時系列解析と予測
– 因果性と反転可能性 –

2002MM100 棚瀬 暁俊
指導教員　國田　寛

1　はじめに

一般的に時系列における局所的な平均値の変動および分散の変動は存在することが当然と考えられるものであり，分析においては非定常過程を前提としたアプローチがなされるべきである．したがって原時系列より，周期の除去などの方法により，因果的定常過程を導出しそれに対して解析を行う方法がなされる．
因果的なモデルとは，現時点において時系列をモデリングをする際に時系列の過去と現在の情報から構築されるモデルであり，将来の情報から構築されるモデルは非因果的なモデルという．しかし非因果的なモデルはあまり一般的ではなく前者のモデルがより自然なモデルである．しかし常にこのように因果的なモデルが構築できるわけではなく，モデリングをするために重要な条件である因果性や相対的な概念にある反転可能性について考察する．
さらに定常時系列の重要なパラメトリックな族である因果的で反転可能な ARMA(自己回帰移動平均) 過程を拡張した非定常性をもつ確率過程により予測を行い考察する．

2　線形過程

時系列 $\{X_t\}$ は，すべての t について

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j} \tag{1}$$

を表されるとき，線形過程である，ここで $\{Z_t\}$〜WN(0, σ^2) であり，$\{\psi_j\}$ は $\sum_{j=-\infty}^{\infty} \mid \psi_j \mid < \infty$ となる定数の列である．線形過程には定常性をもつ AR 過程，MA 過程，ARMA 過程，非定常性をもつ ARIMA 過程，SARIMA 過程などがあり．以下，予測で取り扱う ARMA 過程，および SARIMA 過程について述べる．

2.1　一変量 ARMA 過程

MA 過程の出力が AR 過程で記述されるような $\{X_t\}$ をもつ場合に ARMA 過程である．$\{X_t\}$ は定常で，すべての t に対して

$$\phi(B)X_t = \theta(B)Z_t \tag{2}$$

が成り立つ．ここで $\phi(\cdot)$ と$\theta(\cdot)$ はそれぞれ p 次と q 次の多項式 $\phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p, \theta(z) = 1 + \theta_1 z + \ldots + \theta_q z^q$ である．この 2 つの多項式は共通因子を持たない．

2.2　SARIMA 過程

周期 s の成分を取り除くために系列 $\{X_t\}$ にラグ s の差分を取ることが有効である．もし ARMA(p, q) 過程 $\phi(B)Y_t = \theta(B)Z_t$ を差分系列 $Y_t = (1 - B^s)X_t$ にあてはめてみると，原系列のモデルは $\phi(B)(1-B^s)X_t = \theta(B)Z_t$ となる．もし d と D が非負の整数であるとき，$\{X_t\}$ が周期 s の季節 ARIMA(p, d, q) $\times$ (P, D, Q)$_s$ 過程であるとは，その差分系列 $Y_t = (1 - B)^d(1 - B^s)^D X_t$ が

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)Z_t,\ \{Z_t\} \sim \mathrm{WN}(0, \sigma^2) \tag{3}$$

のような因果的 ARMA 過程であるという．ここで $\phi(z) = 1-\phi_1 z-\cdots-\phi_p z^p, \Phi(z) = 1-\Phi_1 z-\cdots-\Phi_P z^P$, $\theta(z) = 1+\theta_1 z+\cdots+\theta_q z^q, \Theta(z) = 1+\Theta_1 z+\cdots+\Theta_Q z^Q$, である．

3　ARMA(p, q) 過程の因果性

ARMA(p, q) 過程 $\{X_t\}$ が因果的であるとは，あるいは $\{Z_t\}$ の因果的であるとは，$\sum_{j=0}^{\infty} \mid \psi_j \mid < \infty$ および

$$\text{すべての } t \text{ について } X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j} \tag{4}$$

を満たす定数 $\{\psi_j\}$ が存在することをいう．因果性は

$$\text{すべての } \mid z \mid \leq 1 \text{ について} \phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0 \tag{5}$$

という条件と同値である．

4　ARMA(p, q) 過程の反転可能性

反転可能性は，Z_t が X_s, $s \leq t$ によって表されるということであるが，移動平均多項式によって同じように特徴づけられる．ARMA(p, q) 過程 $\{X_t\}$ が反転可能であるとは，$\sum_{j=0}^{\infty} \mid \pi_j \mid < \infty$ および

$$\text{すべての } t \text{ ついて}\quad Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \tag{6}$$

を満たす定数 $\{\pi_j\}$ が存在することをいう．反転可能性は

$$\text{すべての } \mid z \mid \leq 1 \text{ について} \theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q \neq 0 \tag{7}$$

という条件と同値である．

5　ARMA 過程の予測

因果的な ARMA 過程の予測の問題を考えるには，過程 $\{X_t\}$ そのものではなく，変換された過程

$$W_t = \begin{cases} \sigma^{-1}X_t & t=1,...,\,m \\ \sigma^{-1}\phi(B)X_t & t>m \end{cases} \quad (8)$$

を用いる方がよい．ただし $m = max(p,\,q)$ である．$\theta_0 \equiv 1$, $j > q$ について $\theta_j \equiv 0$ と定義する．さらに $p \geq 1$, $q \geq 1$ と仮定する．ここで $\{X_t\}$ の自己共分散関数 $\kappa(i,\,j)=E(W_i,\,W_j)$, $i,\,j \geq 1$ を求める．

