中国科学: 数学 2020 年 第 50 卷 第 10 期: 1455～1486
SCIENTIA SINICA Mathematica
论 文

带跳高频数据下高维积分波动率矩阵估计

穆燕[1], 周勇[2*]

1. 南京财经大学经济学院统计系, 南京 210023;
2. 华东师范大学统计交叉科学研究院和统计学院, 统计与数据科学前沿理论及应用教育部重点实验室, 上海 200062
E-mail: muyan1111@126.com, yzhou@amss.ac.cn

收稿日期: 2019-01-24; 接受日期: 2019-09-17; 网络出版日期: 2020-06-24; * 通信作者
国家自然科学基金 (批准号: 71931004, 91546202 和 71901118) 资助项目

摘要 资产的联合波动率矩阵是资源配置和风险管理的重要统计量, 对其准确估计是金融统计和风险度量中的热点问题之一. 本文在带有市场信息的微观结构噪声下, 研究带跳对数价格的积分波动率矩阵估计问题. 在多资产价格观察不同步下, 当资产数和样本量都趋向于无穷时, 利用不重叠区间方法和稀疏性特征提出高维积分波动率矩阵的 4 种估计方法, 其收敛速度可以达到已存在高维积分波动率矩阵估计的最优收敛速度. 同时, 所提出的调整估计量具有相合性和半正定性. 模拟研究中比较了这些估计量的优缺点, 并将其应用于上海证券指数数据的实证研究中.

关键词 高频数据 跳 市场微观结构 高维波动率矩阵 半正定性

MSC (2010) 主题分类 97K80, 91G70

1 引言

如今的社会是一个信息化、数字化的社会, 互联网、物联网和云计算技术的迅猛发展, 使得数据充满着全天下. 与此同时, 数据也成为一种新的自然资源, 亟待人们对其加以合理、高效、充分地应用, 使之给人们的生活、工作带来更大的效益和价值. 在这类背景下, 数据的数目不仅以指数形式递增, 而且数据的结构越来越趋于复杂化, 这就赋予 “大数据” 不同于以往通俗 “数据” 的内在结构. 随着电子信息技术在国际证券市场上越发地普及应用, 人们可以十分便捷地记录和存储高频数据和超高频数据 (日内以分钟、秒或分笔为采集频率). 因此, 整个金融市场数据呈现样本量大、资产数多的特点. 如何充分利用这些海量样本数据帮助我们更好地理解市场、更有效管理市场风险, 已经成为经济界、金融界以及统计界研究学者关心的热点问题.

现代金融的投资趋向于多元化发展, 由单一资产到众多资产, 从国内市场到国际市场不断地扩展, 即便在同一个市场中也会考虑在不同板块和不同行业间进行资金配置, 所以建立各项资产之间有效的波动关系对于资产选择和优化配置来说至关重要. Markowitz 最早于 1952 年在二维分析框架下说明

英文引用格式: Mu Y, Zhou Y. High-dimensional integrated volatility matrix estimation for high-frequency financial data with jumps (in Chinese). Sci Sin Math, 2020, 50: 1455–1486, doi: 10.1360/N012019-00030

© 2020《中国科学》杂志社

了波动率矩阵在资产投资组合选择模型中扮演十分重要的角色; 在优化投资组合时, 联合波动率矩阵是非常重要的一个输入值, 其估计的优良与否直接影响投资组合的风险.

现实的金融资产组合中通常涉及多达数百个甚至上千个资产, 因此涉及的波动率矩阵通常是一个高维情形. 目前与高维波动率矩阵估计有关的主导模型有两大类: 因子模型和连续扩散模型.

因子模型的好处在于其将众多资产的信息进行重新组合, 找出影响变量的共同结构因子, 进而达到降维的目的, 从某种意义上降低了高维矩阵的估计难度; 另一方面, 它通过旋转使得因子变量更具有可解释性. 最早由 Fan 等 [1] 在不假设波动率矩阵的稀疏性情形下, 直接对观察的对数价格数据建立因子模型, 然后通过 Plug-in 最小二乘法估计因子负载矩阵. 但是他们所考虑的误差协方差矩阵是对角矩阵, 进而得到波动率矩阵估计. 后来 Fan 等 [2] 考虑了在因子已知、误差协方差矩阵是稀疏的情形下, 通过门限截断方法获得误差协方差矩阵进行积分波动率矩阵估计, 但在实际中寻找因子并非一件容易的事情. 因此, Fan 等 [3] 又考虑因子未知的情形, 利用门限方法提出了 "POET" (principal orthogonal complement thresholding) 估计量, 并将其应用到投资组合进行资产配置. 然而以上的文献所考虑的因子都是离散的形式, 于是, Fan 等 [4] 考虑了可观测的有限个连续的已知 Itô 半鞅因子过程估计高维积分波动率矩阵, 并将其应用到投资组合进行资产配置; 后来, Aït-Sahalia 和 Xiu [5] 在文献 [4] 的基础上考虑未知的连续的 Itô 半鞅因子模型, 给出了新的高维波动率矩阵估计量及其应用. 鉴于前面所获得的协方差估计量不具有稳健性, Fan 等 [6] 在因子已知、误差协方差矩阵是稀疏的情形下, 利用 Huber 损失函数获得了具有稳健性的协方差矩阵估计. Dai 等 [7] 在因子和异质性成分为半鞅、负载矩阵为常数的条件下分别讨论了因子已知负载未知、因子未知负载已知以及二者皆未知的情形下高维协方差矩阵的估计及其渐近性质. 但是在因子模型下计算因子得分时, 采用最小二乘法有时可能会失效; 其次, 因子模型通常假设同一天的波动率矩阵是一个常数矩阵, 并且所用的金融数据大都是低频数据.

Wang 和 Zou [8] 及 Tao 等 [9] 认为一天的波动率矩阵不应该是常数矩阵, 与时间是有关系的, 但是鉴于点波动率矩阵的时变性和记忆性, 很难被直接估计出来, 所以利用积分波动率矩阵估计代替对点波动率矩阵的估计. 同时在连续扩散模型下, 他们发现当估计大量资产的积分波动率矩阵时, 已存在的已实现波动率矩阵估计量具有很差的表现行为. 事实上, 当资产数和样本量同阶趋向于无穷时, 对于独立同分布的随机过程, 高维积分波动率矩阵的估计量是不相合的. 因此, 当资产数 (p) 和样本量 (n) 趋向于无穷时 (这里的 $p/n \to 0$), Wang 和 Zou [8] 使用高频数据在稀疏性的条件下通过非参的方法首次提出了高维积分波动率矩阵的相合估计量 ARVM (averaging realized volatility matrix); 并且证明了该相合估计的收敛速度依赖于样本量, 收敛速度达到 1/6. Tao 等 [9] 对 ARVM 估计量进行了改进, 利用多阶段 PT (previous-tick) 方法和门限技巧, 提出了另一个高频金融数据积分波动率矩阵的相合估计量 MSRVM (multi-scale realized volatility matrix), 并证明了在简单加性微观结构噪声的情形下, 该估计量的收敛速度达到 1/4. Kim 等 [10] 总结了近几年利用高频数据估计高维积分波动率矩阵的文献, 最终研究发现对于估计量 KRVM (kernel realized volatility matrix)、PRVM (pre-averaging realized volatility matrix) 和 MSRVM, 这些矩阵的每一个元素的收敛速度达到 1/4, 但是这些矩阵并不能保证半正定性. 尽管 KRVPM (kernel realized volatility positive semi-definite matrix) 和 PRVPM (pre-averaging realized volatility positive semi-definite matrix) 具有半正定性, 但它们的收敛速度降到了 1/5. 换句话说, 收敛速度与半正定性二者不可兼得. 穆燕等 [11] 在带有交易信息的微观结构噪声下获得了逐点收敛速度更高 (达到 1/3) 且具有半正定性的高维积分波动率矩阵估计量 AERVM (averaging estimated-price realized volatility matrix).

综上所述, 大部分文献中仅用到资产的价格数据进行建模研究, 然而大数据时代, 除价格数据外,

像交易量、交易率、买卖价差等都会被记录下来, 这些数据或多或少都影响着资产之间的联合波动, 因此充分利用这些数据估计众多资产的波动率矩阵是十分必要的. 于是, 本文将市场微观噪声作为市场交易信息的参数函数, 即将这些交易信息作为协变量构建市场微观结构噪声. Li 等[12] 用该思想估计单个资产的积分波动率, 并获得了非常好的收敛速度. 虽然穆燕等[11] 在带有交易信息的微观结构噪声下, 获得了收敛速度更高的积分波动率矩阵估计 AERVM, 但是他们考虑的是连续价格模型.

目前大量文献研究表明金融数据的跳是普遍存在的, 例如, Fan 和 Wang[13] 通过实际的高频样本数据检测出跳的存在性, 然后对波动率进行分析; Andersen 等[14] 利用高频数据在带跳的价格模型下利用最近邻替代技术提出了单资产的相合估计 MinRV 和 MedRV; 唐勇和林欣[15] 基于高频数据视角利用上证综指和深圳成指检验了跳和协同跳跃的存在性, 并发现这两个市场发生正跳和负跳的次数相当; Kong 等[16] 检验了高频数据的纯跳过程. 但是他们考虑的仅是单个资产的积分波动率情形. 尽管吴奔和张波[17] 在 Li 等[12] 的基础上对已实现波动率进行改进, 给出了判别跳跃的方法, 进而获得了更好的积分波动率估计, 但是他们是通过引入 Gauss 混合模型来刻画资产的对数价格. 因此, 本文着重在带跳的价格模型下研究众多资产波动率矩阵的高频估计.

本文余下内容的结构如下: 第 2 节具体介绍本文所研究的观察数据模型和价格模型; 所提出的估计方法及估计量的调整将在第 3 节展现; 第 4 节建立估计量的相合性及其收敛速度; 第 5 节对所提出的估计量与已存在的高维波动率矩阵估计量进行模拟比较; 第 6 节将所提出的高维波动率矩阵估计量应用于实际高频上证指数数据; 第 7 节给出本文结论, 并提出一些可能的研究方向; 所有的定理假设和证明将会在附录 A 中给出.

2 模型假设

2.1 带有市场交易信息的微观结构噪声

考虑 p 个资产, 设 $X_i(t)$ 是第 i 个资产在时刻 t 的真实的对数价格, $i=1,\ldots,p$. 假设高频数据 $Y_i(t_{i,l})$ 是第 i 个资产在时间 $t_{i,l}$ $(l=1,\ldots,n_i)$ 被观察到的对数价格. 高频数据通常受市场微观结构噪声污染, 以至于 $Y_i(t_{i,l})$ 是真实对数价格 $X_i(t_{i,l})$ 被污染的观察对数价格. 一般假设

$$\begin{aligned} &Y_i(t_{i,l}) = X_i(t_{i,l}) + \varepsilon_i(t_{i,l}), \quad i=1,\ldots,p, \quad l=1,\ldots,n_i, \\ &\varepsilon_i(t_{i,l}) = g(\boldsymbol{Z}_i(t_{i,l}); \boldsymbol{\theta}_i), \end{aligned} \tag{2.1}$$

其中 $\boldsymbol{Z}_i(t_{i,l})$ 是市场交易信息集, 其包含交易量、交易类型、交易率和买卖范围等; $\boldsymbol{\theta}_i$ 是一个有限维参数. 在第 5 节模拟研究中, 我们将会讨论 g 函数的具体形式 (包括线性结构和非线性结构) 的模拟情形.

注 2.1 在实波动率文献中, 通常假设市场微观结构噪声 $\varepsilon_i(t_{i,l})$ 是独立同分布的, 且独立于真实的对数价格过程; 也有部分学者研究其结构, 例如, Hansen 和 Lunde[18] 及 Kalnina 和 Linton[19] 考虑单变量的微观结构噪声具有序列相关性、且与真实对数价格过程也是相关的情形. Kim 等[10] 认为噪声具有内生性, 进而获得高维波动率矩阵的估计, 并证明了在此模型下多种已存在的估计量的收敛速度是不发生变化的.

2.2 带跳的价格模型

设 $\boldsymbol{X}(t) = (X_1(t),\ldots,X_p(t))^{\mathrm{T}}$ 是 p 个资产在时间 t 的真实对数价格, 标准的无套利条件要求资

产的价格过程是一个半鞅过程, 即在常规的域流概率空间 $(\Omega, \mathcal{F}, (\mathcal{F})_{t\geqslant 0}, \mathrm{P})$ 上, 假设 $\boldsymbol{X}(t)$ 服从半鞅过程:

$$\boldsymbol{X}(t) = \boldsymbol{X}^C(t) + \boldsymbol{X}^D(t), \quad t \in [0, T^*], \tag{2.2}$$

其中 $\boldsymbol{X}^C(t)$ 和 $\boldsymbol{X}^D(t)$ 分别表示价格模型的连续部分和不连续部分, 它们的具体形式为

$$\begin{aligned}
\boldsymbol{X}^C(t) &= \boldsymbol{X}_0 + \int_0^t \boldsymbol{\mu}_s ds + \int_0^t \boldsymbol{\sigma}_s^{\mathrm{T}} d\boldsymbol{B}_s, \quad t \in [0, T^*], \\
\boldsymbol{X}^D(t) &= \int_0^t \boldsymbol{J}_t \times d\boldsymbol{N}_t,
\end{aligned}$$

其中 $\boldsymbol{J}_t$ 和 $\boldsymbol{N}_t$ 分别记为 p 个资产的跳的大小和次数, 这里假设 $\boldsymbol{N}_t$ 是 p 维的 Poisson 过程, 其强度为 $\boldsymbol{\lambda}$; "$A \times B$" 表示向量 A 和 B 的对应元素相乘, 所获得的结果仍是向量. $\boldsymbol{\mu}_t$ 是 p 维的漂移向量, $\boldsymbol{B}_t = (B_{1t}, \ldots, B_{pt})$ 是标准的 p 维 Brown 运动 (即 B_{it} 是独立的标准 Brown 运动), $\boldsymbol{\sigma}_t$ 是 càdlàg 适应的 $p \times p$ 协方差矩阵. $[0, T^*]$ 是观测区间, 不妨假设 $T^* = 1$, 那么价格的连续部分 $\boldsymbol{X}^C(t)$ 波动率为

$$\boldsymbol{\gamma}(t) = (\boldsymbol{\gamma}_{ij}(t))_{1\leqslant i,j\leqslant p} = \boldsymbol{\sigma}_t^{\mathrm{T}} \boldsymbol{\sigma}_t, \tag{2.3}$$

且它的二次变差为

$$\begin{aligned}
[\boldsymbol{X}^C, \boldsymbol{X}^C]_t &= \sum_{t_l} (\boldsymbol{X}^C(t_{l+1}) - \boldsymbol{X}^C(t_l))^{\mathrm{T}} (\boldsymbol{X}^C(t_{l+1}) - \boldsymbol{X}^C(t_l)), \\
\lim_{\Delta\to 0} [\boldsymbol{X}^C, \boldsymbol{X}^C]_t &= \int_0^t \boldsymbol{\gamma}(s) ds = g\bigg(\int_0^t \gamma_{ij}(s) dsg\bigg)_{1\leqslant i,j\leqslant p}, \quad t \in [0, 1],
\end{aligned}$$

其中 $\Delta = \max\{t_{l+1} - t_l, l = 1, 2, \ldots, n\}$. 因为 $\boldsymbol{X}^C(t)$ 二次变差的极限是积分波动率矩阵, 跟传统的连续价格模型结论一样, 金融资产的点波动率矩阵具有较强的记忆性和时变性, 很难精准地获得, 所以通常利用资产的积分波动率矩阵估计代替对资产点波动率矩阵的准确估计, 即本文在带跳的价格模型下估计目标是

$$\boldsymbol{\Gamma} = (\Gamma_{ij})_{1\leqslant i,j\leqslant p} = \int_0^1 \boldsymbol{\gamma}(s) ds = g\bigg(\int_0^1 \gamma_{ij}(s) dsg\bigg)_{1\leqslant i,j\leqslant p}.$$

通常估计积分波动率矩阵将面临如下 4 大问题:

(1) 非同步性. 在实际中, 由于市场信息到达不同投资者的时间存在差异和交易市场的摩擦等因素导致不同的证券对交易信息的反映存在时滞, 进而致使不同的证券具有不同的交易频率, 因此, 非同步性的处理直接影响着积分波动率矩阵的估计效果.

(2) 微观结构噪声. 从交易平台所获得的高频金融数据往往是被污染过后的观测数据, 真实的数据是无法被直接观测到的, 而它们之间的差便是噪声, 因此准确的噪声刻画对积分波动率矩阵的估计相当关键.

(3) 提取价格连续部分. 观测的价格过程包含连续部分和跳跃部分, 只有将离散的跳跃部分从价格过程中提取出去, 才能获得价格过程的连续部分, 进而准确地估计积分波动率矩阵.

(4) 估计量的相合性和半正定性. 估计量的相合性和半正定性是衡量积分波动率矩阵准确估计的重要标准.

因此, 本文将克服以上问题, 在带有市场信息的微观结构噪声和带跳的价格模型下对积分波动率矩阵进行估计.

3 积分波动率矩阵的估计

3.1 参数估计

根据参数形式的市场微观结构噪声 (2.1) 和对数价格模型 (2.2), 首先用已有的市场交易信息估计参数 $\boldsymbol{\theta}_i$ $(i=1,\ldots,p)$. 记

$$\Delta X_i(t_{i,l}) = X_i(t_{i,l}) - X_i(t_{i,l-1}), \quad \Delta Y_i(t_{i,l}) = Y_i(t_{i,l}) - Y_i(t_{i,l-1}),$$
$$\Delta g(\boldsymbol{Z}_i(t_k);\boldsymbol{\theta}_i) = g(\boldsymbol{Z}_i(t_k);\boldsymbol{\theta}_i) - g(\boldsymbol{Z}_i(t_{k-1});\boldsymbol{\theta}_i),$$

则有

$$\Delta Y_i(t_{i,l}) = \Delta g(\boldsymbol{Z}_i(t_{i,l});\boldsymbol{\theta}_i) + \Delta X_i(t_{i,l}).$$

Li 等[12] 之所以可以转换 "噪声" 与 "信号" 的角色, 是因为在连续扩散模型中 $\Delta X_i(t_{i,l}) \sim N(0,\sigma_i^2/n_i)$, 然而本文考虑的是半鞅过程, 虽然 $\Delta X_i(t_{i,l})$ 不一定是均值为 0 的正态分布, 但是 $\Delta X_i^C(t_{i,l}) \sim N(0,\sigma_i^2/n_i)$, 也就是说,

$$\Delta Y_i^C(t_{i,l}) = \Delta g(\boldsymbol{Z}_i(t_{i,l});\boldsymbol{\theta}_i) + \Delta X_i^C(t_{i,l}),$$

其中 $Y_i^C(t_{i,l}) := Y_i(t_{i,l})I(|Y_i(t_{i,l})| \leqslant C)$. 由第 4 节的假设知资产跳跃次数是有限的, 那么上式几乎处处成立. 因此, 如果 $\Delta X_i^C(t_{i,l})$ 独立于 $\Delta g(\boldsymbol{Z}_i(t_{i,l});\boldsymbol{\theta}_i)$, 那么 $\boldsymbol{\theta}_i$ 的极大似然估计 $\widehat{\boldsymbol{\theta}}_i$ 为

$$\widehat{\boldsymbol{\theta}}_i = \arg\min_{C,\boldsymbol{\theta}_i} Q_{n_i}(Y_i^C, \boldsymbol{Z}_i, \boldsymbol{\theta}_i), \quad i=1,\ldots,p,$$
$$Q_{n_i}(Y_i, \boldsymbol{Z}_i, \boldsymbol{\theta}_i) = \frac{1}{2}\sum_{l=1}^{n_i}(\Delta Y_i^C(t_{i,l}) - \Delta g(\boldsymbol{Z}_i(t_{i,l});\boldsymbol{\theta}_i))^2.$$

该参数估计满足一定的假设条件 (见假设 A.1), 可知极大似然估计量 $\widehat{\boldsymbol{\theta}}_i$ 收敛到 $\boldsymbol{\theta}_i$, 且 $\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i = O_p(1/n)$, 具体证明参见文献 [12, 定理 1]. 从而得到估计

$$\Delta \widehat{X}_i^C(t_{i,l}) = \Delta Y_i^C(t_{i,l}) - \Delta g(\boldsymbol{Z}_i(t_{i,l});\widehat{\boldsymbol{\theta}}_i). \tag{3.1}$$

根据已实现波动率的定义, 直接便可获得积分波动率矩阵估计.

