시계열분석팀 2주차 교안 ⌛

[목차]

1. 1주차 복습

(1) 정상성

① $\mathrm{E}(\mathrm{y_t}) = \mu < \infty$: 평균

② $\mathrm{Var}(\mathrm{y_t}) = E(y_t - \mu)^2 = \gamma_0 < \infty$: 분산

③ $\mathrm{Cov}(\mathrm{y_t}, \mathrm{y_{t+k}}) = E(y_t - \mu)(y_{t+k} - \mu) = \gamma_k < \infty \quad \forall k$: 자기공분산

➔ 시간t에 관계없이 평균과 분산이 일정하며, 자기공분산γ_k 가 시차(k)에만 의존함

(2) 정상화

① 시계열 자료의 그래프를 그려본 후 비정상 시계열인지 확인

② 분산이 일정하지 않은 경우 분산안정화 (로그변환, 제곱근 변환, Box-cox변환)

③ 분산 변환 후 추세/계절성을 갖는다면 회귀/평활/차분을 통해 정상화

④ 시계열의 정상부분(stationary component)와 비정상 부분을 분해

⑤ 비정상을 유발하는 성분(추세, 계절성)을 제거하여 정상 시계열을 획득

(3) 백색잡음

- $Y_t \sim WN(0, \sigma_Y^2)$

① $E(Y_t) = 0$

② $Var(Y_t) = \sigma^2$

③ $Cov(Y_t, Y_{t+k}) = 0$ (약정상성조건에 '공분산이 0'이라는 조건이 추가됨)

- 백색잡음은 가장 대표적인 정상시계열
- 많은 확률과정들이 백색잡음으로부터 생성될 수 있다는 측면에서 매우 중요

2. 모형의 필요성

정상성을 만족시키는 Y_t 의 공분산행렬

$$\Gamma = \begin{pmatrix} Cov(Y_1, Y_1) & Cov(Y_1, Y_2) & \dots & Cov(Y_1, Y_n) \\ Cov(Y_2, Y_1) & Cov(Y_1, Y_2) & \dots & Cov(Y_2, Y_n) \\ \vdots & \vdots & \ddots & \\ Cov(Y_n, Y_1) & Cov(Y_n, Y_.) & \dots & Cov(Y_n, Y_n) \end{pmatrix}$$

$$= \begin{pmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \dots & \gamma(n-2) \\ \vdots & \vdots & \ddots & \\ \gamma(n-1) & \gamma(n-2) & \dots & \gamma(0) \end{pmatrix}$$

만약 백색잡음 과정이라면, 대각선 $\gamma(0)$ 을 제외한 나머지가 모두 0이 된다.

따라서 백색잡음 과정이라면 모델링이 필요 없다.

백색잡음 과정이 아니라면, 공분산 $\gamma(1), \gamma(2), \dots \gamma(n-1)$ 모두 추정해야 하는 모수가 된다.

특정모형을 통해 모수를 추정할 수 있으며, 이는 Y_t 의 ACF, PACF그래프를 보고 결정하게 된다.

3. ACF, PACF

(1) ACF(Autocorrelation Function) : 자기상관 함수

✓ 자기공분산함수 : 시간에 따른 상관의 척도

$$\gamma(k) = \gamma_k = Cov(X_t, X_{t+k}) = E[(X_t - \mu)(X_{t+k} - \mu)]$$

✓ 자기상관함수(autocorrelation function)

$$\rho(k) = \rho_k = Corr(X_t, X_{t+k}) = \frac{Cov(X_t, X_{t+k})}{\sqrt{Var(X_t)}\sqrt{Var(X_{t+k})}} = \frac{\gamma(k)}{\gamma(0)},$$

$$\text{이때} \gamma_0 = Var(X_t) = E[(X_t - \mu)^2]$$

✓ 자기 공분산함수와 자기상관함수의 성질

1) $\gamma_0 = Var(Z_t);\ \rho_0 = 1$

2) $|\gamma_k| \le \gamma_0; |\rho_0| \le 1, k = 1, 2, 3, \dots$

3) $\gamma_k = \gamma_{-k}$; $\rho_k = \rho_{-k}, k = 1, 2, 3, \ldots$

✓ 표본자기공분산함수(sample autocovariance function)

$$\widehat{\gamma_k} = \frac{\sum_{t=1}^{n-k}(X_t - \bar{X})(X_{t+k} - \bar{X})}{n}, k = 0,1,2,\ldots, \quad 단, X = \sum_{t=1}^{n} X_t/n$$

✓ 표본자기상관함수(sample autocorrelation function:SACF)

$$\hat{\rho} = \frac{\widehat{\gamma_k}}{\widehat{\gamma_0}}, k = 0, 1, 2, \ldots \quad , E(\widehat{\rho_k}) = \rho_k$$

➔ 자기상관함수는 시차 k에서 자기상관관계가 존재하는지 나타내는 척도이다. 즉, 서로 다른 두 시점이 상호 연관관계를 나타내는 척도이다.

(2) PACF(Partial Autocorrelation Function) : 편자기상관 함수

✓ **편자기상관함수** : 두 변수를 제외한 모든 변수의 영향을 제거한 상태에서, 두 변수사이의 순수한 상호 연관관계

예시)

X를 껌 판매량, Y를 범죄 발생건수라고 하자. 시간에 따라 관측된 두 변수들 사이의 상관계수를 구해보면 매우 상관이 높은 것으로 나타날 것이다. 이는 두 변수사이에 밀접한 관계가 있어서라기보다는 시간이 지남에 따라 인구가 증가하면서 껌의 판매량이 늘고 범죄발생건수도 증가했기 때문이라고 볼 수 있다. 따라서 X와 Y의 순수한 상관관계를 구하기 위해서는 X와 Y에서 시간의 효과를 제거한 후 상관계수를 구해야 할 것이다. 이를 **부분상관계수**(partial correlation coefficient)라고 한다.

* X: 껌 판매량, Y: 범죄발생건수 , Z : 시간

Z의 효과를 배제한 후의 X와 Y의 부분상관계수

$$\rho_{XY,Z} = \frac{E\{[X - E(X|Z)] \cdot [Y - E(Y|Z)]\}}{\sqrt{E\{[X - E(X|Z)]^2 \cdot E[Y - E(Y|Z)]^2}}$$

조건부 기댓값 $E(X|Z)$ 는 X를 Z에 회귀시킨 최적선형 예측값으로 X가 Z에 의해 설명되는 부분

조건부 기댓값 $E(Y|Z)$ 는 Y를 Z에 회귀시킨 최적선형 예측값으로 Y가 Z에 의해 설명되는 부분

$$X^* = X - E(X|Z): X를\ Z에\ 회귀시킨\ 후의\ 잔차$$

$$Y^* = Y - E(Y|Z): Y를\ Z에\ 회귀시킨\ 후의\ 잔차$$

X^*, Y^*: 원래변수 X,Y가 간직하고 있던 정보 중에서 Z와 무관한 부분

X^*, Y^*의 상관계수 ρ_{X^*,Y^*} 는 변수Z에 관하여 수정한 후의 X와 Y의 부분상관계수 $\rho_{XY,Z}$ 와 같아진다.

