時系列ゼミ
第 3 回

伊藤真道

June 1st 2020

＜今回のお言葉＞
かつて戦った敵と再び相見えるのは激アツ

2 Chap2 つづき

2.3 Introduction to ARMA processes

Def. 2.3.1 ARMA(1,1) process

時系列 $\{X_t\}$ が ARMA(1,1) process であるとは，

1. 定常である (つまり，$\mu(t) = E[X_t], \gamma(t, t+h)$ do NOT depend on the time t and $\gamma(t, t+h)$ is dependent on only the lag h.)
2. $\{Z_t\} \sim WN(0, \sigma^2), \phi + \theta \neq 0$ なる ϕ, θ を用いて，

$$X_t - \phi X_t = Z_t + \theta Z_{t-1} \tag{2.3.1}$$

と表されることである．

線形フィルタ $\phi(B) := 1 - \phi B, \theta(B) = 1 + \theta B$ を用いると，(2.3.1) は，

$$\phi(B)X_t = \theta(B)Z_t \tag{2.3.2}$$

と表すことができる．$|\phi| < 1$ の時，(2.3.1) の定常解は存在する．なぜならば，$\chi(z) := \sum_{j=0}^{\infty} \phi^j z^j (1/\phi(z) = 1/(1-\phi z)$ のテイラー展開) とすると，$|\phi| < 1$ より，$\sum_{j=0}^{\infty} |\phi^j| < \infty$

となる．(2.2.7) 式の結果を用いると，$\chi(B)\phi(B)=1$ であり，また，

$$\begin{aligned}\psi(B) &:= \chi(B)\theta(B) = (\sum_{j=0}^{\infty}\phi^j B^j)(1+\theta B)\\&= (1+\phi B+\phi^2 B^2+\cdots)(1+\theta B)\\&= (1+(\phi+\theta)B\phi^0+(\phi+\theta)B^2\phi+(\phi+\theta)B^3\phi^2+\cdots)\\&= \sum_{j=0}^{\infty}\psi_j B^j\\&(where\ \psi_0=1\ and\ \psi_j=(\phi+\theta)\phi^{j-1}\ for\ j\geq 1)\end{aligned}$$

であることから，(2.3.2) の両辺に $\chi(B)$ をかけると，

$$\begin{aligned}X_t &= \chi(B)\theta(B)Z_t = \psi(B)Z_t\\&= \sum_{j=1}^{\infty}\psi_j Z_{t-j}\\&= Z_t+(\phi+\theta)\sum_{j=1}^{\infty}\phi^{j-1}Z_{t-j}\end{aligned}\tag{2.3.3}$$

となる．ここで，

$$\begin{aligned}&E[(X_t-\sum_{j=0}^{k}\psi_j Z_{t-j})^2]\\&= E[(\psi_{k+1}X_{t-k-1})^2]\\&= \psi_{k+1}^2 E[X_{t-k-1}^2]\\&= \phi^{2k}(\phi+\theta)^2 E[X_{t-k-1}^2]\\&\to 0\ as\ k\to\infty\end{aligned}$$

となることから，X_t は $\sum_{j=0}^{\infty}\psi_j Z_{t-j}(=\psi(B)Z_t)$ に二乗平均収束する．よって，example2.2.1 と同様に $MA(\infty)$ process(2.3.3 式) は，ARMA(1.1) の一意な定常解となる．
¡note¿

今度は，$|\phi|>1$ とする．？？？？ $1/\phi(z)=1/(1-\phi z)<0$ に注意して？？？，テイラー展開すると，

$$\frac{1}{\phi(z)} = -\sum_{j=1}^{\infty}\phi^{-j}z^{-j}$$

と表せることから，この場合も，$|\phi| < 1$ の時と同様の議論ができる．この場合は $\chi(B) = -\sum_{j=1}^{\infty}\phi^{-j}B^{-j}$ とおき，(2.3.2) の両辺にこれをかけて，一意な定常解

$$X_t = \chi(B)\theta(B)Z_t = -\theta\phi^{-1}Z_t - (\theta+\phi)\sum_{j=1}^{\infty}\phi^{-(j+1)}Z_{t+j} \tag{2.3.4}$$

