# 시계열자료분석팀

**5팀**
김민
이수린
김동환
서유진
장다연

# INDEX

# 1

## 1주차 복습

## 시계열 자료

**시계열 자료** *(time series)*

시간에 따라 관측된 자료의 집합으로 관측치 사이 연관성(dependency)을 가짐

시계열 자료의 확률 분포 추정을 위해!

**정상성 가정**

**약정상성**

1. 평균과 분산 일정
2. 자기 공분산(ACVF)이 시차에만 의존

## 시계열 자료

### 시계열 자료 *(time series)*

시간에 따라 관측된 자료의 집합으로 관측치 사이 연관성(dependency)을 가진다.

**비정상 시계열**

**Nonstationary part**

추세, 계절성

**Stationary part**

정상성을 만족하는 오차

# 1 1주차 복습

## 시계열 자료

비정상 시계열

Nonstationary part

추세, 계절성

정상화 : 추세 및 계절성 추정 및 제거

1. 회귀
2. 평활
3. 차분

Stationary part

정상성을 만족하는 오차

## 시계열 자료

비정상 시계열

Nonstationary part
추세, 계절성

→ 정상화 : 추세 및 계절성 추정 및 제거

1. 회귀
2. 평활
3. 차분

Stationary part
정상성을 만족하는 오차

→ 비정상 부분 제거 후 남은 오차항 정상성 검정

1. WN sequence
2. Correlogram(ACF 그래프)

# 1 1주차 복습

## 시계열 자료

비정상 시계열

Nonstationary part

추세, 계절성

정상화 : 추세 및 계절성 추정 및 제거

1. 회귀
2. 평활
3. 차분

Stationary part

정상성을 만족하는 오차

비정상 부분 제거 후 남은 오차항 **정상성 검정**

1. WN sequence
2. Correlogram(ACF 그래프)

**정상화 후 오차항이**
**정상성을 만족하지 않는 경우**

**오차항에 대한 추가적 모델링 필요**

# 시계열 모형!

*오늘의 주제!*

# 2

# 모형의 식별

# 2 모형의 식별

## 시계열 모형의 필요성

자기 공분산 함수*(ACVF)*

$$\gamma_x(h) = Cov(X_t, X_{t+h})$$

분산-공분산 행렬 *(variance-covariance matrix)*

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

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

최종적으로 추정해야하는
분산-공분산 행렬

## 시계열 모형의 필요성

오차항이 백색잡음인 경우

**분산-공분산 행렬** *(variance-covariance matrix)*

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

$$Cov(X_t, X_s) = 0, if\ t \neq s$$

인덱스가 다른 확률변수 간의 상관관계 없음

## 시계열 모형의 필요성

오차항이 백색잡음인 경우

**분산-공분산 행렬** *(variance-covariance matrix)*

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

"0(zero)"

$$Cov(X_t, X_s) = 0, if\ t \neq s$$

대각요소를 제외한 상삼각과 하삼각 요소가 전부 0

→ 추정해야하는 것 : $\sigma^2$

## 시계열 모형의 필요성

오차항이 백색잡음이 아닌 경우

**분산-공분산 행렬** *(variance-covariance matrix)*

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

NOT 0(zero)

확률변수 간의 상관관계가 남아 있기 때문에 분산-공분산 행렬 전체 원소 추정해야함

## 시계열 모형의 필요성

오차항이 백색잡음이 아닌 경우

**분산-공분산 행렬** *(variance-covariance matrix)*

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

NOT 0(zero)

확률변수 간의 상관관계가 남아 있기 때문에 분산-공분산 행렬 전체 원소 추정해야함

시계열 모형!

## 자기상관함수

자기상관함수 ACF *(Auto-Correlation Function)*

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

$\rho_X(h)$ : 시차가 h인 확률변수 간의 상관관계

정상 시계열인 경우,

시점이 변해도 동일한 시차에 대해서 같은 값을 가져야 함

EXAMPLE

$$Cov(X_1, X_2) = \gamma(1) = Cov(X_5, X_6)$$

# 2 모형의 식별

## 자기상관함수

자기 상관 함수의 특징

- $\rho(0) = 1$
- $\rho(-h) = \rho(h)$
- $|\gamma(h)| \leq \gamma(0) \; for \; all \; h \in \mathbb{Z} \rightarrow |\rho(h)| \leq 1$

# 2 모형의 식별

## 자기상관함수

### 자기 상관 함수의 특징

$\rho(0) = 1$

자기 자신과의 상관함수는 항상 1

$\rho(-h) = \rho(h)$

$|\gamma(h)| \leq \gamma(0) \; for \; all \; h \in \mathbb{Z} \rightarrow |\rho(h)| \leq 1$

# 2 모형의 식별

## 자기상관함수

### 자기 상관 함수의 특징

$\rho(0) = 1$

$\rho(-h) = \rho(h)$

자기상관함수는 우함수

$|\gamma(h)| \leq \gamma(0)\ for\ all\ h \in \mathbb{Z} \rightarrow |\rho(h)| \leq 1$

