数理解析研究所講究録
第312巻 1977年 91-108

On an ARIMA process

東工大 理　藤井光昭
矢島美寛

§1. 序

§2, 3では、Box-Jenkins[4]によって定義された、ARIMA processesの表現および予測について論じる。近年ARIMA modelは応用面において、有効な手段として用いられているが、理論的に明確さを欠く点が見受けられる。そこで、最も簡単な、ARIMA(0,1,0) process

$$\nabla x_t \equiv x_t - x_{t-1} = a_t$$

を例にとる。ただし $E a_t = 0$, $E|a_t|^2 = \sigma_a^2 < \infty$, $E a_t \overline{a_s} = 0$, $t \neq s$.
Box-Jenkinsは、この条件を満たすprocessとして、$x_{-\infty} = 0$ の仮定のもとで

$$x_t = \sum_{j=0}^{\infty} a_{t-j}$$

のようなものを想定している。しかし右辺の分散は存在せず、等式を $l.i.m$ の意味で解釈することは不可能である。一方 a_t を $i.i.d$ などとして別の収束（概収束，確率収束）で理解

することも困難である。たとえ何らかの解釈が可能であるとしても、分散の存在しないprocessであるから、予測の際には、平均二乗誤差とは別の基準を設けて、予測量の良し悪しを論じなければならない。

そこで§2では、ARIMA(p,d,q)processesの定義を、正分散有限のprocessesでd回差分を取った時、AR-MA(p,q)processesとなるものとし、この定義のもとでのprocessesの一般的表現を求める。

§3では§2において求めた表現をもつprocessesについては、従来のBox-Jenkinsの予測量は、平均二乗誤差を最小にするという基準のもとでは最良でないことを示す。

§4以下では、AR processesとMA processesのスペクトル密度関数の特徴づけを試みる。

§2. ARIMA processes の表現について

まず、いくつかの仮定，記号および定義を導入する。

(A1) $\qquad E x_t = 0,\ E|x_t|^2 < \infty,\ \forall t.$

(N1) $\qquad \mathfrak{H}(x) = \mathfrak{S}\{x_t;\ t=0,\pm 1,\cdots\},$

$\qquad \mathfrak{H}(x,t) = \mathfrak{S}\{x_s;\ s \leq t\}.$

ただし$\mathfrak{S}$は{ }の中の変数によって生成されるHilbert空間を表わす。

(N2) $P_{\mathfrak{D}(x,t)}$; Projection on $\mathfrak{D}(x,t)$.

(D1) (A1) を満たす process x_t の innovation (定常過程においては white noise) を、

$$a_t = x_t - P_{\mathfrak{D}(x,t-1)} x_t$$

で定義する。(cf. Cramer [5]).

(D2) "Backward Shift operator", B "Difference operator," ∇ を

$$B x_t = x_{t-1},$$

$$\nabla x_t = (1-B) x_t = x_t - x_{t-1}, \quad \forall t$$

で定義する。

(A1) の条件のもとで次の定理を得る。

定理 1

x_t が ARIMA process すなわち

$$(1-\phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)\nabla^d x_t = (1-\theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) a_t$$

ただし (i) $a_t = \int_{-\pi}^{\pi} e^{it\lambda} dZ(\lambda)$, $Z(\lambda)$; orthogonal random measure

$$E|dZ(\lambda)|^2 = \frac{\sigma_a^2}{2\pi} d\lambda$$

(ii) $\phi(B) \equiv 1-\phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \; (\phi_p \neq 0)$

$\theta(B) \equiv 1-\theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \quad (\theta_q \neq 0)$

$\phi(Z)=0$, $\theta(Z)=0$, の根はすべて単位円の外

と表現されるための必要十分条件は、x_t が次の表現をもつことである。

$$x_t = \int_{-\pi}^{\pi} e^{it\lambda} \cdot C(t,\lambda) dZ(\lambda) + \sum_{j=0}^{d-1} t^j \cdot \int_{-\pi}^{\pi} g_j(\lambda) dZ(\lambda) + \sum_{j=0}^{d-1} t^j \cdot \xi_j \quad (2.1)$$

ただし (iii) $dZ(\lambda)=\dfrac{\theta(e^{-i\lambda})}{\phi(e^{-i\lambda})}\cdot d\tilde{Z}(\lambda)$,

