Box-Cox 変数変換を含む多変量 ARMA モデルとその応用について

寺　坂　崇　宏

1．はじめに

経済変数間の関係を，データを用いて実証的に検証する際に，分析の対象となる変数に対して，何らかの変換を施して分析を進めるという方法は，基本的な方法として知られている。例えば，変数に対して平方根の変換を適用したり，あるいは対数変換を施したりしてから，分析を進めるという方法は，頻繁に用いられる方法である。

このような，変数に変換を施して分析を進める理由であるが，１つは統計学的観点から，例えば，Box and Cox（1964）が指摘しているように，変数の分布を正規分布に近づけること，あるいは，計量経済モデルや時系列モデル等の誤差項部分の分布を，できるだけ正規分布に近づけることにある。さらに，誤差項が，独立で同一の分布に従うという仮定を満足させることを目的に，施されることも考えられる。これは，変換が機能した場合，正規分布に基づく統計的推測の理論，あるいは，より一般的な統計的推測の議論の適用の可能性が高まり，分析がより厳密かつ簡潔に進めることができるようになる。

もう１つは，経済学的観点から，経済学の分析で用いられる対数の差分による変化率（成長率，増加率）の表現や，経済モデルに基づく変数間の非線形の関係を表現するために，変換を施すことがある。例えば，経済モデルでよく用いられるコブ＝ダグラス型生産関数や，それを含んだ，より一般の形状のCES 型生産関数は，明らかに変数に対して非線形であるが，Zalembka（1970）のように変数を変換して推定したり，Ramsey and Zarembka（1971）のよう

に，変数変換の代表的なものであるBox-Cox変換を用いて特定化することができる。

本稿では，Terasaka and Hosoya（2008），Hosoya and Terasaka（2009）で検討されている，修正Box-Cox変換と多変量ARMAモデルを組み合わせた時系列モデルを利用して，日本の株価（東証株価指数）と金利（コールレート）との関係を，2006年8月から2009年8月の月次データについて推定し，検討をすることを目的とする。特に，変換を施して分析することによる，残差の正規性について検討をすることを目的とする。一方，変換のパラメータの推計値が，統計学的に有意であるかについての検定は，Hosoya and Terasaka（2009）で，その方法と意味について議論しているが，計算費用の関係から本稿では，検討を加えないことにする。修正Box-Cox変換は，検定の議論で関係してくる。よって実質的にはオリジナルのBox-Cox変換を用いた分析となる。分析に使用したモデルは，変換したプロセスが定常プロセスに従うという時系列モデルを想定している。よって変数としてレベルではなく，それぞれの系列について前月比を取ったものを推定に使用している。

論文の構成であるが，2節でBox-Cox変換およびHosoya and Terasaka（2009）で提案した修正されたBox-Cox変換を伴う多変量ARMAモデルについて簡単に紹介する。3節でこのモデルを用いた株価と金利の分析の方法について説明する。4節で分析結果について検討する。最後に結論を述べる。

2．Box-Cox変換およびBox-Cox変換を伴うARMAモデルについて

変数変換の代表的なものとして，Box and Cox（1964）が提案したBox-Cox変換がある。この変換は，変数yを変換のパラメータλを用いて変数

$$y^{[\lambda]} = \begin{cases} \dfrac{y^{\lambda}-1}{\lambda} & (\lambda \neq 0) \\ \log y & (\lambda = 0) \end{cases} \tag{1}$$

と定義される。Box と Cox は，この変換の目的について，推計に使用するモデルの構造の単純化，つまり，モデルの関数形の線形化，誤差分散の均一化，分布の正規化を達成することであると述べている。Box-Cox 変換は，変換後の変数の取りうる範囲に bound が存在するという特性がある。つまり，もし $\lambda>0$ のときは，$y^{[\lambda]}$の取りうる範囲は$y^{[\lambda]}>-\frac{1}{\lambda}$，$\lambda<0$ のときは，$y^{[\lambda]}<-\frac{1}{\lambda}$ である。この場合，例えば変数そのもの分布に正規分布を仮定する場合や，あるいは，統計モデルにおける攪乱項に正規分布を仮定する場合，正規分布の取りうる範囲は $-\infty$ から $+\infty$ なので，この点において厳密には問題が生じることになる。

