

## ハイブリッドモンテカルロ法による GARCH モデルのベイズ推定

高 石 哲 弥\*

### 1. はじめに

最近の研究より、株式収益率や為替レート収益率の分布は単純な正規分布ではなく、裾野の広がった分布（ファットテイル分布）をしていることが知られている。また、ボラティリティはボラティリティのショックが持続するボラティリティクラスターリングを示すことが明らかになっている。このようなボラティリティの特性を考慮したモデルとして ARCH モデルや GARCH モデルのような統計モデルが提案されている。

統計モデルを利用して実証分析を行う場合、与えられた経済時系列データのもとで統計モデルのパラメータを推定する必要がある。ARCH モデルや GARCH モデルに対するパラメータ推定では最尤法が利用されることが多い。最尤法では、パラメータの推定値は尤度関数を最大化するパラメータの値として得られる。

近年のコンピュータの発展により統計モデルの推定に対してマルコフ連鎖モンテカルロ法（以下では単にモンテカルロ法と呼ぶ）によるベイズ推定が試みられるようになってきている。特に、尤度関数が解析的に求められない場合にモンテカルロ法によるベイズ推定は有効な手法となる。また、尤度関数が計算できる ARCH モデルや GARCH モデル等に対してもモンテカルロ法が試みられてきている[1,2]。

これまでに行われている研究では、ギブス法やメトロポリス法を利用したモンテカルロ法が多く利用されている。本研究では統計モデルのベイズ推定ではまだ用いられたことのないハイブリッドモンテカルロ法を利用し、GARCH モデルに対して行ったベイズ推定の結果を報告する（注：人工的な株価経路の生成にハイブリッドモンテカルロを利用した研究は参考文献[11]を参照のこと。）尚、本研究の内容は The 9th Joint Conference on Information Sciences 2006[3]での発表を基にしている。

---

\* 広島経済大学経済学部助教授

## 2. ハイブリッドモンテカルロ法

ハイブリッドモンテカルロ法は1987年に Duane らによって格子 QCD (Quantum Chromo Dynamics) 計算を効率よく実行するために考案された[4]。格子 QCD 計算では多数の変数(数十万~数百万)からなる多重積分を実行する必要があるが、この多重積分をモンテカルロ法で見積もっている。普通、モンテカルロ法では順番に1つの変数を選び、変数値を更新(アップデート)してゆくが、格子 QCD 計算では1つずつアップデートしていくのは困難である。というのも、1つの変数をアップデートする度に巨大行列式(およそ $10^7 \times 10^7$ )を計算しなくてはならず、この方法ではすべての変数を1回アップデートするだけで膨大な計算量を必要としてしまう。Duane らの開発したハイブリッドモンテカルロ法は巧妙にすべての変数を一度のアップデートし、かつ近似のない正確な手法を提供する。(ここではすべての変数を一度にアップデートする方法をグローバルなアップデートと呼ぶことにする。)このハイブリッドモンテカルロ法の出現によって格子 QCD 計算は格段に進歩し、現在も格子 QCD 計算を実行する標準的な手法となっている。

$f(x_1, \dots, x_n)$  の確率分布に従う  $n$  個の確率変数  $x_i$  を考えよう。この確率変数のもとで、ある量  $O(x)$  ( $x_i$  の関数) の期待値は以下で与えられる。

$$\langle O \rangle = \int O(x_1, \dots, x_n) f(x_1, \dots, x_n) dx_1 \cdots dx_n. \quad (1)$$

この積分が実行できれば  $O(x)$  の期待値が求まるが、実際には解析的に求めることができない場合も多い。モンテカルロ計算ではこの積分を次のようにして求める。

まず、 $f(x_1, \dots, x_n)$  の確率分布に従う乱数  $x_i$  を  $k$  個生成する。そして、期待値を以下で近似する。

$$\langle O \rangle \approx \frac{1}{k} \sum_{j=1}^k O(x^j) \quad (2)$$

ここで、 $x^j = (x_1, \dots, x_n)^j$  は  $j$  個目に生成された乱数の組である。この値は式(1)に対する近似値ではあるが、 $k$  の値を大きくとることによって真の期待値に近くすることができる。

モンテカルロ計算ではいかに  $f(x_1, \dots, x_n)$  の確率分布に従う乱数  $x_i$  を効率よく生成されることができかが問題となる。 $f(x_1, \dots, x_n)$  が正規分布のようによく知られた分布であれば簡単であるが、一般には  $f(x_1, \dots, x_n)$  は複雑な関数形をとりその生成は難しくなる。

