

시계열 분석 및 실습

데이터 1

12171676 컴퓨터공학과 이종범

주제 :

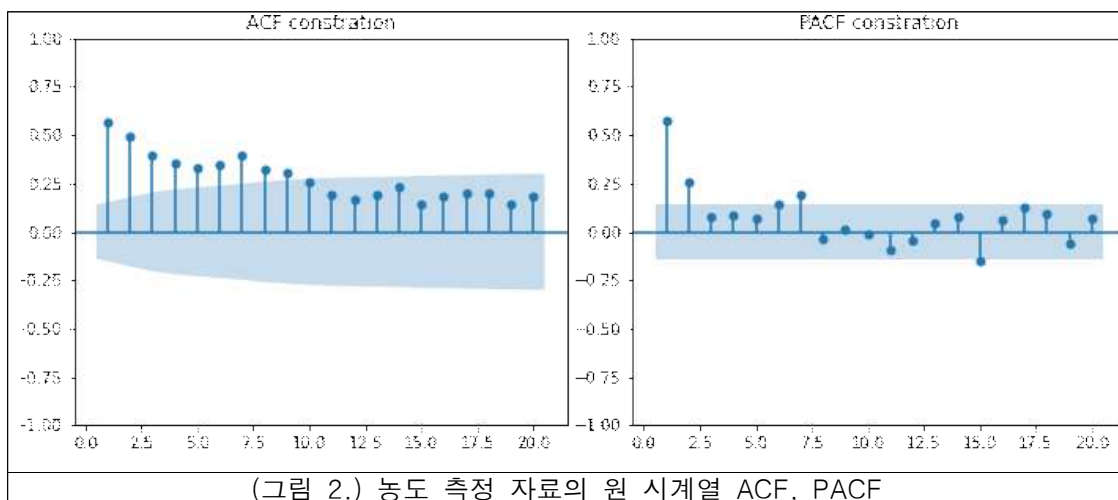
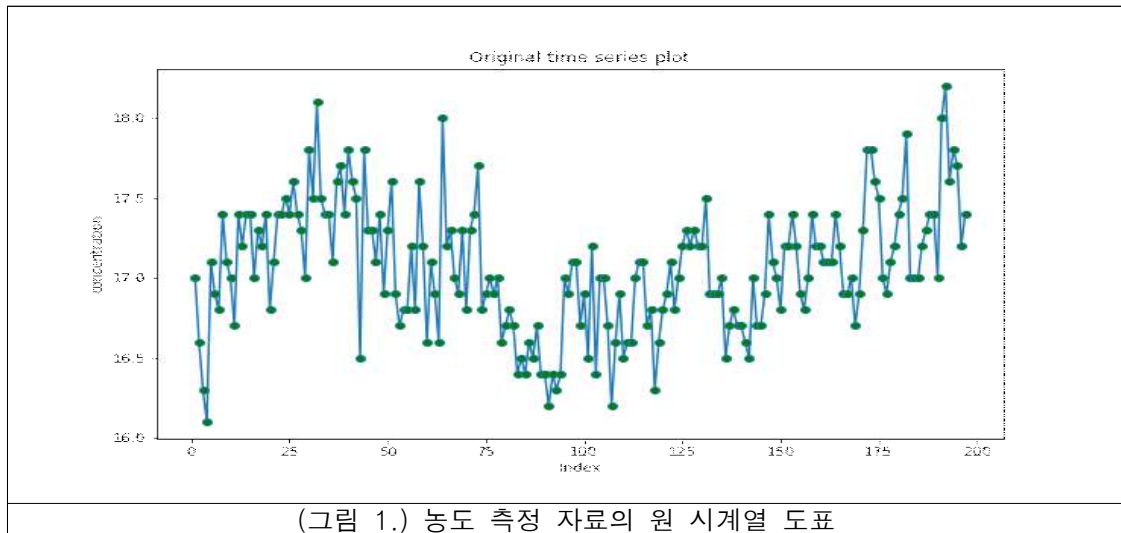
Chemical process concentration readings:
Every two hours ($n = 197$)

최종 분석결과 :

