解　　説

閾値自己回帰モデルの日経平均株価指数への応用

山形大学人文学部　助教授
砂田　洋志

1．はじめに

非線形モデルの中に，GARCHモデル，閾値自己回帰モデル（以後はTARモデルと略記する），Smooth Transition Auto Regressiveモデル（以後はSTARモデルと略記する）を始めとする多様なモデルが含まれていることは，刈屋＝照井(1997)などで示されている。GARCHモデルが実証分析に広く利用されているのに対して，TARモデルやSTARモデルはこれまであまり利用されてこなかったと言えよう。そこで本稿ではTARモデルを紹介すると共に，代表的な金融データの一つと考えられる日経平均株価指数に同モデルを応用する。

TARモデルを用いた日本における先行研究としては，前述した刈屋＝照井(1997)が挙げられる。彼らは伝統的な統計学の立場に基づいて，為替レートや日本の株価（日次変化率）などに対してGARCHモデルやTARモデルを当て嵌めて，その当て嵌まりについて議論している。さらにOECDに加盟する６ヶ国のGDPとアメリカのGNPについては，ベイズ統計学に基づいてパラメータ推定と予測を行っている。TARモデルと似たモデルであるSTARモデルを用いた株価の実証研究として，Sarantis(2001)がある。同論文では日経平均株価指数の対前年度変化率の月次データを，伝統的な統計学に基づいて分析している。

本稿では，日経平均株価指数を推定するに当り，上述した先行研究とは異なり，週次データの変化率を分析する。また，推定方法としてMCMC法に基づいたベイズ推定法を利用する。

本稿の構成であるが，第２節ではTARモデルを説明する。そして第３節ではパラメータの推定方法について説明する。さらに第４節では，パラメータを推定した結果を示す。最後に第５節で結論を述べる。

2．モデルの紹介

TARモデルを紹介する前に，説明の都合上，自己回帰モデルから紹介しよう。自己回帰モデルを数式で記述すれば，

$$y_t=\gamma_0+\gamma_1 y_{t-1}+\gamma_2 y_{t-2}+\cdots+\gamma_p y_{t-p}+u_t \qquad u_t\sim N(0,\sigma^2) \qquad \cdots(1)$$

である。行列を用いて表記すれば，

$$\boldsymbol{Y}=X\gamma+\boldsymbol{u} \qquad \boldsymbol{u}\sim N(\boldsymbol{0},\sigma^2 I) \qquad \cdots(1')$$

である。ただし，YとXは，$Y=\{y_{p+1}, y_{p+2}, y_{p+3},\cdots, y_{n-1}, y_n\}'$，$X=\{\boldsymbol{Y}_{p+1}, \boldsymbol{Y}_{p+2},\cdots, \boldsymbol{Y}_n\}'$，$\boldsymbol{Y}_t=\{1, y_{t-1}, y_{t-2}, y_{t-3},\cdots, y_{t-p}\}$とする。

TARモデルでは状態の変化を考えるので，状態を何種類と考えるかによってモデルが異なってくる。最も単純なのは２つの状態しか考えない場合である。本稿では，状態が２種類のモデルを扱う。このモデルを数式で記述すれば，以下の通りである。

$$y_t=\gamma_0^{(1)}+\gamma_1^{(1)}y_{t-1}+\gamma_2^{(1)}y_{t-2}+\cdots+\gamma_{p1}^{(1)}y_{t-p1}+u^{(1)}_t$$
$$u^{(1)}_t\sim N(0,\sigma_{(1)}^2) \qquad y_{t-d}\leqq r \qquad \cdots(2)$$
$$y_t=\gamma_0^{(2)}+\gamma_1^{(2)}y_{t-1}+\gamma_2^{(2)}y_{t-2}+\cdots+\gamma_{p2}^{(2)}y_{t-p2}+u^{(2)}_t$$
$$u^{(2)}_t\sim N(0,\sigma_{(2)}^2) \qquad r<y_{t-d}$$