## 자기상관함수

### 자기 상관 함수의 특징

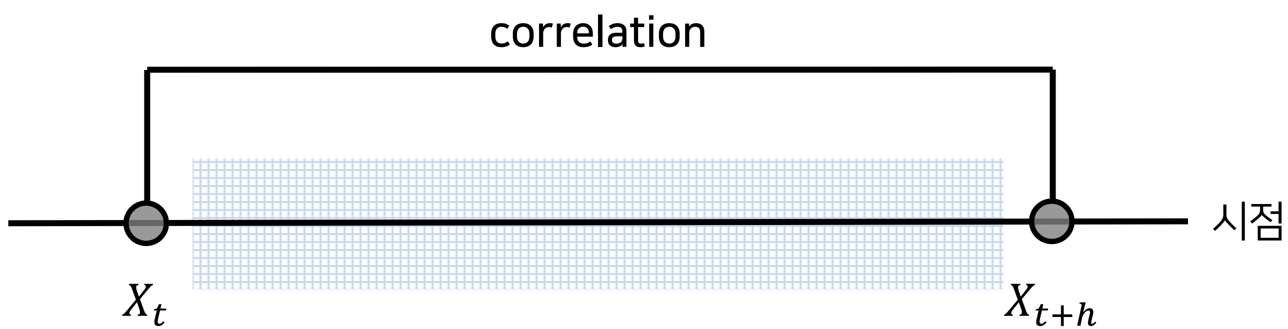
$\rho(0) = 1$

$\rho(-h) = \rho(h)$

$|\gamma(h)| \leq \gamma(0) \; for \; all \; h \in \mathbb{Z} \rightarrow |\rho(h)| \leq 1$

자기상관함수의 절댓값 ≤ 1

## 부분자기상관함수




Corr($X_t$, $X_{t+h}$)에 중간 시점의 영향도 존재

*중간 시점의 영향 제거*

**부분자기상관** *(Partial Auto-Correlation)*

$X_1$과 $X_{k+1}$ 사이에 $X_2, \ldots, X_k$의 영향을 제거했을 때의

조건부 상관관계 (conditional correlation)

## 부분자기상관함수

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

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

중간 시점이 존재하지 않는 경우
자기상관함수와 동일

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

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

중간 시점인 $\{1, X_2, \cdots, X_k\}$만으로 $X_{k+1}$을 설명

$$P_k^* X_1 = \ Best\ Linear\ Predictor\ on\ \{1, X_2, \cdots, X_k\}$$

중간 시점인 $\{1, X_2, \cdots, X_k\}$만으로 $X_1$을 설명

## 부분자기상관함수

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

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

중간 시점이 존재하지 않는 경우
자기상관함수와 동일

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

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

중간 시점인 $\{1, X_2, \cdots, X_k\}$만으로 $X_{k+1}$을 설명

부분자기상관함수 PACF *(Partial Auto-Correlation Function)*

$P_k^* X_{k+1}$과 $P_k^* X_1$을 빼서 $X_1$과 $X_{k+1}$사이에 중간 시점의 영향을 제거한 것

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

## 부분자기상관함수

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

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

$$\vdots$$

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

순서대로 lag =1, lag=2 ··· lag = k인 경우, $\{X_1, \ldots, X_k\}$로 $X_{k+1}$을 설명한 회귀식

**회귀식의 오차를 최소로 하는 추정량**

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

# 2 모형의 식별

## 부분자기상관함수

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

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

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

순서대로 lag =1, lag=2 ⋯ lag = k인 경우, $\{X_1, \dots, X_k\}$로 $X_{k+1}$을 설명한 회귀식

$\{X_{2,} \dots X_k\}$이 설명하는 부분을 제거한 뒤 남은 부분에 대한 $X_1$의 $X_{k+1}$에 대한 영향

= 순수한 $X_1$의 $X_{k+1}$ 사이의 상관관계

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

**회귀식의 오차를 최소로 하는 추정량**

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

# 3

## 선형과정

# 선형과정(Linear Process)

### 선형과정

Time dependency를 설명하기 위해 사용되는 구조로

$\{Z_t\} \sim WN(0, \sigma^2)$들의 선형결합

선형 과정에서의 $\boldsymbol{Z_t}$

: innovation이라 부름

$\sum_j |\psi_j| < \infty$ (absolutely summable) 조건

상수의 합이 발산하지 않도록 만족해야 하는 조건

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

$$\text{where } \psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$$

## 선형과정(Linear Process)

**선형과정**

Time dependency를 설명하기 위해 사용되는 구조로

$\{Z_t\} \sim WN(0, \sigma^2)$들의 선형결합

선형 과정에서의 $\boldsymbol{Z_t}$

: innovation이라 부름

$\sum_j |\psi_j| < \infty$ (absolutely summable) 조건

상수의 합이 발산하지 않도록 만족해야 하는 조건

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

$$\text{where } \psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$$

## 선형과정(Linear Process)

선형 과정의 필요성

1. 공분산 계산 단순화
2. 해석과 추정에 용이
3. 정상 선형과정의 선형 결합 → 정상 선형과정을 다시 만족
4. Wold Decomposition에 의해 약정상성을 만족하는 모든 정상 선형 과정은 선형과정의 선형 결합으로 표현 가능

# 4

## AR 모형

## AR 모형의 정의

AR 모형

$X_t$가 과거 관측값과 현 시점에서의 오차의 함수로 표현되는 모형

**자기회귀**

: 현재를 자신의 과거로 설명

AR(1) 모형

$$\boldsymbol{X_t = \phi_1 X_{t-1} + Z_t} \qquad Z_t \sim WN(0, \sigma^2)$$

AR(p) 모형

$$\boldsymbol{X_t = \underbrace{\phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p}}_{\text{과거의 관측값}} + Z_t}$$