3.2 连续部分估计

目前对于带跳的价格模型估计连续部分的估计量时, 普遍采用截断方法, 例如, Mancino 和 Sanfelici[20]、Aït-Sahalia 和 Jacod[21]、Li 等[12], 他们认为连续部分短时间内价格变化是很小的, 而发生跳跃时短时间的变化幅度较大, 因此, 通过截掉那些价格变化幅度大的, 保留价格变化幅度较小的成分, 便可获得价格的连续部分, 即

$$\Delta \widehat{X}_i'^C(t_{i,l}) = \Delta \widehat{X}_i(t_{i,l})I(|\Delta \widehat{X}_i(t_{i,l})| \leqslant C).$$

但这种截断方法对跳跃并不具有稳健性, 所以, Andersen 等[14] 利用最近邻截断提出了具有稳健性的截断方法: 单边截断 "MinRV" 和双边截断 "MedRV". 该方法认为发生跳跃的次数是有限的, 相邻的两段时间内最多发生一次跳跃, 其估计量形式如下:

$$\mathrm{MinRV}_n = \frac{\pi}{\pi-2}\left(\frac{n}{n-1}\right)\sum_{i=1}^{n-1}\min(|\Delta \widehat{X}_i|, |\Delta \widehat{X}_{i+1}|)^2,$$

$$\mathrm{MedRV}_n = \frac{\pi}{6-4\sqrt{3}+\pi}\left(\frac{n}{n-2}\right)\sum_{i=2}^{n-1}\mathrm{med}(|\Delta\widehat{X}_{i-1}|,|\Delta\widehat{X}_i|,|\Delta\widehat{X}_{i+1}|)^2,$$

这两种方法的截断对带跳的价格模型具有稳健性. 本文将分别利用以上 3 种方法以及参数估计的方法来估计高维积分波动率矩阵.

3.3 已实现波动率矩阵

对于高频金融数据的非同步性, 采用 Zhang[22] 的前值法 (previous tick) 进行处理, 假设对于固定的整数 m, 取 τ_r $(r=1,\ldots,m)$ 作为预先取定的样本频率. 记 $\boldsymbol{\tau}=\{\tau_r, r=1,\ldots,m\}$, 对于资产 i, 定义前值法为

$$\tau_{i,r}=\max\{t_{i,l}\leqslant\tau_r, l=1,\ldots,n_i\},\quad r=1,\ldots,m,$$

其中 $t_{i,l}$ 表示资产 i 第 l 笔交易时间. 也就是说对于事先确定的任意固定 τ_r $(r=1,\ldots,m)$, 取不大于 τ_r 且离 τ_r 最近一次资产 i 发生的交易时间 $\tau_{i,r}$ 作为资产 i 在 τ_r 时发生的交易. 这些 $\tau_{i,r}$ $(i=1,\ldots,p)$ 虽然并不相同, 但是它们与 τ_r 都是非常接近的, 因此认为它们是同时发生的, 进而达到高频数据的同步化目的. 图 1 可以帮助我们更好地理解前值法.

预先选定样本频率 $\boldsymbol{\tau}$ 通常选用等距的网格点形式, 对于固定的 m, 存在 $K=[n/m]$ 类不重叠的等距网格,

$$\boldsymbol{\tau}^k=\left\{\frac{r}{m},\ r=1,\ldots,m\right\}+\frac{k-1}{n}=\left\{\frac{r}{m}+\frac{k-1}{n}, r=1,\ldots,m\right\},$$

其中 $k=1,\ldots,K$; n 是平均样本大小,

$$n=\frac{1}{p}\sum_{i=1}^{p}n_i.$$

基于预先选定样本频率 $\boldsymbol{\tau}$ (通常选用等距的网格点形式), 对每个样本频率 $\boldsymbol{\tau}^k$, 定义资产 i 和 j 的已实现协波动率 $\widehat{\Gamma}_{ij}^{(l)}(\boldsymbol{\tau}^k)$ 和已实现波动率矩阵 $\widehat{\boldsymbol{\Gamma}}^{(l)}(\boldsymbol{\tau}^k)$, 即

$$\widehat{\Gamma}_{ij}^{(l)}=\frac{1}{K}\sum_{k=1}^{K}\widehat{\Gamma}_{ij}^{(l)}(\boldsymbol{\tau}^k),\quad \widehat{\boldsymbol{\Gamma}}^{(l)}=(\widehat{\Gamma}_{ij}^{(l)})=\frac{1}{K}\sum_{k=1}^{K}\widehat{\boldsymbol{\Gamma}}^{(l)}(\boldsymbol{\tau}^k),\quad l=1,2,3,4,$$

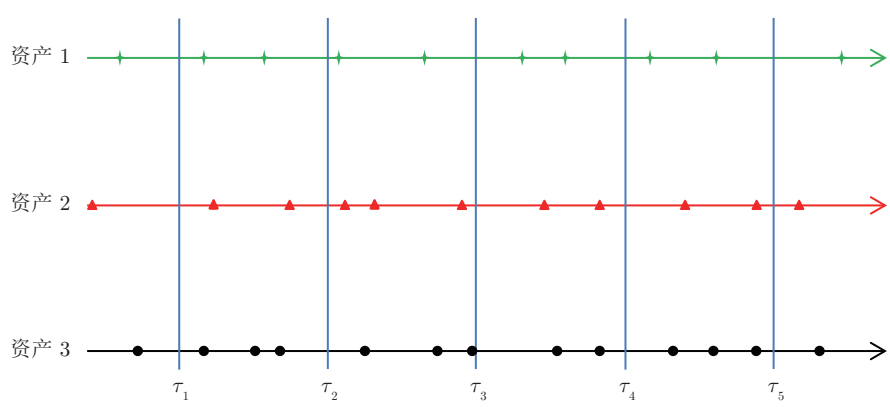

图 1 (网络版彩图) 3 个资产的前值法时间图

其中第 1 种积分波动率矩阵的估计由 (3.1) 直接获得, 即

$$\widehat{\Gamma}_{ij}^{(1)}(\boldsymbol{\tau})=\sum_{r=1}^{m}\Delta\widehat{X}_i^C(\tau_{i,r})\Delta\widehat{X}_j^C(\tau_{j,r}).$$

接下来介绍的 3 个估计量受单资产跳跃模型的启示获得. 第 2 种就是常见的截断方法, 将变化幅度大的认为是发生跳跃的情形, 将其舍掉, 但截断的门限并非固定常数, 而与交易的间隔和参数 ξ 有关, 则第 i 和 j 个资产的协波动率为

$$\widehat{\Gamma}_{ij}^{(2)}(\boldsymbol{\tau})=\sum_{r=1}^{m}[\Delta\widehat{X}_i(\tau_{i,r})I(|\Delta\widehat{X}_i(\tau_{i,r})|\leqslant(\Delta\tau_{i,r})^{\xi})][\Delta\widehat{X}_j(\tau_{j,r})I(|\Delta\widehat{X}_j(\tau_{j,r})|\leqslant(\Delta\tau_{j,r})^{\xi})],$$

其中 $\widehat{X}_i(\tau_{i,r})=Y_i(\tau_{i,r})-g(\boldsymbol{Z}_i(t_{i,l});\widehat{\boldsymbol{\theta}}_i)$. 第 3 种估计方法是对单资产积分波动率估计方法 MinRV 的推广, 定义第 i 和 j 个资产的协波动率为

$$\begin{aligned}\widehat{\Gamma}_{ij}^{(3)}(\boldsymbol{\tau})=&\sum_{r=1}^{m-1}[\Delta\widehat{X}_i(\arg_t\min\{|\Delta\widehat{X}_i(t)|,t=\tau_{i,r},\tau_{i,r+1}\})]\\&\times[\Delta\widehat{X}_j(\arg_t\min\{|\Delta\widehat{X}_j(t)|,t=\tau_{j,r},\tau_{j,r+1}\})].\end{aligned}$$

第 4 种估计方法事实上是第 3 种方法的推广, 因为第 3 种估计量考虑的是单边的影响, 而第 4 种考虑了双边的影响, 因此定义第 i 和 j 个资产的协波动率为

$$\begin{aligned}\widehat{\Gamma}_{ij}^{(4)}(\boldsymbol{\tau})=&\sum_{r=2}^{m-1}[\Delta\widehat{X}_i(\arg_t\operatorname{med}\{|\Delta\widehat{X}_i(t)|,t=\tau_{i,r-1},\tau_{i,r},\tau_{i,r+1}\})]\\&\times[\Delta\widehat{X}_j(\arg_t\operatorname{med}\{|\Delta\widehat{X}_j(t)|,t=\tau_{j,r-1},\tau_{j,r},\tau_{j,r+1}\})],\end{aligned}$$

这里的符号 $\operatorname{med}(a,b,c)$ 表示取中间量. 则已实现波动率矩阵为

$$\widehat{\boldsymbol{\Gamma}}(\boldsymbol{\tau})=(\widehat{\Gamma}_{ij}^{(l)}(\boldsymbol{\tau}))_{i,j=1,\ldots,p},\quad l=1,2,3,4.$$

这 4 种估计量称为 AERVM、TAERVM、MinRVM 和 MedRVM, 其中 MinRVM 和 MedRVM 的定义没有单资产定义的系数是因为 Andersen 等 [14] 进行估计时考虑的噪声是 Gauss 白噪声, 系数是为了对偏差进行修正, 而本文对噪声直接建模估计, 对噪声的刻画更准确, 因而不需要修正. 这里 4 种估计量都是利用 pre-averaging 方法, 目的主要是减少市场微观结构噪声的影响, 进而获得更好的积分波动率矩阵估计.

3.4 估计量的调整

当 p 比较小时, $\widehat{\boldsymbol{\Gamma}}^{(l)}$ $(l=1,2,3,4)$ 通常是 $\boldsymbol{\Gamma}$ 的一个很好的估计量; 然而当 p 比较大时, Wang 和 Zou [8] 举例说明了该估计量的表现是非常糟糕的. 对于常数 $\boldsymbol{\gamma}(t)$, 当 n 和 p 趋向于无穷时, $\widehat{\boldsymbol{\Gamma}}^{(l)}$ $(l=1,2,3,4)$ 都不是相合估计. 因此需要给真实的波动率矩阵 $\boldsymbol{\Gamma}$ 强加一些结构特征. 根据文献 [8,9,23,24] 等中的特殊结构, 本文主要考虑 $\boldsymbol{\Gamma}$ 的稀疏条件假设, 即假设 $\boldsymbol{\Gamma}$ 满足

$$\sum_{j=1}^{p}|\Gamma_{ij}|^{\delta}\leqslant M\pi(p),\quad i=1,\ldots,p,\quad \mathrm{E}[M]\leqslant C,\tag{3.2}$$

其中 M 为正随机变量; $0\leqslant\delta<1$; $\pi(p)$ 是 p 的慢变函数, 如 1、$\ln(p)$ 和 $\ln(\ln(p))$ 等.

注 3.1 稀疏条件假设包含的形式比较丰富, 例如, 满足稀疏条件的稀疏矩阵包括分块对角矩阵、对角退化元素矩阵、每行或每列相对较小的非零元素矩阵以及随机排列前面 3 种矩阵的行和列所得到的矩阵. 此外, 稀疏条件具有一定的实际意义, 它保留了在驱动波动率运动和捕捉市场风险中起到主导性作用的相对少数的市场因子, 如此方便投资者具有针对性地作出投资.

门限调整: 假设 $\boldsymbol{\Gamma}$ 满足稀疏条件, 它重要的元素是 $\boldsymbol{\Gamma}$ 中那些绝对值超过门限的元素. 定义如下:

$$\mathcal{T}_{\varpi}^{T}[\widehat{\boldsymbol{\Gamma}}] = (\widehat{\Gamma}_{ij} I(|\widehat{\Gamma}_{ij}| \geqslant \varpi)),$$

其中 ϖ 是门限调整参数, $I(\cdot)$ 是示性函数.

T- 截断估计量与多数已存在的波动率矩阵估计量一样, 不能保证有限样本的估计量半正定性. 受降维思想的启示, Kim 等 [10] 引入 Hard- 门限调整和 Soft- 门限调整保证了原有半正定矩阵的半正定性. 假设 $\boldsymbol{\Gamma}$ 满足稀疏条件, 定义 Hard- 门限调整:

$$\mathcal{T}_{\varpi}^{H}[\widehat{\boldsymbol{\Gamma}}] = (\widehat{\Gamma}_{ij} I(|\widehat{\Gamma}_{ij}| \geqslant \varpi, i \neq j)) + (\widehat{\Gamma}_{ij} I(|\widehat{\Gamma}_{ij}| \geqslant 0, i = j)),$$

Soft- 门限调整:

$$\mathcal{T}_{\varpi}^{S}[\widehat{\boldsymbol{\Gamma}}] = ((\widehat{\Gamma}_{ij} - \mathrm{sign}(\widehat{\Gamma}_{ij})\varpi) I(|\widehat{\Gamma}_{ij}| \geqslant \varpi, i \neq j)) + (\widehat{\Gamma}_{ij} I(|\widehat{\Gamma}_{ij}| \geqslant 0, i = j)),$$

其中 ϖ 是门限调整参数, $I(\cdot)$ 是示性函数. $\mathrm{sign}(\cdot)$ 是符号函数, 即如果 $\widehat{\Gamma}_{ij} > 0$, 那么 $\mathrm{sign}(\widehat{\Gamma}_{ij}) = 1$; 如果 $\widehat{\Gamma}_{ij} < 0$, 那么 $\mathrm{sign}(\widehat{\Gamma}_{ij}) = -1$; 如果 $\widehat{\Gamma}_{ij} = 0$, 那么 $\mathrm{sign}(\widehat{\Gamma}_{ij}) = 0$.

注 3.2 尽管 Hard- 门限调整估计量 $\mathcal{T}_{\varpi}^{H}[\widehat{\boldsymbol{\Gamma}}]$ 与门限调整估计量 $\mathcal{T}_{\varpi}[\widehat{\boldsymbol{\Gamma}}]$ 定义比较类似, 关键不同在于 $\mathcal{T}_{\varpi}^{H}[\widehat{\boldsymbol{\Gamma}}]$ 保留了那些低于门限的对角元素, 进而保留了原来矩阵的主对角占优的性质, 所以, 该门限截断调整方法不改变矩阵原有的半正定性. 也就是说, 如果估计量 $\widehat{\boldsymbol{\Gamma}}$ 本身是半正定的, 如 KRVPM 和 PRVPM, 通过 Hard- 门限调整过后依然具有半正定性. 然而, 文献 [8,9] 中的带宽调整方法和门限调整方法所获得的估计量是不能够保证半正定性的, 即使调整前的估计量具有半正定性. 类似地, 由于 Soft- 门限调整估计量是 $\widehat{\boldsymbol{\Gamma}}$ 的连续函数, 如果 $\widehat{\boldsymbol{\Gamma}}$ 是半正定矩阵, 则 $\mathcal{T}_{\varpi}^{S}[\widehat{\boldsymbol{\Gamma}}]$ 也是半正定矩阵.

4 渐近理论

为方便叙述所提出估计量的渐近性质, 设 $\boldsymbol{x} = (x_1, \ldots, x_p)^{\mathrm{T}}$ 是 p 维向量, $\boldsymbol{U} = (U_{ij})$ 是 $p \times p$ 的矩阵. 定义 l_d- 范数,

$$\|\boldsymbol{x}\|_d = \left(\sum_{i=1}^{p} |x_i|^d\right)^{1/d}, \quad \|\boldsymbol{U}\|_d = \sup\{\|\boldsymbol{U}\boldsymbol{x}\|_d, \|\boldsymbol{x}\|_d = 1\}, \quad d = 1, 2, \infty.$$

矩阵模 $\|\boldsymbol{U}\|_2$ 是 $\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}$ 的最大特征值的平方根, 且

$$\|\boldsymbol{U}\|_1 = \max_{1 \leqslant j \leqslant p} \sum_{i=1}^{p} |U_{ij}|, \quad \|\boldsymbol{U}\|_\infty = \max_{1 \leqslant i \leqslant p} \sum_{j=1}^{p} |U_{ij}|.$$

对于对称的 $\boldsymbol{U}$, 它的矩阵范数等于最大绝对特征值, $\|\boldsymbol{U}\|_2 \leqslant \|\boldsymbol{U}\|_1 = \|\boldsymbol{U}\|_\infty$. 本文将主要涉及 l_d- 范数, 其中 $d = 1, 2, \infty$.

为建立估计量的渐近理论, 需要一般的假设条件 (同穆燕等 [11] 的假设), 具体见附录. 此外, 本文考虑的对数价格是半鞅过程, 通常假设跳跃的次数是有限的, 即 $\max_{1 \leqslant i \leqslant p} \sum_{r=1}^{n} |\Delta X_i^D(\tau_{i,r})| = O_p(1)$.

该假设在带跳的价格模型中很常见, 如 Li 等[12]、Andersen 等[14]、Asai 和 McAleer[25], 他们所获得的结论都是基于资产价格发生有限次跳跃的假设.

考虑市场微观结构噪声是交易信息的参数函数时, 利用高频金融数据所提出的高维波动率矩阵估计量具有如下几个定理:

定理 4.1 在模型 (2.1)、(2.2) 和假设 A.1–A.3 以及跳跃有限的条件下,

(i) 对任意的 $1 \leqslant i, j \leqslant p$, 协方差估计量 $\widehat{\Gamma}_{ij}^{(l)}$ 满足

$$\mathrm{E}(|\widehat{\Gamma}_{ij}^{(l)} - \Gamma_{ij}|^{\beta}) \leqslant Cn^{-\frac{\beta}{3}}, \tag{4.1}$$

其中 C 是与 n 和 p 无关的常数, $\beta \geqslant 2$ (见证明中假设 A.2);

(ii) 估计量矩阵 $\widehat{\mathbf{\Gamma}}^{(l)}$ 具有半正定性, 其中 $\{\widehat{\mathbf{\Gamma}}^{(l)}, l = 1, 2, 3, 4\}$ 分别表示本节前面所提出的 4 种估计量.

注 4.1 定理 4.1 表明 4 种估计量所获得的结论同不带跳价格模型所获得估计, 逐点收敛速度达到最优, 且估计量本身具有半正定性.

