DOI:10.13546/j.cnki.tjyjc.2010.18.055

基于Copula-GARCH方法的投资组合VaR计算

李育峰[1],严定琪[2],胡海洋[3]

(1.中国人民银行 兰州中心支行,兰州 730000;2.兰州大学 数学与统计学院,兰州 730000;
3.中国人民银行 张掖市中心支行,甘肃 张掖 734000)

摘 要:VaR方法是金融市场风险测量的主流方法。Copula函数广泛的应用于风险管理、投资组合选择、资产定价等金融领域。文章选取五种代表性的Copula并结合带正态分布和学生t分布的GARCH模型描述金融数据,通过Monte Carlo模拟计算投资组合的VaR,并对各种模型的计算能力做了对比,发现Clayton Copula结合GARCH(1,1)-T的模型对VaR的估计最好。

关键词:VaR;Copula;GARCH;Monte Carlo模拟

中图分类号:O29　　文献标识码:A　　文章编号:1002-6487(2010)18-0155-03

0 引言

风险管理的基础和核心是对风险的定量分析和评估。目前,由JP Morgan公司率先提出的VaR方法是金融市场风险测量的主流方法。VaR的含义是"处于风险中的价值",是指市场正常波动下,在一定的概率水平下,某一金融资产或证券组合在持有期末的最大可能损失[1]。

Copula可以解释为"相依函数"或"连接函数",是把多维随机变量的联合分布用其一维边际分布连接起来的函数。我们可以将金融资产风险分解成单个资产的风险和由投资组合产生的风险两部分,其中单个金融资产的风险可以由它们各自的边缘分布来描述,而由投资组合产生的风险则完全由连接它们的Copula 函数来描述[2]。

国内部分学者研究了Copula方法在投资组合风险度量方面的应用。张明恒[3]研究了多资产VaR的Copula计算方法;陈守东等[4]用Copula函数对金融时间序列进行建模,运用Monte Carlo方法计算了投资组合的VaR;吴振翔等[5]应用正态Copula和t-Copula结合GARCH模型计算了组合的VaR。

一般的风险测量模型中,假设金融资产的收益率序列服从正态分布,大量的实证分析证明,金融资产的收益率序列具有尖峰厚尾及异方差性。GARCH模型可以有效地描述金融资产的金融资产的条件异方差性。本文假设组合中单个资产的边缘分布分别服从带正态分布和学生t分布的GARCH模型,而组合中多个资产的联合分布用五种常用Copula函数来描述,通过Monte Carlo模拟计算投资组合的VaR,并对不同模型的VaR计算能力做了比较。

1 Copula-VAR-GARCH模型

1.1 Copula函数

(1)SKlar定理:令$F(x_1,x_2,\cdots,x_n)$为n个随机变量的联合分布,每个随机变量的边缘分布为$F_1(x_1),F_2(x_2),\cdots,F_n(x_n)$,那么,存在一个Copula函数C,满足:$F(x_1,x_2,\cdots,x_n)=C(F_1(x_1),F_2(x_2),\cdots,F_n(x_n))$。

如果$F_1(x_1),F_2(x_2),\cdots,F_n(x_n)$连续,则Copula函数是唯一确定的。

(2)常用Copula函数介绍

①正态Copula函数:

$C_R^{Ga}(u_1,\cdots,u_n)=\Phi_R(\phi^{-1}(u_1),\cdots,\phi^{-1}(u_n))$

其中Φ_R:标准多元正态分布的分布函数,其相关系数矩阵为R

ϕ^{-1}:单维标准正态分布函数的逆函数

②Student-t Copula

$C_{v,R}^{t}(u_1,\cdots,u_n)=t_{v,R}(t_v^{-1}(u_1),\cdots,t_v^{-1}(u_n))$

其中:$t_{v,R}$是标准化的相关系数矩阵为R、自由度为v的多元学生t分布

t_v^{-1}是自由度为v的单维t分布函数的逆函数

③Clayton copula

作者简介:李育峰(1981-),男,甘肃庆阳人,硕士研究生,研究方向:金融工程。
严定琪(1965-),男,湖南岳阳人,副教授,研究方向:金融工程。

$$C_{\alpha}^{Cl}(u_1,\cdots,u_n)=\max[(\sum_{i=1}^{n}u_i^{-\alpha}-d+1)^{-1/\alpha},0]$$