현시점에서의 오차항

## AR 모형의 특성방정식

AR(p) 모형

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

후향 연산자로 표현

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

$Z_t$에 대한 식으로 변환

$$\boldsymbol{Z_t} = \left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right) X_t = \boldsymbol{\phi(B) X_t}$$

## AR 모형의 특성방정식

AR(p) 모형

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

후향 연산자로 표현

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

$Z_t$에 대한 식으로 변환

$$\boldsymbol{Z_t} = \left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right) X_t = \boldsymbol{\phi(B) X_t}$$

AR(p)의 특성방정식 $\phi(B)$

## AR 모형의 조건

AR 모형의 조건

정상성 (stationarity)

시계열의 확률적 특성이 시차에만 의존하는 특성

인과성 (casuality)

t 시점의 관측값이 과거 시점의 오차항으로 설명되는 특성

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

# 4 AR 모형

## AR 모형의 조건

AR(1) 모형을 통해 증명

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

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

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

$$\vdots$$

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

### 선형과정이 되기 위한 조건

innovation들의 계수의 합 유한해야 하는

absolutely summable 조건 만족

$\vdots$

기하급수의 수렴조건에 의해

$|\phi_1| < 1$

$\vdots$

$\phi(B) = 0$의 근의 절댓값이

1보다 커야 함

# 4 AR 모형

## AR 모형의 조건

AR(1) 모형을 통해 증명

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

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

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

$$\vdots$$

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

### 선형과정이 되기 위한 조건

innovation들의 계수의 합 유한해야 하는

absolutely summable 조건 만족

⋮

기하급수의 수렴조건에 의해

$$|\boldsymbol{\phi_1}| < \mathbf{1}$$

⋮

$\phi(B) = 0$의 근의 절댓값이

1보다 커야 함

## AR 모형의 조건

AR(1) 모형을 통해 증명

$|\phi_1| < 1$ 조건에서

① 선형 과정의 성질에 따라

정상 시계열의 선형결합은 정상 시계열

② $X_t$가 과거의 값들로 설명됨

### 선형과정이 되기 위한 조건

innovation들의 계수의 합 유한해야 하는

absolutely summable 조건 만족

⋮

기하급수의 수렴조건에 의해

$|\boldsymbol{\phi_1}| < \mathbf{1}$

⋮

$\phi(B) = 0$의 근의 절댓값이

1보다 커야 함

# 4 AR 모형

## AR 모형의 조건

AR(1) 모형을 통해 증명

$|\phi_1| < 1$ 조건에서

① 선형 과정의 성질에 따라

정상 시계열의 선형결합은 정상 시계열

② $X_t$가 과거의 값들로 설명됨

### 선형과정이 되기 위한 조건

innovation들의 계수의 합 유한해야 하는

absolutely summable 조건 만족

기하급수의 수렴조건에 의해

$|\boldsymbol{\phi_1}| < \mathbf{1}$

$\phi(B) = 0$의 근의 절댓값이

1보다 커야 함

**정상성, 인과성 만족**

## AR 모형의 ACF

AR(1) 모형의 ACF

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

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

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

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

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

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

## AR 모형의 ACF

AR(1) 모형의 ACF

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

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

AR 모형은 시차(h)가 커질수록 ACF가 **지수적으로 감소**

($\phi_1$의 절댓값이 1보다 작기 때문)