X^*, Y^* 는 Z와는 무관한 변수 X, Y의 순수한 상관계수를 의미

✓ 부분자기상관계수(partial correlation); PACF ϕ_{kk}

: $Z_t, Z_{t+1}, \dots, Z_{t+k-1}, Z_{t+k}$ 가 관측되었을 때, k시차만큼 떨어진 Z_t, Z_{t+k} 의 순수한 상관관계를 나타냄

➔ Z_t, 와 Z_{t+k} 에서 $Z_{t+1}, Z_{t+2} \dots, Z_{t+k-1}$ 의 효과를 제거한 후의 상관계수

정상확률과정 Z_t, $(Z_t) = 0$

Z_t 를 $Z_{t+1}, Z_{t+2} \dots, Z_{t+k-1}$ 에 회귀시킨 최적선형 예측

$$E(Z_t | Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}) = \alpha_1 Z_{t+1} + \alpha_2 Z_{t+2} + \cdots + \alpha_{k-1} Z_{t+k-1}$$

Z_{t+k} 를 $Z_{t+k-1}, Z_{t+k-2} \dots, Z_{t+1}$ 에 회귀시킨 최적선형 예측

$$E(Z_{t+k} | Z_{t+k-1}, \dots, Z_{t+1}) = \beta_1 Z_{t+k-1} + \beta_2 Z_{t+k-2} + \cdots + \beta_{k-1} Z_{t+1}$$

PACF $\phi_{kk} = Corr\{Z_t^*, Z_{t+k}^*\}$

$$Z_t^* = Z_t - (\alpha_1 Z_{t+1} + \alpha_2 Z_{t+2} + \cdots + \alpha_{k-1} Z_{t+k-1})$$
$$Z_{t+k}^* = Z_{t+k} - (\beta_1 Z_{t+k-1} + \beta_2 Z_{t+k-2} + \cdots + \beta_{k-1} Z_{t+1})$$

단, $\alpha_i, \beta_i \ (1 \le i \le k-1)$ *는*

$E\big(Z_t - (\alpha_1 Z_{t+1} + \alpha_2 Z_{t+2} + \cdots + \alpha_{k-1} Z_{t+k-1})\big)^2$, $E\big(Z_{t+k} - (\beta_1 Z_{t+k-1} + \beta_2 Z_{t+k-2} + \cdots + \beta_{k-1} Z_{t+1})\big)^2$ *를*

최소로하는 최소제곱추정량

✓ 표본부분자기상관함수(sample partial autocorrelation function : SPACF) ; $\widehat{\phi_{kk}}$

4. AR모형(Auto Regressive Model) : 자기회귀 모형

(1) 정의

- 현재의 관측값을 과거 관측값들의 함수형태로 나타냄
- P시점 전까지 관측값의 선형결합으로 표현
- 자기자신을 과거에 회귀시키기 때문에 이런 표현을 쓴다
- ε_t 은 평균 0, 분산σ_ε^2 를 갖는 **백색잡음**

$$Z_t - \mu = \phi_1(Z_{t-1} - \mu) + \phi_2(Z_{t-2} - \mu) + \cdots + \phi_p\big(Z_{t-p} - \mu\big) + \varepsilon_t$$
$$= \sum_{j=1}^{p} \phi_j \big(Z_{t-j} - \mu\big) + \varepsilon_t$$

이때, $\dot{Z}_t = Z_t - \mu$ 라고 한다면,

$$\dot{Z}_t = \sum_{j=1}^{p} \phi_j \dot{Z}_{t-j} + \varepsilon_t$$

- 후항연산자($\mathrm{BX_t} = X_{t-1}$) 를 사용하여 표현 -> 간단하게 표현 가능

$$\phi(B)(Z_t - \mu) = \phi(B)\dot{Z}_t = \varepsilon_t$$

이때, $\phi(\mathrm{B}) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ 라고하고 **특성함수(characteristic function)**라고 부름

✓ AR(p)모형의 조건

- **정상성**과 **인과성**을 만족시켜야 함
- 정상성 : 시계열의 통계적 특성이 시점에 의존하지 않아야 된다는 특성
- 인과성 : t시점의 관측값이 과거 시점의 오차항으로 설명될 수 있다는 특성

➔ p차 모형을 증명하기는 어려운 부분이 있으므로, AR(1)모형을 통해 알아보자

(2) AR(1)

- 마코프(Markov)과정이라도고 불림
- 1시차 전의 확률변수가 1차 자기회귀모형의 독립변수로 사용
- Φ 는 자기회귀모형의 계수로 우리가 자료를 바탕으로 추정해야할 대상
- ε_t는 백색잡음으로서, 모든 시점 t에 대해 $\mathrm{E}(\varepsilon_t) = 0$고 $\mathrm{Var}(\varepsilon_t) = \sigma^2$으로 일정하며, 서로 다른 시점 s, t에서 $\mathrm{Cov}(\varepsilon_t, \varepsilon_s) = 0$을 만족

1) 표현

$$\dot{\mathrm{Z}}_\mathrm{t} = \phi Z_{t-1} + \varepsilon_t \text{ 또는 } \mathrm{Z_t} = \delta + \phi Z_{t-1} + \varepsilon_t$$

$$\phi(\mathrm{B}) = 1 - \phi B \text{ 를 이용한다면 } \phi(B)\dot{Z}_t = \varepsilon_t$$

2) 인과성 조건

표현식을 다시 써보면

$$\dot{\mathrm{Z}}_\mathrm{t} = \phi\left(\phi \dot{Z}_{t-2} + \varepsilon_{t-1}\right) + \varepsilon_t$$

$$= \phi^2 \left(\phi \dot{Z}_{t-3} + \varepsilon_{t-2}\right) + \varepsilon_t + \phi\varepsilon_{t-1}$$

$$\vdots$$

$$= \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \cdots = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$$

① $|\phi| < 1$

$Z_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$이므로, 과거시점의 선형결합으로 표현이 된다.

➔ 인과성 충족

② $|\phi| = 1$

$$Z_t = Z_{t-1} + \varepsilon_t \quad or\ Z_t = -Z_{t-1} + \varepsilon_t$$

➔ 확률보행모델(Randomwalk Process)이 된다. (정상성조건과 함께 설명)

③ $|\phi| > 1$

$$Z_{t+1} = \phi_1 Z_t + \varepsilon_{t+1}$$

$$\phi_1 Z_t = Z_{t+1} - \varepsilon_{t+1}$$

$$Z_t = \frac{1}{\phi_1} Z_{t+1} - \frac{1}{\phi_1}\varepsilon_{t+1} = \frac{1}{\phi_1}\left(\frac{1}{\phi_1} Z_{t+2} - \frac{1}{\phi_1}\varepsilon_{t+2}\right) - \frac{1}{\phi_1}\varepsilon_{t+1}$$

$$= \frac{1}{\phi_1}\left(\frac{1}{\phi_1} Z_{t+3} - \frac{1}{\phi_1}\varepsilon_{t+3}\right)\frac{1}{\phi_1^2}\varepsilon_{t+2} - \frac{1}{\phi_1}\varepsilon_{t+1} = \cdots$$

$$\left(\frac{1}{\phi_1}\right)^{m} Z_{t+j} - \sum_{j=1}^{m}\left(\frac{1}{\phi_1}\right)^{j} \varepsilon_{t+j}$$

➔ $m \to \infty$ 이면 전자항은 0으로 수렴하여 후자항만 남게된다. 후자항이 과거 오차항이 아니라 미래의 오차항에만 의존하므로 인과성을 충족시키지 못한다.