ある確率分布に従う乱数を発生させる方法としてメトロポリス法[5]が知られて

いる。この方法は比較的簡単に任意の確率分布に従う乱数を発生することができる。メトロポリス法では以下のステップに従って乱数を発生させる。

- (i)  $x_l$  に対して新しい候補  $x_l^{new}$  を選択する。
- (ii) 新しい候補  $x_l^{new}$  を以下の確率で採択する。

$$P = \min\left[\frac{f(x_1, \dots, x_l^{new}, \dots, x_n)}{f(x_1, \dots, x_l, \dots, x_n)}, 1\right]. \quad (3)$$

- (iii) 他の  $x_l$  に対しても上記のステップを繰り返す。
- (iv) (i)から(iii)を繰り返して乱数列を生成する。

ハイブリッドモンテカルロ法の特徴は複数の確率変数を一度にアップデートすることが可能なことである。上記のメトロポリス法も原理的には複数の確率変数を一度にアップデート可能であるが、普通、式(3)での確率が小さくなり実用上は一度にアップデートするのは困難である。

ハイブリッドモンテカルロ法では、ハミルトニアンが現れるが、これは以下のように定義される。まず、 $f(x_1, \dots, x_n)$  を式(4)のように書き直す。

$$f(x_1, \dots, x_n) = \exp(\ln f(x_1, \dots, x_n)). \quad (4)$$

これを利用して式(1)を変形すると、

$$\langle O \rangle = \int O(x_1, \dots, x_n) \exp(\ln f(x_1, \dots, x_n)) dx_1 \cdots dx_n \quad (5)$$

となる。ここで、 $x_i$  に共役な運動量  $p_i$  を導入して

$$\langle O \rangle = \int O(x_1, \dots, x_n) \exp\left(-\frac{1}{2} \sum_j p_j^2 + \ln f(x_1, \dots, x_n)\right) dx_1 \cdots dx_n dp_1 \cdots dp_n / Z, \quad (6)$$

とする。ここで、 $Z$  は規格化定数で

$$Z = \int \exp\left(-\frac{1}{2} \sum_j p_j^2 + \ln f(x_1, \dots, x_n)\right) dx_1 \cdots dx_n dp_1 \cdots dp_n \quad (7)$$

である。(注： $f(x_1, \dots, x_n)$  は確率分布として規格化されているとして扱ってきたが、一般に規格化されていなければ式(7)となる。)一見すると運動量  $p_i$  の導入は式(5)の値を変えてしまうように思うかもしれないが、 $O(x)$  は運動量  $p_i$  によらないので式(5)と式(6)の値は同じとなる。

ハミルトニアンは

$$H(x, p) = \frac{1}{2} \sum_j p_j^2 - \ln f(x) \quad (8)$$

で定義される。ハミルトニアンを利用すると式(6)は

$$\langle O \rangle = \int O(x) \exp(-H(x,p)) dx dp / Z \quad (9)$$

と表される。ここで、 $x=(x_1, \dots, x_n), p=(p_1, \dots, p_n)$  と省略して書いてある。以下でもこのような省略形を適宜利用する。

ハイブリッドモンテカルロ法による乱数列の生成は以下のステップによって実行される。

(i)新しい乱数の候補をハミルトンの運動方程式を解くことによって選択する。ハミルトンの運動方程式は

$$\frac{\partial H}{\partial x_i} = -\dot{p}_i \quad (10)$$

$$\frac{\partial H}{\partial p_i} = \dot{x}_i \quad (11)$$

で与えられる。この方程式は一般には解析的に解けないので leapfrog 法で数值的に解く。(注：数值的に解くだけならさまざまな方法が利用できるが、ハイブリッドモンテカルロ法では leapfrog 法のような時間反転性と体積保存を満たしている積分法しか利用できない。)

(ii)  $\Delta H = H(x_{new}) - H(x)$  とする。新しい候補  $x_{new}$  を次の確率で選択する。

$$P = \min[\exp(-\Delta H), 1]. \quad (12)$$

(iii)上記ステップを繰り返す。

ハイブリッドモンテカルロ法でのアイデアは上記のようにハミルトニアンを導入し、ハミルトンの運動方程式を解くことによって新しい乱数の候補を選択することである。ハミルトンの運動方程式はハミルトニアンを保存するので、正確に解けば  $\Delta H = 0$  である。実際には数值的に解くので、 $\Delta H$  はゼロではないが、leapfrog 法でのステップサイズを小さくとることによって  $\Delta H$  を小さな値とすることができる。そのためグローバルなアップデートにもかかわらず式(12)での採択率を高く取ることができる。