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

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

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

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

## AR 모형의 PACF

$X_{k+1}$을 AR(p) 모형을 통해 추정

BLP 회귀식에 OLS 적용한 것과 같은 결과

$$\alpha(p) = \phi_p \quad \& \quad \alpha(k) = 0 \quad if\ k > p$$

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

시차가 p를 넘어가서 상관관계 없음

# 4 AR 모형

## AR 모형의 PACF

$X_{k+1}$을 AR(p) 모형을 통해 추정

BLP 회귀식에 OLS 적용한 것과 같은 결과

$$\boxed{\alpha(p) = \phi_p} \quad \& \quad \alpha(k) = 0 \quad if\ k > p$$

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

시차가 p를 넘어가서 상관관계 없음

### AR 모형의 PACF 특징

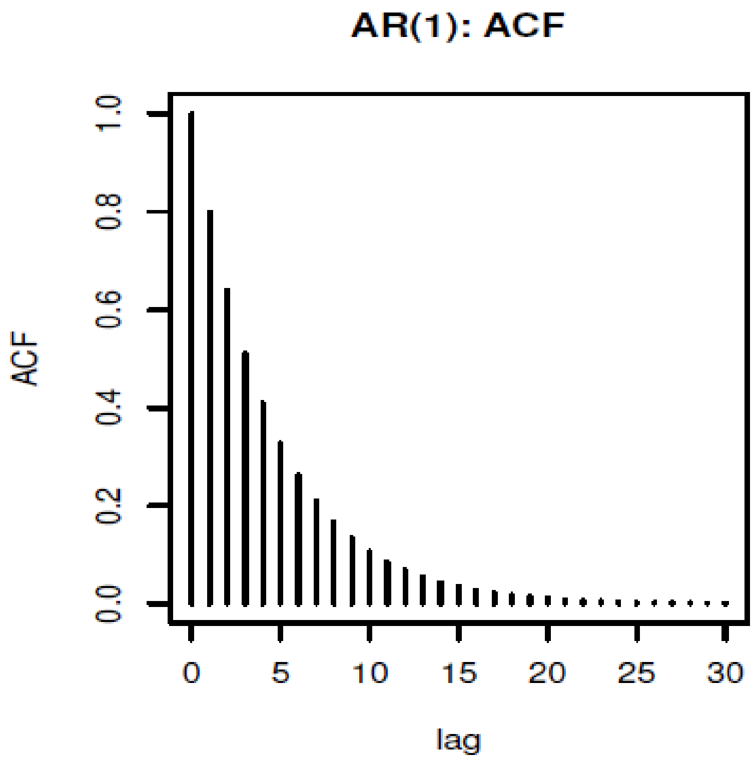
① 시차 p까지만 존재

② 시차가 p를 넘어가면 PACF=0

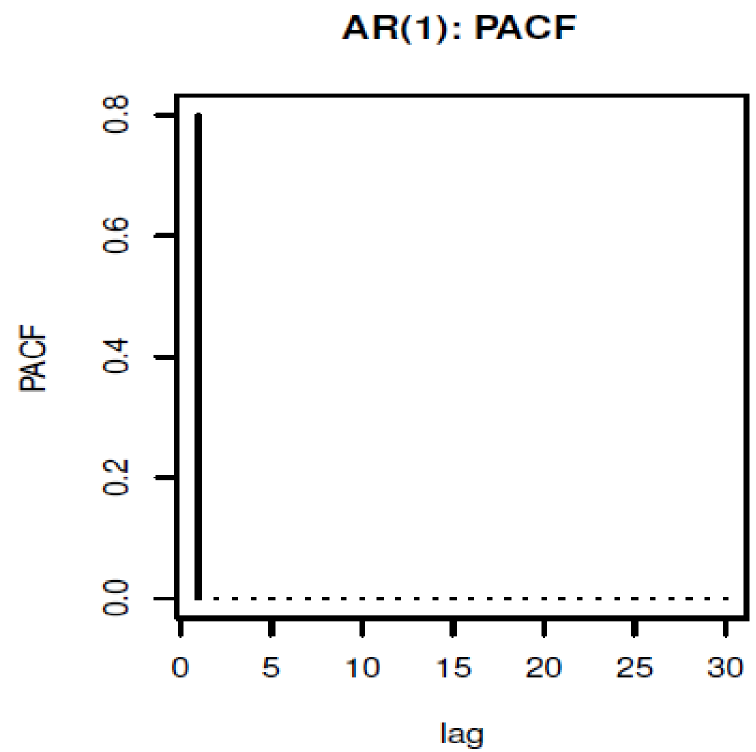
시차 p 이후에 절단됨

## AR(1)의 ACF & PACF

AR(1) ACF




AR(1) PACF




ACF는 시차가 증가함에 따라
지수적으로 감소

PACF는 시차 1 이후로 절단

# 5

# MA 모형

## MA 모형의 정의

MA 모형 (moving Average)

$X_t$가 과거 시점의 오차항들의 함수로 표현되는 모형

$Z_t \sim WN(0, \sigma^2)$를 따르는 innovations로 표현

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

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

## MA 모형의 특성방정식

MA(q) 모형

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

후향 연산자로 표현

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

$$X_t = \left(1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q\right) Z_t = \boldsymbol{\theta}(\boldsymbol{B})\boldsymbol{Z_t}$$

# 5 MA 모형

## MA 모형의 특성방정식

MA(q) 모형

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

후향 연산자로 표현

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

$$X_t = \left(1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q\right) Z_t = \boldsymbol{\theta}(\boldsymbol{B})\boldsymbol{Z_t}$$

MA(q)의 특성방정식 $\boldsymbol{\theta}(\boldsymbol{B})$

## MA 모형의 조건

MA 모형의 조건

정상성 (stationarity)

시계열의 확률적 특성이 시차에만 의존하는 특성

인과성 (casuality)

t 시점의 관측값이 과거 시점의 오차항으로 설명되는 특성

가역성 (invertibility)

t 시점의 오차항이 과거 시점의 관측값으로 설명되는 특성

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \quad for\ all\ t, \qquad \sum_{j=0}^{\infty} |\pi_j| < \infty$$