3) 정상성 조건

정상성 조건 : 1) 평균이 일정, 2) 분산이 일정, 3) 공분산이 시차에만 의존

분산을 구해보면,

$$\gamma_0 = Var(Z_t) = E[(Z_t - \mu)^2] = E\left(\dot{Z}_t^2\right), \quad (E(\varepsilon_t \varepsilon_s) = 0, t \neq s)$$

$$= \sum_{j=0}^{\infty} \phi^{2j} E\left(\varepsilon_{t-j}^2\right) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \phi^{2j}$$

자기공분산을 구해보면,

$$\gamma_k = Cov(Z_t, Z_{t+k}) = E[(Z_t - \mu)(Z_{t+k} - \mu)] = E(\dot{Z}_t \dot{Z}_{t+k})$$
$$= E\left[\left(\sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}\right)\left(\sum_{i=0}^{\infty} \phi^i \varepsilon_{t+k-i}\right)\right]$$
$$= \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \phi^j \phi^{j+k} = \phi^k \sigma_\varepsilon^2 (= \phi^k \gamma_0)$$

AR(1)모형이 정상시계열이기 위해선, 분산과 공분산이 유한이어야 한다.

즉, $\sum_{j=0}^{\infty} \phi^{2j} < \infty$ 이어야 하므로, $|\phi| < 1$ 를 만족해야 한다.

이 조건이 만족되면, $\gamma_0 = \frac{\sigma_\varepsilon^2}{(1-\phi^2)}$, $\gamma_k = \phi^k \gamma_0$, $k = 1, 2, \ldots$

이 되므로, γ_k 가 시차 k만의 함수가 된다.

➔ AR(1)의 정상성조건 : $|\phi| < 1$

➔ $1 - \phi_1 B = 0$ 을 B에 관해서 풀면 특성함수의 해는 $B = \frac{1}{\phi_1}$ 이 되고, $|\phi_1| < 1$ 은 $\left|\frac{1}{\phi_1}\right| > 1$ 과 같아진다. 이는 '$\phi(B) = 0$ 의 근의 절대값이 1보다 커야한다'와 동치이다.

<참고> 확률보행과정(Random walk Process)

- 만약 $\phi = 1$ 이면 $Z_t = Z_{t-1} + \varepsilon_t$, $(\varepsilon_t \sim WN(0, \sigma^2))$, $Z_0 = 0)$ 이되며, 이는 확률보행과정(Random Walk)라고 한다.
- 원점($Z_0 = 0$)에서 출발하여 $Z_t = \varepsilon_t$ 이고, $Z_t = \sum_{i=1}^{t} \varepsilon_i$ 는 t시간후의 위치를 나타냄
- 시점 t에서 어떤 사람이 임의의 방향으로 움직이는 보폭이라고 생각할 수 있음
- 대표적 비정상시계열

✓ 평균 : 일정

$$E(Z_t) = E(Z_{t-1} + \varepsilon_t) = E(Z_{t-2} + \varepsilon_{t-1} + \varepsilon_t)$$
$$= \cdots = E(Z_0 + \varepsilon_t + \varepsilon_{t-1} + \cdots + \varepsilon_1) = 0$$

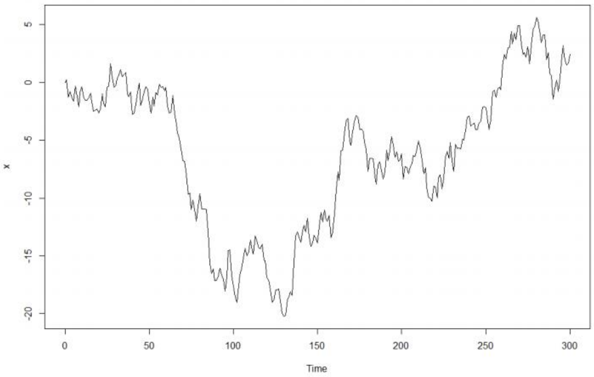
✓ 분산 : t에 의존

$$Var(Z_t) = Var\left(\sum_{i=1}^{t} \varepsilon_i\right) = t\sigma_\varepsilon^2$$

✓ 자기공분산 : t에 의존

$$\gamma_k = Cov(Z_t, Z_{t+k}) = E\left[\left(\sum_{i=1}^{t} \varepsilon_i\right)\left(\sum_{j=1}^{t+k} \varepsilon_j\right)\right] = E\left(\sum_{i=1}^{t} \varepsilon_i^2\right) = t\sigma_\varepsilon^2$$

➔ 분산과 자기공분산이 시간의 함수이므로 정상확률과정이 아님



4) AR(1)의 ACF

AR(1) 표현식에서 양변에 $\dot{Z}_{t-k}$ 를 곱한 후 기대값을 취한다

$$\mathrm{E}\left(\dot{Z}_t \dot{Z}_{t-k}\right) = \phi E\left(\dot{Z}_{t-1}\dot{Z}_{t-k}\right) + \mathrm{E}(\varepsilon_t Z_{t-k})$$

$$\mathrm{k} \geq 1 \text{ 일 때}, \mathrm{E}(\varepsilon_t Z_{t-k}) = 0 \quad \left(Z_{t-k} = \mu + \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-k-j} \textit{ 이므로}, \varepsilon_t \textit{와는 관련이 없음}\right)$$

$$\gamma_k = \phi\gamma_{k-1} = \phi^2\gamma_{k-2} = \cdots = \phi^k\gamma_0 = \phi^k \frac{\sigma_\epsilon^2}{1-\phi^2}, \quad k \geq 1$$

$$\text{k-시차 ACF : } \rho_k = \frac{\gamma_k}{\gamma_0} = \phi^k, \quad k \geq 1$$

➔ 이때, 정상시계열이라면 $|\phi| < 1$ 이다.