を得る．

もし，$\phi = \pm 1$ なら，(2.3.1) の定常解は存在しない．その結果として，上記の定義では，$\phi = \pm 1$ である ARMA(1,1) process は存在しない．まとめると，

- $\phi \neq \pm 1$ ならば，ARMA(1,1) の定常解は存在する
- $|\phi| < 1$ の時
  この時，一意な定常解は (2.3.3) の形で与えられる．この時，X_t は Z_s，$s \leq t$ の現在の値と以前の値によって表現されているため，$\{Z_t\}$ の **causal** もしくは，**causal function** と呼ばれる．
- $|\phi| > 1$ の時
  この時，一意な定常解は (2.3.4) の形で与えられる．X_t は $Z_s, s \geq t$ によって表されているため, この解は **noncausal** である．

¡note¿

Z_t が $X_s, s \leq t$ で表現されうる時，**invertible** といい，この性質を **invertibility** と呼ぶ．今度は，(2.3.1) 式の ARMA(1,1) が，$|\theta| < 1$ の時，invertible であることを示す．$\xi(z) := \sum_{j=0}^{\infty}(-\theta)^j z^j (1/\theta(z) = 1/(1+\theta z)$ のテイラー展開) とすると，$|\theta| < 1$ から，$\sum_{j=1}^{\infty}|(-\theta)^j| < \infty$ である．(2.2.6) 式から，ここも $\xi(B)\theta(B) = 1$ であり，(2.3.2) 式の両辺に $\xi(B)$ をかけることで，

$$Z_t = \xi(B)\phi(B)X_t = \pi(B)X_t$$

とかける．ここで，

$$\begin{aligned}
\pi(B) &= (1 - \theta B + (-\theta)^2 B^2 + (-\theta)^3 B^3 + \cdots)(1 - \phi B) \\
&= (1 - (\theta+\phi)B - (\theta+\phi)B^2(-\theta)^1 - (\theta+\phi)B^3(-\theta)^2 + \cdots) \\
&= \sum_{j=0}^{\infty}\pi_j B^j \\
& where\ \pi_0 = 1\ and\ \pi_j = -(\theta+\phi)B^j(-\theta)^{j-1}\ for\ j \geq 1
\end{aligned}$$

であるから，

$$Z_t = X_t - (\theta + \phi)\sum_{j=1}^{\infty}(-\theta)^{j-1}X_{t-j} \tag{2.3.5}$$

と表せる．よって，$|\theta| < 1$ の時，Z_t は $X_s, s \leq t$ で表すことができる，つまり invertible である．同様の議論によって，$|\theta| > 1$ の時，ARMA(1,1) process は，

$$Z_t = -\phi\theta^{-1} + (\theta + \phi)\sum_{j=1}^{\infty}(-\theta)^{-j-1}X_{t+j} \tag{2.3.6}$$

のように表されることがわかる．よってこの時，**noninvertible** である．まとめると，

- $|\theta| < 1$ の時，ARMA(1,1) process は invertible.
- $|\theta| > 1$ の時，ARMA(1,1) process は noninvertible

Remark 2.1. $\theta = \pm 1$ の時でも，*ARMA(1,1) process* は *invertible* であるが，綺麗* な形で表せないので，以降 *invertible* というときは，$|\theta| < 1$ であるとする．

Remark 2.2. *ARMA(1,1) process* が $|\theta| > 1$ で *noncausal*，もしくは，*noninvertible* であるときには，X_t が *causal* で *noninvertible* となるようなホワイトノイズ W_t を見つけられる．ゆえに，二次の観点からは，*causal* かつ *invertible* な *ARMA(1,1)* のみを考えるので十分である．より高次の観点からも同様である．*(？？？？？)*

¡note¿

2.4 Properties of the Sample Mean and Autocorrelation Function

2.4.1 Estimation of μ

定常過程の平均 μ のモーメント推定量は，標本平均

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n}X_i \tag{2.4.1}$$

であり，これは不偏推定量である．標本平均の平均二乗誤差は，

$$
\begin{aligned}
(Var(\bar{X}_n) =)E[(\bar{X}_n - \mu)^2] &= n^{-2}E[(\sum_{i=1}^{n}(X_i - \mu))^2] \\
&= n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j) \\
&= n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma(i-j) \\
&= n^{-2}\sum_{i-j=-n}^{n}(n - |i-j|)\gamma(i-j) \ (\text{は?????}) \\
&= n^{-1}\sum_{h=-n}^{n}(1 - \frac{|h|}{n})\gamma(h) \qquad (2.4.2)
\end{aligned}
$$

