

偏肥尾分布的 APARCH 模型研究

张家平^{1,2}

(1. 华南理工大学 金融工程研究中心, 广州 510006; 2. 华侨大学 数量经济研究院 福建 泉州 362021)

摘要: 金融收益率时间序列分布的不对称性已被视为金融市场的一个“典型事实”,但现有文献对金融时间序列偏度的关注相对较少。由于受到估计方法和计算效率的约束,波动率模型通常假定条件分布为正态分布或对称肥尾分布,如 t 分布和广义误差分布,偏度建模比肥尾建模要复杂得多。通过运用芬南德兹—斯蒂尔方法,在对称分布的基础上构造出有偏肥尾分布。在假定 APARCH 模型的残差项服从这种偏肥尾分布的情况下,研究波动率的拟合效果与预测能力。实证结果显示,与残差项服从对称分布的波动率模型相比,有偏性对波动率模型的估计和预测都有显著影响,向前预测的步数越多,影响越明显。

关键词: 条件分布; APARCH; 肥尾分布; 时间序列

中图分类号: F830.91 **文献标志码:** A **文章编号:** 1009-4971(2010)02-0078-06

一、前言

波动率是金融资产定价和风险管理的关键输入变量,波动率建模是定量金融中一个非常活跃的研究领域。金融资产收益率的经验分布显示出众所周知的特征,即典型事实,包括波动聚集、尖峰肥尾和杠杆效应。波动率建模就是要充分刻画这些特征。由恩格尔提出的标准 ARCH 模型和伯勒斯勒夫进一步扩展的 GARCH 模型,能较好刻画上述部分典型事实而成为刻画波动率最受欢迎的模型。标准的 ARCH 和 GARCH 模型可以构造尖峰肥尾的无条件分布,但不能消除条件分布的尖峰肥尾^[1],即扰动项偏离正态分布,而且不能描述冲击的杠杆效应等吸引了大量学者对该模型进行改进和拓展。

一般沿着两个方向进行来改进 GARCH 模型。一是改进条件均值方程和条件方差方程的设定。斯瓦特提出了 TS-GARCH 模型^[2], 格罗斯特恩等人提出了 GJR-GARCH 模型^[3], 黑金斯和波拉提出了 N-ARCH 模型^[4], 扎克恩提出

了 T-ARCH 模型^[5], 这些模型通过改变条件方差方程的设定,使之能反映杠杆效应。杰维克提出的 Log-ARCH 模型^[6], 使得波动率总是满足为正的条件的。而我们考察的是更为一般的,并包含上述模型设定形式的 APARCH 模型^[7]。二是采用非正态分布刻画扰动项的尖峰肥尾。条件均值方程的扰动性偏离正态分布,一个自然的拓展就是采用肥尾分布来构建 GARCH 模型。文献中运用最广泛的肥尾分布是伯勒斯勒夫对扰动项引入的学生 t 分布和尼尔逊采用的广义误差分布(GED)^[8~9]。①但这些分布都是对称分布,不能刻画金融收益率分布的不对称性。芬南德兹—斯蒂尔提出对于任何连续、单模式和对称的分布 $f(z)$, 可以通过改变分布两边的尺度(scale)来引入有偏性,从而构建出偏正态分布、偏 t 分布和偏 GED 分布^[10]。坎贝尔和汉森构建偏 t 分布来刻画条件偏度^[11~12], 实证研究发现引入条件偏度非常重要并会影响条件方差的可持续性。哈维和斯蒂克研究了条件偏度对资产定价的影响,建议把条件偏度作为一个定价风险因子^[13]。

我国学者对金融资产收益率的条件分布进

①研究者尝试采用了很多其他分布形式,如双指数分布、双 Weibull 分布、双曲分布等,有兴趣者可参考 Bond(2000)的文献综述。

收稿日期:2009-12-26

基金项目:广东省教育厅项目(08JDTDXM79006)

