

# マーケット・マイクロストラクチャー・ノイズがある場合のボラティリティ推定

金谷 太郎

## I 序論

ボラティリティはファイナンスで最も重要視される量で，特に計量ファイナンスの分野では中心的な研究対象である。それは金融資産の収益率の分散，あるいは標準偏差で定義され，金融資産がもつリスクを計測する際に利用され，ときにはリスクそのものを示すものとして捉えられる。定義上，ボラティリティは直接観測できず，それが時間を通じて変化していくような現象をモデル化するには，ARCH (Autoregressive Conditional Heteroscedasticity) 型モデルやSV (Stochastic Volatility) モデルなどで潜在変数として間接的にモデル化されるのが主流であった。この種のボラティリティ変動モデルに関しては，Ghysels, Harvey, and Renault [1996], Morimune [2007] など多くの優れたサーベイ論文がある。

しかし，近年，金融高頻度データの利用可能性が拡大したことによって事情が変わった。Andersen, Bollerslev, Diebold, and Labys [2003] などの一連の研究によって，本来潜在変数である日次ボラティリティの代理変数として高頻度データから算出したリアライズド・ボラティリティ (Realized Volatility, あるいは Realized Variance, 以下 RV) を用いて，ボラティリティを直接モデル化することの有用性が

指摘された<sup>1)</sup>。これ以降，RV を用いたボラティリティのモデル化に関する研究とともに，RV が真のボラティリティをいかに効率よく推定できるかに関する研究が飛躍的に増大した。前者のRVのモデル化を特に取り上げた日本語のサーベイ論文としては森棟[2007]や渡部[2007]などがあるので，本稿では後者のRVやその修正推定量が真のボラティリティをどれだけ正確に推定することができるかという研究に焦点を当てる。

この研究課題は古くから研究されてきた確率過程の2次変動の理論に基づけば，一見非常に簡単な問題にみえる。すなわち，RVの精度を上げるには，可能な限り高頻度なデータを使えばよいというわけである。しかし，実際の金融データはマーケット・マイクロストラクチャー・ノイズ (以下MMN) を含むと考えられるため，あまりに高頻度なデータを使うことは避けるべきとされてきた。それゆえ，実証研究者はMMNの影響を受けないと考えられるデータの頻度を，経験則や主観的な判断によって決定しなければならなかったのである。

Zhou [1996] は「対数価格+ノイズ」というモデルで為替レートの収益率系列の一次の負の相関を表現し，その結果としてRVに上方バイアスが生じることを明らかにした。ノイズにやや強い仮定はあったものの，その仮定の下でバ

1) 冒頭で触れた通り，ボラティリティは収益率の分散として定義される場合と，その平方根の標準偏差として定義される場合，またその両者として定義される場合がある。ボラティリティを収益率の標準偏差として分散とはっきり区別する場合，Realized VolatilityはRealized Varianceの平方根であるが，本稿では両者を分散として扱い区別せず，いずれにしてもRVと略すことにする。

イアス修正推定量とその分散の振る舞いを明らかにした。この研究が、その後の「対数価格+ノイズ」というタイプのモデルを使ったMMN分析の先駆けとなった。そもそもMMNをどのようにモデル化するかは、議論の分かれるところであるが、本稿ではこの「対数価格+ノイズ」という現在計量ファイナンスの分野で主流となっているモデルと、その応用を中心にとりあげることにする。

Zhou [1996]以降の研究の流れを簡単にまとめると次のようになる。Bandi and Russell [2008]はよりノイズの仮定を緩め、その場合のRVの有限標本の平均二乗誤差(MSE)を導出し、そのMSEを最小にするデータの頻度を決定する方法を確立した。この研究によって、それまで研究者が主観的に決めざるをえなかったRVに使うデータの頻度の決定に、理論的根拠が与えられた。また、そのMSEを計算する際に必要なノイズのモーメントを計算する方法も示されており、その後の同種の研究に対して重要な貢献となっている。

Zhang, Mykland, and Ait-Sahalia [2005]は2種類の異なるデータ頻度で計算されたRVを使って、ノイズがある場合でも一致性をもつ推定量を提案した。この推定量はTwo Scale Estimator (以下TS推定量)とよばれ、高頻度データをバイアス修正、低頻度データをサブサンプリングとよばれる手法を介して分散の収束に寄与するように構成されている。また、Hansen and Lunde [2006]は、Newey and West [1987]の長期分散推定量を基にして、Kernel-Based Estimatorとよばれる推定量を提唱している。この推定量はTS推定量とほぼ同じであり、両者の漸近的な振る舞いは全く変わらない。さらに、推定量の理論的な分析に加え、ダウ・ジョーンズ工業平均株価指数に属する30種類の株価のMMNに関する実証研究がまとめられており、ノイズと株価の相関、あるいはノイズの系列相関など、独立ノイズの仮定では

