中国科学: 数学 2022 年 第 52 卷 第 11 期: 1333～1360
SCIENTIA SINICA Mathematica
《中国科学》杂志社
SCIENCE CHINA PRESS
论 文

金融高频高维数据的波动率矩阵估计: 基于 GARCH-Itô 分组因子模型

高维清[1,2], 吴奔[1,2*], 张波[1,2*]

1. 中国人民大学应用统计科学研究中心, 北京 100872;
2. 中国人民大学统计学院, 北京 100872
E-mail: gaowq143@ruc.edu.cn, wuben@ruc.edu.cn, mabzhang@ruc.edu.cn

收稿日期: 2020-10-20; 接受日期: 2021-07-07; 网络出版日期: 2022-04-07; * 通信作者
国家自然科学基金 (批准号: 71873137 和 71471173) 和中国人民大学科学研究基金 (中央高校基本科研业务费专项资金资助) (批准号: 21XNLG08) 资助项目

摘要 在高频金融数据分析中, 高维波动率矩阵的估计和预测十分具有挑战性, 当金融资产存在自然的分组结构时, 此问题尤为突出. 为此, 本文提出一种新的 GARCH-Itô 分组因子模型, 将对数价格序列表示为共同因子、分组因子以及异质项, 并通过将离散的广义自回归条件异方差 (generalized autoregressive conditional heteroskedasticity, GARCH) 结构嵌入特征值过程的波动率中, 实现刻画数据波动率动态的目的. 本文利用伪极大似然法得到模型的参数估计, 建立极限理论, 模拟研究表明其良好的有限样本性质. 在实证研究中, 利用上海证券交易所主板及深圳证券交易所创业板的股票高频价格数据, 对比了不分组的模型及波动率矩阵的非参数多尺度已实现波动率 (multi-scale realized volatility, MSRV) 估计, 对比结果显示本文模型具有更好的波动率预测效果.

关键词 高维波动率矩阵 高频数据 混频模型 分组因子模型

MSC (2020) 主题分类 62H25, 91G10, 97K80

1 引言

资产配置是指, 投资者根据投资需求在不同类别资产上分配投资资金, 其目的是提高一定风险水平下的收益. 在进行资产配置时, 投资者通常需要考虑各资产的风险、收益、综合评分和规模等因素 (参见文献 [1]), 从而做出合理的投资决策. 例如, 研究表明, 在股票投资中, 按市值和市净率等因素选择股票 (参见文献 [2]), 组合低风险低收益和高风险高收益的股票, 可以让投资组合获得相比单一资产更大的风险回报. 因此, 在资产配置中常按某种因素将备选股票分门别类, 并按类别分散投资. 我们称这样选择的股票为具有特定分组结构的资产. 金融资产的波动率反映了资产收益的不确定性, 利用权重与波动率矩阵构成的二次型衡量了一种资产配置的投资风险. 所以, 资产配置问题的核心要素之一

英文引用格式: Gao W Q, Wu B, Zhang B. High-dimensional volatility matrix estimation with high-frequency financial data: The GARCH-Itô grouped factor model (in Chinese). Sci Sin Math, 2022, 52: 1333–1360, doi: 10.1360/SSM-2020-0304

© 2022《中国科学》杂志社

是估计具有分组结构的资产的波动率矩阵. 随着金融市场的日益繁荣与计算机技术的进步, 市场信息的获取、金融数据的分析与资产交易的进行都愈发便捷. 投资者在进行资产配置时, 不再集中于少量的资产上, 所考虑的市场信息也不再局限于低频的信息. 高频数据带来的大量观测可以帮助我们更好地理解收益的协方差结构并且可以迅速地调整投资组合, 且现代投资组合通常涉及大量资产 (参见文献 [3, 4]). 因此, 基于海量数据的大规模金融资产的波动率研究, 即高维、高频金融资产波动率矩阵估计问题, 也逐渐引起了人们的重视 (参见文献 [5]).

在分析具有分组结构的高维资产问题时, 因子模型不失为一种合适的选择. Sharpe [6] 提出了广为人知的资本资产定价模型 (capital asset pricing model, CAPM), 将市场风险作为单一因子, 利用大盘期望收益率与无风险利率之差构造超额收益率, 从而衡量单个股票的期望收益率与超额收益率的关系. Fama 和 French [7] 提出了三因子模型, 用市场风险、市值风险和账面市值比风险构造三因子, 来解释个股的超额收益率. Fama 和 French [8] 又进一步地引入盈利水平风险因子和投资水平风险因子, 将三因子模型推广至五因子模型. 另一方面, 研究者们也从纯数据驱动的角度来寻找因子, 将因子看成是需要估计的不可观测的变量或者隐变量, 从而提出统计因子模型 (参见文献 [9, 10]). 相比于金融因子模型, 统计因子模型虽然可能存在着金融解释模糊的缺点, 但也省去了精心挑选合适因子这一步骤 (参见文献 [11]). 当资产具有明显的分组结构时, 我们可以依据其分组结构来挑选或估计因子, 从而对资产进行良好刻画 (参见文献 [12]). 例如, Bekaert 等 [13] 将金融因子分为全球因子与地区因子, 来分析全球范围内的股票收益; Goyal 等 [12] 通过纽约证券交易所 (New York Stock Exchange, NYSE) 和纳斯达克证券交易所 (National Association of Securities Dealers Automated Quotations, NASDAQ) 的例子说明了交易所因子的必要性. 进一步地, Ando 和 Bai [14] 提出了分组因子模型的一般形式并给出了一些理论性质.

随着金融数据分析进入高频时代, 新的建模思路与统计工具也相继被提出. 一般地, 我们将以天或者更长时间段为单位的数据称为低频金融数据, 而将以分钟甚至更短时间段为单位的数据称为高频金融数据. 相比于传统的低频数据, 高频数据提供了更丰富的信息, 为我们展示了资产价格变化的更为复杂的动态过程. 对高频数据的统计分析和建模有助于我们更好地理解订单流、流动性和价格动态之间的相互作用 (参见文献 [5]). 在高频金融数据分析中, 资产的价格常被描述为连续时间 Itô 扩散过程, 而价格的波动常用积分波动率来描述, 即一段时间 (通常是一个交易日) 内波动的累积. 于是, 资产的波动率问题常常归结为积分波动率的估计. 由于市场微观结构噪声的存在 (参见文献 [15]), 积分波动率的估计并不是一个平凡的问题, 学者们已经提出了一系列的估计量, 如双尺度已实现波动率 (two-time scale realized volatility, TSRV) [15, 16]、MSRV [17]、核已实现波动率 (kernel realized volatility, KRV) [18]、伪极大似然估计量 (quasi-maximum likelihood estimator, QMLE) [19, 20] 和预平均已实现波动率 (pre-averaging realized volatility, PRV) [21] 等. 另外, 在实际中, 金融资产的价格有时存在很短时间内变化较大的特征, 扩散过程无法完全刻画这种性质. 于是, 就有学者提出利用跳扩散来对资产价格进行建模. 研究表明, 将资产价格的变化分为连续项和跳跃项可以提高波动率预测的精度 (参见文献 [22–25]). 对于含跳跃项的波动率估计问题, Fan 和 Wang [26] 利用 Wang [27] 提出的小波分析方法处理了跳跃项, 并改进了 TSRV 和 MSRV 统计量. Zhang 等 [28] 也提出了通过小波分析处理含噪声的高频数据跳跃项. 国内学者提出了通过非参数方法和 Gauss 混合模型处理高频数据跳跃项 (参见文献 [29, 30]). 例如, 穆燕和周勇 [31] 提出了带跳高频数据的 4 种高维积分波动率估计方法.

目前, 基于小规模资产的高频数据与基于大量资产的低频数据的研究都趋于成熟, 越来越多的研究针对基于大量资产和高频数据的波动率矩阵估计问题展开. 例如, Wang 和 Zou [32]、Tao 等 [33, 34] 通过假设积分波动率矩阵的稀疏性估计高维高频数据的波动率; Fan 等 [35]、Aït-Sahalia 和 Xiu [36]、Pelger [37]

以及 Kim 和 Fan [38] 则通过高频因子模型来估计高维高频数据的协方差矩阵. 其中, Kim 和 Fan [38] 提出的 GARCH-Itô 因子模型是 GARCH-Itô 模型 (参见文献 [39]) 在高维问题上的应用. 相比于传统的高频波动率估计量, GARCH-Itô 模型将高频的点波动嵌入低频的 GARCH 结构之中, 能够很好地刻画波动过程的聚集效应, 从而增加了波动率的可预测性. 在前人研究的基础上, 本文提出一种新的 GARCH-Itô 分组因子模型. 该模型假设资产的对数价格具有分组因子结构, 并用高低频数据结合的思想对因子过程的波动率进行建模, 以期在横向上区分不同资产波动特征的异同, 在纵向上刻画资产波动率随时间的变化. 相比于 Kim 和 Fan [38] 的 GARCH-Itô 因子模型, 本文所提出的模型具有以下优点: 第一, 该模型考虑了资产的真实价格在一些短时间内存在着较大的变动, 在建模时引入了对跳跃项的处理, 更准确地把握了资产价格的特征; 第二, 该模型引入了分组因子的概念, 将不同波动率特质的资产进行分组处理, 具有更高的可解释性. 本文的贡献在于, 促进了 GARCH-Itô 模型的理论与应用发展, 将 Kim 和 Fan [38] 提出的伪极大似然估计拓展至分组因子模型上, 提出了相应的共同因子与分组因子的个数估计方法及模型求解算法; 同时, 解决了分组因子模型中存在的特定问题, 如因子个数估计的一致性以及分组因子与共同因子的渐近正交性, 从而保证该模型在实际应用中是可靠的.

在国内, 实证研究说明我国股市中规模小的公司与规模大的公司的股票有不同的特征, 天然存在着组结构. 从波动率角度看, 小盘股相对大盘股表现得更加活跃, 小公司股票具有换手率高、股价波动率大的特点 (参见文献 [40]). 从收益率角度看, 中国股市存在着 "规模效应", 规模小的公司的收益率往往高于规模大的公司 (参见文献 [40–43]). 上述研究结论为本文模型的应用提供了一个良好的切入点, 在实证研究中, 我们根据规模因素选取波动率低、收益率低的大规模公司股票以及波动率高、收益率高的小规模公司股票进行分析. 实证结果表明, GARCH-Itô 分组因子模型在对资产的波动率矩阵的刻画上具有良好的表现, 由于分组因子的引入, 模型的解释性与预测性都有所提高.

本文余下内容的结构如下: 第 2 节给出 GARCH-Itô 分组因子模型的定义. 第 3 节提出模型参数的伪极大似然估计, 并给出相应的理论性质. 第 4 节利用数值模拟说明该模型的有效性和稳定性. 第 5 节将该模型应用于我国股市的实证研究之中. 第 6 节总结全文. 所有证明在附录中给出.

2 模型建立

在介绍模型之前, 首先给出下文需要用到的符号表示:

记 $\mathrm{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} A_{ii}$ 为 $n \times n$ 矩阵 $\boldsymbol{A}$ 的迹; $\|\boldsymbol{A}\|_F = \sqrt{\mathrm{tr}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})}$ 为矩阵 $\boldsymbol{A}$ 的 Frobenius 范数; $\|\boldsymbol{A}\|_2 = \sqrt{\max_{1\leqslant i\leqslant n}|\lambda_i|}$ 为矩阵 $\boldsymbol{A}$ 的 2- 范数, 其中 λ_i 是 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的特征值; 记 $\|\boldsymbol{A}\|_{\max} = \max_{i,j}|A_{ij}|$; $\mathrm{diag}(\boldsymbol{a})$ 为对角元素由向量 $\boldsymbol{a} = (a_1, \ldots, a_p)^{\mathrm{T}}$ 构成的对角矩阵; $\|\boldsymbol{a}\|_2 = \sqrt{\sum_{i=1}^{p} a_i^2}$ 为向量 $\boldsymbol{a}$ 的 2- 范数; $\|\boldsymbol{a}\|_\infty = \max_i |a_i|$ 为向量 $\boldsymbol{a}$ 的无穷范数; $\boldsymbol{a} \odot \boldsymbol{b} = (a_1b_1, \ldots, a_pb_p)^{\mathrm{T}}$ 为向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的 Hadamard 积. 记 $(\Omega, (\mathcal{F}_t)_{t\in T}, \mathcal{F}, \mathrm{P})$ 为一个带域流的概率空间, 其中 $(\mathcal{F}_t)_{t\in T}$ 是由 t 时刻之前市场信息诱导生成的 σ 代数, 本文所考虑的所有随机过程均基于此概率空间定义.

2.1 GARCH-Itô 分组因子模型

假设存在 p 个具有分组结构的金融资产 $i = 1, \ldots, p$, 并记第 i 个资产在 t 时刻的对数价格为 $X_i(t)$. 令 $\boldsymbol{X}(t) := (X_1(t), \ldots, X_p(t))^{\mathrm{T}}$, 这 p 个金融资产可以分为 K 组, 每组有 p_k 个资产, 即 $p = \sum_{k=1}^{K} p_k$.

考虑如下混频分组因子模型:

$$d\boldsymbol{X}(t)=\boldsymbol{\mu}dt+\sum_{k=0}^{K}\boldsymbol{B}^{(k)}(d)d\boldsymbol{F}^{(k)}(t)+d\boldsymbol{U}(t),\quad t\in[d-1,d),\tag{2.1}$$

其中, $\boldsymbol{\mu}$ 为模型的漂移项系数; $\boldsymbol{F}^{(k)}(t):=(F_1^{(k)}(t),\ldots,F_{r_k}^{(k)}(t))^{\mathrm{T}}$ $(k=0,\ldots,K)$ 表示 K 组因子过程, 记各组因子的个数为 r_k $(k=0,\ldots,K)$, 并令 $\tilde{r}=\sum_{k=0}^{K}r_k$; $\boldsymbol{B}^{(k)}(d):=(\beta_{i,j}^{(k)}(d))_{p\times r_k}$ $(k=0,\ldots,K)$ 表示相应的因子载荷过程; $\boldsymbol{U}(t)$ 表示异质过程; $d=1,\ldots,n$ 表示低频观测时刻, 如每日等.

为了刻画资产的分组结构, 我们将因子过程 $\boldsymbol{F}^{(k)}(t)$ $(k=0,\ldots,K)$ 分两部分考虑: $\boldsymbol{F}^{(0)}(t)$ 表示对所有资产均有影响的因子, $\boldsymbol{F}^{(k)}(t)$ $(k=1,\ldots,K)$ 表示仅对第 k 组资产有影响的因子. 假设因子过程满足

$$d\boldsymbol{F}^{(k)}(t)=(\boldsymbol{V}^{(k)}(t))^{\mathrm{T}}d\boldsymbol{W}^{(k)}(t),\quad k=0,\ldots,K,\tag{2.2}$$

其中, $\boldsymbol{W}^{(k)}(t):=(W_1^{(k)}(t),\ldots,W_{r_k}^{(k)}(t))$ 是 r_k 维标准 Brown 运动, 且对不同的 k 独立; $\boldsymbol{V}^{(k)}(t)$ 是 $r_k\times r_k$ 维波动率矩阵. 同时, 假设异质过程 $\boldsymbol{U}(t)$ 满足

$$d\boldsymbol{U}(t)=\boldsymbol{\kappa}^{\mathrm{T}}(t)d\tilde{\boldsymbol{W}}(t)+d\boldsymbol{J}(t),\tag{2.3}$$

其中, $\tilde{\boldsymbol{W}}(t)$ 是与 $\boldsymbol{W}^{(k)}(t)$ 独立的 p 维标准 Brown 运动, $\boldsymbol{\kappa}(t)$ 为 $p\times p$ 维波动率矩阵, $\boldsymbol{J}(t)$ 为 p 维跳跃过程. 于是, 模型 (2.1)–(2.3) 将金融资产的对数价格分解为 3 部分: 反映市场整体的特征的 "共同因子"、反映组内特征的 "分组因子" 以及反映资产独有特征的异质项. 相应地, 资产的波动率特征也由这 3 部分共同刻画: 记资产对数价格 $\boldsymbol{X}(t)$ 在第 d 天连续项积分波动率矩阵为

$$\boldsymbol{\Psi}(d):=\boldsymbol{\tau}(d)+\boldsymbol{\Sigma}(d),$$

其中,

$$\boldsymbol{\tau}(d)=\sum_{k=0}^{K}\boldsymbol{\tau}^{(k)}(d)=\sum_{k=0}^{K}\boldsymbol{B}^{(k)}(d)\int_{d-1}^{d}(\boldsymbol{V}^{(k)}(t))^{\mathrm{T}}\boldsymbol{V}^{(k)}(t)dt(\boldsymbol{B}^{(k)}(d))^{\mathrm{T}}$$

为因子过程的积分波动率,

$$\boldsymbol{\Sigma}(d)=\int_{d-1}^{d}\boldsymbol{\kappa}^{\mathrm{T}}(t)\boldsymbol{\kappa}(t)dt$$

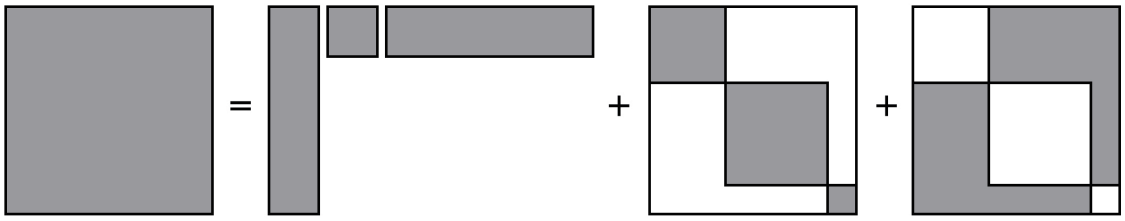
为异质过程的积分波动率. 进一步地, 不妨记第 i 个资产所属分组为 k_i, $\boldsymbol{\Sigma}^S(d)=\Sigma_{ij}(d)\mathbb{1}_{k_i=k_j}$. 则异质过程积分波动率可以按照组结构分成 $\boldsymbol{\Sigma}^S(d)$ 和 $\boldsymbol{\Sigma}(d)-\boldsymbol{\Sigma}^S(d)$ 两部分. 图 1 展示了连续项积分波动率 $\boldsymbol{\Phi}(d)$ 的具体分解形式.

为保证模型的可识别性, 接下来给出适当的假设. 为此, 对于任意 $k=0,\ldots,K$, 记 $\boldsymbol{B}^{(k)}(d)$ 的列向量为 $\boldsymbol{\beta}_j^{(k)}(d)$, $j=1,\ldots,r_k$.

