2010 年 5 月
第 29 卷 第 3 期

数理统计与管理
Journal of Applied Statistics and Management

May 2010
Vol. 29 No. 3

文章编号：1002–1566 (2010) 03 – 0544 – 06

GARCH 模型估计方法选择及对上证指数的应用

黄达　王汉生

(北京大学光华管理学院, 北京 100871)

摘要：本文介绍了对 ARCH/GARCH 模型的两种估计方法：准极大似然估计和极小绝对偏差估计，并提出了一种基于自助法 (Bootstrap) 对估计方法的选择。在厚尾程度不同的情况下进行了模拟分析，表明对于一个具体的数据，该选择法能够自动选择较优的估计方法。并用该方法对上海证券交易所 A 股和 B 股的股价指数进行了分析，印证了上海股市 B 股收益率的尾部厚于 A 股收益率尾部。

关键词：广义自回归条件异方差模型；准极大似然估计；极小绝对偏差估计；厚尾；自助法；估计方法选择

中图分类号：F830, O212　　　　**文献标识码**：A

Estimation Selection of GARCH Model and Application on Shanghai Stock Price Index

HUANG Da　WANG Han-sheng

(Guanghua School of Management, Peking University, Beijing 100871, China)

Abstract: In this paper, we introduce two estimations on GARCH model: one is Quasi-Maximum Likelihood Estimation (QMLE), the other is the Least Absolute Deviation Estimator (LADE). We also suggest an estimation selection method based on bootstrap strategy. We present the effect of this selection method by models with tails of different heaviness and show the method's power in selecting a better estimation. We also implement our methods on the real data of Shanghai Stock Price Index, and confirm that the yield's volatility of B share is heavier than its counterpart of A share.

Key words: GARCH, quasi-maximum likelihood estimation, least absolute deviation estimator, heavy tail, bootstrap, estimation selection

0 引言

ARCH 模型和 GARCH 模型 [1−2] 自问世 20 多年来，已经在金融时间序列的波动性分析领域获得极大的成功。其分析的对象横向上已经从最初的单个波动率序列扩展到向量形式的

收稿日期：2008 年 4 月 16 日　　　　**收到修改稿日期**：2008 年 10 月 15 日

基金项目：本研究部分受到国家自然科学基金资助，项目批准编号 10771006。

波动率序列 [3-4] 以及矩阵形式波动率序列 [5]，纵向上从二阶矩 (波动率) 扩展到一般的 γ 阶矩 [6]。

传统上对此类模型的估计，先假定残差项为高斯 (正态) 分布，再利用极大似然估计求得具体的参数值。然而，实证研究表明金融数据普遍存在着厚尾性 [7-8]。王明进和陈奇志 [9] 通过描述分析对亚洲股市也得出了类似的结论。在这种情况下，再沿用基于正态假定的极大似然估计，就会出现问题。

针对正态残差的不适应性，许多人提出将关于残差的假设改为某些厚尾的分布，如 t- 分布 [10]，帕累托分布 [7-8]，一般误差分布 [11] 等形式，再用相应的准极大似然估计 (Quasi-Maximum Likelihood Estimator，QMLE) 得到参数。这样做虽然解决了厚尾的问题，但是在估计的渐近性上存在一些问题。Hall 和 Yao[12] 指出当 GARCH 模型中残差四阶矩存在时，准极大似然估计以 $n^{1/2}$ 的速度收敛到渐近正态。当残差四阶矩不存在时，QMLE 存在着两个问题，第一是渐近分布不再是正态的，具体是什么样的则不仅依赖于残差的分布类型，还依赖于具体的残差分布的参数；第二是估计的收敛速度慢于 $n^{1/2}$。对于 t 分布来说，自由度不容易确定：如果太大，则分布接近标准正态分布，难以表现出厚尾性，如果太小则容易出现绝对值的矩不存在的现象；注意：当 $\epsilon_t \sim t(d)$ 时，有 $E|\epsilon_t|^d = \infty$。帕累托分布的二阶矩不存在，而基于一般误差分布的极大似然估计的渐近性质则非常复杂，目前尚无较清晰的结论。