捉えきれない現象が報告されている。

さらに、Barndorff-Nielsen, Hansen, Lunde, and Shephard [2008]は、上述のKernel-Based Estimatorの標本自己共分散に対する重み付けを一般化することによって、より広いクラスの推定量を考察した。その一般化された推定量はRealized Kernel (以下RK)と名付けられ、その漸近分布が導出されている。それによって、重み付けを決めるカーネル関数の形状と真のボラティリティへの収束速度の関係を整理している。Sun [2006]ではRKよりも一般的なクラスの推定量を定義し、有限標本において最適な推定量を導出している。ただし、導出の過程で、かなり強いノイズの仮定や推定量のクラスの制限が必要である。

本稿の構成は以下の通りである。次節ではRVの定義とその理論的根拠となる2次変動の理論を紹介する。第Ⅲ節では「対数価格+ノイズ」というタイプのモデルを導入し、RVに及ぼすノイズの影響をみる。第Ⅳ節ではRVを修正し、ノイズの影響を回避するさまざまな手法とそれらの性質を概観する。第Ⅴ節では、結論として今後のこの分野の研究の展望などに簡単に触れる。

## Ⅱ 2次変動とRV

この節では導入として、RVの理論的根拠である2次変動の理論を概観するために、資産価格がMMNの影響を受けずに観測される場合を考える。 $p_t$ を時刻 $t$ における金融資産の真の対数価格とし、それがブラウン運動セミマルチンゲールというクラスの確率過程にしたがっているとすると、単純化のためにドリフト項は考慮しないことにする。高頻度データを扱う文献ではしばしばこの単純化が採用される。高頻度収益率を想定した微小区間ではドリフト項はブラウン運動の項に対して無視できるほど小さくな

るからである。

$$p_t = p_0 + \int_0^t \sigma_s dW_s,$$

ただし、 $\sigma_t$  は cadlag (左連続で右極限をもつ) 確率過程で、 $W$  は標準ブラウン運動である。瞬間的な収益率の分散  $V(dp_t) = \sigma_t^2$  をスポット・ボラティリティとよび、これが興味の対象になる場合もあるが、ここでは区間  $[0, 1]$  における積分値を推定の対象にする。これを Integrated Volatility (あるいは Integrated Variance, 以下 IV) とよぶ。すなわち、 $IV = \int_0^1 \sigma_t^2 dt$  である。時間の単位は日であることが多く、このとき IV は日次ボラティリティなどによばれる。

時刻  $0 = t_0, t_1, \dots, t_n = 1$  の対数価格の系列を  $\{p_{t_i}\}_{i=0}^n$  とする。また、簡単化のために等間隔データの場合、すなわち  $t_i - t_{i-1} = 1/n$  の場合を考える。等間隔データの仮定は特に有限標本での MSE の表現やノイズの系列相関の定義に影響を与える。本稿では 1 変数のボラティリティ推定を主に紹介するので、共分散推定の場合、異なる金融資産が同時に観測されないことが原因となって生じる非同相バイアスを考慮する必要がないため、この簡単化を採用する。対数価格の一階差分  $x_i = p_{t_i} - p_{t_{i-1}}$  は、期間  $[t_{i-1}, t_i]$  の収益率をあらわす。このとき、

$$p \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n x_i^2}{RV^{(n)}} = \int_0^1 \sigma_t^2 dt, \quad (1)$$

が成立する。(1) の左辺を 2 次変動という。その有限和の部分  $\sum_{i=1}^n x_i^2$  を Realized Volatility とよび、IV の推定量として使う。このノイズを考慮しない場合の RV を、 $RV^{(n)} = \sum_{i=1}^n x_i^2$  と表記することにする。このとき RV の漸近分布が知られており、

$$\sqrt{n}(RV^{(n)} - IV) \xrightarrow{L} MN(0, 2IQ),$$

ただし、 $IQ = \int_0^1 \sigma_t^4 dt$  となる。IQ は Integrated Quarticity とよばれる量で、 $(1/3)\sum_{i=1}^n x_i^4$

で一致推定できる。MN は混合正規分布をあらわし、分散が確率変数であることをあらわす。詳細は Barndorff-Nielsen and Shephard [2002] を参照されたい。MMN を考慮しない場合、言い換えれば真の資産価格データが直接観測できる場合は、データ数を増やせば増やすほど RV の推定誤差分散を  $1/n$  のオーダーで減らすことができる。

