数理解析研究所講究録
418巻 1981年 48-56

Autoregressive Integrated Moving Average (p, d, q) モデルについて

東工大 理　藤井光昭

1. 序論

Box-Jenkins により〔2〕において論じられている Autoregressive Integrated Moving Average (p, d, q)（以下において〔2〕に従って ARIMA(p, d, q) と略）モデルは時系列解析の応用分野において現在広く用いられているが，その数学的な取扱いにおいて種々あいまいな点があり，推定理論等を数学的に展開するためにはこれらを明確にしておく必要がある。〔4〕においてこのような方向での試みをおこなったが，ここではこの考えと方法をさらに拡張して ARIMA(p, d, q) モデルやそれに関連する概念の定義をおこない，これを用いてモデルの未知パラメータの推定と初期値の取扱いについて論じる。

この節では，以降において用いるため，準備として記号および〔4〕の結果のなかで必要な部分だけまとめておく。

tを整数の値をとる時間パラメータとし、X_t を $t \geqq -d+1$ で定義された実数値をとる確率過程で $EX_t^2 < +\infty$ であるものとする。ここで d はある正整数である。∇は差分をとるオペレータをあらわすものとし、$\nabla X_t = X_t - X_{t-1}$, $\nabla^l X_t = \nabla(\nabla^{l-1} X_t)$ $(l \geqq 2)$ で定義されるものとする。確率変数 Z の $\{X_j ; -d+1 \leqq j \leqq 0\}$ による条件つき平均値を $E(Z/\{X_j ; -d+1 \leqq j \leqq 0\}) = \hat{E}Z$ のようにあらわすことにする。このとき $t \geqq 1$ である任意な t について $Y_t = \nabla^d X_t$ は $\hat{E}Y_t = 0$, $\hat{E}Y_t Y_{t+h} = R_h$ となるという意味で弱定常過程であるものとする。X_tが $t \geqq -d+1$ でGauss過程であり、Y_t が $t \geqq 1$ において弱定常であるとすれば、Y_t $(t \geqq 1)$ は $\{X_j ; -d+1 \leqq j \leqq 0\}$ と独立になる。

つぎに実数値をとり $EZ_t^2 < +\infty$ である確率過程 Z_t について $\{Z_t ; t_1 \leqq t \leqq t_2\}$ で生成される Hilbert 空間を $L^2(Z ; t_1, t_2)$ のようにあらわすことにする。そして確率変数 Z_t の $L^2(Z ; t_1, t_2)$ 上への射影を $P_{L^2(Z ; t_1, t_2)} Z_t$ のように書くことにし、そして $\hat{E}Z_t^2$ を $\|Z_t\|^2$ のようにあらわすことにする。

white noise a_t $(t \geqq 1)$ を以下のように構成することにする。

$$a_1 = Y_1 / \|Y_1\|$$

$$a_t = (Y_t - P_{L^2(Y ; 1, t-1)} Y_t) / \|Y_t - P_{L^2(Y ; 1, t-1)} Y_t\|$$

$(t \geqq 2)$.

いま R_k は $\sum_{k=0}^{\infty}|R_k|<+\infty$ をみたすものとし，R_k を用いて $f(\lambda)=\sum_{k=-\infty}^{\infty} R_k \cos 2\pi k\lambda$ を定義する。このとき関数 $f(\lambda)$ が $\int_{-\frac{1}{2}}^{\frac{1}{2}}|\log f(\lambda)|d\lambda<+\infty$ であれば上記の構成は可能である。構成法から任意な t $(t \geqq 1)$ について $L^2(Y;1,t)=L^2(a;1,t)$ であることがわかる。

2. Autoregressive Integrated Moving Average (p, d, q) 過程について

$Y_t=\nabla^d X_t$ $(t \geqq 1)$ において d は Y_t が弱定常となる最小の正整数とする。これから Y_t が Autoregressive Moving Average (p, q)（以下において〔2〕に従って ARMA(p, q) と略）過程であることの定義を p, q に注意しておこなう。通常の定常過程においては，周波数領域での解析から〔3〕等において論じられ，〔1〕，〔5〕においても論じられているが，ここでは少し異った観点から考察することにする。

