

高频数据日内波动特征的函数型分析

马晓波¹ 冯凌秉² 李 玮³

(1. 中南财经政法大学新华金融保险学院, 湖北 武汉 430064; 2. 中国人民大学统计学院, 北京 100872;

3. 中南财经政法大学统计与数学学院, 湖北 武汉 430073)

【摘要】近年来,对金融市场的资产收益率的波动轨迹的研究日益增多,文章引入函数型波动过程这一方法以研究中国股市的日内波动交易轨迹。收益的波动率可被视作一个平滑的函数型过程和一个白噪声过程的组合。在本文的模型中,我们首先从每一交易日的对数收益率过程中剔除掉漂移项,来获得波动率过程,随后利用函数型波动过程和函数型主成分分析一起刻画上证综指5分钟收益率序列的日内波动特征。

【关键词】金融市场; 日内高频收益率; 函数型波动; 函数型数据

一、引言

由于对金融资产收益率波动轨迹的建模和预测研究对资产配置、衍生品定价及风险管理等领域 (Poon and Granger, 2003) 有着重要作用,因此这方面的相关研究一直是近年来一个热门的研究方向。一些研究者曾将函数型数据分析理论应用于金融市场的日内波动和交易模式的研究中 (Müller, Sen and Stadtmüller, 2007)。他们假设收益率序列的观测值由一个潜在但未知的随机过程生成,而这些观测值可看作是该过程的实现值。在金融时间序列的波动性分析中,一个主要问题是波动过程通常是时变而且无法直接观测到的 (Bollerslev and Zhou, 2006)。因此,一种能够合适且精确地计算出金融市场波动性的模型就显得尤为必要。众所周知,波动率是指金融市场中价格偏离光滑均值轨迹的程度,它本身也具有时变和随机等性质,因此我们可以用函数型波动过程来描述波动轨迹 (Müller, Stadtmüller and Yao, 2006; Müller, Sen and Stadtmüller, 2007)。而且函数型数据分析 (FDA) 这一方法可以被应用研究金融收益波动及其交易模式。

由于函数型波动过程这一概念是由基于方差函数的非参数模型发展出来的 (Yu and Jones, 2004; Fan, 2005), 因此它本质上是基于非参数模型的,而且其假设条件与普通的时间序列波动模型相比更弱,只要求波动轨迹平滑并且存在二阶矩。本文模型是建立在 Müller, Sen and Stadtmüller (2007) 这一论文的基础上。它指出函数型波动过程基于收益率的可重复观测序列,并且该模型假设观测轨迹是由可重复且独立同分布的未知随机过程的实现值生成,这也是该模型的一个基本假设。该模型的主要任务即是将收益率日内波动轨迹模式的特征问题转换到函数型数据分析的框架下进行分析。本文中,我们还试图根据函数型主成分分析 (FPCA) 方法获得收益率日内波动过程的函数型主成分得分来表示日内波动的特征。

二、理论方法

(一) 函数型波动过程

本文选择将扩散模型中的数据离散化,以适用于实际中的离散观测数据。故定义尺度化对数收益率 (scaled log-return) 及其扩散项如下:

$$\begin{aligned} R_1(t, \Delta) &= \frac{1}{\sqrt{\Delta}} \int_t^{t+\Delta} \beta(v) dv - \beta(t) \sqrt{\Delta} \\ R_2(t, \Delta) &= \frac{1}{\sqrt{\Delta}} \int_t^{t+\Delta} \theta(v) dW(v) - \theta(t) W_\Delta(t) \end{aligned} \quad (1.1)$$

于是对于日内高频数据 $Y_\Delta(t)$, 模型可以改写为如下形式:

$$\begin{aligned} Y_\Delta(t) &= \frac{1}{\sqrt{\Delta}} \int_t^{t+\Delta} \tilde{\mu}(v) dv + \frac{1}{\sqrt{\Delta}} \int_t^{t+\Delta} \tilde{\sigma}(v) dW(v) \\ &= \tilde{\mu}(t) \sqrt{\Delta} + \tilde{\sigma}(t) W_\Delta(t) + R_1(t, \Delta) + R_2(t, \Delta) \end{aligned} \quad (1.2)$$