假设 2.1 (1) 对于任意分组 $k=1,\ldots,K$ 以及任意资产 $i=1,\ldots,p$, 当 $k_i\neq k$ 时, $\beta_{i,j}^{(k)}(d)=0$, $j=1,\ldots,r_k$; 对于任意分组 $k=1,\ldots,K$ 以及 $j=1,\ldots,r_0$, 至少存在一个资产 i, 使得 $k_i=k$ 且 $\beta_{i,j}^{(0)}(d)\neq 0$.

(2) $\boldsymbol{B}^{(k)}(d)$ $(k=0,\ldots,K)$ 是 $\mathcal{F}_{d-1}$ 可测的, $\boldsymbol{\beta}_j^{(k)}(d)$ $(k=0,\ldots,K,\ j=1,\ldots,r_k)$ 互相正交, 且满足

$$\|\boldsymbol{\beta}_j^{(k)}(d)\|_2=1.$$



$\boldsymbol{\Psi}(d)$ $\quad$ $\boldsymbol{\tau}^{(0)}(d)$ $\quad$ $\sum_{k=1}^{K}\boldsymbol{\tau}^{(k)}(d)+\boldsymbol{\Sigma}^{S}(d)$ $\quad$ $\boldsymbol{\Sigma}(d)-\boldsymbol{\Sigma}^{S}(d)$

图 1 连续项积分波动率 $\boldsymbol{\Phi}(d)$ 的具体分解形式

(3) 对于任意 $k=0,\ldots,K$, $(\boldsymbol{V}^{(k)}(t))^{\mathrm{T}}\boldsymbol{V}^{(k)}(t)=\mathrm{diag}(\lambda_1^{(k)}(t),\ldots,\lambda_{r_k}^{(k)}(t))$, 且 $\lambda_1^{(k)}(t)>\cdots>\lambda_{r_k}^{(k)}(t)$.

假设 2.1(1) 说明了分组因子对应的因子载荷只在组内资产上非 0, 共同因子对应的因子载荷在每组上至少有 1 个资产非 0, 即分组因子仅影响同一组内的资产价格而共同因子对每一组资产的价格都会产生影响. 假设 2.1(2) 和 2.1(3) 则保证了因子本身是可以识别的. 在假设 2.1 的条件下, 因子载荷与积分波动率矩阵的特征向量具有一一对应关系, 通过下述命题给出:

命题 2.1 在假设 2.1 的条件下, $\boldsymbol{\beta}_j^{(k)}(d)$ $(k=0,\ldots,K,\ j=1,\ldots,r_k)$ 是 $\boldsymbol{\tau}(d)$ 的特征向量, 且对应的特征值为 $\Lambda_j^{(k)}(d):=\int_{d-1}^{d}\lambda_j^{(k)}(t)dt$, $k=0,\ldots,K$, $j=1,\ldots,r_k$.

在定义了资产价格的分组因子模型之后, 进一步定义因子波动率的混频结构, 从而提出 GARCH-Itô 分组因子模型.

定义 2.1 (GARCH-Itô 分组因子模型) 若一组资产的对数价格 $\boldsymbol{X}(t):=(X_1(t),\ldots,X_p(t))^{\mathrm{T}}$ 满足 GARCH-Itô 分组因子模型, 则对于任意 $k=0,\ldots,K$, $j=1,\ldots,r_k$, 当 $t\in[d-1,d)$ 时, $\lambda_j^{(k)}(t)$ 具有参数形式 $\lambda_j^{(k)}(t):=\lambda_j^{(k)}(t,\boldsymbol{\theta})$, 且具有如下结构:

$$d(\boldsymbol{\beta}_j^{(k)}(d))^{\mathrm{T}}\boldsymbol{X}(t)=(\boldsymbol{\beta}_j^{(k)}(d))^{\mathrm{T}}\boldsymbol{\mu}dt+\sqrt{\lambda_j^{(k)}(t,\boldsymbol{\theta})}dW_j^{(k)}(t)+d(\boldsymbol{\beta}_j^{(k)}(d))^{\mathrm{T}}\boldsymbol{U}(t),$$

其中,

$$\begin{cases}\lambda_j^{(0)}(t,\boldsymbol{\theta})=\lambda_j^{(0)}([t],\boldsymbol{\theta})+(t-[t])\{pw_j^{(0)}+(\gamma_j^{(0)}-1)\lambda_j^{(0)}([t],\boldsymbol{\theta})\}\\ \qquad\qquad+\sum\limits_{l=1}^{r_0}\eta_{j,l}^{(0)}\Big\{\int_{[t]}^{t}\sqrt{\lambda_l^{(0)}(s,\boldsymbol{\theta})}dW_j^{(0)}(s)\Big\}^2,\\ \lambda_j^{(k)}(t,\boldsymbol{\theta})=\lambda_j^{(k)}([t],\boldsymbol{\theta})+(t-[t])\{p_kw_j^{(k)}+(\gamma_j^{(k)}-1)\lambda_j^{(k)}([t],\boldsymbol{\theta})\}\\ \qquad\qquad+\sum\limits_{l=1}^{r_0}\eta_{j,l}^{(k)}\Big\{\int_{[t]}^{t}\sqrt{\lambda_l^{(0)}(s,\boldsymbol{\theta})}dW_j^{(0)}(s)\Big\}^2\\ \qquad\qquad+\sum\limits_{l=r_0+1}^{r_0+r_k}\eta_{j,l}^{(k)}\Big\{\int_{[t]}^{t}\sqrt{\lambda_l^{(k)}(s,\boldsymbol{\theta})}dW_j^{(k)}(s)\Big\}^2,\quad k=1,\ldots,K,\end{cases}\tag{2.4}$$

这里, 当 t 不为整数时, $[t]$ 表示 t 的整数部分, 当 t 为整数时, $[t]=t-1$; $\boldsymbol{\theta}:=(\boldsymbol{w},\boldsymbol{\gamma},\boldsymbol{\eta})$ 表示模型的参数.

注 2.1 (2.4) 中共同因子受其他共同因子往期的影响, 分组因子受同组其他分组因子及共同因子往期的影响. 而由于不同因子之间协积分波动率为 0, 即在当期上, 共同因子之间、分组因子之间以

及共同因子和分组因子之间都是不相关的, 我们可以将这种不相关性理解为在时间横截面上的不相关性, 不与 (2.4) 的动态变化规律相冲突.

注 2.2 若记 $\boldsymbol{\eta}_j^{(0)} := (\eta_{j,1}^{(0)}, \ldots, \eta_{j,r_0}^{(0)})$, $\boldsymbol{\eta}_{j,A}^{(k)} := (\eta_{j,1}^{(k)}, \ldots, \eta_{j,r_0}^{(k)})$, $\boldsymbol{\eta}_{j,B}^{(k)} := (\eta_{j,r_0+1}^{(k)}, \ldots, \eta_{j,r_0+r_k}^{(k)})$, $k = 1, \ldots, K$, 且 $\boldsymbol{\eta}^{(0)} := (\boldsymbol{\eta}_1^{(0)}, \ldots, \boldsymbol{\eta}_{r_0}^{(0)})^{\mathrm{T}}$, $\boldsymbol{\eta}_A^{(k)} := (\boldsymbol{\eta}_{1,A}^{(k)}, \ldots, \boldsymbol{\eta}_{r_k,A}^{(k)})^{\mathrm{T}}$, $\boldsymbol{\eta}_B^{(k)} := (\boldsymbol{\eta}_{1,B}^{(k)}, \ldots, \boldsymbol{\eta}_{r_k,B}^{(k)})^{\mathrm{T}}$, 则有

$$\boldsymbol{\eta} := \begin{pmatrix} \boldsymbol{\eta}^{(0)} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \boldsymbol{\eta}_A^{(1)} & \boldsymbol{\eta}_B^{(1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\eta}_A^{(K)} & \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{\eta}_B^{(K)} \end{pmatrix}. \tag{2.5}$$

则可将 (2.4) 改写为矩阵形式

$$\boldsymbol{\lambda}(t, \boldsymbol{\theta}) = \boldsymbol{\lambda}([t], \boldsymbol{\theta}) + (t - [t])\{\boldsymbol{p}\boldsymbol{w} + (\boldsymbol{\gamma} - I_{r^*})\boldsymbol{\lambda}([t], \boldsymbol{\theta})\} + \boldsymbol{\eta} \int_{[t]}^{t} \boldsymbol{N}(s, \boldsymbol{\theta}) d\boldsymbol{W}(s) \odot \int_{[t]}^{t} \boldsymbol{N}(s, \boldsymbol{\theta}) d\boldsymbol{W}(s),$$

其中,

$$\begin{aligned} &\boldsymbol{w} = (w_1^{(0)}, \ldots, w_{r_0}^{(0)}, \ldots, w_{r_K}^{(K)})^{\mathrm{T}}, \quad \boldsymbol{\gamma} = \mathrm{diag}(\gamma_1^{(0)}, \ldots, \gamma_{r_0}^{(0)}, \ldots, \gamma_{r_K}^{(K)}), \\ &\boldsymbol{\lambda}(t, \boldsymbol{\theta}) = (\lambda_1^{(0)}(\boldsymbol{t}, \boldsymbol{\theta}), \ldots, \lambda_{r_0}^{(0)}(t, \boldsymbol{\theta}), \ldots, \lambda_{r_K}^{(K)}(t, \boldsymbol{\theta}))^{\mathrm{T}}, \quad \boldsymbol{N}(t, \boldsymbol{\theta}) = \mathrm{diag}(\sqrt{\boldsymbol{\lambda}(t, \boldsymbol{\theta})}), \\ &\boldsymbol{W}(t) = (\boldsymbol{W}^{(0)}(t), \ldots, \boldsymbol{W}^{(K)}(t))^{\mathrm{T}}, \quad \boldsymbol{p} = \mathrm{diag}(p, \ldots, p, p_1, \ldots, p_K). \end{aligned}$$

注 2.3 当 $K = 0$ 即不考虑分组因子时, 且当 $\boldsymbol{J}(t) = \mathbf{0}$ 即不考虑跳跃项时, 模型退化为 Kim 和 Fan [38] 提出的 GARCH-Itô 因子模型.

假设 2.2 (平稳性假设) 假设参数矩阵 $\boldsymbol{\eta}$ (定义参见 (2.5)) 可逆, 且满足 $\|\boldsymbol{\eta}\|_2 < 1$. 对于任意 $k = 1, \ldots, K$, $j = 1, \ldots, r_k$, $0 < \gamma_j^{(k)} < 1$. 进一步地, 记 $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta}^{-1}(\mathrm{e}^{\boldsymbol{\eta}} - I_{\tilde{r}} - \boldsymbol{\eta})(\mathrm{diag}(\boldsymbol{\gamma}) - I_{\tilde{r}}) + \mathrm{e}^{\boldsymbol{\eta}} - I_{\tilde{r}}$, 则

$$\|\tilde{\boldsymbol{\eta}} + \mathrm{diag}(\boldsymbol{\gamma})\|_2 < 1$$

且 $I_{\tilde{r}} - \tilde{\boldsymbol{\eta}} - \mathrm{diag}(\boldsymbol{\gamma})$ 可逆.

假设 2.2 可以看作文献 [39, 命题 1] 中条件的推广. 引入了这一假设之后, 可以对因子积分波动率矩阵的特征值 $\Lambda_j^{(k)}(d, \boldsymbol{\theta})$ 进行进一步的分解. 记 $\boldsymbol{\Lambda}(d, \boldsymbol{\theta}) = \int_{d-1}^{d} \boldsymbol{\lambda}(t, \boldsymbol{\theta}) dt$, 有如下命题成立.

命题 2.2 在假设 2.1 和 2.2 的条件下, 对于 $d = 1, \ldots, n$,

$$\boldsymbol{\Lambda}(d, \boldsymbol{\theta}) = \boldsymbol{h}(d, \boldsymbol{\theta}) + \boldsymbol{D}(d) \quad \text{且} \quad \mathrm{E}\{\boldsymbol{\Lambda}(d, \boldsymbol{\theta}) \mid \mathcal{F}_{d-1}\} = \boldsymbol{h}(d, \boldsymbol{\theta}) \quad \text{a.s.},$$

其中 $\boldsymbol{h}(d, \boldsymbol{\theta}) = \boldsymbol{p}\boldsymbol{\eta}^{-2}(\mathrm{e}^{\boldsymbol{\eta}} - I_{\tilde{r}} - \boldsymbol{\eta})\boldsymbol{w} + \boldsymbol{\eta}^{-1}\tilde{\boldsymbol{\eta}}\boldsymbol{\lambda}(d-1, \boldsymbol{\theta})$.

注 2.4 假设 2.2 保证了序列 $\boldsymbol{h}(d, \boldsymbol{\theta})$ 的平稳性. 根据上述命题 2.2, 易知 $\boldsymbol{h}(d, \boldsymbol{\theta})$ 是 $\mathcal{F}_{d-1}$ 可测的, 且在 $\mathcal{F}_{d-1}$ 的条件下, $\boldsymbol{h}(d, \boldsymbol{\theta})$ 与 $\boldsymbol{\Lambda}(d, \boldsymbol{\theta})$ 的条件均值相等. 从而, 当我们将观测时刻限制在低频上, 即 $t = 1, \ldots, n$ 时, 模型因子具有向量广义自回归条件异方差 (vector generalized autoregressive conditional heteroskedasticity, VGARCH) 结构

$$\begin{aligned} \mathrm{E}\{\boldsymbol{\Lambda}(d, \boldsymbol{\theta}) \mid \mathcal{F}_{d-1}\} = {} & \boldsymbol{\eta}^{-1}(\mathrm{e}^{\boldsymbol{\eta}} - I_{\tilde{r}})\boldsymbol{w} + \boldsymbol{\gamma}\mathrm{E}\{\boldsymbol{\Lambda}(d-1, \boldsymbol{\theta}) \mid \mathcal{F}_{d-2}\} \\ & + \tilde{\boldsymbol{\eta}} \int_{d-2}^{d-1} \boldsymbol{N}(s) d\boldsymbol{W}(s) \odot \int_{d-2}^{d-1} \boldsymbol{N}(s) d\boldsymbol{W}(s) \quad \text{a.s.} \end{aligned}$$

GARCH-Itô 分组因子模型假设了资产价格由一个分组因子模型所驱动, 资产的波动率受到共同因子、分组因子和异质项的影响. 进一步地, 因子过程对应的瞬时波动率的特征 $\lambda_j^{(k)}(t,\boldsymbol{\theta})$ 被假设具有高低频数据相结合的 GARCH-Itô 参数化结构. 具体而言, 在整数时间点的低频信息上模型因子服从 VGARCH 结构, 而在整数时间点之间的高频信息上模型是一个连续时间 Itô 过程. 通过将高频信息嵌入低频的 GARCH 模型中, GARCH-Itô 分组因子模型能够抓住各组资产的波动率聚集效应, 从而对波动率矩阵进行良好的预测. 在我们的模型刻画中, 共同因子和分组因子过程对应的波动率的参数化动态表示并不相同, 我们假设共同因子对应的波动率只受共同因子的影响, 这与 Kim 和 Fan [38] 的假设相同; 而分组因子对应的波动率只受共同因子和此组对应的分组因子影响, 不受不同组的分组因子的影响. 由于漂移项 $\boldsymbol{\mu}$ 对估计高频已实现波动率的影响可以忽略不计 (参见文献 [39]), 为便于分析, 后文中取漂移项 $\boldsymbol{\mu}=\mathbf{0}$.

3 参数估计

本节提出 GARCH-Itô 分组因子模型的伪极大似然估计, 讨论相应的理论性质, 并提出基于本模型的资产对数价格积分波动率矩阵的估计与预测方法. 假设共有 n 个低频观测时刻, 每两个低频观测时刻之间存在着 m 个高频观测时刻. 记 n 个低频观测时刻为 $t_{1,0},\ldots,t_{n,0}$, $t_{d,0}$ 与 $t_{d+1,0}$ 之间的高频观测时刻为 $t_{d,0}<t_{d,1}<\cdots<t_{d,m}<t_{d,m+1}=t_{d+1,0}$. 假设观测是等间隔的, 并令 $t_{d,0}=d-1$, $d=1,\ldots,n$, 且记高频观测间隔 $\Delta:=t_{d,g}-t_{d,g-1}=1/m$. 假设市场微观结构噪声存在, 且与资产潜在对数价格 $\boldsymbol{X}$ 相互独立, 则在 $t_{d,g}$ 时刻观测到的资产的真实对数价格 $\boldsymbol{Y}(t_{d,g})$ 满足

$$\boldsymbol{Y}(t_{d,g})=\boldsymbol{X}(t_{d,g})+\boldsymbol{\epsilon}(t_{d,g}),$$

其中, 噪声项 $\boldsymbol{\epsilon}(t_{d,g})$ 具有 0 均值, 且对于不同的资产 i 相互独立.