Box and Cox（1964）は，この変換を次のモデル

$$y^{[\lambda]}=X\beta+\varepsilon \tag{2}$$

対して適用して議論している。ここで$y^{[\lambda]}$は $n\times 1$ の従属変数ベクトル，X は $n\times k$ の説明変数ベクトル，β は $k\times 1$ のパラメータベクトル，そして ε は $n\times 1$ の同一分散で互いに独立な正規分布に従う攪乱項であるとする。このモデルは，従属変数をパラメータで変換することで，攪乱項部分の同一分散の正規分布に従うということを仮定している。

このモデルを推定する 1 つの方法として，Box and Cox（1964）では，最尤法を議論している。つまり，(2)式の尤度関数

$$L(\theta,\lambda)=\frac{1}{(2\pi)^{\frac{n}{2}}\sigma^{n}}\exp\left\{(y^{[\lambda]}-X\theta)'(y^{[\lambda]}-X\theta)\right\}|J(\lambda;y)| \tag{3}$$

をパラメータ θ，λ について最大化する。ここで，$J(\lambda;y)$ は変換のヤコビアンに対応する部分で，

$$J(\lambda;y)=\prod_{i=1}^{n}\left|\frac{dy_i^{[\lambda]}}{dy_i}\right| \tag{4}$$

である。ただし，尤度関数の最大化において，パラメータ θ，λ について同時

に推定することについては述べられていない。その後，例えばDavidson and MacKinnon（1982）が議論しているように，(3)式の最適化では，同時推定することが望ましいことが明らかになっている。

Terasaka（2005）では，Box-Cox変換を多変量ARMAモデルに適用させたモデルを提案している。さらに，Terasaka and Hosoya（2008）及びHosoya and Terasaka（2009）ではBox-Coxの変換のboundを取り払った1つの修正Box-Cox変換を提案し，この変換を伴う多変量ARMAモデルを提示して，その漸近的な特性および数値推定法，有限標本における数値評価およびその応用例を示している。この変換を伴う多変量ARMAモデルは次のように書ける。

$$\sum_{j=0}^{a} A(j)\, y^{[\lambda]}(t-j) = \mu + \tau t + \sum_{k=0}^{b} B(k)\, \varepsilon(t-k) \tag{5}$$

ここで，原系列に対して修正Box-Cox変換を施した系列 $\{y^{[\lambda]}(t),\ t\in\mathbb{Z}\}$ はトレンド定常過程に従うプロセスで，$A(0)=B(0)=I_m$ で $y^{[\lambda]}(t)=(y_1{}^{[\lambda_1]}(t),\ldots,y_m{}^{[\lambda_m]}(t))'$, $\varepsilon(t)=(\varepsilon_1(t),\ldots,\varepsilon_m(t))'$, $A(j)$, $B(k)$ は $m\times m$ 定数行列，定数項 μ は $1\times m$ ベクトル，トレンド項 τ は $1\times m$ ベクトルとする。また，特性多項式 $\det\{\sum_{j=0}^{a}A(j)z^j\}$ と $\det\{\sum_{k=0}^{b}B(k)z^j\}$ の根は単位円の外にあり，$\{y^{[\lambda]}(t)\}$ は反転可能なGaussianプロセスとする。また，特性多項式の根には共通の解を有さないものとする。

