클린업 3주차

5팀 시계열자료분석팀

염예빈
한유진
이재현
박세령
이정우

INDEX

1

2주차 복습

자기상관계수(Auto-Correlation Funciton)

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

이때 $\gamma_0 = Var(X_t) = E\left[(X_t - \mu)^2\right]$

표본자기상관함수(SACF)

$$\widehat{\rho_{(k)}} = \frac{\widehat{\gamma_{(k)}}}{\widehat{\gamma_{(0)}}}$$

$\rho(k) = \ Corr(X_t, X_{t+k})$

자기상관함수(ACF)는

시차 k에서 자기상관관계가

존재하는지 나타내는 척도

서로 다른 두 시점의

상호 연관관계를 나타냄

PACF(Partial Autocorrelation Function) : 편자기상관함수

PACF

:두 변수를 제외한 모든 변수의 영향을 제거한 상태에서,

두 변수사이의 순수한 상호연관관계

AR 모형

자기 회귀 모형

= Auto Regressive Model

자기 자신을 과거 시점에 회귀

현재의 자료 ← 설명O ← 과거의 자료

\+

현재의 자료 ← 설명X ← 오차항

ACF/PACF AR/MA/ARMA 모형의 적합

MA 모형

이동 평균 모형
= Moving Average Model

현재의 관측값이 현재와 과거의 설명해 주지 못하는 부분(오차)의 선형결합으로 표시되는 모형

현재의 자료 ← 선형결합 ← 과거의 자료의 오차

ARMA 모형

자기 회귀 이동 평균 모형
= **A**uto **R**egressive **M**oving **A**verage Model

AR 모형의 차수 p, MA 모형의 차수 q인 모형

AR과 MA를 동시에 사용한다면 추정해야 할 모수의 개수를 줄일 수 있다!

ACF/PACF AR/MA/ARMA 모형의 적합

정리

정상성, 인과성 가역성 모두 만족 해야해!!

AR + MA = ARMA

"$\Phi(B) = 0$ 의 근의 절댓값" > 1

"$\theta(B) = 0$ 의 근의 절댓값" > 1

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

ACF/PACF AR/MA/ARMA **모형의 적합**

모형 적합 Flow

정상 시계열 자료

모형의 식별 — 차수 p , q 결정

모수의 추정 — 최대가능도추정법, 최소제곱, 적률추정법

모형의 진단 — 잔차분석, 과대적합진단

NO

YES

예측모형으로 선택

2

ARIMA

ARIMA

Auto-Regressive Integrated Moving Average

정상화 방법 중 차분 후

ARMA(p,q) = ARIMA(p,d,q)

즉, 차분과정까지 포함한 모델!!

Trend(추세) + ARMA(p,q)형태가 있을 때 사용!

ARIMA의 표현

ARIMA에서 I는 무슨 뜻인가요?

I는 누적(integrated)의 의미

$\phi(B)(1-B)^{q}$ $u) = \theta(B)\varepsilon_t$

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

$\phi_B(1-B)Z_t = \phi(B)\varepsilon_t$에서

$W_t = (1-B)Z_t$이라고 한다면, $W_t = Z_t - Z_{t-1}$

$$Z_t = W_t + Z_{t-1} = W_t + (W_{t-1}+Z_{t-2}) = W_t + W_{t-1} + (W_{t-2}+Z_{t-3})$$

$= \cdots = Z_0 + \sum_{j=1}^{t} W_j$ $\therefore$ Z_t부분은 W_t의 누적합이 된다.

ARIMA의 표현

ARIMA(p, d, q) :

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

$$\left(1-\phi_1\mathrm{B}-\phi_2\mathrm{B}^2-\cdots-\phi_{\mathrm{p}}\mathrm{B}^{\mathrm{p}}\right)(1-\mathrm{B})^{\mathrm{d}}Z_t$$

$$=\left(1+\theta_1B+\theta_2B^2+\cdots+\theta_qB^q\right)\varepsilon_t$$

ARIMA의 표현

ARIMA(p, d, q) :

$$\phi(B)(1-B)^d(Z_t-\mu)=\theta(B)\varepsilon_t$$

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

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

cf) ARIMA(p, 0, q) = ARMA(p, q)

ARIMA(p, d, q)

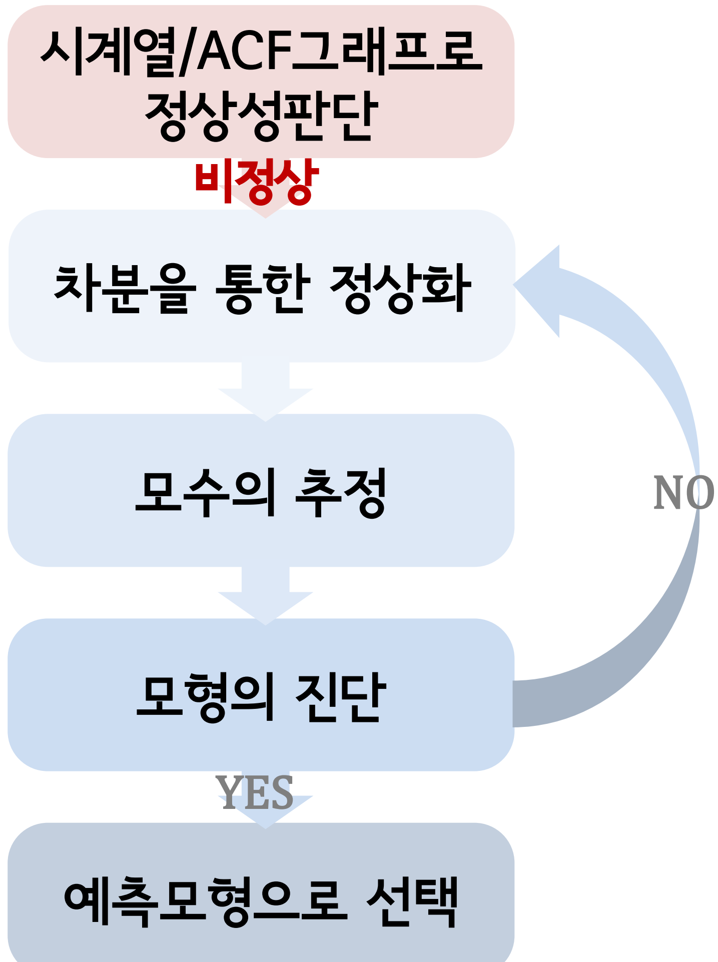
p : AR의 차수

d : 차분을 몇 번 했는가!

q : MA의 차수

cf) ARIMA(p, 0, q) = ARMA(p, q)