➔ $\phi > 0$ 인경우에는 지수적으로 감소

➔ $\phi < 0$인 경우에는 양의 값과 음의 값을 번갈아가지며 지수적으로 감소

5) AR(1)의 PACF

① k=1

$$Z_t = \phi_1 Z_{t-1} + \varepsilon_t$$

$$\phi_{11} = \rho(1)$$

② k=2

$$\phi_{22} = Corr\left[\left(Z_t - \widehat{Z_t}\right), \left(Z_{t-2} - \widehat{Z_{t-2}}\right)\right] = Corr[(Z_t - \phi_1 Z_{t-1}), (Z_{t-2} - \phi_1 Z_{t-1})]$$

$$= \frac{\mathrm{Cov}[(Z_t - \phi_1 Z_{t-1}), (Z_{t-2} - \phi_1 Z_{t-1})]}{\sqrt{Var(Z_t - \phi_1 Z_{t-1}) \cdot Var(Z_{t-2} - \phi_1 Z_{t-1})}}$$

$$= \frac{\mathrm{Cov}[(Z_t - \phi_1 Z_{t-1}), (Z_{t-2} - \phi_1 Z_{t-1})]}{\sqrt{E[(Z_t - \phi_1 Z_{t-1})^2] \cdot E[(Z_{t-2} - \phi_1 Z_{t-1})^2]}} = \frac{\mathrm{Cov}[(Z_t - \phi_1 Z_{t-1}), (Z_{t-2} - \phi_1 Z_{t-1})]}{\sqrt{(\gamma_0 - 2\phi_1\gamma_1 + \phi^2\gamma_0)^2}}$$

$$= \frac{\mathrm{Cov}[(Z_t - \phi_1 Z_{t-1}), (Z_{t-2} - \phi_1 Z_{t-1})]}{\gamma_0 - 2\phi_1\gamma_1 + \phi^2\gamma_0} = \frac{\gamma_2 - \phi_1\gamma_1 - \phi_1\gamma_1 + \phi^2\gamma_0}{\gamma_0 - 2\phi_1\gamma_1 + \phi^2\gamma_0}$$

$$= \frac{\rho_2 - 2\phi_1\rho_1 + \phi_1^2}{1 - 2\phi_1\rho_1 + \phi_1^2} = \frac{\phi_1^2 - 2\phi_1^2\rho_1 + \phi_1^2}{1 - 2\phi_1^2 + \phi_1^2} = 0$$

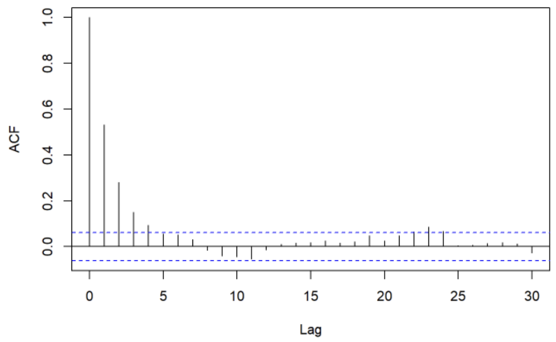
③ k=3이후

$$\phi_{kk} = 0$$

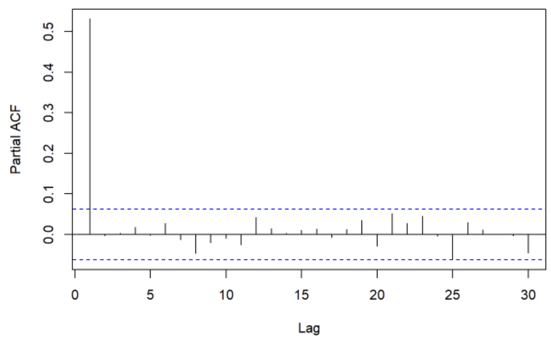
➔ $\phi_{kk} = \begin{cases} \rho_1 = \phi, & k = 1 \\ 0, & k = 2 \end{cases}$

➔ 시차1에서만 ϕ 의 부호에 따라 0이 아니고 2이상의 시차에서는 0이 됨

[AR(1)모형의 ACF]



[AR(1)모형의 PACF]



<참고>

Z_{k+1}를 $Z_1, Z_2, \dots, Z_k$로 표현하는 회귀하는 식 $Z_{k+1} = \phi_{k1}Z_k + \phi_{2k}Z_{k-1} + \cdots\ \phi_{kk}Z_1$을 생각해보자

오차항 $(Z_{k+1}$-$\phi_{k1}Z_k - \phi_{2k}Z_{k-1} - \cdots\ \phi_{kk}Z_1)$의 제곱 의 평균을 최소로하는 $\phi' = (\phi_{k1}, \phi_{k2}, \dots, \phi_{kk})$를 구해 그 값을 원래 회귀식에 넣으면

$\mathrm{argmin}E(Z_{k+1}$-$\phi_{k1}Z_k - \phi_{2k}Z_{k-1} - \cdots\ \phi_{kk}Z_1)^2$

$$\widehat{Z_{k+1}} = \phi_{k1}Z_k + \phi_{2k}Z_{k-1} + \cdots\ \phi_{kk}Z_1$$

각 항의 계수인 $\phi_{k1}, \phi_{k2}, \dots, \phi_{kk}$ 의 의미는 나머지 항이 고정되어 있을 때 Z_{k+1}과 Z들의 관계이다. 회귀식 해석과 상통한다.

만약, AR(p)모형을 생각한다면, $Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots\ \phi_p Z_{t-p} + \varepsilon_t$이다. 애초에 Z_{k+1}을 p시점 이전 값인 $Z_{k+1-p}, \dots\ Z_k$로 밖에 표현할 수 없으므로

$$\widehat{Z_{k+1}} = \phi_{k1}Z_k + \phi_{k2}Z_{k-1} + \cdots\ \phi_{kp}Z_{k+1-p} + 0Z_{k-p} + \cdots 0Z_1 \text{ 이 된다.}$$

따라서 p이후로 PACF은 0이된다.

(3) AR(p) 모형

1) 표현

$$Z_t - \mu = \phi_1(Z_{t-1} - \mu) + \phi_2(Z_{t-2} - \mu) + \cdots + \phi_p(Z_{t-p} - \mu) + \varepsilon_t$$
$$\phi(B)(Z_t - \mu) = \phi(B)\dot{Z}_t = \varepsilon_t$$

- ε_t 는 백색잡음

2) 정상성 조건

- $\phi(X) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p = 0$ 의 근들이 단위원 밖에 있어야함
- $\phi(X) = 0$ 의 근의 절댓값이 1보다 커야함

3) ACF, PACF

- ACF : 지수함수 또는 사인함수와 같은 곡선의 형태를 가져 점차 감소하는 모양을 가짐
- PACF : p차 이후 모두 0이 되어 절단된 모양을 가짐

5. MA모형(Moving Average Model) : 이동평균 모형

(1) 정의

- 현재의 관측값을 과거시점오차들의 함수형태로 나타냄
- q시점 전까지의 오차들의 선형결합으로 표현
- θ 는 이동평균과정의 계수로 자료를 바탕으로 추정해야할 대상
- 이때, $\{\varepsilon_t\} \sim WN(0, \sigma^2)$ 이고 θ 는 각 항의 계수를 의미

$$\dot{Z}_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q}$$

후항연산자를 사용하여 표현하면,

$$Z_t = \varepsilon_t - B\theta_1\varepsilon_t - B^2\theta_2\varepsilon_t - \cdots - B^q\theta_q\varepsilon_t$$
$$= (1 - B\theta_1 - B^2\theta_2 - \cdots - B^q\theta_q)\varepsilon_t$$
$$Z_t = \theta(B)\varepsilon_t$$

이때, $\theta(B) = 1 - B\theta_1 - B^2\theta_2 - \cdots - B^q\theta_q$ 를 특성함수라고 한다.

