

一种带有自权重的积分波动率的非参估计*

李翠霞, 郭二林, 包美娟

(兰州大学数学与统计学院, 甘肃 兰州 730000)

摘要: 为了对带有自权重的积分波动率 $\int_0^1 f(X_t) \sigma_t^2 dt$ 进行估计, 定义了一个新的非参数估计量, 并利用“已实现”波动率的方法, 证明了该估计量是积分 $\int_0^1 f(X_t) \sigma_t^2 dt$ 的一致估计量, 同时还得到此估计量的渐近正态分布以及学生化形式, 从而可对该积分做区间估计或假设检验。

关键词: 自权重; 积分波动率; 非参估计; 渐近正态分布

中图分类号: O211.4 O211.63 **文献标志码:** A **文章编号:** 0529-6579(2013)01-0055-04

A Nonparametric Estimate of Integrated Cross-Volatility

LI Cuixia, GUO Erlin, BAO Meijuan

(School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, China)

Abstract: A nonparametric estimator is proposed for the class of integrated cross volatilities of the form $\int_0^1 f(X_t) \sigma_t^2 dt$, where f is a continuous function, σ_t^2 is the instantaneous cross volatility of continuous semimartingale X . Using “Realized Volatility”, the asymptotic properties, which include consistency and asymptotic normality are obtained. A studentized version has been given and this can be used to construct confidence interval and do hypothesis testing.

Key words: cross; integrated volatility; nonparametric estimate; asymptotic normality

波动率也称为易变性, 是对金融市场风险程度的估计。从统计角度看, 它是以复利计的标的资产投资回报率的标准差。从经济意义上解释, 产生波动率的主要原因来自以下三个方面: 宏观经济因素对某个产业部门的影响, 即所谓的系统风险; 特定的事件对某个企业的冲击, 即所谓的非系统风险; 投资者心理状态或预期的变化对股票价格所产生的作用。

目前, 金融高频数据的模型与分析在理论与实际中都引起了广泛的关注, 尤其对股票、期货及其衍生物的对数价格的积分波动率, 即 $\int_0^2 \sigma_t^2 dt$ 的分析在组合管理、风险分析等方面都引起了广泛的关注。波动性是金融市场最为重要的特性之一。金融

资本市场是一个受多种因素影响的复杂系统, 在一些时间段内显得非常平静, 而在另外一些时间段内剧烈波动。描述波动性的时变特性是非常重要的。因为, 第一, 波动性是证券组合理论、资本资产定价模型(CAPM)、套利定价模型(APT)及期权定价公式的核心变量。第二, 它与市场的不确定性和风险直接相关, 是体现金融市场质量和效率的最简洁和最有效的指标之一。此外, 波动性对企业的投资与财务杠杆决策、消费者的消费行为和模式、经济周期及相关宏观经济变量等都具有重要影响。因此, 提高波动性的估计和预测精度一直是金融经济学研究的重要课题。近年来, 随着计算机与通讯

* 收稿日期: 2012-06-27

基金项目: 兰州大学中央高校基本科研业务费专项资金资助项目(lzujbky-2012-179; lzujbky-2012-180)

作者简介: 李翠霞(1981年生), 女, 讲师; 通讯作者: 郭二林; E-mail: guoerlin@lzu.edu.cn

技术的飞速发展,极大地降低了数据记录与存储的成本,从而使得对高频率的金融数据进行研究成为可能。所谓高频数据,是指以小时、分钟或者秒为采集频率的数据。

一般而言,金融市场中的信息是连续的影响证券市场价格的运动过程的,数据的采集频率越低,信息丢失越多;反之,采集频率越高,信息丢失则越少。因此,金融高频数据一出现,就成为了金融领域的研究热点。20世纪90年代后期以来,国际上兴起了对金融高频时间序列的研究热潮。高频时间序列研究的代表人物是 Andersen 和 2003 年诺贝尔经济学奖得主 Engle 的学生 Bollerslev,他们对高频时间序列的波动率采用“已实现”波动率 (Realized Volatility, RV) 的全新概念来度量^[1-5]。

