

# 経済データの時系列分析

2002MM029 猪子 恵司

指導教員 木村 美善

## 1 はじめに

現在, 自己の責任において資産管理運用を行う必要が生まれつつある. そのためインターネットでの環境が整備されて以来, ネット株が急速に広まっている. しかしながら, 金融商品はリスクを伴うものばかりであり, 未来を予測する力が必要となってくる. 本研究では株式投資のための種々のデータを用い, よりの確な未来予測を行うためのモデルの作成と検証を行っていくことにより時系列分析を学んでいくことを目的とする.

## 2 データについて

日経平均プロフィール (<http://www.nikkei.co.jp/nkave/>) からトヨタ自動車の純資産倍率, 株価収益率など計 16 種類の日次データを収集し, 分析に悪影響を及ぼすと考えられるもの (前日比など) を除いたデータを分析に用いた. ここで  $x_1 \sim x_6$  はそれぞれ売買高, 純資産倍率, 株価収益率, 株式益回り, 配当利回り, 株価に相当する.

## 3 重回帰分析

株価と種々のデータを用いて, 株価に対してどのデータが, どの程度深く関係し, どのような影響を与えているかを探るため重回帰分析を行った.

### 3.1 重回帰モデル

複数の説明変数が被説明変数に影響すると考えられる場合が数多く存在する. 二つ以上の説明変数がある場合を重回帰分析と呼ぶ. 重回帰分析を行うとき重要になってくるのが, 変数選択である. 今回は与えられた全ての説明変数の組み合わせに対して回帰モデルを作成し, そのモデル群の中から決定係数, AIC (赤池情報量規準) などの統計量に基づき, モデルを選択する総当たり法を用いて変数選択をした.

自由度調整済み決定係数と AIC の二つの統計量を今回の重回帰モデル選択に利用した. その結果, 自由度調整済み決定係数  $Adj.R^2 = 0.9846$ ,  $AIC = 1138$  で  $x_6 = \beta_1 x_1 + \beta_2 x_3 + \beta_3 x_4 + \beta_4 x_5$  というモデルが選択された ([5],[7] 参照).

### 3.2 ダービン・ワトソン検定

時系列データを用いているので, 系列相関の問題が出てくる可能性がある. そこでダービン・ワトソン検定を行った結果, 検定統計量  $d = 2.047$  となった, 有意水準  $\alpha$  を 5% とするとデータ数は 100 で, 説明変数の数は  $p = 4$  であるので,  $d_L = 1.59$ ,  $d_U = 1.76$  となり,  $d_U \leq d \leq 4 - d_U$  となるので  $H_0$  は有意水準 5% で採択され, 相関関係はないと考えてよい ([7] 参照).

## 3.3 分析結果

	係数	標準誤差	t 値	p 値
定数項	1900	1215	1.542	0.13
$X_1$	0.079	0.03	2.904	0.004
$X_3$	260	39.12	6.576	$2.6 \times 10^{-9}$
$X_4$	680	486.3	1.398	0.17
$X_5$	-4100	2363	-1.721	0.09
$R^2 = 0.9852$		$Adj.R^2 = 0.9846$		$n = 100$

## 3.4 考察

$x_1, x_3, x_4$  の回帰係数はそれぞれプラスの値を取り, 売買高・株価収益率・株式益回りは株価に対してプラスの要因といえる.  $x_5$  はその逆で, 回帰係数はマイナスの値を取るのので, 配当利回りは株価に対してマイナスの要因といえる. また, 回帰診断の結果を見ると, このモデルに問題ないことがわかる.

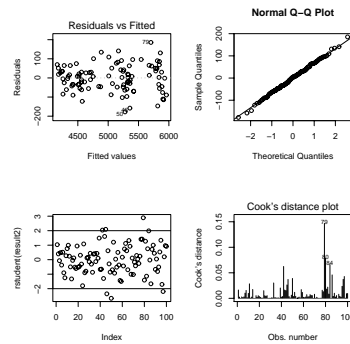


図 1 回帰診断

## 4 時系列分析

株価の動向を予測するための手がかりを得るために時系列分析を行った.

### 4.1 定常時系列

時系列データが定常過程に従うと仮定することにより, そのデータの持つ変動特性について有意義な情報を見出すための手がかりが得られる. そこでトヨタ自動車の株価が定常過程に従う変数かどうかを調べてみた.

トヨタ自動車の株価  $S_t$  について, 1 期間ラグと 10 期間ラグでの相関係数を調べた結果, それぞれ 0.991, 0.926 となり, 株価は比較的強い正の相関関係があることがわかった. 基本的に時系列データが定常的であるならば, その自己相関係数は, 期間ラグが大きくなると急速に 0 に近づく. したがって, トヨタ自動車の株価は定常的でない判断される ([6],[7],[8] 参照).

