第 5 回　1 変量時系列モデルの定式化と推定（7.1.3, 7.2.2–7.2.3）

村澤 康友

2020 年 10 月 27 日

今日のポイント

1. $\{\Delta^d y_t\}$ が共分散定常なら $\{y_t\}$ を d 次の和分（単位根）過程という．$\{\Delta y_t\}$ が WN なら $\{y_t\}$ をランダム・ウォークという．(p,d,q) 次の自己回帰和分移動平均（ARIMA）過程は，任意の t について $\phi(\mathrm{L})\left(\Delta^d y_t - \mu\right) = \theta(\mathrm{L})w_t$．ただし $\phi(\mathrm{L}),\theta(\mathrm{L})$ はラグ多項式で $\{w_t\}$ は WN.
2. AR モデルの係数の OLS 推定量は不偏でないが一致性をもつ．
3. ARMA モデルの尤度関数は予測誤差分解で計算する．時系列の同時 pdf を尤度とする ML 法を厳密な ML 法，初期値を所与とした条件つき pdf を尤度とする ML 法を条件つき ML 法という．
4. ARMA モデルの次数は AIC·SBIC·HQC 等のモデル選択基準で選ぶ．

目次

1 ARIMA 過程 1
- 1.1 差分と和分 1
- 1.2 和分（単位根）過程（p. 128） . . . 2
- 1.3 ARIMA 過程（p. 138） 2

2 AR モデルの OLS 推定 2
- 2.1 OLS 推定量 2
- 2.2 有限標本特性 2
- 2.3 漸近特性 2

3 正規 ARMA モデルの ML 推定 3
- 3.1 最尤（ML）法 3
- 3.2 予測誤差分解 3
- 3.3 厳密な ML 法（p. 137） 4
- 3.4 条件つき ML 法 4

4 次数選択 4
- 4.1 仮説検定とモデル選択 4
- 4.2 Kullback–Leibler 情報量 4
- 4.3 モデル選択基準（p. 132） 4

5 今日のキーワード 5

6 次回までの準備 5

1 ARIMA 過程

1.1 差分と和分

$\{x_t\}$ を $t=1$ から始まる数列とする．

定義 1. $\{x_t\}$ の差分（階差）は $\{\Delta x_t\}$.

定義 2. $\{x_t\}$ の和分は $t \geq 1$ について

$$S_t := x_1 + \cdots + x_t$$

注 1. $t \geq 1$ について

$$\begin{aligned}\Delta S_t &:= S_t - S_{t-1} \\ &= x_1 + \cdots + x_t - (x_1 + \cdots + x_{t-1}) \\ &= x_t\end{aligned}$$

すなわち和分は差分の逆の演算．離散空間上の差分と和分の関係は，連続空間上の微分と積分の関係に相当．

1.2　和分（単位根）過程（p. 128）

$\{y_t\}$ を確率過程とする．

定義 3. $\left\{\Delta^d y_t\right\}$ **が共分散定常なら** $\{y_t\}$ **を** d **次の和分（単位根）過程という．**

注 2. $\mathrm{I}(d)$ と書く．$\mathrm{I}(0)$ ＝共分散定常．

注 3. $\mathrm{I}(d)$ は $\mathrm{I}(0)$ に変換して分析する．

定義 4. $\{\Delta y_t\}$ **がホワイト・ノイズなら** $\{y_t\}$ **をランダム・ウォークという．**

注 4. $\{w_t\}$ を WN とすると，任意の t について

$$\Delta y_t = w_t$$

AR(1) は，任意の t について

$$\phi(\mathrm{L})(y_t - \mu) = w_t$$

ただし $\phi(\mathrm{L}) := 1 - \phi\mathrm{L}$．ランダム・ウォークは $\mu := 0$，$\phi(\mathrm{L}) := \Delta = 1 - \mathrm{L}$ の AR(1)．このとき $\phi(z) := 1 - z = 0$ の根は $z = 1$（単位根）．

1.3　ARIMA 過程（p. 138）

$\{y_t\}$ を $\mathrm{I}(d)$ とする．

定義 5. (p, d, q) **次の自己回帰和分移動平均（*autoregressive integrated moving average, ARIMA*）過程は，任意の** t **について**

$$\begin{aligned}\phi(\mathrm{L})\left(\Delta^d y_t - \mu\right) &= \theta(\mathrm{L}) w_t \\ \{w_t\} &\sim \mathrm{WN}\left(\sigma^2\right)\end{aligned}$$