### 3. GARCH モデル

株価や為替レート収益率にはボラティリティクラスティングというボラティリティのショックが持続する現象が見られる。このボラティリティのショックの持続性を考慮したモデルとして Engel による ARCH モデルがある[6]。ARCH モデルのボラティリティ  $\sigma_t$  は

$$\sigma_t^2 = \omega + \sum_{i=1}^m \alpha_i y_{t-i}^2, \quad (13)$$

で与えられ、このときの時系列は  $y_t = \sigma_t \varepsilon_t$  で生成される。ここで、 $\varepsilon_t$  は平均ゼロ、標準偏差 1 の正規分布に従う確率変数である。また、 $\omega > 0$ 、 $\alpha_i \geq 0$  である。m は ARCH モデルの次数を表し、m 次のモデルは ARCH(m) モデルと言われる。

GARCH モデルは ARCH モデルを含む一般化されたモデルで Bollerslev によって提案された [7]。GARCH(m,n) モデルは

$$y_t = \sigma_t \varepsilon_t, \\ \sigma_t^2 = \omega + \sum_{i=1}^m \alpha_i y_{t-i}^2 + \sum_{j=1}^n \beta_j \sigma_{t-j}^2, \quad (14)$$

で表される。ここで、 $\omega > 0$ 、 $\alpha_i \geq 0$ 、 $\beta_i \geq 0$  である。

本研究では、GARCH(1,1) モデルを利用する。このとき、式(14)は

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (15)$$

となる。従って、本研究での推定すべきパラメータは  $\alpha$ 、 $\beta$ 、 $\omega$  となる。

#### 4. ベイズ推定

ベイズ推定では推定されるデータ  $y$  の下でのパラメータ  $\theta = (\alpha, \beta, \omega)$  の事後分布  $\pi(\theta|y)$  は以下で与えられる。

$$\pi(\theta|y) \propto f(y|\theta) \pi(\theta). \quad (16)$$

ここで、 $f(y|\theta)$  は次で与えられる尤度関数である。

$$f(y|\theta) = \prod_{t=1}^n \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{y_t^2}{\sigma_t^2}\right). \quad (17)$$

$\pi(\theta)$  は  $\theta$  の事前分布である。もし、 $\theta$  についての事前情報が与えられているならばそれを利用して  $\pi(\theta)$  が与えられるが、本研究では事前情報は与えられていないとして定数とする。

パラメータ  $\theta$  のベイズ推定値は以下のように確率分布  $f(y|\theta)$  のもとでの  $\theta$  の期待値  $\langle \theta \rangle$  として与えられる。

$$\langle \theta \rangle = \int \theta \pi(\theta|y) d\theta. \quad (18)$$

モンテカルロ計算では確率分布  $\pi(y|\theta)$  に従う  $\theta$  を生成して

$$\langle O \rangle \approx \frac{1}{k} \sum_{j=1}^k \theta_j, \quad (19)$$

として  $\theta$  を推定する。

ハイブリッドモンテカルロ法のハミルトニアンは  $\theta$  に共役な運動量  $P_\theta$  を導入して

$$H(\theta, p) = \sum_{\theta} p_{\theta}^2 / 2 - \ln f(y | \theta), \quad (20)$$

となる。この式(20)が式(8)に対応するものであり、このハミルトニアンを使ってハイブリッドモンテカルロ法を実行する。

#### 4. ハイブリッドモンテカルロシミュレーション

本研究では、GARCH(1,1)モデルのパラメータとして  $\alpha=0.3$ ,  $\beta=0.5$ ,  $\omega=0.1$  を選び、時系列を  $t=5000$  まで生成し、その時系列データに対してハイブリッドモンテカルロ法を利用したベイズ推定を実行した。そして、求められたパラメータ値が時系列データの生成に用いた  $\alpha=0.3$ ,  $\beta=0.5$ ,  $\omega=0.1$  と一致するかどうかを調べた。

図1は  $\alpha=0.3$ ,  $\beta=0.5$ ,  $\omega=0.1$  をパラメータとする GARCH(1,1) モデルで生成した時系列データである。

時系列データ数による推定の精度を見るために、3つの時系列データセットに対して推定を実行した。3つの時系列データセットは A:  $t=500$  までのデータ(データ

