学位請求論文要旨

金融データの長期記憶モデルによる分析

一橋大学大学院経済学研究科博士課程
山口　圭子

1　はじめに

ボラティリティはリスク管理あるいは派生商品の価格付けにおいて重要な変数である。ボラティリティは変動するということが知られているが, 直接観測可能ではないので, その変動を捉えるモデルとして, ARCH(autoregressive conditional heteroskedasticity) 型モデルや SV(stochastic volatiltiy) モデルが考えられてきた。

ところが, 近年, 高頻度データが利用可能となり, 高頻度データから計算される Realized Volatility (RV) がモデルに依存しない高精度な推定量であるとして注目を集めている。

ボラティリティには RV のほかにもオプション価格から計算されるインプライドボラティリティ(IV) がある。IV の算出にはブラックショールズ (BS) モデルが広く使われているが, ボラティリティを満期まで一定であると仮定するので, ボラティリティが変動する場合, BS モデルを用いて計算したインプライドボラティリティ（BSIV）にはバイアスが生じる。そこで, 原資産価格のボラティリティの変動を許容するモデルフリーインプライドボラティリティ (MFIV) が注目を集めている。実際, シカゴオプション取引所 (Chicago Board Options Exchange) のボラティリティ・インデックス (VIX) では, S&P500 指数の MFIV が採用されている。

RV や IV, 特に RV には長期記憶性があることが知られている。本論文は, RV や IV(BSIV, MFIV) を長期記憶モデルにより分析したものである。

長期記憶性とは, 自己相関関数 $\rho(h)$ が

$$\sum_{h=1}^{\infty}|\rho(h)| = \infty \tag{1}$$

を満たすことをいう。例えば, 矢島 (2003) に詳しい解説がある。

パラメトリックモデルとしては, 次の ARFIMA(p, d, q) モデル (Hosking (1981), Granger and Joyeux (1980)) が挙げられる。

$$\Phi(L)(1-L)^d X_t = \Psi(L)Z_t \tag{2}$$

ここで, L はラグオペレータであり,

$$\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p, \qquad \Psi(L) = 1 - \psi_1 L - \cdots - \psi_q L^q$$

は $|\omega| \leq 1$ に対して $\Phi(\omega) \neq 0, \Psi(\omega) \neq 0$ であるとする。X_t が定常ならば

$$\rho(h) = O(h^{2d-1}), \quad h \to \infty \tag{3}$$

となり, $d < \frac{1}{2}$ ならば定常であり, $0 < d$ ならば長期記憶性を持つことがわかる。

Chen and Deo (2006) では, ARFIMA モデルにおいて短期変動をとらえるための AR や MA の次数選択を誤ると, 長期変動を表す d も一致性を失う場合があることを示した。そこで, 短期変動のモデル化の影響を受けずに長期変動を抽出することを目的としたセミパラメトリックモデルが注目されている。

スペクトル密度関数が原点近傍で以下の (4) 式のように近似できるとする。

$$f_X(\lambda) \sim G_X|\lambda|^{-2d} \quad \text{as} \quad \lambda \to 0 \tag{4}$$

ただし, $G_X > 0$ は定数である。(4) 式を用いて構成された推定量をナローバンド推定量という。(4) 式は原点の近傍のみで成立するので, サンプルサイズが増加するとともに, 利用

する周波数の上限を 0 に収束させていく。Geweke and Porter-Hudak (1983)(以下, GPH と省略), Kunsch (1987), Robinson (1995a,b) 以降, さまざまな方向に研究が進められた。詳しくは Velasco (2006) を参照。

Moulines and Soulier (1999), Hurvich and Brodsky (2001) らの提案したブロードバンド推定量は GPH 推定がもとになっているが, スペクトル密度関数を

$$f_X(\lambda) = |2\sin(\lambda)|^{-2d} f^*(\lambda) \tag{5}$$

として, 原点の近傍だけでなく高周波の部分を含めるものである。$\log f^*(\lambda)$ を cos 関数系 $\{h_j\}$ で

$$\log f^*(\lambda) = \sum_{j=0}^{\infty} \theta_j h_j \tag{6}$$

のように展開して, サンプルサイズ $T \to \infty$ に応じて $j \to \infty$ として利用する。なお, この $f^*(\lambda)$ を展開する考え方は, ナローバンド推定量のバイアス修正 (Andrews and Guggenberger (2003), Andrews and Sun (2004), Guggenberger and Sun (2006)) やシグナル・プラス・ノイズモデル (Sun and Phillips (2003), Hurvich et al. (2005), 本論文第 2 章参照) と密接にかかわっている[1] 。以上は周波数領域でのセミパラメトリックモデルである。