# 5 MA 모형

## MA 모형의 조건

MA(1) 모형을 통해 증명

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

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

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

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

무한등비급수의 합 형태

**선형과정이 되기 위한 조건**

innovation들의 계수의 합 유한해야 하는

absolutely summable 조건 만족

⋮

기하급수의 수렴조건에 의해

$|\phi_1| < 1$

⋮

$\theta(B) = 0$의 근의 절댓값이

1보다 커야 함

## MA 모형의 조건

MA(1) 모형을 통해 증명

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

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

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

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

무한등비급수의 합 형태

가역성 만족

### 선형과정이 되기 위한 조건

innovation들의 계수의 합 유한해야 하는

absolutely summable 조건 만족

기하급수의 수렴조건에 의해

$$|\boldsymbol{\phi_1}| < \mathbf{1}$$

$\theta(B) = 0$의 근의 절댓값이

1보다 커야 함

# 5 MA 모형

## MA 모형의 ACF

$$\gamma_X(h) = \sum_j \psi_j \psi_{j+h} \sigma^2$$

백색잡음 WN(0, $\sigma^2$)를 innovation으로 갖는 선형과정의 ACVF

백색잡음은 확률 변수 간 상관관계가 없어야 하므로

인덱스가 다른 경우에는 공분산이 0

인덱스가 겹치는 innovations에 대해서만

계수의 합을 구하면 됨

# 5 MA 모형

## MA 모형의 ACF

$h = 0$

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

$$\boldsymbol{\rho(0) = 1}$$

$h = 1$

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

$$\boldsymbol{\rho(1) = \frac{\theta}{1+\theta^2}}$$

$h \geq 2$

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

$$\boldsymbol{\rho(2) = 0}$$

## MA 모형의 ACF

$h = 0$

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

$$\rho(0) = 1$$

MA(q) 모형은 시차 h가 q보다 커지면

ACF가 절단되는 특성을 가짐

$h = 1$

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

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

$h \geq 2$

시차 h가 q보다 커지면 인덱스가

겹치는 innovation 쌍이 없기 때문

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

$$\rho(2) = 0$$

# 5 MA 모형

## MA 모형의 PACF

MA(1) 모형의 PACF

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

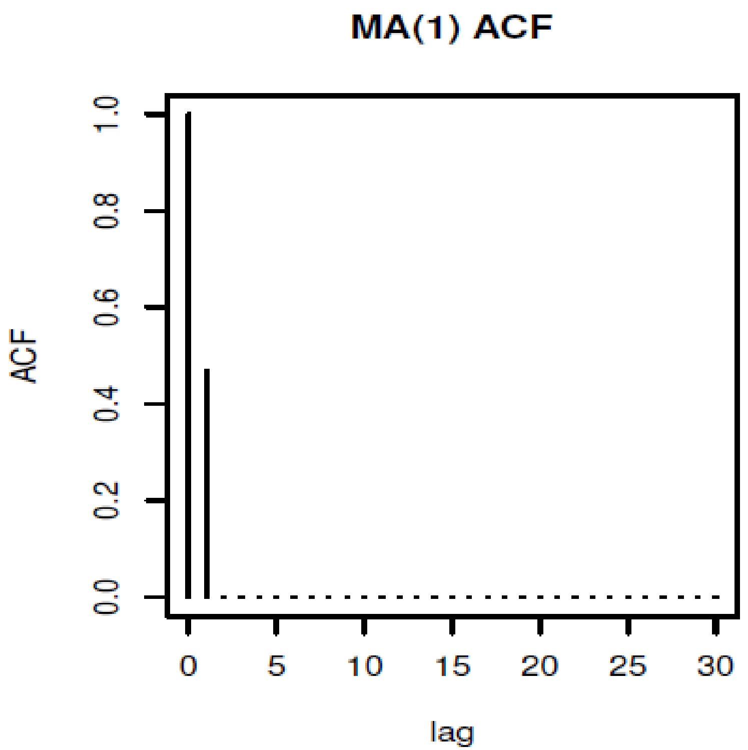
Crammer 공식 사용

시차 k가 커질수록 PACF는 점점 0으로 수렴

# 6 MA 모형

## MA(1)의 ACF & PACF

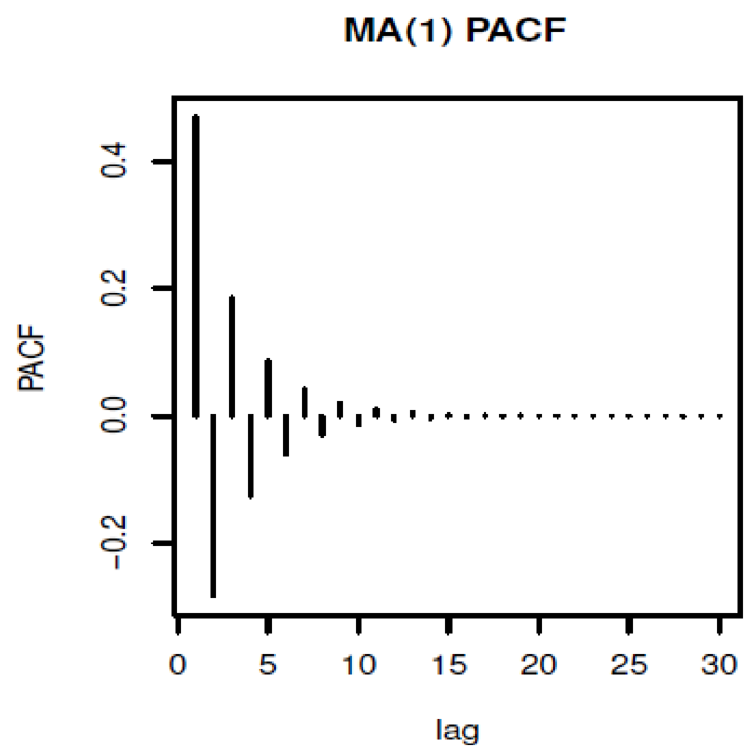
MA(1) ACF




ACF는 시차(lag) **1까지만 존재**

2부터 절단

MA(1) PACF




PACF는 시차가 커질수록

점점 **0으로 수렴**

# 6

# ARMA 모형

## ARMA 모형의 정의

### ARMA 모형

AR 과 MA 를 동시에 포함하는 모형

p와 q는 각각 AR, MA의 차수임

$z_t \sim WN(0, \sigma^2)$ 일 때, 다음과 같이 표현

ARMA(1,1)

$$x_t - x_{t-1} = z_t + z_{t-1}$$

ARMA(p,q)

$$x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - \cdots - \phi_p x_{t-p}$$
$$= z_t + \theta_1 z_{t-1} + \theta_2 z_{t-2} + \cdots + \theta_q z_{t-q}$$