### III マーケット・マイクロストラクチャー・ノイズ

Zhou [1996] は、為替収益率系列の一次の負の相関を表現し、その結果としての RV の上方バイアスの原因を明らかにするために、以下のような「対数価格+ノイズ」というモデルを導入した。実際に観測される対数価格  $\tilde{p}_t$  はノイズ  $\epsilon_t$  を伴って観測されると考えられ、

$$\tilde{p}_t = p_t + \epsilon_t,$$

である。観測される収益率を  $y_i = \tilde{p}_{t_i} - \tilde{p}_{t_{i-1}}$ 、また、ノイズの差分を  $\eta_i = \epsilon_{t_i} - \epsilon_{t_{i-1}}$  と書くことにすると、

$$y_i = x_i + \eta_i$$

である。

ここではひとまず、独立ノイズの仮定の下で、何が起こるかみてみることにする。独立ノイズとは、ノイズ  $\epsilon_{t_i}$  が対数価格  $p_t$  と自らの系列  $\epsilon_{t_j}$  ( $j \neq i$ ) の両方に対して独立なノイズのことをいう。また、ノイズの平均はゼロ、分散は  $\omega^2$  で一定であるとする。このとき  $E(y_i y_{i-1}) = -\omega^2$  となり、観測される収益率は一次の負の系列相関をもつ。もし資産価格過程がセミマルチンゲールであれば、収益率には系列相関はないはずである。しかし、このような高頻度収益率データの負の相関の例は、Zhou [1996] の為替レート例のほか、Dacorogna, Gencay, Muller, Olsen, and Pictet [2001] などで多数報告されている。

次に MMN があるときの RV の振る舞いをみる。観測可能な  $\{y_i\}_{i=1}^n$  を使って計算した RV を、 $RV^{(n)} = \sum_{i=1}^n y_i^2$  と書くことにする。このとき RV は

$$RV^{(n)} = RV_{*}^{(n)} + 2 \sum_{i=1}^n x_i \eta_i + \sum_{i=1}^n \eta_i^2 \quad (2)$$

の3つの項に分けることができる。独立ノイズの場合、(2)の第三項の期待値である  $2n\omega^2$  がバイアスとなる。これは一定区間に観測されるデータ数  $n$  に比例して、RV の上方バイアスは大きくなることを意味する。同時期に Delattre and Jacod [1997] は丸め誤差がある場合に RV の上方バイアスが無限に大きくなるという結果を得ている。丸め誤差は金融資産の場合、価格の離散性にあたり MMN のうちのひとつとして考えられる。図1は、Bandi and Russell [2008] に掲載されている IBM 社の株価の Volatility signature plot (または Variance signature plot, 以下 VSP) である。VSP とは横軸に使うデータの頻度、縦軸に対応する頻度で計算された RV の平均をプロットしたもので、RV 関連の文献ではバイアスの程度を視覚的に捉えるため

にしばしば用いられる<sup>2)</sup>。この図でも、あまりに高頻度なデータを使ってしまうと、RV の値が極端に大きくなっていく様子が見てとれる。

しかし、Hansen and Lunde [2006] は、ダウ・ジョーンズ工業平均株価指数に属する30種類の株価に関して、全ての銘柄の VSP が必ずしも図1のような形状にならないことを報告している。例えば、図1とは全く逆で、高頻度になるほど RV の値が小さくなるような VSP、すなわち下方バイアスを示す銘柄もあった。そのような現象を説明するには、独立ノイズでは不十分で、ノイズの仮定をより一般的にしなければならない。(2)の第三項の期待値は、系列相関を考慮してもコーシー・シュワルツの不等式により非負であるので、この下方バイアスが生じるためには、第二項の真の収益率  $y_i$  とノイズの差分  $\eta_i$  に負の相関がなければならない。また、独立ノイズによるバイアスを修正する推定量(後述(3)の  $RV_{*}^{(n)}$ )を使っても、バイアスが修正しきれていないという実証結果も報告されており、ノイズには系列相関があると結論付けられている。

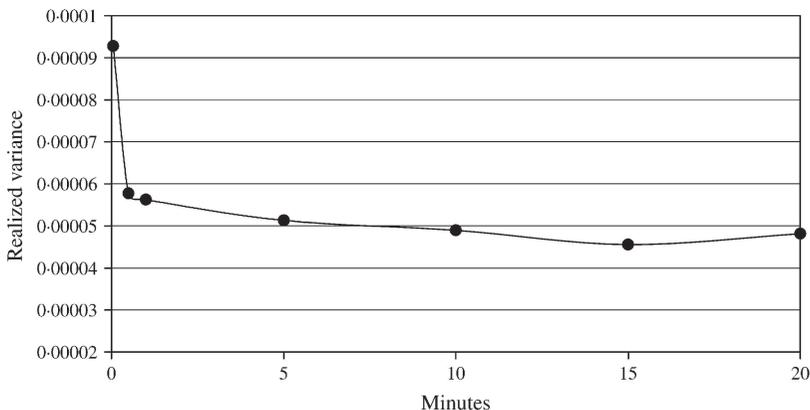


図1 IBM の Volatility signature plot

Bandi and Russell [2008] より転載。ただし、データ頻度ゼロ分は quote-to-quote データで計算された RV の平均という意味である。