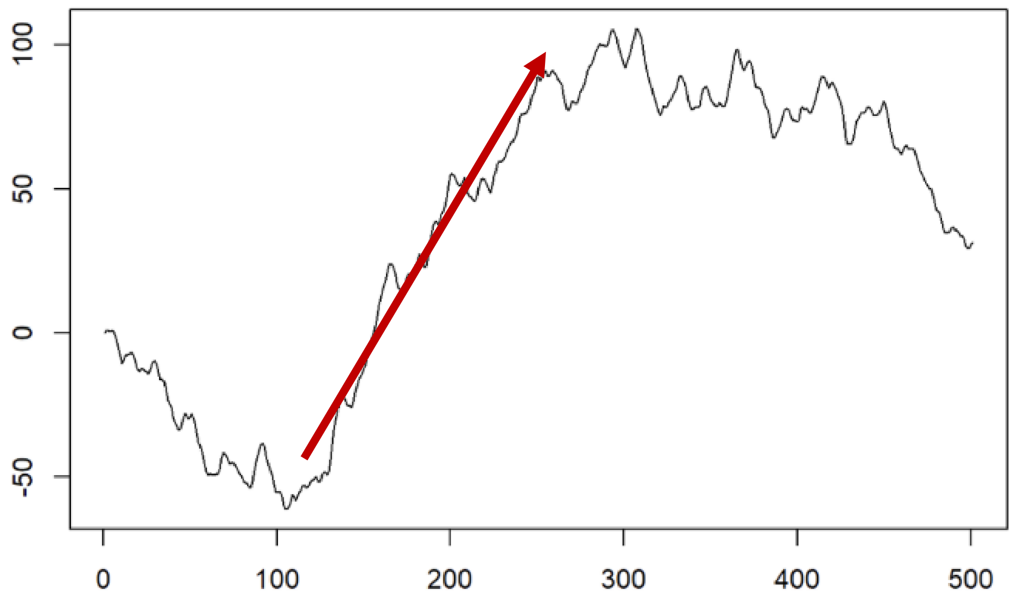
ARIMA 모형의 적합



ARIMA 모형의 적합

시계열/ACF그래프로
정상성판단

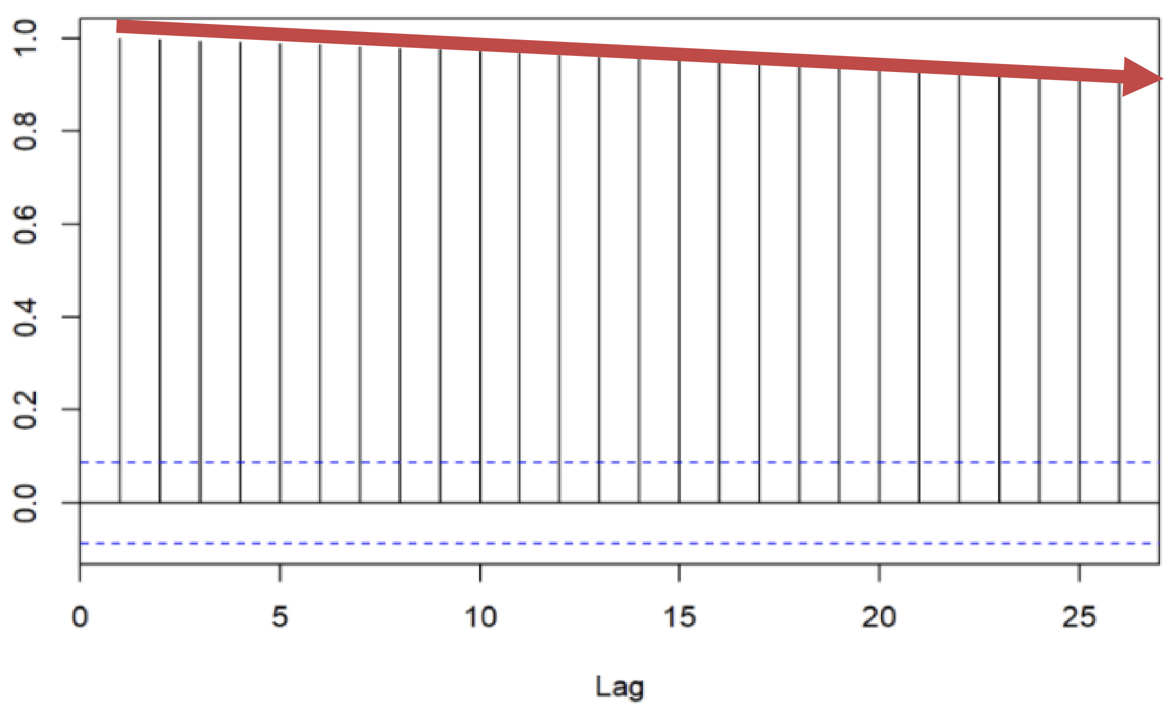
시계열 plot을 보았을 때, 추세가 존재하는 것 처럼 보임!



ARIMA 모형의 적합

시계열/ACF그래프로
정상성판단

추세가 존재한다면, ACF가 천천히 감소



ARIMA 모형이 필요해 보임!!

ARIMA 모형의 적합

일반적으로 d=1 or 2 낮은 차수를 많이 사용.
과대 차분 위험때문!!!

차분을 통한 정상화

차분한 시계열이 ARMA(p,q) 과정을 따르는지 살펴보고
ARIMA(p,d,q)에 적합!

ARIMA 모형의 적합

일반적으로 d=1 or 2 낮은 차수를 많이 사용.
과대 차분 위험때문!!!

차분을 통한 정상화

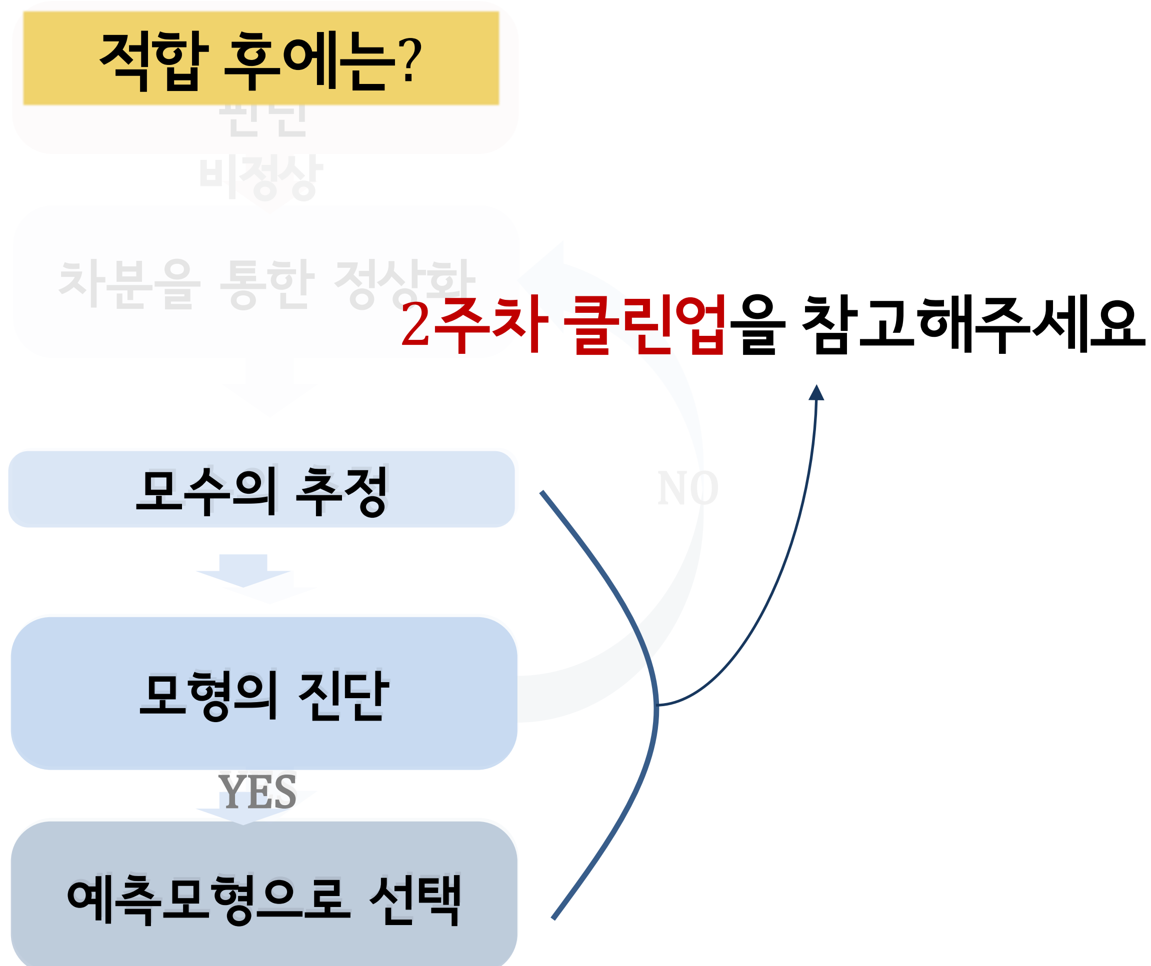
과대 차분이란?