$\{X_t ; t_1 \leqq t \leqq t_2\}$, $\{a_s ; s_1 \leqq s \leqq s_2\}$ で生成される Hilbert 空間を $L^2((Y;t_1,t_2),(a;s_1,s_2))$ のようにあらわしていく。

(i) $t \geqq 2$ である任意な t について $\|P_{L^2(Y;1,t-1)} Y_t\|=0$ のとき $Y_t=\|Y_1\| a_t$ となり直交する確率変数列となる。

以下において $\|P_{L^2}(Y;1,t-1)Y_t\| \neq 0$ となる t が存在するものとする。定常性によりある t_a において $\|P_{L^2}(Y;1,t_a-1)Y_{t_a}\| \neq 0$ となれば，任意な t，$t \geqq t_a$，について $\|P_{L^2}(Y;1,t-1)Y_t\| \neq 0$ となる。

(ii) $1 \leqq k \leqq p-1$ のすべての k について

$$P_{L^2}(Y;t-k,t-1)Y_t = P_{L^2}(Y;1,t-1)Y_t \qquad (t \geqq k+1)$$

ということが成り立たず，

$$P_{L^2}(Y;t-p,t-1)Y_t = P_{L^2}(Y;1,t-1)Y_t \qquad \cdots\cdots\cdots (1)$$

が任意な t $(t \geqq p+1)$ に対して成り立つとき X_t を次数 (p,d) の Autoregressive Integrated（ARI(p,d) と略）過程とよぶことにする。

もし有限な正整数 p をどのようにとっても (1) が成り立たないとき，

$$P_{L^2}(a;t-1,t-1)Y_t = P_{L^2}(Y;1,t-1)Y_t \qquad \cdots\cdots\cdots (2)$$

が任意な t $(t \geqq 3)$ に対して成り立つならば X_t を次数 $(d,1)$ の Integrated Moving Average（IMA$(d,1)$ と略）過程という。もし (2) が $t \geqq 3$ について成り立たず，$1 \leqq k \leqq p-1$ である k について

$$P_{L^2}((Y;t-k,t-1),(a;t-1,t-1))Y_t = P_{L^2}(Y;1,t-1)Y_t \quad \cdots\cdots (3)$$

が $t \geqq k+1$ で成立せず，

$$P_{L^2}((Y;t-p,t-1),(a;t-1,t-1))Y_t = P_{L^2}(Y;1,t-1)Y_t \quad \cdots\cdots (4)$$

が任意な t $(t \geqq p+1)$ について成立するとき，　X_t を次数

$(p, d, 1)$ の Autoregressive Integrated Moving Average 過程とよぶ。

以下 上と同様のことをくりかえしていく。一般的には次のようになる

(iii) q を正整数とするときどのような正整数 k をとっても

$$P_{L^2((Y;t-k,t-1),(a;t-q+1,t-1))}Y_t = P_{L^2(Y;1,t-1)}Y_t$$

が $t \geqq \max(k, q-1)+1$ であるすべての t について成り立つということが生じないものとする。

このときもし

$$P_{L^2(a;t-q,t-1)}Y_t = P_{L^2(Y;1,t-1)}Y_t \qquad \cdots\cdots\cdots (5)$$

がすべての t $(t \geqq q+1)$ に対して成り立つならば X_t を次数 (d, q) の Integrated Moving Average (IMA(d,q)と略) 過程という。

つぎに (5) が $t \geqq q+1$ であるすべての t について成り立つということがなく, $1 \leqq k \leqq p-1$ のすべての k について

$$P_{L^2((Y;t-k,t-1),(a;t-q,t-1))}Y_t = P_{L^2(Y;1,t-1)}Y_t$$

が $t \geqq \max(k, q)+1$ であるすべての t に対して成り立つということも成立せず,

$$P_{L^2((Y;t-p,t-1),(a;t-q,t-1))}Y_t = P_{L^2(Y;1,t-1)}Y_t$$

がすべての t, $t \geqq \max(p, q)+1$, に対して成立するとき X_t を次数 (p, d, q) の Autoregressive Integrated Moving

