GARCH を使った金融時系列データのモデル化に関する研究

2008MI136 水野浩孝　　2008MI269 山元一真
指導教員　石崎文雄

1 はじめに

株式や為替などのリターンの標準偏差はボラティリティ (volatility) と呼ばれる．ボラティリティは投資リスクの表すひとつの指標であり，オプションの世界においては価格を決定する要因のひとつとなっている．ボラティリティの変化の特徴を理解し，予想する事は，リスク管理をする上で非常に重要となる．旧来の資産価格過程モデルにおいては，ボラティリティが一定であると仮定されているモデルが多い．例えば，ユーロピアン・オプションの理論価格の導出に標準的に使われている Black-Scholes モデル等も，資産価格過程を幾何ブラウン運動と仮定している．幾何ブラウン運動は，ボラティリティが一定の確率過程である．しかしながら，近年のファインスに関する実データの計量分析の多くの結果は，ボラティリティは時間とともに確率的に変動すると考える方が適切であることを示している [1]．株式や為替レートなどのリターンはボラティリティが過去のリターンに依存して変動し，ボラティリティが上昇（低下）するとしばらくボラティリティが高い（低い）期間が続くことが実証的に知られている．このようなボラティリティの時間的変動を取り込んだ資産価格過程モデルとして，ボラティリティの変動を明示的に定式化できるボラティリティ変動モデルに注目が集まっている．ボラティリティ変動モデルの代表的なものに ARCH (autoregressive conditional heteroskedasticity) モデルを一般化した GARCH (generalized ARCH) モデルがある．

本研究は，金融時系列データのボラティリティ変動を明示的にモデル化し，実データとの比較を行い GARCH モデルの有用性を調べる事を目的としている．本研究では，ダウ平均株価に着目し，ダウ平均株価のボラティリティ変動過程をボラティリティ変動モデルのひとつである GARCH によりモデル化することを考える．統計処理ソフトウェア R[6] を利用し，情報量基準として AIC[2] を考えることで，モデル化を行う．

2 ボラティリティ変動モデル

金融時系列においてしばしば観察される現象として時系列の変動が大きくなるとしばらく変動の大きい時期が持続し，変動が小さくなるとしばらく変動の小さい時期が持続する．これはボラティリティ・クラスタリングと言われ，これを表現するモデルとして ARCH モデルとそれを一般化した GARCH モデルがある．本稿では GARCH を用いてモデル化を行うが，まず ARCH モデル，GARCH モデルの定式化と特徴を要約する．

2.1 ARCH

ARCH(autoregressive conditional heteroskedasticity) モデル [4] は Engle によって提案されたモデルであり，ボラティリティ変動の持続性を考慮し，ボラティリティσ ^2_tを過去の収益率の予期できないショックの 2 乗 $\varepsilon^2_{t-1}, \varepsilon^2_{t-2}, ..., \varepsilon^2_{t-q}$ の線形関数として次のように安定化される．

$$\sigma^2_t = \omega + \sum_{j=1}^{q} \alpha_j \varepsilon^2_{t-j},$$

$$\omega > 0, \quad \alpha_j \geqq 0 \quad j = 1, 2, ..., q(\omega\text{は定数})$$

ω は定数項であり，価格変動のスケールによってとる値が変動する．そして α_j は残差項の係数であり，過去の残差との依存によりその値が変動する．

2.2 GARCH

GARCH (generalized ARCH) モデル [3] は Bollerslev によって提案され，ARCH モデルの過去の収益率の予期できないショックの 2 乗に，過去のボラティリティを加えたモデルであり次のように定式化される．

$$\sigma^2_t = \omega + \sum_{i=1}^{p} \beta_i \sigma^2_{t-i} + \sum_{j=1}^{q} \alpha_j \varepsilon^2_{t-j},$$

$$\omega > 0, \beta_i, \alpha_j \geqq 0, i = 1, 2, ..., p; j = 1, 2, ..., q(\omega\text{は定数})$$

この式の右辺第 3 項は過去の残差にも依存するという自己相関項を表しており，この項が ARCH モデルを拡張した部分である．定数項 ω，残差項の係数 α_j は ARCH モデルと同じであり，β_i は過去の残差に依存する自己相関項の係数である．