是一个非对称 Copula，对下尾处的变化十分敏感，可用于描述具有下尾相关特性的变量之间的相关关系。

④Gumbel copula

$$C_{\alpha}^{Gu}(u_1,\cdots,u_n)=\exp\{-[\sum_{i=1}^{n}(-\ln u_i)^{\alpha}]^{1/\alpha}\}$$

也是非对称 Copula，它表示更强的上尾相关性。

⑤Frank copula

$$C_{\alpha}^{Fr}(u_1,\cdots,u_n)=-\frac{1}{\alpha}\ln\{1+\frac{\Pi_{i=1}^{n}(e^{-\alpha u_i}-1)}{(e^{-\alpha}-1)^{n-1}}\}$$

是一个对称 Copula,其中的参数 α 控制着变量间的相关程度,$\alpha>0$ 表示两变量正相关，$\alpha<0$ 表示两变量负相关。

1.2 Copula-VAR-GARCH 模型

时变向量$(x_{1,t},\cdots,x_{n,t})$中的每个变量 $x_{i,t}$ 可以用一个由截距项 μ_i、p 阶自回归项和一个误差项 $\sqrt{h_{i,t}}\eta_{i,t}$ 组成的 VAR 模型来描述：

$$x_{i,t}=\mu_i+\sum_{k=1}^{n}\sum_{l=1}^{p}\phi_{i,k,l}x_{k,t-l}+\sqrt{h_{i,t}}\eta_{i,t}$$

假设标准化残差 $\eta_{i,t}$ 是均值为 0、方差为 1 的随机变量，一般建模中假设其服从标准正态分布、标准化的 t 分布或广义误差分布。而 $h_{t,i}$ 可以用 Bollerslev(1994)等提出的 GARCH 模型来描述：

$$h_{i,t}=\omega_i+\sum_{j=1}^{p}\beta_{i,j}\eta_{i,t-j}^{2}+\sum_{k=1}^{q}\alpha_{i,k}h_{i,t-k}$$

大量的实证研究表明，GARCH(1,1)模型就可以比较好的来描述金融时间序列的异方差性，因此本文的实证分析中选取GARCH(1,1)模型。

设上述方程中的 $\eta_{i,t}$ 有联合分布函数 $H(\eta_{1,t},\cdots,\eta_{n,t};\theta)$，$\theta$ 是分布函数的参数向量，由 Sklar 定理有：

$$H(\eta_{1,t},\cdots,\eta_{n,t};\theta)=C(F_{1,t}(\eta_{1,t};\delta_1),\cdots,F_{n,t}(\eta_{n,t};\delta_n);\gamma)$$

其中的 $C(\cdot;\gamma)$ 是 copula 函数，$F_{1,t}(\eta_{1,t};\delta_1),\cdots,F_{n,t}(\eta_{n,t};\delta_n)$ 是边际分布函数，γ 是 copula 函数的参数，$\delta_1,\cdots,\delta_n$ 是边际分布的参数。

2 VaR 的计算与后验测试

2.1 VaR 的 Monte Carlo 计算

(1)模拟每种资产在每个时刻的收益率序列$\{y_{1,t},\cdots,y_{2,t}\}$，路径数为 N：

①从所估计的 Copula 中产生 n 个随机数$(u_{1,t},\cdots,u_{n,t})$；

②用所估计的边际分布函数的逆函数将上述的随机数转化为标准化的对数回报序列：

$$(\eta_{1,t},\cdots,\eta_{n,t})=(F_1^{-1}(u_{1,t};\hat{\delta}_1),\cdots,F_n^{-1}(u_{n,t};\hat{\delta}_n))$$

③应用 VAR-GARCH 模型估计的均值和方差，得到回报向量：

$$(y_{1,t},\cdots,y_{n,t})=(\hat{\mu}_{1,t}+\eta_{1,t}\sqrt{\hat{h}_{1,t}},\cdots,\hat{\mu}_{n,t}+\eta_{n,t}\sqrt{\hat{h}_{n,t}})$$