3.1 伪极大似然估计

记 $\hat{\boldsymbol{\Psi}}(d)$ 为资产过程 $\boldsymbol{X}(t)$ 在第 d 天的积分波动率矩阵 $\boldsymbol{\Psi}(d)$ 的估计, $\hat{\Lambda}_j^{(k)}(d)$ $(k=0,\ldots,K)$ 为 $\hat{\boldsymbol{\Psi}}(d)$ 的第 k 组特征值, $\hat{\boldsymbol{\beta}}_j^{(k)}$ $(j=1,\ldots,r_k)$ 为对应的 r_k 个特征向量. 根据 (2.4) 可知, 因子积分波动率矩阵 $\boldsymbol{\tau}(d)$ 的特征值 $\Lambda_j^{(k)}(d,\boldsymbol{\theta})$ 有如下迭代关系:

$$\begin{cases}\Lambda_j^{(0)}(d,\boldsymbol{\theta})=\dfrac{pw_j^{(0)}+(\gamma_j^{(0)}+1)\lambda_j^{(0)}(d-1,\boldsymbol{\theta})}{2}+\displaystyle\sum_{l=1}^{r_0}\eta_{j,l}^{(0)}\rho_l^{(0)}(d),\\ \Lambda_j^{(k)}(d,\boldsymbol{\theta})=\dfrac{p_kw_j^{(k)}+(\gamma_j^{(k)}+1)\lambda_j^{(k)}(d-1,\boldsymbol{\theta})}{2}+\displaystyle\sum_{l=1}^{r_0}\eta_{j,l}^{(k)}\rho_l^{(0)}(d)+\sum_{l=r_0+1}^{r_0+r_k}\eta_{j,l}^{(k)}\rho_l^{(k)}(d),\quad k=1,\ldots,K,\end{cases}$$

其中,

$$\begin{aligned}\rho_l^{(k)}(d):=\int_{d-1}^{d}\Bigg\{&\int_t^t d(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\boldsymbol{X}(s)-\int_{[t]}^t d(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\boldsymbol{J}(s)\\&-\int_{[t]}^t(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\boldsymbol{\kappa}^{\mathrm{T}}(s)d\tilde{\boldsymbol{W}}(s)\Bigg\}^2dt,\quad k=0,\ldots,K.\end{aligned}$$

由于潜在对数价格 $\boldsymbol{X}(t)$ 是不可观测的, 我们用真实对数价格 $\boldsymbol{Y}(t_{d,g})$ 来作为 $t_{d,g}$ 时刻的潜在对数价格 $\boldsymbol{X}(t_{d,g})$ 的估计, 并估计出资产对数价格第 d 天累计到 $t_{d,g}$ 时刻的跳跃 $\boldsymbol{JV}(t_{d,g})$ 以及第 d 天的累

计跳跃 $\boldsymbol{JV}(d)$. 结合 Euler 法, 将积分离散化, 得到特征值 $\Lambda_j^{(k)}(d,\boldsymbol{\theta})$ 的近似:

$$\begin{cases}\tilde{\Lambda}_j^{(0)}(d,\boldsymbol{\theta})=\dfrac{pw_j^{(0)}+(\gamma_j^{(0)}+1)\hat{\lambda}_j^{(0)}(d-1,\boldsymbol{\theta})}{2}+\displaystyle\sum_{l=1}^{r_0}\eta_{j,l}^{(0)}\hat{\rho}_l^{(0)}(d),\\ \tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})=\dfrac{p_kw_j^{(k)}+(\gamma_j^{(k)}+1)\hat{\lambda}_j^{(k)}(d-1,\boldsymbol{\theta})}{2}+\displaystyle\sum_{l=1}^{r_0}\eta_{j,l}^{(k)}\hat{\rho}_l^{(0)}(d)+\sum_{l=r_0+1}^{r_0+r_k}\eta_{j,l}^{(k)}\hat{\rho}_l^{(k)}(d),\quad k=1,\ldots,K,\end{cases}$$

其中, $\hat{\rho}_l^{(k)}(d)=\sum_{g=1}^m[(\hat{\boldsymbol{\beta}}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)-\boldsymbol{JV}(t_{d,g})\}]^2\Delta$, $k=0,\ldots,K$, 且

$$\begin{cases}\hat{\lambda}_j^{(0)}(d,\boldsymbol{\theta})=pw_j^{(0)}+\gamma_j^{(0)}\hat{\lambda}_j^{(0)}(d-1,\boldsymbol{\theta})+\displaystyle\sum_{l=1}^{r_0}\eta_{j,l}^{(0)}((\hat{\boldsymbol{\beta}}_l^{(0)}(d))^{\mathrm{T}}\{\boldsymbol{Y}(d)-\boldsymbol{Y}(d-1)-\boldsymbol{JV}(d)\})^2,\\ \hat{\lambda}_j^{(k)}(d,\boldsymbol{\theta})=p_kw_j^{(k)}+\gamma_j^{(k)}\hat{\lambda}_j^{(k)}(d-1,\boldsymbol{\theta})+\displaystyle\sum_{l=1}^{r_0}\eta_{j,l}^{(k)}((\hat{\boldsymbol{\beta}}_l^{(0)}(d))^{\mathrm{T}}\{\boldsymbol{Y}(d)-\boldsymbol{Y}(d-1)-\boldsymbol{JV}(d)\})^2\\ \qquad\qquad+\displaystyle\sum_{l=r_0+1}^{r_0+r_k}\eta_{j,l}^{(k)}((\hat{\boldsymbol{\beta}}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{Y}(d)-\boldsymbol{Y}(d-1)-\boldsymbol{JV}(d)\})^2,\quad k=1,\ldots,K.\end{cases}\tag{3.1}$$

从而, 我们提出 GARCH-Itô 分组因子模型对应的伪极大似然函数如下:

定义 3.1 伪极大似然函数 $\hat{\boldsymbol{L}}_n(\boldsymbol{\theta})$ 满足

$$\hat{\boldsymbol{L}}_n(\boldsymbol{\theta})=\sum_{k=0}^K\sum_{j=1}^{r_k}\hat{L}_{n,j}^{(k)}(\boldsymbol{\theta}),$$

其中,

$$\hat{L}_{n,j}^{(k)}(\boldsymbol{\theta})=-\frac{1}{2n}\sum_{d=1}^n\left(\log\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})+\frac{\hat{\Lambda}_j^{(k)}(d)}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}\right),\quad k=0,\ldots,K,\quad j=1,\ldots,r_k.$$

记 $\hat{\boldsymbol{\theta}}$ 为极大化伪极大似然函数 $\hat{\boldsymbol{L}}_n(\boldsymbol{\theta})$ 得到的参数估计, 即

$$\hat{\boldsymbol{\theta}}=\arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\hat{\boldsymbol{L}}_n(\boldsymbol{\theta}).\tag{3.2}$$

为方便起见, 下文将该估计量称为 GARCH-Itô 估计量, 将基于该估计量实现的积分波动率矩阵及其特征值的估计 (预测) 都称为基于 GARCH-Itô 方法的估计 (预测).

3.2 极限理论

假设 3.1 对于整数 $a\geqslant 1$,

(1) $\max_{1\leqslant d\leqslant n}\mathrm{E}(\|\hat{\boldsymbol{\Psi}}(d)-\boldsymbol{\Psi}(d)\|_F^a)\leqslant Cp^am^{-a/4}$, $\max_{1\leqslant d\leqslant n}\mathrm{E}(\|\boldsymbol{JV}(d)-\int_{d-1}^d d\boldsymbol{J}(t)\|_2^a)\leqslant Cp^{a/2}m^{-a/4}$, $\max_{1\leqslant d\leqslant n}\mathrm{E}(\|\boldsymbol{JV}(t_{d,g})-\int_{d-1}^{t_{d,g}}d\boldsymbol{J}(t)\|_2^a)\leqslant Cp^{a/2}m^{-a/4}$;

(2) $\max_{1\leqslant d\leqslant n}\mathrm{E}(\|\hat{\boldsymbol{\Psi}}(d)\|_F^a)\leqslant Cp^a$, $\max_{1\leqslant d\leqslant n}\mathrm{E}(\|\boldsymbol{\Psi}(d)\|_F^a)\leqslant Cp^a$, $\max_{1\leqslant d\leqslant n}\mathrm{E}(\|\boldsymbol{\epsilon}(t_{d,g})\|_\infty)\leqslant C$, $\max_{1\leqslant d\leqslant n}\mathrm{E}(\|\boldsymbol{JV}(d)\|_2^a)\leqslant Cp^{a/2}$, $\max_{1\leqslant d\leqslant n}\mathrm{E}(\|d\boldsymbol{J}(t)\|_2^a)\leqslant Cp^{a/2}$;

(3) $\min_{j'\neq j}\{\Lambda_j^{(k)}(d)-\Lambda_{j'}^{(k)}(d)\}\geqslant Cp_k$, $k=0,\ldots,K$;

(4) 存在常数 $b\in(0,1/2)$, 使得 $p_k=O_p(p^{1/2-b})$, $k=1,\ldots,K$.

假设 3.1(1) 要求积分波动率的估计和累计跳跃项的估计的收敛速度为 $O_p(m^{-1/4})$, 我们选取的估计量 MSRV [17] 和跳跃项估计量 $\boldsymbol{JV}(d)$ [26] 都满足该假设. 假设 3.1(2) 中 $d\boldsymbol{J}(t)$ 是跳跃项的微分. 假设 3.1(3) 是近似因子结构的假设, 其中 $p_0 = p$. 我们在后文证明中需要利用文献 [38, 定理 3], 而此定理的证明需要假设 3.1(3). 假设 3.1(4) 要求每组资产个数的量级是相当的, 且隐含了组数随着资产个数的增长而增加的条件, 在后文特征向量的渐近正交性证明中需要用到此假设.

假设 3.2 (稀疏性假设) 异质过程的第 d 天积分波动率矩阵 $\boldsymbol{\Sigma}(d)$ 满足稀疏条件

$$\max_{1\leqslant d\leqslant n, 1\leqslant i\leqslant p}\sum_{j=1}^{p}|\Sigma_{ij}(d)|^{\delta}\{\Sigma_{ii}(d)\Sigma_{jj}(d)\}^{(1-\delta)/2}\leqslant M_{\sigma}s_p,$$

其中, $\delta\in[0,1)$, M_σ 是有界正随机变量, s_p 是随 p 变化的稀疏测度且 $s_p = o(p^{1/2-b})$, 不妨假设 $s_p = O(\log p)$, 且满足对于整数 $\alpha\geqslant 1$,

$$\max_{1\leqslant d\leqslant n}\mathrm{E}(\|\boldsymbol{\kappa}^{\mathrm{T}}(d)\boldsymbol{\kappa}(d)\|_2^{3\alpha})\leqslant Cs_p^{3\alpha}.$$

假设 3.2 与文献 [38, 第 3.1 小节] 中模型结构相同. 此假设允许异质项对应的不同资产之间存在相关性, 但只能存在较弱的相关性. 其目的是控制异质项对资产价格刻画的影响, 使得资产价格变化主要由因子项驱动.

定理 3.1 若假设 2.1、2.2、3.1 和 3.2 成立, 且 $n\geqslant 2+4\tilde{r}$, $b>0$, $m\to\infty$, $m^{-1/2}\log p\to 0$, 则伪似然估计 $\hat{\boldsymbol{\theta}}$ 收敛到真实参数 $\boldsymbol{\theta}^*$:

$$\|\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^*\|_{\max} = O_p(m^{-1/4}(\log p)^{1/2}+p^{-b}+(s_p)^{1/2}p^{b/2-1/4}).$$

定理 3.1 指出, $\hat{\boldsymbol{\theta}}$ 的收敛速度受到高频观测数 m、资产数 p 及异质过程波动矩阵的稀疏测度 s_p 所影响.

注 3.1 当不存在分组时, 通过拟似然估计的参数以 $O_p(m^{-1/4}+(s_p/p)^{1/2})$ 的速度收敛到真实参数.

注 3.2 定理 3.1 是在初值 $\hat{\lambda}_j^{(k)}(0,\boldsymbol{\theta})$ 与真实的初值相等的情形下推导的. 在实际应用中, 一般先给出初值的估计, 再估计模型参数 $\boldsymbol{\theta}$. 此时, 初始值选取的偏差对参数估计会造成一定的影响, 具体体现在参数的收敛速度需要增加额外的 n^{-1} 项. 该结论已在文献 [39] 中有类似的讨论, 故在此不作详述, 感兴趣的读者可查阅相关文献.

3.3 算法实现

本小节给出伪极大似然法的具体算法实现.

第 1 步 利用多维情形的 MSRV 算法估计波动率矩阵 $\boldsymbol{\Psi}(d)$. 具体地, 首先通过阈值法估计跳跃项, 并得到剔除跳跃项的值 $\tilde{\boldsymbol{Y}}(t_{d,g})$ (参见文献 [26]); 接着采用下式得到估计量 $\hat{\boldsymbol{\Psi}}_{\mathrm{MSRV}}(d)$:

$$\hat{\boldsymbol{\Psi}}_{\mathrm{MSRV}}(d)=\sum_{g=1}^{M}a_g[\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Y}}]_d^{(H_g)}+\varsigma\{[\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Y}}]_d^{(H_1)}-[\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Y}}]_d^{(H_M)}\},$$

其中, M 是一个整数, 且与 $m^{1/2}$ 同阶; $H_g = g+C$, C 是 $m^{1/2}$ 的整数部分;

$$a_g=\frac{12(g+C)(g-M/2-1/2)}{M(M^2-1)},$$

$$\varsigma = \frac{(M+C)(C+1)}{(m+1)(M-1)},$$
$$[\tilde{\boldsymbol{Y}}, \tilde{\boldsymbol{Y}}]_d^{(H)} = \frac{1}{H}\sum_{g=1}^{m-H}\{\tilde{\boldsymbol{Y}}(t_{d,g+H}) - \tilde{\boldsymbol{Y}}(t_{d,g})\}\{\tilde{\boldsymbol{Y}}(t_{d,g+H}) - \tilde{\boldsymbol{Y}}(t_{d,g})\}^{\mathrm{T}}.$$

第 2 步 计算共同因子特征值与特征向量: 对 $\hat{\boldsymbol{\Psi}}_{\mathrm{MSRV}}(d)$ 进行特征分解, 选取前 r_0 个特征值作为共同因子对应的特征值 $\hat{\Lambda}_j^{(0)}(d)$, 并得到相应的特征向量 $\hat{\boldsymbol{\beta}}_j^{(0)}(d)$. 从而得到共同因子对应的波动率矩阵 $\boldsymbol{\tau}^{(0)}(d)$ 的估计

$$\hat{\boldsymbol{\tau}}^{(0)}(d) = \sum_{j=1}^{r_0}\hat{\Lambda}_j^{(0)}(d)\hat{\boldsymbol{\beta}}_j^{(0)}(d)(\hat{\boldsymbol{\beta}}_j^{(0)}(d))^{\mathrm{T}}.$$

第 3 步 计算分组因子特征值和特征向量: 将残差阵 $\hat{\boldsymbol{\varepsilon}}(d) := \hat{\boldsymbol{\Psi}}_{\mathrm{MSRV}}(d) - \hat{\boldsymbol{\tau}}_0(d)$ 按组结构进行分块, 得到 $\hat{\boldsymbol{\varepsilon}}^S(d) = \hat{\varepsilon}_{ij}(d)\mathbb{1}_{k_i=k_j}$. 并逐块进行特征值分解, 得到各组因子对应的特征值 $\hat{\Lambda}_j^{(k)}(d)$, $j = 1, \ldots, r_k$, $k = 1, \ldots, K$. 将特征向量按组结构补 0, 扩充成 p 维的向量, 作为分组因子对应特征向量的估计 $\hat{\boldsymbol{\beta}}_j^{(k),S}(d)$.

第 4 步 利用上述特征值和特征向量, 根据 (3.2) 得到参数的伪极大似然估计.

上述算法实现的前提是确定共同因子的个数 r_0 和分组因子的个数 r_k, $k = 1, \ldots, K$. 然而, 在实际问题中, 因子的个数往往是未知的. 本文基于 Aït-Sahalia 和 Xiu [36] 的研究, 提出因子个数的估计

$$\hat{r}_k = \underset{1\leqslant j\leqslant R_k}{\arg\min}\sum_{d=1}^{n}(p_k^{-1}\hat{\Lambda}_j^{(k)}(d) + jg_k(m,p)) - 1, \quad k = 0, \ldots, K,$$

其中, $g_k(m,p)$ 是与 m 和 p 有关的惩罚函数, 用来防止因子个数的过度选取. 在应用中, 应选取适当的 $g_k(m,p)$, 使得对于 $j \leqslant r_k$, 保证 $p_k^{-1}\hat{\Lambda}_j^{(k)}(d)$ 占据主导地位, 而在 $j > r_k$ 时, $jg_k(m,p)$ 占据主导, 从而正确地估计出 r_k. R_k 为一个较大的常数, 在本文中取为 10. 定理 3.2 通过对惩罚函数 $g_k(m,p)$ 施加具体限制, 得到此估计量的良好性质.

定理 3.2 当假设 3.2 成立, 且 $p_k = o(p)$, $m^{-1/2}\log p \to 0$, $g_0(m,p) \to 0$, $g_k(m,p) \to 0$, $g_0(m,p)(m^{-1/4}(\log p)^{1/2} + p_k/p)^{-1} \to \infty$, $g_k(m,p)(m^{-1/4}(\log p_k)^{1/2} + s_p/p_k)^{-1} \to \infty$ 时,

$$\mathrm{P}(\hat{r}_k = r_k) \to 1, \quad k = 0, \ldots, K.$$

具体地, 本文选取

$$g_0(m,p) = 0.02\hat{\Lambda}_{10}^{(0)}(d)\left(m^{-1/4}(\log p)^{1/2} + \frac{p_k}{p}\right)^{1/2},$$
$$g_k(m,p) = 0.02\hat{\Lambda}_{10}^{(k)}(d)\left(m^{-1/4}(\log p_k)^{1/2} + \frac{s_p}{p_k}\right)^{1/2},$$

易见 $g_0(m,p)$ 和 $g_k(m,p)$ 的选取满足定理 3.2 的条件.

在一般的因子分析中, 因子往往被要求具有正交性. 本文提出的模型也遵循这一条件. 从算法实现的角度来看, 根据模型假设与特征值分解的算法性质, 易知上述算法满足: (1) 同组内的分组因子对应的特征向量相互正交; (2) 不同组的分组因子对应的特征向量相互正交. 然而, 共同因子与分组因子对应特征向量之间的正交性并不显然, 下述定理 3.3 给出了相关结论.

定理 3.3 当假设 3.1 和 3.2 成立, 且 $m \to \infty$, $p \to \infty$, $p^b m^{-1/4} \to 0$ 时, $\hat{\boldsymbol{\beta}}_j^{(k),S}(d)$ 与 $\hat{\boldsymbol{\beta}}_{j'}^{(0)}(d)$ $(k=1,\ldots,K)$ 存在渐近正交性,

$$(\hat{\boldsymbol{\beta}}_j^{(k),S}(d))^{\mathrm{T}}\hat{\boldsymbol{\beta}}_{j'}^{(0)}(d) = O_p(p^b m^{-1/4} + (\log p)^{1/2} m^{-1/4} + p^{-b} + s_p p^{b-1/2}), \quad k=1,\ldots,K.$$

由定理 3.3 可知, 如果希望得到共同因子与分组因子对应特征向量之间的渐近正交性, 模型需要更强的假设, 要求高频观测数 m 具有资产个数 p 的多项式阶的增长速度, 这比定理 3.1 (参数估计的一致性) 和 3.2 (因子个数估计的一致性) 的条件更强. 这相当于要求更大的高频样本量.