ただし $\phi(\mathrm{L}) := 1 - \phi_1\mathrm{L} - \cdots - \phi_p\mathrm{L}^p$，$\theta(\mathrm{L}) := 1 - \theta_1\mathrm{L} - \cdots - \theta_q\mathrm{L}^q$．

注 5. $\mathrm{ARIMA}(p, d, q)$ と書く．

2　AR モデルの OLS 推定

2.1　OLS 推定量

時系列 $(y_0, \ldots, y_T)$ に平均 0 の AR(1) モデルを仮定する．すなわち $t = 1, \ldots, T$ について

$$\begin{aligned}y_t &= \phi y_{t-1} + w_t \\ \{w_t\} &\sim \mathrm{WN}\left(\sigma^2\right)\end{aligned}$$

ただし $|\phi| < 1$．ϕ の OLS 推定量を $\hat{\phi}_T$ とすると

$$\hat{\phi}_T = \frac{\sum_{t=1}^T y_{t-1} y_t}{\sum_{t=1}^T y_{t-1}^2}$$

2.2　有限標本特性

定理 1. 一般に

$$\mathrm{E}\left(\hat{\phi}_T\right) \neq \phi$$

証明. 式変形すると

$$\begin{aligned}\hat{\phi}_T &= \frac{\sum_{t=1}^T y_{t-1} y_t}{\sum_{t=1}^T y_{t-1}^2} \\ &= \frac{\sum_{t=1}^T y_{t-1}(\phi y_{t-1} + w_t)}{\sum_{t=1}^T y_{t-1}^2} \\ &= \phi + \frac{\sum_{t=1}^T y_{t-1} w_t}{\sum_{t=1}^T y_{t-1}^2}\end{aligned}$$

第 2 項の期待値は

$$\begin{aligned}&\mathrm{E}\left(\frac{\sum_{t=1}^T y_{t-1} w_t}{\sum_{t=1}^T y_{t-1}^2}\right) \\ &= \sum_{t=1}^T \mathrm{E}\left(\frac{y_{t-1} w_t}{\sum_{t=1}^T y_{t-1}^2}\right) \\ &= \mathrm{E}\left(\frac{y_0 w_1}{\sum_{t=1}^T y_{t-1}^2}\right) + \cdots + \mathrm{E}\left(\frac{y_{T-1} w_T}{\sum_{t=1}^T y_{t-1}^2}\right) \\ &= \mathrm{cov}\left(\frac{y_0}{\sum_{t=1}^T y_{t-1}^2}, w_1\right) + \cdots \\ &\quad + \mathrm{cov}\left(\frac{y_{T-1}}{\sum_{t=1}^T y_{t-1}^2}, w_T\right)\end{aligned}$$

w_1 は $y_1, \ldots, y_{T-1}$ と相関するので第 1 項は一般に 0 でない．同様に他の項も一般に 0 でない． □

注 6. 無作為標本でないため OLS 推定量は不偏でない．

2.3　漸近特性

定理 2.

$$\operatorname*{plim}_{T\to\infty} \hat{\phi}_T = \phi$$

証明. 式変形すると

$$\begin{aligned}\hat{\phi}_T &= \phi + \frac{\sum_{t=1}^T y_{t-1} w_t}{\sum_{t=1}^T y_{t-1}^2} \\ &= \phi + \frac{(1/T)\sum_{t=1}^T y_{t-1} w_t}{(1/T)\sum_{t=1}^T y_{t-1}^2}\end{aligned}$$

エルゴード定理より

$$\operatorname*{plim}_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}y_{t-1}w_t=\mathrm{E}(y_{t-1}w_t)$$
$$\operatorname*{plim}_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}y_{t-1}^2=\mathrm{E}\left(y_{t-1}^2\right)$$

漸近演算より

$$\begin{aligned}\operatorname*{plim}_{T\to\infty}\hat{\phi}_T&=\phi+\frac{\operatorname{plim}_{T\to\infty}(1/T)\sum_{t=1}^{T}y_{t-1}w_t}{\operatorname{plim}_{T\to\infty}(1/T)\sum_{t=1}^{T}y_{t-1}^2}\\&=\phi+\frac{\mathrm{E}(y_{t-1}w_t)}{\mathrm{E}\left(y_{t-1}^2\right)}\end{aligned}$$

$\mathrm{E}(y_{t-1}w_t)=\operatorname{cov}(y_{t-1},w_t)=0$ より第 2 項は 0. □

注 7. すなわち OLS 推定量は一致性をもつ．また漸近正規性も証明できる（省略).