(iv) $$C(t,\lambda)=\begin{cases}\displaystyle\sum_{j=0}^{t-1}e^{-ij\lambda}\Big(\sum_{K_{d-1}=1}^{j+1}\cdots\sum_{K_2=1}^{K_3}\sum_{K_1=1}^{K_2}\cdot 1\Big) & t>0\\ \displaystyle(-1)^{d}\sum_{j=d}^{|t|}e^{-ij\lambda}\Big(\sum_{K_{d-1}=1}^{j-(d-1)}\cdots\sum_{K_2=1}^{K_3}\sum_{K_1=1}^{K_2}\cdot 1\Big) & t\leq 0\end{cases}$$

ここで $d=1$ の時、( )内は 1 とする。また

$t\leq 0$ において $|t|<d$ のときは $C(t,\lambda)=0$ とする

(v) $g_j(\lambda)\in L^2(F)$ ここで $dF(\lambda)=E|dZ(\lambda)|^2$ $\forall j$,

(vi) $\xi_j\in L^2(\Omega,P)$,

$\xi_j\perp\mathfrak{H}(\nabla^d X)=\left\{\int_{-\pi}^{\pi}h(\lambda)dZ(\lambda)\,;\ \int_{-\pi}^{\pi}|h(\lambda)|^2dF(\lambda)<\infty\right\}$.

〔証明の概略〕

（十分性） 明らか

（必要性） もし異なる2つの processes が

$$\nabla^d X_t=\int_{-\pi}^{\pi}e^{-it\lambda}dZ(\lambda),$$

$$\nabla^d Y_t=\int_{-\pi}^{\pi}e^{-it\lambda}dZ(\lambda),$$

を満たすとする。すると

$$\nabla^d(X_t-Y_t)\equiv 0$$

したがって一般の ARIMA process は

$\nabla^d Y_t=\int_{-\pi}^{\pi}e^{-it\lambda}dZ(\lambda)$ の特殊解と

$\nabla^d W_t\equiv 0$ の一般解の

和 $X_t=Y_t+W_t$ となる。W_t は確率変数を係数とする t の $(d-1)$ 次の多項式となる。一方特殊解を求める際には、Y_t を Oscillatory process (Mandrekar〔7〕, Priestley〔8〕)

$$y_t=\int_{-\pi}^{\pi}e^{it\lambda}\cdot C(t,\lambda)dZ(\lambda),$$

ただし $C(t,\lambda)\in L^2(F)$, $^{\forall}t$

と仮定し、$C(t,\lambda)$の関数形を求める問題に帰着させる。□

ら(2.1)において

$$y_t=\int_{-\pi}^{\pi}e^{it\lambda}\cdot C(t,\lambda)dZ(\lambda)$$

$$w_t=\sum_{j=0}^{d-1}t^j\cdot\left(\int_{-\pi}^{\pi}g_j(\lambda)dZ(\lambda)+\xi_j\right)$$

であり、w_tは"stochastic trend"を表わす部分と言える。(ただし $\int_{-\pi}^{\pi}g_j(\lambda)dZ(\lambda)+\xi_j\in\bigcap_t\mathfrak{H}(x,t)$とは一般には言えない。)ここで見易くするため、$y_t$の部分に着目し、time-domainで表現した簡単な例を与える。

<u>ex.1</u> $\nabla x_t=a_t$ $(0,1,0)$

$$y_t=\begin{cases}\sum_{u=1}^{t}a_u & t>0\\ 0 & t=0\\ -\sum_{u=t+1}^{0}a_u & t<0\end{cases}$$

<u>ex.2</u> $(1-\phi B)\nabla x_t=a_t$ $(1,1,0)$

$$y_t=\begin{cases}\sum_{u=1}^{t}\sum_{j=-\infty}^{0}\phi^j a_{u-j} & t>0\\ 0 & t=0\\ -\sum_{u=t+1}^{0}\sum_{j=-\infty}^{0}\phi^j a_{u-j}\end{cases}$$

<u>Remark 1.</u>

(2.1)において $t>0$ の場合を例にとると、

$$y_t=\begin{cases}\int_{-\pi}^{\pi}e^{it\lambda}\cdot\dfrac{(1-e^{-it\lambda})}{(1-e^{-i\lambda})}\cdot\dfrac{(1-\theta_1e^{-i\lambda}-\cdots-\theta_q e^{-iq\lambda})}{(1-\phi_1e^{-i\lambda}-\cdots-\phi_p e^{-ip\lambda})}\cdot d\tilde{Z}(\lambda) & d=1\\ \int_{-\pi}^{\pi}e^{it\lambda}\cdot\dfrac{(1-(t+1)e^{-it\lambda}+te^{-i(t+1)\lambda})}{(1-e^{-i\lambda})^2}\cdot\dfrac{(1-\theta_1e^{-i\lambda}-\cdots-\theta_q e^{-iq\lambda})}{(1-\phi_1e^{-i\lambda}-\cdots-\phi_p e^{-ip\lambda})}d\tilde{Z}(\lambda) & d=2\end{cases}$$