이미 정상화가 된 경우에도 차분을 시도 하는 것!

과대 차분이란?
이미 정상화가 된 경우에도 차분을 시도 하는 것!

차분한 시계열이 ARMA(p,q) 과정을 따르는지 살펴보고
ARIMA(p,d,q)에 적합!

ARIMA 모형의 적합


적합 후에는?
비정상
차분을 통한 정상화
2주차 클린업을 참고해주세요!
NO
모수의 추정
모형의 진단
YES
예측모형으로 선택

3

SARIMA

계절성을 가지는 시계열

$$Y_t = S_t + I_t$$

1. 결정적 계절성

항상 식을 통해 도출되는 S_t

$$S_t = a_0 + \sum_{j=1}^{k} a_j \cos(\lambda_j t) + b_j \sin(\lambda_j t))$$

회귀, 평활 등을 통하여
계절성분 S_t를 분해하여 제거

2. 확률적 계절성

$(1 - B^d)S_t = X_t$를 만족시키는
확률과정(이때, X_t는 정상적 확률과정)

SARIMA 모형

순수 SARIMA

시계열이 순수하게 계절형일 경우 때 적합할 수 있는 모형

$$\mathbf{ARIMA(P, D, \mathit{Q})_{\mathit{s}}}$$

$$\Phi(\mathrm{B}^s)(1-B^s)^D Z_t = \delta + \Theta(B^s)u_t \quad \text{혹은} \quad \Phi(\mathrm{B}^s)(1-B^s)^D Z_t = \Theta(B^s)\varepsilon_t$$

$$\Phi(\mathrm{B}^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_P B^{Ps}$$

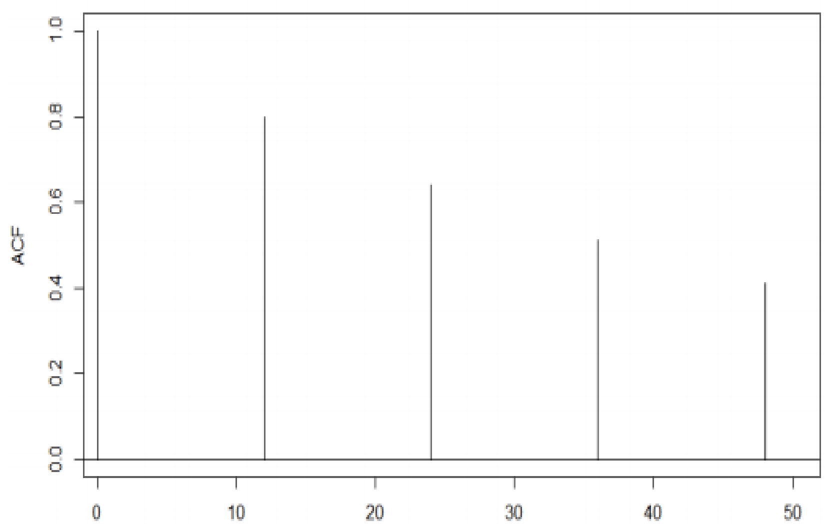
$$\Theta(\mathrm{B}^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \cdots - \Theta_Q B^{Qs}$$

- 현재의 관측치를 계절성을 고려하여 $\mathrm{Z_{t-s}}, Z_{t-2s}, \ldots$들로 설명
- 해당 주기 관측치는 P개를 고려하며, 해당 주기 오차항은 Q개를 고려
- 오차 ε_t는 백색잡음 , D번 계절차분

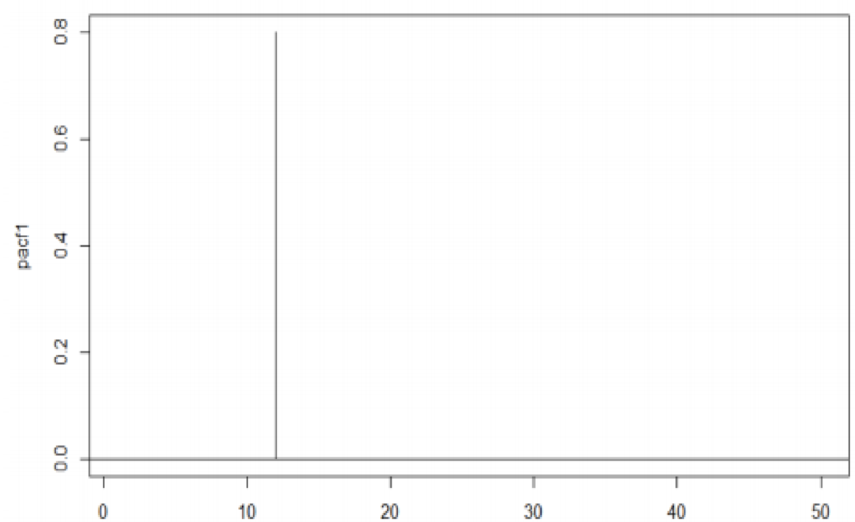
순수 SARIMA

- 비계절 요소에 대한 부분을 전혀 고려하지 않기에, **사용이 제한적**
- 오차가 백색잡음이므로, 계절주기에 해당되지 않는 시차의 **ACF는 0**

[$\mathrm{ARIMA}(1,0,0)_{12}$ ACF]

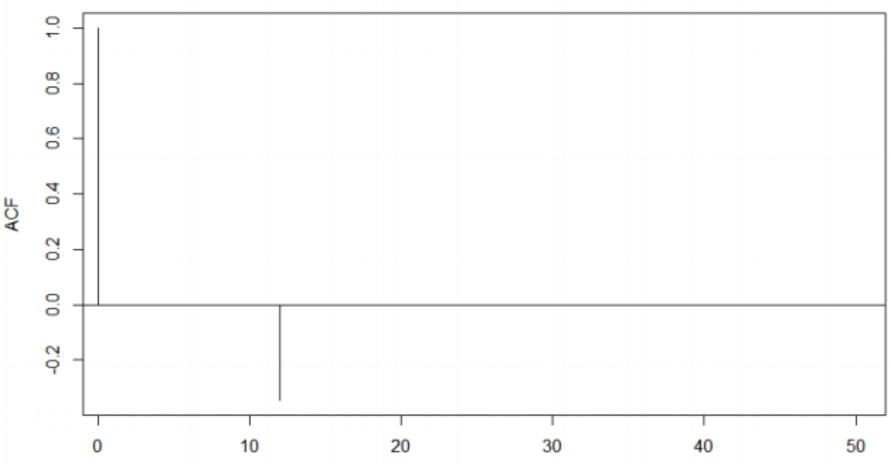


[$\mathrm{ARIMA}(1,0,0)_{12}$ PACF]

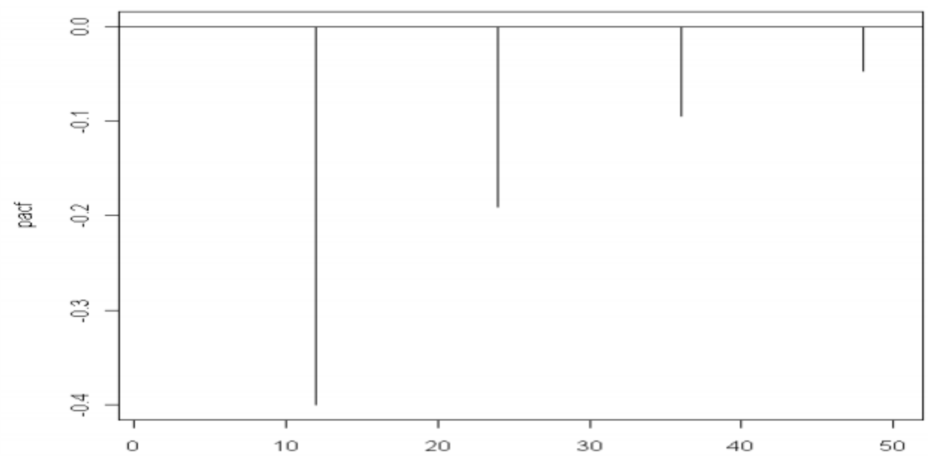


순수 SARIMA – ACF, PACF

[$\mathrm{ARIMA}(0,0,1)_{12}$ ACF]