ここで，$\gamma(h) \to 0 \ as \ h \to \infty$ ならば，(2.4.2) の右辺は 0 に収束する．つまり，このとき，標本平均は μ に二乗平均収束する．もし，$\sum_{h=-\infty}^{\infty}|\gamma(h)| < \infty$ なら，(2.4.2) 式から，$\lim_{n\to\infty} nVar(\bar{X}_n) = \sum_{|h|<\infty}\gamma(h)$ を得る．これらを以下にまとめる．

Proposition 2.4.1

もし，$\{X_t\}$ が定常な時系列であり，その平均，自己共分散をそれぞれ $\mu, \gamma(\cdot)$ で表す．このとき，$n \to \infty$ とするとき，

$$
Var(\bar{X}_n) = E[(\bar{X}_n - \mu)^2] \to 0 \ if \ \gamma(n) \to 0
$$

$$
nE[(\bar{X}_n - \mu)^2] \to \sum_{|h|<\infty}\gamma(h) \ if \sum_{|h|<\infty}|\gamma(h)| < \infty
$$

また，時系列が Gaussian ならば，

$$
n^{1/2}(\bar{X}_n - \mu) \sim N\left(0, \sum_{|h|<\infty}(1 - \frac{|h|}{n})\gamma(h)\right)
$$

となる．これから，簡単に μ の信頼区間を構成できる．

多くの時系列モデル，特に線形なものと ARMA モデルにおいて，$\bar{X}_n$ は漸近的に平均 μ，分散 $n^{-1}\sum_{|h|<\infty}\gamma(h)(=: \nu)$ の正規分布に従う．この分布による μ の 95% 信頼区間は，

$$
(\bar{X}_n - 1.96\nu^{1/2}/\sqrt{n}, \bar{X}_n + 1.96\nu^{1/2}/\sqrt{n}) \qquad (2.4.3)
$$

となる．普通，ν はわからないので，その場合は推定値で置き換えなければならない．

¡note¿

2.4.2 Estimation of $\gamma(\cdot)$ and $\rho(\cdot)$

標本 ACVF，ACF は

$$\hat{\gamma}(h) = n^{-1}\sum_{t=1}^{n-|h|}(X_{t+|h|}-\bar{X}_n)(X_t-\bar{X}_n) \tag{2.4.4}$$

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \tag{2.4.5}$$

で定義された．これらは，n^{-1} を $(n-h)^{-1}$ と置き換えても不偏推定量とはならない．とは言っても一般的な仮定の下，さらに標本数が大きい場合は，ほとんど不偏に近くなるらしい．今度は，標本 ACVF 行列

$$\hat{\Gamma}_k = \begin{bmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & \cdots & \hat{\gamma}(k-1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) & \cdots & \hat{\gamma}(k-2) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{\gamma}(k-1) & \hat{\gamma}(k-2) & \cdots & \hat{\gamma}(0) \end{bmatrix} \tag{1}$$

の性質について考える．$\hat{\Gamma}_k$ は非負定値 (nonnegative definite) である．以下これを確認していく．まず，$k<m$ に対して，$\hat{\Gamma}_m$ が非負定値なら，$\hat{\Gamma}_k$ も非負定値であることに注意する．次に，$k \geq n$ を仮定し，$Y_i = X_i - \bar{X}_n$ として，

$$T = \begin{bmatrix} 0 & \cdots & 0 & 0 & Y_1 & Y_2 & \cdots & Y_k \\ 0 & \cdots & 0 & Y_1 & Y_2 & \cdots & Y_k & 0 \\ \vdots & & & & & & & \vdots \\ 0 & Y_1 & Y_2 & \cdots & Y_k & 0 & \cdots & 0 \end{bmatrix}$$