定理 4.2 假设 $\mathbf{\Gamma}$ 满足稀疏条件, 在模型 (2.1)、(2.2) 和假设 A.1–A.3 以及跳跃有限的条件下,

(i) 高维波动率矩阵估计量 $\widehat{\mathbf{\Gamma}}^{(l)}$ 的门限调整满足

$$\|\mathcal{T}_{\varpi}^{T}[\widehat{\mathbf{\Gamma}}^{(l)}] - \mathbf{\Gamma}\|_2 \leqslant \|\mathcal{T}_{\varpi}[\widehat{\mathbf{\Gamma}}^{(l)}] - \mathbf{\Gamma}\|_{\infty} = O_P(\pi(p)\varpi^{1-\delta});$$

(ii) 高维波动率矩阵估计量 $\widehat{\mathbf{\Gamma}}^{(l)}$ 的 Hard- 门限调整满足

$$\|\mathcal{T}_{\varpi}^{H}[\widehat{\mathbf{\Gamma}}^{(l)}] - \mathbf{\Gamma}\|_2 \leqslant \|\mathcal{T}_{\varpi}^{H}[\widehat{\mathbf{\Gamma}}^{(l)}] - \mathbf{\Gamma}\|_{\infty} = O_P(\pi(p)\varpi^{1-\delta}),$$

并且估计量 $\mathcal{T}_{\varpi}^{H}[\widehat{\mathbf{\Gamma}}^{(l)}]$ 具有半正定性;

(iii) 高维波动率矩阵估计量 $\widehat{\mathbf{\Gamma}}^{(l)}$ 的 Soft- 门限调整满足

$$\|\mathcal{T}_{\varpi}^{S}[\widehat{\mathbf{\Gamma}}^{(l)}] - \mathbf{\Gamma}\|_2 \leqslant \|\mathcal{T}_{\varpi}^{S}[\widehat{\mathbf{\Gamma}}^{(l)}] - \mathbf{\Gamma}\|_{\infty} = O_P(\pi(p)\varpi^{1-\delta}),$$

并且估计量 $\mathcal{T}_{\varpi}^{S}[\widehat{\mathbf{\Gamma}}^{(l)}]$ 具有半正定性, 其中 $l = 1, 2, 3, 4$; 门限参数 $\varpi = n^{-1/3}p^{2/\beta}h_{n,p}$, $h_{n,p}$ 是任意一个以很慢的速度收敛到无穷的序列, 例如, $h_{n,p} = \ln\ln(n \wedge p)$.

注 4.2 同普通门限截断调整相比, 定理 4.2 进一步表明 Hard- 和 Soft- 门限截断方法对估计量进行截断调整没有改变估计的相合性和收敛速度, 同时还保证了原有估计量的半正定性.

5 模拟研究

5.1 模拟模型

本节通过数值模拟检验在带跳的价格模型下所提出的积分波动率矩阵估计量 MinRVM、MedRVM 和修正的 AERVM 估计量 TAERVM 的统计性质. 具体考虑下面 3 种形式的微观结构噪声模型[12], 分别为

$$g_1(V_{t_k}, I_{b/s}(t_k); \alpha, \beta) = \alpha I_{b/s}(t_k) + \frac{\beta I_{b/s}(t_k) V_{t_k}}{\Delta t_k},$$

$$g_2(V_{t_k}, I_{b/s}(t_k); \alpha, \beta^+, \beta^-) = \begin{cases} \alpha I_{b/s}(t_k) + \dfrac{\beta^+ I_{b/s}(t_k) V_{t_k}}{\Delta t_k}, & \text{买家发起}, \\ -\alpha I_{b/s}(t_k) - \dfrac{\beta^- I_{b/s}(t_k) V_{t_k}}{\Delta t_k}, & \text{卖家发起}, \end{cases}$$

$$g_3(V_{t_k}, I_{b/s}(t_k); \beta, \gamma) = I_{b/s}(t_k) \log\left(\gamma + \frac{\beta I_{b/s}(t_k) V_{t_k}}{\Delta t_k}\right).$$

在模拟当中, 假定 $\alpha = 1.875 \times 10^{-4}$, $\beta = \beta^+ = 0.75 \times 10^{-12}$, $\beta^- = 0.30 \times 10^{-12}$ 和 $\gamma = 1 + \alpha = 1.0001875$. 模型 g_1 同 Almgren 和 Chriss [26] 所考虑的买卖发起方和交易率的线性影响; 模型 g_2 考虑了交易率和非对称的买卖发起方的影响; 模型 g_3 考虑交易率的非线性影响. Keim 和 Madhavan [27] 发现对于买方噪声模型关于交易率是凹的, 而对于卖家而言噪声模型关于交易率是凸的.

价格模型设定为 $\boldsymbol{\mu}_t = \mathbf{0}$ 的模型 (2.2), 即

$$\begin{aligned} \boldsymbol{X}(t) &= \boldsymbol{X}^C(t) + \boldsymbol{X}^D(t) \\ &= \boldsymbol{X}_0 + \int_0^t \boldsymbol{\sigma}_s^{\mathrm{T}} d\boldsymbol{B}_s + \int_0^t \boldsymbol{J}_s \times d\boldsymbol{N}_s, \quad t \in [0,1], \end{aligned}$$

其中 $\boldsymbol{B}_t = (B_{1t}, \ldots, B_{pt})$ 是标准 p 维的 Brown 运动; $\boldsymbol{\gamma}(t) = (\boldsymbol{\gamma}_{ij}(t))_{1 \leqslant i,j \leqslant p} = \boldsymbol{\sigma}_t^{\mathrm{T}} \boldsymbol{\sigma}_t$. 设 $\boldsymbol{\gamma}_t$ 的对角元素服从 $(0, 0.1)$ 上的均匀分布, 给定 $\boldsymbol{\gamma}_t$ 对角元素, 定义非对角元素为

$$\gamma_{i,j}(t) = \kappa(t)^{|i-j|} \sqrt{\gamma_{i,i}(t)\gamma_{j,j}(t)},$$

其中过程

$$\kappa(t) = \frac{\mathrm{e}^{2u(t)} - 1}{\mathrm{e}^{2u(t)} + 1}, \quad du(t) = 0.03[0.64 - u(t)]dt + 0.118u(t)dW_{\kappa,t},$$

$$W_{\kappa,t} = \sqrt{0.96} W_{\kappa,t}^0 - 0.2 \sum_{i=1}^p \frac{B_{it}}{\sqrt{p}};$$

$W_{\kappa,t}^0$ 独立于 $\boldsymbol{B}_t$ 的一维标准 Brown 运动, 这些参数的设定主要参见文献 [8,9].

交易类型 $I_{b/s}$ 是通过概率为 1/2 的 Bernoulli 过程生成的, 认为买卖双方是以相同的概率发起交易的. 对于交易量的模型参见文献 [12], 建模如下:

$$V_{t_k} = H_{t_k} + e_{t_k},$$

其中 H_{t_k} $(k = 0, 1, \ldots, n)$ 是相互独立的、服从均值为 $\int_{t_{k-1}}^{t_k} h_t dt$、标准差为 5,000 的 Gamma 分布的随机变量, $h_t = (0.025(t - 0.5)^2 + 0.05) \times 10^8$; $\{e_{t_k}\}$ 服从 ARMA(1,1) 模型

$$e_{t_k} = \phi_1 e_{t_{k-1}} + \psi_1 u_{t_{k-1}},$$

其中 $\phi_1 = 0.5$, $\psi_1 = 0.5$, $\{u_{t_k}\}$ 是独立同分布的、服从均值为 0 标准差为 100 的正态分布. 跳跃部分 $\boldsymbol{J}_s$ 表示资产在时间 s 时跳跃的大小, 服从均值为 0 标准差为 0.2 的正态分布; 而 $\boldsymbol{N}_s$ 表示资产在时间 s 时是否发生跳跃, 其服从跳跃强度为 $\lambda = 0.02$ 的 Poisson 分布. 虽然这里跳跃的大小的波动较大, 这主要是结合了我国股市涨停和跌停 (股价下降 10% 和上升 10%) 的特点和高频对数价格波动的实际所确定的.

5.2 模拟结果

在模拟研究中, 本文将利用所提出的方法估计一天的积分波动率矩阵, 取 n 为 1,000, p 为 3 或 200, 整个模拟过程重复 200 次. 为了检验本文所提出估计量的估计效果, 将分析比较它们的均方差 (mean squared error, MSE), 这里的积分波动率矩阵估计量的均方差采用估计量与真实积分波动率矩阵作差的 l_2- 范数扩大一百倍所得.

表 1–3 给出了 3 只资产分别在 3 种不同的市场微观结构噪声模型下估计量的均方差比较, 其中 $\widehat{\mathbf{\Gamma}}^{(1)}$、$\widehat{\mathbf{\Gamma}}^{(2)}$ 、$\widehat{\mathbf{\Gamma}}^{(3)}$ 和 $\widehat{\mathbf{\Gamma}}^{(4)}$ 分别代表估计量 AERVM、TAERVM、MinRVM 和 MedRVM, T、H 和 S 分别表示前面估计量的普通门限截断调整、Hard- 门限截断调整和 Soft- 门限截断调整.

从表 1–3 可以看出在带跳价格模型下, 当资产数较少 (即 $p=3$) 时, 总体来看, 估计量 AERVM 的整体均方差最小, 表现效果最好, 尤其在低噪声水平下, 这主要是因为该估计量是通过对参数进行估计直接获得, 没有因技术操作产生额外的误差. 另外 3 种方法是直接从市场的对数价格与微观结构噪声估计之间的差入手, 对跳进行处理, 操作程序较为复杂, 最终导致估计效果略差. 其中估计量 TAERVM

表 1 3 只资产在噪声为 g_1 下积分波动率矩阵估计量 MSE 比较

噪声水平	$\widehat{\mathbf{\Gamma}}^{(3)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(4)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(1)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(2)}$	T	H	S
$\kappa=0.001$	2.21	2.24	2.24	2.24	2.22	2.25	2.25	2.25	1.25	1.12	1.12	1.12	1.29	1.42	1.42	1.35
$\kappa=0.100$	1.53	1.66	1.66	1.66	1.54	1.66	1.66	1.66	1.44	1.46	1.46	1.39	1.68	1.88	1.88	1.88
$\kappa=0.400$	2.17	2.42	2.42	2.42	2.14	2.41	2.41	2.41	0.73	0.72	0.72	0.73	0.53	0.74	0.74	0.74
$\kappa=0.800$	2.07	2.19	2.19	2.19	2.07	2.20	2.20	2.19	0.64	0.65	0.65	0.65	0.60	0.74	0.74	0.74
$\kappa=1.200$	1.34	1.39	1.39	1.39	1.70	1.73	1.73	1.74	0.87	0.75	0.75	0.75	0.75	0.84	0.84	0.84
$\kappa=1.600$	1.64	1.93	1.93	1.93	1.65	1.95	1.95	1.94	1.02	0.67	0.67	0.68	0.89	0.76	0.76	0.76

表 2 3 只资产在噪声为 g_2 下积分波动率矩阵估计量 MSE 比较

噪声水平	$\widehat{\mathbf{\Gamma}}^{(3)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(4)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(1)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(2)}$	T	H	S
$\kappa=0.001$	2.22	2.25	2.25	2.25	2.23	2.26	2.26	2.26	1.25	1.11	1.11	1.11	1.28	1.42	1.42	1.35
$\kappa=0.100$	1.54	1.63	1.63	1.63	1.54	1.63	1.63	1.63	1.43	1.43	1.43	1.37	1.67	1.82	1.82	1.82
$\kappa=0.400$	2.18	2.42	2.42	2.42	2.15	2.41	2.41	2.41	0.73	0.70	0.70	0.70	0.53	0.75	0.75	0.75
$\kappa=0.800$	2.07	2.20	2.20	2.20	2.07	2.20	2.20	2.20	0.64	0.61	0.61	0.61	0.60	0.75	0.75	0.75
$\kappa=1.200$	1.34	1.39	1.39	1.39	1.70	1.72	1.72	1.72	0.87	0.77	0.77	0.77	0.74	0.84	0.84	0.83
$\kappa=1.600$	1.65	1.95	1.95	1.94	1.66	1.96	1.96	1.96	0.99	0.73	0.73	0.73	0.90	0.75	0.75	0.75

表 3 3 只资产在噪声为 g_3 下积分波动率矩阵估计量 MSE 比较

噪声水平	$\widehat{\mathbf{\Gamma}}^{(3)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(4)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(1)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(2)}$	T	H	S
$\kappa=0.001$	2.22	2.26	2.26	2.26	2.23	2.27	2.27	2.27	1.25	1.11	1.11	1.11	1.27	1.32	1.32	1.30
$\kappa=0.100$	1.54	1.66	1.66	1.66	1.54	1.66	1.66	1.66	1.43	1.46	1.46	1.39	1.67	1.82	1.82	1.83
$\kappa=0.400$	2.17	2.37	2.37	2.37	2.14	2.34	2.34	2.34	0.74	0.73	0.73	0.73	0.53	0.73	0.73	0.73
$\kappa=0.800$	2.07	2.18	2.18	2.18	2.08	2.19	2.19	2.19	0.64	0.63	0.63	0.63	0.60	0.74	0.74	0.74
$\kappa=1.200$	1.34	1.40	1.40	1.40	1.67	1.41	1.41	1.41	0.86	0.76	0.76	0.76	0.73	0.79	0.79	0.79
$\kappa=1.600$	1.64	1.94	1.94	1.93	1.65	1.95	1.95	1.95	1.03	0.67	0.67	0.67	0.90	0.76	0.76	0.76

是通过简单粗暴的截断方法处理跳跃, 而估计量 MinRVM 和 MedRVM 则是通过相邻的数据替换跳跃的数据处理跳跃. 结果显示估计量 TAERVM 表现次之, MinRVM 和 MedRVM 略差, 由此可见通过简单粗暴的截断方法处理跳跃更好.

从噪声水平来看, 从表 1–3 容易发现估计量在低噪声水平下普遍估计较差. 例如, 当 $\kappa = 0.001$ (相当于无噪声) 时, 所有估计量的均方差都偏大. 由此可见在无噪声的情形下, 烦琐的降噪运算 (预平均方法) 反而增加了估计量的均方差. 此外, 随着噪声水平的增大, 各估计量的估计效果不稳定. 从门限截断调整方法来看表 1–3, 各种门限截断调整方法都不起作用, 甚至该操作使得估计变得更差. 这主要是因为资产数太少, 通常门限截断调整方法是针对众多资产的估计问题所提出来的. 因此, 下面将展开对更多资产的市场波动率矩阵的估计, 由于资产数的增加, 对波动率矩阵估计的难度增加, 同时也需要考虑波动率矩阵的稀疏结构.

当资产数 $p = 200$ 时, 表 4–6 在 3 种不同的市场微观结构噪声模型和半鞅对数价格模型下分别给出了估计量的均方差比较. 同表 1–3 相比, 随着资产数的增加, 相应估计量的均方差的整体数值都有

表 4　200 只资产在噪声为 g_1 下积分波动率矩阵估计量 MSE 比较

噪声水平	$\widehat{\mathbf{\Gamma}}^{(3)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(4)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(1)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(2)}$	T	H	S
$\kappa = 0.001$	9.02	6.05	6.05	6.31	9.02	6.06	6.06	6.35	8.73	5.49	5.49	5.80	8.34	5.10	5.10	5.30
$\kappa = 0.100$	9.06	5.99	5.99	6.30	9.09	6.02	6.02	6.34	8.78	5.58	5.58	5.85	8.51	4.99	4.99	5.28
$\kappa = 0.400$	9.00	6.17	6.17	6.47	9.00	6.23	6.23	6.49	8.84	5.59	5.59	5.94	8.40	4.99	4.99	5.25
$\kappa = 0.800$	8.91	6.22	6.22	6.48	8.91	6.23	6.23	6.48	8.85	5.82	5.82	6.02	8.29	5.10	5.10	5.29
$\kappa = 1.200$	9.10	6.18	6.18	6.44	9.13	6.21	6.21	6.47	8.94	5.99	5.99	6.15	8.47	4.93	4.93	5.27
$\kappa = 1.600$	9.02	6.15	6.15	6.46	9.04	6.22	6.22	6.52	9.09	5.95	5.95	6.22	8.59	5.10	5.10	5.36

表 5　200 只资产在噪声为 g_2 下积分波动率矩阵估计量 MSE 比较

噪声水平	$\widehat{\mathbf{\Gamma}}^{(3)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(4)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(1)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(2)}$	T	H	S
$\kappa = 0.001$	9.02	6.07	6.07	6.33	9.02	6.07	6.07	6.36	8.74	5.51	5.51	5.79	8.35	5.10	5.10	5.32
$\kappa = 0.100$	9.07	6.02	6.02	6.30	9.09	6.03	6.03	6.35	8.78	5.59	5.59	5.87	8.51	4.98	4.98	5.28
$\kappa = 0.400$	9.00	6.16	6.16	6.47	9.01	6.20	6.20	6.48	8.84	5.60	5.60	5.93	8.41	5.00	5.00	5.28
$\kappa = 0.800$	8.91	6.24	6.24	6.48	8.92	6.24	6.24	6.47	8.84	5.80	5.80	6.01	8.29	5.08	5.08	5.28
$\kappa = 1.200$	9.10	6.19	6.19	6.46	9.13	6.21	6.21	6.47	8.94	6.00	6.00	6.17	8.47	4.95	4.95	5.27
$\kappa = 1.600$	9.02	6.16	6.16	6.46	9.04	6.19	6.19	6.52	9.06	5.92	5.92	6.21	8.59	5.10	5.10	5.36

表 6　200 只资产在噪声为 g_3 下积分波动率矩阵估计量 MSE 比较

噪声水平	$\widehat{\mathbf{\Gamma}}^{(3)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(4)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(1)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(2)}$	T	H	S
$\kappa = 0.001$	9.02	6.05	6.05	6.33	9.02	6.06	6.06	6.36	8.73	5.50	5.50	5.78	8.35	5.09	5.09	5.30
$\kappa = 0.100$	9.07	6.01	6.01	6.29	9.09	6.03	6.03	6.35	8.78	5.58	5.58	5.85	8.52	4.96	4.96	5.27
$\kappa = 0.400$	9.00	6.17	6.17	6.47	9.01	6.22	6.22	6.50	8.84	5.60	5.60	5.93	8.41	4.99	4.99	5.24
$\kappa = 0.800$	8.91	6.22	6.22	6.46	8.92	6.23	6.23	6.49	8.86	5.83	5.83	6.01	8.29	5.10	5.10	5.29
$\kappa = 1.200$	9.10	6.19	6.19	6.45	9.13	6.21	6.21	6.49	8.94	6.00	6.00	6.16	8.47	4.91	4.91	5.27
$\kappa = 1.600$	9.02	6.15	6.15	6.46	9.04	6.23	6.23	6.51	9.09	5.99	5.99	6.23	8.60	5.11	5.11	5.35

所增加, 除此之外有 3 点不同: (1) 从整体估计效果来看, 估计量 TAERVM 的估计效果最佳, 这主要是因为随着资产数的增加, 通过参数估计的微观结构噪声所造成的误差发生了累计, 这相比截断技术所带来的操作误差要大很多. (2) 从噪声水平来看, 当资产数很大时, 噪声水平已不再是影响积分波动率矩阵估计好坏的主导因素, 此时资产数占主导影响因素位置. 由此可见, 资产数越大, 积分波动率估计量对噪声水平越稳健. (3) 从截断效果来看, 容易发现调整后的估计量的均方差变小很多, 因而估计效果更好, 这与调整后估计量是真实积分波动率矩阵的相合估计吻合. 同时我们发现表 4–6 中普通门限截断调整与 Hard- 门限截断调整的结果基本一致, 这可能是文献 [8] 中的参数设置已经保证了通过普通门限截断调整后估计依然具有半正定性所致. 此外, Soft- 门限截断调整的效果比普通门限截断调整与 Hard- 门限截断调整的效果略差, 这是因为 Soft- 门限截断调整不仅起到降维的效果, 还对系数进行了调整, 进而产生了额外误差.

由于表 1–6 模拟重复 200 次, 调整参数 ξ、ω 和 K 每次取值不尽相同, 但遵循的原则如下:

$$(\widehat{\xi}, \widehat{K}^{(l)}) = \arg \min_{0<\xi<1/2, 1\leqslant K\leqslant 50} \|\widehat{\mathbf{\Gamma}}^{(l)} - \mathbf{\Gamma}\|_2,$$
$$\widehat{\omega}^{(l)} = \arg \min_{0\leqslant\omega\leqslant 1} \|\widehat{\mathbf{\Gamma}}^{(l)}_{\text{Adjusted}} - \mathbf{\Gamma}\|_2, \quad l = 1, 2, 3, 4.$$

由于 ξ 的选取仅涉及估计量 $\widehat{\mathbf{\Gamma}}^{(2)}$, 为避免重复赘述, 这里仅给出 $p = 200$ 时估计量 $\widehat{\mathbf{\Gamma}}^{(2)}$ 在 g_1 下的参数选取情形, 如表 7, 其中 $\overline{\xi}$、$\overline{\omega}$ 和 $\overline{K}$ 是模拟 200 次的平均取值, 括号里面的数值是调整参数的标准差 (standard error, SE).