✓ MA(q)모형의 조건

- 정상성, 인과성 : 과거시점 오차들의 선형결합이기 때문에 항상 만족
- 가역성(invertibility)라는 새로운 조건을 만족시켜야 함

➔ q차모형을 증명하기는 어려운 부분이 있으므로, MA(1)모형을 통해 알아보자

<참고> MA(q)과정의 정상성

MA(q)과정 :

$$\dot{Z}_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q}$$

✓ 평균

$$E(Z_t) = \mu$$

✓ 분산

$$Var(Z_t) = \sigma^2\left(1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2\right) < \infty$$

✓ 자기공분산

$$\gamma_k = Cov(Z_t, Z_{t+k}) = E[(\varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q})(\varepsilon_{t+k} - \theta_1\varepsilon_{t+k-1} - \cdots - \theta_q\varepsilon_{t+k-q})]$$

이때, $E(\varepsilon_t^2) = \sigma^2$, $E(\varepsilon_s\varepsilon_t) = 0 \ (s \neq t)$ 이기 때문에, 시간t에 따른 함수가 아니다. (k>q이후에는 0이나온다.)

(2) MA(1)

- 1시차전의 오차항이 1차 이동평균모형의 독립변수로 사용
- ε_t는 백색잡음으로서, 모든 시점 t에 대해 $E(\varepsilon_t) = 0$고 $Var(\varepsilon_t) = \sigma^2$으로 일정하며, 서로 다른 시점 s, t에서 $Cov(\varepsilon_t, \varepsilon_s) = 0$을 만족

1) 표현

$$\dot{Z}_t = \varepsilon_t - \theta_1\varepsilon_{t-1}$$

후항연산자를 사용한다면, $= (1 - B\theta)\varepsilon_t$

2) 정상성, 인과성 조건

① 평균

$$E(Z_t) = \mu$$

② 분산

$$Var(Z_t) = E[(\varepsilon_t - \theta_1\varepsilon_{t-1})(\varepsilon_t - \theta_1\varepsilon_{t-1})] = \sigma^2(1 + \theta_1^2) < \infty$$

③ 자기공분산

$$\gamma_k = Cov(Z_t, Z_{t-k}) = E[(\varepsilon_t - \theta_1\varepsilon_{t-1})(\varepsilon_{t-k} - \theta_1\varepsilon_{t-k-1})]$$

$$= \begin{cases} (1+\theta^2)\sigma_\varepsilon^2 = \sigma_Z^2, & k=0 \\ -\theta\sigma_\varepsilon^2, & k=1 \\ 0, & k \geq 2 \end{cases}$$

➔ 정상성을 항상 만족한다

➔ 인과성 : MA는 과거 오차항들의 선형결합이기에 항상 만족한다.

3) 가역성 조건

$$\gamma_k = \begin{cases} (1+\theta^2)\sigma_\varepsilon^2 = \sigma_Z^2, & k=0 \\ -\theta\sigma_\varepsilon^2, & k=1 \\ 0, & k \geq 2 \end{cases}$$

θ 대신에 $\frac{1}{\theta}$를 대입하면, $\gamma_k = \begin{cases} \left(1+\frac{1}{\theta}^2\right)\sigma_\varepsilon^2 = \sigma_Z^2, & k=0 \\ -\frac{1}{\theta\sigma_\varepsilon^2}, & k=1 \\ 0, & k \geq 2 \end{cases}$

이를 바탕으로 ACF를 구해본다면, $\rho_k = \begin{cases} \frac{\theta}{1+\theta^2}, & k=1 \\ 0, & k \geq 2 \end{cases}$ 이다.

$\theta = \theta$ 와 $\theta = \frac{1}{\theta}$ 일 때의 경우 값이 동일하게 나온다.

ACF만을 가지고는 어떤 모형을 적합시키는 것이 좋은지 알 수 없다. 따라서 ACF모형 사이의 일대일 관계를 성립시키기 위해 제약조건을 가해주어야 한다.

즉, 정상성조건과 유사하게 이동평균모수 θ 에 제약조건을 줌으로써 ACF와 모형사이에 일대일 관계가 성립하게끔 해주면 모형 선택이 확실해진다.

이 제약조건이 '가역성'이다.

MA(1)모형의 경우 $\theta(B) = 1 - B\theta = 0$ 의 근의 절대값이 1보다 클 조건인 $|\theta| < 1$ 이 그 조건이다.

➔ MA(1)의 가역성조건 : $|\theta| < 1$

4) MA(1)의 ACF

$$\gamma_k = Cov(Z_t, Z_{t-k}) = E[(\varepsilon_t - \theta_1\varepsilon_{t-1})(\varepsilon_{t-k} - \theta_1\varepsilon_{t-k-1})]$$

$= \begin{cases} (1+\theta^2)\sigma_\varepsilon^2 = \sigma_Z^2, & k=0 \\ -\theta\sigma_\varepsilon^2, & k=1 \\ 0, & k \geq 2 \end{cases}$ 을 γ_0로 나누면,

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} -\dfrac{\theta}{1+\theta^2}, & k = 1 \\ 0, & k \geq 2 \end{cases}$$

➔ 시차1일때만 0이 아니고, 시차 2이상부터 0이 됨

5) MA(1)의 PACF

① k=1

$$Z_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

$$\phi_{11} = \rho(1) = -\frac{\theta}{1+\theta^2}$$

② k=2

$$\phi_{22} = Corr\left[\left(Z_t - \widehat{Z_t}\right)\left(Z_{t-2} - \widehat{Z_{t-2}}\right)\right]$$

$$\frac{\mathrm{Cov}[(Z_t - \rho_1 Z_{t-1})(Z_{t-2} - \rho_1 Z_{t-1})]}{\sqrt{Var(Z_t - \rho_1 Z_{t-1})Var(Z_{t-2} - \rho_1 Z_{t-1})}}$$

$$= \frac{\mathrm{Cov}[(Z_t - \rho_1 Z_{t-1})(Z_{t-2} - \rho_1 Z_{t-1})]}{\sqrt{E(Z_t - \rho_1 Z_{t-1})^2 \cdot E(Z_{t-2} - \rho_1 Z_{t-1})^2}} = \frac{\mathrm{Cov}[(Z_t - \rho_1 Z_{t-1})(Z_{t-2} - \rho_1 Z_{t-1})]}{\sqrt{(\gamma_0 - 2\rho_1\gamma_1 + \rho_1^2\gamma_0)^2}}$$

$$= \frac{\mathrm{E}[(Z_t - \rho_1 Z_{t-1})(Z_{t-2} - \rho_1 Z_{t-1})]}{\gamma_0 - 2\rho_1\gamma_1 + \rho_1^2\gamma_0} = \frac{\gamma_2 - 2\rho_1\gamma_1 + \rho_1^2\gamma_0}{\gamma_0 - 2\rho_1\gamma_1 + \rho_1^2\gamma_0} = \frac{\rho_2 - 2\rho_1^2 + \rho_1^2}{1 - 2\rho_1^2 + \rho_1^2}$$

$$= \frac{-\rho_1^2}{1-\rho_1^2} = \frac{-\left(-\frac{\theta}{1+\theta^2}\right)^2}{1-\left(-\frac{\theta}{1+\theta^2}\right)^2} = \frac{-\theta^2}{1+\theta_2+\theta^4}$$

