시계열자료분석팀

5팀

김민주

박준영

곽동길

강서진

황호성

INDEX

0

1주차 복습

1주차 복습

정리

시계열 자료

관측치들 간 dependency 有

규칙요소

추세 / 순환 / 계절성

불규칙요소

우연 변동

덧셈 분해

$$X_t = m_t + s_t + Y_t$$

정리

시계열 자료

관측치들 간 dependency 有

규칙요소

추세 / 순환 / 계절성

불규칙요소

우연 변동

분석을 위해서는,

시계열 자료가 '정상성'을 만족해야 함!

덧셈 분해

$$X_t = m_t + s_t + Y_t$$

정리 | 정상성

정상성

시계열 자료의 확률적 성질이 '시차'에만 의존

⋮

현실에서는 주로 약정상성을 이용 !

<약정상성의 조건>

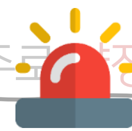
$E[\lvert X_t \rvert]^2 < \infty, \forall t \in Z$	2차적률 有, 시점 t에 관계없이 일정
$E[X_t] = m, \forall t \in Z$	평균 = 상수, 시점 t에 관계없이 일정
$\gamma_x(r,s) = \gamma_x(r+h, s+h),$ $\forall r,s,h \in Z$	공분산은 시차 t에 의존, 시점 t와 무관

정리 | 정상성

정상성

시계열 자료의 확률적 성질이 '시차'에만 의존

현실에서는 주로 약정상성을 이용 !



만족하지 않는다면,

정상화 진행 !

<약정상성의 조건>

$E[\|X_t\|]^2 < \infty, \forall t \in Z$	2차적률 有, 시점 t에 관계없이 일정
$E[X_t] = m, \forall t \in Z$	평균 = 상수, 시점 t에 관계없이 일정
$\gamma_x(r,s) = \gamma_x(r+h, s+h),$ $\forall r,s,h \in Z$	공분산은 시차 t에 의존, 시점 t와 무관

정리 | 정상화

정상화

비정상 시계열을 정상 시계열로 변환

분산 일정 X	평균 일정 X
로그(Log) 변환	회귀 (Regression)
제곱근 (Square Root) 변환	평활 (Smoothing)
Box-Cox 변환	차분 (Differencing)

정리 | 정상화

정상화

비정상 시계열을 정상 시계열로 변환

분산 일정 X

로그(Log) 변환

루트(Square Root) 변환

Box-Cox 변환

평균 일정 X

회귀 (Regression)

평활 (Smoothing)

차분 (Differencing)

남아있는 오차가

정상성을 만족하는지 확인 !

정리 | 정상성 검정

자기공분산함수 (ACVF)

$$\gamma_X(h) = Cov(X_t, X_{t+h}) = E[(X_t - \mu)(X_{t+h} - \mu)]$$

자기상관함수 (ACF)

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = Corr(X_t, X_{t+h}) = \frac{Cov(X_t, X_{t+h})}{\sqrt{Var(X_t)}\sqrt{Var(X_{t+h})}}$$

정리 | 정상성 검정

평균 = 0, 분산 = σ^2

백색잡음 $WN(0, \sigma^2)$

자기상관이 존재하지 않는 시계열

⋮

자기상관 검정	ACF plot
정규성 검정	QQ plot / KS Test / Jarque-Bera Test
정상성 검정	KPSS Test / ADF Test / PP Test

0 1주차 복습

정리 | 정상성 검정

평균 = 0, 분산 = σ^2

백색잡음 $WN(0, \sigma^2)$

자기상관이 존재하지 않는 시계열

모든 백색잡음은 정상 시계열이지만,

모든 정상 시계열이 백색잡음은 아님을 주의

자기상관 검정	[illegible] plot
정규성 검정	QQ plot / KS Test / Jarque-Bera Test
정상성 검정	KPSS Test / ADF Test / PP Test

1

모형의 식별

시계열 모형의 필요성

정상화 과정 이후
남아있는 오차가
WN/IID를 만족해야 함

→

WN/IID를 만족하지 않을 때,
이를 추정하기 위해
시계열 모형 필요

시계열 모형의 필요성

▶ 오차항 Y_t의 공분산 행렬

$$\Gamma = \begin{pmatrix} Cov(Y_1, Y_1) & Cov(Y_1, Y_2) & \cdots & Cov(Y_1, Y_n) \\ \vdots & \vdots & \vdots & \vdots \\ Cov(Y_n, Y_1) & Cov(Y_n, Y_2) & \cdots & Cov(Y_n, Y_n) \end{pmatrix}$$

1 모형의 식별

시계열 모형의 필요성

▶ 오차항 Y_t의 공분산 행렬

$$\Gamma = \begin{pmatrix} Cov(Y_1, Y_1) & Cov(Y_1, Y_2) & \cdots & Cov(Y_1, Y_n) \\ \vdots & \vdots & \vdots & \vdots \\ Cov(Y_n, Y_1) & Cov(Y_n, Y_2) & \cdots & Cov(Y_n, Y_n) \end{pmatrix}$$

⋮

▶ 시차를 이용한 오차항 Y_t의 공분산 행렬

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

정상성 조건 ③ 자기공분산은 시차에만 의존 - 클린업 1주차 참고

시계열 모형의 필요성

▶ 오차항 Y_t가 백색잡음(WN)일 때

$$\Gamma = \begin{pmatrix} \gamma(0) & 0 & \cdots & 0 \\ 0 & \gamma(0) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \gamma(0) \end{pmatrix}$$

만약 오차항 Y_t가 백색잡음이라면,

대각요소를 제외한 모든 요소가 0

시계열 모형의 필요성

▶ 오차항 Y_t가 백색잡음(WN)일 때

$$\Gamma = \begin{pmatrix} \gamma(0) & 0 & \cdots & 0 \\ 0 & \gamma(0) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \gamma(0) \end{pmatrix}$$

만약 오차항 Y_t가 백색잡음이라면,

대각요소를 제외한 모든 요소가 0

➡ 분산만 추정하여 공분산 행렬 구할 수 있음

1 모형의 식별

시계열 모형의 필요성

▶ 오차항 Y_t가 백색잡음(WN)일 때

But 오차항 Y_t가 백색잡음이 아니라면,

확률변수 간 상관관계 존재

▶ 오차항 Y_t가 백색잡음(WN)이 아닐 때

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

1 모형의 식별

시계열 모형의 필요성

▶ 오차항 Y_t가 백색잡음(WN)일 때

But 오차항 Y_t가 백색잡음이 아니라면,

확률변수 간 상관관계 존재

SACVF를 통해 하삼각요소 추정 필요

▶ **오차항 Y_t가 백색잡음(WN)이 아닐 때**

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

시계열 모형의 필요성

◎ 표본자기공분산함수(SACVF)

$$\hat{\gamma}_Y(h) = \frac{1}{n}\sum_{j=1}^{n-h}(Y_j - \bar{Y})(Y_{j+h} - \bar{Y})$$

표본자기공분산함수를 통한 추정의 경우,

시차 h에 따라 추정의 정확도가 달라질 수 있다는 단점 존재

$h \uparrow \rightarrow (n-h) \downarrow$, 정확도 $\downarrow$

시계열 모형의 필요성

◎ 표본자기공분산함수(SACVF)

$$\hat{\gamma}_Y(h) = \frac{1}{n}\sum_{j=1}^{n-h}(Y_j - \bar{Y})(Y_{j+h} - \bar{Y})$$

표본자기공분산함수를 통한 추정의 경우,

시차 h에 따라 추정의 정확도가 달라질 수 있다는 단점 존재

$h\uparrow \rightarrow (n-h)\downarrow$, 정확도 $\downarrow$

남아있는 오차가 IID나 WN이 아닐 때를 추정하기 위해 시계열 모형 필요

시계열 모형의 판단기준

시계열 모형의 판단기준

자기상관함수 ACF
(Autocorrelation Function)

부분자기상관함수 PACF
(Partial Autocorrelation Function)

시계열 모형의 판단기준 | [1] 자기상관함수 (ACF)

자기상관함수 ACF

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = Corr(X_t, X_{t+h}) = \frac{Cov(X_t, X_{t+h})}{\sqrt{Var(X_t)}\sqrt{Var(X_{t+h})}}$$

⋮

시차가 h인 시계열 간의 상관관계 의미

정상성을 만족한다면 시차에만 의존하는 특징

1 모형의 식별

시계열 모형의 판단기준 | [1] 자기상관함수 (ACF)

자기상관함수 ACF

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = Corr(X_t, X_{t+h}) = \frac{Cov(X_t, X_{t+h})}{\sqrt{Var(X_t)}\sqrt{Var(X_{t+h})}}$$

ACF의 특징

1) $p(0) = 1\ (\because \gamma(0) = var(X_t))$

시차가 h인 시계열 간의 상관관계 의미

2) $p(-h) = p(h)$

정상성을 만족한다면 시차에만 의존하는 특징

3) $|\gamma(h)| \leq r(0)\ for\ all\ h\ \in Z\ \rightarrow |p(h)| \leq 1$

시계열 모형의 판단기준 | [2] 부분자기상관함수 (PACF)

부분자기상관함수 PACF

X_t와 X_{t+k}의 상관관계를 구할 때, 그 둘 사이에 존재하는 $\{X_{t+1}, X_{t+2}, \cdots, X_{t+k-1}\}$의 영향을 제외하고 구한 상관계수

부분상관계수

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

1 모형의 식별

시계열 모형의 판단기준 | [2] 부분자기상관함수 (PACF)

부분자기상관함수 PACF

X_t와 X_{t+k}의 상관관계를 구할 때, 그 둘 사이에 존재하는

$\{X_{t+1}, X_{t+2}, \cdots, X_{t+k-1}\}$의 영향을 제외하고 구한 상관계수

더 나은 **부분자기상관함수** 이해를 위해,

예시를 통해 **부분상관계수**를 먼저 알아보자!