2) Andersen, Bollerslev, Diebold, and Labys [2000] によって Volatility signature plot と名付けられて以降、計量ファイナンスの分野で定着したが、Fang [1996] にも同様の図がみられる。

この節の最後に、Bandi and Russell [2008] のRVの最適頻度に関する研究に触れておく。上に述べた通り、MMNがバイアスを引き起こすことはよく知られており、高頻度データ全体を使わずに、ある程度の頻度のデータを使うということは経験的に行われてきた。頻度を下げることによってバイアスは回避できるが、当然推定の分散は大きくなってしまふ。そこで、Bandi and Russell [2008]は「対数価格+ノイズ」というモデルの枠組みの中でRVに使うための最適頻度を導出した。バイアスと分散のトレードオフを考慮して、最適化の基準として有限標本のMSEが採用された。そのMSEを最小にするデータ数は、おおよそ  $n^* \approx (IQ/E(\eta^2)^2)^{1/3}$  となる。つまり、最大限利用できるデータ数  $n$  に対して最適な頻度は、 $n/n^*$  毎のデータを使ってRVを計算することである。ところで、この最適頻度やMSEを評価するためには、ノイズのモーメントを知る必要がある。Bandi and Russell [2008]では、適当な条件の下で

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n y_i^q = E(\eta^q), \quad q=2, 3, 4$$

と一致推定できることが示されている。

#### IV さまざまな修正推定量

本節では、ボラティリティ推定におけるMMNの影響を回避するための、さまざまなRVの修正法に関する研究を紹介する。

Zhou [1996]では、上述の通り独立ノイズが仮定されていたため、収益率には一次の自己相関があった。それゆえ、以下のように、一次の標本自己共分散項を追加することによって、バイアスを修正することが考えられた。

$$RV_{AC1}^{(n)} = RV^{(n)} + \sum_{i=2}^n y_i y_{i-1} + \sum_{i=1}^{n-1} y_i y_{i+1} \quad (3)$$

独立ノイズなら、第二項・第三項の期待値はともに  $-n\omega^2$  なので、 $RV^{(n)}$  のバイアス  $2n\omega^2$  と

打ち消しあふ。また、ボラティリティ一定の条件を追加して、有限標本でのMSEが導出され、それを最小にするデータ数がおおよそ  $\sqrt{3}\sigma^2/(2\omega^2)$  であることも示されている。この推定量もRV同様一致性をもたないが、サブサンプリングという手法を使って、分散が改善されることが示唆されている。例えば、1分収益率のデータをもっていたとして、MMNの影響を考えて15分収益率のデータを使うことを考える。15分収益率のデータは15通りでき、15分RV推定量も15個計算できる。その平均をとるのがサブサンプリングである。

Zhang, Mykland, and Ait-Sahalia [2005]は、そのサブサンプリングを使って一致推定量を提案した。上述のように低頻度収益率データ系列を  $K$  通りつくることを考えたとき、その  $k$  番目の系列を次のように定義する。

$$y_{j,k} = \sum_{i=1}^K y_{i+K(j-1)+k-1}, \quad j=1, 2, \dots, m_k$$

ただし、 $m_k = \lfloor n/K \rfloor$  で  $\cdot$  はガウス記号である。 $k$  番目の低頻度データから計算されたRVを

$$RV^{(K,k)} = \sum_{j=1}^{m_k} y_{j,k}^2,$$

とする。 $\{RV^{(K,k)}\}_{k=1}^K$  のうち  $RV^{(K,1)}$  だけを使うのが、従来の低頻度データRVを使ってMMNを避ける方法であった。しかし、RVのサブサンプリングとは以下のように  $K$  個のRVの平均をとることをいう。すなわち、

$$SRV^{(K)} = \frac{1}{K} \sum_{k=1}^K RV^{(K,k)}, \quad (4)$$

であり、低頻度データを使うことによってバイアスを軽減し、MSEを減少させる効果が期待できる。ただし、あくまでバイアスの「軽減」であり、バイアスはいくらか残るので(5)の第二項による修正が必要である。

また、 $K$  がデータ数に対して適当な程度で大きくなれば、分散を収束させることもできると考えられる。Zhang, Mykland, and Ait-Sahalia [2005]は、(4)の低頻度データを使ったサブサ

ンプリング RV のバイアスを高頻度データ RV を使って修正し、以下の推定量を提案した。

$$TS^{(K)} = SRV^{(K)} - \frac{\bar{m}}{n} RV^{(n)}, \quad (5)$$

ただし、 $\bar{m} = (1/K) \sum_{k=1}^K m_k$  である。この推定量は、高頻度・低頻度の2つの異なる頻度のデータを使ってつくられているので、Two Scale 推定量とよばれている。独立ノイズというやや強い仮定となっているが、MMN の影響下でも一貫性をもつ、初の RV 推定量と認識されている。漸近分布は本稿では省略するが、漸近正規性をもち、 $n^{1/6}$  のオーダーで真のボラティリティ、すなわち IV に収束する。MSE 基準の、最適な  $\bar{m}$  の目安は  $(IQ/(3\omega^2))^{1/3}$  である。Zhang [2006] は、さらに異なる頻度のデータを使うことで、TS 推定量より効率を改善することができることを示している。この推定量は Multi-Scale 推定量 (以下 MS 推定量) とよばれ、収束のオーダーは  $n^{1/4}$  である。これは、Ait-Sahalia, Mykland, and Zhang [2005] がパラメトリックな設定の下で導出した最尤推定量と同じオーダーである。

