2022.03.21.
김은영

# 제4장 ARMA모형의 식별 및 추정

## 4.2 시계열 모형의 추정

- 적률법(Method of Moments)
- 보통선형최소제곱법(Ordinary Least Squares Mehod)
- 비선형최소제곱법(Nonlinear Least Squares Method)
- 최대우도법(Maximum Likelihood Method) ⬅

### 4.2.1 정확한 우도함수

※ 시계열 데이터: 인근 시점값들이 독립이 아니기 때문에 정확한 우도함수의 표현이 어려울 수 있다.

■ 시계열 데이터가 AR(1)모형을 따르는 경우

$$Z_t = \phi Z_{t-1} + a_t\ , \qquad a_t \sim N\left(0,\ \sigma_a^2\right)$$

- 데이터 $(Z_1 = z_1,\ Z_2 = z_2)$가 관측되었다면, 결합확률밀도함수는

$$f_{Z_1, Z_2}\left(z_1, z_2; \phi, \sigma_a^2\right) = f_{Z_1}\left(z_1; \phi, \sigma_a^2\right) f_{Z_2|Z_1}\left(z_2|z_1; \phi, \sigma_a^2\right)$$

여기서, $Z_2|Z_1 = z_1 \sim N\left(\phi z_1,\ \sigma_a^2\right) \ \rightarrow\ f_{Z_2|Z_1}\left(z_2|z_1; \phi, \sigma_a^2\right) = \dfrac{1}{\sqrt{2\pi}\,\sigma_a} \exp\left[-\dfrac{\left(z_2 - \phi z_1\right)^2}{2\sigma_a^2}\right]$

- $(z_1,\ z_2,\ \cdots,\ z_n)$이 관측되는 경우, AR(1)모형의 우도함수는

$$f_{Z_1, \cdots, Z_n}\left(z_1,\ \cdots, z_n; \phi, \sigma_a^2\right) = f_{Z_1}\left(z_1; \phi, \sigma_a^2\right) \prod_{t=2}^{n} f_{Z_t|Z_{t-1}}\left(z_t|z_{t-1}; \phi, \sigma_a^2\right)$$

여기서, $f_{Z_t|Z_{t-1}}\left(z_t|z_{t-1}; \phi, \sigma_a^2\right) = \dfrac{1}{\sqrt{2\pi}\,\sigma_a} \exp\left[-\dfrac{\left(z_t - \phi z_{t-1}\right)^2}{2\sigma_a^2}\right],\ t = 2,\ \cdots, n$

➤ $Z_1$의 분포가 별도로 필요한데, 통상 수렴된 후의 분포를 사용한다.

$$Z_1 \sim N\left(0,\ \frac{\sigma_a^2}{\left(1 - \phi^2\right)}\right)$$

※ 통상적으로 우도함수는 로그변환을 취하는 로그우도함수를 고려하여 최적화한다.

・ AR(1)모형의 로그우도함수

$$\log \mathscr{L}(\phi, \sigma_a^2; z_1, \cdots, z_n) = \log f_{Z_1}(z_1; \phi, \sigma_a^2) + \sum_{t=2}^{n} \log f_{Z_t|Z_{t-1}}(z_t|z_{t-1}; \phi, \sigma_a^2)$$

$$= -\frac{n}{2}\log(2\pi) + \frac{1}{2}\log(1-\phi^2) - \frac{n}{2}\log\sigma_a^2 - \frac{S(\phi)}{2\sigma_a^2}$$

여기서, $S(\phi) = z_1^2(1-\phi^2) + \sum_{t=2}^{n}(z_t - \phi z_{t-1})^2$

### 4.2.2 조건있는 우도함수(Conditional Likelihood Function)

※ 시계열 데이터 $Z_t$들은 서로 독립이 아니기 때문에 이를 바탕으로 결합확률밀도함수를 유도하고 우도함수를 산출하기 어렵다. 대신 오차항인 $a_t$들은 서로 독립이므로 이를 바탕으로 우도함수를 구성하고자 한다.