## ARMA 모형의 정의

### ARMA 모형

ARMA 모형은 AR 과 MA 를 동시에 포함하는 모형

p와 q는 각각 AR,MA의 차수임

$Zt \sim WN(0, \sigma^2)$ 일 때, 다음과 같이 표현

ARMA(1,1) $x_t - x_{t-1} = z_t + z_{t-1}$

ARMA(p,q) $x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - \cdots - \phi_p x_{t-p}$

$= z_t + \theta_1 z_{t-1} + \theta_2 z_{t-2} + \cdots + \theta_q z_{t-q}$

## **통합모형**의 필요성

**시계열 자료를 AR 또는 MA만 이용하여 설명하려면**

**p, q값이 너무 커질 가능성이 있음**

**ARMA를 이용하면 모수를 절약 할 수 있다는 장점이 있음**

## 특성방정식

ARMA(p,q)의 특성 방정식

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

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

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

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

후향 연산자를 활용하면 이와 같이 나타낼 수 있음

## ARMA 모형의 조건

AR, MA모형의 조건을 모두 충족해야 함

정상성

인과성

가역성

EXAMPLE : ARMA(2,2) 모형

$$x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} = z_t + \theta_1 z_{t-1} + \theta_2$$

$$(1 - \phi_1 B - \phi_2 B^2) X_t = (1 + \theta_1 B + \theta_1 B^2) Z_t$$

정상성을 만족하기 위해

$\phi(B)X_t = 0,\ \theta(B)z_t = 0$의 근의 절대값이 1보다 커야 함

# 6 ARMA 모형

## ARMA 모형의 조건

AR, MA모형의 조건을 모두 충족해야 함

정상성 | 인과성 | 가역성

EXAMPLE : ARMA(2,2) 모형

$$x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} = z_t + \theta_1 z_{t-1} + \theta_2$$