つまりARIMA processを強いて、$(1-e^{-i\lambda})^d$を分母に持つスペクトラル表現でしめせば、分子には$(1-e^{-i\lambda})^d$を因数とするtime-dependentな関数が現われることになる。

§3. ARIMA processesの予測について

ARIMA processesは、time-domainにおいて

$$x_{t+l}=\varphi_1 x_{t+l-1}+\cdots+\varphi_{p+d}\,x_{t+l-p-d}+a_{t+l}-\theta_1 a_{t+l-1}-\cdots-\theta_q\,a_{t+l-q},$$

あるいは

$$x_{t+l}=\sum_{j=1}^{\infty}\pi_j\,x_{t+l-j}+a_{t+l}\quad\left(\sum_{j=1}^{\infty}\pi_j=1\right)$$

となる。今$[x_{t+j}]$を$\underset{t}{E}[x_{t+j}]$，すなわち時点tからのwide-sense conditional expectationとする。Box-Jenkinsでは、現時点tからl-step先の値x_{t+l}に対する最良の予測量$\hat{x}_t(l)$は

$$\hat{x}_{t+l}=[x_{t+l}]=\varphi_1[x_{t+l-1}]+\cdots+\varphi_{p+d}[x_{t+l-p-d}]+[a_{t+l}]-\cdots-\theta_q[a_{t+l-q}]\tag{3.1}$$

あるいは

$$\hat{x}_{t+l}=[x_{t+l}]=\sum_{j=1}^{\infty}\pi_j[x_{t+l-j}]+[a_{t+l}]$$

で与えられている。ここで$[x_{t+j}]$は次の規則に従うとされている。

$$[x_{t-j}]=\underset{t}{E}[x_{t-j}]=x_{t-j}\qquad j=0,1,2,\cdots$$

$$[x_{t+j}]=\underset{t}{E}[x_{t+j}]=\hat{x}_t(j)\qquad j=1,2,\cdots$$

$$[a_{t-j}]=E_t[a_{t-j}]=a_{t-j}\qquad j=0,1,2,\cdots$$

$$[a_{t+j}]=\underset{t}{E}[a_{t+j}]=0\qquad j=1,2,\cdots\qquad\text{(cf. [4] Ch.5)}$$

しかし定理1で表現を求めたprocessについては、必ずしも

$$\{a_{t+j}\} = E_t[a_{t+j}] = 0 \qquad j=1,2,\cdots\cdots$$

は成立しない。wide-sense conditional expectationとは、Hilbert 空間におけるProjection を意味する。したがって射影させる部分空間を明確にする必要があるが、予測理論において通常その空間は、過去および現在の変数によって生成される$\mathfrak{H}(x,t)$であろう。ところで、その場合には$P_{\mathfrak{H}(x,t)}a_{t+j}=0\ (j>0)$とは必ずしも言えない。何故なら$a_t,\ t=0,\pm1,\cdots$は$\nabla^d x_t$の innovation process であって、x_tのそれではないからである。換言すれば$\mathfrak{H}(\nabla^d x,t)\subset\mathfrak{H}(x,t)$であるから、$P_{\mathfrak{H}(\nabla^d x,t)}a_{t+j}=0\ (j>0)$は成立しても、$P_{\mathfrak{H}(x,t)}a_{t+j}=0$とは断定できないわけである。ここで現時点を$t=0$とすれば、$\ell$-step 先の値$x_\ell$の最良の予測量は

$$\hat{x}_0(\ell)=P_{\mathfrak{H}(x,0)}x_\ell=\varphi_1\cdot P_{\mathfrak{H}(x,0)}x_{\ell-1}+\cdots+\varphi_{p+d}\cdot P_{\mathfrak{H}(x,0)}x_{\ell-p-d}$$
$$+P_{\mathfrak{H}(x,0)}a_\ell-\theta_1\cdot P_{\mathfrak{H}(x,0)}a_{\ell-1}-\cdots\cdots-\theta_q\cdot P_{\mathfrak{H}(x,0)}a_{\ell-q} \quad (3.2)$$

となる。その際$P_{\mathfrak{H}(x,0)}a_j$は以下のような修正を要する。

定理 2.