つまり，d期前の値(y_{t-d})が属する範囲によってt期の状態，そして自己回帰モデルが変わるのである。自己回帰モデ

ルのラグ数は状態によって変えることも可能であるが，本稿では同一のラグ数$p(=p1=p2)$としたTAR(p)モデルを扱う。

３．パラメータの推定方法について

本稿では，上述したTARモデルのパラメータをMCMC法に基づいてベイズ推定する。推定の際には，MCMC法で広く用いられる計算アルゴリズムである，ギブズ・サンプリング法とM=H法を利用する。まず自己回帰モデルのパラメータを推定するのに必要なフルコンディショナルな事後分布を示し，その後でTARモデルのパラメータのフルコンディショナルな事後分布と提案分布を紹介しよう。

３．１　自己回帰モデルのサンプリング

(1)で記述される自己回帰モデルのパラメータはγとσ^2である。事前分布を$\gamma \mid \sigma^2 \sim \mathrm{N}(\gamma_0, \sigma^2 \mathrm{A}^{-1})$と$\sigma^2 \sim \mathrm{IG}(v_0/2, \lambda_0/2)$とすれば，$\gamma$と$\sigma^2$のフルコンディショナルな事後分布はそれぞれ，以下の通りである。この確率分布に従い，γとσ^2の乱数をそれぞれ多数生成して，それらを平均した値が，γとσ^2に関するベイズ流の推定値となる。

$$f(\gamma \mid \sigma^2, y) \sim \mathrm{N}(N\left((X'X+A)^{-1}(X'X\hat{\gamma}+A\gamma_0)\right), (X'X+A)^{-1}),$$

$$f(\sigma^2 \mid \gamma, y) \sim IG\left(\frac{\tilde{v}}{2}, \frac{\tilde{\lambda}}{2}\right),$$

ただし$\tilde{\lambda} = \lambda_0 + (\gamma-\gamma_0)A^{-1}(\gamma-\gamma_0) + (y-X\gamma)'(y-X\gamma)$，$\tilde{v} = v_0+n+p+1$である。

３．２　TARモデルのサンプリング

状態が２種類のTARモデルの推定では２種類の自己回帰モデルのパラメータ推定をすることになる。状態１に対応する同モデルのパラメータを$\gamma^{(1)}$，$\sigma_{(1)}{}^2$と記述し，状態２に対応するパラメータを$\gamma^{(2)}$，$\sigma_{(2)}{}^2$と記述する。$\gamma^{(1)}$，$\sigma_{(1)}{}^2$そして$\gamma^{(2)}$，$\sigma_{(2)}{}^2$のサンプリングは3.1節の自己回帰モデルの場合を参考にされたい。ただし，事前分布のパラメータ（ハイパーパラメータ）の決定には，自己回帰モデルを最小二乗推定した結果を参考にした。

TARモデル固有のパラメータである閾値rのサンプリング方法と遅れ変数dの選択方法を説明しよう。閾値rのサンプリングは，ランダムウォーク・サンプリング法を利用する。つまり，rの候補r_tは直前の値r_{t-1}を利用して，以下のように生成する。

$r_t = r_{t-1} + \varepsilon_t$

ただし，ε_tは平均０の正規分布に従う確率変数である。そして，r_tを用いて計算した事後密度と，r_{t-1}を用いて計算した事後密度の比をとったものが，採択確率になり，この確率を利用して，r_tかr_{t-1}の何れかを選択する。こうしてrの乱数を数多く生成して，それらを平均した値がベイズ流の推定値となる。r_tのサンプリングに利用する初期値は，刈屋＝照井（1997）で紹介されている，AICを用いてパラメータ推定する方法を適用して得た。

遅れ変数dであるが，まずd=1, d=2, …と固定した上でパラメータを推定する。次にその結果を用いたAICを計算して，その大きさが最小となるdを選択する。ただし，dの候補としては自己回帰のラグ数以下と仮定する。たとえば３期のラグを仮定したTARモデルでは，dの候補として１～３の整数を考えることになる。

４．パラメータの推定

本稿では日経平均株価指数を分析するに当って，1999年３月から2003年12月末の週次データを，パーセント表

示の変化率に直してから分析した。この期間の基本統計量を**表１**に示しておこう。ちなみに，定常性が確保されていることはADF検定により確認してある。

表１　データの基本統計量

期間	平均	標準偏差	データ数
1999.3－2003.12	－0.14823	3.06615	251

４．１　パラメータの推定

TARモデルのパラメータとAICを推定する際には，ラグ数と遅れ変数を固定して推定した。今回はとりあえず，ラグ数の候補として２までを考えた。**表２**には，AICと遅れ変数及びラグ数の関係を示してある。