因为波动率本质上是而非平滑的随机过程,所以我们此处定义了一个内在是平滑的函数型波动过程。如上所述,在以很高的抽样频率记录这些观测值的条件下,样本观测值 $Y_\Delta(t)$ 近似等于随机过程 $\tilde{\sigma}(t) W_\Delta(t)$, 其中, $t_j = j\Delta$, $j=1, 2, \dots, [T/\Delta]$ 。函数型波动过程不仅具有平滑轨迹而且也内在刻画了波动率的特征,此外,它还与参数 Δ 互相独立。

(二) 函数型主成分分析 (FPCA)

在函数型数据分析 (FDA) 中,正则化方法被广泛应用。将拟合优度的计量与粗糙惩罚函数结合在一起是 FDA 的一个基本思想。因此我们将这一思想纳入至 FPCA 中,以得到更优的分析效果。给定任一满足条件 $\|\xi\|_2 = 1$ 的主成分函数 ξ , 我们将样本方差 $\text{var}(\int \xi x_i)$ 与惩罚项 $(1 + \lambda \cdot \text{PEN}_2(\xi))$ 的比值定义为惩罚样本方差,即

$$\text{PCAPSV}(\xi) = \frac{\text{var}(\int \xi x_i)}{\|\xi\|^2 + \lambda \cdot \text{PEN}_2(\xi)} \quad (2.1)$$

在参数 λ 的值保持不变时,由于 $\text{PEN}_2(\xi)$ 的值随着 ξ 的粗糙度的增加而增加,因此惩罚样本方差 $\text{PCAPSV}(\xi)$ 的值会逐渐减小。

三、实证分析

(一) 数据初步整理及函数型波动过程的构造

本文选取我国上证综合指数的5分钟收益率以及相应的成交量作为本文的研究对象,时间选取从2011年1月20日~2011年4月15日,总共120天。我们将5分钟对数收盘价格进行差分得到相应的对数收益率,但是为了避免“隔夜效应

(overnight effect)影响,无论是对日内的收益率数据还是成交量,我们都将舍弃第一个观测值(即9:35时刻)。本文使用带有粗糙度惩罚函数的样条平滑方法以获得每日的日内均值漂移过程,并根据广义交叉验证(general cross validation, GCV)准则对每一条平滑曲线选择相同的平滑参数 λ ,以避免大量冗余重复的计算和平滑曲线欠光滑的可能。在剔除掉均值漂移项后,记作 $Y_{\Delta} = \log(\{R_{\Delta}^i\}^2) - q_0$,令其作为下文函数型波动过程分析的原始波动率。

(二) 日内收益率的波动率的函数型主成分分析

首先,我们需要估计出日内函数型波动过程 $V(t)$ 的均值函数 $\bar{\mu}_v$ 和协方差曲面 $\bar{G}_v(s, t)$,本文分别使用带有粗糙惩罚函数的一元和二元样条平滑方法估计它们。日内函数型波动曲线总体上呈现一个下降的趋势,但是在中午时分波动曲线有一个“凹陷”,并在这之后有一个小“隆起”。Andersen and Bollerslev (1998)认为日内波动率曲线在正午时存在“凹陷”,是因为午餐期间,交易量下降导致收益率波动下降。这一理论也同样适用于中国股市,而且中国股市每天中午11:30~13:00之间会有长达一个半小时的休市时间,这也是中国股市的日内波动曲线在正午存在“凹陷”的主要原因之一。下午开市后,与“隔夜效应”类似,股票的价格形成过程也会受到午间休市这一因素的影响,出现收益率波动性增强的情形,并逐渐下降直至收盘。中国股市的日内收益率波动过程的协方差曲面并非像Müller, Sen and Stadtmüller (2006)的估计结果那样是一个抛物面,反而更像是一个双曲面。协方差曲面透视图表明中国股市日内波动率的协方差曲面对角线上的方差在整个协方差矩阵中占了很大的比重,并且总体上随时间呈现一个递增的趋势,除了分别在中午休市和下午收盘以前的一小段时间波动有所回落以外。这一特征充分表明了日内收益率波动存在异方差的特性。协方差曲面上的最大值大致发生于14:00~14:30,说明这一段时间可能存在较强的波动聚合性。