(3.2)において
$$P_{\mathfrak{H}(x,0)}a_j=a_j \qquad j\le 0$$
$$P_{\mathfrak{H}(x,0)}a_j=P_{\mathfrak{H}\{x_0',x_{-1}',\cdots\cdots,x_{-d+1}'\}}a_j \qquad j>0$$

ここで$\{x_0',x_{-1}',\cdots\cdots,x_{-d+1}'\}$は変数列$\{\cdots\cdots,a_{-t},\cdots,a_{-2},a_{-1},a_0,x_{-d+1},x_{-d+2},\cdots,x_{-1},x_0\}$を直交化した列$\{\cdots,a_{-t},\cdots,a_{-2},a_{-1},a_0,x_{-d+1}',\cdots,x_{-1}',x_0'\}$の最後の$d$個の項とする。

［証明の概略］

(2.1)において $\eta_j = \int_{-\pi}^{\pi} g_j(\lambda)dZ(\lambda) + \xi_j$, $j=0,\cdots,d-1$ とおく。

すると一般の ARIMA process は

$$X_t = \begin{cases} \sum_{u=-\infty}^{0} \psi_{t-u}\, a_{t-u} + \sum_{j=0}^{d-1} t^j \cdot \eta_j, & t>0, \\ \quad 0 \quad + \sum_{j=0}^{t-1} t^j \cdot \eta_j, & t=0,-1,\cdots,-d+1, \\ \sum_{u=-\infty}^{0} \psi_{t-u}\, a_{-u} + \sum_{j=0}^{t-1} t^j \cdot \eta_j, & t\le -d \,. \end{cases}$$

ここで

$$\begin{aligned} \mathfrak{L}(x,0) &= \mathfrak{S}\{\cdots, a_{-t}, \cdots, a_{-1}, a_0, x_{-d+1}, \cdots, x_{-1}, x_0\} \\ &= \mathfrak{L}(a,0) \oplus \mathfrak{S}\{x'_{-d+1}, \cdots, x'_{-1}, x'_0\} \end{aligned}$$

を言う。

(i) (⊃) $\quad \mathfrak{L}(x,0) \supset \mathfrak{L}(\nabla^d x, 0) \supset \{\cdots, a_{-t}, \cdots, a_{-1}, a_0\}$

(ii) (⊂) $\quad \mathfrak{L}(x,0) \subset \mathfrak{S}\{\cdots, a_{-t}, \cdots, a_{-1}, a_0, \eta_{d-1}, \cdots, \eta_1, \eta_0\}$

$\qquad\qquad \mathfrak{S}\{\eta_{d-1}, \cdots, \eta_1, \eta_0\} = \mathfrak{S}\{x_{-d+1}, \cdots, x_{-1}, x_0\}$

を各々用いる。

$\therefore \quad P_{\mathfrak{L}(x,0)}\, a_j = P_{\mathfrak{L}(a,0)}\, a_j + P_{\mathfrak{S}\{x'_{-d+1}, \cdots, x'_{-1}, x_0\}}\, a_j \quad \square$

定理2の証明から、Box-Jenkins の方法では stochastic trend 部分が持つ情報を無視して予測量を構成していることがわかる。したがって stochastic trend を持たないような（少なくとも、任意の j で、$\int_{-\pi}^{\pi} g_j(\lambda)dZ(\lambda) \in \mathfrak{L}(a,0)$）であるような process では、彼等の予測量は最良となるが、そのような仮定はモデルをかなり制限することになろう。

ここで今後の統計的な問題について、二,三考察する。まず $P_{\mathfrak{G}\{x_{-d+1},\dots,x_{-1},x_0\}}a_d\ (d>0)$ の項を評価することに困難を伴う。

今 $d=1$ とする。すると

$$x_0' = x_0 - \sigma_a^{-2}\cdot\sum_{d=-\infty}^{0} E(x_0\bar{a}_d)a_d$$

$$P_{\mathfrak{G}\{x_0'\}} = \frac{E(a_d\bar{x}_0')}{E|x_0'|^2}\cdot x_0'$$

$$= \frac{E(a_d\bar{x}_0)}{E|x_0 - \sigma_a^{-2}\cdot\sum_{d=-\infty}^{0}(x_0\bar{a}_d)a_d|^2}\cdot x_0' \qquad (3\cdot 3)$$