1 모형의 식별

시계열 모형의 판단기준 | [2] 부분자기상관함수 (PACF)

부분자기상관함수 PACF

X_t와 X_{t+k}의 상관관계를 구할 때, 그 둘 사이에 존재하는 $\{X_{t+1}, X_{t+2}, \cdots, X_{t+k-1}\}$의 영향을 제외하고 구한 상관계수

부분상관계수

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

시계열 모형의 판단기준 | [2] 부분자기상관함수 (PACF)

부분상관계수 예시

X: 아이스크림 판매량, Y: 범죄발생건수, Z: 인구수

$Z \uparrow \Rightarrow X \uparrow \,\&\, Y \uparrow \Rightarrow X$ 와 Y의 상관계수 $\uparrow$

Z의 영향을 제거한 X 와 Y의 상관계수를 구해야 함

부분상관계수

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

1 모형의 식별

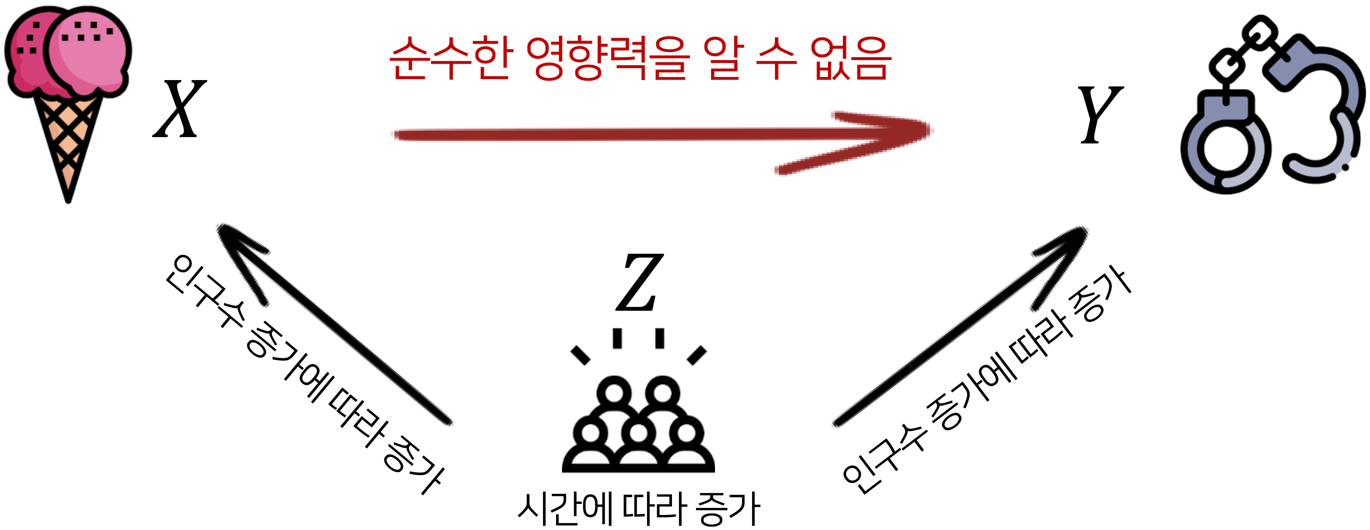
시계열 모형의 판단기준 | [2] 부분자기상관함수 (PACF)

부분상관계수 예시

X: 아이스크림 판매량, Y: 범죄발생건수, Z: 인구수

$Z \uparrow \Rightarrow X \uparrow \& Y \uparrow \Rightarrow X$ 와 Y의 상관계수 $\uparrow$

Z의 영향을 제거한 X 와 Y의 상관계수를 구해야 함



1 모형의 식별

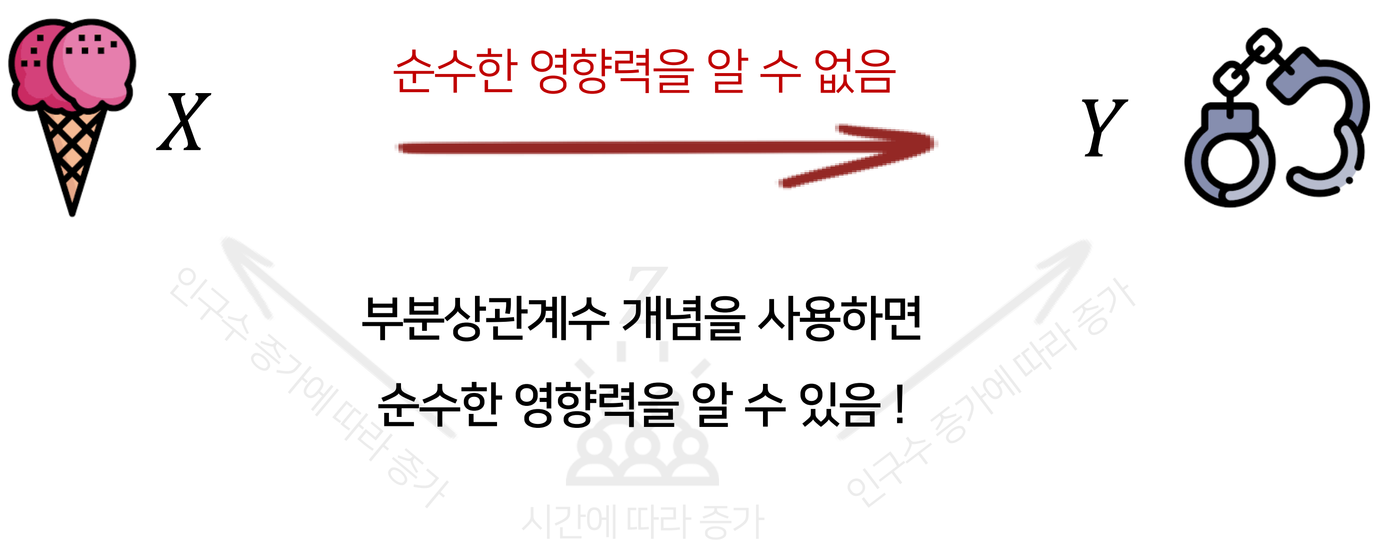
시계열 모형의 판단기준 | [2] 부분자기상관함수 (PACF)

부분상관계수 예시

X: 아이스크림 판매량, Y: 범죄발생건수, Z: 인구수

$Z \uparrow \Rightarrow X \uparrow \& Y \uparrow \Rightarrow X$ 와 Y의 상관계수 $\uparrow$

Z의 영향을 제거한 X 와 Y의 상관계수를 구해야 함



부분상관계수 개념을 사용하면

순수한 영향력을 알 수 있음 !

부분상관계수와 각 항의 의미

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

1) 조건부 기댓값 $E(X|Z)$: X가 Z에 의해 설명되는 부분

2) 조건부 기댓값 $E(Y|Z)$: Y가 Z에 의해 설명되는 부분

3) $X^* = X - E(X|Z)$: X에서 Z의 영향력을 제거하고 남은 부분

4) $Y^* = Y - E(Y|Z)$: Y에서 Z의 영향력을 제거하고 남은 부분

⇒ X와 Y의 부분상관계수: $\rho_{XY,Z} = \rho_{X^*Y^*}$

1 모형의 식별

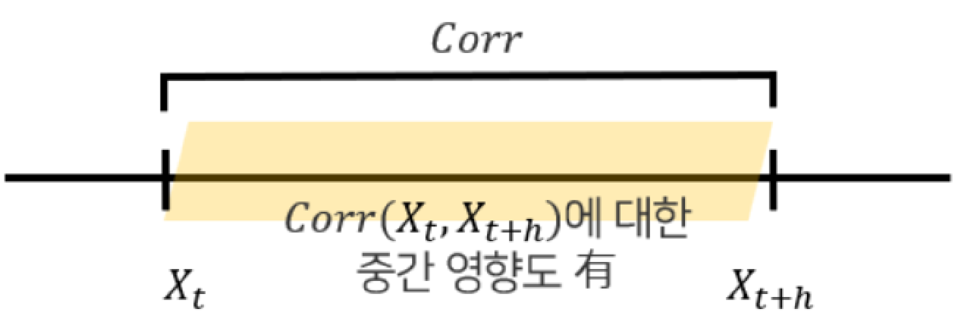
시계열 모형의 판단기준 | [2] 부분자기상관함수 (PACF)

부분자기상관함수 PACF

자기 자신과의 부분상관계수

즉, 중간값들의 효과를 제거한 순수한 X_t와 X_{t+h}의 관계

⋮



시계열 모형의 판단기준 | [2] 부분자기상관함수 (PACF)

부분자기상관함수(PACF)는 일반적으로 $\alpha(k)$, $k = lag$로 표현

부분자기상관함수 PACF 정의

$$\alpha(0) = Corr(X_1, X_1) = 1$$

$$\alpha(1) = Corr(X_1, X_2) = \rho(1)$$

$$\alpha(k) = Corr(X_{k+1} - P_k^* X_{k+1}, X_1 - P_k^* X_1), k \geq 2$$

시계열 모형의 판단기준 | [2] 부분자기상관함수 (PACF)

부분자기상관함수(PACF)는 일반적으로 $\alpha(k)$, $k = lag$로 표현

:

부분자기상관함수 PACF 정의

$$\alpha(0) = Corr(X_1, X_1) = 1$$

$$\alpha(1) = Corr(X_1, X_2) = \rho(1)$$

$$\alpha(k) = Corr(X_{k+1} - P_k^* X_{k+1}, X_1 - P_k^* X_1), k \geq 2$$

$$\Rightarrow P_k^* X_{k+1} = Best\ Linear\ Predictor\ on \{1, X_2, \cdots, X_k\}$$

$$\Rightarrow P_1^* X_{k+1} = Best\ Linear\ Predictor\ on \{1, X_2, \cdots, X_k\}$$