在这篇文章里,我们在离散观察情形下对一个伊藤扩散过程的自权重的波动率进行了估计。假定股票、期权等金融衍生物的价格为 S_t , 则 $X_t = \ln S_t$ 在概率空间 $(\Omega, \mathcal{F}(F_t)_{t \geq 0}, P)$ 下具有形式:

$$dX_t = b_t dt + \sigma_t dW_t, t \geq 0 \quad (1)$$

这里 W 是一个标准布朗运动, b 是一个随机可适过程。以此形式代表股票的对数价格在金融工程中已经得到了广泛的认可。

在本文中我们关心的中心问题就是如何来估计

$$I = \int_0^2 f(X_t) \sigma_t^2 dt \quad (2)$$

其中 $f: \mathbb{R} \rightarrow \mathbb{R}$ 是一个连续函数, σ_t^2 是过程 X 的瞬时波动率, 它被定义为

$$\int_0^1 \sigma_t^2 dt = \int_0^1 d \langle X, X \rangle_t = \int_0^1 \sigma_t^2 d \langle W, W \rangle_t$$

由于它在金融工程中的重要性,大量的学者已经对积分波动率的估计进行了研究^[6-7],但是我们必须强调的一点是,上面提到的工作 [1-7] 都是基于我们这里讨论的特殊情形,即: $f = 1$ 时进行的,而针对一般的函数,对的估计就显得十分重要。一方面,从数学角度上看,有了函数后,对的估计就变得非常复杂。另一方面,在实际工作中,我们往往需要的是一个函数 $f(X)$ 的积分波动率的估计,而不仅仅是的积分波动率的估计。如前面提到的股票价格的对数过程 X_t 可以用伊藤扩散过程来刻画,而实际中我们观察到的仅仅为 $S_t = f(X_t) = e^{X_t}$ 。此时,积分波动率的估计就变成了 $I = \langle X, X \rangle_t = \langle \log S, \log S \rangle_t$, 因此上面文献中提到的关于积分波动率的估计都是不可用的,这时,我们必须寻求新的途径来解决。幸运的是,我们可以利用上面提到的方法,“已实现”波动

率,来构造带有一般函数时的自权重波动率的估计量。

本文中,我们对 I 给出了一个新的非参数估计量 \hat{I}_n , 此估计量被证明是依概率收敛到 I 的,同时对于此估计量,我们还得到了它的渐近正态分布。而针对该渐近正态分布,我们给出了一个学生化的表达式,从而对此估计量可以构造置信区域并做相应的假设检验。最后的模拟数据以及直方图与 QQ 图也说明了我们提出的估计量的收敛是一致有效的。

1 模型假定

假设 1 假设随机过程 b 定义在概率空间 $(\Omega, \mathcal{F}, (F_t)_{t \geq 0}, P)$ 且:

① 过程 b, σ 关于 (F_t) 可适;同时 b 局部有界且 σ 左极右连;

② 对任意的 $t > s > 0$, $W_t - W_s$ 独立于 \mathcal{F}_s 。

假设 X 的观察时刻为 $0 = t_0 < t_1 < \dots < t_n = t$, 不失一般性,我们取 $t = 1$ 。

假设 2 假设过程 X 的观察时间间隔相等, 即: $t_i = \frac{i}{n}, i = 1, \dots, n$ 。因此 $t_i - t_{i-1} = \frac{1}{n}, i = 1, \dots, n$ 。

假设 3 假设函数 f 是 α -李普希兹函数,即存在常数 $C, \kappa > 0$ 以及 $\alpha \in (0; 1]$ 使得,对任意的 x, u 有: $|f(x) - f(u)| \leq C|x - u|^\alpha(1 + |x|^{2\kappa} + |u|^{2\kappa})$ 。

2 主要结果

在本节中,我们考虑在结构 (1) 下对 (2) 中的函数进行估计。我们定义:

$$\hat{I}_n = \sum_{i=1}^n f(X_{\frac{i-1}{n}}) (\Delta_i X)^2 = \sum_{i=1}^n f(X_{\frac{i-1}{n}}) (X_{\frac{i}{n}} - X_{\frac{i-1}{n}})^2 \quad (3)$$

则我们有下面的定理成立:

定理 1 (大数定律) 在假设 1-3 成立的情形下,我们有下面的结论:

$$\hat{I}_n \xrightarrow{P} I$$

证明

$$\begin{aligned} E|\hat{I}_n - I| &= E\left| \sum_{i=1}^n (f(X_{\frac{i-1}{n}}) (\Delta_i X)^2 - \int_{\frac{i-1}{n}}^{\frac{i}{n}} f(X_s) \sigma_s^2 ds) \right| \leq \sum_{i=1}^n E\left| f(X_{\frac{i-1}{n}}) \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} b_s ds \right)^2 \right| + \\ &\quad \sum_{i=1}^n E\left| \int_{\frac{i-1}{n}}^{\frac{i}{n}} (f(X_{\frac{i-1}{n}}) - f(X_s)) \sigma_s^2 ds \right| \end{aligned}$$

$$\leq C \frac{1}{n} + C \frac{1}{n^{\alpha/2}} \rightarrow 0,$$

其中由假设 3，我们有：

$$\begin{aligned} |f(X_{\frac{i-1}{n}}) - f(X_s)| &\leq C |X_{\frac{i-1}{n}} - X_s|^\alpha = \\ &O\left(\left|s - \frac{i-1}{n}\right|^{\alpha/2}\right) \end{aligned}$$

从而，最后一个不等式成立。

现在，我们来证明中心极限定理。为此，我们先给波动过程一个结构上的假定。

假设 4 假设波动函数 $\sigma = \{\sigma_t, t \leq 0\}$ 满足方程：

$$\sigma_t = \sigma_0 + \int_0^t \bar{b}_s ds + \int_0^t \tilde{\sigma}_s d\tilde{W}_s$$

这里， $\bar{b}, \tilde{\sigma}$ 都是可适的左极右连过程，其中 \bar{b} 可料且局部有界， \tilde{W} 是独立于 W 的一个标准布朗运动。

假设 4 是证明中心极限定理时所需的一个标准假设，在中心极限定理中，我们需要一个关于依分布平稳收敛的概念，现在，让我们来简单地叙述一下此概念。

定义 1 假设 $X, X_1, X_2, \dots, X_n, \dots$ 是定义在概率空间 $(\Omega, \mathcal{F}, (F_t)_{t \geq 0}, P)$ 上的一个随机变量序列，当且仅当对任给的集合 $A \in \mathcal{F}$ 和实数，

$$\lim_{n \rightarrow \infty} P(X_n \leq x, A) = P(X \leq x, A)$$

成立时，我们称 X_n 依分布平稳收敛于 X ，并记作 $X_n \xrightarrow{\delta} X$ 。

定理 2 (中心极限定理) 假设 1-4 成立，则我们有

$$\sqrt{n}(\hat{I}_n - 1) \xrightarrow{\delta} \sqrt{2} \int_0^1 f(X_s) \sigma_s^2 dB_s \quad (4)$$

这里 B 是定义在原概率空间的一个扩展上且独立于 \mathcal{F} 的一个标准布朗运动。

证明 记 $\hat{I}_i = f(X_{\frac{i-1}{n}})(\Delta_i X)^2, \mathcal{F}_{\frac{i-1}{n}} = \sigma(X_{\frac{i-1}{n}})$ 。

① 由估计量的定义，知

$$\hat{I}_n = \sum_{i=1}^n (\hat{I}_i - E[\hat{I}_i | \mathcal{F}_{\frac{i-1}{n}}]) + \sum_{i=1}^n E[\hat{I}_i | \mathcal{F}_{\frac{i-1}{n}}]$$

再由假设 4 及伊藤乘积公式，知，对任意的 $s > t$ ，

$$|E(\sigma_s - \sigma_t | \mathcal{F}_t)| \leq C(s - t)$$

现在 $\sum_{i=1}^n E[\hat{I}_i | \mathcal{F}_{\frac{i-1}{n}}] = \sum_{i=1}^n f(X_{\frac{i-1}{n}}) \sigma_{\frac{i-1}{n}}^2 \frac{1}{n} + R_n$ ，

这里 $|R_n| \leq C \frac{1}{n}$ 。因此， $|\sum_{i=1}^n (\hat{I}_i - E[\hat{I}_i | \mathcal{F}_{\frac{i-1}{n}}]) - \int_0^1 f(X_s) \sigma_s^2 ds| \leq C \frac{1}{n}$ 而剩余的部分，记作