表 7 表明在 200 次重复实验中, 估计量 TAERVM 在不同噪声水平下各个参数平均最优值都比较稳健. 除了 Soft- 门限调整的门限参数 $\overline{\omega}_S$ 对噪声水平非常稳健外, 对于低噪声水平和高噪声水平, 各参数的平均估计量的稳定性相对较差, 这主要是由于估计方法本身所造成的系统性误差. 此外, 由于真实的矩阵是通过主对角线绝对占优模拟所得, 因此利用普通门限调整与 Hard- 门限调整效果一致. 同时不难发现 K 的最优值为 2, 虽然 $\overline{K}$ 相对其他参数的 SE 较大, 但是其对噪声水平也比较稳健, 这是因为 K 的取值为非负整数. 事实上对于不同的估计量, 除了 AERVM 的最优值为 $K = 2$、MinRVM 和 MedRVM 的最优值为 $K = 4$ 外, 其他参数的选取方法及稳定性性质基本同表 7, 因此不再重复展示.

为进一步考察估计量, 接下来从估计量的特征值入手展开研究. 判定高维积分波动率矩阵估计量的好坏, 除了检验估计量的均方差, 还有一个很重要的衡量标准就是矩阵的特征值. 因为特征值能反映出起主导作用的因子, 从而能够反映估计的好坏. 在此仅给出 200 只资产在带有市场信息的微观结构噪声 g_1 下估计量调整前后的特征值情形, 其他两种带有市场信息的微观结构噪声情形下估计量的表现基本一致. 图 2–5 分别给出噪声水平 κ 为 0.001、0.4、1.0 和 1.6 时估计量的特征值曲线. 由于 200

表 7 200 只资产在噪声为 g_1 下 $\widehat{\mathbf{\Gamma}}^{(2)}$ 的参数选取

噪声水平	$\overline{\xi}$	$\overline{K}$	$\overline{\omega}_T$	$\overline{\omega}_H$	$\overline{\omega}_S$
$\kappa = 0.001$	0.035 (0.047)	2.020 (0.264)	0.291 (0.121)	0.291 (0.121)	0.031 (0.003)
$\kappa = 0.100$	0.042 (0.031)	1.995 (0.187)	0.401 (0.003)	0.401 (0.003)	0.040 (0.000)
$\kappa = 0.400$	0.043 (0.022)	2.000 (0.142)	0.360 (0.000)	0.360 (0.000)	0.040 (0.000)
$\kappa = 0.800$	0.059 (0.034)	2.010 (0.200)	0.430 (0.000)	0.430 (0.000)	0.040 (0.000)
$\kappa = 1.200$	0.084 (0.051)	1.995 (0.213)	0.322 (0.004)	0.322 (0.004)	0.036 (0.004)
$\kappa = 1.600$	0.076 (0.081)	2.035 (0.353)	0.203 (0.045)	0.203 (0.045)	0.040 (0.000)

只资产的特征值有 200 个, 考虑到起主导作用的结构特征通常排列前几位, 因此图 2–5 给出的是积分波动率矩阵估计量的前 20 个特征值的比较.

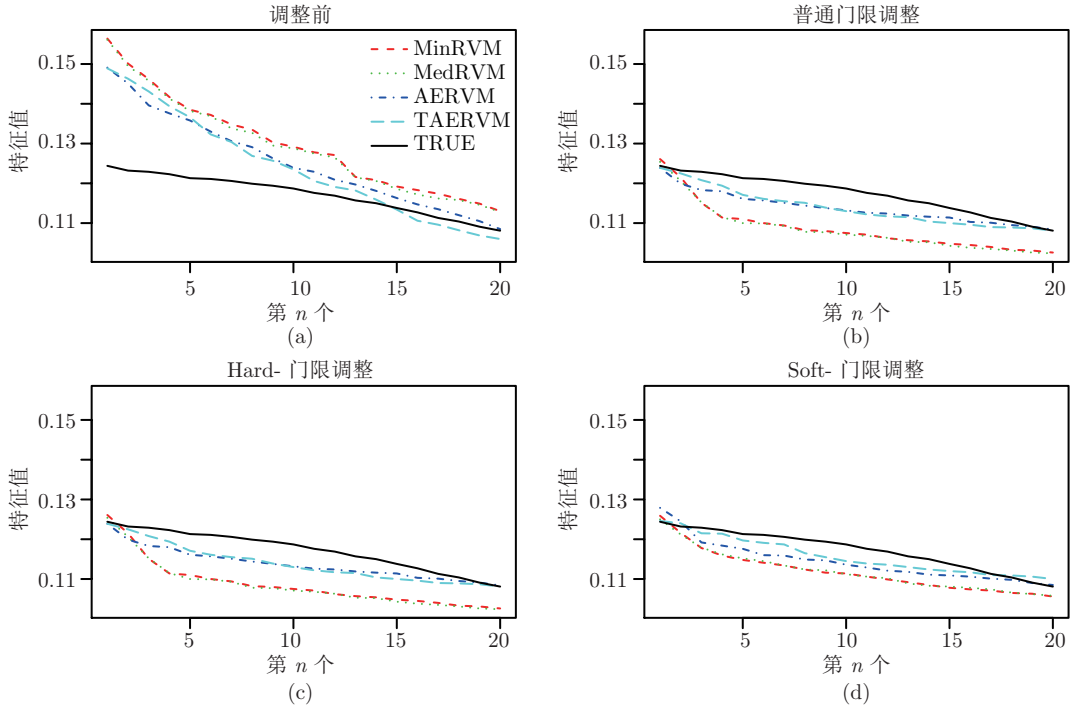

图 2 (网络版彩图) 在 g_1 模型和噪声水平 $\kappa = 0.001$ 下估计量的特征值比较

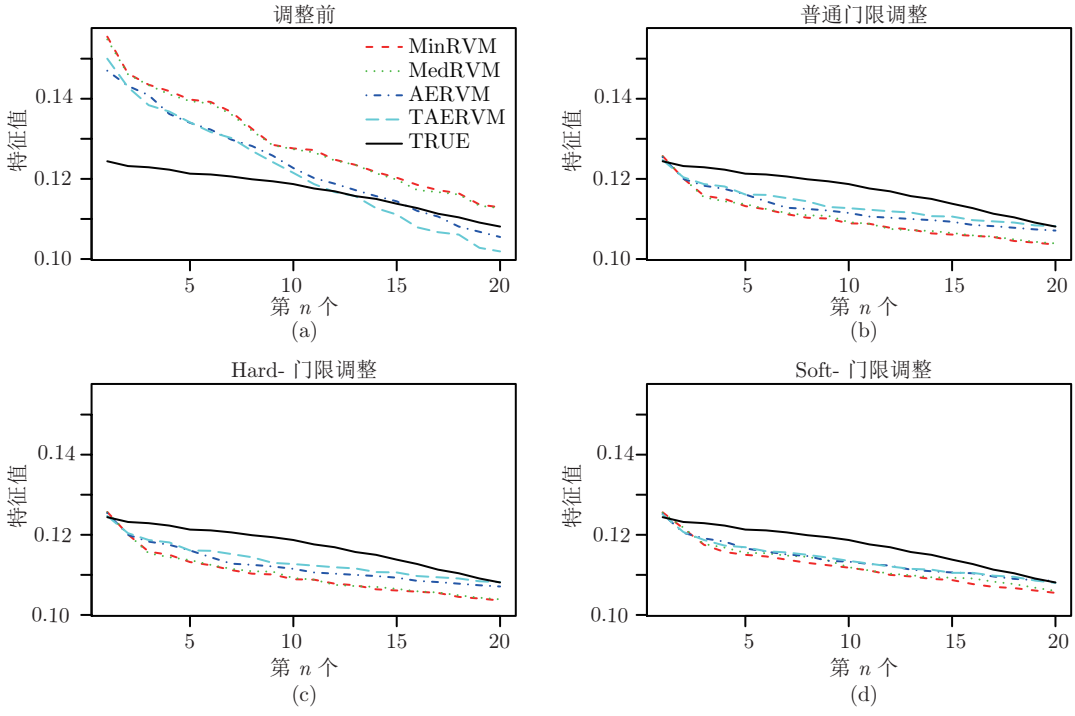

图 3 (网络版彩图) 在 g_1 模型和噪声水平 $\kappa = 0.4$ 下估计量的特征值比较

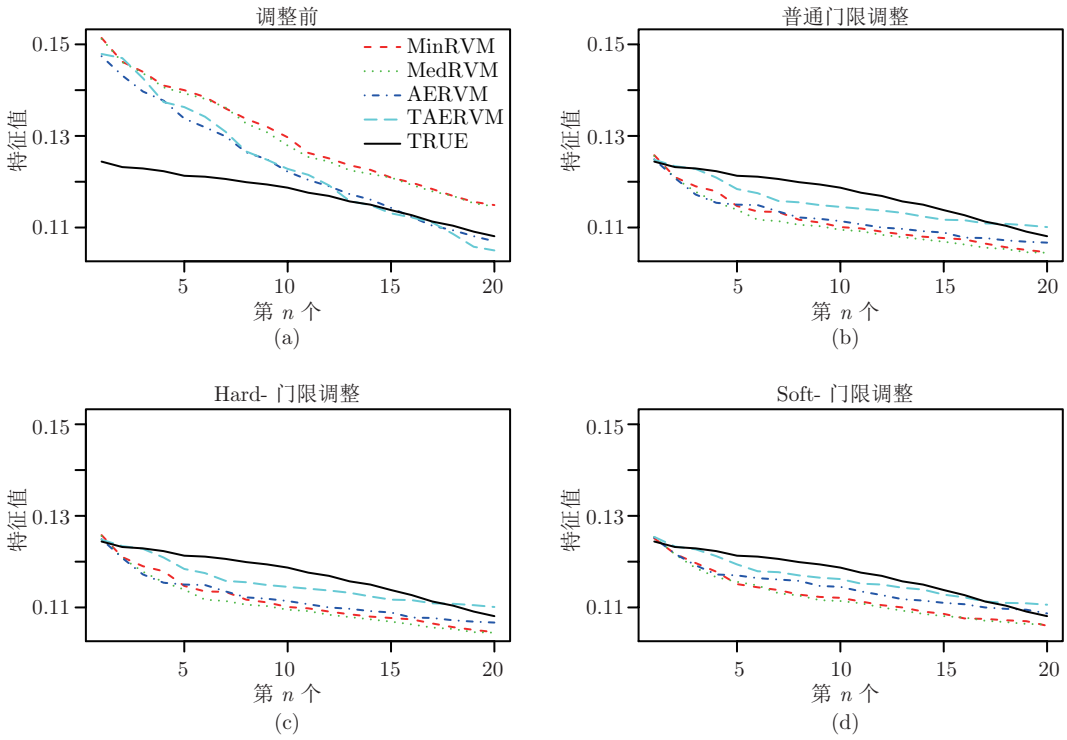

图 4 (网络版彩图) 在 g_1 模型和噪声水平 $\kappa = 1.0$ 下估计量的特征值比较

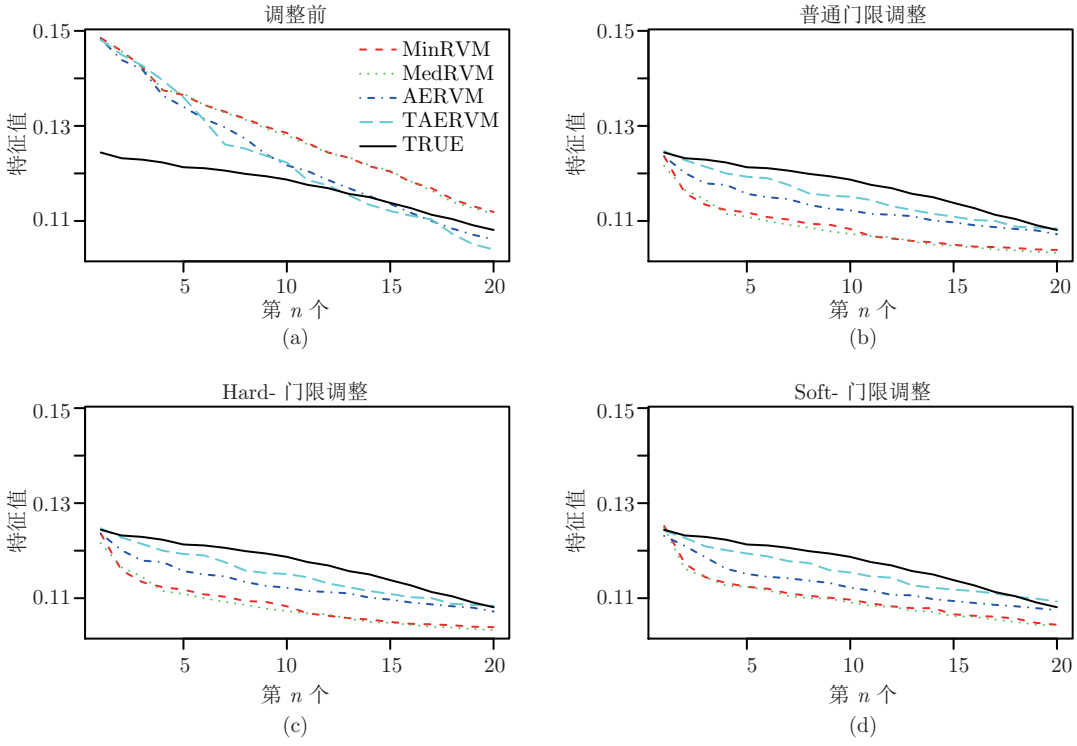

图 5 (网络版彩图) 在 g_1 模型和噪声水平 $\kappa = 1.6$ 下估计量的特征值比较

在 4 种不同的噪声水平下, 图 2–5 中的特征值曲线反应几乎一致, 这表明在带跳的价格模型下, 噪声的影响远小于跳跃对积分波动率矩阵的影响. 在图 2–5 中, 第一张子图是没有对估计量进行门限截断调整的情形, 显然估计量起主导作用的前几项都远高于真实的积分波动率矩阵的特征值曲线, 由此可见调整前估计量的表现都很差, 这与表 4–6 中估计量均方差所获得的结论一致, 同时也符合定理的结论—调整前估计量不相合. 剩下的 3 张子图分别是调整后估计量的特征值曲线, 它们向真实的积分波动率矩阵的特征值曲线靠近了许多. 其中简单粗暴的 TAERVM 的特征值曲线离真实的特征值曲线比较接近, 估计效果最好, 尤其是起主导作用的第一个特征值; 利用微观结构噪声直接获得的估计量 AERVM 略差于 TAERVM, 但优于估计量 MinRVM 和 MedRVM, 而利用邻近处理技术的估计量 MinRVM 和 MedRVM 的特征值曲线估计效果不分伯仲. 显然这些结论与上述表格反映的结论一致.

从具体的调整方法来看, 理论上可以通过观察估计量特征值是否存在负数来判断估计的好坏, 但是由于本节模拟的真实的积分波动率矩阵是严格的主对角占优矩阵, 即使经过普通门限截断也保证了其半正定性, 因此, 图 2–5 显示普通门限调整与 Hard- 门限调整几乎没有差异, 然而在实证研究中, 该方法的缺陷就会暴露出来 (见表 8 和 9). 虽然表 1–6 显示 Soft- 门限调整后的估计量的均方差相比另外两种调整方法的均方差略大, 但是其特征值曲线离真实的更近, 估计效果更好, 同时也说明并非均方差越小, 估计效果越好. 此外, 调整后的估计量起主导作用的特征值离真实的比较接近, 但同时因为截断调整也丢失了很多有用的信息进而低估真实积分波动率矩阵, 导致调整后的估计量微观结构作用弱化, 不像真实的特征值曲线平缓下降. 这无疑对资产的调整方法提出新的挑战.

表 8　2013 年 1 月 4 日 20 只上证指数不同估计量的特征值比较

	$\widehat{\mathbf{\Gamma}}^{(3)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(4)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(1)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(2)}$	T	H	S
1	14.61	11.99	12.79	13.01	14.61	11.99	12.79	13.01	14.61	11.99	13.51	13.02	14.18	11.17	13.16	12.50
2	2.92	3.56	3.08	3.06	2.92	3.56	3.07	3.05	2.92	3.56	3.02	3.06	2.97	3.70	2.97	3.12
3	1.90	2.94	1.98	1.97	1.90	2.94	1.98	1.97	1.91	2.94	2.43	1.97	2.45	3.12	2.45	2.54
4	1.54	2.76	1.68	1.66	1.54	2.76	1.68	1.66	1.54	2.76	1.76	1.66	1.72	3.07	1.78	1.81
5	1.38	2.38	1.50	1.49	1.38	2.38	1.50	1.49	1.38	2.38	1.63	1.49	1.61	2.09	1.61	1.75
6	1.00	1.80	1.11	1.10	1.00	1.80	1.11	1.10	1.00	1.80	1.27	1.10	1.19	1.70	1.19	1.31
7	0.84	1.26	0.88	0.87	0.84	1.26	0.88	0.87	0.84	1.26	1.00	0.87	1.02	1.55	1.04	1.09
8	0.75	1.03	0.84	0.83	0.75	1.03	0.84	0.83	0.75	1.03	0.91	0.83	0.85	1.06	0.86	0.95
9	0.71	0.94	0.78	0.78	0.71	0.94	0.78	0.78	0.71	0.94	0.79	0.78	0.78	1.02	0.78	0.86
10	0.60	0.70	0.63	0.63	0.59	0.70	0.63	0.63	0.60	0.70	0.62	0.63	0.60	0.96	0.60	0.66
11	0.46	0.49	0.57	0.55	0.46	0.49	0.57	0.55	0.46	0.49	0.55	0.55	0.52	0.69	0.52	0.61
12	0.40	0.33	0.51	0.50	0.40	0.33	0.51	0.50	0.40	0.33	0.51	0.50	0.46	0.69	0.48	0.57
13	0.28	0.26	0.37	0.36	0.28	0.26	0.37	0.36	0.28	0.26	0.48	0.36	0.46	0.49	0.46	0.51
14	0.26	0.13	0.35	0.34	0.26	0.13	0.35	0.34	0.26	0.13	0.43	0.34	0.35	0.23	0.32	0.45
15	0.20	−0.01	0.30	0.29	0.20	−0.01	0.30	0.29	0.20	−0.01	0.33	0.29	0.30	0.17	0.30	0.37
16	0.17	−0.17	0.26	0.25	0.17	−0.17	0.26	0.25	0.17	−0.17	0.27	0.25	0.25	0.11	0.24	0.33
17	0.11	−0.22	0.22	0.21	0.11	−0.22	0.22	0.21	0.11	−0.22	0.18	0.21	0.15	−0.12	0.12	0.27
18	0.03	−0.41	0.15	0.14	0.03	−0.41	0.15	0.14	0.03	−0.41	0.15	0.14	0.06	−0.23	0.04	0.16
19	0.03	−0.67	0.10	0.09	0.03	−0.67	0.10	0.09	0.03	−0.67	0.09	0.09	0.03	−0.48	0.03	0.10
20	0.02	−0.89	0.07	0.06	0.02	−0.89	0.07	0.06	0.02	−0.89	0.05	0.06	0.03	−0.99	0.03	0.08

