的度量单位。人们通常用 ρ_{XY},而不用 $\mathrm{Cov}(X,Y)$ 来判断 X 与 Y 的相关程度。

2 个随机变量 X、Y 的线性相关关系,不是普通意义下的关系,而是一种概率意义下的关系。这是因为 X 的取值是随机的,Y 的取值也是随机的。在平面 XOY 内,由 X、Y 所对应的点也不是 1 个普通的点,而是 1 个随机点(X,Y)。因此,所谓随机变量 X 与 Y 具有线性相关关系 $Y=aX+b$,实质上就是随机点(X,Y)在平面 XOY 内的散点分布在直线 $Y=aX+b$ 附近。从散点的分布趋势来看,它们与直线 $Y=aX+b$ 形状相似,这种相似程度的好坏,完全由 ρ_{XY} 的大小来决定。若 ρ_{XY} 越大,则相似的程度越高,线性相关的程度也越高,概率 $P(Y=aX+b)$ 也就越大;否则,结论相反。特别是,当 $\rho_{XY}=1$ 时,随机变量 X 与 Y 的线性相关的程度达到最高,概率 $P(Y=aX+b)$ 达到最大值 1。因此,在理论上 2 个随机变量的相关系数 ρ_{XY} 定义为式(2)。

可以看出,首先 X 与 Y 必须有二阶矩,即各自的方差以及协方差。但是金融市场中出现的不少数据往往是厚尾分布,他们的二阶距——方差是不存在的,有的分布可能连期望都不存在,因此无法用 ρ_{XY} 来反映相关性[3]。但是用概率来反映相关性,它不考虑变量的分布。

设 A、B 2 个事件,若 A、B 相互独立,那么就有

$$P(AB)=P(A)P(B) \tag{3}$$

如果 $P(AB)>P(A)P(B)$,可得到

$$P(B\mid A)=P(AB)/P(A)>P(B)$$

这表明 A 与 B 之间有一种正向的相关性。当 A 发生时,B 发生的可能性变大;同理可得,当 B 发生时,A 发生的可能性也变大。由此可以得到一种条件概率 $P(Y>b\mid X>a)$ 与边缘分布 $P(X>a)$ 之间的关系。可见,用概率大小反映变量之间的相关性比用式(2)计算的相关性系数所反映的变量之间的相关性具有更广泛的实用性和更深刻的意义。

1.2 金融时间序列相关关系检验

协整理论[4] 被认为是 20 世纪 80 年代计量经济学建模理论的一个重大发现,因为它打破了以经济理论为基础的传统建模做法,开辟了从经济变量数据所显示的关系出发确定变量之间关系的新途径。协整理论的经济意义在于它揭示了时间序列变量之间的长期稳定关系,它虽然具有协整关系的变量在短期具有各自的变动规律,但却存在着协调变化的趋势。

协整检验的首要任务就是检验时间序列的单整性,即检验 1 个非平稳序列经过差分后能否变为 1 个平稳序列。协整检验的目的在于检验非平稳时间序列之间是否存在长期稳定关系。

Granger(1969)提出的因果检验方法认为: 若在包含了变量 X 和 Y 过去信息的条件下, 对 Y 的预测效果要好于只单独由 Y 的过去信息对 Y 的预测, 则称 X 是 Y 的格兰杰原因, 否则称之为非格兰杰原因。因此, 这种检验方法只能对金融时间序列作出定性相关关系的分析, 不能给予定量的说明。目前对金融市场相关性的分析, 绝大部分要用到金融时间序列的分析方法, 因此对数据的处理最终得到的不是定量的指标结果, 也不能得出较准确的结论。

2　Copula 方法对相关性分析的改进

Sklar(1959)指出, 可以将 1 个联合分布分解为它的 k 个边缘分布和 1 个 Copula 函数, 这个 Copula 函数描述了变量间的相关性。实际上, Copula 函数是一种将联合分布与其各自的边缘分布连接在一起的函数, 因此也被称为连接函数。