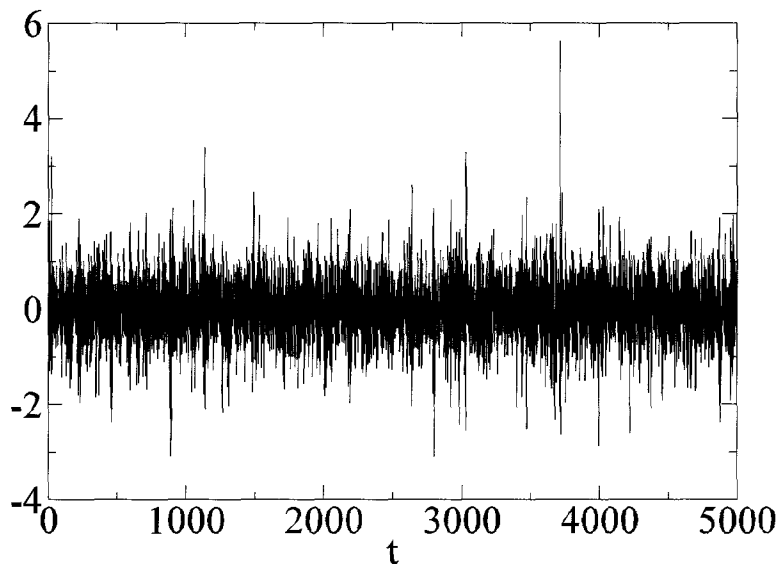


図1 本研究のベイズ推定に用いた GARCH(1,1) データ

数  $N=500$ ), B:  $t=1000$ までのデータ(データ数  $N=1000$ ), C:  $t=5000$ までのデータ(データ数  $N=5000$ )である。データ数が増えればそれだけ推定の精度が上がるのが期待される。

図2から4はハイブリッドモンテカルロ法によって生成したパラメータのモンテカルロ列である。モンテカルロ列のはじめの3000データは初期値依存性を取り除くために省いてある。推定に利用するデータ数が増加するにしたがってモンテカルロ列の揺らぎが小さくなっていくのが見て取れる。

図5は $\alpha$ のモンテカルロ列の自動相関関数である。この結果をみると、時系列データ数が多くなるとモンテカルロ列の相関が小さくなってることが分かる。

図6から8はモンテカルロ列のヒストグラムをプロットしたものである。データ数が少ないとヒストグラムは広がった形をしているが、データ数が増加するに従ってヒストグラムの幅は小さくなりより真の値を中心として広がった形となっている。

推定結果は表1にまとめてある。推定値の()の中の数値は統計誤差, その後の±はヒストグラムの標準偏差を表している。 $\Delta t$ と $\tau$ はそれぞれ leapfrog 法における積分のステップサイズ, 積分長を表している。採択率は式(12)における値である。

N=500

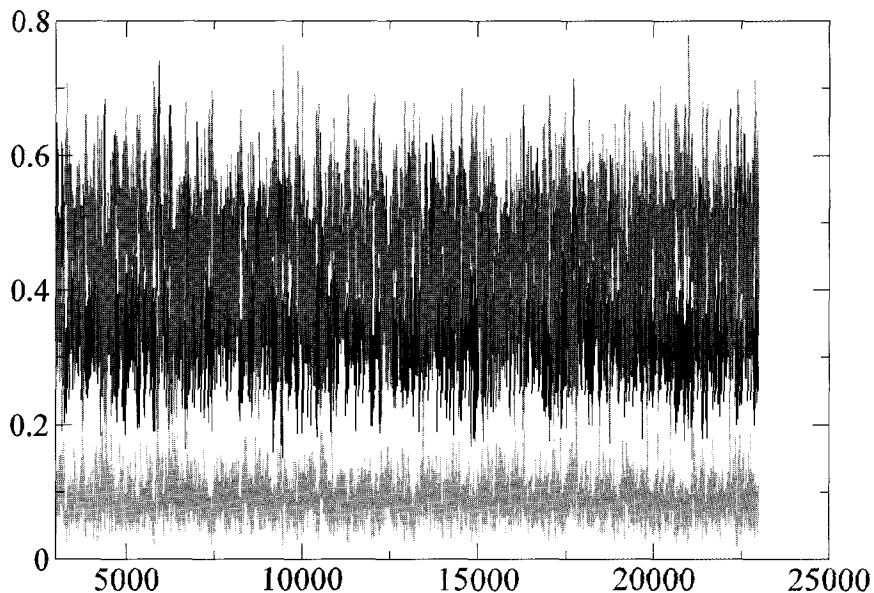


図2 データ数500に対してハイブリッドモンテカルロ法を実行した結果。  
上からそれぞれ  $\beta$ ,  $\alpha$ ,  $\omega$  のモンテカルロ列に対応する。