从某种程度上来说，基于正态分布的极大似然估计可以视为一种最小二乘估计。众所周知，最小二乘估计对于残差中的极值非常敏感。于是，当我们有理由相信残差为厚尾时，为了避免极值的影响，一般采用较为稳健的最小一乘估计来作估计。基于同样的想法，为了克服 QMLE 处理 GARCH 模型残差厚尾时的弊端，Peng 和 Yao[13] 提出了极小绝对变差估计 (Least Absolute Deviation Estimator, LADE) 来估计 GARCH 模型，该估计方法在残差厚尾情况下效果十分稳健。相对于 QMLE, LADE 在残差二阶矩存在条件下，其渐近分布即为正态，这一条件要弱于 QMLE 渐近到正态分布所要求的残差四阶矩存在的条件。此外，LADE 的收敛速度为 $n^{1/2}$，也快于 QMLE。此处理论证明可参见 Peng 和 Yao[13]。

然而，任何估计方法的优劣都是相对的，LADE 与 QMLE 也是一样。当面对具体数据时，我们并不知道残差的真实分布是什么。那么，我们应该如何依据数据来选择一个相对较优的估计方法呢？Huang 等 [14] 用模拟的方法给出了一种基于估计精度的方法在 QMLE 和 LADE 之间做出选择，并指出该方法对残差非对称分布情况依然有效。本文则将通过 boostrap 方法来估计 QMLE 和 LADE 的预测平均绝对误差的方法，来在 QMLE 和 LADE 中作选择。

本文接下来的部分安排如下：在第 1 节，我们介绍 QMLE 和 LADE 并提出了选择方法；第 2 节进行了数值模拟和真实数据分析；第 3 节为结论。

1 估计方法及其选择

广义自回归条件异方差模型 (Generalized AutoRegressive Conditional Heteroscedastic, GARCH) 定义为：

$$X_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 \equiv \sigma_t(\theta)^2 = c + \sum_{i=1}^{p} b_i X_{t-i}^2 + \sum_{j=1}^{q} a_j \sigma_{t-j}^2, \tag{1}$$

其中模型的阶数 $p \geq 1, q \geq 0$，常数项 $c > 0$。b_j 与 a_j 为非负未知参数。$\theta = (c, b_1, \cdots, b_p, a_1, \cdots, a_q)^\top$，$\{\epsilon_t\}$ 为一列均值为 0，方差为 1 的独立同分布随机变量。需要特别注意的是，此处 ϵ_t 的具体分布未知。显然，当 $q = 0$ 时，即为一个自回归条件异方差 (AutoRegressive Conditional Heteroscedastic, ARCH) 模型。本文中，我们仅考虑严平稳条件下的 GARCH 过程。在 $EX_t^4<\infty$ 时，(1) 为严平稳的充要条件是：

$$\sum_{i=1}^{p} b_i+\sum_{j=1}^{q} a_j<1. \tag{2}$$

此时，模型有唯一的严平稳解，且 $EX_t=0$, $\text{var}(X_t)=c/(1-\sum_{i=1}^{p} b_i-\sum_{j=1}^{q} a_j)$; 此处可参阅 Fan 和 Yao[15] 中定理 4.4。

1.1 估计方法

传统上我们假定残差项服从某一分布，并用拟极大似然估计 (QMLE) 来估计参数。QMLE 的具体定义为：

$$\hat{\theta}=\arg\min_{\theta}\sum_{t=\nu+1}^{n}\Big(\frac{X_t^2}{\sigma_t(\theta)^2}+\log(\sigma_t(\theta)^2)\Big). \tag{3}$$

此估计方法在残差为正态分布、t- 分布、广义高斯分布等情况下都适用，具体的讨论请参见 Fan 和 Yao[11] 书中 4.2.3(a) 节。

虽然已经有了设定残差项为某些厚尾分布的方法来修正假定残差为正态时的问题，但是这一方法仍然有其内在的问题。因为在实际数据分析中，模型的残差分布往往是未知的，无论我们设定其为哪种特定分布或分布族，我们便同时排除了其他所有分布/分布族的可能，这需要冒着模型假设错误的危险。而将各个分布族逐一尝试，则过程会非常的繁琐，代价太大。所以针对此类修正的这一缺陷，Peng 和 Yao[13] 提出从估计方法的角度出发来处理厚尾情况。他们指出可以对 GARCH 模型作出稍许修改，即将残差平方的均值 (即残差的二阶矩) 为 1 改为令残差平方的中位数为 1。这一改动相当于将原始的 GARCH 模型残差项乘以一个非负的常量 k。沿用 (1) 中的定义，此时我们的模型变为：

$$X_t=\sigma_t\eta_t,\qquad \sigma_t^2\equiv\sigma_t(\theta)^2=k^2c+\sum_{i=1}^{p}k^2b_iX_{t-i}^2+\sum_{j=1}^{q}a_j\sigma_{t-j}^2, \tag{4}$$