作者简介:张家平(1973-),男,湖北宜昌人,博士研究生,讲师,从事资产定价、风险管理和实证金融研究。

行了比较研究。高见有对残差项分别服从正态分布、t 分布和 GED 分布的 GARCH 模型进行对比分析,发现 t 的拟合较好^[14]。李聪比较 GARCH、EGARCH 和 PARCH 在上述三种分布下的估计,并用于计算 VaR,发现 GED 分布较好^[15]。魏宇等考察了在正态分布、偏正态分布、GED 和偏 GED 下随机波动率模型的估计与拟合检验,发现 GED 和偏 GED 效果较好^[16]。目前还没有文献系统考察条件偏度对波动率预测的影响和实证研究引入条件偏度后对条件波动率预测的改进。

本文主要考察更一般的 APARCH 模型,在扰动项分别服从偏 t 分布 (ST) 和偏广义误差分布 (SGED) 下对我国股市收益的尖峰肥尾和有偏性的刻画^[16],^①并与正态分布、t 分布和 GED 的模型结果进行比较,分析不同分布假定下的 APARCH 对实际市场波动率的刻画精度和适用范围。发现引入条件偏度对波动率的估计与预测均有显著的改进作用。

二、模型设定

1. 偏分布的构造

芬南德兹和斯蒂尔提出一种构建偏分布相当一般化的方法,对于任何连续、单模式和对称的分布 $f(z)$,可以通过改变分布两边的尺度来引入有偏性^[10]。

$$f(z | \beta) = \frac{2}{\beta + 1/\beta} \left[f(\beta z) H(-z) + f\left(\frac{z}{\beta}\right) H(z) \right],$$

$$\text{其中 } H(z) = \frac{1 + \text{sign}(z)}{2} \quad (1)$$

其中 $\rho < \beta < \infty$ 是偏度参数,反映分布偏斜的方向和程度。 $\beta = 1$ 生成对称分布,即 $f(z | \beta = 1) = f(z)$ 。 $\beta < 1$ 时右偏, $\beta > 1$ 时左偏。偏分布的各阶矩决定于原对称分布的各阶矩 M_i 和 β ,计算方法如式(2)。

$$\begin{aligned} \mu_\beta &= M_1(\beta - 1/\beta) \\ \sigma_\beta &= (M_2 - M_1^2)(\beta^2 + 1/\beta^2) + 2M_1^2 - M_2 \quad (2) \end{aligned}$$

$$M_r = 2 \int_0^\infty x^r f(x) dx, \quad r > 2$$

根据式(1)和式(2),我们可以构建 ST、SGED 等偏分布并计算其各阶矩。以 GED 为例,^②标准广义误差的密度函数表示如下:

$$f(z | \nu) = \frac{\nu}{\lambda_\nu^{1+1/\nu} \Gamma(1/\nu)} e^{-\frac{|z|}{\lambda_\nu} \frac{1}{\nu}} \left(\frac{2^{-2/\nu} \Gamma(1/\nu)}{\Gamma(3/\nu)} \right)^{\frac{1}{2}} \quad (3)$$

其中 $\rho < \nu \leq \infty$ (为形状参数,控制 GED 的高度和肥尾。 $\nu = 2$ 时 GED 退化为正态分布 $\nu = 1$ 时为 Laplace 分布 $\nu \rightarrow \infty$ (时是区间为 $[-2, 2]$ 的均匀分布 $\nu < 2$ 时为尖峰分布 $\nu > 2$ 时为宽峰分布 (platykurtotic)。

由式(1)构造标 SGED 的密度函数表示为:

$$f(z | \nu, \beta) = \frac{\nu}{2\theta \cdot \Gamma(1/\nu)} \cdot \exp\left(-\frac{|z - 2\beta \cdot A \cdot S(\beta)| \nu}{[1 - \text{sign}(z - 2\beta \cdot A \cdot S(\beta) - 1)\beta] \theta^\nu}\right) \quad (4)$$