- 모형 : IMA(1,1)
- 파라미터 : $\theta_1 = -0.6994$, *no constant*

1. Model identification

- 원 시계열 자료를 활용한 해석



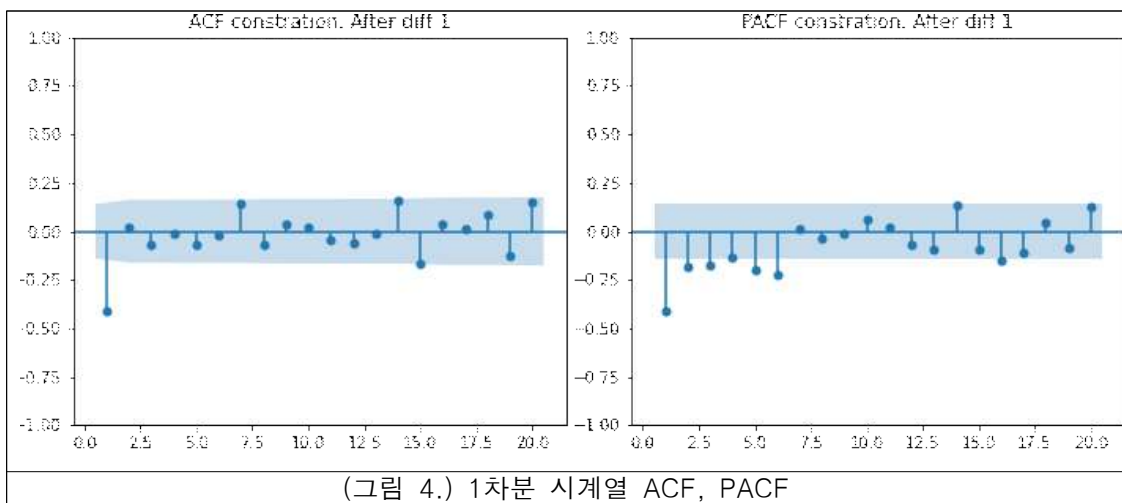
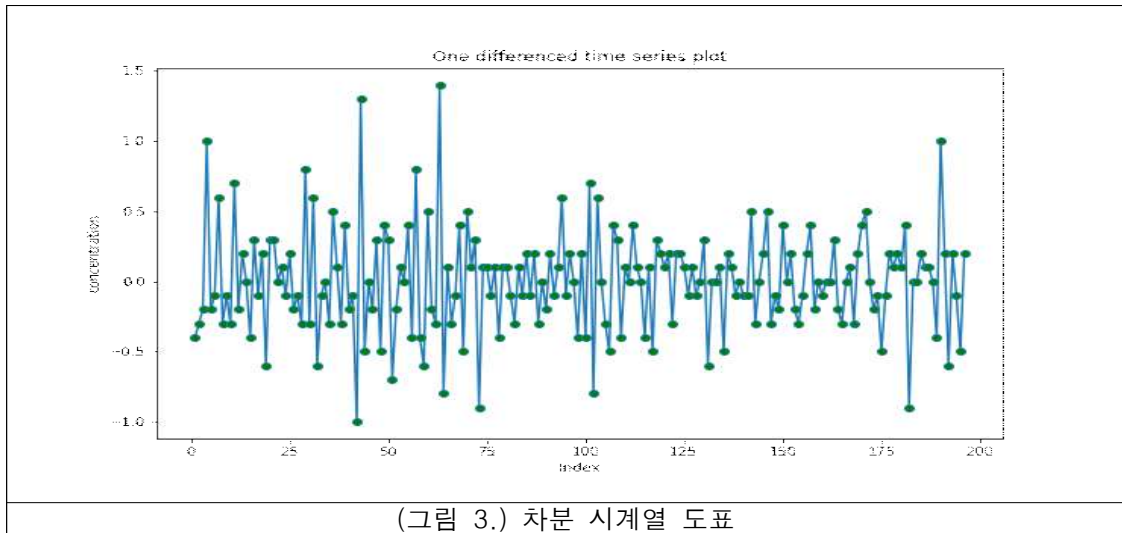
해석 :

- (1) 그림 1.을 통해, 시간에 따라 분산이 달라져 비정상 시계열로 추정
- (2) 그림 2.의 ACF가 매우 천천히 감소하는 특징을 보인다.비정상 시계열 추정
- (3) PACF, ACF 모두 7 주기마다 상승하는 특징을 보인다.