表 9 2013 年 1 月 8 日 20 只上证指数不同估计量的特征值比较

	$\widehat{\mathbf{\Gamma}}^{(3)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(4)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(1)}$	T	H	S	$\widehat{\mathbf{\Gamma}}^{(2)}$	T	H	S
1	11.54	9.56	10.11	10.28	11.53	9.55	10.10	10.27	11.54	9.56	10.52	10.28	11.04	8.63	10.98	9.75
2	2.09	2.79	2.20	2.18	2.09	2.79	2.20	2.18	2.09	2.79	2.09	2.19	2.15	2.67	2.13	2.22
3	1.52	2.46	1.60	1.59	1.52	2.46	1.60	1.59	1.52	2.46	1.61	1.59	1.74	2.61	1.88	1.82
4	1.28	1.67	1.37	1.36	1.28	1.67	1.37	1.36	1.28	1.67	1.26	1.36	1.37	2.28	1.52	1.48
5	1.04	1.55	1.13	1.11	1.04	1.55	1.13	1.11	1.04	1.55	1.12	1.12	1.26	1.77	1.22	1.34
6	0.96	1.48	1.07	1.06	0.96	1.48	1.07	1.06	0.96	1.48	0.94	1.06	1.16	1.57	1.13	1.27
7	0.80	1.25	0.86	0.85	0.80	1.25	0.86	0.85	0.80	1.25	0.79	0.85	0.95	1.47	1.02	1.02
8	0.68	1.08	0.75	0.74	0.68	1.08	0.75	0.74	0.68	1.08	0.68	0.74	0.83	1.21	0.82	0.89
9	0.55	0.79	0.61	0.60	0.55	0.79	0.61	0.60	0.56	0.79	0.54	0.60	0.68	1.08	0.68	0.75
10	0.51	0.65	0.54	0.53	0.51	0.65	0.54	0.53	0.51	0.65	0.50	0.53	0.58	0.79	0.57	0.68
11	0.48	0.56	0.53	0.52	0.48	0.56	0.53	0.52	0.48	0.57	0.49	0.52	0.52	0.64	0.52	0.53
12	0.39	0.56	0.50	0.49	0.39	0.56	0.50	0.49	0.39	0.56	0.39	0.49	0.48	0.56	0.40	0.53
13	0.34	0.31	0.42	0.41	0.34	0.31	0.41	0.41	0.34	0.31	0.30	0.41	0.39	0.42	0.36	0.45
14	0.28	0.22	0.37	0.36	0.28	0.22	0.37	0.36	0.29	0.22	0.27	0.36	0.34	0.23	0.30	0.42
15	0.22	0.01	0.28	0.27	0.22	0.01	0.28	0.27	0.22	0.01	0.23	0.27	0.25	0.10	0.24	0.30
16	0.20	−0.01	0.27	0.26	0.20	−0.01	0.27	0.26	0.20	−0.01	0.19	0.26	0.23	0.03	0.23	0.29
17	0.09	−0.13	0.19	0.18	0.09	−0.12	0.19	0.18	0.09	−0.13	0.09	0.18	0.12	−0.11	0.11	0.21
18	0.08	−0.34	0.14	0.13	0.08	−0.34	0.14	0.13	0.08	−0.34	0.08	0.13	0.10	−0.29	0.10	0.14
19	0.02	−0.62	0.12	0.11	0.02	−0.62	0.12	0.11	0.02	−0.62	0.04	0.11	0.04	−0.51	0.04	0.12
20	0.02	−0.77	0.06	0.05	0.02	−0.77	0.06	0.05	0.02	−0.77	0.01	0.05	0.03	−0.90	0.01	0.07

6 实证研究

为了考察高维积分波动率矩阵估计量的行为, 本文利用 20 只上海证券指数 1) (包含上证指数) 在 2013 年 1 月 4 日和 1 月 8 日两天的以 5 秒记录的交易数据研究各种积分波动率矩阵估计量的特征值. 本节将在微观结构噪声 g_1 (g_2 和 g_3 类似) 下比较本文所提出的估计量 AERVM、TAERVM、MinRVM 和 MedRVM 的特征值.

由于上证指数数据记录的非同步性, 因此采用前值法 [22] 进行同步化. 异步性的存在性导致两天的 20 只资产的总体样本量也是有差异的, 分别是 55,660 和 56,100; 为保证所有估计量两天起主导作用的因子处于一个量级, 同时考虑到模拟中估计量 AERVM 和 TAERVM 在 $K=2$ 时取得最小均方差, MinRVM 和 MedRVM 在 $K=4$ 时取得最小均方差, 于是在此取相同的值. 为了保证稀疏后起主导作用的因子至少能达到总体的 25%, 两天的门限值 ϖ 分别定为 0.48 和 0.45, 具体见表 8 和 9.

由于在实证中没有真实值做参考, 直接根据表的数值大小无法判断估计量的好坏, 但是我们知道真实的积分波动率矩阵是一个半正定矩阵, 基于这一点结合表 8 和 9 发现本文所引用的 Hard- 门限截断调整和 Soft- 门限截断调整后估计量的所有特征值都是正的, 保证了原有矩阵的半正定性, 而普通门限截断调整后的估计量的特征值中存在负数, 无法保证估计量的半正定性, 这与前面定理的理论

1) 20 只上海证券指数: 上证材料、上证地企、上证工业、上证公用、上证海外、上证沪企、上证金融、上证可选、上证全指、上证龙头、上证民企、上证能源、上证消费、上证新兴、上证商品、上证信息、上证央企、上证国企、上证周期、上证指数.

结果一致.

此外, 表 8 和 9 利用实际数据获得波动率矩阵估计的最大特征值与第二大特征值的相差较大, 而模拟中不存在这个问题. 这主要是因为在模拟中, 资产的对数价格的模型设置是随机的, 没有个性特征, 从而 p 个资产的地位是等价的, 积分波动率矩阵的特征值大小差异不明显. 然而现实中的资产不仅具有自身的特性, 而且与其他资产可能存在某种正向或反向的内在联系, 通过计算积分波动率矩阵就会凸显出那些在市场上占主导地位资产的作用. 众所周知, 上证指数包含上海各个行业的指数信息, 起绝对主导性作用, 因此表 8 和 9 中所有估计量的最大特征值与其他特征值会出现如此悬殊的特征.

由于实际市场中不同资产可能受到一个共同因素来驱动变化 (见图 6 和 7), 大幅的波动几乎是同时进行的, 由此可见引起跳跃的原因来自同一类因素. 剔去跳跃波动的部分, 剩余连续部分的波动率矩阵是否具有稀疏性呢? 为此分别考察了 2013 年 1 月 4 日和 8 日两天的 20 只指数的积分波动率矩阵的稀疏程度与解释程度的关系, 如图 8 所示. 为方便论述, 此处采用截断方法, 截掉 20 只指数的价格变化幅度大的部分, 其中临界值 $C = 2.3 \times 10^{-4}$. 图 8 显示, 随着稀疏程度的降低 (矩阵中非零元素的增加), 矩阵的可解释程度也越来越高. 当矩阵的稀疏性达到 50% 时, 矩阵可解释性高达 70% 以上, 因此我们认为稀疏性的假设具有一定的合理性.

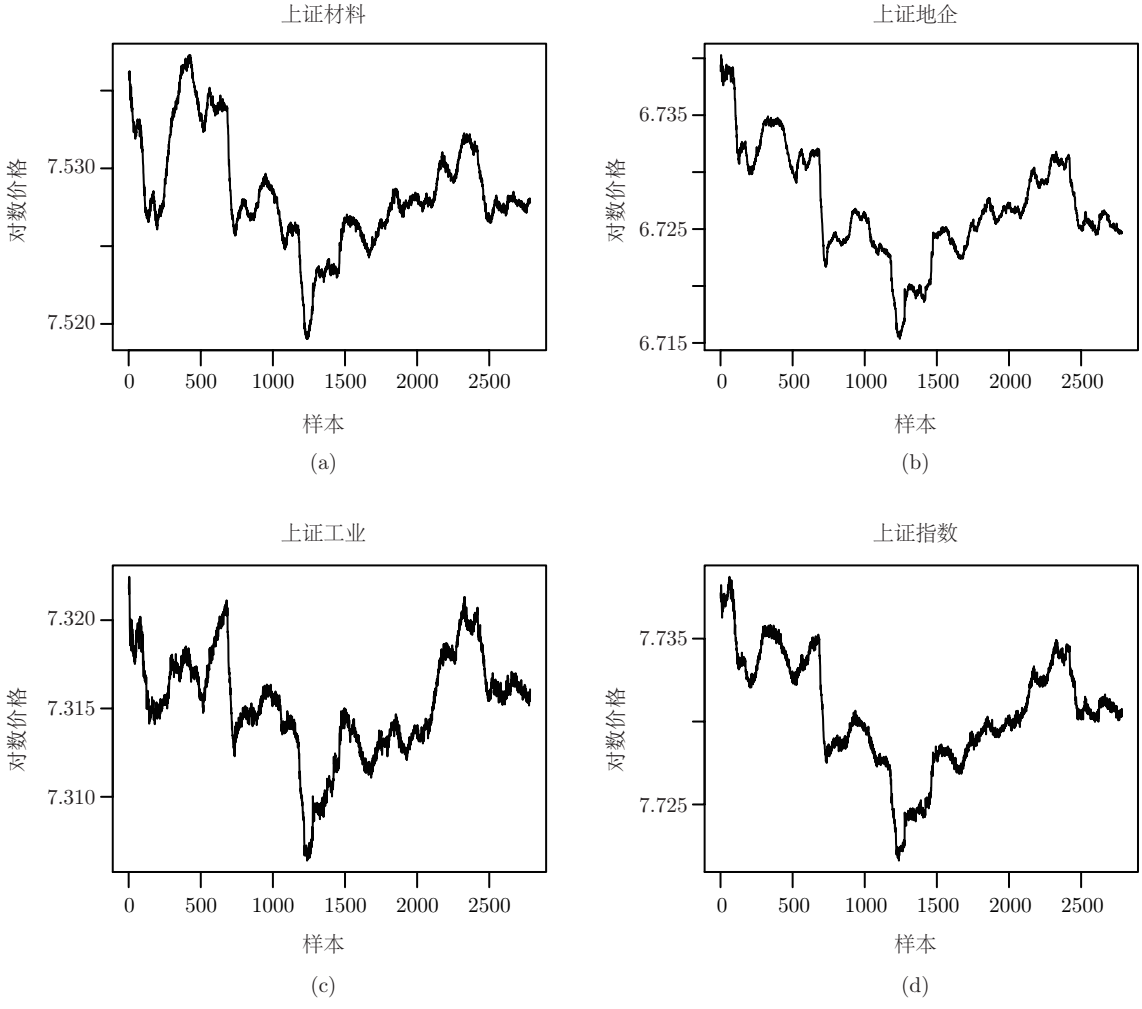

图 6　2013 年 1 月 4 日 4 只指数的数据序列图

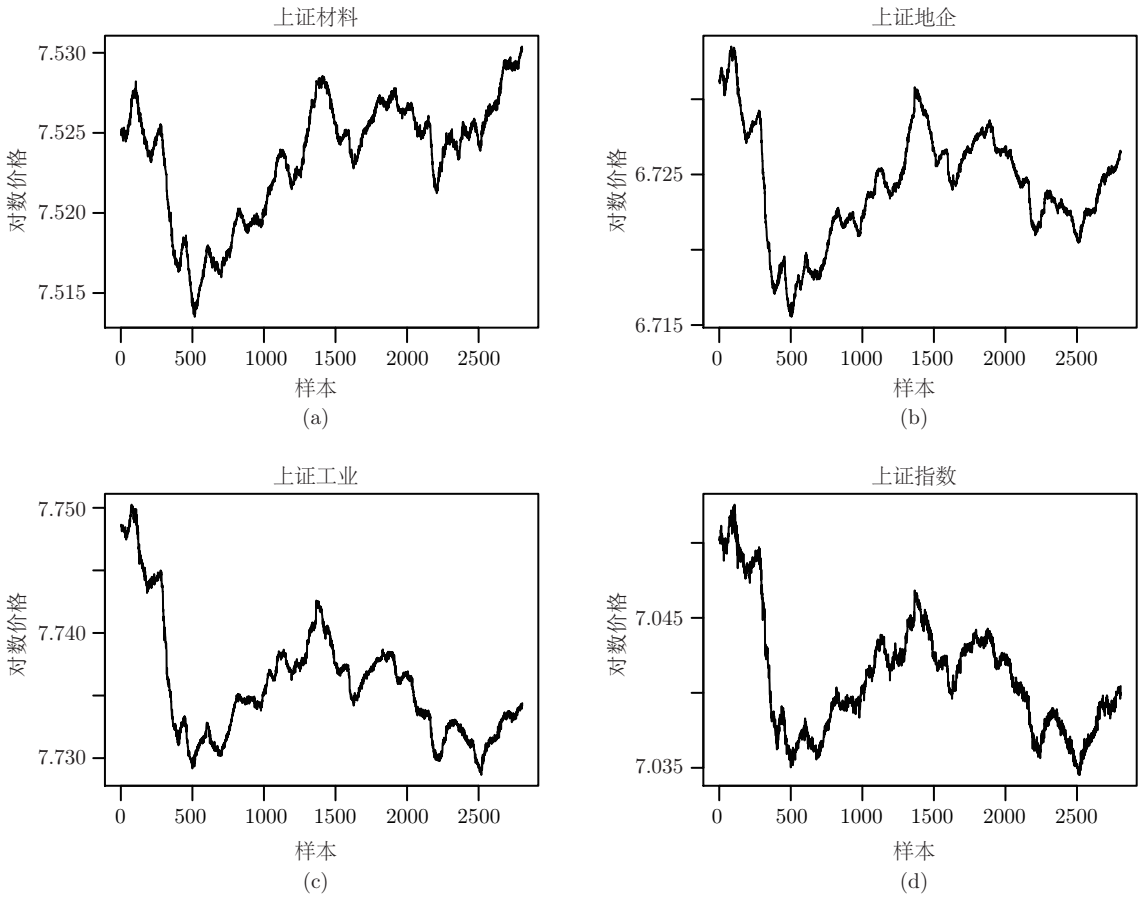

图 7　2013 年 1 月 8 日 4 只指数的数据序列图

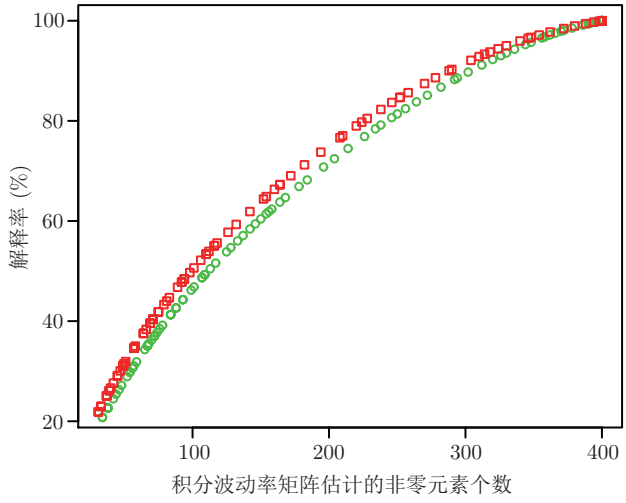

图 8　(网络版彩图) 2013 年 1 月 4 日和 1 月 8 日积分波动率矩阵稀疏程度与解释率的关系. 菱形: 2013 年 1 月 4 日 20 只指数的积分波动率矩阵的稀疏程度与解释率的关系; 圆形: 2013 年 1 月 8 日 20 只指数的积分波动率矩阵的稀疏程度与解释率的关系. 这里的解释率表示 (调整后的估计的所有元素求和)/(调整前的积分波动率估计的所有元素求和)

7 结论

本文充分利用市场交易信息, 考虑带有交易信息的市场微观结构噪声模型, 在单资产跳跃的基础上提出了 4 种高维波动率矩阵估计方法; 从理论上证明了在一定的假设下, 我们的高维积分波动率矩阵估计量是相合的, 收敛速度达到样本量 $n^{1/3}$, 优于目前已有的高维积分波动率矩阵估计量. 从模拟中发现本文提出的 4 种方法对噪声水平非常稳健, 其中简单粗暴的截断方法处理跳跃效果最佳. 实证研究表明引入 Hard- 门限调整和 Soft- 门限调整保证了原有估计量的半正定性.

事实上, 本文研究发现在带跳的价格模型时, 3 种门限截断调整因为截断弱化了微观结构的作用, 进而导致估计量皆低估了真实积分波动率矩阵, 这将迫使我们探讨新的调整方法. 此外, 虽然因子模型假设波动率矩阵是常数, 但是其不需要假设真实波动率矩阵的稀疏性, 该模型可以将跳跃看成一种因子进行处理, 感兴趣的读者可以深入探究.

致谢 作者非常感谢审稿专家提出的宝贵意见和建议.

参考文献

1 Fan J Q, Fan Y Y, Lv J. High dimensional covariance matrix estimation using a factor model. J Econometrics, 2008, 147: 186–197
2 Fan J Q, Liao Y, Mincheva M. High dimensional covariance matrix estimation in approximate factor models. Ann Statist, 2011, 39: 3320–3356
3 Fan J Q, Liao Y, Mincheva M. Large covariance estimation by thresholding principal orthogonal complements. J R Stat Soc Ser B Stat Methodol, 2013, 75: 603–680
4 Fan J Q, Furger A, Xiu D. Incorporating global industrial classification standard into portfolio allocation: A simple factor-based large covariance matrix estimator with high-frequency data. J Bus Econom Statist, 2016, 34: 489–503
5 Aït-Sahalia Y, Xiu D. Using principal component analysis to estimate a high dimensional factor model with high-frequency data. J Econometrics, 2017, 201: 384–399
6 Fan J Q, Wang W C, Zhong Y Q. Robust covariance estimation for approximate factor models. J Econometrics, 2019, 208: 5–22
7 Dai C X, Lu K, Xiu D C. Knowing factors or factor loadings, or neither? Evaluating estimators of large covariance matrices with noisy and asynchronous data. Chicago Booth Research Paper No. 17-02, 2017
8 Wang Y, Zou J. Vast volatility matrix estimation for high-frequency financial data. Ann Statist, 2010, 38: 943–978
9 Tao M, Wang Y, Chen X. Fast convergence rates in estimating large volatility matrices using high-frequency financial data. Econom Theory, 2013, 29: 838–856
10 Kim D, Wang Y Z, Zou J. Asymptotic theory for large volatility matrix estimation based on high-frequency financial data. Stochastic Process Appl, 2016, 126: 3527–3577
11 穆燕, 苑慧玲, 周勇. 高频金融数据的高维积分波动率矩阵估计. 中国科学: 数学, 2018, 48: 319–344
12 Li Y Y, Xie S Y, Zheng X. Efficient estimation of integrated volatility incorporating trading information. J Econometrics, 2016, 195: 33–50
13 Fan J Q, Wang Y Z. Multi-scale jump and volatility analysis for high-frequency financial data. J Amer Statist Assoc, 2007, 102: 1349–1362
14 Andersen T G, Dobrev D, Schaumburg E. Jump-robust volatility estimation using nearest neighbor truncation. J Econometrics, 2012, 169: 75–93
15 唐勇, 林欣. 考虑共同跳跃的波动建模: 基于高频数据视角. 中国管理科学, 2015, 23: 46–53
16 Kong X B, Liu Z, Jing B Y. Testing for pure-jump processes for high-frequency data. Ann Statist, 2015, 43: 847–877
17 吴奔, 张波. 交易信息、跳跃发现与波动估计. 统计研究, 2017, 34: 109–119
18 Hansen P R, Lunde A. Realized variance and market microstructure noise. J Bus Econom Statist, 2006, 24: 127–161
19 Kalnina I, Linton O. Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error. J Econometrics, 2008, 147: 47–59
20 Mancino M E, Sanfelici S. Robustness of Fourier estimator of integrated volatility in the presence of microstructure noise. Comput Statist Data Anal, 2008, 52: 2966–2989
21 Aït-Sahalia Y, Jacod J. Estimating the degree of activity of jumps in high frequency data. Ann Statist, 2009, 37: 2202–2244

22 Zhang L. Estimating covariation: Epps effect, microstructure noise. J Econometrics, 2011, 160: 33–47
23 Bickel P J, Levina E. Regularized estimation of large covariance matrices. Ann Statist, 2008, 36: 199–227
24 Bickel P J, Levina E. Covariance regularization by thresholding. Ann Statist, 2008, 36: 2577–2604
25 Asai M, McAleer M. The impact of jumps and leverage in forecasting covolatility. Econometric Rev, 2017, 36: 638–650
26 Almgren R, Chriss N. Optimal execution of portfolio transactions. J Risk, 2000, 3: 5–39
27 Keim D B, Madhavan A. The upstairs market for large-block transactions: Analysis and measurement of price effects. Rev Financ Stud, 1996, 9: 1–36

附录 A 定理的证明

为了建立估计量的渐近性质, 作如下假设.