$$\text{其中 } \theta = \Gamma(1/\nu)^{1/2} \Gamma(3/\nu)^{-1/2} S(\beta)^{-1}$$

$$S(\beta) = \sqrt{1 + 3\beta^2 - 4A^2\lambda^2}$$

$$A = \Gamma(2/\nu) \Gamma(1/\nu)^{-1/2} \Gamma(3/\nu)^{-1/2}$$

形状参数 ν 与 GED 的相同,偏度参数 $\beta > 0$,当 $\nu = 2$ 和 $\beta = 1$ 时 SGED 退化为正态分布。

2. APARCH(p, q) 模型设定

在 GARCH 模型的框架下,Ding 和 Granger 等人对条件方差方程的设定如下,

$$\varepsilon_t = z_t \sigma_t \varepsilon_t \square D_\theta(0, 1)$$

$$\sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta \quad (5)$$

其中 $\delta > 0$,参数 γ_i 反映冲击的杠杆效应,如果 γ_i 大于 0 (或小于 0),说明负冲击比正冲击的效应要大 (或小)。 $D_\theta(0, 1)$ 表示扰动项或残差项的标准条件分布, θ 是控制分布偏度和峰度的参数。当 $\omega > 0, \sum_i \alpha_i \kappa_i + \sum_j \beta_j < 1$ 其中 $\kappa_i = E(|z| + \gamma_i z)^\delta$ 时,稳定解存在。当 $\gamma \neq 0$ 和 $\delta \neq 2$ 时, κ_i 值取决于扰动项分布的假设。

APARCH 模型包容性很强,许多其他

^①一些分布如稳定分布和广义双曲分布能同时刻画肥尾和有偏性,但这些分布需要改进算法以提高计算效率,是进一步研究的方向。

^②SN、ST 的构造方法类似,为节省篇幅,本文不列出。由于正态分布不能刻画肥尾,一般不采用 SN。

GARCH 类模型都是 APARCH 模型的特例,对应关系如表 1 所示。

表 1 APARCH 模型与其特例的对应关系

APARCH 模型的参数			对应模型	主要的提出者及提出年份
δ	γ_i	β_i		
2	0	0	ARCH	Engle(1982)
2	0		GARCH	Bollerslev(1986)
1	0		TS - GARCH	Taylor(1986)、Schwert(1989)
2			GJR - GARCH	Glosten, Jaganathan 和 Runkle(1993)
1			T - GARCH	Zakoian(1993)
	0	0	N - ARCH	Higgins 和 Bera(1992)
-0			Log - ARCH	Geweke(1986)、Pentula(1986)

三、实证分析

1. 数据来源与统计分析

本文实证分析的数据是上证综指 1997 年 1 月 2 日至 2009 年 9 月 30 日的日指数价格,共

3 059 个观测值。我们采用指数价格对数的一阶差分作为连续复利收益率。从图 1 可以看出,指数价格从 2006 年开始上升,在 2007 年 10 月下跌,持续整个金融危机期间,收益率序列具有显著的波动时变性和聚集性,在危机期间,波动显著增强;对收益率序列进行统计分析,分析结果如表 2 所示。

表 2 指数收益率统计分析结果

	均值	最大值	最小值	标准差	偏度	峰度	ADF	LM
收益率	0.0004	0.0940	-0.0992	0.0175	-0.2105	7.2110	-14.88*	262.83*

注：“*”表示在置信水平为 1% 下是显著的。

从表 2 可以看出,收益率的平均值为正,但接近于 0。最大值和最小值反映了涨跌停板制度对波动的限制。收益率分布不是对称的,偏度为负说明落在左尾部的观测值要比右尾部的多。