時間領域でセミパラメトリックモデルを考えたものとしては, Bhansali et al. (2006) が挙げられる。ARFIMA$(p, d, 0)$ モデルにおいて, サンプルサイズ $T \to \infty$ に応じて $p \to \infty$ とするものである。時間領域ではバンドを扱えないので, ナローバンド推定とは別の方法を考えなければならない。ナローバンド推定量は, 原点の近傍でスペクトル密度関数の長期変動を表す部分は発散するのに対し, 短期変動の部分は有限であることを利用すればよく, 理論的に扱いやすいのに対して, 時間領域では長期変動と短期変動の区別は (1.1) 式にもあるように $\rho(h) \to \infty$ の挙動をみなければならない。無限期先のことを取り扱うのは理

[1] $f^*(\lambda)$ を cos 関数ではなく, 多項式で展開している。

論的に簡単ではない。実際, Bhansali et al. (2006) の方法を用いるには, d の一致推定値があらかじめ必要である。以上のことから, 現状ではセミパラメトリックモデルにおいては周波数領域が主流である。他にも周波数領域と時間領域の中間である wavelet 領域でも考えられている。(詳しくは田中 (2006) を参照。) もっとも, 時間領域は研究されはじめたばかりであり, wavelet 領域は徐々に研究がすすんでいるがまだ解明されていない統計的性質が多く (例えば, wavelet 領域での GPH 型推定量の漸近正規性 (Moulines et al. (2007)) など), 今後の発展に期待したい。

以上, 簡単に d の推定についてみてきた。上記のものをはじめとしてさまざまな推定量が提案されてきたが, そのなかで生き残っている推定量は, それぞれそれなりに良い点があり, 分析目的やデータの性質によって使い分けすることが望ましい。本論文では, 予測を行う場合は信頼区間の狭い ARFIMA モデルを用いた (第 4 章)。それに対して, ロバストに推定・検定を行いたい場合にはセミパラメトリックモデルを用いた (第 2, 5 章)。

以下では各章の分析結果を簡単にまとめる。なお, 第 2 章は山口 (2007), 第 3 章は 2007 年 9 月の統計学会での報告, 第 4 章は山口 (2008) の研究成果をまとめたものである。第 5 章は本論文のために新たに書いた。また, 第 5 章で行った共和分検定の解説が本論文の最後に付録としてまとめられている。

2　本論文の構成と各章の概略

第 2 章 長期記憶過程におけるノイズの有無の検定

本章では, ある変数が長期記憶過程に従うシグナルとホワイト・ノイズに従うノイズの和として表されるとした場合に, ノイズが存在するかどうかを検定するための新たな統計量をセミパラメトリックモデルの枠組みで考えた。同様な統計量に Sun and Phillips (2003)

が考えた統計量があるので，モンテカルロ実験によって，彼らの統計量とここで新たに考えた統計量の性質を調べた。また，RV は長期記憶過程に従うことが指摘されているので，日経 225 株価指数の 5 分ごとのリターンを用いて計算した RV と 1 分ごとのリターンを用いて計算した RV に対してノイズが存在するかどうか検定を行った。その結果，5 分ごとのリターンを用いて計算した RV より 1 分ごとのリターンを用いて計算した RV のほうがより多く棄却された。

第 3 章 日経 225 株価指数インプライドボラティリティの計算方法に関して

本章では，Jiang and Tian (2005, 2007) と VIX の 2 つの方法で日経 225 株価指数オプション価格からを計算し，それらの統計的な性質について分析しつつ，両者の比較を行った。その結果，VIX よりも MFIV のほうが標準偏差が大きく，また最大値はより大きく，最小値はより小さく，Jiang and Tian (2007) の指摘と整合的な結果が得られた。

第 4 章 日経 225 株価指数インプライドボラティリティの情報量 ～“Realized Volatility”予測の観点から～

インプライドボラティリティは現在から満期までの投資家の予測を表していると考えられるので，将来の RV に対して，過去や現在の RV と異なる情報を有している可能性が考えられる。Koopman et al. (2005) は，S&P500 指数に GARCH モデルや SV モデルに BSIV を外生変数として加えて予測パフォーマンスが改善するかどうかを調べている。渡部・山口 (2006) では日経 225 株価指数 RV を ARFIMAX モデルでモデル化し，それに BSIV を外生変数として加えたものも調べている。本章では BSIV だけでなく，3 章で計算した VIX や MFIV も BSIV の代わりに外生変数として加えて，予測パフォーマンスを比較した。MFIV