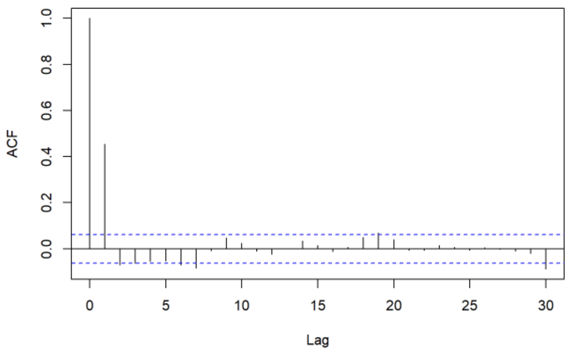
같은 방식으로 계산하면 다음과 같은 식을 얻을 수 있다.

$$\phi_{kk} = \frac{-(-\theta)^k}{1+\theta_2+\cdots+\theta^{2k}} = \frac{-(-\theta)^k(1-\theta^2)}{1-\theta^{2(k+1)}}, \quad k \geq 1$$

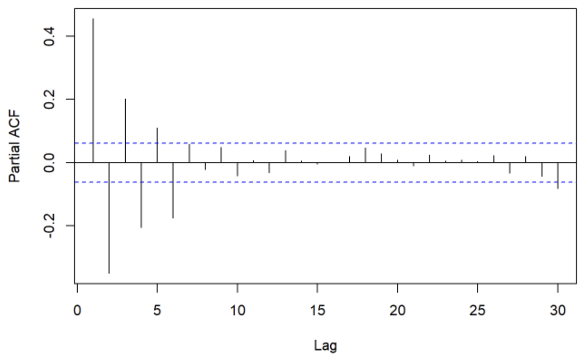
➔ 시차가 증가할수록 지수적으로 감소하는 모양을 가짐

➔

[MA(1)모형의 ACF]



[MA(1)모형의 PACF]



(3) MA(q)

1) 표현

$$\dot{Z}_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q}$$
$$Z_t = \theta(B)\varepsilon_t$$
이때, $\theta(B) = 1 - B\theta_1 - B^2\theta_2 - \cdots - B^q\theta_q$

- ε_t 는 백색잡음

2) 가역성조건

- $\theta(B) = 1 - B\theta_1 - B^2\theta_2 - \cdots - B^q\theta_q = 0$ 의 근이 단위원 밖에 있어야 함
- $\theta(B) = 0$ 의 근의 절댓값이 1보다 커야함

3) ACF, PACF

- ACF : q차 이후 모두 0이 되어 절단된 모양을 가짐
- PACF : 지수적으로 감소 또는 소멸하는 삼각함수의 혼합모양을 가짐

6. AR모형과 MA모형의 쌍대성

(1) 쌍대성(Duality)

- 유한차수의 AR모형을 무한차수의 MA모형으로, 유한차수의 MA모형을 무한차수의 AR모형을 나타낼 수 있는 성질
- 이때 AR은 정상성, 인과성을 만족하고, MA는 가역성을 만족함

➔ AR(1), MA(1)모형으로 이를 확인해보자

1) AR(1) → MA(∞)

$$\mathrm{Z_t} = \phi_1 Z_{t-1} + \varepsilon_t$$

$$(1-\phi_1 B)Z_t = \varepsilon_t$$

$$\mathrm{Z_t} = \frac{1}{1-\phi_1 B}\varepsilon_t$$

이때, 정상성을 만족한다면 $|\phi_1| < 1$이므로 등비급수로 표현가능

$$\mathrm{Z_t} = (1 + \phi_1 B + \phi_1^2 B^2 + \phi_1^3 B^3 + \cdots)\varepsilon_t$$

$$= \varepsilon_t + \phi_1\varepsilon_{t-1} + \phi_1^2\varepsilon_{t-2} + \phi_1^3\varepsilon_{t-3} + \cdots = MA(\infty)$$

2) MA(1) → AR(∞)

$$\mathrm{Z_t} = \varepsilon_t - \theta\varepsilon_{t-1} = (1-\theta_1 B)\varepsilon_t$$

$$\frac{1}{1-\theta_1 B}Z_t = \varepsilon_t$$

이때 가역성을 갖는다면, $|\theta_1| < 1$이므로 등비급수로 표현가능

$$(1 + \theta_1 \mathrm{B} + \theta_1^2 \mathrm{B}^2 + \theta_1^3 \mathrm{B}^3 + \cdots)\mathrm{Z_t} = \varepsilon_t$$

$$\mathrm{Z_t} = -\theta_1 \mathrm{Z_{t-1}} - \theta_1^2 \mathrm{Z_{t-2}} - \theta_1^3 \mathrm{Z_{t-3}} + \cdots + \varepsilon_t = \mathrm{AR}(\infty)$$

(2) 정리

① 유한차수의 정상 AR과정은 무한차수의 MA과정으로 표현 가능하고, 유한차수의 가역성을 갖는 MA과정은 무한차수의 AR과정으로 표현 가능하다.

② 유한차수의 AR과정의 ACF와 유한차수의 MA과정의 PACF는 **지수적으로 감소하는 형태**를, 유한차수의 AR 과정의 PACF와 유한차수의 MA과정의 ACF는 **절단된 형태**를 갖는다.

③ 유한차수의 AR과정에 대해선 가역성 조건은 필요하지 않으나 정상성을 위해서는 **$\boldsymbol{\phi}(\mathbf{B}) = \mathbf{0}$ 의 근들이 단위원 밖에 존재해야 한다**는 조건이 필요하다. 반면에, 유한차수의 MA과정에 대해선 정상성 조건은 필요하지 않으나 가역성을 위해서는 **$\boldsymbol{\theta}(\mathbf{B}) = \mathbf{0}$의 근들이 단위원 밖에 존재해야 한다**는 조건이 필요하다.

7. ARMA모형 : 자기회귀이동평균모형

- 시계열 데이터를 순수한 자기회귀모형이나 이동평균모형으로만 설명하려면 p나 q값이 너무 커질 수 있는 문제가 발생
- 추정해야할 모수의 개수가 많아지면 일반적으로 추정의 효율성이 떨어지고 해석도 쉽지 않다.