Hansen and Lunde [2006] は、前節で触れたように、独立ノイズの仮定では彼らの実証結果を捉えきれないことを示した。そこで、Newey and West [1987] の自己相関に対してロバストな長期分散推定量を応用して、(3) の一期ラグ項に加えて、 $H$  期までのラグをバートレット・ウェイトで重み付けて追加した。

$$HL^{(H)} = \gamma_0 + \sum_{h=1}^H \left(1 - \frac{h-1}{H}\right) (\gamma_h + \gamma_{-h}), \quad (6)$$

ただし、 $\gamma_q = \sum_{i=1}^n y_i y_{i-q}$  であり、 $\gamma_0 = RV^{(n)}$ 、また  $H=1$  のときは  $HL^{(1)} = RV_{AC1}^{(n)}$  である。この推定量は Kernel-Based Estimator とよばれ、TS 推定量とは異なるモチベーションで考案されたが、両者はほぼ同じ推定量であることがわかっており、漸近的性質は同一である。

Barndorff-Nielsen, Hansen, Lunde, and Shephard [2008] は (6) の自己共分散項への重

み付けを一般化した以下の推定量を分析した。

$$RK^{(H)} = \gamma_0 + \sum_{h=1}^H k\left(\frac{h-1}{H}\right) (\gamma_h + \gamma_{-h}), \quad (7)$$

ただし、 $k(x)$  は定義域が  $[0, 1]$  で、確率変数ではない。この推定量を Realized Kernel といい、既出の推定量を全て含んでいる。漸近分布はカーネル関数  $k(x)$  の形状によって異なる。

表1に、主なカーネル関数と RK の収束のオーダー、効率をまとめた。ここでの効率は漸近分散の相対的な大きさをあらわす。上段の3つのカーネルは、 $k(0)=1$ 、 $k(1)=0$  を満たす。 $H = c_0 n^{2/3}$  ( $c_0$  は推定可能な定数) としたとき、RK が IV に  $n^{1/6}$  のオーダーで収束する。Bartlett カーネルを使った RK は、上述の通り、TS 推定量・Kernel-Based Estimator と同じ推定量である。

中段のカーネルには、 $k'(0)=k'(1)=0$  という条件が加わる。このクラスの推定量は、 $H = c_0 n^{1/2}$  としたとき、 $n^{1/4}$  のオーダーで収束する。Ait-Sahalia, Mykland, and Zhang [2005] が、ボラティリティ一定、正規独立ノイズなどのパラメトリックな設定の下で導出した最尤推定量と、同じオーダーを達成している。Turkey-Hanning カーネルの次数を増やしていくと、最尤推定量の効率 8 に近づく。なお、Zhang [2006] の MS 推定量は、RK の形で書くと Cubic カーネルをもつことがわかっている。

下段は、定義域に  $[0, \infty)$  をとるカーネル、すなわち、無限のラグの自己共分散項  $\gamma_h + \gamma_{-h}$  をとることが可能なカーネルである。これらの推定量は、 $H = (\omega/IQ)n^{1/2}$  としたとき、 $n^{1/4}$  のオーダーで収束する。特に Optimal カーネルを使ったとき、その漸近分散の1次の項は、最尤推定量のそれと同じである。しかし、無限ラグの自己共分散項を実用する際、有限個で打ち切ることによる近似の問題や計算量の問題がある。実際、中段の有限の Turkey-Hanning 等の漸近効率が十分下限の 8 に近いので、あえて下段のカーネルを使うことにメリットを感じられ