またこの式は，時点 t 以前の情報が与えられれば時点 t におけるボラティリティ σ^2_t が決定されることを示していることからこのモデルは条件付分散不均一モデルであり，ボラティリティ σ^2_t に関して自己回帰型モデルになっている．

2.3 情報量基準

実データをモデル化する時，統計的検定が容易に行えるような理想的な関数形になっていないことが多い．そのため予想の効率性など現実的な評価に基づいた方法が提案されることになった．真の確率分布にどの程度近いかを表す指標として，カルバック・ライブラーの情報量 (Kullback-Leibler divergence)[5] を少し書き直したもの

$$I(p, q) = \sum_{j=1}^{N} p_j \log p_i - \sum_{j=1}^{N} q_j \log q_i$$

を導入する。すべての j に対して $p_j = q_j$ ならば $I(p,q) = 0$ となりそれ以外の場合 $I(p,q) > 0$ である．このことは，$I(p,q)$ の最小値を与える q が真の確率分布であることを示す．また，この式は右辺第１項が定数となり，第 2 項が $logq_i$ の期待値を表し，平均対数尤度と呼ばれている．平均対数尤度が大きいほど真の分布に近いことになる．明らかにモデル係数が多いほど，真の分布にいくらでも接近する．しかし適切なモデル化を行うためには最大対数尤度とパラメータ数とのバランスをとる必要があり，このために導入された量が情報量規準である．

情報量規準の代表が AIC [2] で，最大対数尤度とパラメータ数を用いて，-2 ×最大対数尤度 +2 ×パラメータ数，と定義される．「最大対数尤度」は当てはめの良さを表すが，「パラメータ数」は逆の傾向を示す．そのため，AIC を最小化することで，両者のバランスをとることができる．

3　シミュレーションの準備

3.1　アドオンパッケージのインストール

R では，アドオンパッケージを CRAN と称されるパッケージをインストールし，自分の欲しい機能を拡張することができる．そのインストール方法は実に簡単で，R のコンソール上でコードを入力することでインストールを行うことが出来る．

今回のシミュレーションでは，garch 関数と garch.sim 関数を使うため，次の 2 つのパッケージ・tseries ・TSA を予めインストールをしておく．

3.2　データの読み込み

garch() 関数では，当てはめに予め差を取ったデータを使うため，ダウ平均株価の階差を取っておく必要がある．今回は，Excel を用いて階差を算出し，2007 年 9 月〜2009 年 9 月までの株価の階差を 1 行 500 列の CSV ファイル (dow.csv) として出力したものを用意した．

先ず，dow.csv のデータを R で引用できるようにするために，csv ファイルの読み込みを行っておく．

4　シミュレーション

本節では，R によるモデル推定のシミュレーション実験について説明する．シミュレーション実験においては，tseries パッケージの garch() 関数で金融時系列データの GARCH への当てはめを行い，その後 garch() 関数により得られたパラメータを TSA パッケージ garch.sim() 関数に当てはめることによりシミュレーションを行い，それを折れ線グラフにプロットする．

先ず，先ほど読み込んだ DOW の残差データを tseries の garch() 関数に当てはめる．

シミュレーションの流れを図で大まかに説明すると，図 1 のような流れでシミュレーションを行うことになる．garch() 関数は，時系列データを GARCH モデルに対してあてはめる関数である．

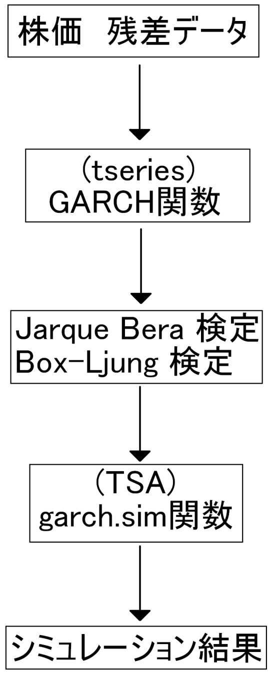

図 1　シミュレーションの大まかな流れ

当てはめを行った後に，Jarque Bera 検定，Box-ljung 検定を行い，適切なボラティリティとレートの値を設定する．

summary() 関数を用いることにより，この両方の検定の結果を次のように出力することができる．

検定結果

```
Diagnostic Tests:
Jarque Bera Test
data: Residuals
X-squared = 7.5265, df = 2, p-value = 0.02321
Box-Ljung test
data: Squared.Residuals
X-squared = 7.4542, df = 1, p-value = 0.006329
```