- 자귀회귀부분과 이동평균부분이 동시에 사용된다면 추정해야할 모수의 개수를 줄일 수 있을 것

➔ 자기회귀이동평균(autoregressive-moving average : ARMA)

(1) 표현

- ARMA(p,q)모형
- AR의 정상성과 MA의 가역성 조건을 둘다 만족해야 한다.

$$\mathrm{Z_t} = \phi_1 Z_{t-1} + \cdots + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}$$

$$\left(1 - \phi_1 B - \cdots - \phi_p B^p\right) Z_t = \left(1 - \theta_1 B - \cdots - \theta_q B^q\right) \varepsilon_t$$

$$\phi(\mathrm{B})\mathrm{Z_t} = \theta(B)\varepsilon_t$$

이때, $\phi(\mathrm{B}) = 1 - \phi_1 B - \cdots - \phi_p B^p$ 이고, $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ 이다.

(2) ARMA(1,1)

$$\mathrm{Z_t} = \phi_1 Z_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

$$(1 - \phi \mathrm{B})\mathrm{Z_t} = (1 - \theta B)\varepsilon_t$$

<참고>
ARMA(1,0) = AR(1)
ARMA(0,1) = MA(1)

(3) ARMA모형의 조건

- 정상성과 가역성을 만족해야한다.

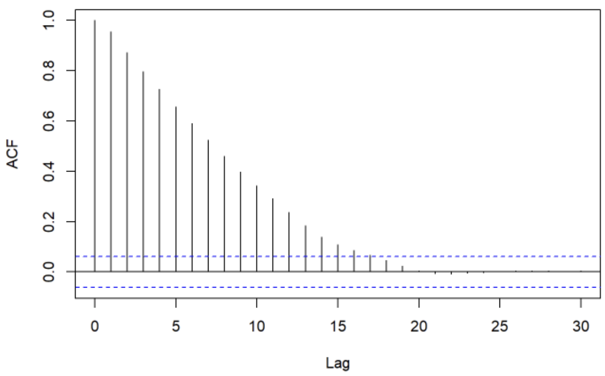
➔ 정상성을 만족하기 위해 $\phi(\mathrm{B}) = 0$ 의 근이 1보다 커야한다.

➔ 가역성을 만족하기 위해 $\theta(\mathrm{B}) = 0$ 의 근이 1보다 커야한다.

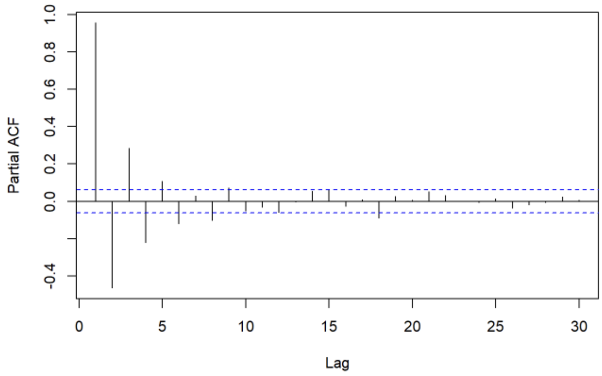
(4) ARMA모형의 ACF, PACF

- ACF : q+1 이후 지수적으로 감소하거나 싸인함수 형태로 소멸하는 모양을 가짐
- PACF : p+1 이후 지수적으로 감소하거나 싸인함수 형태로 소멸하는 모양을 가짐

[ARMA(1,1)모형의 ACF]



[ARMA(1,1)모형의 PACF]



8. 정리

(1) 모형의 정상성과 가역성

	AR(p)	MA(q)	ARMA(p,q)
정상성	조건필요	자체만족	조건필요
가역성	자체만족	조건필요	조건필요

(2) 모형의 ACF와 PACF 패턴

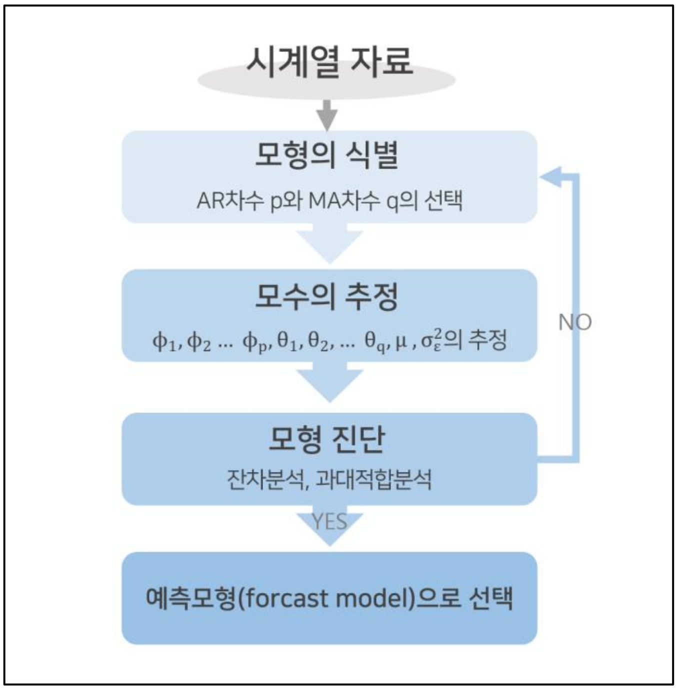
	AR(p)	MA(q)	ARMA(p,q)
ACF	지수적으로 감소	q+1차부터 절단	q+1시점부터 지수적으로 감소
PACF	p+1차부터 절단	지수적으로 감소	p+1시점부터 지수적으로 감소

9. 모형 적합 절차

(1) 흐름

- 모형의 식별(model identification)단계 : 모형의 차수를 결정하는 과정
- 모수의 추정(model estimation)단계 : 식별된 모형의 모수들을 추정하는 과정

  : $\phi_1, \phi_2 \dots \phi_p, \theta_1, \theta_2, \dots \theta_q$, 평균수준 μ, 백색잡음 과정의 분산 σ_ε^2
- 모형의 진단(diagonostics)단계 : 잠정모형(tentative model)의 타당성 검토과정



(2) 모형의 식별

- ARMA(p,1)모형의 p와 q 를 결정하는 단계
- 간결의 원칙(principle of parsimony) 추정할 모수의 개수가 증가하면 최종 예측모형이 복잡해질 뿐만 아니라 추정의 효율성도 떨어 지므로 될 수 있으면 간단한 모형을 선호
- 시계열그림과 표본상관도표(sample correlogram) 이용, 비정상시계열은 정상시계열로 변환(분산안정화변환, 적절한 차분)
- ACF와 PACF를 보고 차수 결정
- 모형 선택의 기준으로 사용되는 통계량(AIC, BIC, SBC) 이용 (여러 모형의 AIC값을 계산해서 가장 작은 값을 갖는 모형 선택)

✓ $\mathrm{AIC} = -2\log \mathrm{L} + 2\mathrm{m}$

✓ $\mathrm{BIC} = -2\mathrm{logL} + \mathrm{mlogn}\ (\mathrm{m} = \mathrm{p} + \mathrm{q} + 1)$