假设 A.1 对于 $i=1,\ldots,p$, 有

(i) 漂移向量 $\boldsymbol{\mu}_t$ 和点波动率矩阵 $\boldsymbol{\sigma}_t$ 中的每个元素都是局部有界的, 且对角线元素满足

$$\inf_{t\in[0,1]}\sigma_{ii}(t)>0,\quad \text{a.s.};$$

(ii) 对任意的 $\tau_{i,r}\in[0,1]$, $r=1,\ldots,m$, 在给定 $\mathcal{F}_{\tau_{i,r}}$ 下, $\boldsymbol{Z}_{\tau_{i,r}}$ 和 $\Delta X_{\tau_{i,r}}$ 是相互独立的;

(iii) $\max_{i,r}|\boldsymbol{Z}_{\tau_{i,r}}|=O_p(1)$, 其中 $|\boldsymbol{Z}|$ 表示向量 $\boldsymbol{Z}$ 的 Euclid 距离;

(iv) 对于 $q\in\mathbb{N}$, $\boldsymbol{\theta}_i$ 的参数空间 Θ 在 $\mathbb{R}^q$ 上是紧集, 且 $\boldsymbol{\theta}_i$ 的邻域 $\mathcal{N}(\boldsymbol{\theta}_{i0})\subset\Theta$, $g(\boldsymbol{z}_i;\boldsymbol{\theta}_i)$ 关于 $\boldsymbol{\theta}_i$ 是二阶连续可微的;

(v) 对所有的 $\boldsymbol{\theta}_{i1},\boldsymbol{\theta}_{i2}\in\mathcal{N}(\boldsymbol{\theta}_{i0})$, $|g(\boldsymbol{z}_i;\boldsymbol{\theta}_{i1})-g(\boldsymbol{z}_i;\boldsymbol{\theta}_{i2})|\leqslant L_0(\boldsymbol{z}_i)|\boldsymbol{\theta}_{i1}-\boldsymbol{\theta}_{i2}|$, 其中 $L_0(\boldsymbol{z}_i)$ 是局部有界的;

(vi) 对所有的 $\boldsymbol{\theta}_i\in\mathcal{N}(\boldsymbol{\theta}_{i0})$, $|\frac{\partial g}{\partial\boldsymbol{\theta}_i}(\boldsymbol{z}_i;\boldsymbol{\theta}_i)|\leqslant f(\boldsymbol{z}_i)$, 其中 $f(\boldsymbol{z}_i)$ 是局部有界的;

(vii) $|\frac{\partial g}{\partial\boldsymbol{\theta}_i}(\boldsymbol{z}_i;\boldsymbol{\theta}_{i1})-\frac{\partial g}{\partial\boldsymbol{\theta}_i}(\boldsymbol{z}_i;\boldsymbol{\theta}_{i2})|\leqslant L_1(\boldsymbol{z}_i)|\boldsymbol{\theta}_{i1}-\boldsymbol{\theta}_{i2}|$, 其中 $L_1(\boldsymbol{z}_i)$ 是局部有界的;

(viii) 对任意的 $\varepsilon>0$, 当 $n\to\infty$ 时, 有

$$\inf_{|\boldsymbol{\theta}_i-\boldsymbol{\theta}_{i0}|\geqslant\varepsilon}\sum_{r=1}^{m}|\Delta g(\boldsymbol{Z}_i(\tau_{i,r});\boldsymbol{\theta}_i)-\Delta g(\boldsymbol{Z}_i(\tau_{i,r});\boldsymbol{\theta}_{i0})|^2\to\infty,\quad \text{a.s.};$$

(ix) $\|[\frac{1}{m}\sum_{r=1}^{m}\Delta\frac{\partial g}{\partial\boldsymbol{\theta}_i}(\boldsymbol{Z}_i(\tau_{i,r});\boldsymbol{\theta}_{i0})\Delta\frac{\partial g}{\partial\boldsymbol{\theta}_i}(\boldsymbol{Z}_i(\tau_{i,r});\boldsymbol{\theta}_{i0})]^{-1}\|_2=O_p(1)$, 其中对任意的矩阵 A, $\|A\|$ 表示 l_2- 范数;

(x) 对任意的 $k,l=1,\ldots,q$, 且 $\boldsymbol{\theta}_{i1},\boldsymbol{\theta}_{i2}\in\mathcal{N}(\boldsymbol{\theta}_{i0})$, 有

$$\left|\frac{\partial^2 g}{\partial\theta_{ik}\partial\theta_{il}}(\boldsymbol{z}_i;\boldsymbol{\theta}_{i1})-\frac{\partial g}{\partial\theta_{ik}\partial\theta_{il}}(\boldsymbol{z}_i;\boldsymbol{\theta}_{i2})\right|\leqslant L_2(\boldsymbol{z}_i)|\boldsymbol{\theta}_{i1}-\boldsymbol{\theta}_{i2}|,$$

其中 $L_2(\boldsymbol{z}_i)$ 是局部有界的;

(xi) 对任意的 $k,l=1,\ldots,q$, $\boldsymbol{\theta}_i\in\mathcal{N}(\boldsymbol{\theta}_{i0})$, 有 $|\frac{\partial^2 g}{\partial\theta_{ik}\partial\theta_{il}}(\boldsymbol{z}_i;\boldsymbol{\theta}_i)|\leqslant L_3(\boldsymbol{z}_i)$, 其中 $L_3(\boldsymbol{z}_i)$ 是局部有界的.

假设 (ii) 类似于通常的回归假设, 在模型 (2.1) 下, 相当于假设资产的下一次瞬时交易仅依赖最新的交易市场信息. 当高频交易发生时, 尤其在算法交易的过程中, 资产的买卖与否以及交易量完全由过去交易决定, 因此该假设是合理的. 在极大似然估计 (MLE) 文献中, (iv)–(xi) 是常规假设; (viii) 类似于 MLE 的可识别假设条件; (ix) 与 Fisher 信息矩阵的可逆条件一致. 以上的假设主要是为了保证极大似然估计量 $\widehat{\boldsymbol{\theta}}_i$ 收敛到 $\boldsymbol{\theta}_i$, 具体参见文献 [12, 定理 4.1].

假设 A.2 对于 $\beta\geqslant 2$, 模型 (2.2) 中的漂移项和 $\boldsymbol{X}(t)$ 的点波动率矩阵 (2.3) 的对角元素分别满足 $\max_{1\leqslant i\leqslant p}\max_{0\leqslant t\leqslant 1}\mathrm{E}[|\mu_i(t)|^{2\beta}]<\infty$, $\max_{1\leqslant i\leqslant p}\max_{0\leqslant t\leqslant 1}\mathrm{E}[|\gamma_{ii}(t)|^{\beta}]<\infty$.

假设 A.3 存在常数 C_1 和 C_2 使得 $\max_{1\leqslant i\leqslant p}\frac{n_i}{n}\leqslant C_1$, $\max_{1\leqslant i\leqslant p}\max_{1\leqslant l\leqslant n_i}|t_{i,l}-t_{i,l-1}|\leqslant\frac{C_2}{n}$, 其中 $n=(n_1+\cdots+n_p)/p$, n_i 是第 i 个资产的观察样本量, $t_{i,l}$ 是第 i 个资产第 l 个观察样本的记录时间.

以上两个假设与传统的连续扩散过程的假设一样, 虽然减少了对噪声矩条件的要求, 但是为估计微观结构噪声增加了假设 A.1.

接下来证明文中的定理, 为方便叙述, 本文针对不同的估计量分别进行定理 4.1 和 4.2 的证明. 记 C 是一般的常数, 它的值与 n 和 p 无关, 并且其随着不同的情形改变, 但仍为常数. 当 n 和 p 趋向于无穷时, O_p 记为概率的阶数.

附录 A.1 关于 $\widehat{\boldsymbol{\Gamma}}^{(1)}$ 定理的证明

由 $\widehat{\boldsymbol{\Gamma}}^{(1)}$ 的定义知,

$$\begin{aligned}|\widehat{\Gamma}_{ij}^{(1)}-\Gamma_{ij}|^\beta&=\left|\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m\Delta\widehat{X}_i^C(\tau_{i,r}^k)\Delta\widehat{X}_j^C(\tau_{j,r}^k)-\Gamma_{ij}\right|^\beta\\&=\left|\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m[\Delta Y_i(\tau_{i,r}^k)I(|\Delta Y_i(\tau_{i,r}^k)|\leqslant\widehat{C}_i)-\Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\widehat{\boldsymbol{\theta}}_i)]\right.\\&\quad\left.\times[\Delta Y_j(\tau_{j,r}^k)I(|\Delta Y_j(\tau_{j,r}^k)|\leqslant\widehat{C}_j)-\Delta g(\boldsymbol{Z}_j(\tau_{j,r}^k);\widehat{\boldsymbol{\theta}}_j)]-\Gamma_{ij}\right|^\beta\\&\leqslant C\left\{\left|\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m[\Delta X_i^C(\tau_{i,r}^k)\Delta X_j^C(\tau_{j,r}^k)-\Gamma_{ij}]\right|^\beta+|H_1|^\beta+|H_2|^\beta+|H_3|^\beta\right\}.\end{aligned}$$

上式中的最后一步由凸函数性质获得. 首先,

$$\begin{aligned}|H_1|^\beta=&\left|\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m[\Delta X_i^D(\tau_{i,r}^k)I(|\Delta Y_i(\tau_{i,r}^k)|\leqslant\widehat{C}_i)-\Delta X_i^C(\tau_{i,r}^k)I(|\Delta Y_i(\tau_{i,r}^k)|>\widehat{C}_i)\right.\\&+(\Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\boldsymbol{\theta}_i)-\Delta\widehat{g}(\boldsymbol{Z}_i(\tau_{i,r}^k);\widehat{\boldsymbol{\theta}}_i))I(|\Delta Y_i(\tau_{i,r}^k)|\leqslant\widehat{C}_i)\\&\left.-\Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\widehat{\boldsymbol{\theta}}_i)I(|\Delta Y_i(\tau_{i,r}^k)|>\widehat{C}_i)]\Delta X_j^C(\tau_{j,r}^k)\right|^\beta\\\leqslant&\,C(H_{11}+H_{12}+H_{13}+H_{14}).\end{aligned}$$

同样利用凸函数性质易知上式的不等式成立, 其中

$$\begin{aligned}H_{11}&=\left|\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m\Delta X_i^D(\tau_{i,r}^k)I(|\Delta Y_i(\tau_{i,r}^k)|\leqslant\widehat{C}_i)\Delta X_j^C(\tau_{j,r}^k)\right|^\beta\\&=\left|\frac{1}{K}\sum_{(k,r)\in M}\Delta X_i^D(\tau_{i,r}^k)\Delta X_j^C(\tau_{j,r}^k)\right|^\beta\leqslant\frac{C}{K^\beta}\sum_{(k,r)\in M}|\Delta X_j^C(\tau_{j,r}^k)|^\beta,\end{aligned}\tag{A.1}$$

这里 $M=\{(t,r)\mid\Delta X_i^D(\tau_{i,r}^k)I(|\Delta Y_i(\tau_{i,r}^k)|\leqslant\widehat{C}_i)\neq0,k=1,\ldots,K,r=1,\ldots,m\}$, 由跳跃的有限性知 $\#\{M\}$ (集合 M 的个数) 是一个有限常数. 类似地,

$$H_{12}=\left|\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m\Delta X_i^C(\tau_{i,r}^k)I(|\Delta Y_i(\tau_{i,r}^k)|>\widehat{C}_i)\Delta X_j^C(\tau_{j,r}^k)\right|^\beta$$

$$\leqslant \frac{1}{K^\beta} \sum_{(k,r)\in N} |\Delta X_i^C(\tau_{i,r}^k)\Delta X_j^C(\tau_{j,r}^k)|^\beta, \tag{A.2}$$

其中 $N = \{(t,r) \mid |\Delta Y_i(\tau_{i,r}^k)| > \widehat{C}_i, k = 1,\ldots,K, r = 1,\ldots,m\}$, 由跳跃的有限性知 $|\Delta Y_i(\tau_{i,r}^k)| > \widehat{C}_i$ $(k = 1,\ldots,K,\ r = 1,\ldots,m)$ 发生无穷多次的概率为 0, 因此, $\#\{N\}$ 是一个有限常数. 接下来计算 $|H_1|^\beta$ 上界的第三部分, 即

$$\begin{aligned} H_{13} &= \left|\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m [\Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\boldsymbol{\theta}_i) - \Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\widehat{\boldsymbol{\theta}}_i)] I(|\Delta Y_i(\tau_{i,r}^k)| \leqslant \widehat{C}_i)\right|^\beta \\ &\leqslant \left(\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m |\Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\boldsymbol{\theta}_i) - \Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\widehat{\boldsymbol{\theta}}_i)|\right)^\beta \\ &\leqslant C\frac{1}{K^\beta}\left(\sum_{k=1}^K\sum_{r=1}^m |\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i|\right)^\beta. \end{aligned} \tag{A.3}$$

上式中的不等式分别由绝对值的性质和假设 A.1(v) 得到. 类似地有

$$H_{14} = \left|\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m \Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\widehat{\boldsymbol{\theta}}_i) I(|\Delta Y_i(\tau_{i,r}^k)| > \widehat{C}_i)\Delta X_j^C(\tau_{j,r}^k)\right|^\beta \leqslant \frac{C}{K^\beta}\sum_{(k,r)\in N} |\Delta X_j^C(\tau_{j,r}^k)|^\beta. \tag{A.4}$$

由 (A.1)–(A.4) 易知,

$$\begin{aligned} \mathrm{E}|H_1|^\beta &\leqslant \frac{C}{K^\beta}\Bigg[\mathrm{E}\Bigg(\sum_{(k,r)\in M} |\Delta X_j^C(\tau_{j,r}^k)|^\beta\Bigg) + \mathrm{E}\Bigg(\sum_{(k,r)\in N} |\Delta X_i^C(\tau_{i,r}^k)\Delta X_j^C(\tau_{j,r}^k)|^\beta\Bigg) \\ &\quad + \mathrm{E}\Bigg(\sum_{k=1}^K\sum_{r=1}^m |\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i|\Bigg)^\beta\Bigg] \\ &\leqslant \frac{C}{K^\beta}(m^{-\beta/2} + m^{-\beta} + n^{1-\beta}). \end{aligned}$$

同理可得

$$\mathrm{E}|H_2|^\beta \leqslant \frac{C}{K^\beta}(m^{-\beta/2} + m^{-\beta} + n^{1-\beta}).$$

接下来计算 $\mathrm{E}|H_3|^\beta$ 的量级, 首先计算

$$\begin{aligned} |H_3|^\beta = \Bigg|&\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m [\Delta X_i^D(\tau_{i,r}^k) I(|\Delta Y_i(\tau_{i,r}^k)| \leqslant \widehat{C}_i) - \Delta X_i^C(\tau_{i,r}^k) I(|\Delta Y_i(\tau_{i,r}^k)| > \widehat{C}_i) \\ &+ (\Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\boldsymbol{\theta}_i) - \Delta\widehat{g}(\boldsymbol{Z}_i(\tau_{i,r}^k);\widehat{\boldsymbol{\theta}}_i)) I(|\Delta Y_i(\tau_{i,r}^k)| \leqslant \widehat{C}_i) \\ &- \Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\widehat{\boldsymbol{\theta}}_i) I(|\Delta Y_i(\tau_{i,r}^k)| > \widehat{C}_i)] \\ &\times [\Delta X_j^D(\tau_{j,r}^k) I(|\Delta Y_j(\tau_{j,r}^k)| \leqslant \widehat{C}_j) - \Delta X_j^C(\tau_{j,r}^k) I(|\Delta Y_j(\tau_{j,r}^k)| > \widehat{C}_j) \\ &+ (\Delta g(\boldsymbol{Z}_j(\tau_{j,r}^k);\boldsymbol{\theta}_j) - \Delta\widehat{g}(\boldsymbol{Z}_j(\tau_{j,r}^k);\widehat{\boldsymbol{\theta}}_j)) I(|\Delta Y_j(\tau_{j,r}^k)| \leqslant \widehat{C}_j) \\ &- \Delta g(\boldsymbol{Z}_j(\tau_{j,r}^k);\widehat{\boldsymbol{\theta}}_j) I(|\Delta Y_j(\tau_{j,r}^k)| > \widehat{C}_j)]\Bigg|^\beta. \end{aligned}$$

根据上式的对称性, 理论上应该分为 10 个部分进行计算, 鉴于该计算过程基本与计算 $|H_1|^\beta$ 的过程类似, 经计算最后得 $\mathrm{E}|H_3|^\beta \leqslant \frac{C}{K^\beta}$.

定理 4.1 的证明 取 $m=O(n^{2/3})$, 则 $K=O(n^{1/3})$, 综上可知,

$$\begin{aligned}\mathrm{E}|\widehat{\Gamma}_{ij}^{(1)}-\Gamma_{ij}|^{\beta}&\leqslant C\left[\mathrm{E}\left|\frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{m}\Delta X_i^C(\tau_{i,r}^k)\Delta X_j^C(\tau_{j,r}^k)-\Gamma_{ij}\right|^{\beta}+\mathrm{E}|H_1|^{\beta}+\mathrm{E}|H_2|^{\beta}+\mathrm{E}|H_3|^{\beta}\right]\\&\leqslant Cn^{-\beta/3}.\end{aligned}$$

因此定理 4.1(i) 结论得证. 接下来证明估计量的半正定性.

对任意非零向量 $\boldsymbol{x}$, 有

$$\begin{aligned}\boldsymbol{x}^{\mathrm{T}}\widehat{\boldsymbol{\Gamma}}^{(1)}\boldsymbol{x}&=\frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{m}\boldsymbol{x}^{\mathrm{T}}[\Delta\widehat{\boldsymbol{X}}^C(\tau_r^k)][\Delta\widehat{\boldsymbol{X}}^C(\tau_r^k)]^{\mathrm{T}}\boldsymbol{x}\\&=\frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{m}\{\boldsymbol{x}^{\mathrm{T}}[\Delta\widehat{\boldsymbol{X}}^C(\tau_r^k)]\}\{\boldsymbol{x}^{\mathrm{T}}[\Delta\widehat{\boldsymbol{X}}^C(\tau_r^k)]\}^{\mathrm{T}}\\&\geqslant 0.\end{aligned}$$

因此定理 4.1(ii) 结论得证. □

为方便证明定理 4.2, 首先介绍如下引理:

引理 A.1 假设 $\boldsymbol{\Gamma}$ 满足稀疏条件, 则在模型 (2.1)、(2.2) 和假设 A.1–A.3 以及跳跃有限的条件下, 有

$$\max_{1\leqslant i,j\leqslant p}|\widehat{\Gamma}_{ij}-\Gamma_{ij}|=O_p(\varpi),\tag{A.5}$$

$$\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}I(|\Gamma_{ij}|\geqslant\varpi)=O_p(\pi(p)\varpi^{-\delta}),\tag{A.6}$$

$$\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}I(|\widehat{\Gamma}_{ij}|\geqslant\varpi,|\Gamma_{ij}|<\varpi)=O_p(\pi(p)\varpi^{-\delta}),\tag{A.7}$$

$$\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\Gamma_{ij}|I(|\Gamma_{ij}|<\varpi)=O_p(\pi(p)\varpi^{1-\delta}).\tag{A.8}$$

证明 利用 Markov 不等式和定理 4.1 中的 (4.1) 得

$$\begin{aligned}\mathrm{P}\Big(\max_{1\leqslant i,j\leqslant p}|\widehat{\Gamma}_{ij}-\Gamma_{ij}|>\varpi\Big)&\leqslant\sum_{0\leqslant i,j\leqslant p}\mathrm{P}(|\widehat{\Gamma}_{ij}-\Gamma_{ij}|>\varpi)\\&\leqslant\frac{Cp^2}{\varpi^{\beta}n^{\beta/3}}=\frac{C}{h_{n,p}^{\beta}}\to 0,\quad n,p\to\infty.\end{aligned}$$

从而结论 (A.5) 得证.

