論文の内容の要旨

論文題目　Estimation and statistical testing of the correlation between latent processes
(潜在確率過程間の相関の推定及び統計的検定)

氏　　名　木村　晃敏

本論文では, 後に述べるような表現を持つ確率過程 $\mathbb{Y}$ の観測データから, 潜在確率過程である $\mathbb{X}$ の相関推定、およびその統計的検定を扱う.

1　先行研究

金融資産価格のボラティリティは, 重要なリスク指標とみなされている. 実現ボラティリティは, 累積ボラティリティのよく知られた一致推定量で, その収束レートや漸近混合正規性についてもよく研究されており, 価格過程の高頻度サンプリングにより, 正確に推定できる. ふたつの金融資産価格の共分散や相関も、重要なリスク指標であり, 特にポートフォリオのリスク管理において重要である. 実現共分散, 実現相関についても, 高頻度サンプリングのもとで, 同様の漸近的性質を持つ.

実は, 二つの異なる株式のリターンの間の実現相関は, サンプリング頻度が上昇するにつれて, 減少することが知られている ([2]). この現象は, 取引タイミングの非同期性, マーケットマイクロストラクチャーによって引き起こされると考えられている.

金融市場において, ふたつの金融資産の取引のタイミングは, めったに一致しないため, 実現共分散を用いるには, 何かしらの方法で同期化することが必要となる. しかし, そのような推定量は, 同期化したサンプリングの幅を, 元の取引タイミングと比較して小さくすると, 深刻なバイアスを持つことが知られている ([5]). このような同期化を避けるため, 非同期共分散推定の方法が開発されてきている. 代表的なものに, フーリエ変換に基づくアプローチ ([7], [8]) と, 累積的共分散推定量 ([5, 3, 4], [9]) に基づく方法がある.

マーケットマイクロストラクチャーは, しばしば潜在価格過程にノイズを加えることでモデル化される. このモデリングは多くの成功をおさめ, またノイズ除去のテクニックも開発されてきた. 代表的なものとして, サブサンプリング法 ([12], [11]), プレアベレージング法 ([10], [6]), が挙げられる. また, 取引の非同期性と, マーケットマイクロストラクチャーを同時に扱う研究も数多くある.

マーケットマイクロストラクチャーの原因として, ビッドアスクスプレッド, 離散化誤差, 情報の非対称性が考えられている. 一方, このような加法的ノイズモデリングと, マーケットマイクロストラクチャーの関係性については, それほど詳しく説明されていない.

情報技術の発展により, 取引の観測時刻はますます正確に観測できるようになり, 今日では, 注文の超高頻度データも利用できるようになった. 近年, 超高頻度サンプリングにおいて, マーケットマイクロストラクチャーを, ノイズとしてではなく, リミットオーダーブックのダイナミクスとして捉える研究が行われている. 多くの研究では, リミットオーダーブックを, 点過程を用いてモデリングするアプローチが採用されており, 具体的には, ポアソン過程, ホークス過程, 二重確率的ポアソン過程を利用している. これらのアプローチは, 取引の非同期性を自然に扱うこともできる.

本研究では, 二重確率的ポアソン過程を含むモデルを扱う. まず, 間接的に高頻度観測される二つの潜在確率過程の相関に対する推定量を提案し, その漸近挙動を調べる.

2　モデルの概要, 相関推定, 漸近的性質

$\mathcal{B} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ をフィルター付き確率空間, $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,T]}$ をそのフィルターとする. $\mathbb{X} = (X^1, X^2)$ を $\mathcal{B}$ 上の $\mathbb{R}^2$-値 伊藤過程で,

$$\mathbb{X}_t = \mathbb{X}_0 + \int_0^t \mathbb{X}_s^0 \, ds + \int_0^t \mathbb{X}_s^1 \, dw_s \quad (t \in [0,T]), \tag{2.1}$$

と表現されるものとする. ただし, w は r-次元 $\mathcal{F}$-Wiener 過程, $\mathbb{X}_0$ は $\mathbb{F}_0$-可測確率変数, $\mathbb{X}^0$ は 2 次元 $\mathbb{F}$-適合過程, $\mathbb{X}^1$ は $\mathbb{R}^2 \otimes \mathbb{R}^r$-値 $\mathbb{F}$-適合過程とする. また, $\mathcal{B}$ 上の, 2 次元確率過程 $\mathbb{Y}^n = (Y^{n,1}, Y^{n,2})$ は,

$$\mathbb{Y}_t^n = \mathbb{Y}_0^n + \int_0^t a_n \mathbb{X}_s \, ds + \mathbb{M}_t^n \quad (t \in [0,T]),$$

の分解を持つものとする. ただし, a_n は, n に依存する正の数で, $a_n \to \infty$ $(n \to \infty)$ とし, $\mathbb{M}^n$ は 2 次元確率過程で, $\mathbb{M}_0^n = 0$ とする.

ここで, 観測時刻 $\Pi = (t_k)_{k=0,\ldots,b_n}$ $(0 = t_0 < t_1 < \cdots < t_{b_n} = T)$ に対し, $I_k = [t_{k-1}, t_k)$, $h_k = t_k - t_{k-1}$ とする. また, $b_n \to \infty$ $(n \to \infty)$ を仮定する.

このとき, 我々の目標は, $\mathbb{Y}$ の観測データから, 直接は観測されない $\mathbb{X}$ の相関に対する推定量を構成し, その $n \to \infty$ の下での漸近挙動を調べることとなる.