Average過程という。

このとき

$$P_{L^2}((Y;t-p,t-1),(a;t-q,t-1))Y_t = -\sum_{k=1}^{p}\phi_k Y_{t-k} + \sum_{l=1}^{q}\theta_l a_{t-l}$$

とあらわすことができる。このことは

$$\sum_{k=0}^{p} \phi_k Y_{t-k} = \sum_{l=0}^{q} \theta_l a_{t-l} \quad (\phi_0=1) \qquad \cdots\cdots(6)$$

を意味する。$\{\phi_l\}$, $\{\theta_l\}$ はともに無関係な定数である。

上述の考察を周波数領域での考察との関連で調べてみる。同伴多項式を $\Phi(Z)=\sum_{k=0}^{p} \phi_k Z^{p-k} = \prod_{k=1}^{p} (Z-\alpha_k)$, $\Theta(Z)=\sum_{l=0}^{q}\theta_l Z^{q-l} = \prod_{l=1}^{q} (Z-\beta_l)$ とおく。$\{\alpha_k\}$, $\{\beta_l\}$ は複素数値である。オペレータ D を $DY_t=Y_{t-1}$, $D^jY_t = D(D^{j-1})Y_t$ $(j\geq 2)$ で定義する。このとき(6)は

$$\prod_{k=1}^{p}(1-\alpha_k D)Y_t = \prod_{l=1}^{q} (1-\beta_l D)a_t$$

となる。今 $\{\alpha_k\}$, $\{\beta_l\}$ のなかに共通の値があるものとする。簡単のためにそれを $\alpha_1=\beta_1$ とする。このとき

$$(1-\alpha_1 D)\left\{\prod_{k=2}^{p}(1-\alpha_k D)Y_t - \prod_{l=2}^{q}(1-\beta_l D)a_t\right\}=0$$

となる。一般に U_t を弱定常過程とし $(1-\alpha_1 D)U_t=0$ とする。このときもし $|\alpha_1|\neq 1$ であれば $U_t=0$ (L^2 の意味で)となる。従ってもし $|\alpha_1|\neq 1$ を仮定すれば $\prod_{k=2}^{p}(1-\alpha_k D)Y_t = \prod_{l=2}^{p}(1-\beta_l D)a_t$ となりこれを用いて

$$P_{L^2}(Y;1,t-1)Y_t = -\left\{\prod_{k=2}^{q}(1-\alpha_k D)-1\right\}Y_t + \left\{\prod_{l=2}^{q}(1-\beta_l D)-1\right\}a_t$$

とあらわすことができて, X_t は ARIMA $(p-1, d, q-1)$ 過程

であることを示す。$|\alpha_1|=1$ のときは $\sum_{k=0}^{\infty}|R_k|=+\infty$ となりここでの前提条件をみたさない。

3. ARIMA(p, d, q) 過程におけるパラメータの最尤推定量について

X_t を Gauss 過程であるものとし，標本として $\{Y_t; 1\leqq t\leqq T\}$ が得られたものとする。ここで X_t は ARIMA(p, d, q) 過程であると仮定する。このとき $t\geqq \max(p, q)+1$ に対して（以下において簡単のため $p\geqq q$ と仮定）

$$\sum_{k=0}^{p}\phi_k Y_{t-k}=\sum_{\ell=0}^{q}\theta_\ell a_{t-\ell}\ ,\quad (\phi_0=1)$$

とあらわされているとすれば

$$\theta_0 a_t=\sum_{k=0}^{p}\phi_k Y_{t-k}-\sum_{\ell=1}^{q}\theta_\ell a_{t-\ell}$$

