Rによる時系列解析
ーARMAモデルとGARCHモデルー

2003MM104 高松 ななみ

指導教員: 田中 豊

1 はじめに

現在、様々な分野において、時系列データは扱われている。時系列解析とは、時間と共にに不規則に変動しているように見える時系列データが、時間とどのような規則性を持って変化するかを分析し、また将来どのようになるかを予測することである。この研究は、分散の均一性を仮定した自己回帰移動(ARMA)モデル及び分散の不均一性を組み込んだGARCHモデルの２つのモデル族を中心とした時系列解析の方法論について学び、分散が均一に考えられる太陽黒点数データと分散が不均一と考えられるダウ平均株価データの分析を行うことを目的としている。

2 データについて

Ｒに組み込まれている、スイスの連邦観測所で集められた、1700年から1988年までの１年間に観測された太陽黒点相対数（sunspot.year）の計289個のデータと、ダウ・ジョーンズ社が算出する1997年7月1日から1999年4月9日までのダウ平均株価指数の日次の収益率の計464個(参考文献[2]より)の２つのデータを使用した。

3 モデル

3.1 自己回帰移動平均モデル(ARMAモデル)

ARMA(p,q)過程の方程式は、ある時点のデータはそれ以前のデータで推定できる、AR(p)モデルと、過去の誤差に影響されるモデル、MA(q)モデルを組みあわせたものである。

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=1}^{q} \theta_i Z_{t-i} + Z_t, \{Z_t\} \sim WN(0, \sigma^2)$$

残差Z_tは、期待値ゼロ、分散一定のホワイトノイズである。係数ϕ_iをゼロとおいた時は、移動平均モデルMAモデルになり、また係数θ_iがゼロの時には自己回帰モデルARになることから自己回帰移動平均モデルARMAは自己回帰モデルや移動平均モデルより一般的である。また、トレンドを取り除くために、差分をとる機能を加えてd階の差分を含むARIMA(p,d,q)に一般化される。

3.2 GARCHモデル

ARMAモデルの残差はホワイトノイズという前提であったが、為替レートや株価などの金融資産の時系列データなどは，残差はホワイトノイズの性質を満たさず、残差の分散は一様ではなく分散が変動するという特徴がある。分散の不均一性を組み込み、また今期の分散が過去の分散に依存することをモデルに組み込んだ時系列モデルがARCHモデルである。条件付き分散が過去の分散だけでなく、過去の残差の２乗にも依存するというARCHモデルを拡張したのがGARCHモデルである。GARCH(p,q)過程を

$$Z_t = \sqrt{h_t} e_t \quad , \{e_t\} \sim IID(0, 1)$$

と表わされる。ここで、h_t は、

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Z_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}^2$$

で定義される。
（ただし、$\alpha_0 > 0, \alpha_i, \beta_j \geq 0, i = 1, ..., p, j = 1, ..., q$)

4 時系列データの解析

4.1 太陽黒点

プロット図で、著しい上下非対称性がみられるので、対数変換と差分２の変換を行う。

4.1.1 標本自己相関と標本偏自己相関

データの相関やモデルの次数を決定する手がかりとして、標本自己相関関数と標本偏自己相関関数を用いる。

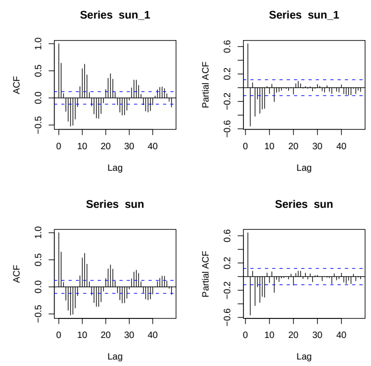

図 1: 標本自己相関関数と標本偏自己相関関数プロット図

図1の左上の図は、著しい上下非対称性がみられる。左下の図は、データを差分変換した標本自己相関関数である。データに見られる１０年程度の繰り返しに対応して、標本自己相関関数にも１０年程度周期的な変動がみられるが、その振幅は、ラグの増加とともに次第に減衰している。また、ラグが大きくなるにつれてACFの値が急速に減衰するので、時系列が、短期的従属であることを示している。ARMAの当てはめにはよい兆候である。**図1**の右下の図では、8(もしくは、12)より後のラグで偏自己相関関数の値が全て信頼区間$\pm 1.96/\sqrt{269}$ にあることから、ＡＲ(8)、AR(12)モデルが当てはまりが良いことを示唆している。

4.1.2　定常時系列時系列への当てはめ

AIC=479.6162で、最小となり、ARMA(4,4)モデルが選択された。10個のデータを予測し、その結果と実数値を以下に示す。実線が観測値で、点線が予測値である。