と定義すると，

$$\hat{\Gamma}_k = n^{-1}TT'$$

となる．よって，$\forall a \in \mathbb{R}^k$ に対して，

$$a'\hat{\Gamma}_k a = n^{-1}(a'T)(T'a) = n^{-1}\|T'a\|_2^2 \geq 0$$

と表せることから，標本 ACV 行列 $\hat{\Gamma}_k$，標本 AC 行列 $\hat{R}_k = \hat{\Gamma}_k/\gamma(0)$ は非負定値となることが確認できた．たまーーーに，標本 ACVF, ACF の n^{-1} を $(n-h)^{-1}$ に置き換えた定義を見ることがあるが，このとき，$\hat{\Gamma}_k, \hat{R}_k$ は非負定値とならない可能性がある (らしい)．

¡note¿

観測データ $X_1, X_2, \ldots, X_n$ 以外の情報がなければ，$h > n$ なる h に対して妥当な $\gamma(h), \rho(h)$ の推定を行えない．また，例え，$h \leq n$ としても，例えば，$h = n-1$ の時の ACVF を計算するときのペア (X_{t+h}, X_t) は一つしか得られないように，h が n よりわずかに小さい場合は，その推定値 $\hat{\gamma}(h), \hat{\rho}(h)$ は信頼できるものではない．Jenkins(1976) では，

- $n \geq 50$
- $h \leq n/4$

であるべきであると主張されている．
¡note¿

標本 ACF$\hat{\rho}(h)$ は，適切なモデルを選択するのに重要な役割を持っていることは以前に述べた (e.g. chap1 の最後の方の例，時系列が iid かどうかの検定)．$\rho(h)$ に関する推論を行うためには，推定量 $\hat{\rho}(h)$ の標本分布が必要であるが，これは，時系列がかなり単純な形でも計算しづらい．そのため，大標本の際は正規分布で近似できることを利用する．線形モデル，特に ARMA モデルに対しては，$\hat{\boldsymbol{\rho}}_k = (\hat{\rho}(1), \hat{\rho}(2), \ldots, \hat{\rho}(k))'$ とすると，

$$\hat{\boldsymbol{\rho}}_k \approx N(\boldsymbol{\rho}_k, n^{-1}W)$$

が成り立つ．ここで，$\boldsymbol{\rho}_k = (\rho(1), \rho(2), \ldots, \rho(k))'$ であり，共分散行列 W の (i,j) 要素は Bartlett の公式

$$\begin{aligned} w_{ij} = \sum_{k=-\infty}^{\infty} \{ & \rho(k+i)\rho(k+j) + \rho(k-j)\rho(k+j) + 2\rho(i)\rho(j)\rho^2(k) \\ & - 2\rho(i)\rho(k)\rho(k+j) - 2\rho(j)\rho(k)\rho(k+j)\} \end{aligned}$$

で与えられるものである．上手い具合に因数分解すると，

$$
\begin{aligned}
w_{ij} = \sum_{k=1}^{\infty} &\{\rho(k+i) + \rho(k-i) - 2\rho(i)\rho(k)\} \\
&\times \{\rho(k+j) + \rho(k-j) - 2\rho(j)\rho(k)\}
\end{aligned} \tag{2.4.10}
$$

となる．こっちの方が計算上は便利かも．

Taylor expansion(exponential power series)

今回の担当章では，テイラー展開がいっぱい出てきた (よっ，久しぶりだな！学部の授業以来だね，元気してた？)．テイラー展開は以下のテイラーの定理によっている．

テイラーの定理

$k \geq 1$ を整数とし，関数 $f : \mathbb{R} \to \mathbb{R}$ が点 $a \in \mathbb{R}$ で k 階微分可能であるとする．このとき，

$$
\begin{aligned}
f(x) = f(a) + f'(a)(x-a) &+ \frac{f^{(2)}}{2!}(a)(x-a)^2 + \cdots + \frac{f^{(k)}}{k!}(x-a)^k + h_k(x)(x-a)^k \\
&\lim_{x \to a} h_k(x) = 0
\end{aligned}
$$

を満たす関数 $h_k : \mathbb{R} \to \mathbb{R}$ が存在する．この $h_k(x)$ はペアノの剰余項と呼ばれる．

テイラー展開は，非線形関数の最適化の際によく利用されるニュートン法の「ちゅきぴ」とも言える．ニュートン法は，局所的にテイラー展開によって二次近似を行っているからだ．