表２　ラグ数p，遅れ変数dとAICの関係

p,d	$(p,d)=(1,1)$	$(p,d)=(2,1)$	$(p,d)=(2,2)$
AIC	1259.63	1257.78	1250.44

以上から，TARモデルの中で最適な（AICが最小の）ものは，$p=2,d=2$のモデルとなった。ちなみに，DICや周辺尤度に基づいて選んでも同じモデルが選ばれた。Lopes(2004)ではAIC,DIC,BICを利用している。$p=2,d=2$のTARモデルのパラメータを推定した結果を**表３**に示しておこう。

表３　TARモデルの推定結果

パラメータ	事後平均	事後標準偏差	自己相関
$\gamma_0^{(1)}$	－0.60230	0.23199	－0.09560
$\gamma_1^{(1)}$	－0.06257	0.06569	0.00688
$\gamma_2^{(1)}$	－0.20247	0.08522	－0.01789
$\sigma_{(1)}^2$	9.23368	0.91858	0.01897
$\gamma_0^{(2)}$	2.81151	1.28121	0.01311
$\gamma_1^{(2)}$	0.03930	0.07571	－0.05213
$\gamma_2^{(2)}$	－0.56096	0.29721	－0.00421
$\sigma_{(2)}^2$	7.18539	1.46237	0.06118
r	2.73338	0.04604	0.68798

注：サンプリングは６千回行い，推定には後の４千回のサンプルを利用した。
rの採択率は0.2465である。

４．２　推定結果の考察

推定の結果，AICが最も小さいモデルとして，$p=2,d=2$のTARモデルが選ばれた。このTARモデルの閾値は2.73と推定されたが，これは全データの上位約20%に位置する値である。回帰係数については，定数項$\gamma_0^{(1)}$, $\gamma_0^{(2)}$と状態１におけるy_{t-2}の係数$\gamma_2^{(1)}$が有意であった。また，状態２におけるy_{t-2}の係数$\gamma_2^{(2)}$もほぼ有意と言える。

５．結論

本稿では，これまであまり利用されることのなかったTARモデルを紹介すると共に，同モデルを用いて1999年３月から2003年12月という期間の日経平均株価指数の週次データを分析した。具体的には，同指数の変化率をパーセント表示にして，$p=2,d=2$のTARモデルのパラメータをベイズ推定した。今後さらに，TARモデルやSTARモデルを株式市場の分析に応用してみることが待たれる。

参考文献

[1] 刈屋武昭・照井伸彦, 1997,『非線形経済時系列分析法とその応用』, 岩波書店.

[2] Lopes, H. F. and Salazar, E., 2006, “Time Series Mean Level and Stochastic Volatility Modeling Smooth Transition Autoregressions: A Bayesian Approach,” Advance in Econometrics, Part B, Vol. 20, pp225-238.

[3] 中妻照雄, 2003,『ファイナンスのためのMCMC法によるベイズ分析』, 三菱経済研究所.

[4] Sarantis, N., 2001, “Nonlinearity, Cyclical Behaviour and Predictability in Stock Markets: International Evidence,” International Journal of Forecasting, Vol. 17, pp459-482.

１．ＲＮプライム指数構成銘柄の一部変更

キリンビバレッジ(株)（コード：2595）株式の整理ポストへの割当てに伴い，平成18年７月13日（木）より，Russell/Nomura Primeインデックス（ＲＮプライム指数）構成銘柄から当該株式を除外することについての発表がありましたので，お知らせいたします。

２．日経株価指数300の構成銘柄の一部入替え

(株)日本経済新聞社より，日経株価指数300構成銘柄である「大和工商リース（コード：9762）」が完全子会社化に伴い上場廃止となるため，平成18年７月26日（水）に同銘柄を除外し，「イオンクレジットサービス（コード：8570）」を補充するとの発表がありましたので，お知らせします。

１．日経株価指数300

除外銘柄	補充銘柄	実施日
大和工商リース（9762）	イオンクレジットサービス（8570）	平成18年７月26日（水）