(3.3)において、x_0' および $E|x_0'|^2$ は $\{x_t, t\le 0\}$, $\{R(t,s); t\le 0, s\le 0\}$. ($R(t,s)\equiv E x_t\bar{x}_s$)によって表現可能であるが $E(a_d\bar{x}_0)$ は、時点 d までの covariance に依存する。したがって時点 $t=0$ からこの項を評価するには、process に関する事前情報が必要となる。一方特定の t,s に対して $R(t,s)$ は定常の場合と異なり $R(t,s)\neq R(t-s)$ であるから、ここでも process の構造に仮定をおかなければ、$E(x_0\bar{a}_d)\ (d<0)$, $E|x_0'|^2$ の良い推定量を得ることはできない。

また根本的な問題として、実際には有限個の観測値を用いて予測をしなければならない。この時、定常過程においても AR-MA モデルと infinite AR モデルのいずれを選択すべきかの問題が生じる。その議論は非定常な過程についても考慮する必要がある。今 $x_0, x_{-1}, \dots, x_{-m+1}$ の値が与えられたとする。x_1 の値を予測するひとつの方法は、

$$E|x_1 - (\alpha_0 x_0 + \alpha_1 x_{-1} + \cdots + \alpha_{m-1}x_{-m+1})|^2$$

を最小にする$(\alpha_0,\alpha_1,\cdots,\alpha_{m-1})$を求めるものである。これらの係数は、定常過程のYule-Walker equationに対応する、

$$R(1,0)=\alpha_0 R(0,0)+\alpha_1 R(-1,0)+\cdots\cdots+\alpha_{m-1}R(-m+1,0)$$
$$R(1,-1)=\alpha_0 R(0,-1)+\alpha_1 R(-1,-1)+\cdots\cdots+\alpha_{m-1}R(-m+1,-1)$$
$$\vdots$$
$$R(1,-m+1)=\alpha_0 R(0,-m+1)+\alpha_1 R(-1,-m+1)+\cdots\cdots+\alpha_{m-1}R(-m+1,-m+1)$$

を解いて求めることができる。しかしこの方程式に推定量$\hat{R}(t,s)$を代入して$\hat{\alpha}_i$を導く時、$\hat{R}(t,s)$の推定に関する障害は、上に述べたとおりである。一方(3.1)式を用いる場合、過去のinnovation $a_j\,(j\le 0)$を求めることさえ難しい。Box-Jenkinsでは$a_j=0$として、予測値を計算している。しかしこの方法を用いるとすれば、$P_G\{x_{-d+1},\cdots,x_{-1},x_0\}a_j\,(j>0)$を評価する意味はない。そこでまず$a_j\,(j\le 0)$を求めることが必要だが、そのためにはinfinite AR表現に、inversionし、

$$x_t-\sum_{j=1}^{\infty}\pi_j x_{t-j}=a_t$$

を用いる。今、データは$x_0,x_{-1},\cdots,x_{-m+1}$だけであるから、$\pi_j$の0への収束が非常に遅い時、誤差の分散(かりに$t=0$とすれば、$-\sum_{j=m}^{\infty}\pi_j x_{-j}$の分散)は大きくなる。これに比して$P_G\{x_{-d+1},\cdots,x_{-1},x_0\}a_j\,(j>0)$の分散が小さければ、この項を評価する意義はうすれる。したがって応用上では、この二つの量の比較をまず考えなければならない。

§4. AR process と MA process のスペクトル密度関数のある表現

ARMA モデルのあてはめをおこなおうとするとき $(p+q)$ 個の係数を用いてモデルを表現するにしても，AR$(p+q)$，MA$(p+q)$，ARMA(p, q) というモデルをあてはめることが可能であり，どれを用いるのが良いかはその目的や問題とする定常過程の構造によって異なるであろう。ここではスペクトル密度の推定という目的に限定することにし，上に述べた問題を論じるための一つの試みをおこなう。本来この問題は推定論上の観点も含めて論ずべきであると考えられるが，ここではその第一段階として AR モデルと MA モデルを分離して考え，それぞれから得られるスペクトル密度関数の特徴を論じることにする。