[$\mathrm{ARIMA}(0,0,1)_{12}$ PACF]



- ARMA 모형의 ACF와 PACF와 비슷함
- 단, **계절주기 차수에서만** 자기상관성을 가짐

승법 SARIMA

$$\mathrm{Y}_{\mathrm{j+12t}} - \Phi_1 Y_{j+12(t-1)} - \cdots - \Phi_P Y_{j+12(t-P)} = U_{j+12t} + \Theta_1 U_{j+12(t-1)} + \cdots + \Theta_Q U_{j+12(t-Q)}$$

$$\Phi(\mathrm{B}^{12}) Y_t = \Theta(B^{12}) U_t , \qquad U_t \sim WN(0, \sigma^2)$$

$\mathrm{U_t}$ 가 백색잡음이 아니라 ARMA 모형을 따를 수도 있음!

$$\mathrm{U_t} \sim ARMA(p, q)$$

$$\phi(\mathrm{B})\mathrm{U_t} = \theta(B)\varepsilon_t, \qquad \mathrm{U_t} = \phi(B)^{-1}\theta(B)\varepsilon_t$$

$$\Phi(\mathrm{B}^s) Y_t = \Theta(B^s)\phi(B)^{-1}\theta(B)\varepsilon_t$$

순수 SARIMA / 승법 SARIMA의 오차항

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

순수 SARIMA

$$\Phi(B^{12})Y_t = \Theta(B^{12})\underline{U_t}$$

$U_t \sim ARMA(p, q)$

승법 SARIMA

$$\Phi(B^{12})Y_t = \Theta(B^{12})\underline{\phi(B)^{-1}\theta(B)\varepsilon_t}$$

승법 SARIMA

$$\mathrm{U_t} \sim ARMA(p,q)$$

$$\phi(\mathrm{B})\mathrm{U_t} = \theta(B)\varepsilon_t, \qquad \boxed{\mathrm{U_t} = \phi(B)^{-1}\theta(B)\varepsilon_t}$$

$$\Phi(\mathrm{B}^s)Y_t = \Theta(B^s)\phi(B)^{-1}\theta(B)\varepsilon_t$$

$\mathrm{Y_t}$가 d번 차분, D번 계절차분을 마친 상태의 시계열이라면,

$$\mathrm{Y_t} = (1-B)^d(1-B^s)^D Z_t$$

$$\phi(\mathrm{B})\Phi(\mathrm{B}^{12})(1-B)^d(1-B^s)^D Z_t = \theta(B)\Theta(B^{12})\varepsilon_t$$

승법 SARIMA

순수 SARIMA와 다르게, 승법 SARIMA는 비계절적 요소를 고려하는 모형
순수 SARIMA모형의 오차가 ARIMA를 따름

$$ARIMA(P,D,Q)_s \ \& \ U_t \sim ARIMA(p,d,q)$$

순수SARIMA에 ARIMA가 곱하기로 붙어서 '승법'이라는 용어 사용

$$\boldsymbol{SARIMA(p,d,q)(P,D,Q)_s}$$

$$\phi(\mathrm{B})\Phi(\mathrm{B}^s)(1-B)^d(1-B^s)^D Z_t = \theta(B)\Theta(B^s)\varepsilon_t$$

$$\Phi(\mathrm{B}^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_P B^{Ps}, \Theta(\mathrm{B}^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \cdots - \Theta_Q B^{Qs}$$
$$\phi(\mathrm{B}) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p, \theta(\mathrm{B}) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$$

승법 SARIMA

$$\boldsymbol{ARIMA(p,d,q)(P,D,Q)_s}$$

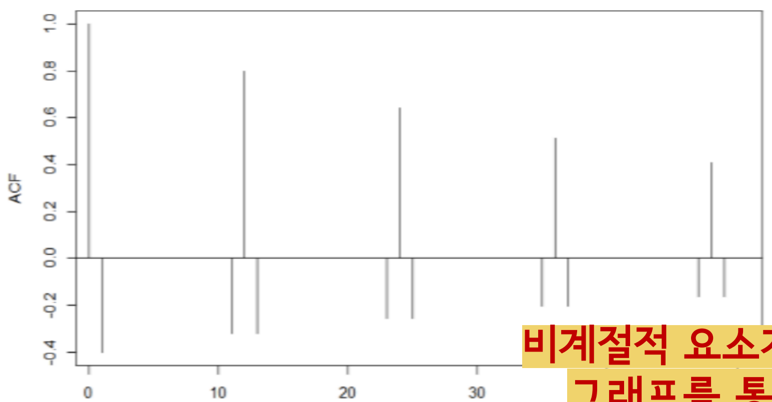
$$\phi(\mathrm{B})\Phi(\mathrm{B}^s)(1-B)^d(1-B^s)^D Z_t = \theta(B)\Theta(B^s)\varepsilon_t$$

비계절적 요소 고려

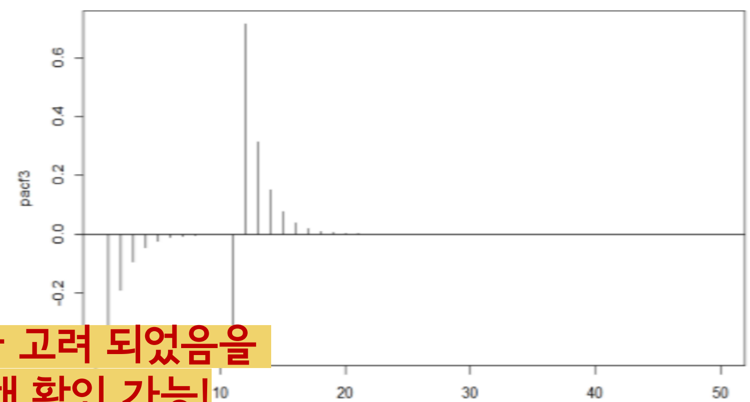
- **계절성과 추세를 고려한 모형**
- ARIMA 모형처럼 예측시점과 가까운 과거 자료를 이용할 뿐만 아니라, 원자료의 주기적 특성을 반영하여 그 이전 주기의 자료를 추가적으로 활용
- **순수SARIMA보다 비계절 모형과 혼합된 모형이 모수 절약의 원칙에 따라 더 자주 사용됨**

승법 SARIMA – ACF, PACF

[$ARIMA(0,0,1)(1,0,0)_{12}$ ACF]

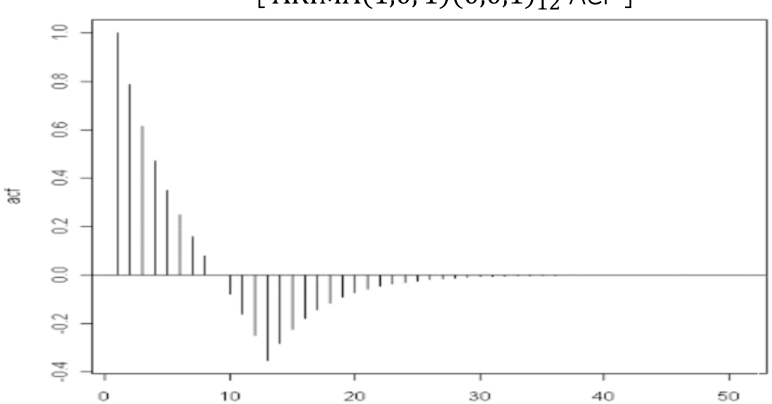


[$ARIMA(0,0,1)(1,0,0)_{12}$ PACF]