추가적으로 **식별성**을 갖춰야 함

정상성을 만족하기 위해

$\phi(B)X_t = 0,\ \theta(B)z_t = 0$의 근의 절대값이 1보다 커야 함

## ARMA 모형의 조건

### 식별성이란?

AR, MA모형의 조건을 모두 충족해야 함

**주어진 파라미터 조합에 대해 하나의 모형만이 대응 되는 특성**

$$(1-\phi B) = (1+\theta B) \Leftrightarrow \phi = -\theta$$

정상성 인과성 가역성

위와 같은 경우, ARMA(1,1) 모형의 식은 $X_t = Z_t$
ARMA(1,1)인지 WN인지 식별할 수 없는 문제가 생김

추가적으로 식별성을 갖춰야 함

**$\phi + \theta \neq 0$ 의 식별성에 대한 조건이 필요.**

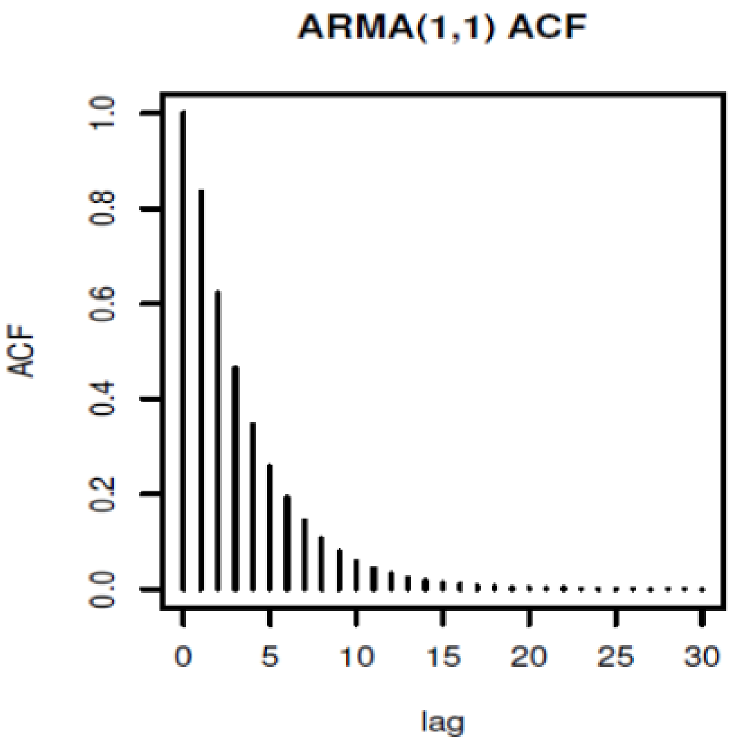
정상성을 만족하기 위해

$\phi(B)X_t = 0,\ \theta(B)z_t = 0$의 근의 절대값이 1보다 커야 함

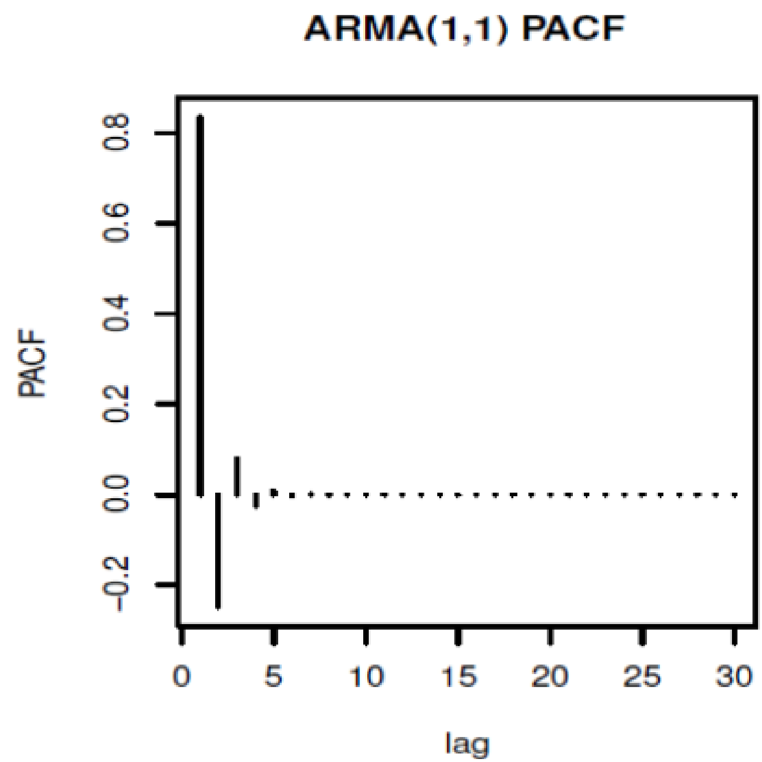
# 6 ARMA

## ARMA(1,1) ACF & PACF

ARMA(1,1) ACF




ARMA(1,1) PACF




ARMA모형의 경우 ACF와 PACF 모두 지수적으로 감소함으로

모형 식별을 위해 추가적인 정보가 더 필요함

# 7

# 모형의 적합 절차

# 7 모형의 적합 절차

## 적합 절차 순서

오차항을 모델링하기 위한 적절한 모델을
선택하기 위한 모형의 적합 절차

| | | | |
|---|---|---|---|
| 모형식별 | 모수 추정 | 모형 진단 | 예측 |

# 7 모형의 적합 절차

## 모형 식별

ARMA 모형은 ACF와 PACF 모두 지수적으로 감소하기 때문에
ACF, PACF만으로 모형을 식별하는데 한계가 있음

AIC, AICC, BIC 등의 Information Criteria를 참고하여 식별 가능

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

IC는 모형의 적합도와 복잡도를 함께 고려

## Information Criteria

AIC (Akaike Information Criteria)

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

AICC (AIC bias corrected)

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

BIC(Bayesian information Criteria

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

IC를 최소화 하는 p,q값을 얻게 됨

## Information Criteria

## AIC vs AICC

둘 다 모델의 적합도와 복잡도를 고려한 모델 선택 지표

AIC (Akaike Information Criteria)

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

AIC는 iid가정 때문에

시간적 의존성(Time dependency) 을 설명 x

AICC (AIC bias corrected)

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

BIC(Bayesian information Criteria

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

시간적 의존성(Time dependency) 에 대해

보정(correction)을 한 지표로 AICC 만들어짐

IC를 최소화 하는 p,q값을 얻게 됨

## 모수 추정

### 최대가능도추정법 (MLE)

관측된 시계열의 결합확률밀도함수의 가능도 함수를 최대화하는 모수를 추정량으로 구함

### 최소제곱법 (LSE)

오차 제곱합이 가장 작게 되도록 모수의 추정량을 구함

### 적률추정법 (MME/MoM)

모집단의 적률을 상응하는 표본의 적률로 대체한 후, 방정식을 풀어 모수의 추정량을 구함

7 모형의 적합 절차

모수 추정

## MLE 수렴 속도 올리기

일반적으로 MLE의 성능이 가장 좋지만 MLE를 구하기 위해선

최적화가 필요하며 최적화 알고리즘에는 초기값이 필요

최소제곱법 (LSE)

오차 제곱합이 가장 작게 되도록 모수의 추정량을 구함

LSE나 MME와 같은 추정값을 초기값으로 이용하면

MLE의 수렴속도가 빨라질 수 있음

# 7 모형 적합 절차

## 모형의 적합성 진단

### 모수에 대한 검정

모형의 조건을 만족하는가?

↓

① 정상성 & 가역성

:모수의 절대값이 1보다 작은지

② 식별성 : 모수의 합이 0이 아닐 것

### 잔차에 대한 검정

추세, 계절성, 이상치가 없는가?

↓

① 자기상관성 유무

: 잔차에 대한 ACF, PACF

: Ljung-Box test/McLeod-Li test

② 정규성 : QQplot / Jarque-Bera test

## 예측

기본적으로 t+h시점에 대한 예측 추정값은 {1, $\boldsymbol{X_1}$, $\boldsymbol{X_2}$, . . .}의 선형결합이라고 가정

$$\boldsymbol{PnXn{+}h} = \boldsymbol{a0} \cdot \boldsymbol{1} + \boldsymbol{a1} \cdot \boldsymbol{Xn} + \boldsymbol{a2} \cdot \boldsymbol{Xn{-}1} + \cdots + \boldsymbol{an} \cdot X_1$$

# 7 모형 적합 절차

## 예측

기본적으로 t+h시점에 대한 예측 추정값은 $\{1, \boldsymbol{X_1}, \boldsymbol{X_2}, . . .\}$의 선형결합이라고 가정

$$\boldsymbol{PnXn{+}h} = \boldsymbol{a0} \cdot \boldsymbol{1} + \boldsymbol{a1} \cdot \boldsymbol{Xn} + \boldsymbol{a2} \cdot \boldsymbol{Xn{-}1} + \cdots + \boldsymbol{an} \cdot X_1$$