2.1　Copula 定义及重要定理 Sklar 定理

令 F 为具有边缘分布 $F_1, \cdots, F_N$ 的联合分布函数, 那么存在 1 个 Copula 函数 C, 且满足

$$F(x_1, \cdots, x_n, \cdots, x_N)=C[F_1(x_1), \cdots, F_n(x_n), \cdots, F_N(x_N)] \tag{4}$$

若 $F_1, \cdots, F_N$ 连续, 则 C 唯一确定; 反之, 若 $F_1, \cdots, F_N$ 为一元分布, 那么由式(4)定义的函数 F 是边缘分布 $F_1, \cdots, F_N$ 的联合分布函数。

通过 Copula 函数 C 的密度函数 c 和边缘分布 $F_1, \cdots, F_N$, 可以求出的 N 元分布函数 $F(x_1, \cdots, x_n, \cdots, x_N)$ 的密度函数为

$$f(x_1, \cdots, x_n, \cdots, x_N)=c[F_1(x_1), \cdots, F_n(x_n), \cdots, F_N(x_N)]\prod_{n=1}^{N} f_n(x_n) \tag{5}$$

其中, $c(u_1, \cdots, u_n, \cdots, u_N)=\dfrac{\partial(u_1, \cdots, u_n, \cdots, u_N)}{\partial u_1, \cdots, \partial u_n, \cdots, \partial u_N}$, f_n 是边缘分布 F_n 的密度函数[5]。

定理 1: 对随机变量 $(X_1, \cdots, X_n, \cdots, X_N)$ 作严格的单调增变换, 相应的 Copula 函数不变[6-7], 即

若 $\partial h_n(x)/\partial x>0$, 则

$$C_{X_1, \cdots, X_n, \cdots, X_N}=C_{h_1(X_1), \cdots, h_n(X_n), \cdots, h_N(X_N)}$$

由 Sklar 定理可知, Copula 函数为求取联合分布函数提供了一条便捷的通道。另外 Sklar 定理的重要性还在于它提供了一条在不研究边缘分布的情况下分析多元分布相依结构的途径。

由定理 1 可知, Copula 函数导出的相关性指标是严格单调增变换下的相关性, 比线性相关的范围要宽。

2.2　基于 Copula 函数的相关性分析与度量方法

根据上面的定理介绍几种基于 Copula 函数的一致性和相关性测度[8-10]。

2.2.1　Kendall 相关系数 τ

设 (x_1, y_1)、(x_2, y_2) 是独立同分布的向量, 令

$$\tau=P[(x_1-x_2)(y_1-y_2)>0]-P[(x_1-x_2)(y_1-y_2)<0] \tag{6}$$

于是 τ 就度量了 x 与 y 变化的一致性程度。可以证明

$$\tau=2P[(x_1-x_2)(y_1-y_2)>0]-1 \tag{7}$$

从式(7)看出, τ 在 $[-1, 1]$ 之间, 设 (x_1, y_1) 相应的连接函数是 $C(u, v)$, Schwettzer 和 Wolff (1981)证明了 τ 可由相应的 Copula 函数得到

$$\tau=4\int_0^1\int_0^1 C(u, v)\mathrm{d}C(u, v)-1 \tag{8}$$

从 τ 的构造可以看出, x_1-x_2 反应了 x 的变化, 而 y_1-y_2 则反应了 y 的变化。

若 $(x_1-y_1)\cdot(x_2-y_2)>0$, 说明它们变化是一致的;

若 $(x_1-y_1)\cdot(x_2-y_2)<0$, 说明它们的变化是相反一致的, 因此, τ 就反应了变化一致与否的程度。

综上所述, 可以得到: $\tau=1$, 表示 x 的变化与 y 的变化完全一致, 所以二者正相关; $\tau=-1$, 表示 x 的变化与 y 的反向变化完全一致, 所以二者负相关; $\tau=0$, 表示 x 的变化与 y 的变化一半是一致的, 一半是相反一致的, 所以不能判断二者是否相关。