* BLP(Best Linear Predictor)는 뒤에서 다루겠지만,

지금은 $\{1, X_2, \cdots, X_k\}$가 X_1과 X_{k+1}에 미치는 영향 정도로 생각

시계열 모형의 판단기준 | [2] 부분자기상관함수 (PACF)

부분자기상관함수 PACF의 주요 아이디어

중간값들의 영향력을 선형회귀로 추정하여 제거하는 것

$X_1, \cdots, X_{k+1}$을 이용한 X_{k+1} 회귀식

$$X_{k+1} = \phi_{11} X_k + \epsilon_{k+1}$$

$$X_{k+1} = \phi_{21} X_k + \phi_{22} X_{k-1} + \epsilon_{k+1}$$

$$\vdots$$

$$X_{k+1} = \phi_{k1} X_k + \phi_{k2} X_{k-1} + \cdots + \phi_{kk} X_1 + \epsilon_{k+1}$$

시계열 모형의 판단기준 | [2] 부분자기상관함수 (PACF)

부분자기상관함수 PACF의 주요 아이디어

중간값들의 영향력을 선형회귀로 추정하여 제거하는 것

▶ $X_1, \cdots, X_{k+1}$을 이용한 X_{k+1} 회귀식

$$X_{k+1} = \phi_{11} X_k + \epsilon_{k+1}$$

$$X_{k+1} = \phi_{21} X_k + \phi_{22} X_{k-1} + \epsilon_{k+1}$$

$$\vdots$$

$$X_{k+1} = \phi_{k1} X_k + \phi_{k2} X_{k-1} + \cdots + \phi_{kk} X_1 + \epsilon_{k+1}$$

시계열 모형의 판단기준 | [2] 부분자기상관함수 (PACF)

▶ $X_1, \cdots, X_{k+1}$을 이용한 X_{k+1} 회귀식

$$X_{k+1} = \phi_{11} X_k + \epsilon_{k+1}$$

$$X_{k+1} = \phi_{21} X_k + \phi_{22} X_{k-1} + \epsilon_{k+1}$$

$$\vdots$$

$$X_{k+1} = \phi_{k1} X_k + \phi_{k2} X_{k-1} + \cdots + \phi_{kk} X_1 + \epsilon_{k+1}$$

⋮

▶ 오차항을 최소화 하는 방법으로 X_{k+1}의 추정값 BLP 구하기

$$\hat{X}_{k+1} = arg\min_{\phi} E(X_{k+1} - \phi_{k1} X_k - \phi_{k2} \phi_{k-1} - \cdots - \phi_{kk} X_1)^2$$

1 모형의 식별

시계열 모형의 판단기준 | [2] 부분자기상관함수 (PACF)

회귀식의 계수인 ϕ_{kk}는 $\{X_2, \cdots, X_k\}$가 고정되어 있을 때,

X_{k+1}과 X_1간의 선형적인 상관관계를 나타내는 수치

⋮

따라서 PACF는 다음과 같이 정의

$$\alpha(k) = \phi_{kk},\ k \geq 1$$

▶ 오차항을 최소화 하는 방법으로 X_{k+1}의 추정값 BLP 구하기

$$\hat{X}_{k+1} = arg\min_{\phi} E(X_{k+1} - \phi_{k1}X_k - \phi_{k2}\phi_{k-1} - \cdots - \phi_{kk}X_1)^2$$

2

선형 과정

2 선형 과정

Linear process (선형 과정)

선형 과정 Linear process

$\{Z_t\} \sim WN(0, \sigma^2)$들의 선형결합으로 표현된 X_t가 선형 과정

⋮

선형 과정식

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$$

* 이때 선형결합의 계수는 $\sum_j |\psi_j| < \infty$ $(absolutely\ summable)$ 조건을 만족해야 함

Linear process(선형 과정)

시계열을 선형과정으로 표현하는 이유?

후향연산자를 이용한 선형 과정 표현

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j} = \psi(B) Z_t,\ where\ \psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$$

$$\because \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j} = \sum_{j=-\infty}^{\infty} \psi_j B^j Z_t = \left(\sum_{j=-\infty}^{\infty} \psi_j B^j\right) Z_t$$

1) 공분산 계산이 편리함

$$Cov(a_1Z_1 + a_2Z_2 + a_3Z_3, b_2Z_2 + b_3Z_3) = a_2b_2 + a_3b_3$$

$$Cov(a_1Z_1, a_2Z_2 + a_3Z_3\ ) = 0$$

위 식과 같이 index가 동일한 경우만 계산해 공분산 구할 수 있음

2) 해석이 용이하고 추정방법이 잘 발달됨 (Regression/Linear algebra)

3) **정상 확률 과정의 선형 결합이 다시 정상 확률 과정 조건을 만족함**

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$$

4) 약한 의미의 정상 확률 과정이 선형 과정의 합과 결정적 과정으로 표현됨 (Wold decomposition)

* 이때 선형결합의 계수는 $\sum_j |\psi_j| < \infty$ $(absolutely\ summable)$ 조건을 만족해야 함

3

AR 모형

정의

AR 모형 Auto-Regressive Model

현 시점의 관측값을 과거 시점의 관측값과

현 시점의 오차의 선형결합 형태로 표현한 모형

$Z_t \sim WN(0, \sigma^2)$

$$AR(1): X_t = \phi_1 X_{t-1} + Z_t$$

$$AR(p): X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t$$

AR(p)에서 p는 몇 시점 전까지의 관측값을 사용했는지를 나타내는 모수

3 AR 모형

정의

AR 모형 Auto-Regressive Model

현 시점의 관측값을 과거 시점의 관측값과

현 시점의 오차의 선형결합 형태로 표현한 모형

관측값을 자기 자신의 과거로 회귀시킨다는 의미에서

'자기회귀 모형' 이라고 부름

$$Z_t \sim WN(0, \sigma^2)$$

$$AR(1): X_t = \phi_1 X_{t-1} + Z_t$$

$$AR(p): X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t$$

AR(p)에서 p는 몇 시점 전까지의 관측값을 사용했는지를 나타내는 모수

3 AR 모형

특성방정식

P시점 전까지지의 관측값으로 나타낸 선형결합(X_t)은
후향연산자를 사용해 표현가능

$X_{t-1} = BX_t,$ 에서 B가 후향연산자

$$
\begin{aligned}
X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t \\
&= \phi_1 B X_t + \phi_2 B^2 X_t + \cdots + \phi_p B^p X_t + Z_t \\
&= (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q) Z_t
\end{aligned}
$$

$$Z_t = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) X_t$$

3 AR 모형

특성방정식

P시점 전까지의 관측값으로 나타낸 선형결합(X_t)은

후향연산자를 사용해 표현가능

$X_{t-1} = BX_t,$ 에서 B가 후향연산자

Z_t 에 대해 표현

$$Z_t = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) X_t$$

특성방정식

P시점 전까지의 관측값으로 나타낸 선형결합(X_t)은
후향연산자를 사용해 표현가능

$X_{t-1} = BX_t,$ 에서 B가 후향연산자

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t$$

AR(p)의 특성방정식(= $\phi(B)$)

$$= \phi_1 B X_t + \phi_2 B^2 X_t + \cdots + \phi_p B^p X_t + Z_t$$

$$= (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q) Z_t$$

$$Z_t = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) X_t$$

AR(p)를 나타낸 식들 정리

정의를 이용해 나타낸 식

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t$$

후향연산자를 이용해 나타낸 식

$$X_t = \phi_1 B X_t + \phi_2 B^2 X_t + \cdots + \phi_p B^p X_t + Z_t$$

특성방정식을 이용해 나타낸 식

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

AR 모형의 조건 : 정상성과 인과성

AR 모형의 두 가지 조건

1. 정상성(Stationarity) : 시계열의 확률적 특성이 시점에 의존하지 않는 특성

2. 인과성(Causality) : t 시점의 관측값이 과거시점의 오차항으로 설명되는 특성

$$\rightarrow \psi_j = 0, \forall j < 0 \leftrightarrow X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

선형과정 X_t가 위 조건을 만족할 때, X_t 는 인과성을 가짐!

AR 모형의 조건 : 정상성과 인과성

AR 모형의 두 가지 조건

1. 정상성(Stationarity) : 시계열의 확률적 특성이 시점에 의존하지 않는 특성
2. 인과성(Causality) : t 시점의 관측값이 과거시점의 오차항으로 설명되는 특성

$$\rightarrow \psi_j = 0, \forall j < 0 \leftrightarrow X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

선형과정 X_t가 위 조건을 만족할 때, X_t 는 인과성을 가짐!

AR 모형이 어떻게 정상성과 인과성을 만족하는지

AR(1) 모형을 통해 알아보자!

AR 모형의 조건 : 정상성과 인과성

$$
\begin{aligned}
X_t &= \phi_1 X_{t-1} + Z_t \\
&= \phi_1(\phi_1 X_{t-2} + Z_{t-1}) + Z_t \\
&= \phi_1^2 X_{t-2} + \phi_1 Z_{t-1} + Z_t \\
&= \phi_1^2(\phi_1 X_{t-3} + Z_{t-2}) + \phi_1 Z_{t-1} + Z_t \\
&= \phi_1^3 X_{t-3} + \phi_1^2 Z_{t-2} + \phi_1 Z_{t-1} + Z_t \\
&\quad \vdots \\
&= \phi_1^{M+1} X_{t-M-1} + \sum_{j=0}^{M} \phi_1^j Z_{t-j}
\end{aligned}
$$

AR(1) 식을 위와 같이 과거 시점의 관측값과 오차들의 선형결합으로 정리할 수 있음

But 인과성을 만족하기 위해서는 오차항만으로 관측값을 설명해야함

AR 모형의 조건 : 정상성과 인과성

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + Z_t \\ &= \phi_1(\phi_1 X_{t-2} + Z_{t-1}) + Z_t \\ &= \phi_1^2 X_{t-2} + \phi_1 Z_{t-1} + Z_t \\ &= \phi_1^2(\phi_1 X_{t-3} + Z_{t-2}) + \phi_1 Z_{t-1} + Z_t \\ &= \phi_1^3 X_{t-3} + \phi_1^2 Z_{t-2} + \phi_1 Z_{t-1} + Z_t \\ &\vdots \\ &= \phi_1^{M+1} X_{t-M-1} + \sum_{j=0}^{M} \phi_1^j Z_{t-j} \end{aligned}$$

인과성의 만족 여부는 ϕ_1의 범위에 따라 달라지는데,

식을 통해 확인해보자!