계수 $\{a_0, a_1, . . . , a_n\}$ 에 대한 추정은

MSPE (Mean Squared Prediction Error)를 최소화하는 값을 사용

$$\boldsymbol{MSPE} = \boldsymbol{E}[\boldsymbol{Xn{+}h} - \boldsymbol{PnXn{+}h}\ ]^2$$

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

# 7 모형 적합 절차

## 예측

기본적으로 t+h시점에 대한 예측 추정값은 $\{1, \boldsymbol{X_1}, \boldsymbol{X_2}, ...\}$의 선형결합이라고 가정

$$\boldsymbol{PnXn{+}h} = \boldsymbol{a0} \cdot \boldsymbol{1} + \boldsymbol{a1} \cdot \boldsymbol{Xn} + \boldsymbol{a2} \cdot \boldsymbol{Xn}-\boldsymbol{1} + \cdots + \boldsymbol{an} \cdot X_1$$

계수 $\{a_0, a_1, ..., a_n\}$ 에 대한 추정은

MSPE (Mean Squared Prediction Error)를 최소화하는 값을 사용

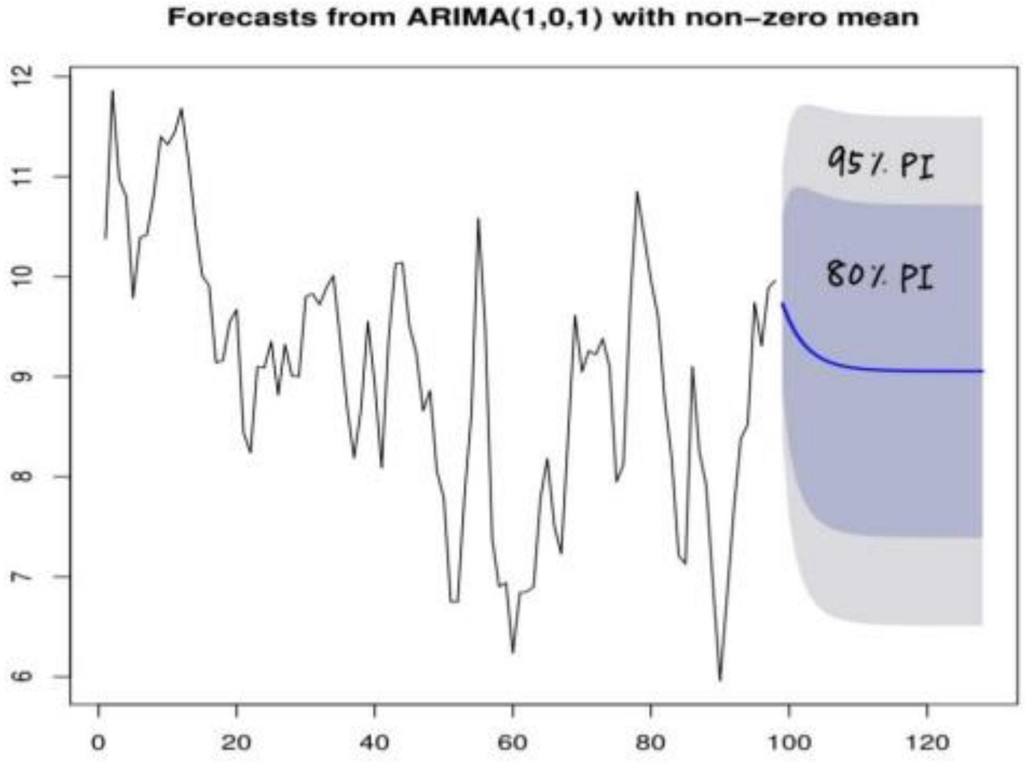
계산 방식은 회귀분석에서의 LSE계산 방법과 동일함

하지만 직접계산은 사실상 불가능하고 컴퓨터를 이용

## 예측

MSPE를 이용하면 PI(Prediction Interval)도 계산 가능

$$P_n X_{n+h} \pm 1.96 \cdot \sqrt{MSPE}$$



## 부록 - 단위근 검정

### 단위근 검정

데이터가 정상성을 만족하는지 않는지 검정하는 방법 중 하나

클린업 1주차에서 정상성에 대한 검정법으로 살펴본

Kpss, ADF, PP test가 단위근 검정법의 일종

# 7 모형 적합 절차

## 부록 - 단위근 검정

EXAMPLE : AR(p) 모형

$$x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - \cdots - \phi_p x_{t-p}$$
$$= \left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_\mathrm{p} B^\mathrm{p}\right) = Z_t$$

정상성을 만족하기 위해서 $\phi(B)$=0의 모든 근의 절댓값이 1보다 커야함

근의 크기가 1인 근을 **단위근**

## 부록 - 단위근 검정

### 단위근 검정

$H_0 : \phi = 1 \ vs \ H_1 : |\phi| < 1$

귀무가설이 기각되고 대립가설이 채택되면

정상 시계열임을 알 수 있음

모형의 진단에서 모형의 조건에 대한 만족 여부를 확인하고자 할 때,

단위근 검정 사용 가능

# 다음 주 예고

1. ARIMA
2. SARIMA
3. 이분산 시계열 모형
4. 시계열 데이터 with ML/DL

# 감사합니다!