(3) 모수의 추정

- 최대가능도추정법 : $Z_1, Z_2, Z_3, \dots$ 의 결합확률밀도함수인 가능도함수(likelihood function)를 최대로 하는 모수의 추정량을 구하는 방법
- 최소제곱추정법 : 오차제곱합을 최소로 하는 추정법
- 적률추정법 : 모집단의 적률에 대응되는 표본적률의 방정식을 풀어 추정량을 구하는 방법

$$\mu = E[Z_t] \text{의 적률추정량 : } \hat{\mu} = \overline{Z} = \sum_{t=1}^{n} Z_t / n$$

$$\sigma_z^2 = r_0 = E(Z_t - \mu)^2 \text{의 적률추정량 : } \hat{r}_0 = \frac{1}{n} \sum_{t=1}^{n} (Z_t - \overline{Z})^2$$

(4) 모형의 진단

① 잔차분석 : 잔차를 이용하여 모형에 대한 가정이 옳은지 알아보(1주차 클린업)

② 과대적합진단

- 잠정모형에 모수를 추가하여 더 많은 개수의 모수를 포함하는 모형을 적합
- 추가된 모수가 유의하다고 판정되거나 잠정모형의 모수추정값이 과대적합 후의 모수추정값과 큰 차이가 있거나 과대적합된 모형의 잔차들의 분산이 잠정모형의 잔차들의 분산보다 작은 경우 -> 추가된 모수들의 설명력이 있다고 판단, 잠정모형을 새로운 모형으로 대체
- ex) AR(1)모형이 잠정모형으로 선택된 경우 AR(2) 또는 AR(1,1)모형을 자료에 적합

- $\phi(B)(Z_t - \mu) = \theta(B)\varepsilon_t$ 의 기본가정 : $\phi(B)$ 와 $\theta(B)$ 이 공통요인을 포함하지 않음
- 모수과잉(parameter redundancy)을 방지하기 위해 AR과 MA항을 동시에 추가하면 안됨
  ➔ 두개의 모수가 동시에 추가될 경우 공통요인(common factor) 또는 거의 공통인 요인이 모형 의 AR 부분과 MA 부분에 존재할 수 있기 때문에 공통항을 서로 상쇄시켜야 모수과잉의 현상이 일어나지 않음

10. R실습

11. 부록

(1) 단위근

- $Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \cdots - \phi_p Z_{t-p} = \varepsilon_t$ 또는 $(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) Z_t = \varepsilon_t$
- 단, ε_t는 서로 독립이며 평균이 0, 분산이 σ^2인 백색잡음
- 정상성 조건 : $\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) = 0$ 의 근(root)의 절댓값이 1보다 큼
- 비정상 확률과정 : 1보다 크지 않은 근이 존재하는 경우
- 단위근(unit root) : $\phi(B) = 0$ 의 근 중 크기가 1인 근

단위근을 갖는 비정상 확률과정의 예 :

$AR(1) \quad Z_t - \phi Z_{t-1} = \epsilon_t$

정상성 조건 : $\boldsymbol{\phi}(B) = (1 - \phi B) = 0$의 근의 절대값이 1보다 커야 하므로 $|\phi| < 1$

만약 $\phi = 1$, 즉, 단위근을 갖는 경우

$$Z_t - Z_{t-1} = \epsilon_t \Rightarrow Z_t = \sum_{i=1}^{t} \epsilon_t \text{ (단 } Z_0 = 0)$$

➔ 단위근이 존재하는 불안정적인 시계열을 그대로 사용하면 표본 수가 증가함에 따라 회귀계수의 t-값도 증가하여 상관관계가 없 는 변수간에도 매우 강한 상관관계가 있는 것으로 나타나는 **가성회귀**(허구적 회귀 : spurious regression)의 문제가 발생함.

➔ 단위근 검정 : ϕ = 1인지 여부를 판단하는 문제를 단위근 검정이라고 한다. 따라서 단위근 검정은 $H_0 : \phi = 1$를 검정하는 문제

➔ 그래서 단위근검정을 통해서 검정을 통해 -> 차분을 하여 정상성조건을 만족시켜서 - > 정상성을 만족하는 확률과정으로 바꾼다.

(2) Durbin-Levinson 알고리즘

회귀모형을 이용하여 $PACF$를 구하는 방법

종속변수 Z_{t+k} 를 $Z_{t+k-1}, Z_{t+k-2}, \ldots, Z_t$에 회귀시킨 회귀식

$$Z_{t+k} = \phi_{k1} Z_{t+k-1} + \phi_{k2} Z_{t+k-2} + \cdots + \phi_{kk} Z_t + \epsilon_{t+k}$$

양변에 $Z_{t+k-j} \, (j \geq 1)$ 를 곱한 후 기대값을 구하여 보면

$$\gamma_j = \phi_{k1}\gamma_{j-1} + \phi_{k2}\gamma_{j-2} + \cdots + \phi_{kk}\gamma_{j-k}$$

양변을 γ_0 로 나누면

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \cdots + \phi_{kk}\rho_{j-k}$$

$j = 1, 2, \ldots, k$에 대하여

$$\begin{cases} \rho_1 = \phi_{k1}\rho_0 + \phi_{k2}\rho_1 + \cdots + \phi_{kk}\rho_{k-1} \\ \rho_2 = \phi_{k1}\rho_1 + \phi_{k2}\rho_0 + \cdots + \phi_{kk}\rho_{k-2} \\ \quad \vdots \\ \rho_k = \phi_{k1}\rho_{k-1} + \phi_{k2}\rho_{k-2} + \cdots + \phi_{kk}\rho_0 \end{cases}$$

연립방정식을 Cramer 공식을 이용

$$\phi_{11}=\rho_1, \quad \phi_{22}=\frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}=\frac{\rho_2-\rho_1^2}{1-\rho_1^2}, \quad \phi_{33}=\frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}$$

$$\phi_{kk}=\frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}}$$

표본부분자기상관함수(sample partial autocorrelation function: $SPACF$**)** $\hat{\phi}_{kk}$

위의 ϕ_{kk} 식에서 ρ_j 대신에 $SACF$ $\hat{\rho}_j$ 를 대입하여 얻음

k 가 커질 경우 행렬식의 계산을 하는 대신
Durbin-Levinson 알고리즘, Durbin(1960), 이용

$$\hat{\phi}_{11}=\hat{\rho}_1$$

$$\hat{\phi}_{k+1,k+1}=\frac{\hat{\rho}_{k+1}-\sum_{j=1}^{k}\hat{\phi}_{kj}\hat{\rho}_{k+1-j}}{1-\sum_{j=1}^{k}\hat{\phi}_{kj}\hat{\rho}_j}$$

$$\hat{\phi}_{k+1,j}=\hat{\phi}_{kj}-\hat{\phi}_{k+1,k+1}\hat{\phi}_{k,k+1-j}, \qquad j=1,2,\ldots,k$$