AR(1) 식을 위와 같이 과거 시점의 관측값과 오차들의 선형결합으

But 인과성을 만족하기 위해서는 오차항만으로 관측값을

AR 모형의 조건 : 정상성과 인과성

$i)\ |\phi_1| < 1$

$$X_t = \phi_1^{M+1}\, X_{t-M-1} + \sum_{j=0}^{M} \phi_1^j Z_{t-j}$$

$M \to \infty$이면 $\phi_1^{M+1} X_{t-M-1} \to 0$ 으로 수렴, 나머지 부분은 $\sum_{j=0}^{M} \phi_1^j Z_{t-j}$이 됨

즉, 오차항(Z_t) 의 선형결합만 남아 인과성 만족

또한 정상 시계열의 선형결합은 여전히 정상 시계열임을 확인했으므로 정상성 역시 만족!

AR 모형의 조건 : 정상성과 인과성

$ii)\ |\phi_1| = 1$

$$X_t = X_{t-1} + \varepsilon_t \quad or \quad X_t = X_{t-1} + \varepsilon_t$$

대표적인 비정상 확률 과정 중 하나인 확률보행과정(random walk process)

$|\phi_1| = 1$일 때는 인과성과 정상성 모두 만족하지 못함

AR 모형의 조건 : 정상성과 인과성

$iii)\ |\phi_1| > 1$

$$X_t = \phi_1 X_{t-1} + Z_t = \cdots = \phi_1^{M+1} X_{t-M-1} + \sum_{j=0}^{M} \phi_1^j Z_{t-j}$$

위 식에서 $\phi_1^{M+1} X_{t-M-1}$ 부분을 제거하지 못해 인과성을 만족하지 못함!

AR 모형의 조건 : 정상성과 인과성

정리하자면, AR모형은 정상성과 인과성이 모두 만족되는

$|\phi_1| < 1$ 이 성립할 때만 사용할 수 있음

그리고 이는 '$\phi(B) = 0$의 근의 절댓값이 1보다 커야한다'와 동치

위 식에서 $\phi_1^{M+1}X_{t-M-1}$ 부분을 제거하지 못해 인과성을 만족하지 못함!

AR 모형의 ACF

계산의 편의를 위해 $E(X_t) = 0$ 가정

AR(1) 모형을 이용한 ACF 계산

[1] ACF식을 유도하기 위해 양변에 X_{t-h} 곱하기

$$X_t = \phi_1 X_{t-1} + Z_t$$

$$X_t X_{t-h} = \phi_1 X_{t-1} X_{t-h} + Z_t X_{t-h}$$

AR 모형의 ACF

계산의 편의를 위해 $E(X_t) = 0$ 가정

AR(1) 모형을 이용한 ACF 계산

[1] ACF식을 유도하기 위해 양변에 X_{t-h} 곱하기

$$X_t = \phi_1 X_{t-1} + Z_t$$

$$X_t X_{t-h} = \phi_1 X_{t-1} X_{t-h} + Z_t X_{t-h}$$

[2] 기댓값을 취해 ACF 식 구하기

$$\gamma(h) = \phi_1 \gamma(h-1) + Cov(Z_t, X_{t-h}) = \phi_1 \gamma(h-1)$$

$$(\because Cov(Z_t, X_{t-h}) = Cov(Z_t,\ \phi_1^{M+1} X_{t-M-1} + \textstyle\sum_{j=0}^{M} \phi_1^j Z_{t-h-j} = 0)$$

$$\gamma(h) = \phi_1(\phi_1 \gamma(h-2)) = \cdots = {\phi_1}^h \gamma(0)$$

$$\frac{\gamma(h)}{\gamma(0)} = {\phi_1}^h = \rho(h)$$

AR 모형의 ACF

AR(1)모형을 이용한 ACF 계산

[1] ACF식을 유도하기 위해 양변에 X_{t-h}를 곱함

[2] 기댓값을 취해 ACF 식을 구함

$$\gamma(h) = \phi_1\gamma(h-1) + Cov(Z_t, X_{t-h}) = \phi_1\gamma(h-1)$$

$$(\because Cov(Z_t, X_{t-h}) = Cov(Z_t,\ \phi_1^{M+1}X_{t-M-1} + \textstyle\sum_{j=0}^{M}\phi_1^{j}Z_{t-j} = 0)$$

$$\gamma(h) = \phi_1(\phi_1\gamma(h-2)) = \cdots = {\phi_1}^h\gamma(0)$$

앞서 정상성을 만족하기 위해서 $|\phi_1| < 1$가 되어야 함을 확인했기 때문에,

h가 커짐에 따라 AR 모형의 ACF는 **지수적으로 감소함**을 알 수 있음

$$\frac{\gamma(h)}{\gamma(0)} = {\phi_1}^h = \rho(h)$$

AR 모형의 PACF

AR(p) 모형을 이용한 PACF 유도

AR(p)는 X_{k+1}을 p 시점 이전 값들로만 표현한 것

$$\hat{X}_{k+1} = \phi_1 X_k + \phi_2 X_{k-1} + \cdots + \phi_p X_{k+1-p}$$

$$\hat{X}_{k+1} = \phi_1 X_k + \phi_2 X_{k-1} + \cdots + \phi_p X_{k+1-p} + 0X_{k-p} + \cdots + 0X_1$$

AR 모형의 PACF

AR(p)는 X_{k+1}을 p 시점 이전 값들로만 표현한 것

AR(p) 모형을 이용한 PACF 유도

$$\hat{X}_{k+1} = \phi_1 X_k + \phi_2 X_{k-1} + \cdots + \phi_p X_{k+1-p}$$

$$\hat{X}_{k+1} = \phi_1 X_k + \phi_2 X_{k-1} + \cdots + \phi_p X_{k+1-p} + 0X_{k-p} + \cdots + 0X_1$$

$$\alpha(0) \coloneqq 1$$

$$\alpha(p) = \phi_p$$

$$\alpha(k) = 0 \;\; if \;\; k > p$$

$\alpha(0)$의 경우 중간 시점이 없기 때문에!

AR 모형의 PACF

AR(p)모형을 이용한 PACF 유도

AR(p)는 X_{k+1}을 p 시점 이전 값

$$\hat{X}_{k+1} = \phi_1 X_k + \phi_1 X_k + \cdots + \phi_p X_{k+1-p}$$

$$\hat{X}_{k+1} = \phi_1 X_k + \phi_1 X_k + \cdots + \phi_p X_{k+1-p} + 0X_{k-p} + \cdots + 0X_1$$

즉, AR 모형의 PACF는 p 시차 전 까지만 존재하며, p 이후로는 모두 0이 됨

이를 "AR(p) 모형의 PACF는 시차 p 이후에 절단된다"라고 표현함

$$\alpha(0) \coloneqq 1$$

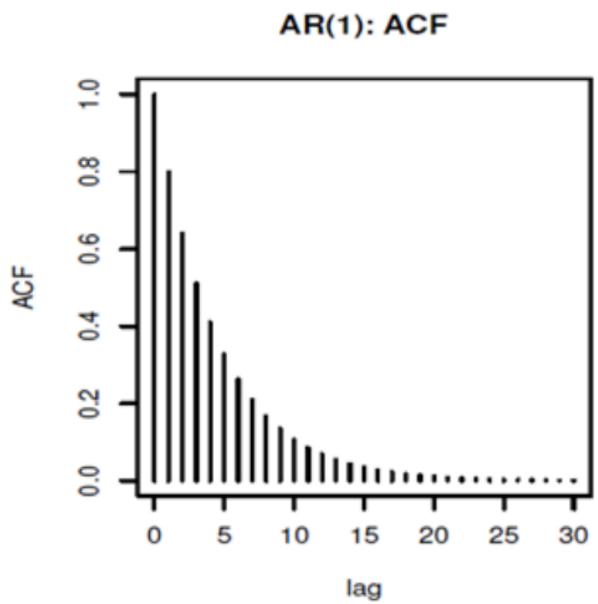
$$\alpha(p) = \phi_p$$

$$\alpha(k) = 0 \;\; if \;\; k > p$$

$\alpha(0)$의 경우 중간 시점이 없기 때문에!