## 4.2 AR モデル

近似的に定常過程に従うとみなせるようなデータに変換でき、定常時系列モデルを仮定して未来の株価変動を予測することが可能になる場合があるかもしれない。

株価のようにトレンドをもつ時系列データを、近似的に定常時系列データに変換して分析するために、株価そのものではなく収益率を用いることが行われている。

ここで収益率とは

$$x_t = \log S_t - \log S_{t-1}$$

で表されるもののことを言う。

この株価収益率に対して AR モデル (1 変量自己回帰モデル) を適用して、分析を行った ([2],[7] 参照)。

## 4.3 ARIMA 分析

階差をとった系列に ARMA モデルをあてはめるとき、それを ARIMA モデル (自己回帰和分移動平均モデル) と呼ぶ。AR モデル同様に株価収益率を ARIMA モデルで分析をした。モデルの選択には  $p, d, q$  をそれぞれ、 $1 \leq p \leq 10, 1 \leq d \leq 5, 1 \leq q \leq 5$  の範囲で動かし、AIC が最小となるものを選んだ。その結果、ARIMA(3, 1, 4) モデルが AIC = -486.2805 で選択された ([1] 参照)。

## 4.4 残差解析

AR, ARIMA 両モデルについてリュング・ボックス検定を行ったところ、それぞれ  $K = 10$  のとき  $\hat{Q} = 12.1696$  と  $\tilde{Q} = 7.3249$  となり、 $\chi_{9,0.05}^2 = 16.92 > \hat{Q}$  と  $\chi_{3,0.05}^2 = 7.815 > \tilde{Q}$  となった。このことから、AR, ARIMA 両モデルの残差に自己相関はないといえる。次に残差の正規性をチェックをするために、QQ プロットを行った。その結果、どちらのモデルも直線上に並んでおらず、残差の正規性は疑わしいことがわかった。次にプロット図を示しておく ([6],[8] 参照)。

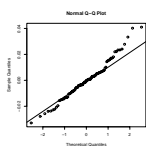


図2 AR QQ

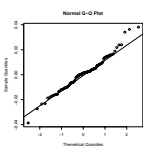


図3 ARIMA QQ

## 4.5 考察

QQ プロットの結果から、どちらのモデルも不適切なモデルであることがわかった。分析結果はそれを如実に表しているといえる。さらに AR モデルでは、推定値と実測値を比べてみた。その結果 AR モデルは当てはまりが悪いことが見て取れた。これは株価・株価収益率の分散が一定でないことが考えられる。このことから AR 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。

## 4.6 GARCH 分析

AR モデルより、株価・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 モデルによる分析

を考慮することにした。GARCH モデルの基本的特徴は、条件付分散が時間とともに変化する点にある。条件付分散が変化することは、非線形モデルとなることを意味する。

モデルの選択は  $(p, q)$  を  $0 \leq p \leq 3, 1 \leq q \leq 3$  の範囲で動かし、この中で最も AIC の値が小さいものを選んだ。その結果 GARCH(1, 1) モデルが AIC = -325.8043 で選択された。

トヨタ自動車の株価収益率データに GARCH(1, 1) モデルをあてはめると、

$$\begin{cases} \hat{x}_t = -0.003623919 + \hat{\epsilon}_t \\ \sigma_{t-1}^2 = 0.00006196 + 0.1866\epsilon_{t-1}^2 + 0.5487\sigma_{t-2}^2 \end{cases}$$

となる。推定結果は  $\epsilon_{t-1}^2 = (x_{t-1} - \mu)^2$  より  $\sigma_{t-2}^2$  の方が大きい効果を持つことを示している。すなわち、条件付き分散は、前期の 2 乗収益率だけでなく過去の多くの 2 乗収益率に影響されていると思われる ([2],[6],[10] 参照)。

## 5 おわりに

本研究をまとめると、株価と種々のデータの関係をつかむために重回帰分析を行った。その結果、 $x_6 = x_1 + x_3 + x_4 + x_5$  というモデルが最良だと判断され、売買高、株価収益率、株式益回りがそれぞれ株価に対してプラスの要因であり、配当利回りは株価に対してマイナスの要因だということがわかった。

次に株価の動向を予測するための手がかりを得るために時系列分析を行った。AR モデル、ARIMA モデルでは有益な情報が引き出せなかったが、GARCH モデルでは条件付き分散は、前期の 2 乗収益率だけでなく過去の多くの 2 乗収益率に影響されている、という結果を得ることができた。

本研究では、株価変動の中の本質的な性質を理解するにはいたらなかった。それは株価変動は数値データ以外の様々な情報や、投資家の心理的要因により大きく変化するという性質を持っているためだと考えられる。

## 参考文献

- [1] 廣松毅・浪花貞夫：経済時系列分析，朝倉書店 (1990)。
- [2] 刈屋武昭・矢島美寛・田中勝人・竹内啓：経済時系列の統計 ～その数理的基礎～，岩波書店，(2003)。
- [3] 間瀬茂・神保雅一・鎌倉稔成・金藤浩司：工学のためのデータサイエンス入門，数理工学社 (2004)。
- [4] 岡田昌史：The R Book データ解析環境 R の活用事例集，九天社 (2004)。
- [5] 佐和隆光：回帰分析，朝倉書店 (1979)。
- [6] 田中勝人：計量経済学，岩波書店 (1998)。
- [7] 津田博史：株式の統計学，朝倉書店 (1994)。
- [8] 山本拓：経済の時系列分析，創文社 (1988)。
- [9] R による統計処理：  
<http://aoki2.si.gunma-u.ac.jp/R/>。
- [10] SFC ホームページ：  
<http://web.sfc.keio.ac.jp/~kogure/>。