多少のずれは生じているが、ARMA(4,4)モデルでは、ほぼ同じ振る舞いをしていることより、ARMA(4,4)モデルの当てはまりがいいといえる。

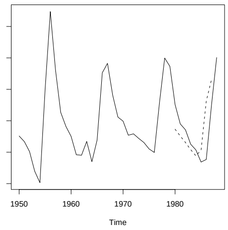

図 2: ARMA(4,4)に当てはめた予測図

4.2　ダウ平均株価

4.2.1　プロット図

まず視覚的にデータをとらえるために、プロット図を示す。

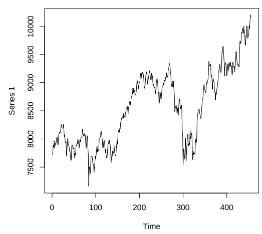

図 3: ダウ平均株価のプロット図

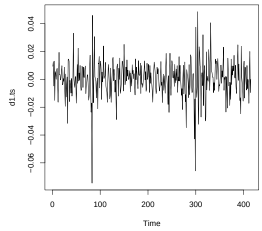

図 4: 株価収益率のプロット図

図4を見ると、n=100付近とn=300付近を境にして、変動の分散が大きく変化していることが分かる。よって、高いボラティリティが続く期間と低いボラティリティが続く時期の双方があることを示している。また、**図3**株価のデータのプロットと**図4**収益率のデータのプロットの図を比較すると、収益率に変換することにより定常的な時系列データになったと思われるのでこの株価収益率に対して、ARMAモデルなどの当てはめを行っていく。

4.2.2　ARMAモデルへの当てはめ

ARモデルでは、0次のAIC値が-2446.6で、次数0が選択され、MAモデルでは、1次のAIC値が-2444.665で、次数1が選択、また、ARMAモデルでは、AIC値は-2445.167でARMA(2,5)が選択された。以下の図から、どのモデルも当てはまりが悪いことが見て取れる。推定値の動きをみると、**図5**では、観測値がほとんど一定であり、**図6**では、周期性は表現できたが、平均が一定である。よって、以上のモデルでは株価収益率の変動を表現しきれなかったのではないかと考えられる。株価・株価収益率の分散が一定でないことが原因であると示唆される。実線が予測値で、点線が観測値である。

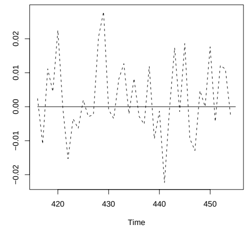

図 5: AR(0),MA(1)の予測図

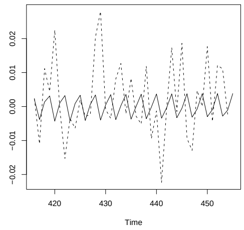

図 6: ARMA(2,5)の予測図

4.2.3　GARCHモデル

4.2.2の結果より、株価や株価収益率の分散が一定でない可能性が示唆された。そこで、GARCH関数を用いて、株価収益率の分析をした。残差$\tilde{Z}_t = Y_t - \hat{a}$に対する最小のAICCのGARCHモデルはGARCH(1,1)であり係数の推定値は、

$$\hat{\alpha}_0 = 0.00001211, \hat{\alpha}_1 = 0.1353, \hat{\beta}_1 = 0.7939$$

で、AIC値は3260.086であった。

$$h_t = 0.00001211 + 0.1353 Z_{t-1} + 0.7939 h_{t-1}$$

となる。適合度を測るために、下図にGARCH(1,1)モデルの予測図を**図7**に示す。

GARCHモデルに反映したい、データの大きな変動もしくは、小さな変動のあとに似たような大きさの変動が続くという傾向を示すことができた。また、**図7**の変動は**図4**に対応して、変動するボラティリティを明確に反映している。従って、最適なモデルに当てはめられたと考えられる。

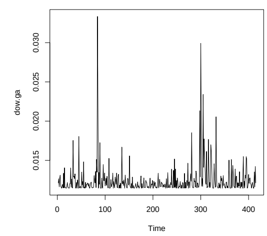

図 7: GARCH(1,1)モデルに当てはめた予測図

5　終わりに

分散が一様のものと分散にばらつきがあるものには違いがあった。太陽黒点ではもう少し違った分析も試みれたらもっと良好な結果が得られたかもしれない。また時系列解析において、将来を予測する為に必要な最適なモデルの作成が、いかに難しく、重要であるか痛感されられたが、興味深いものでもあった。

参考文献

[1] 村田友恵：Rにおける時系列解析,南山大学数理情報学部数理科学科卒業論文,2005.

[2] ダウ・ジョーンズ・ニューズワイヤーズ `http://www1.dowjones.jp/Scripts/web/word/get_djia.shtml`.