3 AR 모형

AR 모형의 ACF & PACF



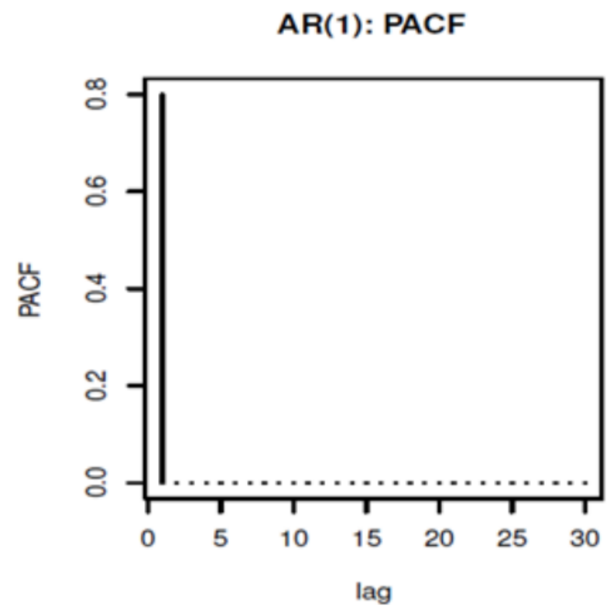
지수적으로 감소

p(위 그림에서는 1)이후에 절단

3 AR 모형

AR 모형의 ACF & PACF

지수적으로 감소



p(위 그림에서는 1)이후에 절단

4

MA 모형

4 MA 모형

정의

MA(Moving Average) 이동평균모형

현 시점의 관측값을 과거 시점의 오차항만을 이용해 설명하는 모형

$$MA(1): X_t = Z_t + \theta_1 Z_{t-1}$$

$$MA(q): X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q}$$

$Z_t \sim WN(0, \sigma^2)$

4 MA 모형

특성방정식

MA모형도 AR모형과 같이 후향연산자를 이용해 표현가능

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q}$$

$$= Z_t + \theta_1 B Z_t + \theta_2 B^2 Z_t + \cdots + \theta_q B^q Z_t$$

$$= (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q) Z_t$$

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

특성방정식

MA모형도 AR모형과 같이 후향연산자를 이용해 표현가능

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q}$$

$$= Z_t + \theta_1 B Z_t + \theta_2 B^2 Z_t + \cdots + \theta_q B^q Z_t$$

$$= (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q) Z_t$$

Z_t 에 대해 표현

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

특성방정식

MA모형도 AR모형과 같이 후향연산자를 이용해 표현가능

MA(q)의 특성방정식(= $\theta(B)$)

$$(1-\theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q)$$

MA 모형의 조건 : 정상성, 인과성, 가역성

MA 모형의 세 가지 조건

1. 정상성(Stationarity) : 시계열의 확률적 특성이 시점에 의존하지 않는 특성
2. 인과성(Causality) : t 시점의 관측값이 과거시점의 오차항으로 설명되는 특성
3. 가역성(Invertibility) : t 시점의 오차항이 과거시점의 관측값으로 설명되는 특성

   → $\sum_{j=0}^{\infty}|\pi_j| < \infty$ 인 $\{\pi_j\}$ 가 존재하며, 아래 식을 만족할 때 가역성을 가짐

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \; for \; all \; t$$

Z_t에 대한 식을 X_t에 대한 식으로 표현할 수 있는지를 따지는 조건이라고 이해하면 편함!

MA 모형의 조건 : 정상성, 인과성, 가역성

MA모형은 백색잡음을 만족하는 오차의 선형결합이기 때문에,

자연스럽게 인과성과 정상성을 만족

↓

가역성의 조건 만족 여부만 확인하면 됨!

1. 정상성(Stationarity) : 시계열의 확률적 특성이 시점에 의존하지 않는 특성
2. 인과성(Causality) : t 시점의 관측값이 과거시점의 오차항으로 설명되는 특성
3. 가역성(Invertibility) : t 시점의 오차항이 과거시점의 관측값으로 설명되는 특성

→ $\sum_{j=0}^{\infty}|\pi_j| < \infty$ 인 $\{\pi_j\}$ 가 존재하며, 아래 식을 만족할 때 가역성을 가짐

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \; for \; all \; t$$

Z_t에 대한 식을 X_t에 대한 식으로 표현할 수 있는지를 따지는 조건이다

MA(1) 모형을 통한 가역성 조건 확인

$$X_t = Z_t + \theta Z_{t-1} = (1+\theta B)Z_t$$

$$(1+\theta B)^{-1}X_t = Z_t$$

$$(1+\theta B)^{-1} = \frac{1}{1-(-\theta B)} = 1 - \theta B + \theta B^2 - (\theta B)^3 + \cdots$$

⋮

MA(1) 식을 위와 같이 무한등비급수의 합 형태로 표현할 수 있음

따라서 가역성 조건 역시 $|\theta| < 1$일 때만 성립하며,

이는 특성방정식 $\theta(\mathrm{B}) = 0$의 근이 절댓값 1보다 커야 한다는 조건과 동치

MA 모형의 ACF

✓ MA(1) 모형의 ACF 계산 과정

⋮

[1] MA(1) 모형 양변에 X_{t-h} 곱하기

$$X_{t-h}X_t = X_{t-h}(Z_t + \theta_1 Z_{t-1}) = X_{t-h}Z_t + \theta_1 X_{t-h}Z_{t-1}$$

[2] 기댓값 취하기

$$\gamma(\mathrm{h}) = Cov(X_{t-h}, Z_t + \theta_1 Z_{t-1}) = Cov(Z_{t-h}, +\theta_1 Z_{t-h-1}, Z_t + \theta_1 Z_{t-1})$$

MA 모형의 ACF

✓ MA(1) 모형의 ACF 계산 과정

⋮

[1] MA(1) 모형 양변에 X_{t-h} 곱하기

$$X_{t-h}X_t = X_{t-h}(Z_t + \theta_1 Z_{t-1}) = X_{t-h}Z_t + \theta_1 X_{t-h}Z_{t-1}$$

[2] 기댓값 취하기

$$\gamma(\mathrm{h}) = Cov(X_{t-h}, Z_t + \theta_1 Z_{t-1}) = Cov(Z_{t-h}, +\theta_1 Z_{t-h-1}, Z_t + \theta_1 Z_{t-1})$$

MA 모형의 ACF

MA(1) 모형의 ACF 계산 과정

⋮

[3] h의 범위에 따른 $\gamma(h)$ 확인

i) h = 0

$$\gamma(0) = Cov(Z_t, +\theta_1 Z_{t-1}, Z_t + \ \theta_1 Z_{t-1}) = \sigma^2 + \theta_1^2\sigma^2 = (1 + \theta_1^2)\sigma^2$$

ii) h = 1

$$\gamma(1) = Cov(Z_{t-1}, +\theta_1 Z_{t-2}, Z_t + \ \theta_1 Z_{t-1}) = \theta_1\sigma^2$$

iii) h ≥ 2

$$\gamma(\mathrm{h}) = Cov(Z_{t-h}, +\theta_1 Z_{t-h-1}, Z_t + \ \theta_1 Z_{t-1}) = 0$$

4 MA 모형

MA 모형의 ACF

✓ MA(1) 모형의 ACF 계산 과정

⋮

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

ACF는 시차 q = 1이후에 **단절 !**

MA 모형의 PACF

Crammer 공식을 이용한
MA 모형의 PACF에 대한 자세한 증명은 시계열 2주차 교안 참고!