Jarque Bera 検定では，時系列データの正規性を検定することができ，Box-Ljung 検定では時系列データの独立性を検定することができる．

ボラティリティ 0〜4，レート 1〜4 の全ての組み合わせを実行し，その中からそれぞれの検定で有意水準 5% を超える組み合わせ，つまり p-value が 0.05 を超える組み合わせ選出する．

有意水準が 5% を超えた組み合わせの中から，AIC が最も低くなる組み合わせを選択する．AIC の出力については AIC() 関数を用いて出力を行う．

以上の操作を行ったところ，有意水準を満たし AIC

が最も低くなった組み合わせ (1,4) を使い，シミュレーションを行う．

シミュレーションには，TSA パッケージの garch.sim() 関数を用いる．この関数は，tseries の garch() 関数により得られたパラメータを使うことで GARCH によるシミュレーションを行うことが可能な関数である．

先ず, ボラティリティ 1, レート 4 の時のパラメータを取得するために再び summary 関数実行する．すると，次ぎのような出力結果を得ることができる．

出力結果

Coefficient(s):
Estimate Std. Error t value Pr(¿—t—)
a0 1.849e-05 6.897e-06 2.681 0.00735 **
a1 1.171e-08 4.082e-02 0.000 1.00000
a2 4.585e-03 4.150e-02 0.110 0.91202
a3 9.905e-02 5.721e-02 1.731 0.08340 .
a4 2.135e-01 7.757e-02 2.752 0.00593 **
b1 6.446e-01 6.882e-02 9.366 ¡ 2e-16 ***
先ほどの結果より，
a0=1.849e-05
a1=1.171e-08
a2=4.585e-03
a3=9.905e-02
a4=2.135e-01
b1=6.446e-01

がパラメータとなる．次に garch.sim 関数を使い a 部分を alpha，b 部分を beta として

シミュレーション

```
>sim=garch.sim(alpha=c(a0,a1,a2,a3,a4)
,beta=b1,n=506)
```

と入力することにより GARCH によるシミュレーションを行うことができる．

元の時系列データ（DOW）と，このシミュレーションの結果の例を線グラフにプロットしたものを図 2，図 3 に示した．

このシミュレートでは 2 年間全体のデータからボラティリティを出しているので、ボラティリティの推移をみることができない．似通っているかどうかを判断するためにはボラティリティの推移を見なくてはならない．

そこで，次の章ではダウ平均株価の残差データを一定の区間ごとに区切ったデータを用いてボラティリティを求めていく．

5　GARCH の有用性の確認

GARCH モデルの有用性を確認するには，一定の区間ごとに区切ったデータでシミュレーションを行い，それぞれのボラティリティを求め，ボラティリティの推移が

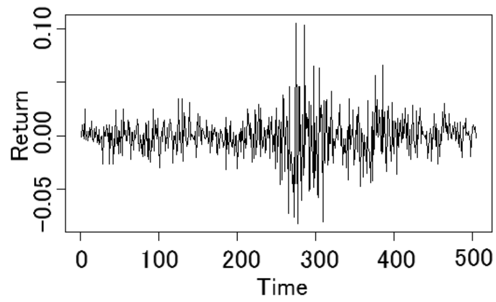

図 2　2007 年 9 月〜2009 年 9 月のダウ平均株価の残差

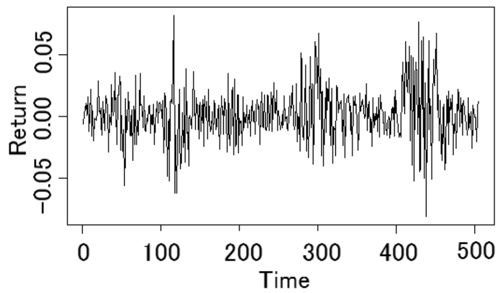

図 3　シミュレーション結果の残差

元データとシミュレーション結果とで似通っていることを確認する必要がある．

5.1　ボラティリティの推移の確認

ダウ平均株価のデータを 25 分割し，20 個ずつのデータでシミュレーションを行いボラティリティの推移を確認し，ダウ平均株価の残差と、シミュレーション結果の残差を折れ線グラフにプロットしたものを図 4.，図 5 に掲載する．