まずは, 共分散推定量を, $\alpha, \beta = 1, 2$ に対し,

$$S_n^{\alpha\beta} = \sum_{k=2}^{b_n} \Big(\frac{Y_{t_k}^{n,\alpha} - Y_{t_{k-1}}^{n,\alpha}}{a_n h_k} - \frac{Y_{t_{k-1}}^{n,\alpha} - Y_{t_{k-2}}^{n,\alpha}}{a_n h_{k-1}} \Big) \Big(\frac{Y_{t_k}^{n,\beta} - Y_{t_{k-1}}^{n,\beta}}{a_n h_k} - \frac{Y_{t_{k-1}}^{n,\beta} - Y_{t_{k-2}}^{n,\beta}}{a_n h_{k-1}} \Big)$$

のように定義する. さらに, 相関推定量を $C_n^{12} = S_n^{12} / \sqrt{S_n^{11} S_n^{22}}$ として定義する. 共分散推定量は, $\mathbb{Y}$ の観測データからは推定されない a_n に依存しているが, 相関推定量は, a_n に依存していないことに注意し, これを主に扱う.

我々は, 確率過程が有限の期間で観測されていて, 背後の確率過程の軌道が漸近的に推定できるような状況で, 相関推定量の一致性と漸近混合正規性 (混合正規分布への安定収束) を証明する. これは例えば, 市場が活発で, 多くの取引, 注文がなされているような状況であり, その際, a_n と b_n の $n \to \infty$ におけるバランスが重要となる.

3　漸近分散推定量, 統計的検定

推定量の漸近混合正規性は, スチューデント化された推定量の漸近正規性を意味し, 漸近分散推定量を構成できれば, 区間推定や, 統計的仮説検定に応用できる. そこで, 我々は, 相関推定量の漸近分散推定量を [1] に基づくものと, [4] に基づくものの, 二種類を提案し, それらの一致性を, 相関推定量の漸近混合正規性に必要な条件の下で示した.　また, [1] に基づく漸近分散推定量は, より強い条件の下で, より良い収束レートを持つことを示した.

4 シミュレーション, 実データへの応用

漸近分散推定量のふるまいを, 強度過程に相関を持つ, 二つの二重確率的ポアソン過程のシミュレーションに基づいて, 平均二乗誤差や, 区間推定の一種である, 信頼区間の被覆確率を用いて調べる. その結果, 前者の漸近分散推定量の漸近的ふるまいが良好であることが示される.

実データへの応用を考えると, a_n は, イベントの発生回数（例えば, 株式の取引回数, 注文回数など) は決まっていることから, 固定されているものと考えられ, 選べるパラメータは, b_n のみとなる. b_n は, 大きくなれば漸近的性質の仮定に反すると考えられ, 一方で小さくなれば, 収束が悪くなる. そこで, 観測幅 T/b_n と, 信頼区間の被覆確率の関係をプロットし, 各種パラメータに対し, 適切な観測幅はどのようにふるまうか調べる. また, 検定の検出力と観測幅の関係についても調べる. その結果, 適切な観測幅は主に, 二つのイベントの観測期間全体での発生回数のうち少ないほう, および相関の強さを表すパラメータに依存することが示唆される.

以上のことを踏まえたうえで, TOPIX Core30 の構成銘柄に関して, 分析を行う. TOPIX Core30 は, 東京証券取引所の一部上場銘柄のうち, 時価総額, 流動性の特に高い 30 銘柄で構成された株価指数である. まず, 注文データに対する相関推定量とその受容域の, 観測幅を動かした下でプロットする. その後, シミュレーションの観察を念頭に置いて, 適切な観測幅を決定する. さらに, グラフィカルモデルをデータに適用し, 株式の注文データ同士の関係を観察する.

参考文献

[1] Ole E. Barndorff-Nielsen and Neil Shephard. Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, 72(3):885–925, 2004.

[2] Thomas W Epps. Comovements in stock prices in the very short run. *Journal of the American Statistical Association*, 74(366a):291–298, 1979.

[3] Takaki Hayashi and Nakahiro Yoshida. Asymptotic normality of a covariance estimator for non-synchronously observed diffusion processes. *Annals of the Institute of Statistical Mathematics*, 60(2):367–406, 2008.

[4] Takaki Hayashi and Nakahiro Yoshida. Nonsynchronous covariation process and limit theorems. *Stochastic Processes and their Applications*, 121(10):2416–2454, 2011.

[5] Takaki Hayashi, Nakahiro Yoshida, et al. On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli*, 11(2):359–379, 2005.

[6] Jean Jacod, Yingying Li, Per A Mykland, Mark Podolskij, and Mathias Vetter. Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic processes and their applications*, 119(7):2249–2276, 2009.

[7] Paul Malliavin and Maria Elvira Mancino. Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics*, 6(1):49–61, 2002.

[8] Paul Malliavin, Maria Elvira Mancino, and Maria Cristina Recchioni. A non-parametric calibration of the HJM geometry: an application of Itô calculus to financial statistics. *Japanese Journal of Mathematics*, 2(1):55–77, Mar 2007.

[9] Per A Mykland. A Gaussian calculus for inference from high frequency data. *Annals of finance*, 8(2-3):235–258, 2012.

[10] Mark Podolskij, Mathias Vetter, and Margit Sommer. Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Available at SSRN 950344*, 2007.

[11] Lan Zhang et al. Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli*, 12(6):1019–1043, 2006.

[12] Lan Zhang, Per A Mykland, and Yacine Aït-Sahalia. A tale of two time scales. *Journal of the American Statistical Association*, 2012.