それは置いといて，本編では，$1/(1-\phi z)$ や $1/(1+\theta z)$ のテーラー展開があった．とりあえず，$f(z) = 1/(1-\phi z)$, $|\phi| < 1$ のテイラー展開やってみる．難しい事を考えず n 階微分してみると，

$$
\begin{aligned}
f'(z) &= \frac{-1}{(1-\phi z)^2} \times (-\phi) = \frac{\phi}{(1-\phi z)^2} \\
f''(z) &= \frac{-2\phi}{(1-\phi z)^3} \times (-\phi) = \frac{2\phi^2}{(1-\phi z)^3} \\
f'''(z) &= \frac{-6\phi^2}{(1-\phi z)^4} \times (-\phi) = \frac{6\phi^3}{(1-\phi z)^4} \\
&\vdots \\
f^{(n)} &= \frac{n!\phi^n}{(1-\phi z)^{n+1}}
\end{aligned}
$$

ここで，$a = 0$ とすると，$f^{(k)}(0) = k!\phi^k$, $k = 0, 1, 2, \ldots, n$ であり，この関数の $a = 0$ の近傍における n 次のテイラー近似 $\chi_n(z)$ は，

$$
\chi_n(z) = \sum_{j=0}^{n} \frac{f^{(j)}(0)}{j!} z^j = \sum_{j=0}^{n} \phi^j z^j
$$

これの $n \to \infty$ としたものが，本編の $\chi(z)$ である．

¡note¿

コンボリューション (畳み込み，合成積, convolution)

私はコンボリューションがわからぬ．確率変数の和の pdf の導出や，微積分の問題でいくつか解いた事はあるが意味がわからない．計算自体はできるが，結局何をやってんだかさっぱりわからん．以下，高校数学の美しい物語 (https://mathtrain.jp/tatami) を参考にした．

畳み込みの意味

- 添字の和が一定である部分の関数値をかけて足し合わせたもの
- 少しずつ逆方向に添字をスライドさせながら掛け合わせたもの．

それでは，連続 ver. と離散 ver. の二つの定義を紹介する．

連続 ver.(畳み込み積分)

二つの連続関数 $f(x), g(x)$ から，以下のような新しい関数

$$h(x) = \int_{-\infty}^{\infty} f(x)g(t-x)dx$$

を作る操作のこと．

離散 ver.(畳み込み和)

定義域が整数値であるような二つの関数 (or 数列)a_n, b_n から，以下のような新しい数列

$$c_n = \sum_{t=-\infty}^{\infty} a_n b_{t-n}$$

を作る操作のこと．

例を見てみよう．

例 2.1. 多項式 $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ と $b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$ の積の，x^k の係数は，添字の和が k となる a_i, b_j，$(i, j = 0, 1, 2, \ldots, i + j = k)$ の畳み込み和となる．例えば，$(a_2 x^2 + a_1 x + a_0)(b_4 x^4 + b_3 x^3 + b_2 x^2 + b_1 x + b_0)$ の x^4 の係数は，$a_2 b_2 + a_1 b_3 + a_0 b_4$ となる．

例 2.2. 独立な確率変数 X, Y の分布がわかっているとき，$T := X + Y$ の分布は，離散の場合

$$P(T = t) = P(X + Y = t) = \sum_k P(X = k)P(Y = t - k)$$

連続の場合，

$$g(t) = \int_{-\infty}^{\infty} f_1(x) f_2(t - x) dx$$

で与えられる．独立でない場合は，それぞれ，

$$P(T = t) = \sum_k P(X = k, Y = t - k)$$

$$g(t) = \int_{-\infty}^{\infty} f(x, t - x) dx$$

で与えられる．連続の場合は，$t = x + y, u = y$ と変数変換して導出する．

前回の I 氏の発表担当分の (2.2.6) 式

$$\psi_i = \sum_{k=-\infty}^{\infty} \alpha_k \beta_{j-k} = \sum_{k=-\infty}^{\infty} \beta_k \alpha_{j-k}$$

も典型的な畳み込みである．その後の

$$\begin{aligned} \psi(B) &:= \phi(B)\pi(B) = (1 - \phi B)(\sum_{j=0}^{\infty} \phi^j B^j) \\ &= \sum_{j=0}^{\infty} \phi^j B^j - \sum_{j=1}^{\infty} \phi^j B^j = \phi^0 B^0 = 1 \end{aligned}$$

は関係ない (私が確認したかっただけ)．