其中 $\eta_t=k\epsilon_t$, $k>0$, $Median(\eta_t^2)=1$。由对数函数的单调性可知 η_t^2 的中位数为 1 等价于 $\log(\eta_t^2)$ 的中位数为 0; 而 $X_t^2=\sigma_t^2\eta_t^2$ 等价于 $\log(\eta_t^2)=\log(X_t^2)-\log(\sigma_t(\theta)^2)$; 所以 η_t^2 的中位数为 1 等价于 $\log(X_t^2)-\log(\sigma_t(\theta)^2)$ 的中位数等于 0。此时便可以用极小绝对偏差估计 (LADE) 来估计模型参数。LADE 的定义是：

$$\tilde{\theta}=\arg\min_{\theta}\sum_{t=\nu+1}^{n}\big|\log(X_t^2)-\log(\sigma_t(\theta)^2)\big|. \tag{5}$$

如果我们假设 $e_t=\log(\epsilon_t^2)$ 服从密度函数为 $0.5\lambda\exp(-\lambda|x|)$ 的拉普拉斯分布时 (λ 为一个正值常数), LADE 则也可以视为一种极大似然估计。

注意 理论上来说，一个 GARCH 模型等价于一个 ARCH(∞) 模型，但是由于样本量总是有限的，所以实际运算中的 $\sigma_t(\theta)^2$ 为一个截断的值，我们记为 $\tilde{\sigma}_t(\theta)^2$。对于一个平稳的 GARCH 时间序列，随着时间间隔的拉大，前后期波动率之间的相关系数会趋于 0，所以对计算的影响可以忽略不计。

1.2 估计方法的选择

和不同残差假定的模型类似；对于具体的数据，估计方法也有适合与不适合之分。那么应该如何在上述两种估计中作出选择呢？

显然，如果一个估计比另一个估计的绝对误差要小，那么我们自然认为前者要优于后者。由于 $\hat{\theta}$ 和 $\tilde{\theta}$ 是基于不同的模型参数假设得到的解，为了保证它们之间绝对误差的可比性，我们定义 $\beta=(\beta_1,\cdots,\beta_{p+q})^\top=(b_1/c,\cdots,b_p/c,a_1,\cdots,a_q)^\top$，由上一节的讨论不难看出，此时的 β 在两种模型假设下是一样的。令 $\hat{\beta}=(\hat{\beta}_1,\cdots,\hat{\beta}_{p+q})^\top=(\hat{b}_1/\hat{c},\cdots,\hat{b}_p/\hat{c},\hat{a}_1,\cdots,\hat{a}_q)^\top$；用类似的形式，我们可以定义 $\tilde{\beta}$。Huang 等 [14] 给出了一个通过比较 $\sum_{i=1}^{p+q}|\hat{\beta}_i-\beta_i|$ 和 $\sum_{i=1}^{p+q}|\tilde{\beta}_i-\beta_i|$ 的相对大小的选择方法。但是，在处理实际数据时真实的参数值 β_i 是未知的，因此该方法处理真实的金融数据时有其局限性。我们建议使用自助法对来估计平均绝对误差 (Mean Absolute Deviation, MAE)，以选择合适的估计方法。

当我们用 QMLE 求得参数 $\hat{\theta}$ 后，便可以得到残差项 $\breve{\epsilon}_t=X_t/\tilde{\sigma}_t(\hat{\theta})$，$\tilde{\sigma}_t(\theta)$ 的定义即上一节最后所讨论的被截断的 $\sigma_t(\theta)^2$ 的平方根。再将 $\breve{\epsilon}_t$ 标准化，使得其均值为 0，方差为 1，记标准化后的残差为 $\hat{\epsilon}$。对 $-\infty<t<\infty$，定义 $X_t^*=\sigma_t^*\epsilon_t^*$，其中

$$(\sigma_t^*)^2=\hat{c}+\sum_{i=1}^{p}\hat{b}_i(X_{t-i}^*)^2+\sum_{j=1}^{q}\hat{a}_j(\sigma_{t-j}^*)^2,\qquad -\infty<t<\infty,$$