MA(1) 모형의 PACF

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

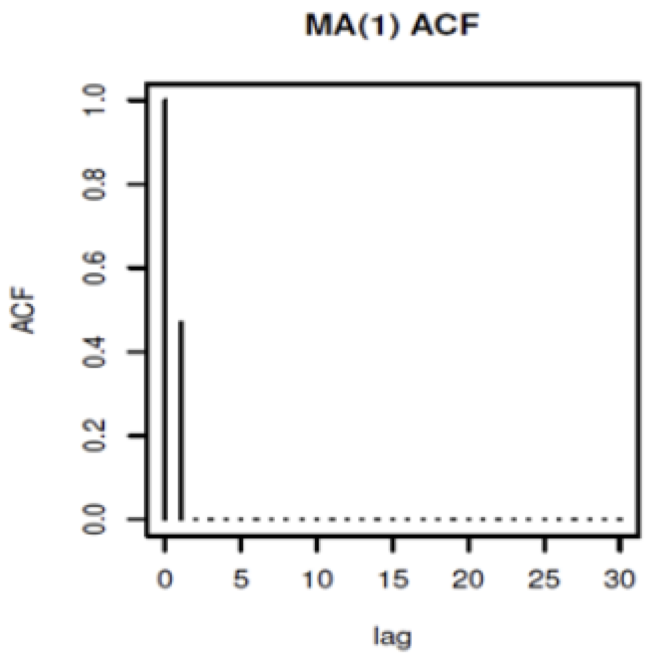
⋮

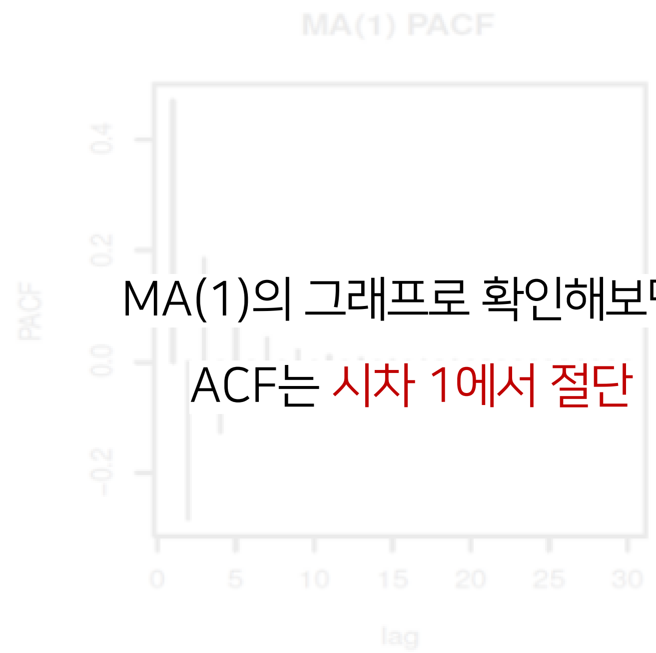
MA모형은 $|\theta| < 1$일 때 성립하므로,

위 식은 시차 k가 커질수록 0에 수렴

4 MA 모형

MA 모형의 ACF & PACF

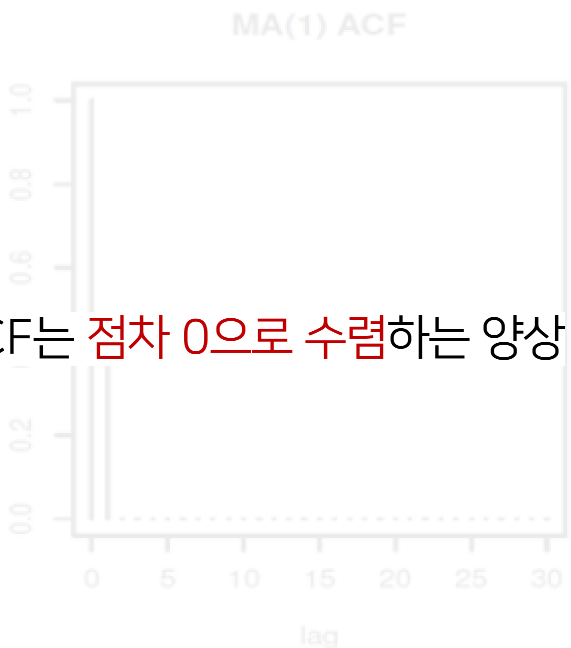




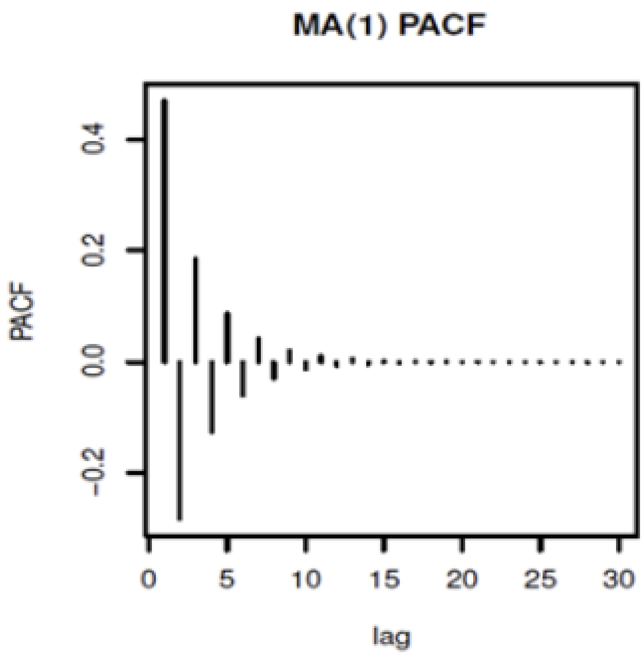
MA(1)의 그래프로 확인해보면,

ACF는 시차 1에서 절단

MA 모형의 ACF & PACF



PACF는 점차 0으로 수렴하는 양상



5

ARMA 모형

AR 모형과 MA 모형의 쌍대성

AR 모형과 MA 모형의 쌍대성

유한차수의 AR 모형은 무한차수의 MA 모형,

유한차수의 MA 모형은 무한차수의 AR 모형으로 표현 가능

→ 서로에 의해 표현될 수 있는 특성

1) AR(1) → MA(∞)

$$X_t = \phi_1 X_{t-1} + Z_t$$

$$= \phi_1(\phi_1 X_{t-2} + Z_{t-1}) + Z_t = \phi_1^2 X_{t-2} + \phi_1 Z_{t-1} + Z_t$$

$$= \phi_1^2 \, (\phi_1 X_{t-3} + Z_{t-2}) + \phi_1 Z_{t-1} + Z_t = \phi_1^3 X_{t-3} + \phi_1^2 Z_{t-2} + \phi_1 Z_{t-1} + Z_t$$

$$= \cdots = \phi_1^{M+1} X_{t-M-1} + \sum_{j=0}^{M} \phi_1^j Z_{t-j}$$

AR 모형과 MA 모형의 쌍대성

AR 모형과 MA 모형의 쌍대성

유한차수의 AR 모형은 무한차수의 MA 모형,

유한차수의 MA 모형은 무한차수의 AR 모형으로 표현 가능

→ 서로에 의해 표현될 수 있는 특성

1) AR(1) → MA(∞)

AR 모형의 인과성 조건에 따라,

$X_t = \sum_{j=0}^{\infty} \phi_1^j Z_{t-j}$로 MA 모형이 됨

$$= \phi_1(\phi_1 X_{t-2} + Z_{t-1}) + Z_t = \phi_1^2 X_{t-2} + \phi_1 Z_{t-1} + Z_t$$

$$= \phi_1^2(\phi_1 X_{t-3} + Z_{t-2}) + \phi_1 Z_{t-1} + Z_t = \phi_1^3 X_{t-3} + \phi_1^2 Z_{t-2} + \phi_1 Z_{t-1} + Z_t$$

$$= \cdots = \phi_1^{M+1} X_{t-M-1} + \sum_{j=0}^{M} \phi_1^j Z_{t-j}$$

5 ARMA 모형

AR 모형과 MA 모형의 쌍대성

2) MA(1) → AR(∞)

$$X_t = \theta_1 B Z_t + Z_t$$

$$X_t = (1 + \theta_1 B) Z_t \rightarrow Z_t = \frac{1}{1 - (-\theta_1 B)} X_t$$

$$\therefore (1 - \theta_1 B + \theta_1^2 B^2 - \cdots) X_t = Z_t$$

$$\Longrightarrow X_t - \theta_1 B X_t + \theta_1^2 B^2 X_t - \cdots = Z_t$$

$$X_t = \theta_1 B X_t - \theta_1^2 B^2 X_t + \cdots + Z_t$$

MA 모형을 현재시점까지 관측값과 현재시점의 오차항으로 표현 가능

→ AR 모형

5 ARMA 모형

ARMA 모형을 사용하는 이유

AR 모형 또는 MA 모형만으로 시계열 자료를 설명할 경우,

모수 p나 q가 너무 커질 위험 존재

→ 효율성이 떨어지고, 해석이 어렵다는 단점

ARMA 모형은

AR(p) 모형과 MA(q) 모형을 모두 포함

모수의 절약이 가능

ARMA 모형

ARMA 모형

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots \theta_q Z_{t-q}$$

ARMA 특성방정식

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots \theta_q Z_{t-q}$$

$$X_t - \phi_1 B X_t - \phi_2 B^2 X_t - \cdots - \phi_p B^p X_t = Z_t + \theta_1 B Z_t + \theta_2 B^2 Z_t + \cdots \theta_q B^q Z_t$$

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) X_t = (1 + \theta_1 B + \theta_2 B^2 + \cdots \theta_q B^q) Z_t$$

$$\phi(B) X_t = \theta(B) Z_t$$

ARMA 모형

ARMA 모형

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots \theta_q Z_{t-q}$$

AR 모형의 특성방정식

MA 모형의 특성방정식

ARMA 특성방정식

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots \theta_q Z_{t-q}$$

$$X_t - \phi_1 B X_t - \phi_2 B^2 X_t - \cdots - \phi_p B^p X_t = Z_t + \theta_1 B Z_t + \theta_2 B^2 Z_t + \cdots \theta_q B^q Z_t$$

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) X_t = (1 + \theta_1 B + \theta_2 B^2 + \cdots \theta_q B^q) Z_t$$

$$\phi(B) X_t = \theta(B) Z_t$$

ARMA 모형의 조건

AR & MA 모형의 조건 세 가지

1. 정상성 2. 인과성 3. 가역성

모두 만족해야 함

+

식별성

주어진 파라미터 조합에 대해 단 하나의 모형이 대응되는 특성

ARMA 모형 | 식별성

$$[X_t = Z_t]$$

: $(1-\phi B) = (1+\theta B) \Leftrightarrow \phi = -\theta$를 만족하는 ARMA (1,1)