根据稀疏性条件 (3.2) 知,

$$\begin{aligned}\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}I(|\Gamma_{ij}|\geqslant\varpi)&\leqslant\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}\frac{|\Gamma_{ij}|^{\delta}}{\varpi^{\delta}}I(|\Gamma_{ij}|\geqslant\varpi)\\&\leqslant\varpi^{-\delta}\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\Gamma_{ij}|^{\delta}=O_p(\pi(p)\varpi^{-\delta}).\end{aligned}\tag{A.9}$$

从而结论 (A.6) 得证.

利用定理 4.1 中的 (4.1) 结论易知,

$$\begin{aligned}\mathrm{P}\bigg(\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}I\bigg\{|\widehat{\Gamma}_{ij}-\Gamma_{ij}|>\frac{\varpi}{2}\bigg\}>0\bigg)&\leqslant \mathrm{P}\bigg(\max_{1\leqslant i,j\leqslant p}|\widehat{\Gamma}_{ij}-\Gamma_{ij}|>\frac{\varpi}{2}\bigg)\\&\leqslant \frac{Cp^2}{\varpi^{\beta}n^{\beta/3}}=\frac{C}{h_{n,p}^{\beta}}\to 0,\quad n,p\to\infty.\end{aligned}$$

因此,

$$\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}I\bigg(|\widehat{\Gamma}_{ij}-\Gamma_{ij}|\geqslant\frac{\varpi}{2}\bigg)=o_p(1). \tag{A.10}$$

结合 (A.9) 和 (A.10) 可得

$$\begin{aligned}&\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}I\{|\widehat{\Gamma}_{ij}|\geqslant\varpi,|\Gamma_{ij}|<\varpi\}\\&\leqslant\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}I\bigg\{|\widehat{\Gamma}_{ij}|\geqslant\varpi,|\Gamma_{ij}|\leqslant\frac{\varpi}{2}\bigg\}+\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}I\bigg\{|\widehat{\Gamma}_{ij}|\geqslant\varpi,\frac{\varpi}{2}<|\Gamma_{ij}|<\varpi\bigg\}\\&\leqslant\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}I\bigg\{|\widehat{\Gamma}_{ij}-\Gamma_{ij}|\geqslant\frac{\varpi}{2}\bigg\}+\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}I\bigg\{|\Gamma_{ij}|>\frac{\varpi}{2}\bigg\}\\&\leqslant o_p(1)+\bigg(\frac{\varpi}{2}\bigg)^{-\delta}M\pi(p)\\&=O_p(\pi(p)\varpi^{-\delta}).\end{aligned}$$

从而结论 (A.7) 得证.

根据稀疏性条件 (3.2), 进行简单放缩有

$$\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\Gamma_{ij}|I(|\Gamma_{ij}|\leqslant\varpi)=\varpi^{1-\delta}\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\Gamma_{ij}|^{\delta}=O_p(\pi(p)\varpi^{1-\delta}),$$

从而结论 (A.8) 得证. □

定理 4.2 的证明 根据 L_2- 范数和 L_∞- 范数的三角不等式, 有

$$\begin{aligned}\|\mathcal{T}_{\varpi}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathbf{\Gamma}\|_2&\leqslant\|\mathcal{T}_{\varpi}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathcal{T}_{\varpi}[\mathbf{\Gamma}]\|_2+\|\mathcal{T}_{\varpi}[\mathbf{\Gamma}]-\mathbf{\Gamma}\|_2\\&\leqslant\|\mathcal{T}_{\varpi}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathcal{T}_{\varpi}[\mathbf{\Gamma}]\|_\infty+\|\mathcal{T}_{\varpi}[\mathbf{\Gamma}]-\mathbf{\Gamma}\|_\infty.\end{aligned} \tag{A.11}$$

由引理 A.1 中的 (A.8) 知 (A.11) 的第二项满足

$$\|\mathcal{T}_{\varpi}[\mathbf{\Gamma}]-\mathbf{\Gamma}\|_\infty=\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\Gamma_{ij}|I(|\Gamma_{ij}|<\varpi)=O_p(\pi(p)\varpi^{1-\delta}). \tag{A.12}$$

经过简单的代数运算, (A.11) 的第一项化为

$$\|\mathcal{T}_{\varpi}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathcal{T}_{\varpi}[\mathbf{\Gamma}]\|_\infty=\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\widehat{\Gamma}_{ij}^{(1)}I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi)-\Gamma_{ij}I(|\Gamma_{ij}|\geqslant\varpi)|$$

$$
\begin{aligned}
&\leqslant \max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\widehat{\Gamma}_{ij}^{(1)}-\Gamma_{ij}|I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi,|\Gamma_{ij}|\geqslant\varpi)\\
&\quad+\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\widehat{\Gamma}_{ij}^{(1)}|I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi,|\Gamma_{ij}|<\varpi)\\
&\quad+\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\Gamma_{ij}|I(|\widehat{\Gamma}_{ij}^{(1)}|<\varpi,|\Gamma_{ij}|\geqslant\varpi)\\
&=: I_1+I_2+I_3.
\end{aligned}
$$

注意到上式的第一项

$$
\begin{aligned}
I_1&=\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\widehat{\Gamma}_{ij}^{(1)}-\Gamma_{ij}|I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi,|\Gamma_{ij}|\geqslant\varpi)\\
&\leqslant\max_{1\leqslant i,j\leqslant p}|\widehat{\Gamma}_{ij}^{(1)}-\Gamma_{ij}|\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}I(|\Gamma_{ij}|\geqslant\varpi)\\
&=O_p(\pi(p)\varpi^{1-\delta}).
\end{aligned}\tag{A.13}
$$

最后一步是由引理 A.1 中的 (A.5) 和 (A.6) 得到.

由绝对值不等式及引理 A.1 中的 (A.5)、(A.7) 和 (A.8) , 有

$$
\begin{aligned}
I_2&=\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\widehat{\Gamma}_{ij}^{(1)}|I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi,|\Gamma_{ij}|<\varpi)\\
&\leqslant\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}(|\widehat{\Gamma}_{ij}^{(1)}-\Gamma_{ij}|+|\Gamma_{ij}|)I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi,|\Gamma_{ij}|<\varpi)\\
&\leqslant\max_{1\leqslant i,j\leqslant p}|\widehat{\Gamma}_{ij}^{(1)}-\Gamma_{ij}|\sum_{j=1}^{p}I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi,|\Gamma_{ij}|<\varpi)+\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\Gamma_{ij}|I(|\Gamma_{ij}|<\varpi)\\
&=O_p(\pi(p)\varpi^{1-\delta}).
\end{aligned}\tag{A.14}
$$

类似地,

$$
\begin{aligned}
I_3&=\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\Gamma_{ij}|I(|\widehat{\Gamma}_{ij}^{(1)}|<\varpi,|\Gamma_{ij}|\geqslant\varpi)\\
&\leqslant\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}(|\widehat{\Gamma}_{ij}^{(1)}-\Gamma_{ij}|+|\widehat{\Gamma}_{ij}^{(1)}|)I(|\widehat{\Gamma}_{ij}^{(1)}|<\varpi,|\Gamma_{ij}|\geqslant\varpi)\\
&\leqslant\max_{1\leqslant i,j\leqslant p}|\widehat{\Gamma}_{ij}^{(1)}-\Gamma_{ij}|\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}I(|\Gamma_{ij}|\geqslant\varpi)+\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}\varpi I(|\Gamma_{ij}|\geqslant\varpi)\\
&=O_p(\pi(p)\varpi^{1-\delta}).
\end{aligned}\tag{A.15}
$$

综合 (A.13)–(A.15) 中的 I_1、I_2 和 I_3 的量级可以得到

$$
\|\mathcal{T}_{\varpi}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathcal{T}_{\varpi}[\mathbf{\Gamma}]\|_{\infty}=O_p(\pi(p)\varpi^{1-\delta}).\tag{A.16}
$$

因此, 根据 (A.11)、(A.12) 和 (A.16), 定理 4.2(i) 得证, 接下来证明定理 4.2(ii).

类似于定理 4.2(i), 有

$$\begin{aligned}\|\mathcal{T}_{\varpi}^{H}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathbf{\Gamma}\|_2 &\leqslant \|\mathcal{T}_{\varpi}^{H}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathcal{T}_{\varpi}^{H}[\mathbf{\Gamma}]\|_2+\|\mathcal{T}_{\varpi}^{H}[\mathbf{\Gamma}]-\mathbf{\Gamma}\|_2\\ &\leqslant \|\mathcal{T}_{\varpi}^{H}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathcal{T}_{\varpi}^{H}[\mathbf{\Gamma}]\|_\infty+\|\mathcal{T}_{\varpi}^{H}[\mathbf{\Gamma}]-\mathbf{\Gamma}\|_\infty. \end{aligned}\tag{A.17}$$

显然上式的第二部分满足

$$\begin{aligned}\|\mathcal{T}_{\varpi}^{H}[\mathbf{\Gamma}]-\mathbf{\Gamma}\|_\infty &= \max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\Gamma_{ij}|I(|\Gamma_{ij}|\geqslant\varpi,i\neq j)+\Gamma_{ij}|I(|\Gamma_{ij}|\geqslant 0,i=j)-\Gamma_{ij}|\\ &\leqslant \max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\Gamma_{ij}|I(|\Gamma_{ij}|\leqslant\varpi)=O_p(\pi(p)\varpi^{1-\delta}). \end{aligned}\tag{A.18}$$

最后一个不等式是由引理 A.1 中的 (A.8) 得到. 如果能证明 $\|\mathcal{T}_{\varpi}^{H}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathcal{T}_{\varpi}^{H}[\mathbf{\Gamma}]\|_\infty$ 和 $\pi(p)\varpi^{1-\delta}$ 在概率意义上同阶, 那么结论得证. 事实上, 根据引理 A.1 经过简单的迭代, 有

$$\begin{aligned}\|\mathcal{T}_{\varpi}^{H}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathcal{T}_{\varpi}^{H}[\mathbf{\Gamma}]\|_\infty &= \max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\widehat{\Gamma}_{ij}^{(1)}I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi,i\neq j)+\widehat{\Gamma}_{ij}^{(1)}I(|\widehat{\Gamma}_{ij}|\geqslant 0,i=j)\\ &\quad -\Gamma_{ij}I(|\Gamma_{ij}|\geqslant\varpi,i\neq j)-\Gamma_{ij}I(|\Gamma_{ij}|\geqslant 0,i=j)|\\ &\leqslant \max_{1\leqslant i\leqslant p}\left\{\sum_{j\neq i}|\widehat{\Gamma}_{ij}^{(1)}I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi)-\Gamma_{ij}I(|\Gamma_{ij}|\geqslant\varpi)|+|\widehat{\Gamma}_{ii}^{(1)}-\Gamma_{ii}|\right\}\\ &\leqslant \max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\widehat{\Gamma}_{ij}^{(1)}-\Gamma_{ij}|I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi,|\Gamma_{ij}|\geqslant\varpi)\\ &\quad+\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\widehat{\Gamma}_{ij}^{(1)}|I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi,|\Gamma_{ij}|<\varpi)\\ &\quad+\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\Gamma_{ij}|I(|\widehat{\Gamma}_{ij}^{(1)}|<\varpi,|\Gamma_{ij}|\geqslant\varpi)+O_p(n^{-1/3})\\ &=: I_1+I_2+I_3+O_p(n^{-1/3}), \end{aligned}$$

结合 (A.13)–(A.15) 中的 I_1、I_2 和 I_3 的量级可以得到

$$\|\mathcal{T}_{\varpi}^{H}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathcal{T}_{\varpi}^{H}[\mathbf{\Gamma}]\|_\infty=O_p(\pi(p)\varpi^{1-\delta}).\tag{A.19}$$

根据 (A.17)–(A.19), 定理 4.2(ii) 结论得证, 接下来证明定理 4.2(iii).

类似于前面的 (A.17), 有

$$\begin{aligned}\|\mathcal{T}_{\varpi}^{S}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathbf{\Gamma}\|_2 &\leqslant \|\mathcal{T}_{\varpi}^{S}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathcal{T}_{\varpi}^{S}[\mathbf{\Gamma}]\|_2+\|\mathcal{T}_{\varpi}^{S}[\mathbf{\Gamma}]-\mathbf{\Gamma}\|_2\\ &\leqslant \|\mathcal{T}_{\varpi}^{S}[\widehat{\mathbf{\Gamma}}^{(1)}]-\mathcal{T}_{\varpi}^{S}[\mathbf{\Gamma}]\|_\infty+\|\mathcal{T}_{\varpi}^{S}[\mathbf{\Gamma}]-\mathbf{\Gamma}\|_\infty. \end{aligned}\tag{A.20}$$

根据 $\mathcal{T}_{\varpi}^{S}[\widehat{\mathbf{\Gamma}}^{(1)}]$ 的定义及引理 A.1 中的 (A.6) 和 (A.8), 有

$$\|\mathcal{T}_{\varpi}^{S}[\mathbf{\Gamma}]-\mathbf{\Gamma}\|_\infty=\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|[\Gamma_{ij}-\mathrm{sign}(\Gamma_{ij})\varpi]I(|\Gamma_{ij}|\geqslant\varpi,i\neq j)+\Gamma_{ij}I(|\Gamma_{ij}|\geqslant 0,i=j)-\Gamma_{ij}|$$

$$
\begin{aligned}
&\leqslant \max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|\Gamma_{ij}|I(|\Gamma_{ij}|<\varpi)+\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}\varpi I(|\Gamma_{ij}|\geqslant\varpi)\\
&=O_p(\pi(p)\varpi^{1-\delta}).
\end{aligned}\tag{A.21}
$$

经过简单的推导, 结合引理 A.1 中的 (A.6) 和 (A.7), 有

$$
\begin{aligned}
&\|\mathcal{T}_{\varpi}^{S}[\widehat{\boldsymbol{\Gamma}}^{(1)}]-\mathcal{T}_{\varpi}^{S}[\boldsymbol{\Gamma}]\|_{\infty}\\
&=\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}|[\widehat{\Gamma}_{ij}^{(1)}-\operatorname{sign}(\widehat{\Gamma}_{ij}^{(1)})\varpi]I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi,i\neq j)\\
&\quad+\widehat{\Gamma}_{ij}^{(1)}I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant 0,i=j)-\Gamma_{ij}I(|\Gamma_{ij}|\geqslant 0,i=j)\\
&\quad-[\Gamma_{ij}-\operatorname{sign}(\Gamma_{ij})\varpi]I(|\Gamma_{ij}|\geqslant\varpi,i\neq j)|\\
&\leqslant\max_{1\leqslant i\leqslant p}\left\{\sum_{j\neq i}|\widehat{\Gamma}_{ij}^{(1)}I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi)-\Gamma_{ij}I(|\Gamma_{ij}|\geqslant\varpi)|+|\widehat{\Gamma}_{ii}^{(1)}-\Gamma_{ii}|\right\}\\
&\quad+\varpi\max_{1\leqslant i\leqslant p}\sum_{j\neq i}[I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi)+I(|\Gamma_{ij}|\geqslant\varpi)]\\
&\leqslant\|\mathcal{T}_{\varpi}^{H}[\widehat{\boldsymbol{\Gamma}}^{(1)}]-\mathcal{T}_{\varpi}^{H}[\boldsymbol{\Gamma}]\|_{\infty}+\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}\varpi I(|\Gamma_{ij}|\geqslant\varpi)+\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}\varpi I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi)\\
&\leqslant O_p(\pi(p)\varpi^{1-\delta})+2\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}\varpi I(|\Gamma_{ij}|\geqslant\varpi)+\max_{1\leqslant i\leqslant p}\sum_{j=1}^{p}\varpi I(|\widehat{\Gamma}_{ij}^{(1)}|\geqslant\varpi,|\Gamma_{ij}|<\varpi)\\
&=O_p(\pi(p)\varpi^{1-\delta}).
\end{aligned}\tag{A.22}
$$

由 (A.20)–(A.22), 定理 4.2(iii) 得证. □

附录 A.2 关于 $\widehat{\boldsymbol{\Gamma}}^{(2)}$ 定理的证明

为了证明定理 4.1, 首先介绍如下引理:

引理 A.2 在模型 (2.1)、(2.2) 和假设 A.1–A.3 以及跳跃有限的条件下,

(i) 对每一个固定的 K 和 m, 依概率 1, 有

$$
|\Delta\widehat{X}_i(\tau_{i,r}^k)|I(|\Delta\widehat{X}_i(\tau_{i,r}^k)|>(\Delta\tau_{i,r}^k)^{\xi})=|\Delta X_i(\tau_{i,r}^k)+b_{i,r}^k|I([\tau_{i,r-1}^k,\tau_{i,r}^k]\text{ 存在跳})
$$

对任意的 $i=1,\ldots,p$, $k=1,\ldots,K$, $r=1,\ldots,m$ 成立, 其中 $b_{i,r}^k=\Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\boldsymbol{\theta}_i)-\Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\widehat{\boldsymbol{\theta}}_i)$.

(ii) $\frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{m}\Delta\widehat{X}_i^C(\tau_{i,r}^k)\Delta\widehat{X}_j^C(\tau_{j,r}^k)=\frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{m}\Delta X_i^C(\tau_{i,r}^k)\Delta X_j^C(\tau_{j,r}^k)+o_p(m^{-1/2})$.

证明 (i) 由简单的代数运算知,

$$
\begin{aligned}
|\Delta\widehat{X}_i(\tau_{i,r}^k)|&=|\Delta X_i(\tau_{i,r}^k)-[\Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\widehat{\boldsymbol{\theta}}_i)-\Delta g(\boldsymbol{Z}_i(\tau_{i,r}^k);\boldsymbol{\theta}_i)]|\\
&=|\Delta X_i(\tau_{i,r}^k)+b_i(\tau_{i,r}^k)|\leqslant|\Delta X_i(\tau_{i,r}^k)|+|b_i(\tau_{i,r}^k)|.
\end{aligned}
$$

由假设 A.1(v) 知, 对任意的 i、k 和 r, 存在 $C>0$ 使得 $|b_i(\tau_{i,r}^k)|\leqslant C|\widehat{\boldsymbol{\theta}}_i-\boldsymbol{\theta}_i|$. 利用文献 [12, 定理 4.1] 可以得到对任意的 $\xi\in(0,1/2)$, 依概率 1 有 $|b_i(\tau_{i,r}^k)|\leqslant(\Delta\tau_{i,r}^k)^{\xi}/2$ 对所有的 i、k 和 r 成立, 在该情形下, 如果 $|\Delta\widehat{X}_i(\tau_{i,r}^k)|>(\Delta\tau_{i,r}^k)^{\xi}$, 那么 $|\Delta X_i(\tau_{i,r}^k)|>(\Delta\tau_{i,r}^k)^{\xi}/2$.