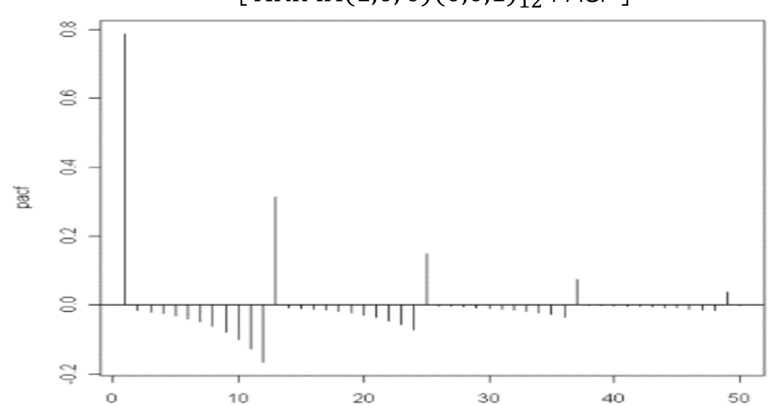


비계절적 요소가 고려 되었음을 그래프를 통해 확인 가능!

[$ARIMA(1,0,1)(0,0,1)_{12}$ ACF]



[$ARIMA(1,0,0)(0,0,1)_{12}$ PACF]

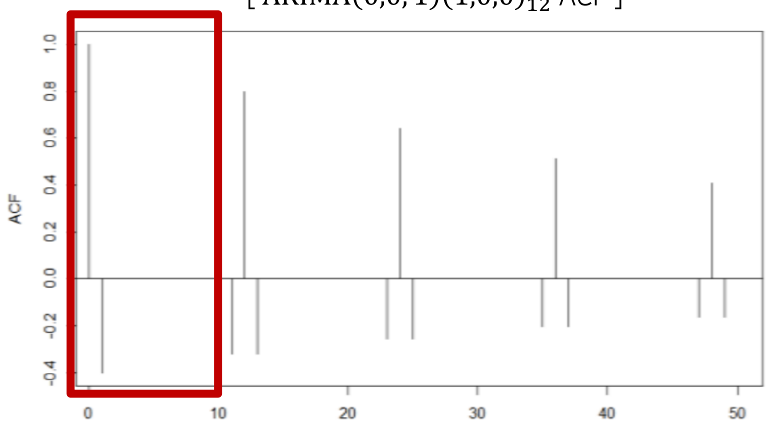


승법 SARIMA – ACF, PACF

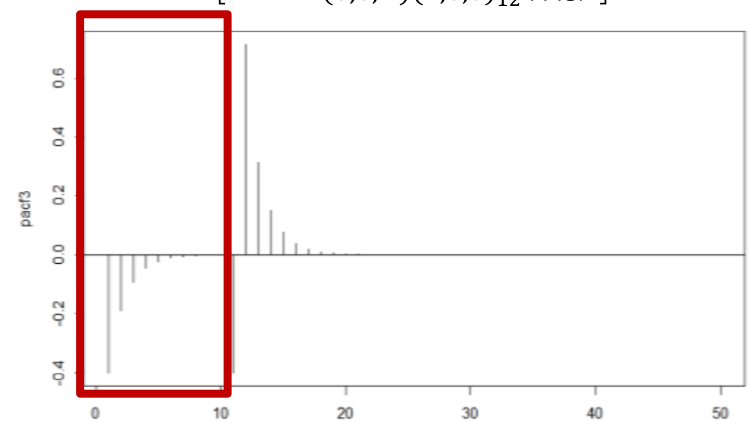
어떻게
확인하는데요..??!

승법 SARIMA – ACF, PACF

[ARIMA(0,0, 1)(1,0,0)$_{12}$ ACF]



[ARIMA(0,0, 1)(1,0,0)$_{12}$ PACF]



[SARIMA(1,0, 1)(0,0,1)$_{12}$ ACF]

[SARIMA(1,0, 0)(0,0,1)$_{12}$ PACF]

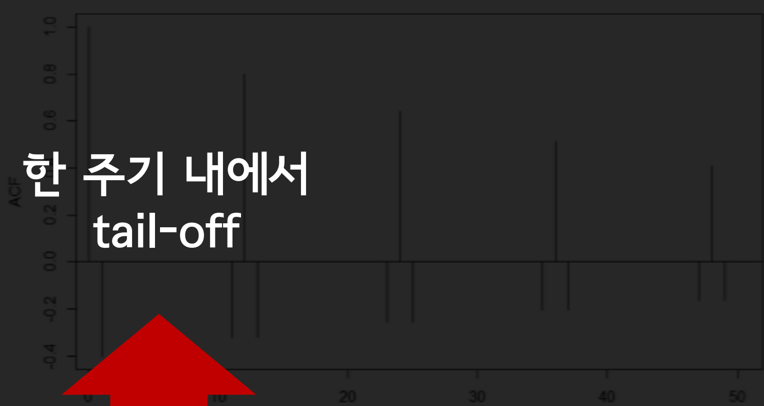
한 주기 내에서
cut-off

한 주기 내에서
tail-off

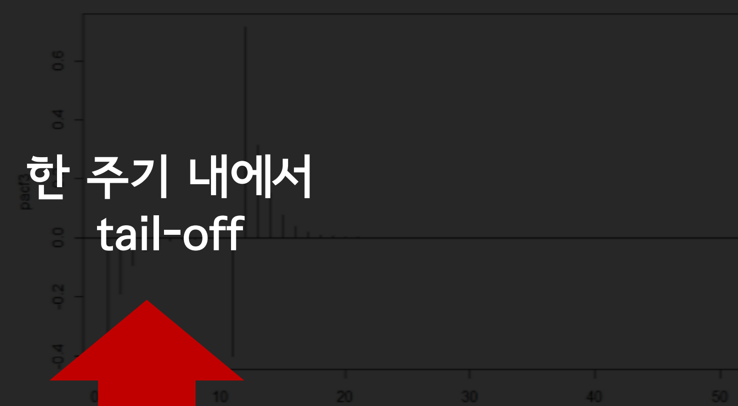
계절성을 가지는 시계열 | 순수 SARIMA | **승법 SARIMA**

승법 SARIMA – ACF, PACF

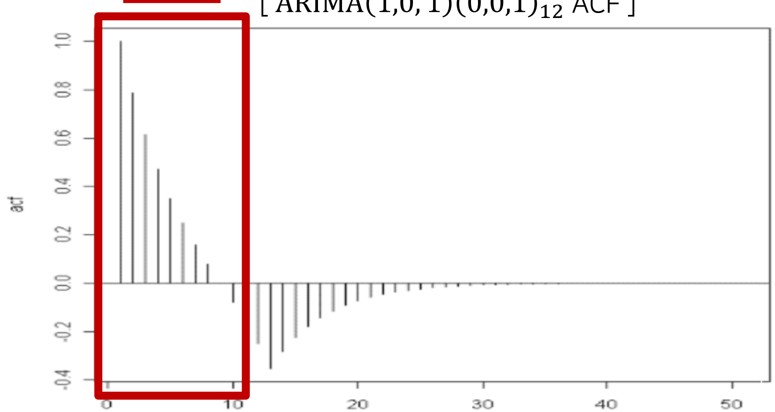
[$SARIMA(0,0,1)(1,0,0)_{12}$ ACF]