3　正規 ARMA モデルの ML 推定

3.1　最尤（ML）法

パラメトリックな確率過程を仮定する．母数を θ とし，観測する時系列の同時 pmf・pdf を $f(.;\theta)$ とする.

定義 6. ある母数の下で標本の実現値を観測する確率（密度）を，その母数の**尤度**という.

注 8. $(y_1,\dots,y_T)$ を観測する確率（密度）は

$$f(y_1,\dots,y_T;\theta)$$

これを θ の「尤もらしさ」と解釈する.

定義 7. 標本の pmf・pdf を母数の尤度を表す関数とみたものを**尤度関数**という.

注 9. $L(\theta;y_1,\dots,y_T)$ と書く（$(y_1,\dots,y_T)$ と θ の位置が pmf・pdf と逆).

注 10. $(y_1,\dots,y_T)$ を観測したときの θ の尤度関数は

$$L(\theta;y_1,\dots,y_T):=f(y_1,\dots,y_T;\theta)$$

定義 8. 尤度関数の対数を**対数尤度関数**という.

注 11. $\ell(\theta;y_1,\dots,y_T)$ と書く.

注 12. $(y_1,\dots,y_T)$ を観測したときの θ の対数尤度関数は

$$\begin{aligned}\ell(\theta;y_1,\dots,y_T)&:=\ln L(\theta;y_1,\dots,y_T)\\&=\ln f(y_1,\dots,y_T;\theta)\end{aligned}$$

定義 9. （対数）尤度関数を最大にする解を母数の推定値とする手法を**最尤**（*maximum likelihood, ML*）**法**という.

定義 10. ML 法による推定量を *ML* **推定量**という.

定理 3. *ML 推定量は一般に漸近有効.*

証明. 省略（大学院レベル). □

3.2　予測誤差分解

時系列 $(y_1,\dots,y_T)$ の同時 pdf を $f(.)$，条件つき pdf を $f(.|.)$ で表す.

定理 4 (予測誤差分解)**.** 任意の $(y_1,\dots,y_T)$ について

$$\begin{aligned}&f(y_1,\dots,y_T)\\&=f(y_T|y_{T-1},\dots,y_1)\cdots f(y_2|y_1)f(y_1)\end{aligned}$$

証明. 条件つき pdf の定義より，任意の $(y_1,\dots,y_T)$ について

$$\begin{aligned}&f(y_1,\dots,y_T)\\&=\frac{f(y_1,\dots,y_T)}{f(y_1,\dots,y_{T-1})}\frac{f(y_1,\dots,y_{T-1})}{f(y_1,\dots,y_{T-2})}\cdots\\&\quad\frac{f(y_1,y_2)}{f(y_1)}f(y_1)\\&=f(y_T|y_{T-1},\dots,y_1)f(y_{T-1}|y_{T-2},\dots,y_1)\cdots\\&\quad f(y_2|y_1)f(y_1)\end{aligned}$$