我们对原始波动率 Y_{Δ} 进行函数型主成分分析时,我们需要选择分析所用的主成分个数 K 。选择个数 K 的方法有很多种,比如传统多元统计分析中的碎石图、方差比例,或者舍一曲线交叉验证法(one-curve-leave-out cross-validation) (Rice and Silverman 1991)和伪AIC准则(Yao, Müller and Wang, 2005)等方法。本文使用第一种方法选择 K ,选择结果为 $K=4$ 。四个主成

份大约共解释了总方差的85.94%,其对应的特征根 λ_k 的值分别为19.55、6.64、4.77和3.66。波动轨迹表明日内收益率波动可分为4种各不相同的波动模式,例如波动率在早晨达到波峰、中午波谷、中午休市后的“隆起”,或者波动过程呈现单调递减这4种模式。

四、结论

在本文中,我们提出了一个主要基于函数型数据的函数型主成分分析和函数型典型相关分析过程,以探究股票收益率的波动模式,以及它和成交量之间的关系。我们采用Müller, Stadtmüller and Yao (2006)提出的函数型波动过程的定义,它将日内波动过程嵌入函数型分析的框架中。通过函数型主成分分析,我们发现了中国上证综指日内收益率波动过程中一些独特的特征,比如中午股市休市以后波动过程的小“隆起”、近似双曲面的协方差曲面等。但是中国股市的日内收益率波动过程也具有与其他文献中相似的特征,如由于隔夜效应造成波动率呈现一个递减下降的趋势,及由于午餐时间等原因使得波动过程在中午有一个小“波谷”等。类似的特征也可以在波动过程的函数型主成分分析中找到。

参 考 文 献

- [1] Bollerslev, T. and Zhou, H. Volatility puzzles: a simple framework for gauging return-volatility regression. *Journal of Econometrics*. Vol.131. No.1~2. 2006: 123~150
- [2] Fan, J. A selective overview of nonparametric methods in financial econometrics (with discussion). *Statistical Science*, Vol.20. 2005: 317~357
- [3] Müller, H.G. Sen R. and Stadtmüller, U. Functional data Analysis for Volatility Process. Preprint. 2007
- [4] Müller, H. G. Stadtmüller, U. and Yao, F. Functional variance processes. *Journal of American Statistical Association*. Vol.101. No. 9. 2006: 1007~1018
- [5] Poon, S.H. and Granger, C.W.J. Forecasting volatility in financial markets: A review. *Journal of Economic Literature*. Vol.41. No.2. 2003: 478~539
- [6] Rice, J.A. and Silverman, B.W. Estimating the mean and covariance structure nonparametrically when the data are curves. *Journal of Royal Statistical Society Series B*. Vol.53. No.1. 1991: 233~243

淘商网

中国企业商务第一门户

淘商网(<http://www.tsw100.com>),是全球企业间(B2B)电子商务的领航者,是全球国际贸易领域内领先、最活跃的网上交易市场和商人社区,每日向全球各地企业及商家提供数百万条商业信息。

淘商网隶属中国企业经济类核心期刊、中国企业实战指导期刊《企业导报》杂志,是全球企业发布商机的第一门户,全球企业交流与合作的第一平台。所有商机信息通过《企业导报》和淘商网同步发布,平面媒体和网络媒体强强联合,整合立体权威传播。

淘商网本着“一网淘尽天下商机”的目的,致力于为全球华人提供准确、实用、及时、丰富的商机信息,立足于打造全球领先的“可视化”电子商务平台,整合庞大的商机数据库,准确而高效地为每一位客户提供他们需要的产品信息、供需商机以及涉及其产品或服务的各类行业资讯,帮助他们更加有效、清晰、直观、生动、全方位展示企业形象、推广品牌、宣传产品、及实用技术,为企业提供高性价比的商业视频营销。

淘商网将凭借巨大而翔实的商业信息数据库,便捷而高效的功能和服务,帮助众多供应商和采购商建立联系、提供商业机会,为企业产品进入国内和国际市场开启了一扇方便的电子商务之门,力争成为全球商机网络推广的首选网站。