表1 カーネルの種類とその漸近分散の比較

カーネル名	$k(x)$	オーダー	効率
Bartlett (TS)	$1-x$	$n^{1/6}$	0.76
Epanechnikov	$1-x^2$	$n^{1/6}$	1.31
2nd order	$1-2x+x^2$	$n^{1/6}$	0.68
Cubic (MS)	$1-3x^2+2x^3$	$n^{1/4}$	9.04
5th order	$1-10x^3+15x^4-6x^5$	$n^{1/4}$	10.2
6th order	$1-15x^4+24x^5-10x^6$	$n^{1/4}$	12.1
7th order	$1-21x^5+35x^6-15x^7$	$n^{1/4}$	13.9
8th order	$1-28x^6+48x^7-21x^8$	$n^{1/4}$	15.7
Parzen	$1-6x^2+6x^3(0 \leq x \leq 1/2)$ $2(1-x)^3(1/2 < x \leq 1)$	$n^{1/4}$	8.54
Turkey-Hanning1	$\sin^2(\pi(1-x)/2)$	$n^{1/4}$	9.18
Turkey-Hanning2	$\sin^2(\pi(1-x)^2/2)$	$n^{1/4}$	8.29
Turkey-Hanning5	$\sin^2(\pi(1-x)^5/2)$	$n^{1/4}$	8.07
Turkey-Hanning16	$\sin^2(\pi(1-x)^{16}/2)$	$n^{1/4}$	8.02
Turkey-Hanning $\infty$	$\sin^2(\pi \exp(-x)/2)$	$n^{1/4}$	8.01
Quadratic spectral	$(3/x^2)(\sin(x)/x) - \cos x$	$n^{1/4}$	9.38
Dirichlet (Fourier)	$\sin(x)/x$	$n^{1/4}$	11.7
Fejer	$(\sin(x)/x)^2$	$n^{1/4}$	8.89
Optimal	$(1+x)e^{-x}$	$n^{1/4}$	8.00

Barndorff-Nielsen, Hansen, Lunde, and Shephard [2008] より一部抜粋して転載。

効率（各カーネル関数を使ったときのRKの漸近分散の相対的な大きさ）は同じ収束のオーダー間で比較可能である。Ait-Sahalia, Mykland, and Zhang [2005] の最尤推定量はオーダー  $n^{1/4}$  で効率 8 となり、下限となる。

ない。

また、表1下段のクラスのDirichletカーネルを使ったRKは、Malliavin and Mancino [2002] のフーリエ推定量のRK表現となっている。フーリエ推定量は他の時間領域でのボラティリティ推定量とは異なり、周波数領域での分析により導出された新しい推定量であると認識され、MMNに対してロバストであるというシミュレーションあるいは実証研究の結果が、Oya [2005] や Nielsen and Frederiksen [2008] などで報告されてきた。また、Kanatani [2004b] や Mancino and Sanfelici [2007] によって、RKの形に書き換えられることが示され、ノイズに由来するバイアスを緩和する理由が、理論的に明らかになった。しかし、表1にもあるように、他のカーネルに比べて効率にお

いて劣っているのが、実用上魅力がある推定量であるとは言いがたい。

RKに関しては、そのサブサンプリング版が Barndorff-Nielsen, Hansen, Lunde, and Shephard [2007] において分析され、その実用性が Barndorff-Nielsen, Hansen, Lunde, and Shephard [2009] で議論されており、今後ボラティリティ推定において標準的な手法になることが予想される。

最後に、有限標本での最適化に取り組んでいる Sun [2006] の研究を紹介する。以下のような二次形式は、既存の推定量全てを含む。

$$WRV = \sum_{i=1}^n \sum_{j=1}^n w_{ij} y_i y_j,$$

この推定量のクラスは Kanatani [2004a] において定義され、Weighted Realized Volatility と

名付けられている。Sun [2006] では、この推定量を Quadratic (Form) Estimator とよんでいる。例えば、この表現において、RK は  $|i-j|=h$  毎に同じ重み付けをした特殊形となる。Sun [2006] は、正規独立ノイズ、ボラティリティ一定、さらに推定量を不偏推定量のクラスに限って、有限標本 MSE を最小化するウェイト付けが、

$$w_{ij}^* = \frac{\lambda}{\sqrt{\lambda(\lambda+4)}} \frac{\cosh(\alpha(n+1-|i-j|)) - \cosh(\alpha(n+1-i-j))}{\sinh(\alpha(n+1))},$$

ただし、

$$\lambda = \frac{IV}{n\omega^2}, \quad \alpha = \log\left(\frac{1}{2}\lambda + \frac{1}{2}\sqrt{4\lambda + \lambda^2 + 1}\right),$$

と明示的に書けることを示した。この最適ウェイトを使った推定量は、最良二次不偏推定量 (Best Quadratic Unbiased Estimator, 以下 BQU 推定量) と名付けられている。最適ウェイト  $w_{ij}^*$  には推定すべき対象の IV が入っており、もし IV が既知ならば、収束のオーダーは  $n^{1/2}$  であり、TS 推定量などで計算した推定値を使った場合には、先に述べた RK や最尤推定量と同様、BQU 推定量は  $n^{1/4}$  で収束することも証明されている。

## V 結語

以上、高頻度データを用いたボラティリティの推定問題を、RV とその修正推定量を中心に概観した。まず、近年の高頻度データの利用可能性の拡大が、RV という古典的で簡便な手法に光を当てた。それを受けて、高頻度データを扱う上で避けることができない MMN の問題に、「対数価格+ノイズ」という Zhou [1996] 以降のモデルを用いて、多くの研究者がこの問題に取り組んでいる。本論でみたように、簡単なバイアス修正、最適頻度の導出、一致推定量の構成、ノイズの仮定の拡張、さらに推定の効率を上げる手法の開発、そしてそれらの推定量を一般化した統一理論といった研究が、ここ数