④重复上述过程 j = 1; 2; : : : ; N 次。

(2)计算组合简单回报：

$$R_t^j=\omega_1\cdot\exp\{y_{1,t}\}+\cdots+\omega_n\cdot\exp\{y_{n,t}\}-1,j=1,2,\cdots,N$$

其中 ω_i 表示第 i 种资产在组合中的权重。

(3)计算 95%，97.55%，99%的置信度下未来 m 日的 VaR：

①产生 t 时间的未来 m 日末的累积回报：

$$\bar{R}_t^j=\sum_{k=t+1}^{t+m}R_k^j,j=1,2,\cdots,N$$

②将 t 时间的 N 个 m 日末的累积回报 $\bar{R}_t^j$,j=1,2,…,N 以升序排列，置信度为 c 的 VaR 是其中第$(1-c)\cdot N$ 个数据的相反数。

2.2 VaR 的后验测试

Kupiec 提出了一个基于模型失败比率的检验。假设 T 天的期间内观测到 N 次失败(即损失超过预测值)，我们想知道这个比率是否显著的不同于预测值 1-c，在 T 天内观测到 N 次失败的概率为 $p^N(1-p)^{T-N}$，其中 p 代表任意一天内损失超过预测值的概率，可以用似然率：

$$LR=-2\ln[c^{T-N}(1-c)^N]+2\ln[(1-\frac{N}{T})^{T-N}(\frac{N}{T})^N]$$

检验零假设 $H_0(p=1-c)$,其中的统计量 LR 服从自由度为 1 的 χ^2 分布。

3 实证分析

3.1 数据来源与基本统计特征

本文选取沪深 300 指数和香港恒生指数构成的投资组合(等权重)作为研究对象，数据的时间跨度是从 2005.04.08 至 2008.09.26，由于香港和大陆有不同的节假日，休市开市不一致，故处理数据时，做了相应的调整：对于没有数据的时点以前一交易日的数据代替，共得到 856 个收盘价格数据。用公式 rt = 100(ln pt ln pt 1)调整为百分对数收益率，共得到 855 个数据。为了解两个市场的日收益率的基本特征，表 1 列出了一些基本统计值。

从表 1 中的数据可以看出，日收益率尖峰厚尾的特征很明显，偏度系数不为零，峰度系数远大于 3，利用 Kolmogorov-smirnov 检验可判定收益率序列不服从正态分布。进行滞后为 15 和 20 阶的 Engle 拉格朗日乘子检验，对应的概率值都

表 1 沪深 300 指数与恒生指数日收益率描述性统计特征

	平均值	标准差	偏度	峰度	K-S 检验	ARCH(15)	ARCH(20)
沪深 300	-0.0941	2.1735	0.4688	6.8543	2.69(0)	31.40(0.008)	35.88(0.0159)
恒生指数	-0.0366	1.5554	-0.0277	9.2331	3.37(0)	28.76(0.017)	33.31(0.0312)

表 2 选取的模型列表

模型	Copula	边际分布	模型	Copula	边际分布	模型	Copula	边际分布
N-N	normal	Normal	N-T	Normal	Student-t	T-N	t-Copula	Normal
T-N	t-Copula	Normal	C-N	Clayton	Normal	C-T	Clayton	Student-t
G-N	Gumbel	Normal	G-T	Gumbel	Student-t	F-N	Frank	Normal
F-T	Frank	Student-t						

表 3　　GARCH 模型的参数估计

模型	数据	μ	ω	α	β	DOF
GARCH-N	恒生指数	0.095	0.0235	0.89	0.0992	
GARCH-N	沪深 300	-0.18	0.0494	0.9177	0.0754	
GARCH-T	恒生指数	-0.12	0.0140	0.9219	0.0746	4.8391
GARCH-T	沪深 300	-0.20	0.0284	0.9467	0.0491	4.2428

表 4　　Copula 的参数估计

	N-N	N-T	T-N	T-T	C-N	C-T	G-N	G-T	F-N	F-T
R/α	0.3063	0.2996	0.3051	0.4142	0.2590	0.6926	1.1774	1.3752	2.1584	2.7263
v			26.2154	4.2137						