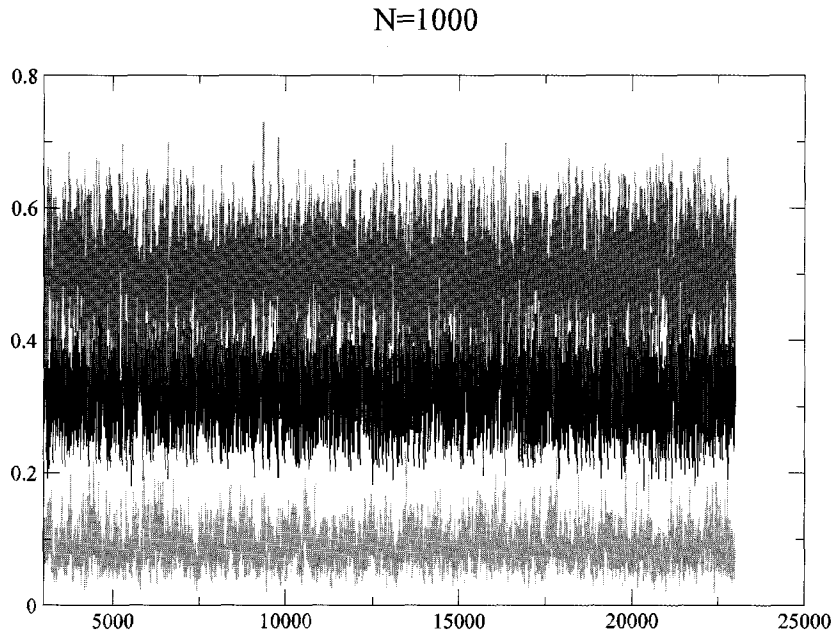


図3 データ数1000に対してハイブリッドモンテカルロ法を実行した結果。上からそれぞれ  $\beta$ ,  $\alpha$ ,  $\omega$  のモンテカルロ列に対応する。

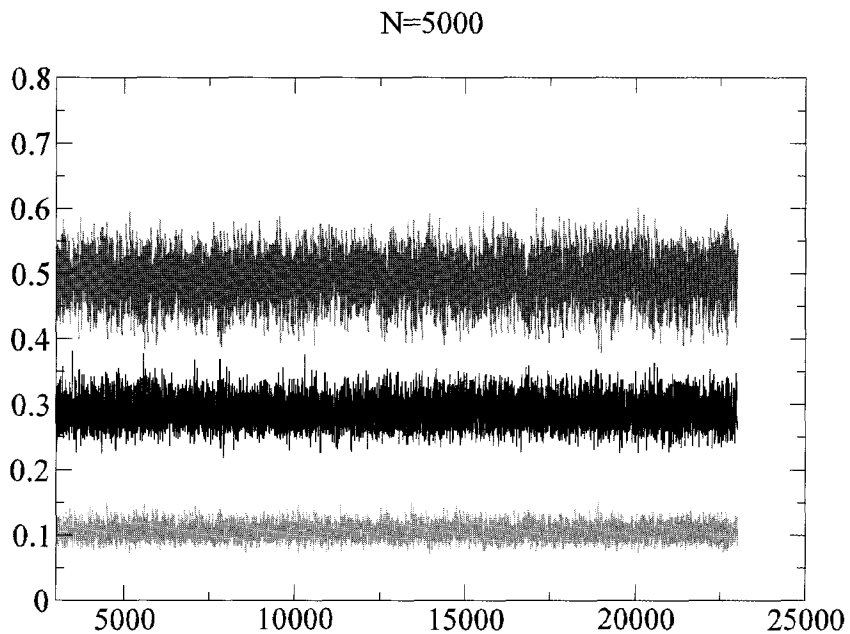


図4 データ数5000に対してハイブリッドモンテカルロ法を実行した結果。上からそれぞれ  $\beta$ ,  $\alpha$ ,  $\omega$  のモンテカルロ列に対応する。