很明显, 对严格单调增的函数 s 与 t, 则有

$$[s(x_1)-s(x_2)][t(y_1)-t(y_2)]>0$$
$$\Leftrightarrow (x_1-x_2)(y_1-y_2)>0$$

所以, τ 值对严格单调增的变换是不变的, 这就充分说明了 τ 作为 (x, y) 的相关指标所具有的优点。

2.2.2　Spearman 相关系数 ρ

设 (x, y) 有联合分布 $H(x, y)$, 它们相应的边缘分布是 $F(x)$ 和 $G(y)$, 而 $(x_0, y_0)\sim F(x)G(y)$, 即 (x_0, y_0) 独立。设 (x, y) 与 (x_0, y_0) 也独立, 令

$$\rho=3\{P[(x-x_0)(y-y_0)>0]\}-P[(x-x_0)(y-y_0)<0] \tag{9}$$

当 (x, y) 的 Copula 函数 $C(x, y)$ 给定后, Schwettzer 和 Wolff(1981)证明了 ρ 可由相应的 Copula 函数给出

$$\rho=12\int_0^1\int_0^1 uv\,\mathrm{d}C(u, v)-3 \tag{10}$$

2.2.3　Schweizer 和 Wolff 的相关系数 σ

Schweizer 和 Wolff 的相关系数 σ 的关系式为

$$\sigma_c = 12\int_0^1\int_0^1 | C(u, v) - uv | du dv \quad (11)$$

2.2.4 秩相关性系数 γ

严格单调增的变换不会改变 x_i 与 y_i 的秩,因为大小顺序关系是不变的,所以它也只与(x, y)的Copula函数 $C(u, v)$相关,Schwettzer 和 Wolff(1981)给出了它的表达式

$$\gamma = 2\int_0^1\int_0^1 (|u+v-1| - |u-v|)dC(u, v) \quad (12)$$

2.2.5 尾部相关系数

令随机变量 X、Y 服从标准均匀分布[11],它们的联合分布函数为Copula函数 C,假定 C 对称,即对所有的 x 和 y, $x\in[0, 1]$, $y\in[0, 1]$,都有

$$C(x, y) = C(y, x)$$

在限定水平 $u\in[0, 1]$, $P\{X>u, Y>u\}>0$,即 $C(1-u, 1-u)>0$,则 X 在条件$\{X>u, Y>u\}$下的分布为

$$F_u(x) \equiv P\{X\leqslant x | X>u, Y>u\} = 1-\frac{C(1-\max(x, u), 1-u)}{C(1-u, 1-u)} \quad (13)$$

定义关于函数 C 的上尾相关 Copula 函数(upper tail dependence copula, UTDC)为

$$C_u^{up}(x, y) \equiv P\{X\leqslant F_u^{-1}(x), Y\leqslant F_u^{-1}(y) | X> u, Y> u\} \quad (14)$$

显然,当 $u\to1$ 时,UTDC 描述了二元随机样本在上尾部的相关结构。

Juri 证明了若 C 是1个阿基米德Copula函数,其母函数 $\varphi\in\Re(\beta>1)$;当 $u\to1$ 时,UTDC 将收敛于1个与 C 具有相同参数的Copula生存函数 $C^{(\beta)}$,即

$$\lim_{u\to1} C_u^{up}(x, y) = C^{(\beta)}(x, y) \quad (15)$$

其中,$C^{(\beta)}$为Copula函数 $C^{(\beta)}$的生存函数,则

$$C^{(\beta)}(x, y) \equiv G_\beta[G_\beta^{-1}(x), G_\beta^{-1}(y)];$$

$$G_\beta(x, y) \equiv \frac{g_1(x, y)-g_\beta(x, y)}{g_1(1, 1)-g_\beta(1, 1)}\geqslant0。$$