である。従って $\{a_t; t\geqq p+1\}$ を用いての尤度関数は

$$l(\boldsymbol{a};\boldsymbol{\phi},\boldsymbol{\theta})=\frac{1}{(2\pi)^{\frac{T-p}{2}}\theta_0^{T-p}}e^{-\frac{1}{2\theta_0^2}\sum_{t=p+1}^{T}\left\{\sum_{k=0}^{p}\phi_k Y_{t-k}-\sum_{\ell=1}^{q}\theta_\ell a_{t-\ell}\right\}^2}$$

となる。ここで $\boldsymbol{a}=(a_{p+1}, a_{p+2},\cdots, a_T)'$，$\boldsymbol{\phi}=(\phi_1,\phi_2,\cdots,\phi_p)'$，$\boldsymbol{\theta}=(\theta_0,\theta_1,\cdots,\theta_q)'$ である。[2] による $\boldsymbol{\phi}$, $\boldsymbol{\theta}$ の最尤推定量の求め方は $\boldsymbol{\phi}$, $\boldsymbol{\theta}$ に順次値を代入して $l(\boldsymbol{a};\boldsymbol{\phi},\boldsymbol{\theta})$ の値を求め，$l(\boldsymbol{a};\boldsymbol{\phi},\boldsymbol{\theta})$ を最大にする $\boldsymbol{\phi}$, $\boldsymbol{\theta}$ を求めるという方法である。ここではこの方法を検討することにする。この方法を用いるときに問題になるのは初期値 a_{p+1-q}, a_{p+2-q},

……, a_{p-1}, a_p の選び方である。〔2〕においてはこれらの値を0とおく方法などが提案されている。しかし上で述べたような定義と表現を用いればつぎのように選ぶことも可能である。

$\{Y_t;\ 1 \leqq t \leqq T\}$ を用いて自己共分散 $\{R_k;\ 0 \leqq k \leqq H\}$ の1つの一致推定量 $\{\hat{R}_k;\ 0 \leqq k \leqq H\}$ を求める。Hはpより大であるTに無関係な正整数である。$1 \leqq j \leqq p$ について

$$a_j = (Y_j - c_1 Y_{j-1} - \cdots - c_{j-1} Y_1) / \| Y_j - c_1 Y_{j-1} - \cdots - c_{j-1} Y_1 \|$$

と表わすことができ, c_1, c_2, ---, c_{j-1} は $\{R_k;\ 0 \leqq k \leqq p\}$ の関数である。この $\{R_k;\ 0 \leqq k \leqq p\}$ を $\{\hat{R}_k;\ 0 \leqq k \leqq H\}$ でおきかえてできた c_1, c_2, $\cdots$, c_{j-1} を $\hat{c}_1$, $\hat{c}_2$, $\cdots$, $\hat{c}_{j-1}$ とする。$\hat{c}_1$, $\hat{c}_2$, ---, $\hat{c}_{j-1}$ を用いて得られる a_j を $\hat{a}_j$ とする。この $\hat{a}_{p+1-q}$, $\hat{a}_{p+2-q}$, $\cdots$, $\hat{a}_p$ を初期値として用いる。

このような初期値の選び方がどのような効果を持つかは今後の検討課題である。Tが十分大きいときには初期値の選び方は Φ, Θ の最尤推定量の構成にあまり影響を持たないことが予想される。Tがあまり大きくないときに問題になるであろう。

References

[1] Akaike, H., (1974), Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes, Ann. Inst. Statist. Math., 26, 363-387.

[2] Box, G. E. P. and Jenkins, J. M., (1970), Time Series Analysis, Forecasting and Control, Holden Day, San Francisco.

[3] Hannan, E. J., (1969), The identification of vector mixed autoregressive-moving average systems, Biometrika, 223-225.

[4] Huzii, M., (1980), On an ARIMA model and estimation of parameters for prediction, Analysing Time Series (ed. O. D. Anderson), North-Holland, Amsterdam, 169-176.

[5] Stigum, B. P., (1975), Asymptotic properties of autoregressive integrated moving average process, Stochastic Processes Appl., 3, 315-344.