以下において x_t は $Ex_t = 0$，$Ex_{t+h}\overline{x_t} = R_h$ であるような弱定常過程とし，前節までと同様に

$$x_t = \int_{-\pi}^{\pi} e^{it\lambda} dZ(\lambda)$$

とあらわされているものとする。

いま，x_t が p 次の AR process であると仮定すれば white noise a_t を用いて

$$\sum_{j=0}^{p} (-\phi_j) x_{t-j} = a_t \qquad (\phi_0 = -1)$$

のように表わされ，x_t のスペクトル密度関数 $f(\lambda)$ は

$$f(\lambda) = \sigma_a^2 / (2\pi |\phi(e^{-i\lambda})|^2)$$

であるが $\phi(B^{-1}) = 0$ の根を $-\alpha_j$ ($1 \leqq j \leqq p$, $|\alpha_j| < 1$) とおくと

$$\phi(e^{-i\lambda}) = \prod_{j=1}^{p} (1+\alpha_j e^{-i\lambda})$$

とあらわすことができる。これから $f(\lambda)$ の特徴や推定を論じるために都合のよい $f(\lambda)$ の表現を求めることにする。

任意な弱定常過程 y_t ($E y_t = 0$) に対し operator $D(\alpha)$ を

$$D(\alpha) y_t = (1+\alpha B) y_t = y_t + \alpha y_{t-1}$$

と定義する。このとき p 次の AR process x_t は

$$D(\alpha_1)D(\alpha_2)\cdots D(\alpha_p) x_t = a_t$$

とあらわすことができる。いま

$$y_t^{(j)} = D(\alpha_{j+1}) D(\alpha_{j+2}) \cdots D(\alpha_p) x_t \qquad (0 \leqq j \leqq p-1)$$

とおくと、$y_t^{(j)}$ も平均値0である弱定常過程となる。$(\sigma^{(j)})^2 = Var(y_t^{(j)})$ とおくと

$$(\sigma^{(j)})^2 = (\sigma^{(j+1)})^2 (1 + |\alpha_{j+1}|^2 + 2 \mathrm{Real}\, \bar{\alpha}_{j+1} \rho^{(j+1)})$$

となる。ただし $\rho^{(j+1)} = E y_t^{(j+1)} \bar{y}_{t-1}^{(j+1)} / E|y_t^{(j+1)}|^2$, $\rho^{(p)} = R_1/R_0$, $\sigma^{(0)} = \sigma_a$, $\rho^{(1)} = -\alpha_1$, $(\sigma^{(p)})^2 = R_0$ である。

この条件のもとで $y_t^{(j)}$ は purely non-deterministic になり、$|\rho^{(j)}| < 1$ ($1 \leqq j \leqq p$) となる。また $\rho^{(j)}$ は $\alpha_1, \alpha_2, \cdots, \alpha_j$ のみの関数となる。

以上により次の結果が得られる。

定理3. p 次の AR process x_t のスペクトル密度関数 $f(\lambda)$

は

$$f(\lambda)=R_0\prod_{j=1}^{p}\left(1+|\alpha_j|^2+2\,\mathrm{Real}\,\bar{\alpha}_j\,\rho^{(j)}\right)\Big/\prod_{j=1}^{p}|1+\alpha_j e^{-i\lambda}|^2$$

とあらわすことができる。

上の表現はモデルのあてはめのためであり，x_tが与えられてその分散R_0は一定にしておき，pを変化させるたびにα_jおよびwhite noiseの分散σ_a^2をそれにあわせてかえていくというものである。

つぎにx_tがq次のMA processであるとする。このとき

$$x_t=\sum_{j=0}^{q}(-\theta_j)\,a_{t-j}\qquad(\theta_0=-1)$$

のようにあらわされ，x_tのスペクトル密度関数$f(\lambda)$は

$$f(\lambda)=(\sigma_a^2/2\pi)\,|\,\theta(e^{-i\lambda})|^2$$

であるが，$\theta(B^{-1})=0$ の根を $-\beta_j$ $(1\leqq j\leqq q,\ |\beta_j|<1)$

とおくと

$$\theta(e^{-i\lambda})=\prod_{j=1}^{q}(1+\beta_j e^{-i\lambda})$$

とあらわすことができる。

このとき

$$x_t=D(\beta_1)D(\beta_2)\cdots D(\beta_q)\,a_t$$

とあらわされる。いま

$$\eta_t^{(j)}=D(\beta_{j+1})D(\beta_{j+2})\cdots D(\beta_q)\,a_t\qquad(0\leqq j\leqq q-1)$$