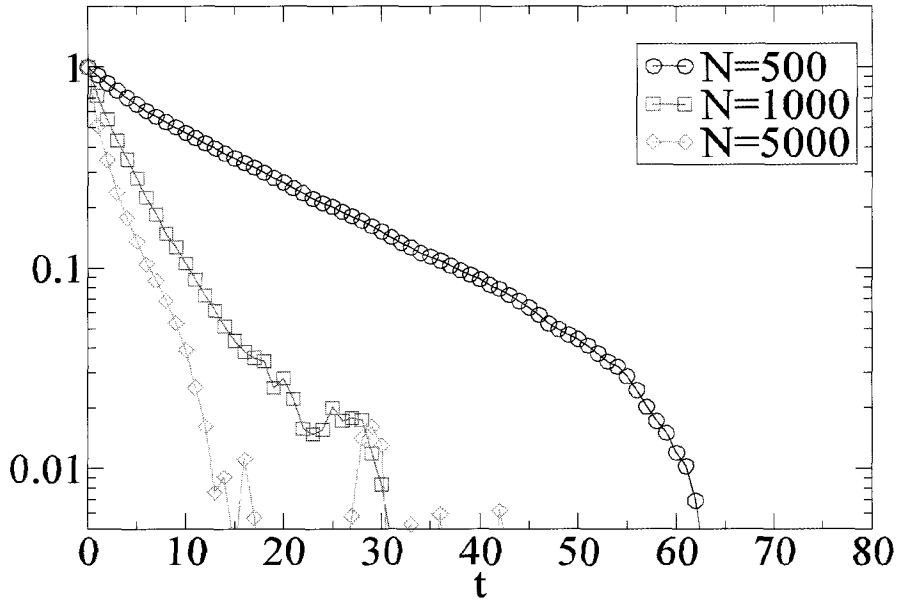


図5  $\alpha$  の自動相関関数

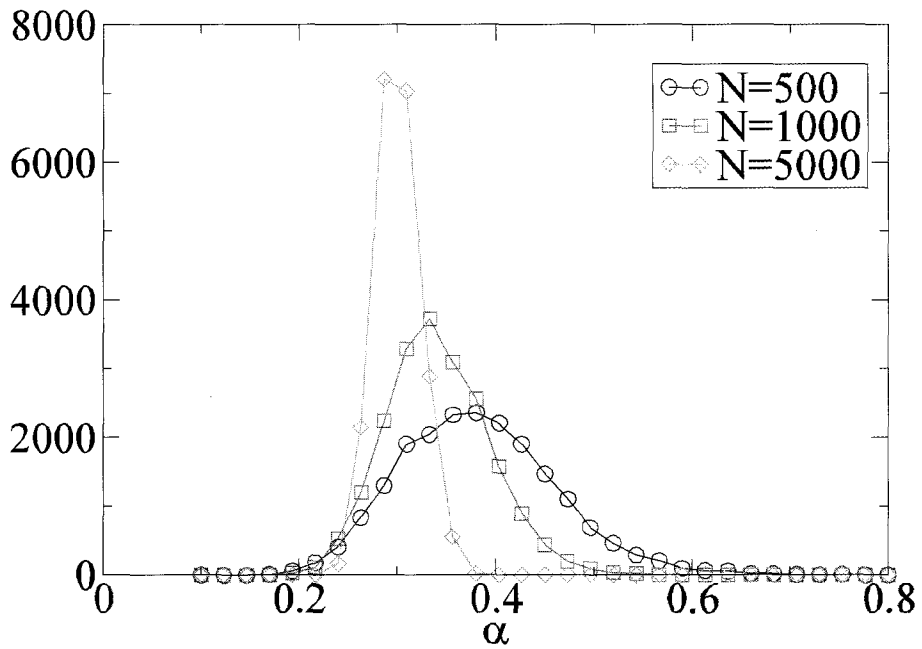


図6  $\alpha$  のヒストグラム

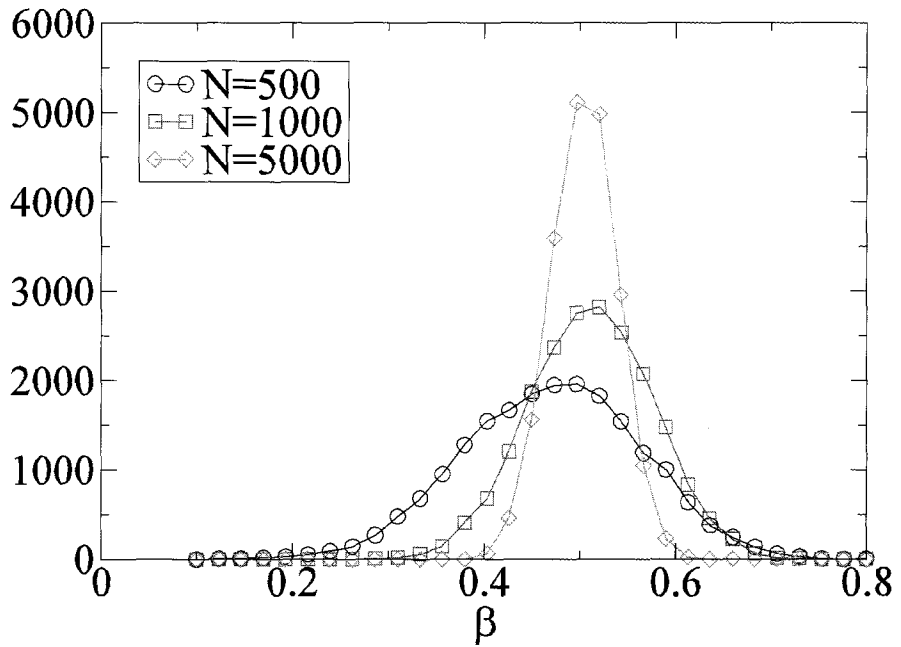
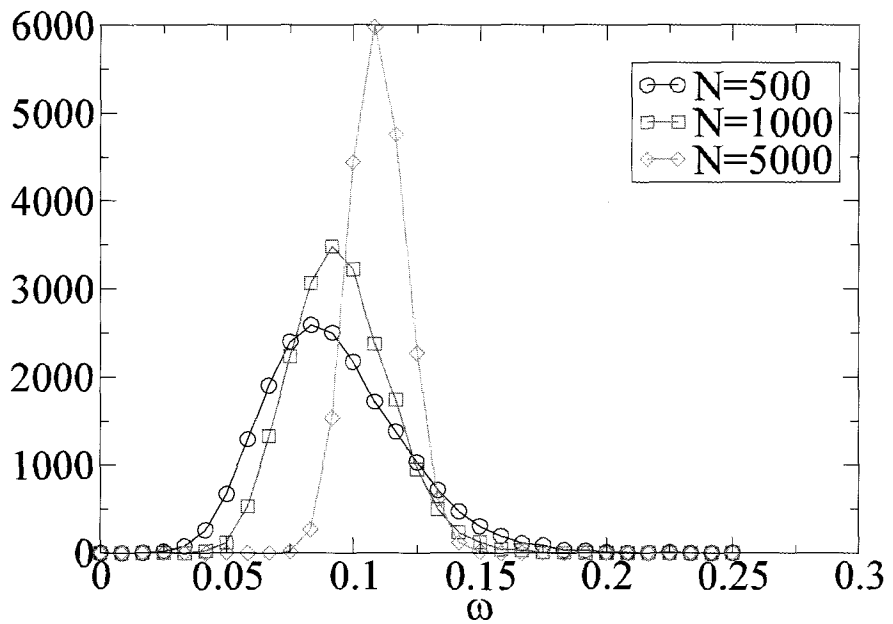
図7  $\beta$  のヒストグラム図8  $\omega$  のヒストグラム

表1 推定結果

	真の値	N=500	N=1000	N=5000
$\alpha$	0.3	0.370(4) $\pm$ 0.08	0.330(2) $\pm$ 0.05	0.2892(3) $\pm$ 0.02
$\beta$	0.5	0.460(4) $\pm$ 0.09	0.500(2) $\pm$ 0.06	0.4928(5) $\pm$ 0.034
$\omega$	0.1	0.0887(8) $\pm$ 0.03	0.0982(5) $\pm$ 0.02	0.1055(2) $\pm$ 0.01
$\Delta t$		0.0002	0.0002	0.0002
$\tau$		0.04	0.04	0.04
採択率		0.82	0.69	0.61

図9は $\alpha$ ,  $\beta$ ,  $\omega$ の推定値をデータ数の依存性がわかるようにプロットした図である。図中の破線は真値を表している。この図からわかるように、データ数が少なければ推定値の誤差が大きいですが、データ数が増加するにしたがって誤差は小さくなり真の値に近づいている。