其中,$G_\beta(x)\equiv G_\beta(x, 1)$; $g_\beta(x, y)\equiv(x^\beta+y^\beta)^{1/\beta}$。

此时,Copula函数 C 与 $C^{(\beta)}$的上尾相关系数相等,即

$$\lambda^{up}(C)2-2^{1/\beta} = \lambda^{up}(C^{(\beta)})$$

同理,可以定义 $u\in[0, 1]$, $C(u, v)>0$ 时,关于函数 C 的下尾相关 Copula 函数(LTDC)

$$C_u^{lo}(x, y) \equiv P\{X\leqslant F_u^{-1}(x), Y\leqslant F_u^{-1}(y) | X\leqslant u, Y\leqslant u\} = C[F_u^{-1}(x), F_u^{-1}(y)]/C(u, u) \quad (16)$$

其中,$F_u(x)\equiv P\{X\leqslant x | X\leqslant u, Y\leqslant u\} = 1-C[\min(x, u), u]/C(u, u)$

显然当 $u\to0$ 时,LTDC 描述了二元随机样本在下尾部的相关结构。Juri 同样证明了若 C 是1个阿基米德Copula函数,其母函数 $\varphi\in\Re_\alpha$, $0<\alpha<\infty$,当 $u\to0$ 时,LTDC 将收敛于1个与函数 C 具有相同参数的 Clayton Copula 函数,即 $\lim_{u\to0} C_u^{lo}(x, y) = C^{cl, \alpha}(x, y)$。

同样地,Copula 函数 C 的下尾相关系数 $\lambda^{lo}(C) = 2^{-1/\alpha} = \lambda^{lo}(C^{cl, \alpha})$。

运用 Copula 方法对金融市场进行相关性分析时,应该首先确定变量的边缘分布,然后选择恰当的Copula,最后根据上述的方法进行度量。

3 结 语

国内外实证研究表明,金融市场中出现的很多数据往往不服从正态分布,例如汇率的日收益率服从 t 分布,不同汇率的分布函数具有不同的自由度参数,而且不少数据往往是厚尾分布,它们的方差有时并不存在,因此不能用线性相关系数来反映相关性。另外,线性相关系数无法捕捉变量间非线性的相关关系,只有当联合分布服从椭圆分布时,联合分布才能由变量间的相关系数和边缘分布唯一确定,而椭圆分布只能反映变量间对称的相关模式。也就是说,线性相关系数和与之对应的椭圆分布只能描述变量间线性的相关程度和对称的相关模式。在现实中,大多数金融市场间的相关关系都是时变的[12],特别是在收益或收益的波动为极值的条件下,相关关系的变化很大,而且存在非对称相关的可能,即在大的正收益或大的负收益发生时,金融市场间的相关程度可能并不相同,因此用它分析存在非线性和非对称关系的变量间相关性时,会产生误导。运用基于Copula的相关性分析方法,克服了以上不足之处,使得我们研究金融市场的相关性更加细致和全面。

目前国内对Copula的研究主要集中在对2个金融市场的相关关系上,而且通常都集中在对相关程度的分析上,忽略了对金融市场间的相关结构或模式的研究。事实上,具有相同相关程度的两对随机变量,可能会有不同的相关模式,因此可以进一步研究金融市场的相关模式、多个金融市场之间的相关关系以及金融微观市场之间的详细情况。

参考文献:

[1] 谢明文. 协方差、相关系数与相关性的关系[J]. 数理统计与管理, 2004, 23(3): 33-36.