年の間に飛躍的に進んだ。しかし、「対数価格+ノイズ」というモデルの下で RV 推定量の精度を上げる方向には研究は進んでいるが、Hansen and Lunde [2006] のような、ノイズの性質を十分に議論するような実証研究はまだ少ない。このことは、その利用可能性が拡大したとはいえ、一部の欧米の研究機関を除いて、十分な高頻度データが必ずしも容易に利用できないことにも原因がある。特に、欧米に比べて、日本の金融高頻度データは利用が難しいのが現状である。しかし、今後高頻度データの利用可能性がより高まり、実証研究が進めば、広範な市場と金融資産の MMN の統計的性質が明らかになるはずである。「対数価格+ノイズ」というデータ記述型のモデルと、既存のマーケット・マイクロストラクチャーの理論との整合性や解釈が議論され、ファイナンスの理論研究においても大きく貢献していくと期待される。マーケット・マイクロストラクチャーに関する理論については O'Hara [1995] によくまとめられている。

本稿では、主に分散推定についての研究を紹介したが、最後に共分散推定について少し触れておく。恒等式  $4 \cdot \text{Cov}(X, Y) = V(X+Y) - V(X-Y)$  から、共分散に関しては、分散の推定がわかれば十分のように考えられがちである。しかし、取引データのような金融高頻度データを扱う場合は、 $X$  と  $Y$  が同時観測されず、 $X+Y$  や  $X-Y$  といったポートフォリオが観測不可能であるので、共分散推定には分散推定にはない固有の問題が含まれることになる。データの同時観測が仮定されている場合で、かつノイズを考えない場合の Realized Covariance (RC) Matrix については Barndorff-Nielsen and Shephard [2004]、ノイズを考える場合の RC の最適頻度については Bandi and Russell [2005] などがある。Hayashi and Yoshida [2005] は、同時観測されない2つの系列の共分散の不偏推定量を提案して、その一致性を示

した。この推定量は非同期バイアスの問題を解決する手法として認識され、後に Cumulative Covariance Estimator (CC 推定量) と名付けられた。Griffin and Oomen [2009] は、上述の Bandi and Russell [2008] が RV の最適頻度を導出したのと同様に、CC 推定量の有限標本の MSE を分析し、CC 推定量を計算する際に最適なデータの頻度を求めた。Voev and Lunde [2007] では、より一般的な状況で CC 推定量のバイアスあるいは分散が考察され、そのサブサンプリング版の有用性が指摘されている。有限標本では、Kanatani [2008] がカーネルの選択法、Kanatani [2009] が CC 推定量のサブサンプリングについて研究している。しかし、全体としては、共分散の場合は観測の非同期性を考慮しなければならないため、特に漸近論において、分散の場合に比べて研究は進んでおらず、その重要性を鑑みれば今後の研究の進展が期待される。

## 謝辞

本稿の作成にあたり森棟公夫教授と京都大学経済研究所・奥井亮氏から頂いた有益なコメントに感謝する。また、日本学術振興会科学研究費(研究課題番号: 21730174)による財政的支援にも感謝する。なお、本稿に含まれる誤りや漏れは全て筆者に属するものである。

## 参考文献

森棟公夫 [2007] 「実現ボラティリティ」『現代経済学の潮流 2007』3-32 ページ。  
 渡部敏明 [2007] 「Realized volatility—サーベイと日本の株式市場への応用」『経済研究 58』(4), 352-373 ページ。  
 Ait-Sahalia, Y., P. Mykland and L. Zhang [2005] “How Often to Sample a Continuous-time Process in the Presence of Market Microstructure Noise,” *Review of Financial Studies*, **18**, pp. 351-416.  
 Andersen, T. G., T. Bollerslev, F. X. Diebold and P. Labys [2000] “Great Realizations,” *Risk*, **13**(3), pp. 105-108.