(5)式を特定化するには，モデルのパラメータである λ, $A(j)$, $B(k)$, μ, τ と，モデルの次数 (a, b) を推定する必要がある。これらの推定は，Hosoya and Terasaka（2009）の方法により行った。ごく簡単に説明すると，次数の決定方法については，Hannan and Rissanen（1982）の方法を活用した。次数 (a, b) が既知で，$y(0)=\cdots=y(-a+1)=0$, $\varepsilon(0)=\cdots=\varepsilon(-b+1)=0$ の下で，(5)式の対数尤度関数は

$$\log(L_T) = -\frac{1}{2}mT\log(2\pi) - \frac{1}{2}T\log\det\Sigma$$

$$
-\frac{1}{2}\sum_{t=1}^{T}\left\{\left(y^{[\lambda]}(t)+\sum_{j=1}^{a}A(j)y^{[\lambda]}(t-j)-\mu-\tau t-\sum_{k=1}^{b}B(k)\varepsilon(t-k)\right)'\times\right.
$$

$$
\left.\Sigma^{-1}\left(y^{[\lambda]}(t)+\sum_{j=1}^{a}A(j)y^{[\lambda]}(t-j)-\mu-\tau t-\sum_{k=1}^{b}B(k)\varepsilon(t-k)\right)\right\}
$$

$$
+\sum_{t=1}^{T}\sum_{l=1}^{m}k_l\,(y_l\,(t),\,\lambda_l) \tag{6}
$$

と書ける。ここで，$k_l\,(y_l\,(t),\,\lambda_l)$ は変数 $y\,(t)$ から $y^{[\lambda]}(t)$ への変換のヤコビアンの l 番目の要素である $\left|\dfrac{\partial {y_l}^{[\lambda]}(t)}{\partial y_l(t)}\right|$ の対数を指す。この関数を λ，$A\,(j)$，$B\,(k)$，μ，τ について，同時に最適化を実施して，最終的なパラメータの推定値を決定した。

3．データ分析

本稿では，Terasaka (2005)，Terasaka and Hosoya (2008) と同様に株価（TOPIX）と金利（コールレート）との関係を，提案したモデル(5)式を用いて推定した。分析には月次データを用いた。株価のデータは，東証株価指数の毎月末のデータを取り，その対前月比のデータを利用した。また，金利のデータは，コールレート（無担保翌日物）の毎月末のデータを分析に使用した。東証株価指数のデータは，東京証券取引所発行の，東証統計月報と東京証券取引所のホームページから，コールレートのデータは，日本銀行のホームページからそれぞれ入手した。対前月比のデータを利用したのは，モデル(5)が変換したプロセスが定常性を満たす必要があるためであるが，これらの原系列は，明らかに定常性を満たさないと考えられるためである。

推定した式であるが，次の 2 種類の式を推定した。1つは，変数変換のパラメータに制約を設けないで式を推定し，もう 1つは，変数変換のパラメータを 0 に制約して式を推定した。後者の式の意味であるが，Box-Cox 変換では，

変換のパラメータを0としたときは，変数は対数に変換される。対前期の比率データに対して対数変換を施したものは，変数の変化率を意味する。つまり，後者の式は，変数の変化率の動きを推計していることになる。すると，2式の推定結果を比較検討することで，変化率の動きを推計している式よりその残差部分が，より正規性を満たすように，変数を変換して式を推定できるのかを検討することができる。

推定期間であるが，2006年8月から2009年8月に設定したのは，本稿の作成時点で，出来るだけ最近の動きについて分析するためと，データ系列を見ると，2005年10月末にコールレートが0％になり，比率データを求めると，2005年11月のデータが異常値になってしまい分析が困難になることと，この後2006年7月まで，コールレートの対前期比の動きが非常に大きいため，分析結果に問題が生じることが予想されたためである。この時期は，日本銀行により金利が非常に低く誘導されていたために，金利のわずかな変化でも，対前月比では大きな動きになってしまうという特殊な時期であり，この時期を入れた長期的な分析をする場合，例えばモデルに構造変化を取り入れる必要があろう。

尚，計算は東北大学情報シナジーセンターの大規模計算システムのSX-aで実施し，計算のためのプログラムはFortranで作成した。

4. 推 計 結 果

推計結果は以下のようになった。はじめに，変数変換に制約を設けて推定した結果を示す。推定された次数はARMA（1,0）で，推定値は，

$$\hat{A}(1)=\begin{pmatrix}-0.382 & -0.0327\\ -1.321 & 0.146\end{pmatrix},\ \hat{\mu}=\begin{pmatrix}-0.0111\\ 0.100\end{pmatrix},\ \hat{\tau}=\begin{pmatrix}0.00167\\ -0.00578\end{pmatrix}$$

$$\Sigma=\begin{pmatrix}3.07\times10^{-3} & -7.89\times10^{-4}\\ -7.89\times10^{-4} & 2.79\times10^{-2}\end{pmatrix}$$