とおくと，$\eta_t^{(j)}$ は平均値 0 である弱定常過程になる。$(\sigma_\eta^{(j)})^2 = \mathrm{Var}(\eta_t^{(j)})$ とおくと，

$$(\sigma_\eta^{(j)})^2 = (\sigma_\eta^{(j+1)})^2 (1 + |\beta_{j+1}|^2 + 2\,\mathrm{Real}\,\overline{\beta_{j+1}}\,\gamma^{(j+1)})$$

となる。ただし $(\sigma_\eta^{(0)})^2 = R_0$，$\eta_t^{(g)} = a_t$，$(\sigma_\eta^{(g)})^2 = \sigma_a^2$，$\gamma^{(j+1)} = E\,\eta_t^{(j+1)}\,\overline{\eta_{t-1}^{(j+1)}} \,/\, E|\eta_t^{(j+1)}|^2$ である。

上の条件のもとで $\eta_t^{(j)}$ は purely nondeterministic process となり，$|\gamma^{(j)}| < 1$ $(1 \leqq j \leqq g)$ となる。また $\gamma^{(j)}$ は $\beta_{j+1}, \beta_{j+2}, \cdots, \beta_g$ のみの関数である。

以上により次の結果が得られる。

<u>定理4</u>　g 次の MA process x_t のスペクトル密度関数 $f(\lambda)$ は

$$f(\lambda) = (R_0/2\pi) \prod_{j=1}^{g} |1 + \beta_j e^{-i\lambda}|^2 \Big/ \prod_{j=1}^{g} (1 + |\beta_j|^2 + 2\,\mathrm{Real}\,\overline{\beta_j}\,\gamma^{(j)})$$

とあらわすことができる。

この表現も x_t が与えられ R_0 が与えられたという形である。いま a_t を white noise で $\sigma_a^2 > 0$ とする。このとき任意な β_j に対して $(\sigma_\eta^{(j)})^2 = E|a_t + (\Sigma\beta_j)a_{t-1} + \cdots|^2 \geqq \sigma_a^2 > 0$ となり，$(\sigma_\eta^{(j)})^2 > 0$ $(0 \leqq j \leqq g-1)$ が成り立つ。

§5. <u>スペクトル密度関数の特徴</u>

§4 において求めた AR process または MA process のスペクトル密度関数の分解した形での表現を用いてその特徴を調べることにする。この分解により特徴は各要素ごとに調べることにする。

(i) α_j または β_j が実数のとき

AR process に対して

$$f_j^A(\lambda) = \left(1 + |\alpha_j|^2 + 2\,\mathrm{Real}\,\bar{\alpha}_j\,\rho^{(j)}\right) \Big/ \left|1 + \alpha_j e^{-i\lambda}\right|^2$$

とおく。$\alpha_j < 0$ のとき $f_j^A(\lambda)$ は単調減少。$\rho^{(1)} = -\alpha_1$ であるから $f_1^A(\lambda) = (1 - \alpha_1^2) / (1 + \alpha_1^2 + 2\alpha_1 \cos\lambda)$ となる。従って

$$\lim_{\alpha_1 \to -1} f_1^A(\lambda) = \begin{cases} +\infty & ;\ \lambda = 0 \\ 0 & ;\ \lambda \neq 0 \end{cases}$$

となる。$\alpha_j > -1$ のとき $f_j^A(0)$ は有界. $f_j^A(0) \neq 0$. 故に α_1 を -1 に近く選ぶことにより $f(0)$ はいくらでも大きな値をとることができ δ 関数のようなスペクトル密度が達成可能である。 $\alpha_j > 0$ のときには $f_j^A(\lambda)$ は単調増加で $\lambda = \pi$ において上と同じ議論が成り立つ。

MA process に対しては $\sigma_a^2 \leqq R_0$ であるから $|\beta_j| \leqq 1$ である限りどのように β_j を選んでも $f(\lambda)$ は有界である。いま

$$f_j^M(\lambda) = |1 + \beta_j e^{-i\lambda}|^2$$

とおくと, $f_j^M(\lambda)$ は $\beta_j > 0$ のとき単調減少で $\beta_j < 0$ のとき単調増加。$\beta_j \to 1$ (または -1) のとき $\lambda = \pi$ (または 0) で 0 に

近い値をとる。

(ii) α_j または β_j が複素数のとき

AR process について考える。α_j が $\phi(B^{-1})=0$ の根なら $\overline{\alpha_j}$ も根である。いま $\alpha_{j'}=\overline{\alpha_j}$ とし，

$$g_j^A(\lambda)=f_j^A(\lambda)\,f_{j'}^A(\lambda)$$

$$=\frac{(1+|\alpha_j|^2+2\,\mathrm{Real}\,\overline{\alpha_j}\,\rho^{(j)})(1+|\alpha_j|^2+2\,\mathrm{Real}\,\alpha_j\,\rho^{(j)})}{|1+\alpha_j e^{-i\lambda}|^2\,|1+\overline{\alpha_j}\,e^{-i\lambda}|^2}$$