3.4 积分波动率矩阵的估计与预测

在得到模型参数的估计 $\hat{\boldsymbol{\theta}}$ 之后, 基于该模型, 我们可以得到各因子特征值的估计 $\tilde{\Lambda}_j^{(k)}(d,\hat{\boldsymbol{\theta}})$ 以及积分波动率矩阵的估计 $\hat{\boldsymbol{\Psi}}(d) = \hat{\boldsymbol{\tau}}(d) + \hat{\boldsymbol{\Sigma}}(d)$. 则异质过程的积分波动率估计 $\hat{\boldsymbol{\Sigma}}(d)$ 满足

$$\hat{\Sigma}_{ij}(d) = \begin{cases} \hat{\hat{\Sigma}}_{ij}(d) \vee 0, & i=j, \\ s_{ij}\{\hat{\hat{\Sigma}}_{ij}(d)\}\mathbb{1}_{\{|\hat{\hat{\Sigma}}_{ij}(d)| \geqslant \varpi_{ij}\}}, & i \neq j, \end{cases}$$

其中,

$$\hat{\hat{\boldsymbol{\Sigma}}}(d) = \hat{\boldsymbol{\Psi}}_{\mathrm{MSRV}}(d) - \sum_{k=0}^{K}\sum_{j=1}^{r_k} \hat{\Lambda}_j^{(k)}(d)\hat{\boldsymbol{\beta}}_j^{(k),S}(d)(\hat{\boldsymbol{\beta}}_j^{(k),S}(d))^{\mathrm{T}};$$

$\hat{\boldsymbol{\beta}}_j^{(0),S}(d) := \hat{\boldsymbol{\beta}}_j^{(0)}(d)$; $s_{ij}(\cdot)$ 为阈值函数, 满足 $|s_{ij}(x) - x| \leqslant \varpi_{ij}$,

$$\varpi_{ij} = \varpi_m \sqrt{(\hat{\hat{\Sigma}}_{ii}(d) \vee 0)(\hat{\hat{\Sigma}}_{jj}(d) \vee 0)},$$

ϖ_m 为与样本量有关的常数, 具体选取方法可参见文献 [44, 定理 3]. 这样, 我们构造的估计量 $\hat{\boldsymbol{\Sigma}}(d)$ 满足稀疏性假设 3.2.

进一步地, 我们对因子特征值及积分波动率矩阵进行预测. 根据命题 2.2 可知, 在给定第 d 天的信息 $\mathcal{F}_d$ 的情形下, $\boldsymbol{h}(d+1,\boldsymbol{\theta})$ 是第 $d+1$ 天的特征值 $\boldsymbol{\Lambda}(d+1,\boldsymbol{\theta})$ 的条件均值, 因此, $\boldsymbol{h}(d+1,\boldsymbol{\theta})$ 的估计

$$\hat{\boldsymbol{h}}(d+1,\hat{\boldsymbol{\theta}}) = \boldsymbol{p}\hat{\boldsymbol{\eta}}^{-2}(\mathrm{e}^{\hat{\boldsymbol{\eta}}} - I_{\tilde{r}} - \hat{\boldsymbol{\eta}})\hat{\boldsymbol{w}} + \hat{\boldsymbol{\eta}}^{-1}\tilde{\boldsymbol{\eta}}\hat{\boldsymbol{\lambda}}(d,\hat{\boldsymbol{\theta}})$$

可以作为第 $d+1$ 天因子特征值的良好预测. 其中, $\hat{\boldsymbol{\lambda}}(d-1,\hat{\boldsymbol{\theta}})$ 为 $\boldsymbol{\lambda}(d-1,\hat{\boldsymbol{\theta}})$ 的近似, 其定义由 (3.1) 给出. 从而, 我们可以利用 $\hat{\boldsymbol{h}}(d+1,\hat{\boldsymbol{\theta}})$ 来预测 $d+1$ 天的积分波动率矩阵

$$\begin{aligned} \tilde{\boldsymbol{\Psi}}(d+1) :&= \tilde{\boldsymbol{\tau}}(d+1) + \tilde{\boldsymbol{\Sigma}}(d+1) \\ &= \sum_{k=0}^{K}\sum_{j=1}^{r_k} \hat{h}_j^{(k)}(d+1,\hat{\boldsymbol{\theta}})\hat{\boldsymbol{\beta}}_j^{(k),S}(d)(\hat{\boldsymbol{\beta}}_j^{(k),S}(d))^{\mathrm{T}} + \tilde{\boldsymbol{\Sigma}}(d+1), \end{aligned}$$

其中, $\tilde{\boldsymbol{\tau}}(d+1)$ 表示第 $d+1$ 天因子过程积分波动率矩阵的预测, 将 $\hat{\boldsymbol{h}}(d+1,\hat{\boldsymbol{\theta}})$ 与第 d 天的特征向量 $\hat{\boldsymbol{\beta}}_j^{(k),S}(d)$ 结合即可求得; $\tilde{\boldsymbol{\Sigma}}(d+1)$ 表示第 $d+1$ 天异质过程积分波动率矩阵的预测, 这需要我们对异质项波动率矩阵施加额外的假设. 例如, 若 $\boldsymbol{\Sigma}(d)$ 为一个鞅过程, 可得到 $\mathrm{E}\{\boldsymbol{\Sigma}(d+1) \mid \mathcal{F}_d\} = \boldsymbol{\Sigma}(d)$. 此时可以用第 d 天的估计值 $\hat{\boldsymbol{\Sigma}}(d)$ 进行预测.

4 模拟研究

4.1 生成数据

在模拟研究中, 假设共有 $p=200$ 个资产, 且组数 $K=2$, 每一组资产的个数为 $p_1=p_2=100$. 假设低频观测天数 $n=100$, 分别考虑高频观测时刻数 $m=200$ 和 $m=1{,}500$ 两种情形. 固定观测时刻 $t_{d,g}=d-1+g/m$, $d=1,\ldots,n$, $g=1,\ldots,m$, 利用本文提出的 GARCH-Itô 分组因子模型, 分别在相应时刻上生成模型的因子过程、异质过程, 从而得到资产潜在对数价格 $\boldsymbol{X}(t_{d,g})$.

对于因子过程, 假设共有 $\tilde{r}=6$ 个因子驱动模型, 其中共同因子数 $r_0=2$, 分组因子数 $r_1=r_2=2$, 模型真实参数 $\boldsymbol{\theta}^*$ 为 $(\boldsymbol{w}^{(0)})^*=(0.3,0.2)$, $(\boldsymbol{w}^{(1)})^*=(0.14,0.1)$, $(\boldsymbol{w}^{(2)})^*=(0.1,0.075)$, $(\boldsymbol{\gamma}^{(0)})^*=(0.2,0.175)$, $(\boldsymbol{\gamma}^{(1)})^*=(0.1,0.15)$, $(\boldsymbol{\gamma}^{(2)})^*=(0.05,0.11)$, $(\boldsymbol{\eta}_1^{(0)})^*=(0.25,0.2)$, $(\boldsymbol{\eta}_2^{(0)})^*=(0.18,0.15)$, $(\boldsymbol{\eta}_{1,A}^{(1)})^*=(0.18,0.06)$, $(\boldsymbol{\eta}_{1,B}^{(1)})^*=(0.18,0.14)$, $(\boldsymbol{\eta}_{2,A}^{(1)})^*=(0.06,0.01)$, $(\boldsymbol{\eta}_{2,B}^{(1)})^*=(0.13,0.1)$, $(\boldsymbol{\eta}_{1,A}^{(2)})^*=(0.14,0.02)$, $(\boldsymbol{\eta}_{1,B}^{(2)})^*=(0.12,0.13)$, $(\boldsymbol{\eta}_{2,A}^{(2)})^*=(0.04,0)$, $(\boldsymbol{\eta}_{2,B}^{(2)})^*=(0.08,0.09)$. $\boldsymbol{\lambda}(t,\boldsymbol{\theta})$ 的初始值为 $\boldsymbol{\lambda}^*(0,\boldsymbol{\theta})=(300,200,80,60,30,10)$. 根据 (2.4) 迭代生成 $\lambda_j^{(k)}(t_{d,g},\boldsymbol{\theta}^*)$. 进一步, 我们通过模拟研究发现 $\boldsymbol{\eta}_A$ 对不同组波动率差异的影响最大, $\boldsymbol{\eta}_B$ 和 $\boldsymbol{w}$ 次之, $\boldsymbol{\lambda}(0,\boldsymbol{\theta})$ 的影响最小, 具体结果见附录. 接着, 根据假设 2.1(1), 生成互相正交的 $\tilde{r}$ 个模长为 1 的特征向量 $\boldsymbol{\beta}_j^{(k)}(d)$.

对于异质过程的连续项, 生成异质项波动率矩阵 $\Sigma_{ij}=0.5^{|i-j|}\sqrt{\Sigma_{ii}\Sigma_{jj}}$, 其中 $\Sigma_{ii}=1$ $(i=1,\ldots,p)$, 再通过 Cholesky 分解得到 $\boldsymbol{\kappa}(t)$. 对于异质过程的跳跃项, 假设各观测时刻发生的跳跃是相互独立的, 跳跃发生的次数服从 Poisson 分布 $P(\lambda)$, $\lambda=1{,}000$. 对于第 i 个资产, 跳跃幅度为 $w_0\mathbb{1}_{w_0>0}+w_i$, 其中, w_0 服从正态分布 $N(0,0.3^2)$, $w_i=0.8\times(|\cos(i)|)$.

假设扰动项 $\boldsymbol{\epsilon}(t_{d,g})$ 服从正态分布 $N(0,10^{-7})$, 资产的真实对数价格 $\boldsymbol{Y}(t_{d,g})=\boldsymbol{X}(t_{d,g})+\boldsymbol{\epsilon}(t_{d,g})$. 在此参数设定下, 因子的平均累计贡献率为 62.5%, 这与我们的实证较为一致.

根据上述步骤生成的资产价格数据具有明显的组间差异. 记 $\hat{\boldsymbol{\Psi}}_{\mathrm{MSRV}}^{(1)}(d)$ 和 $\hat{\boldsymbol{\Psi}}_{\mathrm{MSRV}}^{(2)}(d)$ 分别为第 d 天第一组和第二组利用 MSRV 方法估计出的积分波动率矩阵, 则两组资产的平均积分波动率比值为 $n^{-1}\sum_{d=1}^{n}\mathrm{tr}\{\hat{\boldsymbol{\Psi}}_{\mathrm{MSRV}}^{(1)}(d)\}/\mathrm{tr}\{\hat{\boldsymbol{\Psi}}_{\mathrm{MSRV}}^{(2)}(d)\}=1.831$. 可以看出, 第一组的波动率显著地大于第二组, 这一特征与第 4.2 小节将要分析的实际数据相契合.

4.2 模拟结果

选取前 $n_0=90$ 天为训练集, 后 $n-n_0=10$ 天为测试集, 在训练集上进行参数估计, 重复实验 50 次.

表 1 是 50 次模拟中共同因子和分组因子个数正确估计比例. 随着日内数据的增加, 因子个数估计正确比例上升, 且当 m 取 1,500 时, 共同因子和分组因子个数估计全部正确. 这说明此估计因子个数的方法是有效的.

表 2 和 3 通过偏差的平方 (bias^2)、方差 (Var) 和均方误差 (MSE) 分别衡量模型中共同因子与分组因子相关参数的估计误差. 可以明显看出, 各参数的估计误差都在能够接受的范围之内, 且随着日

表 1 共同因子和分组因子个数正确估计比例

m	$\hat{r}_0$	$\hat{r}_1$	$\hat{r}_2$
200	1.00	0.96	0.84
1,500	1.00	1.00	1.00

表 2　GARCH-Itô 分组因子模型共同因子对应的参数估计误差 ($\times 10^{-4}$)

Class		m	$\hat{w}$	$\hat{\eta}^{(0)}_{\cdot,1}$	$\hat{\eta}^{(0)}_{\cdot,2}$	$\hat{\gamma}$
$\lambda_1^{(0)}$	bias^2	200	3.287	46.609	77.510	12.276
		1,500	3.742	13.164	13.059	0.863
	Var	200	16.452	5.936	16.787	50.244
		1,500	7.480	6.078	6.896	28.121
	MSE	200	19.739	52.544	94.297	62.520
		1,500	11.239	19.242	19.954	28.983
$\lambda_2^{(0)}$	bias^2	200	0.273	33.162	37.182	3.196
		1,500	1.997	5.206	5.094	0.033
	Var	200	10.622	2.295	7.482	53.733
		1,500	3.814	2.705	3.668	27.679
	MSE	200	10.894	35.457	44.664	56.929
		1,500	5.811	7.911	8.762	27.712

表 3　GARCH-Itô 分组因子模型分组因子对应的参数估计误差 ($\times 10^{-4}$)

Class		m	$\hat{w}$	$\hat{\eta}^{(\cdot)}_{\cdot,1}$	$\hat{\eta}^{(\cdot)}_{\cdot,2}$	$\hat{\eta}^{(\cdot)}_{\cdot,3}$	$\hat{\eta}^{(\cdot)}_{\cdot,4}$	$\hat{\gamma}$
$\lambda_1^{(1)}$	bias^2	200	5.029	144.147	2.674	140.663	98.792	26.584
		1,500	12.325	28.989	0.256	53.340	25.487	11.846
	Var	200	6.072	0.878	1.621	6.564	7.288	26.206
		1,500	3.111	1.778	1.075	6.952	7.642	14.803
	MSE	200	11.101	145.025	4.295	147.227	106.080	52.791
		1,500	15.436	30.767	1.331	60.292	33.129	26.649
$\lambda_2^{(1)}$	bias^2	200	1.922	7.687	1.379	87.700	49.415	4.566
		1,500	2.476	0.712	1.073	11.406	1.061	0.596
	Var	200	3.322	0.461	0.780	2.139	4.701	38.500
		1,500	1.590	0.571	0.889	3.269	5.713	15.670
	MSE	200	5.245	8.148	2.178	89.836	54.116	43.061
		1,500	4.066	1.283	1.962	14.675	6.774	16.265
$\lambda_1^{(2)}$	bias^2	200	2.947	72.343	0.659	26.605	91.476	87.316
		1,500	3.358	10.127	0.153	0.285	11.539	7.309
	Var	200	5.315	0.802	0.797	6.286	5.289	33.954
		1,500	1.528	1.095	0.904	6.378	15.317	14.925
	MSE	200	8.262	73.145	1.456	32.890	96.765	121.270
		1,500	4.886	11.222	1.057	6.663	26.856	22.234
$\lambda_2^{(2)}$	bias^2	200	2.352	2.124	1.951	17.389	40.700	2.605
		1,500	0.861	0.088	0.597	0.288	0.366	1.117
	Var	200	3.052	0.286	0.577	2.847	3.363	31.336
		1,500	0.779	0.359	0.565	2.002	6.204	30.277
	MSE	200	5.404	2.410	2.528	20.237	44.063	33.941
		1,500	1.640	0.447	1.162	2.290	6.570	31.394

内数据的增加, 误差大幅度减小. 这说明模型具有良好的有限样本性质.

进一步地, 我们在测试集上对模型的一步预测效果进行评估. 记模型积分波动率矩阵各特征值在第 d 天的相对预测误差为

$$\mathrm{err}_j^{(k)}(d) := \frac{|\hat{h}_j^{(k)}(d,\hat{\boldsymbol{\theta}}) - \Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)|}{\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)}, \tag{4.1}$$

其中 $\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)$ 是因子积分波动率矩阵特征值的真值. 表 4 展示了在不同观测频率下, 在 50 次重复实验中利用本文的 GARCH-Itô 方法得到的各特征值在测试集 $d = n_0+1,\ldots,n$ 上的平均相对预测误差. 我们同时引入了基于 MSRV 方法的特征值预测作为对比, 即利用 MSRV 估计出第 $d-1$ 天的积分波动率矩阵 $\hat{\boldsymbol{\Psi}}_{\mathrm{MSRV}}(d-1)$ 并进行相应的特征值分解, 得到第 $d-1$ 天的共同因子与分组因子对应的特征值 $\hat{\Lambda}_j^{(k)}(d-1)$ 作为第 d 天特征值的预测. 这相当于忽略了数据的 GARCH-Itô 结构, 仅利用分组信息来预测数据的波动率. 作为高频非参数方法, 可以预见 MSRV 的短期预测能力将高于长期. 因此, 我们同时考虑短期 MSRV 预测, 即以 "小时" 为单位 (而不是 "天") 进行向前一步的预测. 假设一个交易日有 4 个小时, 在短期 MSRV 中, 我们选取短期高频数据量为 $m/4$. 根据表 4, 我们发现, 随着日内数据的增加, 利用 GARCH-Itô 方法得到的因子积分波动率对应特征值的平均相对预测误差显著减小, 并且在所有情形下, GARCH-Itô 方法的结果都远好于 MSRV 方法.

为比较分组结构与不分组结构的预测误差, 我们另外构建一个不分组模型, 即假设所有的因子都为共同因子, 并选取与分组模型贡献率相同的因子个数 $r_0 = \tilde{r} = 6$. 在这种情形下, 由于分组模型与不分组模型的因子含义不尽相同, 除前两个共同因子之外, 无法直接比较其预测效果, 但由因子所解释的资产价格整体的波动率 (即因子特征值的和) 是可以相比的. 因此, 定义由部分因子组成的集合 $\mathcal{I}_0$ 解释的波动率在第 d 期的相对预测误差为

$$\mathrm{err}_{\mathcal{I}_0}(d) := \frac{|\sum_{(k,j)\in\mathcal{I}_0}\{\hat{h}_j^{(k)}(d,\hat{\boldsymbol{\theta}}) - \Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)\}|}{\sum_{(k,j)\in\mathcal{I}_0}\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)}, \tag{4.2}$$

其中, $\mathcal{I}_0$ 为因子指标集, $\sum_{(k,j)\in\mathcal{I}_0}\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)$ 代表由指标集 $\mathcal{I}_0$ 中的因子所解释的各资产对数价格的

表 4 GARCH-Itô 方法和 MSRV 方法的因子积分波动率对应特征值的平均相对预测误差, 基于 50 次重复实验的均值 (标准差)

Class	m	GARCH-Itô	MSRV	短期 MSRV
$\lambda_1^{(0)}$	200	0.173 (0.056)	0.279 (0.100)	0.251 (0.044)
	1,500	0.129 (0.028)	0.221 (0.077)	0.220 (0.031)
$\lambda_2^{(0)}$	200	0.167 (0.063)	0.250 (0.088)	0.234 (0.069)
	1,500	0.140 (0.030)	0.245 (0.080)	0.220 (0.031)
$\lambda_1^{(1)}$	200	0.292 (0.103)	0.336 (0.094)	0.442 (0.137)
	1,500	0.197 (0.041)	0.329 (0.085)	0.249 (0.035)
$\lambda_2^{(1)}$	200	0.258 (0.096)	0.288 (0.073)	0.421 (0.130)
	1,500	0.170 (0.039)	0.287 (0.070)	0.236 (0.030)
$\lambda_1^{(2)}$	200	0.283 (0.086)	0.362 (0.113)	0.350 (0.119)
	1,500	0.229 (0.056)	0.369 (0.123)	0.336 (0.057)
$\lambda_2^{(2)}$	200	0.257 (0.072)	0.331 (0.104)	0.262 (0.100)
	1,500	0.194 (0.055)	0.293 (0.099)	0.426 (0.078)

表 5 3 种方法的因子积分波动率的平均相对预测误差, 基于 50 次重复实验的均值 (标准差)

m	分组	不分组	不分组且忽略跳跃项
200	0.227 (0.087)	0.378 (0.109)	2.232 (0.620)
1,500	0.173 (0.037)	0.310 (0.057)	0.802 (0.380)

积分波动率之和. 在这里, 由于分组与不分组模型的前两个因子是相同的, 我们主要考虑分组因子组成的集合的预测效果. 因此, 取 $\mathcal{I}_0 = \{(1,1),(1,2),(2,1),(2,2)\}$ 表示分组因子的指标集. 对于不分组模型, 用指标集 $\mathcal{I}_1 = \{(0,3),(0,4),(0,5),(0,6)\}$ 代替 $\mathcal{I}_0$, 表示后 4 个共同因子的指标集.