W_{n+1} と X_{n+1} の 1 期先予測量は，$\hat{W}_{n+1} = P_nW_{n+1}$ および $\hat{X}_{n+1} = P_nX_{n+1}$ によって与えられる．イノベーションアルゴリズムを過程 $\{W_t\}$ に用いると，

$$\begin{cases} \hat{W}_{n+1} = \sum_{j=1}^{n}\theta_{n,\,j}(W_{n+1-j} - \hat{W}_{n+1-j}), & 1 \leq n < m \\ \hat{W}_{n+1} = \sum_{j=1}^{q}\theta_{n,\,j}(W_{n+1-j} - \hat{W}_{n+1-j}), & n \geq m \end{cases}$$

が得られる．

$\{1,\,X_1,...,\,X_n\}$ による任意の確率変数 Y の最良線形予測量は，$\{1,\,W_1,...,\,W_n\}$ による Y の最良線形予測量と同じであり，この予測量を P_nY で表すことにする．P_n の線形性と (8) 式を用いると，$\hat{W}_t$ を得る．(8) 式よりすべての $t \geq 1$ について，$X_t - \hat{X}_t = \sigma[W_t - \hat{W}_t]$ であることを示している．(8) 式における $(W_j - \hat{W}_j)$ を $\sigma^{-1}(X_j - \hat{X}_j)$ で置き換えて，$\hat{W}_{n+1}$ に適用すると，結果として $\hat{X}_t$ は

$$\begin{cases} \sum_{j=1}^{n}\theta_{n,\,j}(X_{n+1-j} - \hat{X}_{n+1-j}), & 1 \leq n < m \\ \phi_1X_n + \cdots + \phi_pX_{n+1-p} + \\ \quad \sum_{j=1}^{q}\theta_{n,\,j}(X_{n+1-j} - \hat{X}_{n+1-j}), & n \geq m \end{cases} \quad (9)$$

を得る．ここで $\theta_{n,\,j}$ は κ を用いたイノベーションアルゴリズムから求められる．(9) 式は 1 期先予測量 $\hat{X}_2, \hat{X}_3, ...$ を逐次的に定める．

6　SARIMA 過程の予測

はじめの $d + Ds$ 個の観測値 $X_{-d-Ds+1},...,\,X_0$ が $\{Y_t,\,t \geq 1\}$ と無相関であるという仮定のもとで，X_{n+h} の $\{1,\,X_{-d-Ds+1},...,\,X_n\}$ に基づく最良線形予測量 P_nX_{n+h} は

$$P_nX_{n+h} = P_nY_{n+h} + \sum_{j=1}^{d\,+\,Ds} a_jP_nX_{n+h-j} \quad (10)$$

となることがわかる．右辺の第 1 項はちょうど ARMA 過程 $\{Y_t\}$ の $\{1,\,Y_1,...,\,Y_n\}$ による最良線形予測量となっていて，ARMA 過程の予測として計算される．予測量 P_nX_{n+h} は (11) 式から各 $j \geq 1$ に対して $P_nX_{n+1-j} = X_{n+1-j}$ であることに注意すると，$h = 1,\,2, ...$ に対して逐次的に計算される．

7　時系列データの予測

ここで 1995 年〜2002 年就業形態別常用雇用指数－就業形態計 (30 人以上) の月次データ全 96 期 (12 ヶ月 ×8 年分) を扱った．

今回の時系列データは月次データであるから周期 12 として考えるためラグ 12 の差分をとる．しかし，まだ非定常性が見ることができるので，さらにラグ 1 の差分をとることにした．つまり SARIMA 過程における差分系列を $d = 1$, $D = 1$, $s = 12$ として，これは $Y_t = \nabla\nabla_{12}X_t = (1-B)(1-B^{12})X_t$ となる．そして Y_t が因果的で反転可能な ARMA 過程ならば (2) 式を満たす．

プログラム ITSM により ARMA(1, 1) 過程が得られた．ここで我々は $\hat{Y}_{t+h}$, $h = 1,\,2, ...$ を知りたいとする．(9) 式は一本の式

$$\hat{Y}_{t+1} = -0.\,8382Y_t + \theta_{t,\,1}(Y_t - \hat{Y}_t),\ t \geq 1 \quad (11)$$

で表される．$\theta_{t,\,1}$ をイノベーションアルゴリズムで求めることによって $\hat{Y}_{84}$=-0.02221 の値が求められると $\hat{Y}_{85}$ 以降の将来の予測量は

$$\hat{Y}_{t+1} = -0.8382\hat{Y}_t,\ t \geq 84$$

簡単な計算から求めることができる．

我々が求めたいのは実データの予測量，つまり $\hat{X}_{t+1}$, $t \geq 97$, を知りたい．(10) 式のように

$$\hat{X}_t = \hat{Y}_{t-13} + X_{t-1} + X_{t-12} - X_{t-13},\ t \geq 97 \quad (12)$$

の関係式を得ることができる．先に求めた $\hat{Y}_t$, $t \geq 84$, の値を (13) 式に適用することにより，$\hat{X}_{97}, \hat{X}_{98}, ...$ と 1 期先予測量を逐次的に求めることができる．

8　おわりに

本論文では，線形性，定常性に焦点を当てて考察を行った．やはり様々な時系列データを取り扱うとき，より正確な解析，予測をしたいのであれば，より当てはまりのよいモデルを考えなければならない．今後，様々な性質を持つ時系列に対して，どのような確率モデルがよいのか理解を深めていきたい．

参考文献

[1] P．J．ブロックウェル，R．A．デービス著:
逸見功，田中稔，宇佐美嘉弘，渡辺則夫訳:
入門　時系列解析と予測

[2] 1995 年〜2002 年就業形態別常用雇用指数 (2000 年の年平均を 100 としている) －就業形態計 (30 人以上) の月次データ