□

例 1. 時系列 $(y_1,\dots,y_T)$ に平均 0 の正規 AR(1) モデルを仮定する．すなわち $t=1,\dots,T$ について

$$y_t=\phi y_{t-1}+w_t$$
$$\{w_t\}\sim\mathrm{IN}\left(0,\sigma^2\right)$$

ただし $|\phi|<1$．$\operatorname{var}(y_1)=\sigma^2/\left(1-\phi^2\right)$ より

$$y_1\sim\mathrm{N}\left(0,\frac{\sigma^2}{1-\phi^2}\right)$$

また $t=2,\dots,T$ について

$$y_t|y_{t-1},\dots,y_1 \sim \mathrm{N}\left(\phi y_{t-1},\sigma^2\right)$$

予測誤差分解より尤度関数は

$$\begin{aligned}
&L(\phi,\sigma^2;y_1,\dots.y_T)\\
&:= f(y_1,\dots,y_T)\\
&= f(y_T|y_{T-1},\dots,y_1)\cdots f(y_2|y_1)f(y_1)\\
&= \prod_{t=2}^{T}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(y_t-\phi y_{t-1})^2}{2\sigma^2}\right)\\
&\quad \frac{1}{\sqrt{2\pi\sigma^2/(1-\phi^2)}}\exp\left(-\frac{y_1^2}{2\sigma^2/(1-\phi^2)}\right)
\end{aligned}$$

対数尤度関数は

$$\begin{aligned}
&\ell(\phi,\sigma^2;y_1,\dots.y_T)\\
&:= \ln L(\phi,\sigma^2;y_1,\dots.y_T)\\
&= \sum_{t=2}^{T}\left\{-\frac{1}{2}\ln 2\pi-\frac{1}{2}\ln\sigma^2-\frac{(y_t-\phi y_{t-1})^2}{2\sigma^2}\right\}\\
&\quad -\frac{1}{2}\ln 2\pi-\frac{1}{2}\ln\frac{\sigma^2}{1-\phi^2}-\frac{y_1^2}{2\sigma^2/(1-\phi^2)}\\
&= -\frac{T}{2}\ln 2\pi-\frac{T}{2}\ln\sigma^2+\frac{1}{2}\ln\left(1-\phi^2\right)\\
&\quad -\frac{(1-\phi^2)y_1^2}{2\sigma^2}-\frac{1}{2\sigma^2}\sum_{t=2}^{T}(y_t-\phi y_{t-1})^2
\end{aligned}$$

3.3 厳密な ML 法（p. 137）

定義 11. $(y_1,\dots,y_T)$ の同時 pdf を尤度とする ML 法を**厳密な** ***ML*** **法**という．

注 13. 漸近有効だが尤度の計算が複雑（ARMA モデルを状態空間モデルで表現し，カルマン・フィルターで尤度を計算する）．

3.4 条件つき ML 法

定義 12. 初期値 $(y_1,\dots,y_p)$ と $(w_{p-q+1},\dots,w_p)$ を所与とした $(y_{p+1},\dots,y_T)$ の条件つき pdf を尤度とする ML 法を**条件つき** ***ML*** **法**という．

注 14. 初期値の周辺 pdf を省くと漸近有効でないが推定が簡単になる．AR モデルなら条件つき ML 法＝ OLS.

4 次数選択

4.1 仮説検定とモデル選択

予測モデルの次数選択は仮説検定と異なる．

1. 真の次数が無限大なら真のモデルは推定できず仮説検定も無意味.
2. 真の次数が有限でも推定する係数が多いと予測値が不安定になる.

モデル選択基準による次数選択が便利.

4.2 Kullback–Leibler 情報量

確率変数 Y の真の分布を $f_0(.)$，予測モデルの下での分布を $f(.)$ とする.

定義 13. $f(.)$ の $f_0(.)$ に対する *Kullback–Leibler* **情報量**は

$$I(f(.);f_0(.)) := -\mathrm{E}_{f_0}\left(\ln\frac{f(Y)}{f_0(Y)}\right)$$

注 15. $f(.)$ の $f_0(.)$ に対する「距離」を表す．$f(.)=f_0(.)$ なら「距離」は 0 で最小．ただし真の次数が無限大なら $f_0(.)$ は予測に使えない．式変形すると

$$I(f(.);f_0(.)) = -\mathrm{E}_{f_0}(\ln f(Y))+\mathrm{E}_{f_0}(\ln f_0(Y))$$

第 1 項を最小化，すなわち $\mathrm{E}_{f_0}(\ln f(Y))$ を最大化する $f(.)$ が最適な予測モデル．$\mathrm{E}_{f_0}(\ln f(Y))$ は未知なので推定が必要.