とおく。このとき $g_j^A(\lambda)$ の分母は $S_1^2(\cos\lambda+S_2)^2+S_3$ という形に書ける。ただし $S_1=2|\alpha_j|$，$S_2=(1+|\alpha_j|^2)u_j/(2|\alpha_j|^2)$，$S_3=v_j^2(1-|\alpha_j|^2)^2/|\alpha_j|^2$ である。ここで $\alpha_j=u_j+iv_j$ (u_j, v_j は実数) とおいた。類似の表現は [3] にも求められている。$g_j^A(\lambda)$ の最大値を与える λ の値 λ_0 が u_j と v_j の値により $0<\lambda_0<\pi$ となる。そして

$$\lim_{|\alpha_1|\to 1} g_1^A(\lambda_0)=1/v_1^2,\quad \lim_{|\alpha_1|\to 1} g_1^A(\lambda)=0\quad(\lambda\neq\lambda_0)$$

となり $v_1\to 0$ とすることにより $g_1^A(\lambda_0)$ はいくらでも大きな値をとり $\lambda=\lambda_0$ でのラインスペクトルに近いものを達成することが可能である。

$$\lim_{|\alpha_1|\to 0} g_1^A(\lambda)=1$$

で white noise に対応する。

MA process についても同様に

$$g_j^M(\lambda)=|1+\beta_j e^{-i\lambda}|^2\,|1+\overline{\beta_j}\,e^{-i\lambda}|^2$$

とおくと

$$g_j^M(\lambda) = T_1^2(\cos\lambda + T_2)^2 + T_3$$

とあらわされる。ここで T_1, T_2, T_3 はそれぞれ S_1, S_2, S_3 において α_j を β_j, u_j を Real β_j, v_j を Imag β_j でおきかえたものである。$g_j^M(\lambda)$ は $|\beta_j| \leqq 1$ において有界であり $\{1, \cos\lambda, \cos 2\lambda\}$ ではられる線形空間内にある。$|T_2| \leqq 1$ のとき $\cos\lambda_0 = -T_2$ となる λ_0 で最小値をとり, その他のときは 0 または π で最小値をとる。最大値は 0 か π でとる。そして

$$\lim_{|\beta_j| \to 0} g_j^M(\lambda) = 1$$

となり white noise に対応する。

AR proces や MA process の自己相関係数の性質が〔1〕,〔2〕,〔6〕などで論じられているが, ここでも上の議論を更に発展させて AR モデルのあてはめと MA モデルのあてはめのそれぞれの特徴および推定量の性質も考慮に入れた議論を将来展開したいものだと考えている。パラメータの数を一定にしておいた上での AR モデルによる達成可能な $f(\lambda)$ の族と MA モデルにより達成可能な $f(\lambda)$ の族や ARMA モデルによって達成される $f(\lambda)$ の族の関係, それぞれのモデルにおけるパラメータ数と推定量の分散の関係, および平均2乗誤差との関係等が今後の問題である。

References

[1] Anderson, O. D., (1975), The recursive nature of the stationarity and invertibility restraints on the parameters of mixed autoregressive-moving average processes, Biometrika, 62, 704-706.

[2] Anderson, O. D., (1976), On two convex autocorrelation regions for moving average processes, Biometrika, 63, 681-683.

[3] Anderson, T. W., (1971), The statistical analysis of time series, John Wiley & Sons, Inc., New York.

[4] Box, G.E.P. & Jenkins, G. M., (1970), Time series analysis and forcasting and control, Holden Day, San Francisco.

[5] Cråmer, H., (1960), On some class of non stationary stochastic processes, Proc. 4th Berkeley Symp. on Math. Stat. & Prob. 2, 57-78.

[6] Davies, N., Pate, M. B. & Frost, M. G., (1974), Maximum autocorrelations for moving average processes, Biometrika, 61, 199-200.

[7] Mandrekar, V., (1968), A characterization of oscillatory process and their prediction, Proc. Amer. Math. Soc., 32, 280-284.

[8] Priestley, M. B. (1965), Evolutionary spectra and nonstationary processes, J. R. S. S., 25, 153-188.