峰度是正态分布的 2 倍多,表示经验分布存在“尖峰肥尾”典型事实。

为考察收益率序列是否为平稳的,我们进行 ADF 检验,估计结果对应的 p 值几乎为 0,拒绝序列

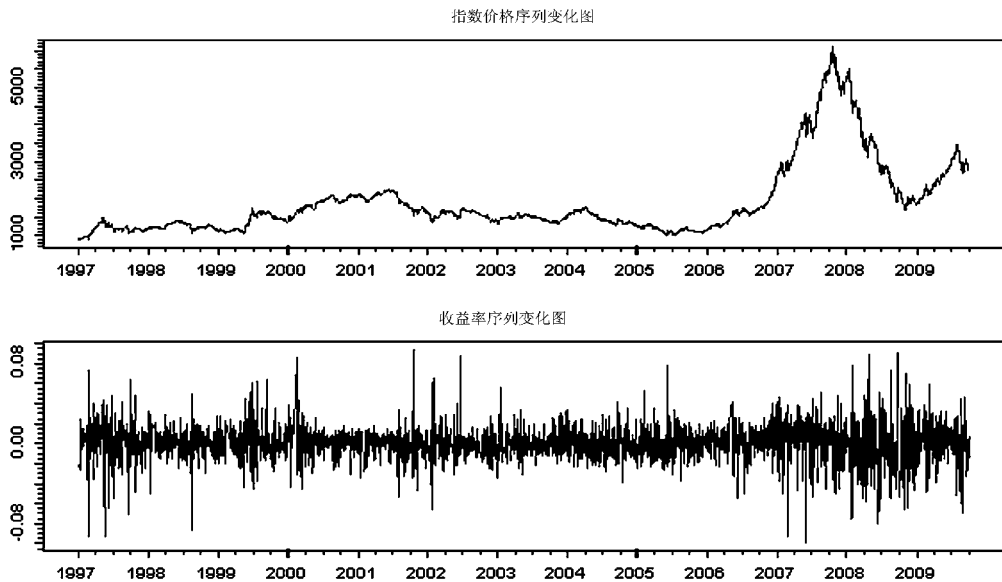


图 1 上证指数价格与收益率时间序列图

存在单位根。LM 检验发现收益率存在显著的 ARCH 效应。

2. 条件分布为 ST、SGED 的 APARCH 模型的参数估计及分析

我们分别用 ST、SGED 来刻画条件均值方程

残差分布的有偏、肥尾等特征。利用最大似然法, 我们可以估计 APARCH - ST、APARCH - SGED 模型的参数, 包括条件分布的参数。为进行模型比较, 我们还估计出常用的基于 N、t、GED 的 APARCH 模型, 结果见表 3。

表 3 不同分布假定下模型的估计结果

模型项目	APARCH - N		APARCH - T		APARCH - GED		APARCH - ST		APARCH - SGED	
	参数	标准差	参数	标准差	参数	标准差	参数	标准差	参数	标准差
μ	0.0018	0.010	0.0178	0.009 +	0.0187	0.009 +	0.0047	0.011	0.0033	0.012
ω	0.0152	0.003*	0.0152	0.004*	0.0146	0.004*	0.0154	0.004*	0.0148	0.004*
α	0.1194	0.011*	0.1226	0.014*	0.1189	0.014*	0.1250	0.014*	0.1228	0.014*
γ	0.1703	0.044*	0.2129	0.056*	0.1829	0.055*	0.2047	0.055*	0.1796	0.054*
β	0.8924	0.011*	0.8889	0.013*	0.8916	0.013*	0.8872	0.013*	0.8889	0.013*
δ	0.8792	0.153*	1.0440	0.191*	0.9988	0.196*	1.0647	0.190*	1.0294	0.192*
ν			5.1197	0.484*	1.2362	0.041*	5.1138	0.486*	1.2295	0.041*
λ							0.9281	0.022*	0.9181	0.018*
似然值	-3135		-3016		-3021		-3011		-3014	
Q(20)	24.12		24.82		24.42		25.14		24.76	
Q ² (20)	20.05		16.72		18.24		16.11		17.11	
AIC	2.0549		1.9772		1.9809		1.9745		1.9767	
BIC	2.0667		1.9909		1.9948		1.9903		1.9924	

注: 1. “*”和“+”表示置信水平分别为 1% 和 5% 时是统计显著的。

2. Q(20) 和 Q²(20) 分别表示标准化残差和标准化残差平方的系列相关 Ljung - Box Q 检验。

从表 3 可以看出, 一是通过对各模型的标准化残差的诊断检验, APARCH(1, 1) 设定充分消除了收益率序列的系列相关和 ARCH 效应, 而且各模型均满足 $\sum_i \alpha_i \kappa_i + \sum_j \beta_j < 1$, 说明条件方差是平稳的。二是除均值项 μ 外, 其他参数都在 1% 置信水平下是统计显著。三是 APARCH - ST 和 APARCH - SGED 的偏度参数(分别为 0.9281、0.9181) 均小于 1, 说明标准化残差分布是左偏的。这与样本偏度 -0.2105 的结果是一致的。四是我们根据对数似然值和信息准则来进行模