対数尤度は8.54であった。次に，変数変換に制約を設けないで推定した結果を

示す。推定された次数は ARMA（1,0）で，推定値は，

$$\hat{\lambda} = \begin{pmatrix} 3.77 \\ 1.58 \end{pmatrix}, \hat{A}(1) = \begin{pmatrix} -0.341 & -0.0679 \\ -0.952 & 0.309 \end{pmatrix},$$

$$\hat{\mu} = \begin{pmatrix} -0.0310 \\ 0.0777 \end{pmatrix}, \hat{\tau} = \begin{pmatrix} 0.00108 \\ -0.00380 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 3.07\times10^{-3} & -7.89\times10^{-4} \\ -7.89\times10^{-4} & 2.79\times10^{-2} \end{pmatrix}$$

であった。対数尤度は9.24であり，制約を設けないで推定したモデルの対数尤度が，設けて推定したモデルの対数尤度より高いので，適切な結果である。

2つの式を推定した際，次数の選択は，Hannan and Rissanen（1982）に基づき，BIC を用いて実施したが，その選択の状況は，次のようになった。表1は，変換に制約を設けて推定した場合，表2は，変換に制約を設けないで推定した場合の結果である。どちらの場合についても，ARMA の次数が（1,1）で選択されていることがわかる。

表1　次数の選択
（制約を設けて推定した場合）

AR 次数	*MA* 次数	BIC
1	0	−7.22
0	1	−7.16
1	1	−7.06
0	2	−7.05
2	0	−7.05

表2　次数の選択
（制約を設けず推定した場合）

AR 次数	*MA* 次数	BIC
1	0	−7.65
2	0	−7.31
1	1	−7.01
3	0	−6.77
2	1	−6.72

Doornik and Hansen（1994）による残差の正規性の検定の統計量を計算したところ，変換のパラメータに0の制約をおいたモデルが11.6で，これは残差が正規分布であるという帰無仮説を有意水準5％で棄却した。他方，制約をお

かないで推定したモデルが1.73で，これは正規分布であるという帰無仮説を有意水準5％で棄却しなかった。よって，このケースでは，変数変換が正規性に改善をもたらしていると考えられ，これらの系列では，統計学的な意味において，変化率のモデルより優れた変換のモデルが存在することが示唆される。ただし，変換のパラメータの値が統計学的に有意であるかどうかについては，Hosoya and Terasaka（2009）で議論している，モンテカルロ法によるワルド検定を利用して検定する必要がある。図1はTOPIXの対前月比の推移と