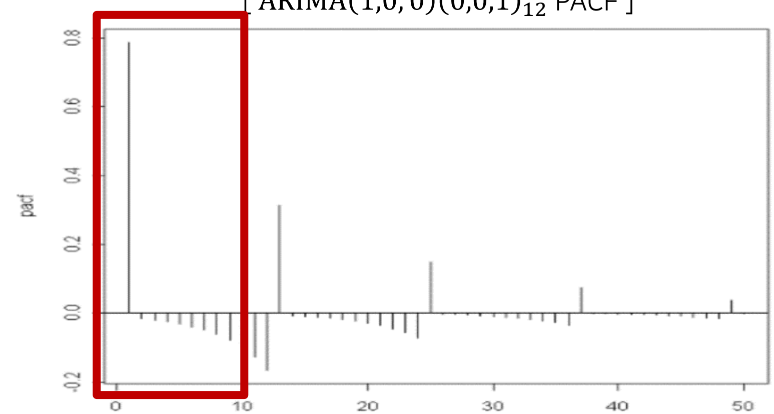
[$SARIMA(0,0,1)(1,0,0)_{12}$ PACF]



[$ARIMA(1,0,1)(0,0,1)_{12}$ ACF]



[$ARIMA(1,0,0)(0,0,1)_{12}$ PACF]



승법 SARIMA

순수 SARIMA

$$ARIMA(P,D,Q)_s$$

$$\Phi(B^s)(1-B^s)^D Z_t = \Theta(B^s)u_t$$

비계절적 요소를 고려하지 않음

승법 SARIMA

$$SARIMA(p,d,q)(P,D,Q)_s$$

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D Z_t = \theta(B)\Theta(B^s)\varepsilon_t$$

오차항 $u_t \sim ARIMA(p,d,q)$

비계절적 요소를 고려함

$$SARIMA(p,d,q)(P,D,Q)_s$$

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D Z_t = \theta(B)\Theta(B^s)\varepsilon_t$$

- 계절성과 추세를 고려한 모형
- ARIMA 모형처럼 예측시점과 가까운 과거 자료를 이용할 뿐만 아니라, 원자료의 주기적 … 자료를 추가적으로 활용
- 순수SARIMA보다 비계절 모형과 혼합된 모형이 모수절약의 원칙에 따라 더 자주 사용됨

4

ARFIMA

ARFIMA

ARFIMA모형은 왜 필요할까?

ARFIMA

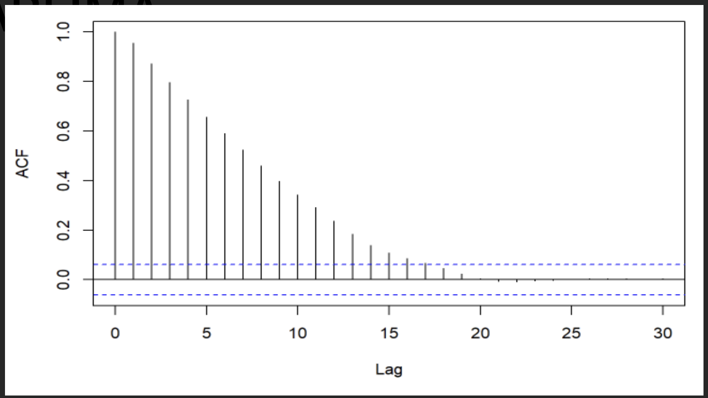
VS

AR

MA

ARMA

ARIMA



기존의 모형들은 ACF 모양이

기하급수적으로 감소

→ 단기억 확률과정

정상성을 만족하지만 ACF가 0으로

느리게 수렴하는 시계열 자료에는

좋은 적합이 아님

AR

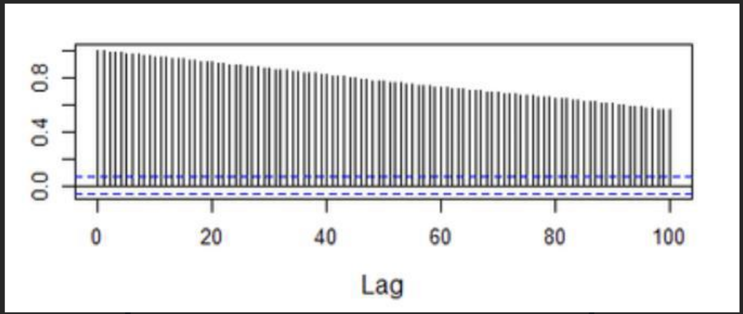
MA

ARMA

ARIMA

ARFIMA

ARFIMA



그 한계점을 보완한 모형이 ARFIMA!

ARFIMA 모형의 ACF는

0으로 느리게 수렴

→ 장기억 확률과정

ARFIMA

Auto **R**egressive **F**ractionally **I**ntegrated **M**oving **A**verage

유리수

ARIMA VS ARFIMA

d : 양의 정수

$0 < d < \frac{1}{2}$

ARFIMA

Auto Regressive Fractionally Integrated Moving Average

$\text{ARFIMA}(p, d, q) : \phi(B)(1-B)^d(Z_t - \mu) = \theta(B)\varepsilon_t$, 이때 $0 < d < \frac{1}{2}$

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

$$= (1-\theta_1 B - \theta_2 B^2 - \cdots - \theta_p B^p)\,\varepsilon_t$$

ARFIMA

Auto Regressive Fractionally Integrated Moving Average

$$\text{ARFIMA(p, d, q)} : \phi(B)(1-B)^{\boxed{d}}(Z_t-\mu) = \theta(B)\varepsilon_t \text{, 이때 } 0 < d < \frac{1}{2}$$

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

1. 장기간 동안의 시계열 종속성 결정
2. -0.5에서 0일 경우 자기상관함이 0이 되는 문제 발생

ε_t

ARFIMA

장기억 확률과정 (long memory process)

자기 상관 함수 $\rho(k) \sim Ck^{2d-1},\ k \to \infty$ (단, $C > 0$)

- ACF 합이 무한대로 발산.

- ACF가 0으로 빠르게 수렴하지 않음.

ARFIMA 모형의 적합

ARFIMA모형의 적합

- ARFIMA에서 d는 최소제곱법, 최대우도추정법으로 찾을 수 있음.
- 이후 ARMA모형과 비슷하게 나머지 모수를 추정.
- R에선 arfima() 함수를 사용할 수 있음.

ARFIMA

```
#arfima 함수에 데이터만 입력하면 알아서 찾아준다!
sun = sunspots
arfima(sun)
```

```
##
## Call:
##   arfima(y = sun)
##
## Coefficients:
##          d      ar.ar1     ma.ma1     ma.ma2
## 0.24676848 0.97439271 0.64294637 0.08996052
```

AR 차수: 1

MA 차수: 2

d값

ARFIMA(1, 0.2468, 2) 적합

- A… 으로 찾을 수 있…
- 이후 ARMA모형과 비슷하게 나머지 모수를 추정한다.
- R에 … fima() 함수를 사용할 수 있다.

5

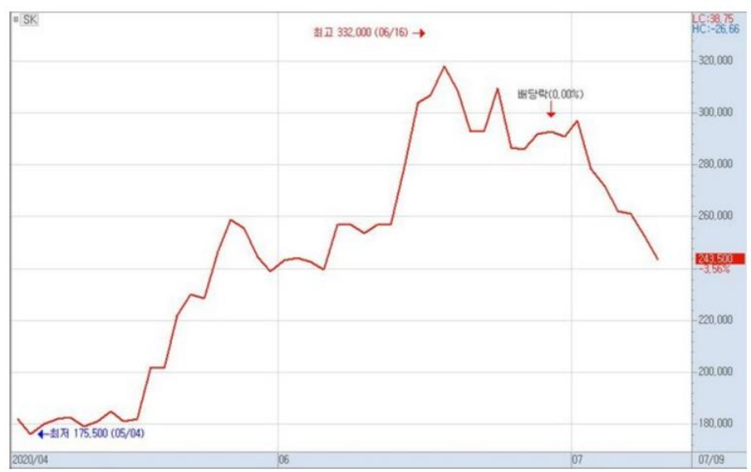
이분산 시계열 모형

이분산 시계열 모형의 필요성

ARIMA와 같은 전통적 시계열 모형

- 평균부분의 움직임에 초점을 맞춤.
- 등분산성을 가정.