型比较, 在不同分布假定下, APARCH - ST 模型的似然值最大, AIC 和 BIC 值最小, 说明该模型在五个模型中拟合效果最好。同时基于 ST 和 SGED 假定的模型拟合效果均优于基于对称分布的模型效果, 说明引入有偏性对模型拟合效果的改进是显著的。

标准化残差的密度函数曲线见图 2, 为便于进行比较, 我们还画出相应的对称分布和正态分布图。从图 2 可以看出, 偏分布的左尾比对照分布的左尾更厚, 能更准确的刻画金融风险。

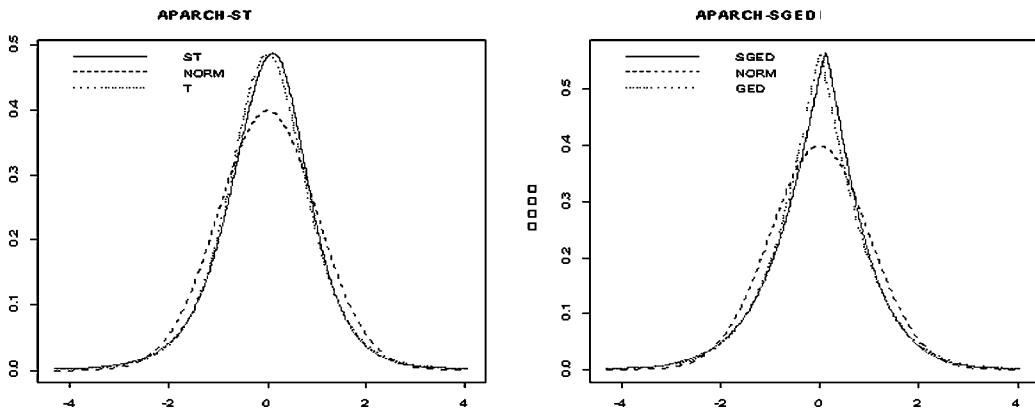


图 2 标准化残差的密度函数、相应的对称分布和正态分布的密度函数图

3. 波动率预测及预测效果比较

预测是研究波动率模型的主要目的之一。根据表 3 ,对称分布中 APARCH - T 的拟合效果最好 ,因此我们选择 APARCH - T、APARCH - ST

和 APARCH - SGED ,分别进行 1、2、5、10、20 步预测 ,并根据均方根误差和平均相对误差来比较各模型的预测效果。计算结果如表 4 所示。

表 4 各模型预测误差估计表

步数	均方根误差			平均相对误差		
	APARCH - T	APARCH - ST	APARCH - SGED	APARCH - T	APARCH - ST	APARCH - SGED
1	2.364494*	2.368118	2.383065	1.67195*	1.674512	1.685081
2	1.968294*	1.970739	1.98037	3.600812	3.580791	3.474898*
5	1.440547*	1.441898	1.444714	8.655478	8.63282	8.35133*
10	1.123832	1.125114	1.122981*	37.73914	37.73093	36.51616*
20	0.904657	0.906272	0.902807*	80.13095	80.5779	77.81576*

注 “*” 表示在该行中的最小值。

从表 4 可以看出 ,根据均方根误差 , APARCH - T 在进行 1、2、5 步预测时误差最小 ,而 APARCH - SGED 在进行 10、20 步预测时是最优的 ;根据平均相对误差 ,除 APARCH - T 在 1 步预测是最优的外 ,其他预测都是 APARCH - SGED 最优。这说明分布偏度对预测的影响模式 ,对于 1 步预测 ,有偏性的影响不显著 ,甚至对称分布的预测效果还要好 ;对于多步预测 ,有偏性的影响是显著的。根据我国股市样本 , APARCH - SGED 设定的预测效果在五个模型中效果最好。

四、结 论

我们考察了在不同分布假定下 ,APARCH 模型的估计与预测 ,特别是在对称分布中引入偏度系数以刻画样本偏度。通过实证研究 ,得出分布的有偏性对模型拟合和预测都有显著影响 ,特别是向前多步预测。这对于刻画金融市场的典型事实、准确进行资产定价、风险度量与管理具有重要意义。