同时, 选取不分组且忽略跳跃项的方法 (即 Kim 和 Fan [38] 提出的方法) 作为对照组. 表 5 展示了 3 种方法在日内高频观测数分别为 $m = 200$ 和 $m = 1{,}500$ 时的因子积分波动率平均相对预测误差. 随着日内数据的增加, 预测误差降低, 且本文提出的方法明显优于其他 2 种方法. 这说明, 当数据存在分组结构时, 分组模型能够明显地提高波动率的预测精度. 同时, 跳跃项对模型的影响不容忽视, 因此在下文的实证分析中首先剔除数据的跳跃项.

5 实证分析

本节选取波动率较低的大规模公司股票与波动率较高的小规模公司股票, 构造具有分组结构的资产组合, 利用 GARCH-Itô 分组因子模型, 对其波动率矩阵进行分析预测. 本节的数据来源于 Wind 数据库. 我们从上证主板与深证创业板各选取 50 支股票, 取其每分钟的对数收盘价, 从而对于每一个交易日, 得到 $m = 237$ 个高频观测. 数据的时间跨度为 2019 年 1 月 2 日至 2019 年 12 月 31 日, 共计 $n = 244$ 个交易日. 这 50 支上证主板股和 50 支深证创业板股的平均股本分别为 93.5 亿和 8.3 亿, 体现了 "大规模公司 - 小规模公司" 的分组结构. 同时, 这两组股票对数价格的平均积分波动率比值为 0.425, 反映了小规模公司股票往往具有相对较高的波动率的特点, 进一步说明分组分析的必要性.

在拟合 GARCH-Itô 分组因子模型之前, 需要确定共同因子与分组因子的个数. 选取与模拟中相同的方法, 共同因子和分组因子个数的估计为 $\hat{r}_0 = 2$ 和 $\hat{r}_1 = \hat{r}_2 = 3$. 此时这 8 个因子在 244 个交易日的平均累计贡献率达到 66.8%. 图 2 展示了特征值相对大小的碎石图. 选取前 9 个月 181 个交易日为训练集, 后 3 个月 63 个交易日为测试集, 在训练集上最大化伪极大似然函数 $\hat{\boldsymbol{L}}_n(\boldsymbol{\theta})$, 得到参数的估计值 $\hat{\boldsymbol{\theta}}$. 与模拟研究类似, 我们评估模型的一步预测效果, 同时与 MSRV 方法进行比较. 不同之处在于, 此时特征值的真值未知, 我们用 MSRV 方法估计的波动率矩阵 $\hat{\boldsymbol{\Psi}}_{\mathrm{MSRV}}(d)$ 的对应特征值 $\hat{\Lambda}_j^{(k)}(d)$ 代替 (4.1) 中的 $\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)$. 图 3 展示了在测试集上应用本文提出的 GARCH-Itô 方法和 MSRV 方法得到的共同因子和分组因子的各特征值的相对预测误差. 根据图 3, 我们发现, 相比 MSRV 方法, 对于所有的特征值, 本文提出的 GARCH-Itô 方法都有更低的平均相对预测误差. 这说明在实际数据中, 资产对数价格的因子过程在一定程度上存在着 GARCH 结构, 即波动率聚集的特性.

为了说明分组分析的必要性, 我们将分组模型与不分组的模型进行比较. 为使两模型具有可比性, 在不分组模型中, 考虑因子数 $r = 5$ 和 $r = 6$, 对应着因子平均累计贡献率分别为 65.9% 和 70.9%. 根据 (4.2), 我们可以比较因子组成的集合 $\mathcal{I}_0$ 所解释的波动率在测试集上的平均预测误差. 其中, 我们利用 $\hat{\boldsymbol{\Phi}}_{\mathrm{MSRV}}(d)$ 分解得到的特征值 $\hat{\Lambda}_j^{(k)}(d)$ 代替未知真实值 $\Lambda_j^{(k)}(d,\theta^*)$. 首先比较所有因子的总体预测效果, 即将 $\mathcal{I}_0$ 取为因子全集. 另外, 分组因子模型的两个共同因子实际上与不分组因子模型的前两个因子相同, 这是由模型的定义和算法决定的; 而对于两类模型中不同的因子, 即分组模型中指标属于 $\mathcal{I}_0 = \{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3)\}$ 的分组因子, 与不分组模型中指标属于

$\tilde{\mathcal{I}}_0 = \{(0,3),(0,4),(0,5)\}$ $(r=5)$ 或 $\tilde{\mathcal{I}}_0 = \{(0,3),(0,4),(0,5),(0,6)\}$ $(r=6)$ 的因子, 两者的预测效果不同. 故我们比较了 I_0 和 $\tilde{I}_0$ 下因子的预测效果. 进一步, 选取只考虑高频数据的 MSRV 方法和小时预测的短期 MSRV 方法作为对照组. 表 6 展示了这 4 种方法在全体因子上和提出共同因子上的因子积分波动率的平均相对预测误差. 本文提出的方法的预测误差低于其余 3 种方法, 且本文提出的方法的优势在剔除共同因子影响后更加明显, 这更能说明我们的分组是有效的.

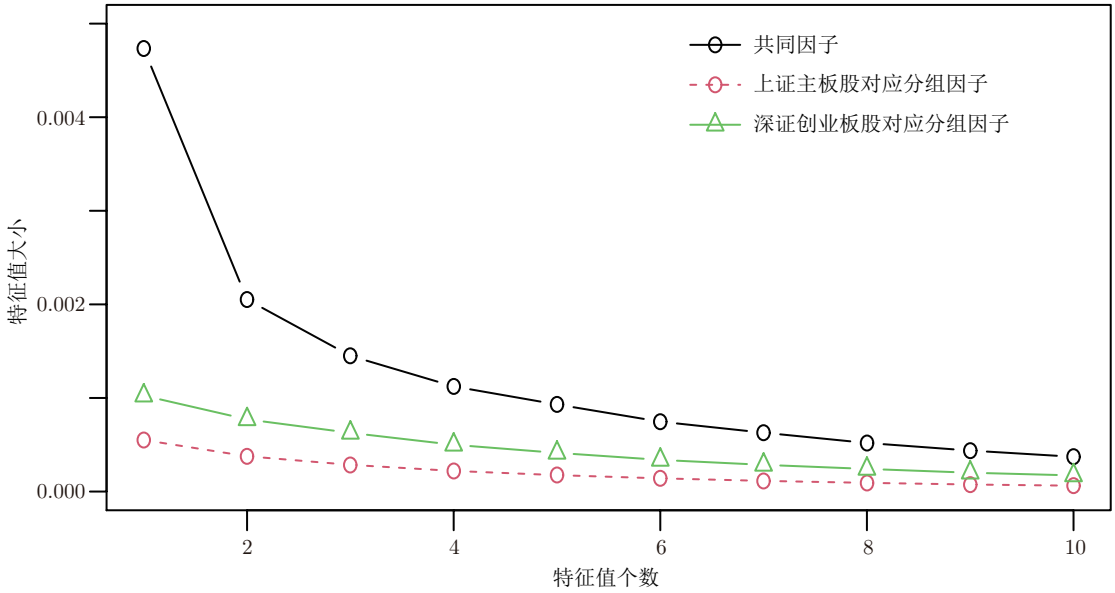


图 2 (网络版彩图) 特征值的相对大小的碎石图

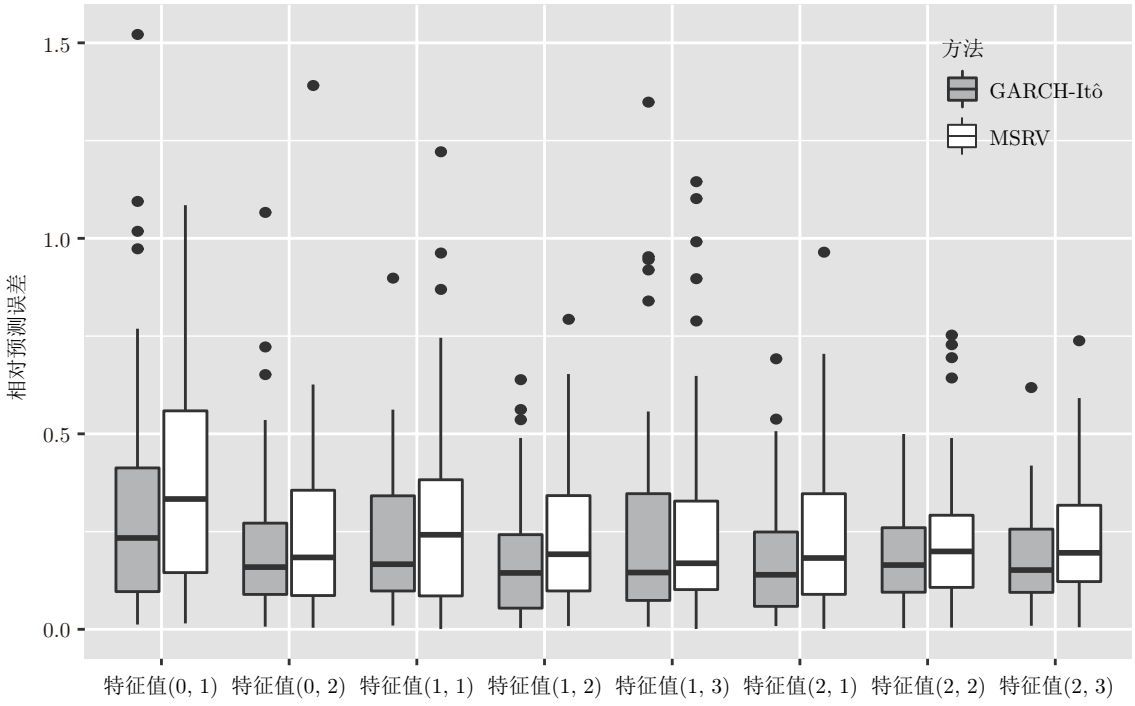


图 3 GARCH-Itô 方法和 MSRV 方法在共同因子和分组因子上积分波动率对应特征值的相对预测误差

表 6 4 种方法的因子积分波动率的平均相对预测误差

	分组	不分组 ($r=5$)	不分组 ($r=6$)	MSRV	短期 MSRV
全体因子	0.186	0.191	0.187	0.236	0.218
剔除共同因子	0.138	0.194	0.220	0.173	0.259

6 总结

在分析高维、高频金融资产数据时, 因子模型能够有效地降低数据的维数, 提取蕴含数据特征的潜在因子; 而 GARCH-Itô 模型则通过将高频数据嵌入低频 GARCH 模型中, 来刻画波动率的聚集效应, 从而提高模型的解释力与预测效果. 当资产具有明显的分组结构时, 各组资产的特征不尽相同, 此时利用分组建模的思想, 更能发挥出因子模型与 GARCH-Itô 模型的优势. 鉴于此, 本文提出了一种新的分组数据的高维、高频数据波动率矩阵的分析方法: GARCH-Itô 分组因子模型. 该模型在引入了共同因子的同时也引入了分组因子, 从而刻画分组效应. 我们提出了模型参数的伪极大似然估计, 并证明了估计量的极限性质. 在模拟研究中, 我们发现该模型具有良好的有限样本性质. 在实证研究中, 我们将公司规模作为分组的依据, 选取了大规模公司和小规模公司的股票资产作为分析对象, 建立 GARCH-Itô 分组因子模型. 通过与积分波动率矩阵的 MSRV 高频估计量以及不分组模型的比较, 我们发现分组模型具有更好的波动率预测能力. 在实际数据中, 金融资产间特征的差异可以体现在方方面面, 从而能够产生多种不同的分组方式, 此时如何确定最为合适的分组, 或比较不同分组的优劣, 都将是我们未来研究的方向.

致谢 特别感谢审稿人对本文提出的宝贵意见.

参考文献

1 Lu L X. Global asset allocation strategy of Chinese investors (in Chinese). Finance Economy, 2008, 8: 104–106 [陆黎雪. 中国投资者的全球资产配置策略. 金融经济, 2008, 8: 104–106]
2 Gao H, Zhao J W. Empirical research for hedge ratio and shares portfolio of Shanghai-Shenzhen 300 shares index futures (in Chinese). J Manag Sci, 2007, 2: 80–90 [高辉, 赵进文. 沪深 300 股指套期保值及投资组合实证研究. 管理科学, 2007, 2: 80–90]
3 Bandi F M, Russell J R, Zhu Y. Using high-frequency data in dynamic portfolio choice. Econometric Rev, 2008, 27: 163–198
4 Cai T T, Hu J, Li Y, et al. High-dimensional minimum variance portfolio estimation based on high-frequency data. J Econometrics, 2020, 214: 482–494
5 Cont R. Statistical modeling of high-frequency financial data. IEEE Signal Process Mag, 2011, 28: 16–25
6 Sharpe W F. Capital asset prices: A theory of market equilibrium under conditions of risk. J Finance, 1964, 19: 425–442
7 Fama E F, French K R. The cross-section of expected stock returns. J Finance, 1992, 47: 427–465
8 Fama E F, French K R. A five-factor asset pricing model. J Financ Econ, 2015, 116: 1–22
9 Lehmann B N, Modest D M. The empirical foundations of the arbitrage pricing theory. J Financ Econ, 1988, 21: 213–254
10 Connor G, Korajczyk R A. Risk and return in an equilibrium APT: Application of a new test methodology. J Financ Econ, 1988, 21: 255–289
11 Bai J, Ng S. Evaluating latent and observed factors in macroeconomics and finance. J Econometrics, 2006, 131: 507–537
12 Goyal A, Perignon C, Villa C. How common are common return factors across the NYSE and Nasdaq? J Financ Econ, 2008, 90: 252–271
13 Bekaert G, Hodrick R J, Zhang X. International stock return comovements. J Finance, 2009, 64: 2591–2626
14 Ando T, Bai J. Multifactor asset pricing with a large number of observable risk factors and unobservable common and

group-specific factors. https://mpra.ub.uni-muenchen.de/52785/1/MPRA_paper_52785.pdf
15 Zhang L, Mykland P A, Aït-Sahalia Y. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. J Amer Statist Assoc, 2005, 100: 1394–1411
16 Zhang L. Estimating covariation: Epps effect, microstructure noise. J Econometrics, 2011, 160: 33–47
17 Zhang L. Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. Bernoulli, 2006, 12: 1019–1043
18 Barndorff-Nielsen O E, Hansen P R, Lunde A, et al. Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. Econometrica, 2008, 76: 1481–1536
19 Aït-Sahalia Y, Fan J, Xiu D. High-frequency covariance estimates with noisy and asynchronous financial data. J Amer Statist Assoc, 2010, 105: 1504–1517
20 Xiu D. Quasi-maximum likelihood estimation of volatility with high frequency data. J Econometrics, 2010, 159: 235–250
21 Jacod J, Li Y, Mykland P A, et al. Microstructure noise in the continuous case: The pre-averaging approach. Stochastic Process Appl, 2009, 119: 2249–2276
22 Aït-Sahalia Y, Jacod J, Li J. Testing for jumps in noisy high frequency data. J Econometrics, 2012, 168: 207–222
23 Andersen T G, Bollerslev T, Diebold F X. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. Rev Econ Stat, 2007, 89: 701–720
24 Barndorff-Nielsen O E, Shephard N. Econometrics of testing for jumps in financial economics using bipower variation. J Financ Econ, 2005, 4: 1–30
25 Corsi F, Pirino D, Renò R. Threshold bipower variation and the impact of jumps on volatility forecasting. J Econometrics, 2010, 159: 276–288
26 Fan J, Wang Y. Multi-scale jump and volatility analysis for high-frequency financial data. J Amer Statist Assoc, 2007, 102: 1349–1362
27 Wang Y. Jump and sharp cusp detection by wavelets. Biometrika, 1995, 82: 385–397
28 Zhang X, Kim D, Wang Y. Jump variation estimation with noisy high frequency financial data via wavelets. Econometrics, 2016, 4: 1–26
29 Zhao H, Qin K J. Jumps in stock prices and macro information release (in Chinese). Statist Res, 2014, 31: 79–89 [赵华, 秦可佶. 股价跳跃与宏观信息发布. 统计研究, 2014, 31: 79–89]
30 Wu B, Zhang B. Trading data, jump detection and estimation of integrated volatility (in Chinese). Statist Res, 2017, 34: 109–119 [吴奔, 张波. 交易信息、跳跃发现与波动率估计. 统计研究, 2017, 34: 109–119]
31 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 [穆燕, 周勇. 带跳高频数据下高维积分波动率矩阵估计. 中国科学: 数学, 2020, 50: 1455–1486]
32 Wang Y, Zou J. Vast volatility matrix estimation for high-frequency financial data. Ann Statist, 2010, 38: 943–978
33 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
34 Tao M, Wang Y, Zhou H H. Optimal sparse volatility matrix estimation for high-dimensional Itô processes with measurement errors. Ann Statist, 2013, 41: 1816–1864
35 Fan J, 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
36 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
37 Pelger M. Large-dimensional factor modeling based on high-frequency observations. J Econometrics, 2019, 208: 23–42
38 Kim D, Fan J. Factor GARCH-Itô models for high-frequency data with application to large volatility matrix prediction. J Econometrics, 2019, 208: 395–417
39 Kim D, Wang Y. Unified discrete-time and continuous-time models and statistical inferences for merged low-frequency and high-frequency financial data. J Econometrics, 2016, 194: 220–230
40 Zou H M, Ouyang L N. Demonstration and analysis of "size effect" in Chinese stock market (in Chinese). J East China Jiaotong Univ (Soc Sci), 2004, 21: 67–70, 74 [邹洪梅, 欧阳令南. 中国股市 "规模效应" 的实证综述及原因分析. 华东交通大学学报, 2004, 21: 67–70, 74]
41 Zhou W, Li Y A. Return correlation, small company effect and market efficiency: An empirical test of Shanghai and Shenzhen stock markets (in Chinese). Modern Economic Sci, 1999, 1: 3–5 [周文, 李友爱. 收益率相关性、小公司效应与市场有效性—对沪、深股票市场的实证检验. 当代经济科学, 1999, 1: 3–5]
42 Wang W, Zhou Y. An empirical study on the size effect and time effect of the stock market in China: Based on Shanghai stock market (in Chinese). Economic Res J, 2002, 10: 16–21, 30, 94 [汪炜, 周宇. 中国股市 "规模效应" 和

"时间效应" 的实证分析—以上海股票市场为例. 经济研究, 2002, 10: 16–21, 30, 94]
43 Yang X, Chen Z H. Empirical studies on three factors asset pricing model in China's stock market (in Chinese). J Quant Tech Econ, 2003, 12: 137–141 [杨炘, 陈展辉. 中国股市三因子资产定价模型实证研究. 数量经济技术经济研究, 2003, 12: 137–141]
44 Fan J, Kim D. Robust high-dimensional volatility matrix estimation for high-frequency factor model. J Amer Statist Assoc, 2018, 113: 1268–1283

附录 A 定理的证明

在给出证明前, 首先给出需要用到的符号表示:

记 $\mathrm{Vec}(\boldsymbol{A}) = (A_{11}, \ldots, A_{1n}, A_{2,1}, \ldots, A_{mn})^{\mathrm{T}}$ 为矩阵 $\boldsymbol{A} = (A_{ij})_{m\times n}$ 的按行拉直向量; 记 $\|\boldsymbol{a}\|_p = (\sum_{i=1}^n |a_i|^p)^{1/p}$ 为向量 $\boldsymbol{a}$ 的 p- 范数, 其中 $p \geqslant 1$.