ARCH



▲SK 주가 그래프. (키움증권 HTS)

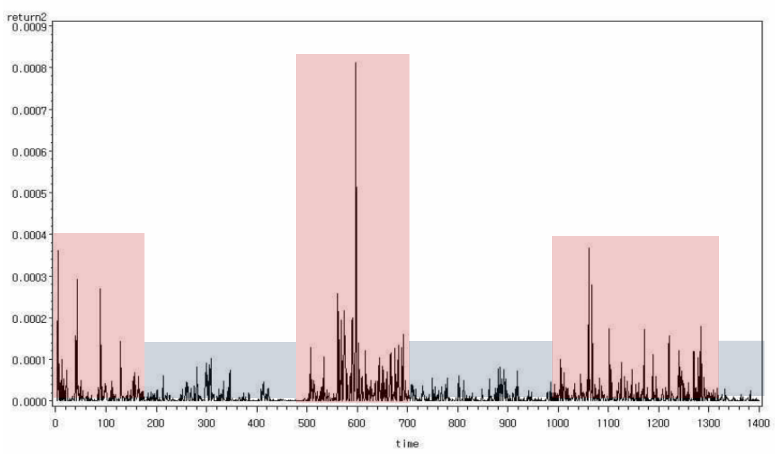
10일(현지 시간) 뉴욕 증시에서 테슬라 주가는 전날 대비 8.97% 오른 1025.05달러로 장을 마쳤다. 이는 종가 기준 테슬라의 사상 최고치다. / 그래프=김은실 디자이너

금융관련 재무시계열 모형은 **불확실성을 의미하는 분산부분**에

관심이 있고, 분산이 과거자료에 의존

→ 전통적 시계열 모형은 좋은 적합이 아니다!

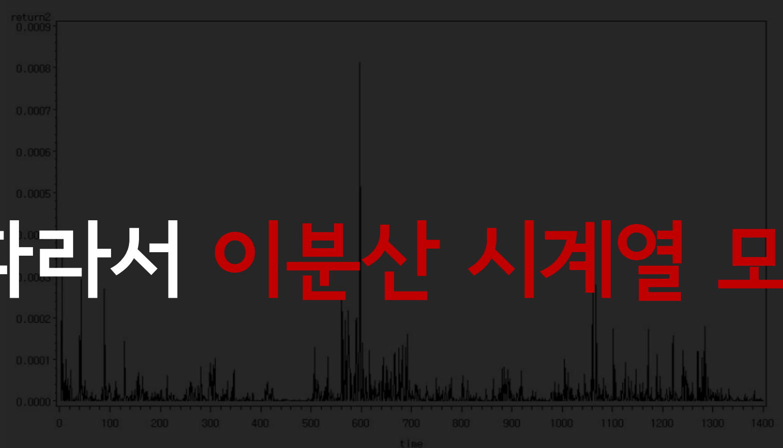
이분산 시계열 모형의 필요성



- **변동성 집중 현상** : 한번 큰 값을 기록하면 상당시간 동안 큰 상태로 지속된 후 다시 작은 값들이 상당기간동안 지속되는 현상.
- 변동성 집중 현상은 **분산에 자기상관성이 존재함**을 보여줌.

ARCH

이분산 시계열 모형의 필요성



따라서 이분산 시계열 모형이 필요!

- **변동성 집중 현상** : 한번 큰 값을 기록하면 상당시간 동안 큰 상태로 지속된 후 다시 작은 값들이 상당기간동안 지속되는 현상.
- 변동성 집중 현상은 **분산에 자기상관성이 존재함**을 보여준다.

ARCH

자기회귀 조건부 이분산 모형
= Auto Regressive Conditional Heteroskedasticity

변동성을 설명하기 위한 비선형 모델

현재 시점의 오차항 변동성 ← 설명O ← 과거 시점의 오차항 제곱

ARCH

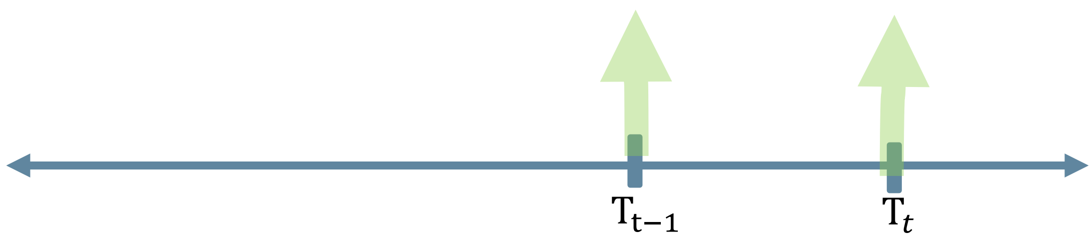
ARCH(p): t시점의 오차항의 변동성을 p시점 전의 오차항들의 제곱으로 설명하는 모형

ARCH

ARCH(p): t시점의 오차항의 변동성을 p시점 전의 오차항들의 제곱으로 설명하는 모형

[ARCH(1)]

$$\alpha_1 \varepsilon_{t-1}^2 + \alpha_0 = \sigma_t^2$$

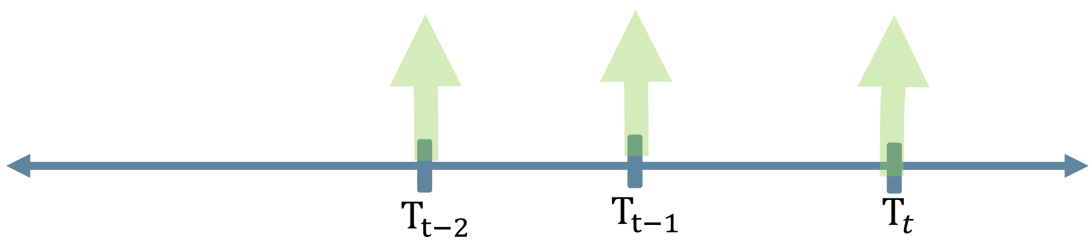


ARCH

ARCH(p): t시점의 오차항의 변동성을 p시점 전의 오차항들의 제곱으로 설명하는 모형

[ARCH(2)]

$$\alpha_2 \varepsilon_{t-2}^2 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_0 = \sigma_t^2$$

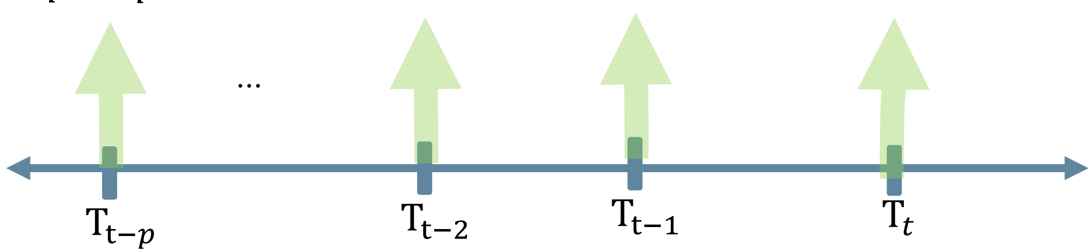


ARCH

ARCH(p): t시점의 오차항의 변동성을 p시점 전의 오차항들의 제곱으로 설명하는 모형

[ARCH(p)]

$$\alpha_p \varepsilon_{t-p}^2 + \cdots + \alpha_2 \varepsilon_{t-2}^2 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_0 = \sigma_t^2$$



ARCH

[표현: ε_t ~ ARCH(p)]