4.3 モデル選択基準（p. 132）

定常過程 $\{Y_t\}$ の 1 期先予測モデルを $f(.;\boldsymbol{\theta})$ とする．$\mathrm{E}(\ln f(Y_t;\boldsymbol{\theta}))$ を最大化する $\boldsymbol{\theta}$ を $\boldsymbol{\theta}^*$ とする．$\mathrm{E}(\ln f(Y_t;\boldsymbol{\theta}^*))$ の推定は 2 つの推定を含む.

1. $\boldsymbol{\theta}^*$ の推定
2. $\boldsymbol{\theta}^*$ を所与とした $\mathrm{E}(\ln f(Y_t;\boldsymbol{\theta}^*))$ の推定

$\boldsymbol{\theta}^*$ が既知ならエルゴード定理より

$$\operatorname*{plim}_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\ln f(Y_t;\boldsymbol{\theta}^*) = \mathrm{E}(\ln f(Y_t;\boldsymbol{\theta}^*))$$

$\boldsymbol{\theta}^*$ の ML 推定量を $\hat{\boldsymbol{\theta}}_T$ とする．$\boldsymbol{\theta}^*$ を $\hat{\boldsymbol{\theta}}_T$ で置き換えると偏りが生じるので修正が必要．未知係数の数を k とする.

補題 1. 任意の $\boldsymbol{\theta}$ と $(y_1,\dots,y_T)$ について

$$\sum_{t=1}^{T}\ln f(y_t;\boldsymbol{\theta})=\ell(\boldsymbol{\theta};y_1,\dots,y_T)$$

証明. 予測誤差分解より

$$f(y_1,\dots,y_T;\boldsymbol{\theta})=\prod_{t=1}^{T}f(y_t;\boldsymbol{\theta})$$

したがって

$$\begin{aligned}\ell(\boldsymbol{\theta};y_1,\dots,y_T)&:=\ln f(y_1,\dots,y_T;\boldsymbol{\theta})\\&=\sum_{t=1}^{T}\ln f(y_t;\boldsymbol{\theta})\end{aligned}$$

□

定義 14. **赤池の情報量基準**（*Akaike's information criterion, AIC*）は

$$\mathrm{AIC}:=-2\ell\left(\hat{\boldsymbol{\theta}}_T;y_1,\dots,y_T\right)+2k$$

注 16. AIC が最小のモデルを選択する．第 2 項は偏りの修正項であり，モデルの大きさに対するペナルティーと解釈できる．

定義 15. 真のモデルを選ぶ確率が $T\to\infty$ で 1 に収束する性質をモデル選択基準の**一致性**という．

注 17. AIC は一致性をもたない（過剰定式化の傾向がある）．

定義 16. ***Schwarz*** **のベイズ情報量基準**（*Schwarz's Bayesian information criterion, SBIC*）は

$$\mathrm{SBIC}:=-2\ell\left(\hat{\boldsymbol{\theta}}_T;y_1,\dots,y_T\right)+k\ln T$$

注 18. $\boldsymbol{\theta}^*$ をベイズ法で推定した場合の周辺尤度 $\mathrm{E}(\ell(\boldsymbol{\theta}^*;y_1,\dots,y_T))$ の近似から得られる．

注 19. SBIC は一致性をもつ．$\ln T>2$ ならモデルの大きさに対するペナルティーが AIC より大きく，AIC より小さいモデルを選択する．

定義 17. ***Hannan–Quinn*** **の基準**（*Hannan–Quinn criterion, HQC*）は

$$\mathrm{HQC}:=-2\ell\left(\hat{\boldsymbol{\theta}}_T;y_1,\dots,y_T\right)+2k\ln\ln T$$

注 20. HQC も一致性をもつ．モデルの大きさに対するペナルティーは AIC と SBIC の中間．

5 今日のキーワード

差分（階差），和分，和分（単位根）過程，ランダム・ウォーク，自己回帰和分移動平均（ARIMA）過程，尤度，尤度関数，対数尤度関数，最尤（ML）法，ML 推定量，予測誤差分解，厳密な ML 法，条件つき ML 法，Kullback–Leibler 情報量，赤池の情報量基準（AIC），一致性，Schwarz のベイズ情報量基準（SBIC），Hannan–Quinn の基準（HQC）

6 次回までの準備

提出 宿題 5

復習 教科書第 7 章 1.3, 2.2–2.3 節

予習 教科書第 7 章 7.4.1 節