附录 A.1 命题 2.2 的证明

我们有

$$\begin{aligned}\boldsymbol{\tau}(d) &= \boldsymbol{B}^{(0)}(d)\int_{d-1}^{d}(\boldsymbol{V}^{(0)}(t))^{\mathrm{T}}\boldsymbol{V}^{(0)}(t)dt(\boldsymbol{B}^{(0)}(d))^{\mathrm{T}} + \sum_{k=1}^{K}\boldsymbol{B}^{(k)}(d)\int_{d-1}^{d}(\boldsymbol{V}^{(k)}(t))^{\mathrm{T}}\boldsymbol{V}^{(k)}(t)dt(\boldsymbol{B}^{(k)}(d))^{\mathrm{T}}\\ &= \sum_{j=1}^{r_0}\int_{d-1}^{d}\lambda_j^{(0)}(t)dt\boldsymbol{\beta}_j^{(0)}(d)(\boldsymbol{\beta}_j^{(0)}(d))^{\mathrm{T}} + \sum_{k=1}^{K}\sum_{j=1}^{r_k}\int_{d-1}^{d}\lambda_j^{(k)}(t)dt\boldsymbol{\beta}_j^{(k)}(d)(\boldsymbol{\beta}_j^{(k)}(d))^{\mathrm{T}}.\end{aligned}$$

由假设 2.1(3) 可知, 对于任意 $k = 0, \ldots, K$, $j = 1, \ldots, r_k$, $(\boldsymbol{\beta}_j^{(k)}(d))^{\mathrm{T}}\boldsymbol{\beta}_j^{(k)}(d) = 1$; 而对于任意 $j' \neq j$ 或 $k' \neq k$, $(\boldsymbol{\beta}_j^{(k)}(d))^{\mathrm{T}}\boldsymbol{\beta}_{j'}^{(k')}(d) = 0$. 由此可知, 对于任意 $k = 0, \ldots, K$, $j = 1, \ldots, r_k$,

$$\boldsymbol{\tau}(d)\boldsymbol{\beta}_j^{(k)}(d) = \int_{d-1}^{d}\lambda_j^{(k)}(t)dt\boldsymbol{\beta}_j^{(k)}(d).$$

这说明, $\boldsymbol{\beta}_1^{(0)}(d), \ldots, \boldsymbol{\beta}_{r_0}^{(0)}(d), \boldsymbol{\beta}_1^{(1)}(d), \ldots, \boldsymbol{\beta}_{r_K}^{(K)}(d)$ 是 $\boldsymbol{\tau}(d)$ 的特征向量, 且对应的特征值为 $\int_{d-1}^{d}\lambda_1^{(0)}(t)dt$, $\ldots, \int_{d-1}^{d}\lambda_{r_0}^{(0)}(t)dt, \int_{d-1}^{d}\lambda_1^{(1)}(t)dt, \ldots, \int_{d-1}^{d}\lambda_{r_K}^{(K)}(t)dt$.

附录 A.2 定理 3.1 的证明

引理 A.1 若 $\boldsymbol{a}$、$\tilde{\boldsymbol{a}}$、$\boldsymbol{b}$ 和 $\tilde{\boldsymbol{b}}$ 是 p 维向量, 且 4 阶矩存在, 则

$$\mathrm{E}(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}\times\tilde{\boldsymbol{a}}^{\mathrm{T}}\tilde{\boldsymbol{b}}) \leqslant \mathrm{E}(\|\boldsymbol{a}\|_2^4)^{1/4}\mathrm{E}(\|\boldsymbol{b}\|_2^4)^{1/4}\mathrm{E}(\|\tilde{\boldsymbol{a}}\|_2^4)^{1/4}\mathrm{E}(\|\tilde{\boldsymbol{b}}\|_2^4)^{1/4}.$$

证明 由 Hölder 不等式可知

$$\begin{aligned}\mathrm{E}(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}\times\tilde{\boldsymbol{a}}^{\mathrm{T}}\tilde{\boldsymbol{b}}) &\leqslant \mathrm{E}(\{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}\}^2)^{1/2}\mathrm{E}(\{\tilde{\boldsymbol{a}}^{\mathrm{T}}\tilde{\boldsymbol{b}}\}^2)^{1/2} \leqslant \mathrm{E}(\|\boldsymbol{a}\|_2^2\|\boldsymbol{b}\|_2^2)^{1/2}\mathrm{E}(\|\tilde{\boldsymbol{a}}\|_2^2\|\tilde{\boldsymbol{b}}\|_2^2)^{1/2}\\ &\leqslant \mathrm{E}(\|\boldsymbol{a}\|_2^4)^{1/4}\mathrm{E}(\|\boldsymbol{b}\|_2^4)^{1/4}\mathrm{E}(\|\tilde{\boldsymbol{a}}\|_2^4)^{1/4}\mathrm{E}(\|\tilde{\boldsymbol{b}}\|_2^4)^{1/4}.\end{aligned}$$

证毕. □

引理 A.2 在假设 2.1、3.1 和 3.2 的条件下,

$$\mathrm{E}(|\hat{\rho}_l^{(0)}(d) - \rho_l^{(0)}(d)|) \leqslant Cp\left\{m^{-1/4} + \left(\frac{s_p}{p}\right)^{1/2}\right\} \quad \text{a.s.}, \quad l = 1, \ldots, r_0,$$

$$\mathrm{E}(|\hat{\rho}_l^{(k)}(d)-\rho_l^{(k)}(d)|)\leqslant C\{m^{-1/4}(\log p)^{1/2}p^{1/2-b}+p^{1/2-2b}+(s_p)^{1/2}p^{1/4-b/2}\}\quad \text{a.s.},\quad l=1,\ldots,r_k,\quad k=1,\ldots,K.$$

证明 记 $\boldsymbol{JJ}(t_{d,g})=\boldsymbol{J}(t_{d,g})-\boldsymbol{J}(d-1)$.

令 $\hat{\rho}_l^{(k)}(d)-\rho_l^{(k)}(d)=(a1)+(b1)+(c1)$, 其中

$$\begin{aligned}(a1)&=\Delta\sum_{g=1}^{m}((\hat{\boldsymbol{\beta}}_l^{(k),S}(d))^{\mathrm{T}}\{\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)-\boldsymbol{JV}(t_{d,g})\})^2\\&\quad-((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{X}(t_{d,g})-\boldsymbol{X}(d-1)-\boldsymbol{JJ}(t_{d,g})\})^2,\\(b1)&=\Delta\sum_{g=1}^{m}((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{X}(t_{d,g})-\boldsymbol{X}(d-1)-\boldsymbol{JJ}(t_{d,g})\})^2\\&\quad-\Delta\sum_{g=1}^{m}((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{X}(t_{d,g})-\boldsymbol{X}(d-1)-\boldsymbol{JJ}(t_{d,g})-\boldsymbol{\kappa}^{\mathrm{T}}(t_{d,g})\tilde{\boldsymbol{W}}(t_{d,g})+\boldsymbol{\kappa}^{\mathrm{T}}(d-1)\tilde{\boldsymbol{W}}(d-1)\})^2,\\(c1)&=\Delta\sum_{g=1}^{m}((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{X}(t_{d,g})-\boldsymbol{X}(d-1)-\boldsymbol{JJ}(t_{d,g})-\boldsymbol{\kappa}^{\mathrm{T}}(t_{d,g})\tilde{\boldsymbol{W}}(t_{d,g})+\boldsymbol{\kappa}^{\mathrm{T}}(d-1)\tilde{\boldsymbol{W}}(d-1)\})^2\\&\quad-\int_{d-1}^{d}((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{X}(t)-\boldsymbol{X}(d-1)-\boldsymbol{J}(t)+\boldsymbol{J}(d-1)-\boldsymbol{\kappa}^{\mathrm{T}}(t)\tilde{\boldsymbol{W}}(t)+\boldsymbol{\kappa}^{\mathrm{T}}(d-1)\tilde{\boldsymbol{W}}(d-1)\})^2dt.\end{aligned}$$

对于 $(a1)$, 容易得出 $(a1)=(d1)+(e1)$, 其中

$$\begin{aligned}(d1)&=\Delta\sum_{g=1}^{m}\{(\hat{\boldsymbol{\beta}}_l^{(k),S}(d))^{\mathrm{T}}-(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\}\{\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)-\boldsymbol{JV}(t_{d,g})\}\\&\quad\times\{(\hat{\boldsymbol{\beta}}_l^{(k),S}(d))^{\mathrm{T}}+(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\}\{\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)-\boldsymbol{JV}(t_{d,g})\},\\(e1)&=\Delta\sum_{g=1}^{m}((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)-\boldsymbol{X}(t_{d,g})+\boldsymbol{X}(d-1)-\boldsymbol{JV}(t_{d,g})+\boldsymbol{JJ}(t_{d,g})\})\\&\quad\times((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)+\boldsymbol{X}(t_{d,g})-\boldsymbol{X}(d-1)-\boldsymbol{JV}(t_{d,g})-\boldsymbol{JJ}(t_{d,g})\}).\end{aligned}$$

我们有 $\|(\hat{\boldsymbol{\beta}}_l^{(k),S}(d))^{\mathrm{T}}\|_2^2=1$, $\|(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\|_2^2=1$, 所以 $\mathrm{E}(\|(\hat{\boldsymbol{\beta}}_l^{(k),S}(d))^{\mathrm{T}}+(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\|_2^a)\leqslant C$. 当 $k=0$ 时, 根据假设 3.1 以及文献 [38, 定理 3.1] 有 $\mathrm{E}(\|(\hat{\boldsymbol{\beta}}_l^{(k),S}(d))^{\mathrm{T}}-(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\|_2^a)\leqslant C\{m^{-a/4}+(s_p/p)^a\}$; 当 $k=1,\ldots,K$ 时, 由于 $p_k=O(p^{1/2-b})$, 根据文献 [36, 定理 4] 有

$$\begin{aligned}\mathrm{E}(\|(\hat{\boldsymbol{\beta}}_l^{(k),S}(d))^{\mathrm{T}}-(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\|_2^a)&\leqslant\frac{\|\hat{\epsilon}^S(d)-(\sum_{k=1}^{K}\boldsymbol{\tau}^{(k)}(d)+\boldsymbol{\Sigma}^S(d))\|_2^a+\|\boldsymbol{\Sigma}(d)-\boldsymbol{\Sigma}^S(d)\|_2^a}{(\lambda_j^{(k)}(d)-\lambda_{j+1}^{(k)}(d))^a}\\&\leqslant C\{m^{-a/4}(\log p)^{a/2}+p^{-ab}+(s_p)^ap^{a(b-1/2)}\},\end{aligned}$$

且 $\mathrm{E}(\|\boldsymbol{Y}^{(k)}(t_{d,g})-\boldsymbol{Y}^{(k)}(d-1)-\boldsymbol{JV}^{(k)}(t_{d,g})\|_2^{2a})=O_p(p_k^a)$, 其中上标 (k) 表示第 k 组. 由引理 A.1 知, 当 $k=0$ 时,

$$\begin{aligned}\mathrm{E}(|(d1)|)&\leqslant C\Delta\sum_{g=1}^{m}\mathrm{E}(\|\hat{\boldsymbol{\beta}}_l^{(k),S}(d)-\boldsymbol{\beta}_l^{(k)}(d)\|_2^4)^{1/4}\mathrm{E}(\|\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)-\boldsymbol{JV}(t_{d,g})\|_2^4)^{1/2}\\&\leqslant Cp\left\{m^{-1/4}+\left(\frac{s_p}{p}\right)\right\}\quad\text{a.s.};\end{aligned}$$

当 $k=1,\ldots,K$ 时,

$$\begin{aligned}\mathrm{E}(|(d1)|) &\leqslant C\Delta\sum_{g=1}^{m}\mathrm{E}(\|\hat{\boldsymbol{\beta}}_l^{(k),S}(d)-\boldsymbol{\beta}_l^{(k)}(d)\|_2^4)^{1/4}\mathrm{E}(\|\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)-\boldsymbol{JV}(t_{d,g})\|_2^4)^{1/2}\\ &\leqslant Cp_k\{m^{-1/4}(\log p)^{1/2}+p^{-b}+s_p p^{b-1/2}\}\quad \text{a.s.}\end{aligned}$$

容易推导出 $(e1)=(e11)-(e12)$, 其中,

$$\begin{aligned}(e11) &= \Delta\sum_{g=1}^{m}((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)-\boldsymbol{X}(t_{d,g})+\boldsymbol{X}(d-1)\})\\ &\quad\times((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)+\boldsymbol{X}(t_{d,g})-\boldsymbol{X}(d-1)-\boldsymbol{JV}(t_{d,g})-\boldsymbol{JJ}(t_{d,g})\}),\\ (e12) &= \Delta\sum_{g=1}^{m}((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{JV}(t_{d,g})-\boldsymbol{JJ}(t_{d,g})\})\\ &\quad\times((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)+\boldsymbol{X}(t_{d,g})-\boldsymbol{X}(d-1)-\boldsymbol{JV}(t_{d,g})-\boldsymbol{JJ}(t_{d,g})\}).\end{aligned}$$

对于 $(e11)$, 由假设 3.1(2) 及文献 [38, 引理 6.1] 中 $\boldsymbol{\epsilon}$ 的性质有

$$\mathrm{E}(|(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)-\boldsymbol{X}(t_{d,g})+\boldsymbol{X}(d-1)\}|^2\mid\boldsymbol{\beta}_l^{(k)}(d))\leqslant C,$$

则当 $k=0$ 时, 利用引理 A.1 得

$$\begin{aligned}\mathrm{E}(|(e11)|) &\leqslant \Delta\sum_{g=1}^{m}\mathrm{E}(\|(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\|_2^4)^{1/4}\mathrm{E}(\|\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)+\boldsymbol{X}(t_{d,g})\\ &\quad-\boldsymbol{X}(d-1)-\boldsymbol{JV}(t_{d,g})-\boldsymbol{JJ}(t_{d,g})\|_2^4)^{1/4}\\ &\quad\times\mathrm{E}(|(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)-\boldsymbol{X}(t_{d,g})+\boldsymbol{X}(d-1)\}|^2)^{1/2}\\ &\leqslant Cp^{1/2}\mathrm{E}(|(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{Y}(t_{d,g})-\boldsymbol{Y}(d-1)-\boldsymbol{X}(t_{d,g})+\boldsymbol{X}(d-1)\}|^2)^{1/2}\\ &\leqslant Cp^{1/2}\quad\text{a.s.}\end{aligned}$$

同理, 当 $k=1,\ldots,K$ 时, $\mathrm{E}(|(e11)|)\leqslant Cp_k^{1/2}$ a.s. 类似地, 对于 $(e12)$, 由假设 3.1(1) 知 $\mathrm{E}(\|\boldsymbol{JV}^{(k)}(t_{d,g})-\boldsymbol{JJ}^{(k)}(t_{d,g})\|_2^a)\leqslant Cp_k^{a/2}m^{-a/4}$, 所以利用引理 A.1 得

$$\mathrm{E}(|(e12)|)\leqslant Cp_k m^{-1/4}\quad\text{a.s.},$$

即 $\mathrm{E}(|(e1)|)\leqslant C(p_k m^{-1/4}+p_k^{1/2})$ a.s. 容易推导出 $(b1)=(f1)+(g1)$, 其中

$$\begin{aligned}(f1) &= 2\Delta\sum_{g=1}^{m}((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{\kappa}^{\mathrm{T}}(t_{d,g})\tilde{\boldsymbol{W}}(t_{d,g})-\boldsymbol{\kappa}^{\mathrm{T}}(d-1)\tilde{\boldsymbol{W}}(d-1)\})((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{X}(t_{d,g})\\ &\quad-\boldsymbol{X}(d-1)-\boldsymbol{JJ}(t_{d,g})\}),\\ (g1) &= \Delta\sum_{g=1}^{m}((\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{\kappa}^{\mathrm{T}}(t_{d,g})\tilde{\boldsymbol{W}}(t_{d,g})-\boldsymbol{\kappa}^{\mathrm{T}}(d-1)\tilde{\boldsymbol{W}}(d-1)\})^2.\end{aligned}$$

利用假设 3.2 和引理 A.1 得

$$\mathrm{E}(|(g1)|)\leqslant\Delta\sum_{g=1}^{m}\mathrm{E}(\|\boldsymbol{\beta}_l^{(k)}(d)\|_2^4)^{1/2}\mathrm{E}(\|\boldsymbol{\kappa}^{\mathrm{T}}(t_{d,g})\tilde{\boldsymbol{W}}(t_{d,g})-\boldsymbol{\kappa}^{\mathrm{T}}(d-1)\tilde{\boldsymbol{W}}(d-1)\|_2^4)^{1/2}\leqslant Cs_p.$$