$\{\nu_t\}$ ~ i.i.d.N(0,1)
$\alpha_0 \geq 0$, $\alpha_j \geq 0$, j = 1, ⋯, p

$$\varepsilon_t = \sigma_t \cdot \nu_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \cdots + \alpha_p \varepsilon_{t-p}^2$$

t시점의 오차항의 변동성을

P시점 전의 오차항들의 제곱으로 설명

ARCH

ARCH(1) : $\varepsilon_t = \sigma_t \cdot \nu_t\ ,\ \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$

$$\varepsilon_t^2 = \sigma_t^2 \cdot \nu_t^2$$

$$= (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2) \cdot \nu_t^2$$

$$= (\alpha_0 + \alpha_1 (\sigma_{t-1}^2 \cdot \nu_{t-1}^2)) \cdot \nu_t^2$$

재귀적으로 계산해보면

$$\cdots$$

$$= \alpha_0 \sum_{j=1}^{n} \alpha_1^j \nu_t^2 \nu_{t-1}^2 \cdots \nu_{t-j}^2 + \alpha_1^{n+1} \nu_t^2 \cdots \nu_{t-n}^2$$

ARCH

[ARCH(1)]

$$\varepsilon_t^2 = \alpha_0 \sum_{j=1}^{n} \alpha_1^j \nu_t^2 \nu_{t-1}^2 \cdots \nu_{t-j}^2 + \alpha_1^{n+1} \nu_t^2 \cdots \nu_{t-n}^2$$

ε_t 가 ν_t 의 비선형적 형태

따라서 ARCH는 비선형 모델!

ARCH

ARCH모형의 검정

귀무가설: $H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$

- LM(largrange multiplier) 검정: **검정통계량 $n \cdot R_{\varepsilon}^2$** 은 점근적으로 $\chi^2(p)$를 따른다는 성질을 활용하여 검정.

- Ljung-Box Q 검정 : ARCH(p) 모형에서 ε_t^2 가 AR(p)모형을 따름을 이용

ARCH

[Ljung Box 검정]

```
Box.test(apple_ret_num^2, lag=4, type="Ljung")
```

ARCH(4)

```
##
##  Box-Ljung test
##
## data:  apple_ret_num^2
## X-squared = 161.84, df = 4, p-value < 2.2e-16
```

귀무가설 기각

→ ARCH모형이 유의하다!

AI

- LM

n ·

검정.

- Ljung-Box Q 검정 : ARCH(p) 모형에서 ε_t^2 가 AR(p)모형을 따름을 이용

ARCH

ARCH모형의 문제점

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \cdots + \alpha_p \varepsilon_{t-p}^2 + \cdots$$

추정해야 할 모수가 많음

→

- 추정량의 정확도가 떨어짐
- 비음조건(non-negative)을 만족하기 어려움

ARCH

ARCH모형의 문제점

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \cdots + \alpha_p \varepsilon_{t-p}^2 + \cdots$$

추정해야 할 모수가 많아지는
문제를 해결하려면?

- 추정량의 정확도가 떨어진다.

일반화된 모형이 필요! → GARCH 모형

…(non-negative)을 만족하기 어렵다.

GARCH

일반화 자기 회귀 이분산 모형
= Generalized Auto Regressive Conditional Heteroskedastic

ARCH 모형을 일반화한 모형

현재 시점의 오차항 변동성 ← 설명 ← 과거 시점의 오차항 제곱, 과거 시점의 변동성

GARCH

[표현: ε_t ~ GARCH(p,q)]

$\{\nu_t\}$ ~ i.i.d.N(0,1)
$\alpha_0 \geq 0,\ \alpha_j \geq 0,\ \beta_j \geq 0$
$j = 1, \cdots, p$

$$\varepsilon_t = \sigma_t \cdot \nu_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + \cdots + \alpha_p\varepsilon_{t-p}^2 + \beta_1\sigma_{t-1}^2 + \beta_2\sigma_{t-2}^2 + \cdots + \beta_p\sigma_{t-q}^2$$

$$= \alpha_0 + \sum_{i=1}^{p} \alpha_i\varepsilon_{t-1}^2 + \sum_{j=1}^{q} \beta_j\sigma_{t-j}^2$$

GARCH

[표현: ε_t ~ GARCH(p,q)]

$\{\nu_t\}$: 가우시안 백색잡음
$\alpha_0 \geq 0,\ \alpha_j \geq 0,\ \beta_j \geq 0$
$j = 1, \cdots, p$

$$\varepsilon_t = \sigma_t \cdot \nu_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + \cdots + \alpha_p\varepsilon_{t-p}^2 +$$ ARCH

$$\beta_1\sigma_{t-1}^2 + \beta_2\sigma_{t-2}^2 + \cdots + \beta_p\sigma_{t-q}^2$$ 추가된 부분

ARCH 추가된 부분

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i\varepsilon_{t-1}^2 + \sum_{j=1}^{q} \beta_j\sigma_{t-j}^2$$

GARCH

[표현: $\varepsilon_t \sim$ GA... ...우시안 백색잡음

$\alpha_j \geq 0, \beta_j \geq 0$

$= 1, \cdots, p$

$\varepsilon_t = \sigma_t \cdot \nu_t$

$\sigma_t^2 = \alpha_0 + \alpha \ldots _{-p} +$ ARCH

$\beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \cdots + \beta_p \sigma_{t-q}^2$ 추가된 부분

$= \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$

GARCH 모형은 왜 ARCH모형보다 추정해야 할 모수가 적을까?

GARCH(1,1)의 예시로 알아보자!

GARCH

[ε_t ~ GARCH(1,1)]

$\{\nu_t\}$~ i.i.d.N(0,1)
$\alpha_0 \geq 0$, $\alpha_j \geq 0$, $\beta_j \geq 0$
j = 1, ⋯, p

$$\varepsilon_t = \sigma_t \cdot \nu_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

$$(1-\beta L)\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

$$\sigma_t^2 = \frac{\alpha_0}{(1-\beta L)} + \frac{\alpha}{(1-\beta L)} \varepsilon_{t-1}^2$$

$$= \frac{\alpha_0}{(1-\beta L)} + \alpha \sum_{j=1}^{\infty} \beta^{j-1} \varepsilon_{t-j}^2$$

GARCH

[ε_t ~ GARCH(1,1)]

$\{\nu_t\}$~ i.i.d.N(0,1)
$\alpha_0 \geq 0,\ \alpha_j \geq 0,\ \beta_j \geq 0$
$j = 1, \cdots, p$

$$\varepsilon_t = \sigma_t \cdot \nu_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

$$(1-\beta L)\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

$$\sigma_t^2 = \frac{\alpha_0}{(1-\beta L)} + \frac{\alpha}{(1-\beta L)} \varepsilon_{t-1}^2$$

$$\sigma_t^2 = \frac{\alpha_0}{(1-\beta L)} + \alpha \sum_{j=1}^{\infty} \beta^{j-1} \varepsilon_{t-j}^2 \rightarrow \text{ARCH}(\infty)$$

GARCH

[ε_t ~ GARCH(1,1)]

$\{\nu_t\}$: 가우시안 백색잡음

$\alpha_0 \geq 0,\ \alpha_j \geq 0,\ \beta_j \geq 0$, …, p

$\varepsilon_t = \sigma_t \cdot \nu_t$

$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$

$(1-\beta L)\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$

$$\sigma_t^2 = \frac{\alpha_0}{(1-\beta L)} + \frac{\alpha}{(1-\beta L)}\varepsilon_{t-1}^2$$

$$= \frac{\alpha_0}{(1-\beta L)} + \alpha \sum_{j=1}^{\infty} \beta^{j-1}\varepsilon_{t-j}^2$$

ARCH에서 많은 모수들을 추정해야 했던 문제를 해결!

$$\mathrm{GARCH}(1,1) = \mathrm{ARCH}(\infty)$$

THANK YOU