Andersen, T. G., T. Bollerslev, F. X. Diebold and P. Labys [2003] “Modeling and Forecasting Realized Volatility,” *Econometrica*, **71**(2), pp. 579-625.  
 Bandi, F. M. and J. R. Russell [2005] “Realized Covariation, Realized Beta and Microstructure Noise,” Working Paper, University of Chicago.  
 Bandi, F. M. and J. R. Russell [2008] “Microstructure Noise, Realized Variance, and Optimal Sampling,” *The Review of Economic Studies*, **75**, pp. 339-369.  
 Barndorff-Nielsen, O., P. R. Hansen, A. Lunde and N. Shephard [2007] “Subsampling Realized Kernels,” *Journal of Econometrics*, forthcoming.  
 Barndorff-Nielsen, O., P. R. Hansen, A. Lunde and N. Shephard [2008] “Designing Realized Kernels to Measure the Ex Post Variation of Equity Prices in the Presence of Noise,” *Econometrica*, **76**(6), pp. 1481-1536.  
 Barndorff-Nielsen, O., P. R. Hansen, A. Lunde and N. Shephard [2009] “Realized Kernels in Practice: Trades and Quotes,” *Econometrics Journal*, forthcoming.  
 Barndorff-Nielsen, O. and N. Shephard [2002] “Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models,” *Journal of the Royal Statistical Society, Series B*, **64**, pp. 253-280.  
 Barndorff-Nielsen, O. and N. Shephard [2004] “Econometric Analysis of Realized Covariation: High Frequency Covariance, Regression and Correlation in Financial Economics,” *Econometrica*, **72**, pp. 885-925.  
 Dacorogna, M. M., R. Gencay U. Muller, R. B. Olsen, and O. V. Pictet [2001] *An Introduction to High-Frequency Finance*, Academic Press, London.  
 Delattre, S. and J. Jacod [1997] “A Central Limit Theorem for Normalized Functions of the Increments of a Diffusion Process in the Presence of Round Off Errors,” *Bernoulli*, **3**, pp. 1-28.  
 Fang, Y. [1996] “Volatility Modeling and Estimation of High-frequency Data with Gaussian Noise,” Ph. d. thesis, MIT, Sloan School of Management.  
 Ghysels, E., A. C. Harvey, and E. Renault [1996] “Stochastic Volatility,” in *Handbook of Econometrics, Vol. 14, Statistical Methods in Finance*, ed. by G. S. Maddala and C. R. Rao. Amsterdam :

- North-Holland, 119-191.
- Griffin, J. E. and R. C. A. Oomen [2009] "Covariance Measurement in the Presence of Non-synchronous Trading and Market Microstructure Noise," Working Paper, University of Warwick.
- Hansen, P. R. and A. Lunde [2006] "Realized Variance and Market Microstructure Noise" (with discussions), *Journal of Business, & Economic Statistics*, **24**, pp. 127-161.
- Hayashi, T. and N. Yoshida [2005] "On Covariance Estimation of Non-synchronously Observed Diffusion Processes," *Bernoulli*, **11**, pp. 359-379.
- Kanatani, T. [2004a] "High Frequency Data and Realized Volatility," Ph. d. thesis, Graduate School of Economics, Kyoto University.
- Kanatani, T. [2004b] "Integrated Volatility Measuring from Unvenly Sampled Observations," *Economics Bulletin*, **3**(36), pp. 1-8.
- Kanatani, T. [2008] "Finite Sample Analysis of Weighted Realized Covariance with Noisy Asynchronous Observations," Working Paper, Kyoto University.
- Kanatani, T. [2009] "Subsampling Cumulative Covariance Estimator," Working Paper, Shiga University.
- Malliavin, P. and M. E. Mancino [2002] "Fourier Series Method for Measurement of Multivariate Volatilities," *Finance and Stochastics*, **6**, pp. 49-61.
- Mancino, M. E. and S. Sanfelici [2007] "Robustness of Fourier Estimator of  $\sigma^2$  in the Presence of Microstructure Noise," *Computational Statistics and Data Analysis*, **52**, pp. 2966-2989.
- Morimune, K. [2007] "Volatility Models," *The Japanese Economic Review*, **58**(1), pp. 1-23.
- Nielsen, M. O. and P. Frederiksen [2008] "Finite Sample Accuracy and Choice of Sampling Frequency in Integrated Volatility Estimation," *Journal of Empirical Finance*, **15**, pp. 265-286.
- Newey, W. and K. West [1987] "A Simple Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, **55**, pp. 703-708.
- O'Hara, M. [1995] *Market microstructure theory*, Oxford: Blackwell Publishers.
- Oya, K. [2005] "Measurement of Volatility of Diffusion Processes with Noisy High Frequency Data," *Proceeding of MODSIM05*, pp. 940-945.
- Sun, Y. [2006] "Best Quadratic Unbiased Estimators of Integrated Variance in the Presence of Market Microstructure Noise," Working Paper, University of California, San Diego.
- Voev, V. and A. Lunde [2007] "Integrated Covariance Estimation Using High-frequency Data in the Presence of the Noise," *Journal of Financial Econometrics*, **5**, pp. 68-104.
- Zhang, L. [2006] "Efficient Estimation of Stochastic Volatility Using Noisy Observations: a Multi-scale Approach," *Bernoulli*, **12**, pp. 1019-1043.
- Zhang, L., P. Mykland and Y. Ait-Sahalia [2005] "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-frequency Data," *Journal of the American Statistical Association*, **100**, pp. 1394-1411.
- Zhou, B. [1996] "High-frequency Data and Volatility in Foreign Exchange Rates," *Journal of Business, & Economic Statistics*, **14**(1), pp. 45-52.