ϵ_t^* 系从标准化后的残差序列 $\{\hat{\epsilon}_t\}$ 中均匀地随机独立抽取。令 $X_1^*,\cdots,X_n^*$ 为基于自助法抽得的样本，$\hat{\beta}^*$ 为该样本的 QMLE。于是我们可以通过由自助法得到的估计 $\hat{\beta}^*$。重复 B 次上述的自助抽样和估计 (B 是一个大的正整数)。我们就可以估计 $\hat{\beta}$ 的平均绝对误差：

$$MAE(\hat{\beta}^*)=E\Big(\sum_{i=1}^{p+q}|\hat{\beta}_i^*-\hat{\beta}_i|\Big|X_1,\cdots,X_n\Big),$$

计算上，这相当于将每次得到的 $\sum_{i=1}^{p+q}|\hat{\beta}_i^*-\hat{\beta}_i|$ 求平均得到。

相应地，LADE $\tilde{\beta}$ 的平均绝对误差

$$MAE(\tilde{\beta}^*)=E\Big(\sum_{i=1}^{p+q}|\tilde{\beta}_i^*-\tilde{\beta}_i|\Big|X_1,\cdots,X_n\Big).$$

也可以用自助法类似地得到。唯一的区别在于当我们“标准化”残差时，令其平方的中位数为 1，而不再是均值为 1。上式中 $\tilde{\beta}^*$ 是重抽样所得样本 $X_1^*,\cdots,X_n^*$ 的 LADE。

选择方法 如果 $MAE(\hat{\beta}^*)\le MAE(\tilde{\beta}^*)$，我们选用 QMLE；否则选用 LADE。

2 数值模拟与实证分析

在本小节，我们将分别通过数值模拟和实证分析的方法来演示上一节给出的选择方法。

2.1 数值模拟

模拟数据的模型分别设定为 ARCH(2) 和 GARCH(1,1)。对于任一模型，我们分别令残差 ϵ_t 服从标准正态分布，自由度为 3 或 4 的 t 分布，以及为使得平方对数为拉普拉斯分布的 ϵ_t。我们将这些分布标准化，以使得它们的前两阶矩分别为 0 和 1。在 ARCH(2) 模型中，我们令 $c=1$，$(b_1,b_2)=(0.7,0.2)$；在 GARCH(1,1) 模型中，我们令 $(b_1,a_1)=(0.2,0.7)$。样本大小为

$n=250$, 500 和 1000。对每种设定，我们分别模拟 100 次。而自助法重抽样的次数 B，我们设为 200。

表 1 列出了通过自助法估计平均绝对误差法判断 LADE 优于 QMLE 的比例。从横向上来看，当厚尾程度比较轻时，LADE 表现逊于 QMLE，而随着残差厚尾程度逐渐增强，LADE 被选中的比例越来越大；这说明我们的选择方法能够很好的在厚尾与非厚尾之间作出正确的选择。从纵向上来看，除了残差为标准正态分布时 LADE 被选中比例随着样本量的增加而减小，对于厚尾的其他三种情况，LADE 较优的比例随着样本量的增加而变大。这说明随着样本量的增加，我们的方法会选中正确估计方法的概率越来越高。

表 1 LADE 优于 QMLE 的比例 (通过自助法估计平均绝对误差法判断)

模型	序列长度	$N(0,1)$	$t(4)$	$t(3)$	Laplace
GARCH(1,1)	250	0.040	0.350	0.540	0.690
	500	0.030	0.360	0.650	0.880
	1000	0.000	0.470	0.800	1.000
ARCH(2)	250	0.140	0.390	0.500	0.800
	500	0.140	0.460	0.610	0.920
	1000	0.050	0.560	0.760	1.000

2.2 实证分析

以下我们便将对上海证券交易所 1997 年 1 月 2 日到 2000 年 12 月 15 日 A 股股价指数 (SHA) 和 B 股价格指数 (SHB) 用我们的选择法选取估计方法。两组数据的所含的样本均为 955 个观测。

表 2 对于上证 A、B 股不同模型的估计参数以及平均绝对误差

股指种类	估计方法	常数项	GARCH 项系数	ARCH 项系数	MAE
上证 A 股	QMLE	0.25	0.68	0.27	5.13
	LADE	0.04	0.83	0.04	5.16
上证 B 股	QMLE	0.60	0.73	0.20	6.23
	LADE	0.20	0.70	0.07	6.21