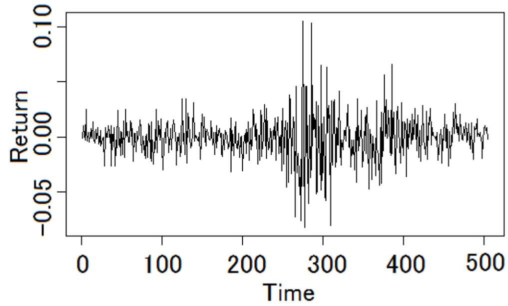

図 4　ダウ平均株価の残差

2 つのグラフを見比べると，小さな変動が起こる箇所と大きな変動が起こる箇所が似通っていることが確認できる．

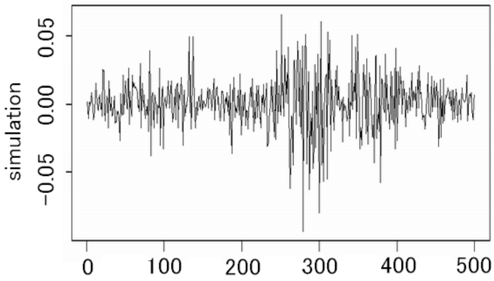

図 5　シミュレーション結果の残差

5.2　検定値と AIC の比較

GARCH に当てはめを行った際の Jarque Bera, Box-Ljung の値と AIC の一覧とボラティリティの推移を表 1 に示した. *1.

	ダウ平均株価	シミュレーション結果
0-20	(0,4)	(0,4)
20-40	(0,3)	(0,1)
40-60	(0,1)	(0,1)
60-80	(0,1)	(0,1)
80-100	(0,2)	(0,1)
100-120	(0,1)	(0,1)
120-140	(0,1)	(0,1)
140-160	(0,1)	(0,1)
160-180	(0,1)	(0,1)
180-200	(0,1)	(0,1)
200-220	(0,1)	(0,1)
220-240	(0,1)	(0,1)
240-260	(0,1)	(0,1)
260-280	(0,1)	(0,1)
280-300	(0,1)	(0,1)
300-320	(0,1)	(0,1)
320-340	(0,1)	(0,1)
340-360	(0,1)	(0,1)
360-380	(0,1)	(0,1)
380-400	(0,1)	(0,1)
400-420	(0,1)	(0,1)
420-440	(0,1)	(0,1)
440-460	(0,1)	(0,1)
460-480	(0,2)	(0,1)
480-500	(0,1)	(0,1)

表 1　ダウ平均株価とシミュレーション結果のボラティリティ推移表

表 1 の, ダウ平均株価とシミュレーション結果それぞれのボラティリティ/レート値の推移を比較してみると, 0.1 である期間が多いが, 0-20 の区間でボラティリティ/レート値が大きくなっており, それ以外では小さい値を維持しているのを見ると, ボラティリティ/レート値の推移はおおよそ似通っていることが確認できる.

*1 V/R 値とはボラティリティ/レート値のことである.

6　おわりに

本研究では統計ソフト R で GARCH を用いて DOW 平均株価のボラティリティをモデル化し, 実際のデータとシミュレーション結果を比較した. ボラティリティには上昇 (低下) するとしばらくボラティリティが高い (低い) 期間が続く特性があり, シミュレーション結果にもその特性が表れていることを確認した. またシミュレーションで得られたデータと実際のデータとのボラティリティの推移が似ているので, GARCH は有用なモデルであるということがわかった. しかし, GARCH はシミュレーション結果が対称形になりやすいという欠点があるが, ボラティリティの非対称性を考慮したモデルである EGARCH, GJR などを用いることで, より現実を反映したシミュレーションを行えると考えられる.

参考文献

[1] 渡部敏明: “ボラティリティ変動モデル”, 朝倉書店 2004.

[2] 松葉育雄: “期記憶過程の統計－自己相関な時系列の理論と方法”, 共立出版 2004.

[3] T.Bollerslev: “Generalized Autoregressive Conditional Heteroskedasticity”, Journal of Econometrics, 31, pp.307-327, 1986.

[4] R.F.Engle,: “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation,”, Econometrica, 50, pp.987-1006, 1982.

[5] T.M.Cover and J.A.Thomas: “Elements of Information Theory, 2nd edition,”, Wiley-Interscience, 2006.

[6] The R Project for Statistical Computing: http://www.r-project.org/