(下转第59页)

护有效竞争，构造企业技术能力发展的制度基石；给企业制造危机意识，加大企业自主技术学习的努力程度；组织企业家协会和行业协会，促进实用技术的推广，促进企业间的相互学习和交流；加强本地企业间联系，营造企业间相互信任与合作的氛围，避免投机性行为的发生；加强本地企业与外部企业尤其是跨国公司的联系，为本地企业加入全球的产业链、有效地获得外部技术创造条件；建立企业与大学和研究机构的合作关系，推动企业之间、企业（产业）与大学之间建立获取新知识的网络结构；对企业的研究发展与技术改造活动，直接提供财政补贴或税收方面的优惠措施；制定基于企业技术能力发展的人力资本政策，增加对与企业技术能力密切相关的教育投资、职工的在职培训投资。

4 3　企业技术能力发展战略

西部企业在提高技术能力方面，要根据自身条件选择适当的技术，并消化、吸收、改进，使之本地化，争取有所创新。西部企业不能以单纯的引进、模仿为满足，应从长远利益出发，力争走出从模仿到模仿创新再到自主创新之路。有条件的地方可以勇敢地选择跨越式的技术切入点，通过合作研究或战略联盟的方式获得新技术，实现技术能力的快速提升。

参考文献：

[1]　魏后凯，孙承平. 我国西部大开发战略实施效果评价[J]. 开发研究，2004，10(3)：21-25.

[2]　安同奥. 企业技术能力发展论[M]. 北京：人民出版社，2004.

[3]　贾根良. 后发优势的演化创新观[J]. 山西大学学报：哲学社会科学版，2004，27(1)：70-75.

[4]　林毅夫，刘志强. 中国的财政分权与经济增长[J]. 北京大学学报：哲学社会科学版. 2000，37(4)：5-17.

[5]　李晚佳. 扩散与梯度转移的困境：对中国经济差距的制度解读[M]. 北京：经济科学出版社，2005.

[6]　东方社奇，耿宏强，张文红. 西部地区经济发展评价体系与评价指标[M]. 北京：社会科学文献出版社，2005.

[7]　魏　江，许庆瑞. 企业技术能力作用于创新效益的经济控制模型研究[J]. 数量经济技术经济研究，1997，12(9)：42-45.

[8]　钱德勒. 企业规模经济与范围经济[M]. 北京：中国社会科学出版社，1999.

（上接第 54 页）

[2]　Schmitz V. Revealing the Dependence Structure between X(1) and X(n)[J]. Journal of Statistical Planning and Inference, 2004, 123(1): 41-47.

[3]　张尧庭. 连接函数(Copula)技术与金融风险分析[J]. 统计研究，2002，11(4)：48-51.

[4]　Engle R F, Granger C W. Cointegration and Error Correction: Representation, Estimation and Testing [J]. Econometrics, 1987(55): 251-286.

[5]　Sklar A. Functions de Répartition àn Dimensions Etlures Merges[J]. Publ Inst Statist of Paris University, 1959(8): 229-231.

[6]　Nelsen R B. An Introduction to Copulas[M]. New-York: Springer Press, 1999.

[7]　Schwettzer B, Wolff E. On Nonparametric Measures of Dependence for Random Variables[J]. Annals of Statistics, 1981(9): 879-885.

[8]　Joe H. Multivariate Models and Dependence Concepts [M]. London: Chapman & Hall, 1997.

[9]　Cossette H, Gaillardetz P, Marceau E. On Two Dependent Individual Risk Models[J]. Insurance: Mathematics and Economics, 2002(30): 153-166.

[10]　Rossi D G. Colman Filtering of Consistent Forward Rate Curves: A Tool to Estimate and Model Dynamically The Term Structure[J]. Journal of Empirical Finance, 2004(11): 277-308.

[11]　Embrechts P, McNeil A, Straumann D. Correlation and Dependence in Risk Management[J]. Properties and Pitfalls, 2002(29): 114-117.

[12]　Scaillet O. A Nonparametric Analysis of Stock Index Return Dependence Through Bivariatc Copulas[J]. European Investment Review, 2002(1): 7-16.