从表 2 中可以看出，由于两种估计方法的模型假定不同，所以得到的参数也不一样。根据选择较小 MAE 的原则，对于上证 A 股的应当选用 QMLE，而上证 B 股应当选用 LADE。这也印证了王明进和陈奇志 [9] 提出的的 IC-GARCH 模型在处理多元波动率情况时得到的上证 B 股的波动率要高于上证 A 股的结论。我们需要注意的是，在 2001 年之前，B 股不对境内投资者发售，B 股市场上主要的交易方为外资，所以其价格动态特征和 A 股不一样。即使同属上海证券交易所，不同股种之间的厚尾性等特征也会不一样。秦宛顺和王永宏 [16] 对造成 A 股和 B 股价格之间的差异的原因有系统的论述。

3 结论

本文中我们介绍了两种 ARCH/GARCH 模型的估计方法，并在扰动项未知的前提下给出了一个针对具体的数据在两种估计中做出选择的方法。相对于以前文献中将模型扰动项设定为服从某一具体分布的做法，避免了模型设定错误的可能性。我们同时通过模拟和实证分析的方法验证了我们选择方法的有效性。金融数据中厚尾性是一个普遍的现象，尤其是在不健

全的亚洲市场，由于政策、信息流动性、投资理念等多种原因，厚尾性更强。对于厚尾数据我们需要一个更优的参数估计方法，但是面对真实数据，难以判断具体的厚尾程度，因此需要有一个方法自动选择一个最优的方法。

本文中的方法不仅可以应用在 ACRH/GARCH 模型处理上，也可以应用在在一般的回归等场合。

[参考文献]

[1] Engle R F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation [J]. Econometrica, 1982, 50: 987–1007.

[2] Bollerslev T. Generalized autoregressive conditional heteroscedasticity [J]. Journal of Econometrics, 1986, 31: 307–327.

[3] Bauwens L, Laurent S and Rombouts J V K. Multivariate GARCH models: A survey [J]. Journal of Applied Econometrics, 2008, to appear.

[4] Wang M, Yao Q. Modelling Multivariate Volatilities: An ad hoc Approach [A]. In Contemporary Multivariate Analysis and Experimental Designs – In Celebration of Professor Kai-Tai Fang's 65th Brithday [C]. Singapore: World Scientific, 2005, 87–97.

[5] Fan J, Wang M & Yao Q. Modelling multivariate volatilities via conditionally uncorrelated components [J]. Journal of Royal Statistical Society, Series B, 2008, 70: 679–702.

[6] Penzer J, Wang M & Yao Q. Approximating volatilities by asymmetric power GARCH functions [J]. Austrialia New Zealand Journal of Statistics, 2008, to appear.

[7] Mittnik S & Rachev S T. Stable Paretian models in finance [M]. New York: Wiley, 2000.

[8] Mittnik S, Rachev S T & Paolella M S. Stable Paretian modeling in finance: Some empirical and theoretical aspects [A]. A Practical Guide to Heavy Tails [C]. Birkhäuser, Boston, 1998: 79–110.

[9] 王明进, 陈奇志. 基于独立成分分解的多元波动率模型 [J]. 管理科学学报, 2006, (10): 56–64.

[10] Bollerslev T. A Conditionally heteroskedastic time series model for speculative prices and rates of return [J]. The Review of Economics and Statistics, 1987, 69: 542–547.

[11] Nelson D. Conditional heteroscedasticity in asset returns: A new approach [J]. Econometrica, 1991, 59: 347–370.

[12] Hall P & Yao Q. Inference in ARCH and GARCH models with heavy-tailed errors [J]. Econometrica, 2003, 71: 285–317.

[13] Peng L & Yao Q. Least absolute deviations estimation for ARCH and GARCH models [J]. Biometrika, 2003, 90: 967–975.

[14] Huang D, Wang H, & Yao Q. Estimating GARCH models: When to use what? [J]. Econometrics Journal, 2008, 11: 1–12.

[15] Fan J and Yao Q. Nonlinear time series: Nonparametric and parametric methods [M]. New York: Springer, 2003.

[16] 秦宛顺, 王永宏. 中国 A 股与 B 股价格差异的实证分析 [J]. 数量经济技术经济研究, 2005, (5): 15–19.