从而有

$$\begin{aligned}\mathrm{E}(|(f1)|) &\leqslant C\Delta\sum_{g=1}^{m}\mathrm{E}(|(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{\kappa}^{\mathrm{T}}(t_{d,g})\tilde{\boldsymbol{W}}(t_{d,g})-\boldsymbol{\kappa}^{\mathrm{T}}(d-1)\tilde{\boldsymbol{W}}(d-1)\}|^2)^{1/2}\\ &\quad\times\mathrm{E}(\|\boldsymbol{\beta}_l^{(k)}(d)\|_2^4)^{1/4}\mathrm{E}(\|\boldsymbol{X}(t_{d,g})-\boldsymbol{X}(d-1)-\boldsymbol{J}\boldsymbol{J}(t_{d,g})\|_2^4)^{1/4}\\ &\leqslant Cs_p^{1/2}p_k^{1/2}\quad\text{a.s.}\end{aligned}$$

对于 $(c1)$,

$$\begin{aligned}\mathrm{E}(|(c1)|)=\sum_{g=1}^{m}\int_{t_{d,g-1}}^{t_{d,g}}&\mathrm{E}(|(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{X}(t_{d,g})-\boldsymbol{X}(t)-\boldsymbol{J}(t_{d,g})+\boldsymbol{J}(t)-\boldsymbol{\kappa}^{\mathrm{T}}(t_{d,g})\tilde{\boldsymbol{W}}(t_{d,g})+\boldsymbol{\kappa}^{\mathrm{T}}(t)\tilde{\boldsymbol{W}}(t)\}\\ &\times(\boldsymbol{\beta}_l^{(k)}(d))^{\mathrm{T}}\{\boldsymbol{X}(t_{d,g})+\boldsymbol{X}(t)-\boldsymbol{J}(t_{d,g})-\boldsymbol{J}(t)-\boldsymbol{\kappa}^{\mathrm{T}}(t_{d,g})\tilde{\boldsymbol{W}}(t_{d,g})-\boldsymbol{\kappa}^{\mathrm{T}}(t)\tilde{\boldsymbol{W}}(t)\\ &-2\boldsymbol{X}(d-1)+2\boldsymbol{J}(d-1)+2\boldsymbol{\kappa}^{\mathrm{T}}(d-1)\tilde{\boldsymbol{W}}(d-1)\}|)dt.\end{aligned}$$

当 $t_{d,g-1}<t<t_{d,g}$ 时, 由假设 3.1(2) 知

$$\begin{aligned}&\mathrm{E}(\|\boldsymbol{X}^{(k)}(t_{d,g})-\boldsymbol{X}^{(k)}(t)-\boldsymbol{J}^{(k)}(t_{d,g})+\boldsymbol{J}^{(k)}(t)-(\boldsymbol{\kappa}^{\mathrm{T}}(t_{d,g})\tilde{\boldsymbol{W}}(t_{d,g}))^{(k)}+(\boldsymbol{\kappa}^{\mathrm{T}}(t)\tilde{\boldsymbol{W}}(t))^{(k)}\|_2^4)\\ &\quad\leqslant C\mathrm{E}\bigg(\bigg\{\int_t^{t_{d,g}}\|\boldsymbol{\Phi}^{(k)}(u)\|_2du\bigg\}^2\bigg)+\mathrm{E}\bigg(\bigg\{\bigg\|\int_t^{t_{d,g}}d\boldsymbol{J}^{(k)}(u)\bigg\|_2^2\bigg\}^2\bigg)+\mathrm{E}\bigg(\bigg\{\int_t^{t_{d,g}}\|(\boldsymbol{\kappa}^{\mathrm{T}}(u)\boldsymbol{\kappa}(u))^{(k)}\|_2du\bigg\}^2\bigg)\\ &\quad\leqslant Cm^{-2}(p_k^2+s_p^2).\end{aligned}$$

由引理 A.1 知

$$\mathrm{E}(|(c1)|)\leqslant Cp_k^{1/2}m^{-1/2}(p_k^2+s_p^2)^{1/4}\leqslant Cp_k^{1/2}m^{-1/2}(p_k^{1/2}+s_p^{1/2}).$$

所以

$$\begin{aligned}\mathrm{E}(|\hat{\rho}_l^{(0)}(d)-\rho_l^{(0)}(d)|)&\leqslant Cp\bigg\{m^{-1/4}+\bigg(\frac{s_p}{p}\bigg)^{1/2}\bigg\}\quad\text{a.s.},\quad l=1,\ldots,r_0,\\ \mathrm{E}(|\hat{\rho}_l^{(k)}(d)-\rho_l^{(k)}(d)|)&\leqslant Cp_k\bigg\{m^{-1/4}(\log p)^{1/2}+p^{-b}+\bigg(\frac{s_p}{p_k}\bigg)^{1/2}\bigg\}\\ &\leqslant C\{m^{-1/4}(\log p)^{1/2}p^{1/2-b}+p^{1/2-2b}\\ &\quad+(s_p)^{1/2}p^{1/4-b/2}\}\quad\text{a.s.},\quad k=1,\ldots,K,\quad l=1,\ldots,r_k.\end{aligned}$$

综上, 引理 A.2 得证. □

进一步地, 在定义 A.1 下有引理 A.3.

定义 A.1 记

$$L_{n,j}^{(k)}(\boldsymbol{\theta}):=-\frac{1}{2n}\sum_{d=1}^{n}\bigg(\log\Lambda_j^{(k)}(d,\boldsymbol{\theta})+\frac{\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)}{\Lambda_j^{(k)}(d,\boldsymbol{\theta})}\bigg).$$

记 $\boldsymbol{L}_n(\boldsymbol{\theta})=\sum_{k=0}^{K}\sum_{j=1}^{r_k}L_{n,j}^{(k)}(\boldsymbol{\theta})$.

引理 A.3 在假设 2.1、2.2 和 3.1 的条件下,

$$\begin{aligned}&\mathrm{E}(|\hat{L}_{n,j}^{(0)}(\boldsymbol{\theta})-L_{n,j}^{(0)}(\boldsymbol{\theta})|)=O\bigg(m^{-1/4}+\bigg(\frac{s_p}{p}\bigg)^{1/2}\bigg),\quad j=1,\ldots,r_0,\\ &\mathrm{E}(|\hat{L}_{n,j}^{(k)}(\boldsymbol{\theta})-L_{n,j}^{(k)}(\boldsymbol{\theta})|)=O(m^{-1/4}(\log p)^{1/2}+p^{-b}+(s_p)^{1/2}p^{b/2-1/4}),\quad j=1,\ldots,r_k,\quad k=1,\ldots,K.\end{aligned}$$

证明 首先, 由于

$$\begin{aligned}\hat{L}_{n,j}^{(k)}(\boldsymbol{\theta})-L_{n,j}^{(k)}(\boldsymbol{\theta})&=-\frac{1}{2n}\sum_{d=1}^{n}\left\{\log\left(\frac{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}{\Lambda_j^{(k)}(d,\boldsymbol{\theta})}\right)+\frac{\hat{\Lambda}_j^{(k)}(d)}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}-\frac{\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)}{\Lambda_j^{(k)}(d,\boldsymbol{\theta})}\right\}\\&=-\frac{1}{2n}\sum_{d=1}^{n}\log\frac{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}{\Lambda_j^{(k)}(d,\boldsymbol{\theta})}-\frac{1}{2n}\sum_{d=1}^{n}\frac{\hat{\Lambda}_j^{(k)}(d)-\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}\\&\quad+\frac{1}{2n}\sum_{d=1}^{n}\frac{\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)\{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})-\Lambda_j^{(k)}(d,\boldsymbol{\theta})\}}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})\Lambda_j^{(k)}(d,\boldsymbol{\theta})},\end{aligned}$$

所以

$$\begin{aligned}&|\hat{L}_{n,j}^{(k)}(\boldsymbol{\theta})-L_{n,j}^{(k)}(\boldsymbol{\theta})|\\&\quad\leqslant\frac{1}{2n}\sum_{d=1}^{n}\left|\log\frac{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}{\Lambda_j^{(k)}(d,\boldsymbol{\theta})}\right|+\frac{1}{2n}\sum_{d=1}^{n}\left|\frac{\hat{\Lambda}_j^{(k)}(d)-\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}\right|+\frac{1}{2n}\sum_{d=1}^{n}\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)\left|\frac{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})-\Lambda_j^{(k)}(d,\boldsymbol{\theta})}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})\Lambda_j^{(k)}(d,\boldsymbol{\theta})}\right|.\end{aligned}$$

记

$$\begin{aligned}(a2)&=\frac{1}{2n}\sum_{d=1}^{n}\left|\log\frac{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}{\Lambda_j^{(k)}(d,\boldsymbol{\theta})}\right|,\\(b2)&=\frac{1}{2n}\sum_{d=1}^{n}\left|\frac{\hat{\Lambda}_j^{(k)}(d)-\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}\right|,\\(c2)&=\frac{1}{2n}\sum_{d=1}^{n}\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)\left|\frac{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})-\Lambda_j^{(k)}(d,\boldsymbol{\theta})}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})\Lambda_j^{(k)}(d,\boldsymbol{\theta})}\right|.\end{aligned}$$

对于 $(a2)$, 由于 $\log(1+x)\leqslant x$, $\forall\, x>-1$, 所以 $\log\{\frac{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}{\Lambda_j^{(k)}(d,\boldsymbol{\theta})}\}\leqslant\frac{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})-\Lambda_j^{(k)}(d,\boldsymbol{\theta})}{\Lambda_j^{(k)}(d,\boldsymbol{\theta})}$.

不妨先考虑共同因子部分 $\Lambda_j^{(0)}(d,\boldsymbol{\theta})$, $j=1,\ldots,r_0$. 容易推导出 $\tilde{\Lambda}_j^{(0)}(d,\boldsymbol{\theta})-\Lambda_j^{(0)}(d,\boldsymbol{\theta})=(a21)+(a22)$, 其中

$$\begin{aligned}(a21)&=\sum_{l=1}^{r_0}\eta_{j,l}^{(0)}\sum_{t=1}^{d-1}(\gamma_j^{(0)})^{d-t-1}\bigg\{((\hat{\boldsymbol{\beta}}_l^{(0),S}(d))^{\mathrm{T}}\{\boldsymbol{Y}(t)-\boldsymbol{Y}(t-1)-\boldsymbol{J}\boldsymbol{V}(t)\})^2\\&\quad-\left\{\int_{t-1}^{t}\sqrt{\lambda_l^{(0)}(s,\boldsymbol{\theta})}dW_l^{(0)}(s)\right\}^2\bigg\},\\(a22)&=\sum_{l=1}^{r_0}\eta_{j,l}^{(0)}\{\hat{\rho}_l^{(0)}(d)-\rho_l^{(0)}(d)\}.\end{aligned}$$

由引理 A.2 及假设 2.2 可知

$$\mathrm{E}(|(a22)|)\leqslant Cp\left\{m^{-1/4}+\left(\frac{s_p}{p}\right)^{1/2}\right\}\quad\text{a.s.}$$

由假设 2.2, 类似引理 A.2 的证明有

$$\mathrm{E}(|(a21)|)\leqslant Cp\left\{m^{-1/4}+\left(\frac{s_p}{p}\right)^{1/2}\right\}\quad\text{a.s.}$$

所以

$$\mathrm{E}(|\tilde{\Lambda}_j^{(0)}(d,\boldsymbol{\theta})-\Lambda_j^{(0)}(d,\boldsymbol{\theta})|)\leqslant Cp\left(m^{-1/4}+\left(\frac{s_p}{p}\right)^{1/2}\right)\quad \text{a.s.}$$

对于分组因子部分 $\Lambda_j^{(k)}(d,\boldsymbol{\theta})$, $j=1,\ldots,r_k$, $k=1,\ldots,K$,

$$\mathrm{E}(|\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})-\Lambda_j^{(k)}(d,\boldsymbol{\theta})|)\leqslant C\{m^{-1/4}(\log p)^{1/2}p^{1/2-b}+p^{1/2-2b}+(s_p)^{1/2}p^{1/4-b/2}\}\quad \text{a.s.},\quad k=1,\ldots,K.$$

又由于平稳性假设 2.2 及 $\Lambda_j^{(k)}(d,\boldsymbol{\theta})$ 和 $\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})$ 的迭代公式, 可知

$$\frac{\Lambda_j^{(k)}(d,\boldsymbol{\theta})}{p_k}=O_p(1),\quad \frac{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}{p_k}=O_p(1),\quad j=1,\ldots,r_k,\quad k=0,\ldots,K,$$

所以当 $k=0$ 时,

$$\mathrm{E}(|(a2)|)\leqslant\frac{1}{2n}\sum_{d=1}^{n}\mathrm{E}\left(\frac{|\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})-\Lambda_j^{(k)}(d,\boldsymbol{\theta})|}{\Lambda_j^{(k)}(d,\boldsymbol{\theta})}\right)\leqslant C\left\{m^{-1/4}+\left(\frac{s_p}{p}\right)^{1/2}\right\}\quad \text{a.s.};$$

当 $k=1,\ldots,K$ 时,

$$\mathrm{E}(|(a2)|)\leqslant C\{m^{-1/4}(\log p)^{1/2}+p^{-b}+(s_p)^{1/2}p^{b/2-1/4}\}\quad \text{a.s.}$$

同理, 对于 $(c2)$, 当 $k=0$ 时, $\mathrm{E}(|(c2)|)\leqslant C\{m^{-1/4}+(s_p/p)^{1/2}\}$ a.s. 当 $k=1,\ldots,K$ 时, $\mathrm{E}(|(c2)|)\leqslant C\{m^{-1/4}(\log p)^{1/2}+p^{-b}+(s_p)^{1/2}p^{b/2-1/4}\}$ a.s. 对于 $(b2)$, 根据假设 3.1 及文献 [38, 定理 3.1] 知, 当 $k=0$ 时有

$$\mathrm{E}(|(b2)|)\leqslant\frac{1}{2n}\sum_{d=1}^{n}\mathrm{E}\left(\frac{|\hat{\Lambda}_j^{(k)}(d)-\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)|}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}\right)\leqslant C\left(m^{-1/4}+\frac{s_p}{p}\right)\quad \text{a.s.}$$

当 $k=1,\ldots,K$ 时, 由于 $m^{-1/2}\log p\to 0, s_p=o(p_k)$, 所以有

$$\begin{aligned}
&|\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)-\hat{\Lambda}_j^{(k)}(d)|\\
&\quad=\left|(\boldsymbol{\beta}_j^{(k)}(d))^{\mathrm{T}}\left(\left(\sum_{k=1}^{K}\boldsymbol{\tau}^{(k)}(d)+\boldsymbol{\Sigma}(d)\right)-\hat{\boldsymbol{\epsilon}}^S(d)\right)\boldsymbol{\beta}_j^{(k)}(d)\right|\\
&\qquad+|(\boldsymbol{\beta}_j^{(k)}(d))^{\mathrm{T}}\hat{\boldsymbol{\epsilon}}^S(d)\boldsymbol{\beta}_j^{(k)}(d)-(\hat{\boldsymbol{\beta}}_j^{(k),S}(d))^{\mathrm{T}}\hat{\boldsymbol{\epsilon}}^S(d)\hat{\boldsymbol{\beta}}_j^{(k),S}(d)|\\
&\quad\leqslant C\left\|\left(\sum_{k=1}^{K}\boldsymbol{\tau}^{(k)}(d)+\boldsymbol{\Sigma}(d)\right)-\hat{\boldsymbol{\epsilon}}^S(d)\right\|_2+C\|\hat{\boldsymbol{\epsilon}}^S(d)\|_2\frac{\|(\sum_{k=1}^{K}\boldsymbol{\tau}^{(k)}(d)+\boldsymbol{\Sigma}(d))-\hat{\boldsymbol{\epsilon}}^S(d)\|_2^2}{p_k^2}\\
&\quad\leqslant C(m^{-1/4}p^{1/2-b}(\log p)^{1/2}+p^{1/2-2b}+s_p)+C\left(m^{-1/2}p^{1/2-b}(\log p)+p^{1/2-3b}+\frac{s_p^2}{p_k}\right)\\
&\quad\leqslant C(m^{-1/4}p^{1/2-b}(\log p)^{1/2}+p^{1/2-2b}+s_p).
\end{aligned}$$

故 $\mathrm{E}(|(b2)|)\leqslant C(m^{-1/4}(\log p)^{1/2}+p^{-b}+s_pp^{b-1/2})$ a.s.

综上, 引理 A.3 得证. □

定义 A.2 记 $\hat{\Phi}_{n,j}^{(k)}(\boldsymbol{\theta})$ 和 $\Phi_{n,j}^{(k)}(\boldsymbol{\theta})$ 分别为 $\hat{L}_{n,j}^{(k)}(\boldsymbol{\theta})$ 和 $L_{n,j}^{(k)}(\boldsymbol{\theta})$ 关于 $\boldsymbol{\theta}$ 的导数, 即

$$\hat{\Phi}_{n,j}^{(k)}(\boldsymbol{\theta}) := \frac{\partial \hat{L}_{n,j}^{(k)}(\boldsymbol{\theta})}{\partial \mathrm{Vec}(\boldsymbol{\theta})}, \quad \Phi_{n,j}^{(k)}(\boldsymbol{\theta}) := \frac{\partial L_{n,j}^{(k)}(\boldsymbol{\theta})}{\partial \mathrm{Vec}(\boldsymbol{\theta})}.$$

记 $\hat{\boldsymbol{\Phi}}_n(\boldsymbol{\theta}) = \sum_{k=0}^{K}\sum_{j=1}^{r_k}\hat{\Phi}_{n,j}^{(k)}(\boldsymbol{\theta})$, $\boldsymbol{\Phi}_n(\boldsymbol{\theta}) = \sum_{k=0}^{K}\sum_{j=1}^{r_k}\Phi_{n,j}^{(k)}(\boldsymbol{\theta})$. 记

$$\nabla\hat{\boldsymbol{\Phi}}_n(\boldsymbol{\theta}) = \frac{\partial \hat{\boldsymbol{\Phi}}_n(\boldsymbol{\theta})}{\partial \mathrm{Vec}(\boldsymbol{\theta})}, \quad \nabla\boldsymbol{\Phi}_n(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{\Phi}_n(\boldsymbol{\theta})}{\partial \mathrm{Vec}(\boldsymbol{\theta})}.$$

引理 A.4 (1) 当 $n \geqslant 2+4\tilde{r}$ 时, $\nabla\hat{\boldsymbol{\Phi}}_n(\tilde{\boldsymbol{\theta}}) \to^p \nabla\boldsymbol{\Phi}_n(\boldsymbol{\theta}^*)$.