は VIX, BSIV と比較して, 1 期先の RV に対してより多くの追加的な情報を含んでいることが分かった。このことは, Jiang and Tian (2007) が VIX の計算方法は近似誤差が大きいと指摘したこと (第 3 章参照) と整合的である。

第 5 章 フラクショナル共和分モデルによるボラティリティの分析

4 章とは別の方法で RV と IV を比較する。Christensen and Prabhala(1998) は以下の式を考えた。

$$y_t = \alpha + \beta x_t + e_t, \quad t = 1, \ldots, T \tag{7}$$

ここで, $y_t = \ln \sigma_{RV}$, $x_t = \ln \sigma_{IV}$ である。(7) 式から 3 つ仮説を検証できるとしている。もし, IV が何らかの情報を有している場合 β はゼロにはならない。そして, IV が RV を不偏予測していれば, $\alpha = 0$, $\beta = 1$ となるはずである。さらに, IV が有効ならば e_t はホワイト・ノイズで他の市場情報と無相関になる。実証結果では, β が 1 より小さいと報告されている。

その理由の 1 つは, IV として BSIV が用いられているが, ブラックショールズモデルではボラティリティを満期まで一定であると仮定するので, ボラティリティが変動する場合, BSIV にはバイアスが生じてしまい, このことが影響しているのではないかということである。Jiang and Tian(2005) では BSIV と MFIV, それぞれで β を推定し, BSIV では有意に 1 より小さいが, MFIV では $\beta = 1$ という帰無仮説を棄却できないという結果が報告されている。

2 つめの理由は, リスクプレミアムの存在である。Bollerslev and Zhou(2006) では, (7) 式で $y_t = \sigma_{RV}$, $x_t = \sigma_{MFIV}$ とし, 対数価格過程 p_t が連続 SV モデルに従っている場合に, リスクプレミアムが存在すると $\beta < 1$ となることを示した (それに対して, α が正か負かは

わからない)。

3つめの理由は，ボラティリティの長期記憶性である。Bandi and Perron(2006), Christensen and Nielsen(2006), Nielsen(2007), Nielsen and Frederiksen(2006) では，RV と BSIV が定常長期記憶過程に従っているとし，定常フラクショナル共和分モデルで分析している。(7) 式の $\{e_t\}$ の和分オーダーを $I(\delta_1)$, $\{x_t\},\{y_t\}$ の和分オーダーを $I(\delta_2)$ とする。$0 \leq \delta_1 < \delta_2 < 0.5$ が成立する場合に，$\{x_t\},\{y_t\}$ は定常フラクショナル共和分関係にあるという。定常フラクショナル共和分モデルの分析でよく利用されるのは，Robinson(1994) が提案した NBLS 推定 (narrow-band least squares estimation) である。原点付近のピリオドグラム m 個

$$\frac{1}{m} + \frac{m}{T} \to 0, \text{ as } T \to \infty \tag{8}$$

での回帰である。通常の OLS とは異なり，説明変数と誤差項に相関があっても $(\frac{m}{T})^{\delta_1-\delta_2}$ consistent である (Robinson and Marinucci(2003))。また，Nielsen(2007) では NBLS 推定を GLS タイプに改良した (δ_1 も同時に推定する)LW 推定が提案されている。

以上のことを踏まえて，本章では (7) 式を日経 225 株価指数の RV と MFIV を用いて，フラクショナル共和分モデルで分析した。日本のデータで分析を行ったのは本論文がはじめてである。また，RV と MFIV を共和分分析しているものはほとんどないが，Andersen et al. (2007) では S&P500 指数を分析している。先行研究と同様に，β は有意に 0 より大きく，1 より小さいという結果が得られた。共和分モデルは長期的な関係をみているので，IV は RV に対して情報を含んでいるが，不偏ではなく，長期的にもリスクプレミアムの影響が示唆される。

また，IV と RV をフラクショナル共和分で分析している先行研究では，共和分が存在するかを事前に確かめずに推定だけ行っているので，事前の分析を行ったところ，共和分関係が

確認された。

原点近傍のピリオドグラムだけを用いるセミパラメトリックなフラクショナル共和分モデルでは, 説明変数と誤差項に相関があっても一致性はあるが, その場合は漸近バイアスが発生する。Robinson (2007) では漸近バイアスが発生しない推定方法が提案されているが, この方法は相関がない場合には利用できない。そこで, 本章では, 説明変数と誤差項に相関がないかどうかの検定を提案した。