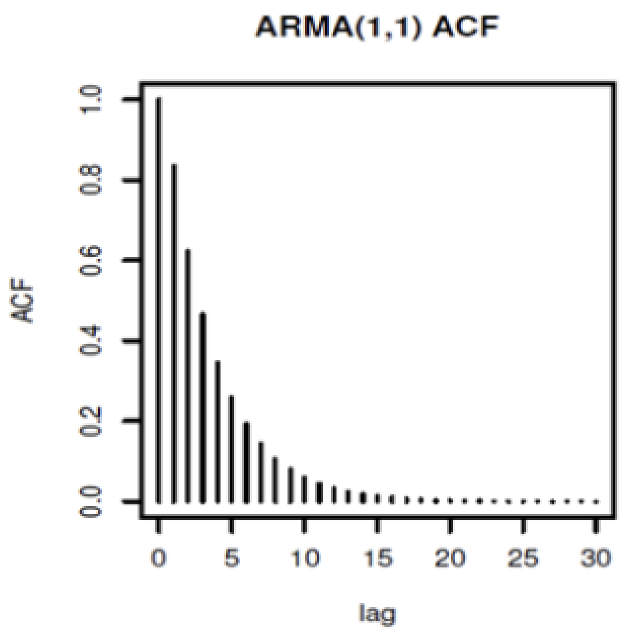
&

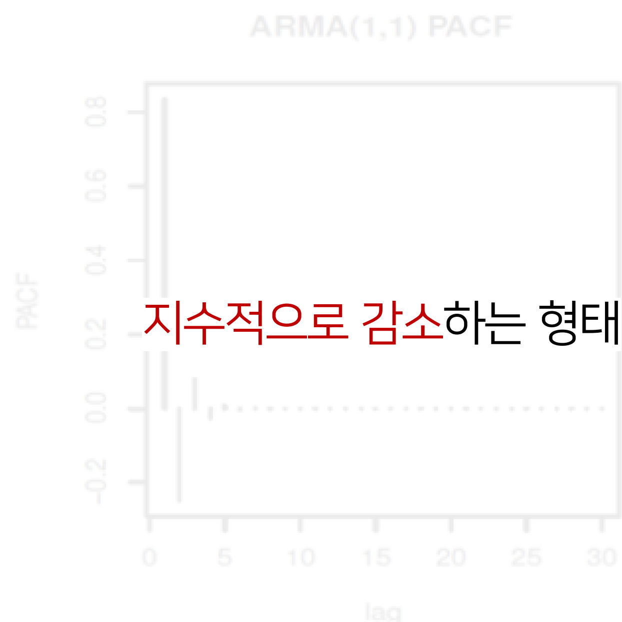
: $WN(0, \sigma^2)$

정상성, 인과성, 가역성 세 조건만으로 모형 식별이 어려움

→ $\phi + \theta \neq 0$ 즉, 식별성 조건 추가 필요

ARMA 모형의 ACF & PACF

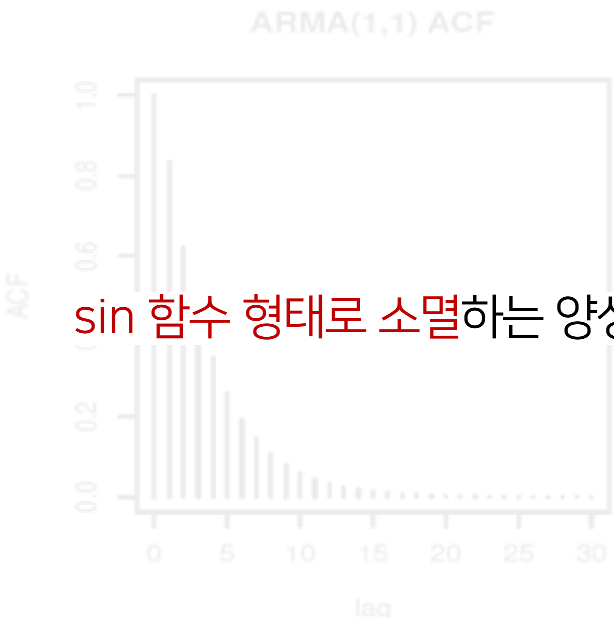




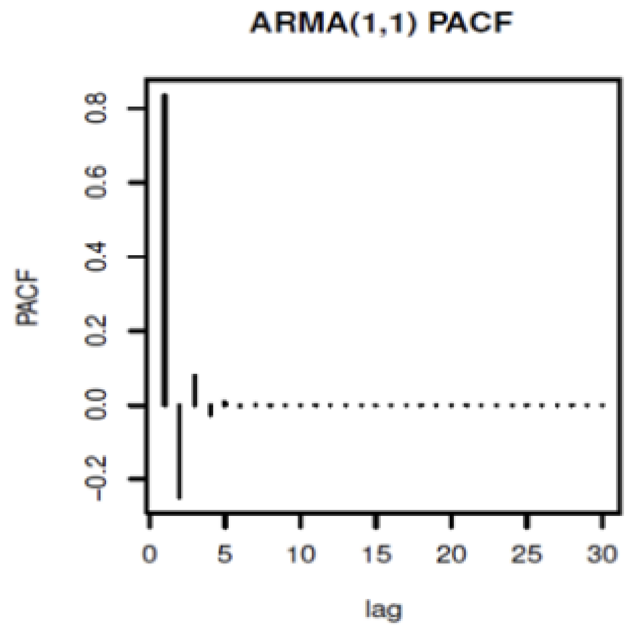
지수적으로 감소하는 형태

5 ARMA 모형

ARMA 모형의 ACF & PACF

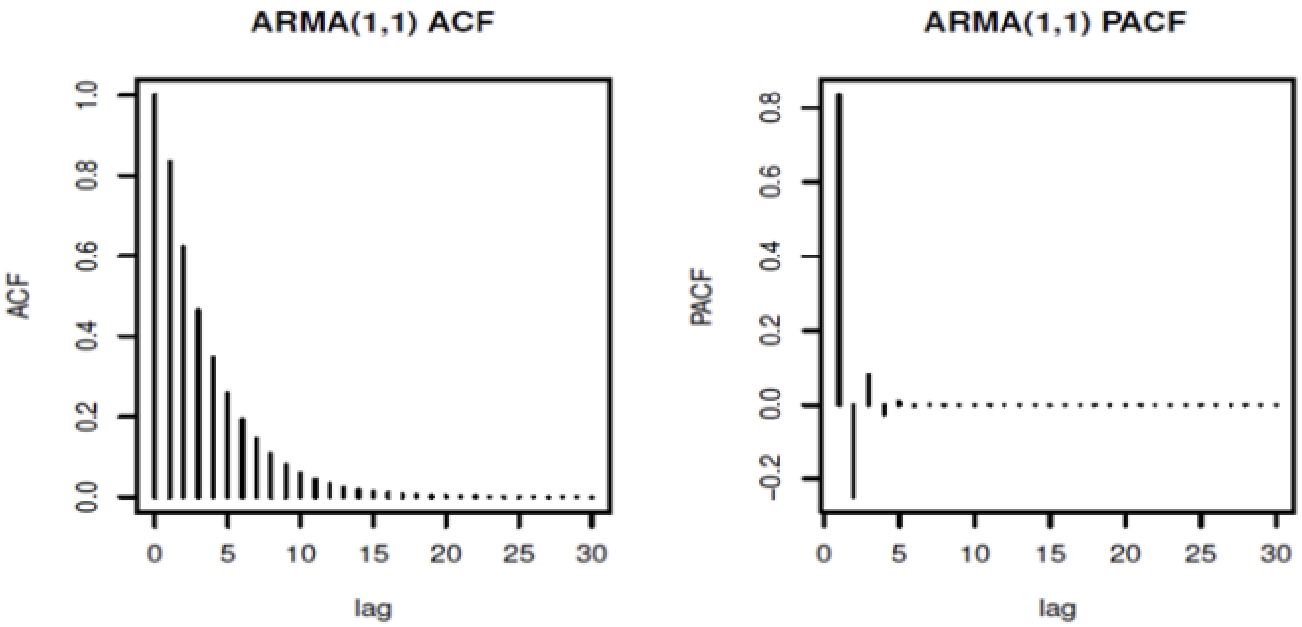


sin 함수 형태로 소멸하는 양상



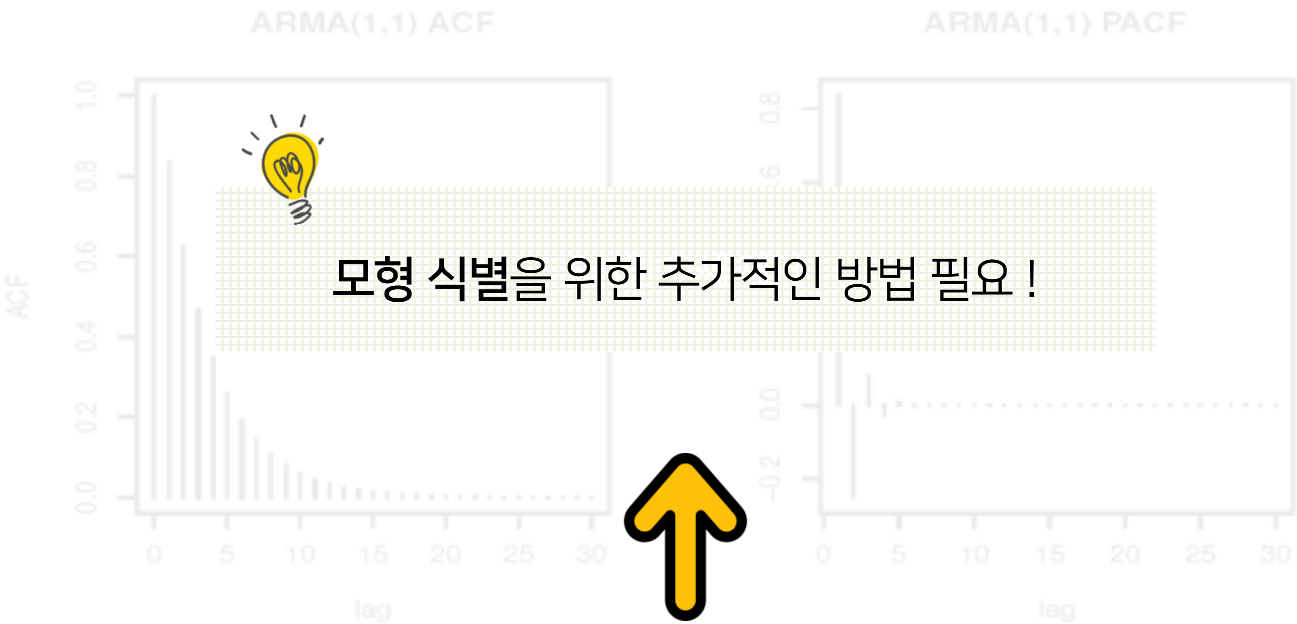
5 ARMA 모형

ARMA 모형의 ACF & PACF



두 그래프 모두 절단되지 않고 지수적으로 감소한다는 한계점 존재

ARMA 모형의 ACF & PACF



모형 식별을 위한 추가적인 방법 필요 !