表 5　　VaR 失败次数及 Kupiec 检验

置信度	95%	97.5%	99%	置信度	95%	97.5%	99%
理论数	42	21	8	理论数	42	21	8
N-N	61(7.7011)	34(6.7822)	18(8.2099)	N-T	32(2.8722)	15(2.038)	4(2.9513)
T-N	60(6.9396)	32(4.9455)	17(6.7347)	T-T	32(2.8722)	14(2.8095)	4(2.9513)
C-N	59(6.2140)	31(4.1236)	16(5.3797)	C-T	34(1.8324)	16(1.4024)	6(0.8041)
G-N	63(9.3296)	35(7.7934)	20(11.4951)	G-T	32(2.8722)	14(2.8095)	5(1.6752)
F-N	62(8.4979)	36(8.8643)	18(8.2099)	F-T	30(4.1719)	13(3.7263)	4(2.9513)

注：括号中的数字是失败次数对应的库皮克检验的 LR 值，自由度为 1 的卡方分布在 10%,5%,1%下的临界值分别是 2.7055,3.8415 和 6.6349.

小于 0.05，这证实了条件异方差的存在，因此适合用 GARCH 模型进行分析。

3.2　模型选取与参数估计

本文选取正态、学生 t、Clayton、Gumbel 和 Frank 这五种 Copula 来描述变量之间的相关结构，选取带正态分布和学生 t 分布的 GARCH(1,1)模型来描述各变量的边际分布，其中的均值方程简化为不带自回归项的形式，即：$x_{i,t}=\mu_i+h_{i,t}\eta_{i,t}$，共有 10 个模型，见表 2。

应用软件 MATLAB R2008a 对恒生指数和沪深 300 指数估计上述模型的参数，结果列于表 3、表 4。从表 3 可以看出，条件均值方程中的常数不为零，沪深 300 指数的均衡收益为负。条件方差方程中的 α 和 β 显著不为零，这表明价格波动在很大程度上由过去的价格震荡和误差决定，日收益率序列具有很强的波动集聚性。另外，α+β 非常接近于 1，这表明波动具有很强的持续性。

3.3　VaR 的后验测试

由上一部分估计的模型参数，按照 Monte Carlo 模拟计算方法，取模拟路径数 N=5000，计算每一天未来 10 日的 VaR，共得到 846 个值。将每一天未来 10 日末的累积回报与 VaR 作对比，分别得到 MaxVaR 与 VaR 的失败次数，然后计算 Kupiec 检验的统计量 LR 值，结果列于表 5 中。

从表 5 中可以看出，凡是边际分布为 GARCH(1,1)-N 的模型对 VaR 的估计表现都不好，失败次数都远大于理论失败次数，并且库皮克检验都予以拒绝，这说明此类模型严重低估投资组合的风险。而边际分布为 GARCH(1,1)-T 的模型对 VaR 的估计都比较好，实际失败次数与理论失败次数比较接近，库皮克检验也都基本通过。其中 Clayton -Copula 结合 GARCH(1,1)-T 的模型对 VaR 估计最为准确，所有估计结果在 10%的水平下都是显著的。

4　结论

本文采用 Copula-GARCH 方法对恒生指数与沪深 300 指数组成的投资组合进行 Monte Carlo 模拟，以度量投资组合的风险，对由不同 Copula 和不同边际分布组成的模型的风险度量能力进行了对比。从实证结果可以看出，边际分布为 GARCH (1,1)-T 的模型对 VaR 的估计比较好，而边际分布为 GARCH(1,1)-N 的模型都存在低估风险的倾向。其中 Clayton -Copula 结合 GARCH (1,1)-T 的模型对 VaR 的估计最为准确，这主要是因为 Clayton Copula 可以很好的描述下尾相关性，而计算资产组合的 VaR 主要关注的也是资产收益率分布的下尾部。

参考文献：

[1]王春峰.金融市场风险管理[M].天津：天津大学出版社，2001.

[2]韦艳花，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.

[3]张明恒.多金融资产风险价值的 Copula 计量方法研究[J].数量经济技术经济研究，2004,(4).

[4]陈守东，胡铮洋，孔繁利.Copula 函数度量风险价值的 MonteCarlo 模拟[J].吉林大学社会科学学报，2006,(3).

[5]吴振翔，陈敏，叶五一，缪柏其.基于 Copula-GARCH 的投资组合风险分析[J].系统工程理论与实践，2006,(3).

(责任编辑/易永生)