由 Lévy 的连续模定理和假设 A.1 知, 对任意充分大的 n_i $(i=1,\ldots,p)$, 每个区间 $[\tau_{i,r-1}^k,\tau_{i,r}^k]$ 最多包含一个跳. 如果在区间 $[\tau_{i,r-1}^k,\tau_{i,r}^k]$ 内确实存在跳, 那么有 $|\Delta X_i(\tau_{i,r}^k)|\geqslant 3/2(\Delta\tau_{i,r}^k)^\xi$, 这意味着 $|\Delta\widehat{X}_i(\tau_{i,r}^k)|\geqslant(\Delta\tau_{i,r}^k)^\xi$. 因此依概率 1 有

$$|\Delta\widehat{X}_i(\tau_{i,r}^k)|I(|\Delta\widehat{X}_i(\tau_{i,r}^k)|>(\Delta\tau_{i,r}^k)^\xi)=|\Delta X_i(\tau_{i,r}^k)+b_i(\tau_{i,r}^k)|I([\tau_{i,r-1}^k,\tau_{i,r}^k]\text{ 存在跳}).$$

(ii) 由 (i) 的结论知依概率 1 有

$$\begin{aligned}
&\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m\Delta\widehat{X}_i(\tau_{i,r}^k)I(|\Delta\widehat{X}_i(\tau_{i,r}^k)|\leqslant(\Delta\tau_{i,r}^k)^\xi)\Delta\widehat{X}_j(\tau_{j,r}^k)I(|\Delta\widehat{X}_j(\tau_{j,r}^k)|\leqslant(\Delta\tau_{j,r}^k)^\xi)\\
&=\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m[\Delta X_i^C(\tau_{i,r}^k)+b_{i,r}^k][\Delta X^j(\tau_{j,r}^k)+b_{j,r}^k]\\
&=\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m[\Delta X_i^C(\tau_{i,r}^k)\Delta X_j^C(\tau_{j,r}^k)+\Delta X_i^C(\tau_{i,r}^k)b_{j,r}^k+\Delta X_j^C(\tau_{j,r}^k)b_{i,r}^k+b_{i,r}^kb_{j,r}^k]\\
&=:\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m\Delta X_i^C(\tau_{i,r}^k)\Delta X_j^C(\tau_{j,r}^k)+I_1+I_2+I_3.
\end{aligned}$$

接下来对 I_1、I_2 和 I_3 分别单独考虑它们的概率量级,

$$\begin{aligned}
\mathrm{E}|I_1|&=\mathrm{E}\left|\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m\Delta X_i^C(\tau_{i,r}^k)b_{j,r}^k\right|\leqslant C\frac{1}{K}(Km)^{-1/2}\sum_{k=1}^K\sum_{r=1}^m\mathrm{E}|\Delta X_i^C(\tau_{i,r}^k)|n^{-1}\\
&\leqslant C\frac{1}{nK}(Km)^{-1/2}\sum_{k=1}^K\sum_{r=1}^m\sqrt{\mathrm{E}|\Delta X_i^C(\tau_{i,r}^k)|^2}\leqslant C\frac{1}{nK}(Km)^{-1/2}\sum_{k=1}^K\sum_{r=1}^m\sqrt{\Delta\tau_{i,r}^k}\\
&\leqslant C\frac{1}{nK^{1/2}}.
\end{aligned}$$

类似地有 $\mathrm{E}|I_2|\leqslant\frac{C}{nK^{1/2}}$.

对于 I_3, 经过简单的代数运算有

$$\mathrm{E}|I_3|=\mathrm{E}\left|\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m b_{i,r}^kb_{j,r}^k\right|\leqslant C\frac{1}{K}(Km)^{-1/2}\sum_{k=1}^K\sum_{r=1}^m n^{-2}\leqslant C\frac{m^{1/2}}{n^2K^{1/2}}.$$

令 $m=O(n^{2/3})$, $K=O(n^{1/3})$, 由 Chebyshev 不等式易得到引理 A.2(ii).

由 $\widehat{\mathbf{\Gamma}}^{(2)}$ 的定义和引理 A.1 易知,

$$\begin{aligned}
&\mathrm{E}|\widehat{\Gamma}_{ij}^{(2)}-\Gamma_{ij}|^\beta\\
&=\mathrm{E}\left|\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m\Delta\widehat{X}_i(\tau_{i,r}^k)I(|\Delta\widehat{X}_i(\tau_{i,r}^k)|\leqslant(\Delta\tau_{i,r}^k)^\xi)\Delta\widehat{X}_j(\tau_{j,r}^k)I(|\Delta\widehat{X}_j(\tau_{j,r}^k)|\leqslant(\Delta\tau_{j,r}^k)^\xi)-\Gamma_{ij}\right|^\beta\\
&=\mathrm{E}\left|\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m\Delta X_i^C(\tau_{i,r}^k)\Delta X_j^C(\tau_{j,r}^k)+o_p(m^{-1/2})-\Gamma_{ij}\right|^\beta\\
&\leqslant\left\{\left[\mathrm{E}\left|\frac{1}{K}\sum_{k=1}^K\sum_{r=1}^m\Delta X_i^C(\tau_{i,r}^k)\Delta X_j^C(\tau_{j,r}^k)-\Gamma_{ij}\right|^\beta\right]^{1/\beta}+o(m^{-1/2})\right\}^\beta\\
&\leqslant Cn^{-\beta/3}.
\end{aligned}$$

倒数第二个不等式是由 Minkowski 不等式得到, 最后一个不等式直接由文献 [11, 定理 4.1] 获得. 因此定理 4.1(i) 结论得证. 接下来证明估计量的半正定性.

对任意非零向量 $\boldsymbol{x}$, 有

$$\boldsymbol{x}^{\mathrm{T}}\widehat{\boldsymbol{\Gamma}}^{(2)}\boldsymbol{x}=\frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{m}\boldsymbol{x}^{\mathrm{T}}[\Delta\widehat{\boldsymbol{X}}'^{C}(\tau_r^k)][\Delta\widehat{\boldsymbol{X}}'^{C}(\tau_r^k)]^{\mathrm{T}}\boldsymbol{x}=\frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{m}\{\boldsymbol{x}^{\mathrm{T}}[\Delta\widehat{\boldsymbol{X}}'^{C}(\tau_r^k)]\}\{\boldsymbol{x}^{\mathrm{T}}[\Delta\widehat{\boldsymbol{X}}'^{C}(\tau_r^k)]\}^{\mathrm{T}}$$
$$\geqslant 0,$$

其中 $\Delta\widehat{\boldsymbol{X}}'^{C}(\tau_r^k)$ 表示向量的每个元素满足 $\Delta\widehat{X}_i(\tau_{i,r}^k)I(|\Delta\widehat{X}_i(\tau_{i,r}^k)|\leqslant(\Delta\tau_{i,r}^k)^{\xi})$, 因此, 定理 4.1(ii) 结论得证. □

对于 $\widehat{\boldsymbol{\Gamma}}^{(2)}$, 定理 4.2 的证明基本与 $\widehat{\boldsymbol{\Gamma}}^{(1)}$ 方法一样, 在此不再详述.

附录 A.3 关于 $\widehat{\boldsymbol{\Gamma}}^{(3)}$ 和 $\widehat{\boldsymbol{\Gamma}}^{(4)}$ 定理的证明

为证明带跳的 $\widehat{\Gamma}_{ij}^{(3)}$ 的相合性, 首先考虑不带跳的情形下 $\widehat{\Gamma}_{ij}^{(3)}$ 的相合性, 因此考虑如下引理:

引理 A.3 在模型 (2.1) 和连续扩散的对数价格过程以及假设 A.1–A.3 下, 有

$$\Delta\widehat{X}_i(\tau_{i,r})=\Delta X_i(\tau_{i,r})+O_p(n^{-1}).$$

证明 由模型 (2.1) 知,

$$\Delta\widehat{X}_i(\tau_{i,r})=\Delta Y_i(\tau_{i,r})-\Delta g(\boldsymbol{Z}_i(t_{i,r});\widehat{\boldsymbol{\theta}}_i)=\Delta X_i(\tau_{i,r})-[\Delta g(\boldsymbol{Z}_i(t_{i,r});\widehat{\boldsymbol{\theta}}_i)-\Delta g(\boldsymbol{Z}_i(t_{i,r});\boldsymbol{\theta}_i)].$$

因此,

$$\begin{aligned}|\Delta\widehat{X}_i(\tau_{i,r})-\Delta X_i(\tau_{i,r})|&=|\Delta g(\boldsymbol{Z}_i(t_{i,r});\widehat{\boldsymbol{\theta}}_i)-\Delta g(\boldsymbol{Z}_i(t_{i,r});\boldsymbol{\theta}_i)|\\&\leqslant[L_0(\boldsymbol{Z}_i(t_{i,r}))+L_0(\boldsymbol{Z}_i(t_{i,r-1}))]|\widehat{\boldsymbol{\theta}}_i-\boldsymbol{\theta}_i|\\&=O_p(n^{-1}).\end{aligned}$$

因此结论得证. □

引理 A.4 在模型 (2.1) 和连续扩散的对数价格过程以及假设 A.1–A.3 下, 对任意的 $1\leqslant i,j\leqslant p$, 协方差估计量 $\widehat{\Gamma}_{ij}^{(3)}$ 满足 $\mathrm{E}(|\widehat{\Gamma}_{ij}^{(3)}-\Gamma_{ij}|^{\beta})\leqslant Cn^{-\frac{\beta}{3}}$, 其中 C 是与 n 和 p 无关的常数.

证明 由于对数价格是连续扩散过程, 因此, 由 MinRV 方法估计积分波动率矩阵 $\widehat{\boldsymbol{\Gamma}}^{(3)}$ 的定义知其第 (i,j) 个元素与真实积分波动率矩阵的第 (i,j) 个元素的差为

$$\begin{aligned}\widehat{\Gamma}_{ij}^{(3)}-\Gamma_{ij}&=\frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{m}\{\Delta\widehat{X}_i(\tau_{i,r})\vee\Delta\widehat{X}_i(\tau_{i,r+1})\}\{\Delta\widehat{X}_j(\tau_{j,r})\vee\Delta\widehat{X}_j(\tau_{j,r+1})\}-\Gamma_{ij}\\&=III_1+\{0\vee III_2\vee III_3\vee III_4\}.\end{aligned}$$

其中 $\vee$ 表示或者的意思, $III_1=\frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{m-1}\Delta\widehat{X}_i(\tau_{i,r})\Delta\widehat{X}_j(\tau_{j,r})-\Gamma_{ij}$. 由文献 [11, 定理 4.1] 证明知,

$$\mathrm{E}|III_1|^{\beta}\leqslant Cn^{-\frac{\beta}{3}}. \tag{A.23}$$

第二项 III_2 与第三项 III_3 同阶, 因此只考虑第二项 III_2, 即

$$III_2=\frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{m}[\Delta\widehat{X}_i(\tau_{i,r+1})-\Delta\widehat{X}_i(\tau_{i,r})]\Delta\widehat{X}_j(\tau_{j,r}).$$

由引理 A.3 和简单的代数运算得

$$
\begin{aligned}
\mathrm{E}|III_2|^\beta &\leqslant C(Km)^{\beta/2-1}K^{-\beta}\sum_{k=1}^{K}\sum_{r=1}^{m-1}\mathrm{E}|[\Delta\widehat{X}_i(\tau_{i,r+1})-\Delta\widehat{X}_i(\tau_{i,r})]\Delta\widehat{X}_j(\tau_{j,r})|^\beta\\
&= C(Km)^{\beta/2-1}K^{-\beta}\sum_{k=1}^{K}\sum_{r=1}^{m-1}\mathrm{E}\{|\Delta\widehat{X}_j(\tau_{j,r})|^\beta\mathrm{E}[|\Delta\widehat{X}_i(\tau_{i,r+1})-\Delta\widehat{X}_i(\tau_{i,r})|^\beta\mid X_j]\}\\
&\leqslant C(Km)^{\beta/2-1}K^{-\beta}\sum_{k=1}^{K}\sum_{r=1}^{m-1}\mathrm{E}\{|\Delta X_j(\tau_{j,r})+O_p(n^{-1})|^\beta\\
&\quad\times\mathrm{E}[|\Delta X_i(\tau_{i,r+1})-\Delta X_i(\tau_{i,r})+O_p(n^{-1})|^\beta\mid X_j]\}\\
&\leqslant C(Km)^{\beta/2-1}K^{-\beta}\sum_{k=1}^{K}\sum_{r=1}^{m-1}\mathrm{E}\Big\{|\Delta X_j(\tau_{j,r})|^\beta m^{-\beta/2}\max_{0\leqslant s\leqslant 1}\mathrm{E}|\gamma_{ii}(s)|^{\beta/2}+o(m^{-\beta/2})\Big\}\\
&\leqslant C(Km)^{\beta/2-1}K^{-\beta}\sum_{k=1}^{K}\sum_{r=1}^{m-1}[m^{-\beta}+o(m^{-\beta})]\Big[\max_{0\leqslant s\leqslant 1}\mathrm{E}|\gamma_{ii}(s)|^{\beta/2}\Big]^2\\
&\leqslant C(Km)^{-\beta/2}=Cn^{-\beta/2},
\end{aligned}
\tag{A.24}
$$

上式中的第二个不等式是由文献 [8, 引理 8.4] 的证明得到; 第三个不等式的推导类似于文献 [8, 引理 8.4] 的证明. 下面证明 III_4 的阶数:

$$
\begin{aligned}
\mathrm{E}|III_4|^\beta &= \mathrm{E}\left|\frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{m}[\Delta\widehat{X}_i(\tau_{i,r+1})\Delta\widehat{X}_j(\tau_{j+1,r})-\Delta\widehat{X}_i(\tau_{i,r})\Delta\widehat{X}_j(\tau_{j,r})]\right|^\beta\\
&= \mathrm{E}\left|\frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{m}[\Delta\widehat{X}_i(\tau_{i,r+1})-\Delta\widehat{X}_i(\tau_{i,r})]\Delta\widehat{X}_j(\tau_{j+1,r})+\Delta\widehat{X}_i(\tau_{i,r})[\Delta\widehat{X}_j(\tau_{j+1,r})-\Delta\widehat{X}_j(\tau_{j,r})]\right|^\beta\\
&\leqslant C(Km)^{-\beta/2},
\end{aligned}
\tag{A.25}
$$

最后一个不等式是由 III_2 的阶数和 III_3 的阶数确定的. 由 (A.23)–(A.25) 和 Minkowski 不等式知,

$$
\mathrm{E}|\widehat{\Gamma}_{ij}^{(3)}-\Gamma_{ij}|^\beta\leqslant C\{[\mathrm{E}|III_1|^\beta]^{1/\beta}+[\mathrm{E}|\{0\vee III_2\vee III_3\vee III_4\}|^\beta]^{1/\beta}\}^\beta\leqslant Cn^{-\beta/3}.
$$

因此引理 A.4 得证. □

注 A.1 该引理对估计量 $\widehat{\Gamma}_{ij}^{(4)}$ 也成立, 证明方法类似.

为区别带跳和不带跳的作用, 下面介绍引理 A.5.

引理 A.5 在模型 (2.1) 和 (2.2) 以及假设 A.1–A.3 下, 带跳和不带跳的过程 X_t 的 MinRV 和 MedRV 所获得的逐点收敛速度是相同的.

证明 在此只考虑 MinRV 的情形, MedRV 的形式类似. 因为跳的次数是有限的, 所以, $|\Delta X_i(\tau_{i,r})| = |X_i(\tau_{i,r})-X_i(\tau_{i,r-1})|$ 和 $|\Delta X_i(\tau_{i,r+1})| = |X_i(\tau_{i,r+1})-X_i(\tau_{i,r})|$ 最多有一个包含跳, 因此, 无论是否发生跳, 估计量的第 (i,j) 个元素中的每一项为下面 4 种可能:

$$
\begin{aligned}
&\{\Delta X_i(\tau_{i,r})\vee\Delta X_i(\tau_{i,r+1})\}\{\Delta X_j(\tau_{j,r})\vee\Delta X_j(\tau_{j,r+1})\}\\
&\quad=\{\Delta X_i(\tau_{i,r})\Delta X_j(\tau_{j,r})\vee\Delta X_i(\tau_{i,r+1})\Delta X_j(\tau_{j,r})\\
&\qquad\vee\Delta X_i(\tau_{i,r})\Delta X_j(\tau_{j,r+1})\vee\Delta X_i(\tau_{i,r+1})\Delta X_j(\tau_{j,r+1})\}
\end{aligned}
$$

$$\leqslant \max\{\Delta X_i^C(\tau_{i,r})\Delta X_j^C(\tau_{j,r}), \Delta X_i^C(\tau_{i,r+1})\Delta X_j^C(\tau_{j,r}),$$
$$\Delta X_i^C(\tau_{i,r})\Delta X_j^C(\tau_{j,r+1}), \Delta X_i^C(\tau_{i,r+1})\Delta X_j^C(\tau_{j,r+1})\}.$$

显然这里通过取邻近替换技术将包含跳跃的价格数据全转换成了连续的价格数据. □

由以上结论和跳的有限性 (不可能连续发生两次跳跃), 可得

$$\begin{aligned}\mathrm{E}|\widehat{\Gamma}_{ij}^{(3)} - \Gamma_{ij}|^\beta = \mathrm{E}\Big|\frac{1}{K}\sum_{k=1}^{K}\sum_{r=1}^{m-1}[&\max\{\Delta X_i^C(\tau_{i,r})\Delta X_j^C(\tau_{j,r}), \Delta X_i^C(\tau_{i,r+1})\Delta X_j^C(\tau_{j,r}),\\ &\Delta X_i^C(\tau_{i,r})\Delta X_j^C(\tau_{j,r+1}), \Delta X_i^C(\tau_{i,r+1})\Delta X_j^C(\tau_{j,r+1})\} - \Delta X_i^C(\tau_{i,r})\Delta X_j^C(\tau_{j,r})]\Big|^\beta\\ &\leqslant Cn^{-\beta/3}.\end{aligned}$$

上式的最后一个不等式是类似于 III_2 的证明. 因此定理 4.1(i) 得证. 由半正定性的定义直接得到定理 4.1(ii).

对于 $\widehat{\mathbf{\Gamma}}^{(4)}$ 的定理 4.1 证明类似于 $\widehat{\mathbf{\Gamma}}^{(3)}$ 的定理 4.1 证明; $\widehat{\mathbf{\Gamma}}^{(3)}$ 和 $\widehat{\mathbf{\Gamma}}^{(4)}$ 的定理 4.2 证明过程与 $\widehat{\mathbf{\Gamma}}^{(1)}$ 的定理 4.2 证明过程类似, 在此不再详述.

High-dimensional integrated volatility matrix estimation for high-frequency financial data with jumps

Yan Mu & Yong Zhou

Abstract The joint volatility matrix of assets is an important statistic for resource allocation and risk management. Accurate estimation of the joint volatility matrix is one of the hot issues in financial statistics and risk measurement. In this paper, we study the integral volatility matrix estimation of logarithmic price data with jumps under microstructure noise including market information. When the prices are not synchronized, and the number of assets and sample size tend to infinity, four estimation methods of high-dimensional integral volatility matrices are proposed by using the non-overlapping interval method and sparse characteristics. The convergence rate can reach the optimal convergence rate of the existing high-dimensional integral volatility matrix estimation. At the same time, the proposed adjusted estimators are consistent and semi-positive definite. The advantages and disadvantages of these estimators are compared in the simulation study. Finally the proposed methods are applied to the empirical study of Shanghai Securities Index data.

Keywords **high-frequency data, jump, market microstructure, high-dimensional volatility matrix, semi-positive definite**

MSC(2010) **97K80, 91G70**

doi: 10.1360/N012019-00030