■ 오차항 $a_t(t=1,\ 2,\ \cdots,\ n)$들이 관측된다고 하면 결합확률밀도함수는

$$f_{a_1, \cdots, a_n}(a_1, \cdots, a_n; \Theta) = \left(2\pi\sigma_a^2\right)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma_a^2}\sum_{t=1}^{n} a_t^2\right]$$

■ ARMA(p,q)모형

$$a_t = \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q} + Z_t - \phi_1 Z_{t-1} - \cdots - \phi_p Z_{t-p}$$

・ 오차항을 산출하려면 아래 초기값이 필요하다.

$$a_* = \left(a_{1-q}, \cdots, a_{-1}, a_0\right)$$

$$z_* = \left(z_{1-p}, \cdots, z_{-1}, z_0\right)$$

➤ $z_1$ 관측 → $a_1 = \theta_1 a_0 + \cdots + \theta_q a_{1-q} + z_1 - \phi_1 z_0 - \cdots - \phi_p z_{1-p}$

➤ $z_2$ 관측 → $a_2 = \theta_1 a_1 + \cdots + \theta_q a_{2-q} + z_2 - \phi_1 z_1 - \cdots - \phi_p z_{2-p}$ ➤ $\cdots$

・ 위의 초기조건 하에서의 로그우도함수는

$$\log \mathscr{L}_*(\phi, \sigma_a^2; z_1, \cdots, z_n) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma_a^2 - \frac{S_*(\Theta)}{2\sigma_a^2}$$

여기서, $S_*(\Theta) = \sum_{t=1}^{n} a_t^2\left(\Theta | z_*, a_*, z_1, \cdots, z_n\right)$

⇒ 로그우도함수 최대화 = $S_*(\Theta)$ 최소화

■ 초기값에 대한 가정

① 과거 오차항을 모두 0으로 둔다: $a_{1-q} = \cdots = a_0 = 0$

② 과거 시계열 값을 모두 표본평균으로 대체한다: $z_{1-p} = \cdots = z_0 = \bar{z}$

### 4.2.3 조건없는 우도함수(Unconditional Likelihood Function)

■ 오차항들을 산출하기 위해, 과거값들을 예측하여 우도함수를 산출하는 방법

· **후방예측(back forecasting)**: 과거값을 예측하는 것

[예] ARMA(p,q)모형을 후방형태로 표현하면

$$Z_t - \phi_1 Z_{t+1} - \cdots - \phi_p Z_{t+p} = b_t - \theta_1 b_{t+1} - \cdots - \theta_q b_{t+q}$$

또는

$$Z_t = \phi_1 Z_{t+1} + \cdots + \phi_p Z_{t+p} + b_t - \theta_1 b_{t+1} - \cdots - \theta_q b_{t+q}$$

■ 조건없는 우도함수(Box et al., 1994)

$$\log \mathscr{L}(\phi, \sigma_a^2; z_1, \cdots, z_n) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma_a^2 - \frac{S(\Theta)}{2\sigma_a^2}$$

$$\text{여기서, } S(\Theta) = \sum_{t=-M}^{n} E^2\left[a_t | \Theta, z_1, \cdots, z_n\right]$$

★ $M$: 과거 시계열 후방예측값의 변화가 매우 작아지는(예컨대 0.005 이하) 충분히 큰 수

**[예 4.6]**

## 4.3 잔차검정

■ 시계열모형: 오차항이 백색잡음이라고 가정

→ '잔차'를 구하여 이를 검정

■ ARMA(p,q)모형 $\Phi_p(B) Z_t = \Theta_q(B) a_t$에서의 잔차: $\hat{a}_t = \hat{\Theta}_q^{-1}(B) \hat{\Phi}_p(B) Z_t$

→ 실제잔차: $\hat{a}_t = z_t - (t$시점 예측값)

■ 잔차가 정규분포를 따르는지 여부: 정규확률도표

■ 등분산성 및 패턴 유무 확인: 잔차 산점도

■ 잔차 시계열이 백색잡음을 따르는지 여부

➤ ACF와 PACF를 구해서 확인

➤ 포트맨토(portmanteau) 검정

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_K = 0$$

· 시차 $k$의 표본자기상관계수: $r_k = \dfrac{\sum_{t=k+1}^{n} \hat{a}_t \hat{a}_{t-k}}{\sum_{t=k+1}^{n} \hat{a}_t^2}$, $k = 1, 2, \cdots, K$

· 검정통계량

① Box and Pierce(1970) : $Q = n \sum_{k=1}^{K} r_k^2 \sim \chi^2(K-p-q)$

② Ljung and Box(1978): $Q = n(n+2) \sum_{k=1}^{K} \dfrac{r_k^2}{n-k} \sim \chi^2(K-p-q)$

[예] 나일강 유량 시계열 데이터에 대한 LB 검정 결과

| 시차 | 12 | 24 | 36 | 48 |
|---|---|---|---|---|
| 카이제곱 | 11.9 | 16.6 | 25.0 | 37.3 |
| 자유도 | 9 | 21 | 33 | 45 |
| p-값 | 0.222 | 0.735 | 0.841 | 0.787 |