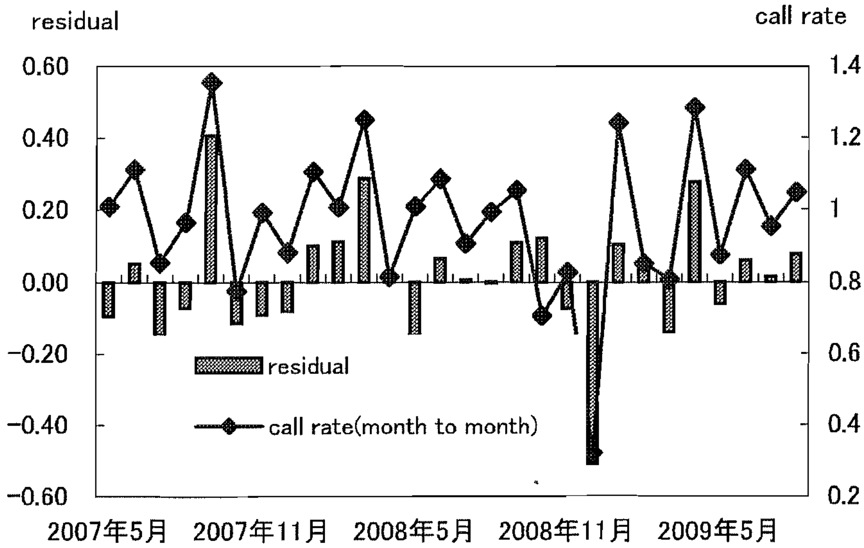

図1　コールレートの推移と残差

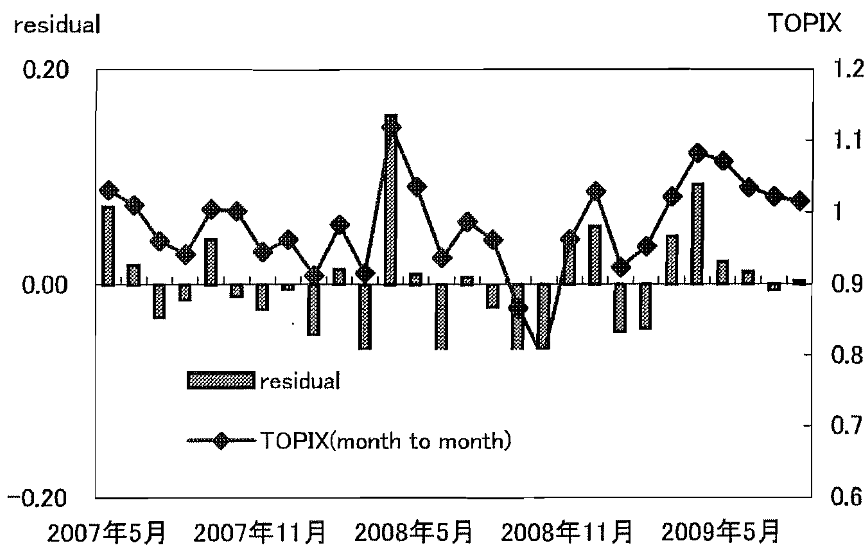

図2　TOPIXの推移と残差

推定されたモデルの残差を，図２はコールレートの対前月比の推移と，推定されたモデルの残差をそれぞれ示している。

いずれの残差の系列においても，変換がある程度機能していることが推察される。ここで，いずれの図においても2007年５月から書かれているのは，Hannan 法に基づく推定を実施する際，2006年８月から2007年４月までの期間を使用する必要があるためである。

5．結　　論

本稿では，変換を伴う ARMA モデルを用いて TOPIX とコールレートの推計を実施した。推計の結果，変換が機能していることを示唆する結果が得られたが，変換のパラメータの推定値が統計学的に有意であるかの検討までは実施していないので，この点については今後の課題である。また，データの推定期間については期間が短いので，この点についても何らかの改善が必要である。

参考文献

Box, G.E.P., Cox, D.R., 1964 An analysis of transformations. Journal of the Royal Statistical Society B 26, 211-252

Davidson, R., MacKinnon, J., 1984 Model specification tests based on artificial regressions. International economic review 25, 485-502.

Doornik, J.A., Hansen, H., 1994 An omnibus test for univariate and multivariate normality. In : Mimeograph. Nuffield College, Oxford.

Hannan, E.J., Rissanen, J., 1982 Recursive estimation of mixed autoregressive-moving average order, Biometrica, 69, 81-94

Hosoya, Y., Terasaka, T., 2009. Inference on transformed stationary time series. Journal of Econometrics 151, 129-139.

Ramsey, J.B., Zalembka, P., 1971 Specification error tests and alternative functional forms of the Aggregate production function. Journal of the American statistical association, 66,471-477

Terasaka, T., 2005. The Box-Cox transformation in the multivariate ARMA model. Annual report of economic society Tohoku University, 66, 791-811.

Terasaka, T., Hosoya, Y., 2007. A modified Box-Cox transformation in the multivariate ARMA model. Journal of Japan Statistical Society, 37, 1-28

Zalembka, P., 1970 On the empirical relevance of the CES production function. The review of economics and statistics. 52, 47-53