## 5. まとめ

本研究ではハイブリッドモンテカルロ法を用いてGARCHモデルに対してベイズ推定を実行した。推定結果は真の値とよく一致しており、ハイブリッドモンテカ

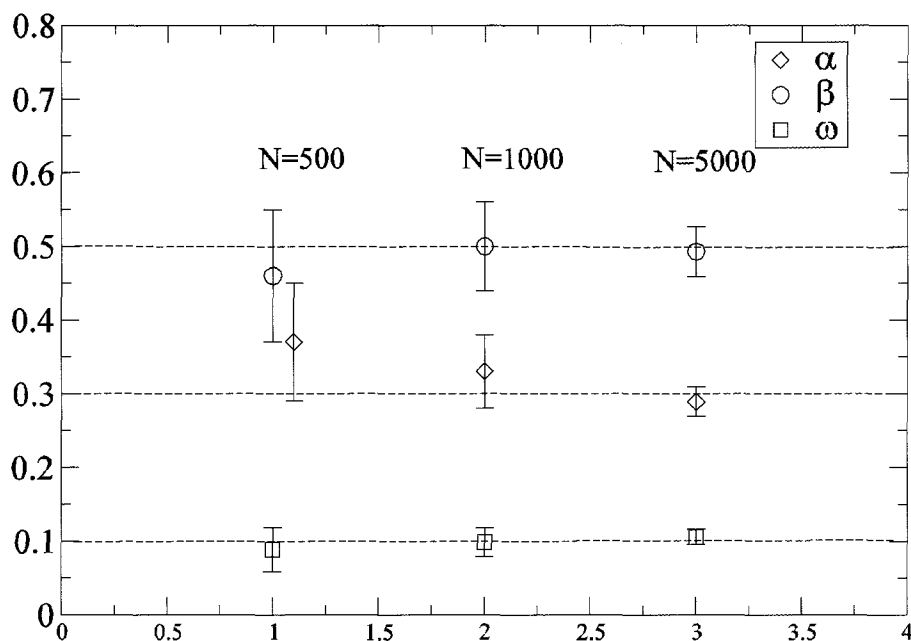


図9 データ数の違いによる推定値の変動

ルロ法が GARCH モデルのような統計モデルの推定に利用できることがわかった。

ハイブリッドモンテカルロ法はメトロポリス法と違い、推定に用いるパラメータのすべてを一度にアップデートすることができる。しかし、今回用いたモデルは3つのパラメータを持つ GARCH(1,1)モデルであり、アップデートを行う変数の数が多いとは言えないので、ハイブリッドモンテカルロ法の利点を生かしているとは言えない。今後はストカスティックボラティリティモデルのようにアップデートを行う変数の数が多いモデルに対してハイブリッドモンテカルロ法が有効であるかどうかを調べる必要があろう。

今回用いたハイブリッドモンテカルロ法は式(10)及び(11)を解くために単純な2次の leapfrog 法を用いたが、改善された積分法を利用することによってハイブリッドモンテカルロ法の有効性を上げられることが知られているので[8,9,10]、今後はそのような積分法も利用して有効性を上げる試みも必要であろう。

本研究での数値計算は統計数理研究所の Altix3700及び SX-6 スーパーコンピュータシステムを用いて実行された。

## 参 考 文 献

- [1] L.Bauwens and M.Lubrano, "Bayesian inference on GARCH models using the Gibbs sampler", *Econometrics J.* 1, (1998), C23-C46.
- [2] T.Nakatsuma, "Bayesian analysis of ARMA-GARCH models", *Journal of Econometrics*, 59 (2005) 57-69.
- [3] T.Takaishi, "Bayesian estimation of GARCH model by hybrid Monte Carlo", *Proceedings of the 9th Joint Conference on Information Sciences 2006*.  
<http://www.atlantis-press.com/publications/aisr/jcis-06/>
- [4] S. Duane, A.D. Kennedy, B.J. Pendleton, D. Roweth, "Hybrid Monte Carlo", *Physics Letters B*195, (1987) 216-222.
- [5] Metropolis et al., "Equations of State Calculations by Fast computing Machines", *Journal of Chemical Physics*, 21 (1953) 1087-1092.
- [6] R.Engel, "Autoregressive conditional heteroscedasticity with estimates of the variance of the UK inflation," *Econometrica* 50 (1982) 987-1008.
- [7] T.Bollerslev, "Generalized autoregressive conditional heteroscedasticity", *J.Econometrics* 51 (1986) 307-327.
- [8] T.Takaishi, "Choice of Integrator in the Hybrid Monte Carlo Algorithm", *Computer Physics Communications* 133 (2000) 6-17.
- [9] T.Takaishi, "Higher Order Hybrid Monte Carlo at Finite Temperature", *Phys. Lett. B*540 (2002) 159-165.

- [10] T.Takaishi and Ph.Forcrand, "Testing and tuning symplectic integrators for Hybrid Monte Carlo algorithm in lattice QCD", Physical Review E 73 (2006) 036706.
- [11] 高石 哲弥, 「インポートランスサンプリングによる株価経路の生成」広島経済大学研究論集 第28巻第4号 (2006年3月) 22-35