$\sum_{i=1}^n M_i$ ，则 $E[M_i | \mathcal{F}_{\frac{i-1}{n}}] = 0$ 。故 M_i 即是我们找到的鞅差。

② 在这一部分，我们来证明

$$\text{Var}(\sqrt{n}(\hat{I}_{n-1})) \xrightarrow{P} 2 \int_0^1 f^2(X_s) \sigma_s^4 ds$$

由条件期望的定义以及独立增量性，我们有

$$n \sum_{i=1}^n E[f(X_{\frac{i-1}{n}})(\Delta_i X)^2 - \int_{\frac{i-1}{n}}^{\frac{i}{n}} f(X_s) \sigma_s^2 ds]^2 | \mathcal{F}_{\frac{i-1}{n}}] =$$

$$n \sum_{i=1}^n E[f^2(X_{\frac{i-1}{n}})(\Delta_i X)^4 + \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} f(X_s) \sigma_s^2 ds\right)^2 -$$

$$2f(X_{\frac{i-1}{n}})(\Delta_i X)^2 \int_{\frac{i-1}{n}}^{\frac{i}{n}} f(X_s) \sigma_s^2 ds | \mathcal{F}_{\frac{i-1}{n}}] =$$

$$n \sum_{i=1}^n E[3f^2(X_{\frac{i-1}{n}}) \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \sigma_s^2 ds\right)^2 + \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} f(X_s) \sigma_s^2 ds\right)^2 -$$

$$- 2f(X_{\frac{i-1}{n}}) \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \sigma_s^2 ds\right) - \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} f(X_s) \sigma_s^2 ds\right) | \mathcal{F}_{\frac{i-1}{n}}] =$$

$$2 \sum_{i=1}^n f^2(X_{\frac{i-1}{n}}) \sigma_{\frac{i-1}{n}}^4 \frac{1}{n} + n \sum_{i=1}^n S_n, \quad (5)$$

利用多维伊藤乘积公式，上面的第三个等式成立。

再利用的定义，我们还有

$$|n \sum_{i=1}^n S_n| \leq n \sum_{i=1}^n |S_n| \leq$$

$$n \sum_{i=1}^n \left[\frac{c}{n^3} + \frac{c}{n^{\alpha/2+3}} + \frac{c}{n^{\alpha/2+2}}\right] = O\left(\frac{c}{n^{\alpha/2}}\right) \quad (6)$$

成立。

联合 (5) 式和 (6) 式，由黎曼积分，即得

$$\text{Var}(\sqrt{n}(\hat{I}_{n-1})) \xrightarrow{P} 2 \int_0^1 f^2(X_s) \sigma_s^4 ds \quad (7)$$

现在联合①和②以及文献 [8] 中的证明，再由鞅中心极限定理，即得此定理成立。现在，我们可以构造方差的一个一致估计量 \hat{C}_n ，即有

$$\hat{C} = \frac{2n}{3} \sum_{i=1}^n f^2(X_{\frac{i-1}{n}})(\Delta_i X)^4 \xrightarrow{P}$$

$$2 \int_0^1 f^2(X_s) \sigma_s^4 ds$$

由此，我们可以得到中心极限定理的一个学生化形式，利用此形式即可以找出相应的置信区间，从而对该估计量做对应的假设检验。

推论 1 假定定理 2 成立，则有

$$\frac{\sqrt{n}(\hat{I}_n - 1)}{\hat{C}_n} \xrightarrow{\delta} N(0, 1) \quad (8)$$

这里 $N(0; 1)$ 是独立于 \mathcal{F} 的一个标准正态分布。

3 模拟结果

我们按一个交易日 6.5 个小时计算，每 20, 5, 3, 1 秒取一次数据，即分别取 $n = 1\ 170, 4\ 680, 7\ 800, 23\ 400$ 时来模拟。在这里，我们把股票的对数价格取作著名的 Ornstein-Uhlenbeck 过程，即

$$X_t = \int_0^t \cos u du + \int_0^t e^{-2(t-u)} dW_u$$

这里, W_u 是一个标准布朗运动。

我们重复运行 1 000 次, 结果包含了相对偏 (relative bias), 标准误差 (standard errors) 与均方误差 (mean square errors)。同时给出的一些 QQ 图也证明了我们提到的定理 2, 即中心极限定理是成立的。