두 그래프 모두 절단되지 않고 지수적으로 감소한다는 한계점 존재

6

모형의 적합 절차

모형 적합 | ① 모형 식별

[1] 사용할 모형과 차수 결정

AR & MA : ACF와 PACF 그래프의 절단되는 특징

→ p와 q 차수까지 쉽게 결정 가능

vs

ARMA : ACF와 PACF 모두 지수적 감소

→ Information Criteria 사용 필요

모형 적합 | ① 모형 식별

[1] 사용할 모형과 차수 결정

AR & MA : ACF와 PACF 그래프의 절단되는 특징

→ p와 q 차수까지 쉽게 결정 가능

vs

ARMA : ACF와 PACF 모두 지수적 감소

→ Information Criteria 사용 필요

$$Information\ Criteria = \{Goodness\ of\ fit + model\ complexity\}$$

가장 작은 IC 값을 가지는 모형을 선택

6 모형의 적합 절차

모형 적합 | ① 모형 식별

Information Criteria

AIC (Akaike Information Criteria)

$$2\ln L_n(\hat{\theta}) + 2(p+q+1)$$

AICC (AIC bias corrected)

$$2\ln L_n(\hat{\theta}) + \frac{2(p+q+1)n}{n-(p+q+1)+1}$$

Bayesian Information Criterion

$$2\ln L_n(\hat{\theta}) + (p+q+1)\ln n$$

모형 적합 | ② 모수 추정

[2] 모수 추정

모수의 차수를 결정했다면, 모수 ϕ, θ, σ^2를 추정해야 함

최대가능도추정법 (MLE)	관측된 시계열의 결합확률밀도함수인 모수의 가능도함수를 최대화하는 모수의 추정량을 구함
최소제곱법 (LSE)	오차의 제곱합이 가장 작게 되도록 하는 모수의 추정량을 구함
적률추정법 (MME/MoM)	모집단의 적률을 상응하는 표본의 적률로 대체한 후, 방정식을 풀어 모수의 추정량을 구함

모형 적합 | ③ 모형 진단

[3] 모형 진단

모형이 적합한지 진단. 모수에 대한 검정/잔차에 대한 검정

모수에 대한 검정

1) 모형의 조건을 만족하는지 확인

- 정상성과 가역성 조건 만족 여부
- 식별성 만족 여부

2) 모수의 유효성 확인

- 모수$\neq 0$인지 확인

모형 적합 | ③ 모형 진단

잔차에 대한 검정

1) 추세, 계절성, 이상치 없는지 확인

2) $\mathrm{WN}(0, \sigma^2)$을 따르는지 확인

- 잔차에 대한 ACF, PACF 그래프
- Ljung-Box test / McLeod-LI test / Different sign test

3) 정규성을 만족하는지 확인

- 잔차의 QQ plot
- Jarque-Bera test

모형 적합 | ④ 예측

[4] 예측

과거의 모든 정보를 알고 있다고 가정하는 infinite한 방법과

알고 있는 자료를 사용해 예측하는 finite한 방법

실제로 주로 사용되는 방식이므로 이를 클린업에서 다룸

가지고 있는 데이터의 **선형결합**을 활용해 미래 예측

$$P_b X_{n+h} = a_0 * 1 + a_1 X_n + a_2 X_{n-1} + \cdots + a_n X_1$$

→ n개의 자료를 가지고 n + h 시점 예측

모형 적합 | ④ 예측

계수 $\{a_0, a_1, \cdots, a_n\}$는

MSPE(Mean Squared Prediction Error)를 최소화하는 방향으로 추정

$$MSPE = \mathbb{E}[X_{n+h} - P_n X_{n+h}]^2$$
$$= \mathbb{E}[X_{n+h} - (a_0 * 1 + a_1 X_n + a_2 X_{n-1} + \cdots + a_n X_1)\,]^2$$

회귀분석 LSE와 동일한 계산 방식

가지고 있는 데이터의 **선형결합**을 활용해 미래 예측

$$P_b X_{n+h} = a_0 * 1 + a_1 X_n + a_2 X_{n-1} + \cdots + a_n X_1$$

→ n개의 자료를 가지고 n + h 시점 예측

7

2주차 정리

7 2주차 정리

정리 | 시계열 모형의 필요성

정상화 과정 이후
남아있는 오차가
WN/IID를 만족해야 함

→

오차가 WN/IID가 아닐 때,
이를 추정하기 위해
시계열 모형 필요

정리 | 선형과정 모형 [1] AR (자기회귀 모형)

자기회귀 모형 AR

$$AR(p):\ X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t$$

$$= \phi_1 B X_t + \phi_2 B^2 X_t + \cdots + \phi_p B^p X_{t-p} + Z_t$$

$$= \phi(B) X_t$$

⋮

$|\phi_1| < 1$ 조건 만족해야 함

후향연산자를 이용해 나타낸 식인 경우,

$\phi(B) = 0$의 근의 절댓값 > 1 만족해야 함

정리 | 선형과정 모형 [1] AR (자기회귀 모형)

AR 모형의 조건

정상성(Stationarity)	시계열의 확률적 특성이 시점에 의존하지 않는 특성
인과성(Causality)	t 시점의 관측값이 과거시점의 오차항으로 설명되는 특성

AR 모형의 ACF & PACF

ACF	지수적으로 감소되는 양상
PACF	p 이후에 절단되는 양상

정리 | 선형과정 모형 [2] MA (이동평균 모형)

이동평균 모형 MA

$$MA(q):\ X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q}$$

$$= Z_t + \theta_1 B Z_t + \theta_2 B^2 Z_t + \cdots + \theta_q B^q Z_t$$

$$= \theta(B) Z_t$$

⋮

$|\theta| < 1$ 조건 만족해야 함

후향연산자를 이용해 나타낸 식인 경우,

$\theta(B) = 0$의 근의 절댓값 > 1 만족해야 함

정리 | 선형과정 모형 [2] MA (이동평균 모형)

MA 모형의 조건

정상성(Stationarity)	시계열의 확률적 특성이 시점에 의존하지 않는 특성
인과성(Causality)	t 시점의 관측값이 과거시점의 오차항으로 설명되는 특성
가역성(Invertibility)	t 시점의 오차항이 과거시점의 관측값으로 설명되는 특성

MA 모형의 ACF & PACF

ACF	q 이후로 절단되는 양상
PACF	0으로 수렴되는 양상

정리 | 선형과정 모형 [3] ARMA (AR+MA)

ARMA 모형의 정의

$$\text{ARMA}(p,q)\text{: } \phi(B)X_t = \theta(B)Z_t$$

ARMA 모형의 조건

정상성(Stationarity)	시계열의 확률적 특성이 시점에 의존하지 않는 특성
인과성(Causality)	t 시점의 관측값이 과거시점의 오차항으로 설명되는 특성
가역성(Invertibility)	t 시점의 오차항이 과거시점의 관측값으로 설명되는 특성
식별성(Identifiability)	$\phi + \theta \neq 0$

정리 | 선형과정 모형 [3] ARMA (AR+MA)

ARMA 모형의 ACF & PACF

ACF	지수적으로 감소되는 양상
PACF	0으로 수렴하는 양상

p와 q 선정에 문제 발생

모형 식별을 위한 추가적인 방법 필요

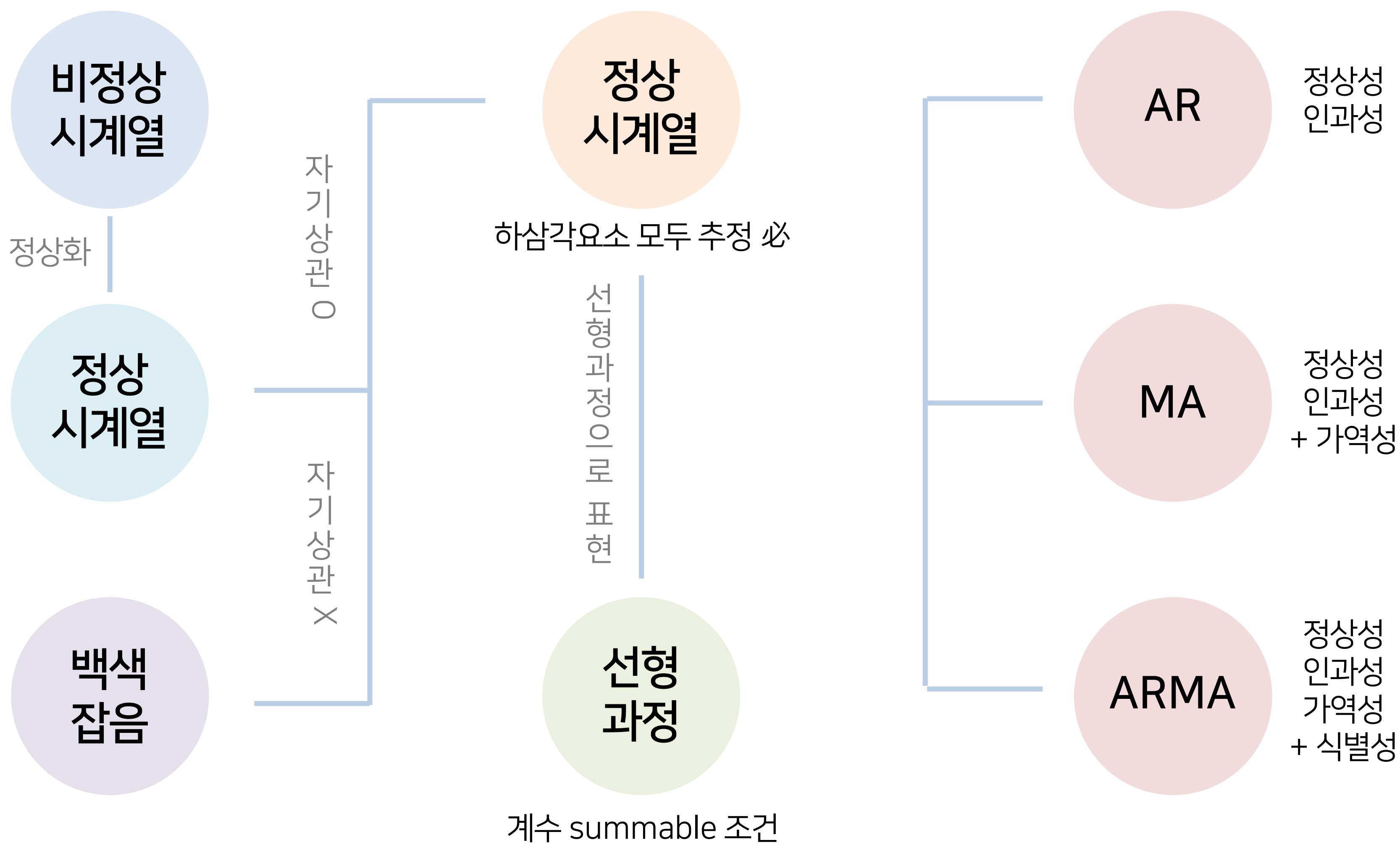
7 2주차 정리

정리 | 모형의 적합 절차

1) 모형 식별		2) 모수 추정		3) 모형 진단		4) 예측
AIC AICC BIC	→	MLE LSE MME/MoM	→	모수 검정 잔차 검정	→	MSPE

! 흐름 정리

시계열 자료 분석 흐름 정리



감사합니다