(2) 对于任意 $q \geqslant 1, d = 1,\ldots,n$, $j = 1,\ldots,r_k$, $k = 1,\ldots,K$, $i,l,h = 1,\ldots,\tilde{r}$,

$$\left\| \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \frac{1}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})} \frac{\partial \tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}{\partial \mathrm{Vec}(\boldsymbol{\theta}_i)} \right\|_q \leqslant C,$$

$$\left\| \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \frac{1}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})} \frac{\partial (\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}))^2}{\partial \mathrm{Vec}(\boldsymbol{\theta}_i)\partial \mathrm{Vec}(\boldsymbol{\theta}_l)} \right\|_q \leqslant C,$$

$$\left\| \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \frac{1}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})} \frac{\partial (\tilde{\Lambda}_{k,j}^{(k)}(d,\boldsymbol{\theta}))^3}{\partial \mathrm{Vec}(\boldsymbol{\theta}_i)\partial \mathrm{Vec}(\boldsymbol{\theta}_l)\partial \mathrm{Vec}(\boldsymbol{\theta}_h)} \right\|_q \leqslant C.$$

(3) 当 $n \geqslant 2+2\tilde{r}$ 时, $-\nabla\boldsymbol{\Phi}_n(\boldsymbol{\theta}^*)$ 几乎处处是一个半正定矩阵.

证明 类似于文献 [38, 命题 6.1 和 定理 3.2], 引理 A.4(1) 由引理 A.3 可证; 引理 A.4(2) 的证明类似于文献 [39, 引理 2]; 引理 A.4(3) 的证明类似于文献 [38, 引理 6.5(1)]. □

定理 3.1 的证明 由于

$$\begin{aligned}
\|\hat{\boldsymbol{\Phi}}_n(\boldsymbol{\theta}^*)\|_\infty &\leqslant \frac{1}{2n}\sum_{d=1}^{n}\sum_{k=0}^{K}\sum_{j=1}^{r_k}\left\| \frac{1}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)}\frac{\partial \tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)}{\partial \mathrm{Vec}(\boldsymbol{\theta})} - \frac{\hat{\Lambda}_j^{(k)}(d)}{(\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*))^2}\frac{\partial \tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)}{\partial \mathrm{Vec}(\boldsymbol{\theta})}\right\|_\infty \\
&\leqslant \frac{1}{2n}\sum_{d=1}^{n}\sum_{k=0}^{K}\sum_{j=1}^{r_k}\left\| \frac{1}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)}\frac{\partial \tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)}{\partial \mathrm{Vec}(\boldsymbol{\theta})}\right\|_\infty \left| \frac{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*) - \hat{\Lambda}_j^{(k)}(d)}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)}\right|,
\end{aligned}$$

又因为

$$|\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*) - \hat{\Lambda}_j^{(k)}(d)| \leqslant |\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*) - \Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)| + |\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*) - \hat{\Lambda}_j^{(k)}(d)|,$$

由引理 A.3 的证明可知, 当 $k=0$ 时,

$$\mathrm{E}(|\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*) - \Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)|) \leqslant Cp\left\{m^{-1/4} + \left(\frac{s_p}{p}\right)^{1/2}\right\} \quad \text{a.s.};$$

当 $k = 1,\ldots,K$ 时,

$$\mathrm{E}(|\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*) - \Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)|) \leqslant C\{m^{-1/4}p^{1/2-b}(\log p)^{1/2} + p^{1/2-2b} + (s_p)^{1/2}p^{1/4-b/2}\} \quad \text{a.s.}$$

类似于引理 A.3 的证明, 当 $k=0$ 时,

$$\mathrm{E}(|\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*) - \hat{\Lambda}_j^{(k)}(d)|) \leqslant Cp\left(m^{-1/4} + \frac{s_p}{p}\right);$$

当 $k=1,\ldots,K$ 时,

$$\mathrm{E}(|\Lambda_j^{(k)}(d,\boldsymbol{\theta}^*)-\hat{\Lambda}_j^{(k)}(d)|)\leqslant C(m^{-1/4}p^{1/2-b}(\log p)^{1/2}+p^{1/2-2b}+s_p).$$

又由于 $(s_p)p^{b-1/2}=o_p(1)$, 所以, 当 $k=0$ 时,

$$|\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)-\hat{\Lambda}_j^{(k)}(d)|=O_p\bigg(p\bigg\{m^{-1/4}+\bigg(\frac{s_p}{p}\bigg)^{1/2}\bigg\}\bigg);$$

当 $k=1,\ldots,K$ 时,

$$|\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)-\hat{\Lambda}_j^{(k)}(d)|=O_p(m^{-1/4}p^{1/2-b}(\log p)^{1/2}+p^{1/2-2b}+(s_p)^{1/2}p^{1/4-b/2}).$$

类似引理 A.3 的证明, 由于 $\frac{\Lambda_j^{(k)}(d,\boldsymbol{\theta})}{p_k}=O_p(1)$, $\frac{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta})}{p_k}=O_p(1)$, $j=1,\ldots,r_k$, $k=0,\ldots,K$, 所以当 $k=0$ 时,

$$\bigg|\frac{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)-\hat{\Lambda}_j^{(k)}(d)}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)}\bigg|=\frac{|\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)-\hat{\Lambda}_j^{(k)}(d)|}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)}=O_p\bigg(m^{-1/4}+\bigg(\frac{s_p}{p}\bigg)^{1/2}\bigg);$$

当 $k=1,\ldots,K$ 时,

$$\bigg|\frac{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)-\hat{\Lambda}_j^{(k)}(d)}{\tilde{\Lambda}_j^{(k)}(d,\boldsymbol{\theta}^*)}\bigg|=O_p(m^{-1/4}(\log p)^{1/2}+p^{-b}+(s_p)^{1/2}p^{b/2-1/4}).$$

由引理 A.4(2) 得

$$\|\hat{\boldsymbol{\Phi}}_n(\boldsymbol{\theta}^*)\|_\infty=O_p(m^{-1/4}(\log p)^{1/2}+p^{-b}+(s_p)^{1/2}p^{b/2-1/4}).$$

又因为

$$\hat{\boldsymbol{\Phi}}_n(\hat{\boldsymbol{\theta}})-\hat{\boldsymbol{\Phi}}_n(\boldsymbol{\theta}^*)=-\hat{\boldsymbol{\Phi}}_n(\boldsymbol{\theta}^*)=\nabla\hat{\boldsymbol{\Phi}}_n(\tilde{\boldsymbol{\theta}})\{\mathrm{Vec}(\hat{\boldsymbol{\theta}})-\mathrm{Vec}(\boldsymbol{\theta}^*)\},$$

由引理 A.4(1)(3) 知, $\|\mathrm{Vec}(\hat{\boldsymbol{\theta}})-\mathrm{Vec}(\boldsymbol{\theta}^*)\|_\infty$ 与 $\|\hat{\boldsymbol{\Phi}}_n(\boldsymbol{\theta}^*)\|_\infty$ 同阶, 故

$$\|\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^*\|_{\max}=O_p(m^{-1/4}(\log p)^{1/2}+p^{-b}+(s_p)^{1/2}p^{b/2-1/4}).$$

定理 3.1 得证. □

附录 A.3　定理 3.2 的证明

不妨记 $f_d(j)=p^{-1}\hat{\Lambda}_j^{(0)}(d)+jg_0(m,p)$, 则由文献 [36, 定理 2] 知, 当 $1\leqslant j\leqslant r_0$ 时, 存在 $K>0$, 使得 $f_d(j)-f_d(r_0+1)>K$; 当 $r_0+1<j\leqslant R_0$ 时, $\mathrm{P}(f_d(j)<f_d(r_0+1))\to 0$.

记 $f(j)=\sum_{d=1}^n f_d(j)$, 则当 $1\leqslant j\leqslant r_0$ 时, 存在 $K>0$, 使得

$$f(j)-f(r_0+1)=\sum_{d=1}^n(f_d(j)-f_d(r_0+1))>K;$$

当 $r_0+1<j\leqslant R_0$ 时,

$$\mathrm{P}(f(j)-f(r_0+1)<0)=\mathrm{P}\bigg(\sum_{d=1}^n(f_d(j)-f_d(r_0+1))<0\bigg).$$

又由于

$$\bigcup_{d=1}^{n}\{f_d(j): f_d(j)-f_d(r_0+1)<0\}\supseteq\{f(j): f(j)-f(r_0+1)<0\}\supseteq\bigcap_{d=1}^{n}\{f_d(j): f_d(j)-f_d(r_0+1)<0\},$$

所以 $\mathrm{P}(f(j)-f(r_0+1)<0)\leqslant \mathrm{P}(f_d(j)-f_d(r_0+1)<0)\to 0$. 综上, $\mathrm{P}(\hat{r}_0=r_0)=\mathrm{P}(f(j)\geqslant f(r_0+1))\to 1$, $1\leqslant j\leqslant R_0$.

附录 A.4 定理 3.3 的证明

由文献 [36, 定理 4] 知, 当 $\frac{1}{m}\log p\to 0$ 时,

$$\left\|\hat{\boldsymbol{\varepsilon}}^S(d)-\left(\sum_{k=1}^{K}\boldsymbol{\tau}^{(k)}(d)+\boldsymbol{\Sigma}^S(d)\right)\right\|_2=O_p(m^{-1/4}p^{1/2-b}(\log p)^{1/2}+p^{1/2-2b}).$$

又由假设 3.1(1) 和 3.2 知

$$\begin{aligned}\left\|\hat{\boldsymbol{\varepsilon}}(d)-\left(\sum_{k=1}^{K}\boldsymbol{\tau}^{(k)}(d)+\boldsymbol{\Sigma}^S(d)\right)\right\|_2&\leqslant\left\|\hat{\boldsymbol{\varepsilon}}(d)-\left(\sum_{k=1}^{K}\boldsymbol{\tau}^{(k)}(d)+\boldsymbol{\Sigma}(d)\right)\right\|_2+\|\boldsymbol{\Sigma}(d)-\boldsymbol{\Sigma}^S(d)\|_2\\&=O_p(p^{1/2}m^{-1/4}+s_p),\end{aligned}$$

故

$$\begin{aligned}\|\hat{\boldsymbol{\varepsilon}}^S(d)-\hat{\boldsymbol{\varepsilon}}(d)\|_2&\leqslant\left\|\hat{\boldsymbol{\varepsilon}}(d)-\left(\sum_{k=1}^{K}\boldsymbol{\tau}^{(k)}(d)+\boldsymbol{\Sigma}^S(d)\right)\right\|_2+\left\|\hat{\boldsymbol{\varepsilon}}^S(d)-\left(\sum_{k=1}^{K}\boldsymbol{\tau}^{(k)}(d)+\boldsymbol{\Sigma}^S(d)\right)\right\|_2\\&=O_p(p^{1/2}m^{-1/4}+m^{-1/4}p^{1/2-b}(\log p)^{1/2}+p^{1/2-2b}+s_p).\end{aligned}$$

不妨记 $\hat{\boldsymbol{\beta}}_j^{(k)}(d)$ $(j=1,\ldots,r_k)$ 是 $\hat{\boldsymbol{\Phi}}_{\mathrm{MSRV}}(d)$ 上对应位置的特征向量, 根据 Davis-Kahan sin θ 理论, 有

$$\begin{aligned}\|\hat{\boldsymbol{\beta}}_j^{(k),S}(d)-\hat{\boldsymbol{\beta}}_j^{(k)}(d)\|_2&\leqslant\frac{\|\hat{\boldsymbol{\varepsilon}}^S(d)-\hat{\boldsymbol{\varepsilon}}(d)\|_2}{\lambda_j^{(k)}(d)-\lambda_{j+1}^{(k)}(d)}\\&=O_p(p^bm^{-1/4}+(\log p)^{1/2}m^{-1/4}+p^{-b}+s_pp^{b-1/2}),\quad k=1,\ldots,K.\end{aligned}$$

又由特征分解性质, 故 $(\hat{\boldsymbol{\beta}}_j^{(k)}(d))^{\mathrm{T}}\hat{\boldsymbol{\beta}}_{j'}^{(0)}(d)=0$, $k\neq 0$, $j=1,\ldots,r_k$, $j'=1,\ldots,r_0$,

$$\begin{aligned}(\hat{\boldsymbol{\beta}}_j^{(k),S}(d))^{\mathrm{T}}\hat{\boldsymbol{\beta}}_{j'}^{(0)}(d)&=(\hat{\boldsymbol{\beta}}_j^{(k),S}(d))^{\mathrm{T}}\hat{\boldsymbol{\beta}}_{j'}^{(0)}(d)-(\hat{\boldsymbol{\beta}}_j^{(k)}(d))^{\mathrm{T}}\hat{\boldsymbol{\beta}}_{j'}^{(0)}(d)\\&=(\hat{\boldsymbol{\beta}}_j^{(k),S}(d)-\hat{\boldsymbol{\beta}}_j^{(k)}(d))^{\mathrm{T}}\hat{\boldsymbol{\beta}}_{j'}^{(0)}(d)\\&\leqslant\|\hat{\boldsymbol{\beta}}_j^{(k),S}(d)-\hat{\boldsymbol{\beta}}_j^{(k)}(d)\|_2\|\hat{\boldsymbol{\beta}}_{j'}^{(0)}(d)\|_2\\&=O_p(p^bm^{-1/4}+(\log p)^{1/2}m^{-1/4}+p^{-b}+s_pp^{b-1/2}),\quad k=1,\ldots,K.\end{aligned}$$

故当 $m\to\infty$, $p\to\infty$, $p^bm^{-1/4}\to 0$ 时, $(\hat{\boldsymbol{\beta}}_j^{(k),S}(d))$ 与 $\hat{\boldsymbol{\beta}}_{j'}^{(0)}(d)$ 渐近正交.

附录 B 模拟补充

以模拟中生成第一组资产价格的参数值 $(\boldsymbol{\lambda}^{(1)}(0,\boldsymbol{\theta}))^*$、$(\boldsymbol{w}^{(1)})^*$、$(\boldsymbol{\gamma}^{(1)})^*$、$(\boldsymbol{\eta}_A^{(1)})^*$ 和 $(\boldsymbol{\eta}_B^{(1)})^*$ 为基准参考值, 第二组在保证剩余参数值与第一组相同的条件下, 分别考虑 $\boldsymbol{\lambda}^{(2)}(0,\boldsymbol{\theta})$、$\boldsymbol{w}^{(2)}$、$\boldsymbol{\gamma}^{(2)}$、$\boldsymbol{\eta}_A^{(2)}$ 和 $\boldsymbol{\eta}_B^{(2)}$ 的取值为第一组二分之一的情形. 表 B1 通过两组的平均资产日积分波动率比值说明各参数对波动率的影响.

表 B1　各参数选择下对应的两组平均资产日积分波动率比值

参数	比值
$\boldsymbol{\lambda}^{(2)}(0,\boldsymbol{\theta}) = 1/2(\boldsymbol{\lambda}^{(1)}(0,\boldsymbol{\theta}))^*, \boldsymbol{w}^{(2)} = (\boldsymbol{w}^{(1)})^*, \boldsymbol{\gamma}^{(2)} = (\boldsymbol{\gamma}^{(1)})^*, \boldsymbol{\eta}_A^{(2)} = (\boldsymbol{\eta}_A^{(1)})^*, \boldsymbol{\eta}_B^{(2)} = (\boldsymbol{\eta}_B^{(1)})^*$	1.013
$\boldsymbol{\lambda}^{(2)}(0,\boldsymbol{\theta}) = (\boldsymbol{\lambda}^{(1)}(0,\boldsymbol{\theta}))^*, \boldsymbol{w}^{(2)} = 1/2(\boldsymbol{w}^{(1)})^*, \boldsymbol{\gamma}^{(2)} = (\boldsymbol{\gamma}^{(1)})^*, \boldsymbol{\eta}_A^{(2)} = (\boldsymbol{\eta}_A^{(1)})^*, \boldsymbol{\eta}_B^{(2)} = (\boldsymbol{\eta}_B^{(1)})^*$	1.225
$\boldsymbol{\lambda}^{(2)}(0,\boldsymbol{\theta}) = (\boldsymbol{\lambda}^{(1)}(0,\boldsymbol{\theta}))^*, \boldsymbol{w}^{(2)} = (\boldsymbol{w}^{(1)})^*, \boldsymbol{\gamma}^{(2)} = 1/2(\boldsymbol{\gamma}^{(1)})^*, \boldsymbol{\eta}_A^{(2)} = (\boldsymbol{\eta}_A^{(1)})^*, \boldsymbol{\eta}_B^{(2)} = (\boldsymbol{\eta}_B^{(1)})^*$	1.102
$\boldsymbol{\lambda}^{(2)}(0,\boldsymbol{\theta}) = (\boldsymbol{\lambda}^{(1)}(0,\boldsymbol{\theta}))^*, \boldsymbol{w}^{(2)} = (\boldsymbol{w}^{(1)})^*, \boldsymbol{\gamma}^{(2)} = (\boldsymbol{\gamma}^{(1)})^*, \boldsymbol{\eta}_A^{(2)} = 1/2(\boldsymbol{\eta}_A^{(1)})^*, \boldsymbol{\eta}_B^{(2)} = (\boldsymbol{\eta}_B^{(1)})^*$	1.456
$\boldsymbol{\lambda}^{(2)}(0,\boldsymbol{\theta}) = (\boldsymbol{\lambda}^{(1)}(0,\boldsymbol{\theta}))^*, \boldsymbol{w}^{(2)} = (\boldsymbol{w}^{(1)})^*, \boldsymbol{\gamma}^{(2)} = (\boldsymbol{\gamma}^{(1)})^*, \boldsymbol{\eta}_A^{(2)} = (\boldsymbol{\eta}_A^{(1)})^*, \boldsymbol{\eta}_B^{(2)} = 1/2(\boldsymbol{\eta}_B^{(1)})^*$	1.236

High-dimensional volatility matrix estimation with high-frequency financial data: The GARCH-Itô grouped factor model

Weiqing Gao, Ben Wu & Bo Zhang

Abstract　In the research area of the high-frequency financial data analysis, estimation and prediction of the high-dimensional volatility matrix are challenging problems, especially when the assets of interest have a naturally given group structure. In order to address this issue, we propose a novel GARCH-Itô grouped factor model, in which we present the log price series of grouped assets with common factors, group-specific factors, and an asset-specific error term. We then embed the discrete GARCH structure into the volatility of the eigenvalue processes to capture the volatility dynamics of the observed price data. We propose a quasi maximum likelihood method for parameter estimation, establish its asymptotic properties and illustrate its good finite-sample performance with simulation. In real data analysis, we compare our method with the ungrouped factor model and the nonparametric high-frequency volatility estimator MSRV using the data from the SSE Main Board and the SZSE ChiNext market, where the proposed method outperforms its competitors in terms of prediction of the volatility matrix.

Keywords　**high-dimensional volatility matrix, high-frequency data, mixed-frequency model, grouped factor model**

MSC(2020)　**62H25, 91G10, 97K80**

doi: 10.1360/SSM-2020-0304