从下表中, 我们可以得出的结论是:

随着 n 的增大, 相对偏 (r. b.), 标准误差 (s. e.) 与均方误差 (m. s. e.) 都逐渐减小, 这和我们的理论结果是一致的。

表 1 n 取不同值时的模拟结果

Table 1 Simulation result for n

| n | (r. b., s. e., m. s. e.) |
|--------|------------------------------|
| 1 170 | (0.001 8 0.041 7, 0.001 7) |
| 4 680 | (0.000 7, 0.020 3, 0.000 4) |
| 7 800 | (-0.000 3, 0.015 8, 0.000 2) |
| 23 400 | (-0.000 1, 0.009 4, 0.000 1) |

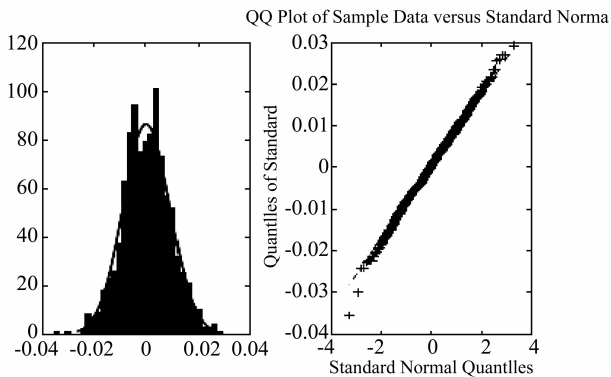


图 1 $n = 23 400$ 时的直方图与 QQ 图

Fig. 1 Histogram and QQ plot for $n = 23 400$

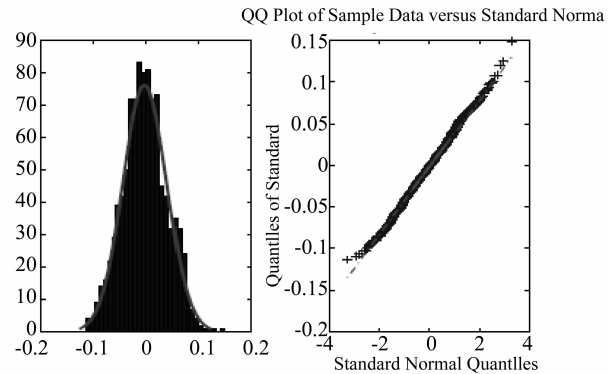


图 2 $n = 1 170$ 时的直方图与 QQ 图

Fig. 2 Histogram and QQ plot for $n = 1 170$

参考文献:

- [1] ANDERSEN T G, BOLLERSLEV T. Answering the skeptics: yes, standard volatility models do provide accurate forecasts [J]. *International Economic Review*, 1998, 39 (4): 885 - 905.
- [2] ANDERSEN T G, BOLLERSLEV T, DIEBOLD F, et al. Exchange rate returns standardized by realized volatility are (nearly) Gaussian [J]. *Multinational Finance Journal*, 2000, 4: 159 - 179.
- [3] ANDERSEN T G, BOLLERSLEV T, DIEBOLD F, et al. The distribution of realized exchange rate volatility [J]. *J Amer Stat Asso*, 2001, 96, 42 - 55.
- [4] ANDERSEN T G, BOLLERSLEV T, DIEBOLD F, et al. The distribution of realized stock return volatility [J]. *J Fina Econ*, 2001, 61: 43 - 76.
- [5] ANDERSEN T G, BOLLERSLEV T, DIEBOLD F, et al. Modeling and forecasting realized volatility [J]. *Econometrica*, 2003, 71(2): 579 - 625.
- [6] A'IT-SAHALIA Y, MYKLAND P. The effects of random and discrete sampling when estimating continuous-time diffusions [J]. *Econometrica*, 2003, 71: 483 - 549.
- [7] BARNDORFF-NIELSEN O, SHEPHARD N. Econometric analysis of realized volatility and its use in estimating stochastic volatility models [J]. *J Royal Stat Soci B*, 2002, 64: 253 - 280.
- [8] PODOLSKIJ M, VETTER M. Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps [J]. *Bernoulli*, 2009, 15: 634 - 658.