运用芬南德兹—斯蒂尔方法构造偏分布函数优点是计算较为简单 ,容易估计 ,同时满足一定条件下各阶矩都存在。一些参数更为复杂的分布 ,如稳定分布、广义双曲分布能同时刻画分布的尖峰肥尾和偏度 ,因此是残差时间序列的备选分布 ,考察它们的拟合效果和预测能力是我们进一步研究的方向。

参考文献:

- [1] TAVARES A B ,CURTO J D ,TAVARES G N. Modeling Heavy tails and Asymmetry using ARCH - type Models with Stable Paretian Distributions [J]. Nonlinear Dynamics ,2008 ,51: 231 - 243.
- [2] Schwert W. Stock Volatility and the Crash of S87 [J]. Review of Financial Studies ,1990 ,(3) :77 - 102.
- [3] GLOSTEN L ,JAGANNATHAN R ,RUNKLE D. On the Relation between Expected Value and the Volatility of the Nominal Excess Return on stocks [J]. Journal of Finance ,1993 ,48: 1779 - 1801.
- [4] HIGGINS M L ,BERA A K. A Class of Nonlinear ARCH Models [J]. International Economic Review ,1992 ,33: 137 - 158.
- [5] ZAKOIAN J M. Threshold Heteroskedasticity Models [J]. Journal of Economic Dynamics and Control ,1994 ,15: 931 - 955.
- [6] GEWEKE J. Modelling of the Persistence of Conditional Variances: A Comment [J]. Econometric Review ,1986 ,5: 57 - 61.
- [7] DING Z ,GRANGER C W J ,ENGLER R F. A Long Memory Property of Stock Market Returns and a New Model [J]. Journal of Empirical Finance ,1993 ,1: 83 - 106.
- [8] BOLLERSLEV T. A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return [J]. Review of Economic Statistics ,1987 ,69: 542 - 547.
- [9] NELSON D. Conditional Heteroskedasticity in Asset Returns: A New Approach [J]. Econometrica ,1991 ,59 (2) : 347 - 370.

- [10] FERNANDEZ C, STEEL M. On Bayesian Modelling of Fat Tails and Skewness [J]. Journal of American Statistical Association, 1998, 93: 359 – 371.
- [11] HARVEY C R, SIDDIQUE A. Autoregressive Conditional Skewness [J]. The Journal of Financial and Quantitative Analysis, 1999, 34(4): 465 – 487.
- [12] HANSEN B E. Autoregressive Conditional Density Estimation [J]. International Economic Review, 1994, 35(3): 705 – 730.
- [13] HARVEY C R, SIDDIQUE A. Conditional skewness in asset pricing tests [J]. Journal of Finance, 2000, LV(3): 1263 – 1295.
- [14] 高见有. 不同残差分布下 GARCH 族模型的分析对比 [J]. 西部金融, 2009, (6): 66 – 67.
- [15] 李聪. 运用 GARCH 族模型在不同分布下对深证综合的 VaR 分析 [J]. 统计与决策. 2006, 221(9): 79 – 80.
- [16] 魏宇, 高隆昌. 基于有偏胖尾分布的随机波动模型估计及其检验 [J]. 系统管理学报. 2008, 17(3): 266 – 272.

Research on APARCH Models with Skewed & Fat – Tailed Distributions

ZHANG Jia-ping^{1 2}

(1. Research Center of Financial Engineering, South China University of Technology, Guangzhou 510006, China; 2 School of Economic, Huaqiao University, Quanzhou 362021, China)

Abstract: The asymmetry of financial return distribution has been regarded as a "typical fact", while less attention is relatively paid to the skewness of financial time distribution in the current literature. For the limitation of estimation methodology and computation performance, volatility models assume that the distribution of residuals is normal distribution, or symmetrical fat – tailed distribution, such as t or generalized error distribution. Modeling skewness is much complicated than modeling fat tail. The author constructs the skewed and fat – tailed distributions based on symmetric distributions by using Fernandez – Steel method. And the author studies the prediction of volatility together with APARCH model. Compared with volatility models with symmetrical distribution, empirical results show that there are significant influences of the skewness on simulation and prediction of volatility models.

Key words: conditional distribution; APARCH; skewed & fat – tailed distribution; time distribution

[责任编辑 张大勇]