s=7 Seasonality 추정

-> 차분을 수행하여, 시계열 정상화 이후에 모형 제시

- 차분 시계열 자료를 활용한 해석



해석 :

- (1) 그림 3.을 통해, 충분히 등분산으로 생각할 수 있다.
- (2) 그림 4. ACF에서 Cut off after lag1 특성이 보인다.
-> IMA(1,1)의 특성
- (3) 그림 4. PACF에서 lag 1 이후 sinusoidal decay의 특성을 보인다.
-> IMA(1,1)의 특성
- (4) 그림 4. ACF, PACF 모두 주기 7마다 상승하는 것이 관측된다.
-> s=7의 계절성

-> IMA(1,1), ARIMA(0,1,1)(0,0,1) s=7 모형 시도

2. Parameter estimation diagnostics

- IMA(1,1)

파라미터 추정

	coeff	P_value
θ_1	-0.6994	0.000...
MSE (residuals)	0.1007	0.000...

(표 1.) IMA(1,1) 모형의 파라미터

Integration이 존재하여, 상수항 제거 후 추정.(소프트웨어 권장사항)

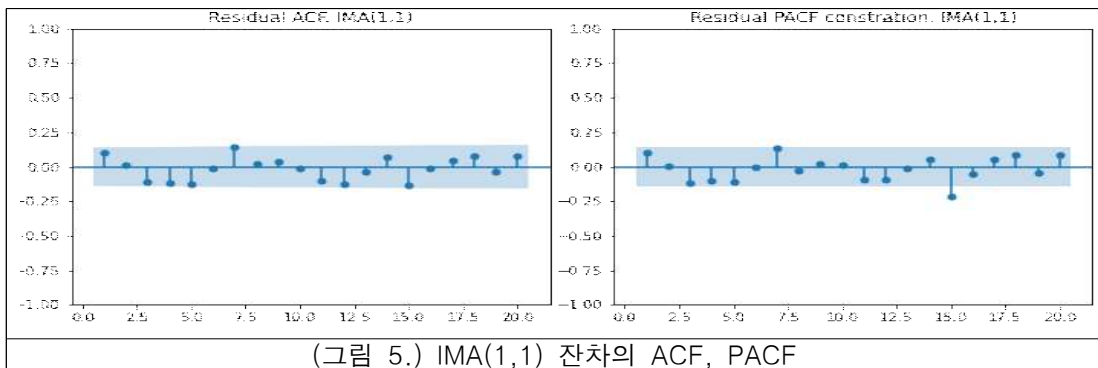
표 1.에서 IMA(1,1) 파라미터는 $\theta_1 = -0.6994$, no constant임을 확인.

잔차 진단

Lag	12	24	36	48
Chi-Square	19.78	30.33	49.92	53.64
P-value	0.071	0.17	0.06	0.266

(표 2.) IMA(1,1) 모형의 Box-Pierce Chi-square statistics

표 2.에서 모형의 잔차가 귀무가설을 채택하므로, 잔차는 랜덤이라 할 수 있다.



(그림 5.) IMA(1,1) 잔차의 ACF, PACF

그림 5.를 통해 잔차가 random임을 확인할 수 있다 하지만, 여전히 7주기마다 상승하는 계절 패턴이 보인다.

- ARIMA(0,1,1)(0,0,1) s=12

파라미터 추정

	coeff	P_value
θ_1	-0.6994	0.000...
θ_{s1}	0.1455	0.085
MSE (residuals)	0.0984	0.000...

(표 3.) ARIMA(0,1,1)(0,0,1) s=12 모형의 파라미터

표 3.에서 MSE는 적게 나왔지만, 유의수준 5%에서 θ_{s1} 이 기각할 수 없음을 볼 수 있다.

잔차 진단

Lag	12	24	36	48
Chi-Square	13	24.56	40.2	42.9
P-value	0.37	0.42	0.28	0.68

(표 4.) ARIMA(0,1,1)(0,0,1) s=12 모형의 Box-Pierce Chi-square statistics

표 4.에서 모형의 잔차가 귀무가설을 채택하므로, 잔차는 랜덤이라 할 수 있다.

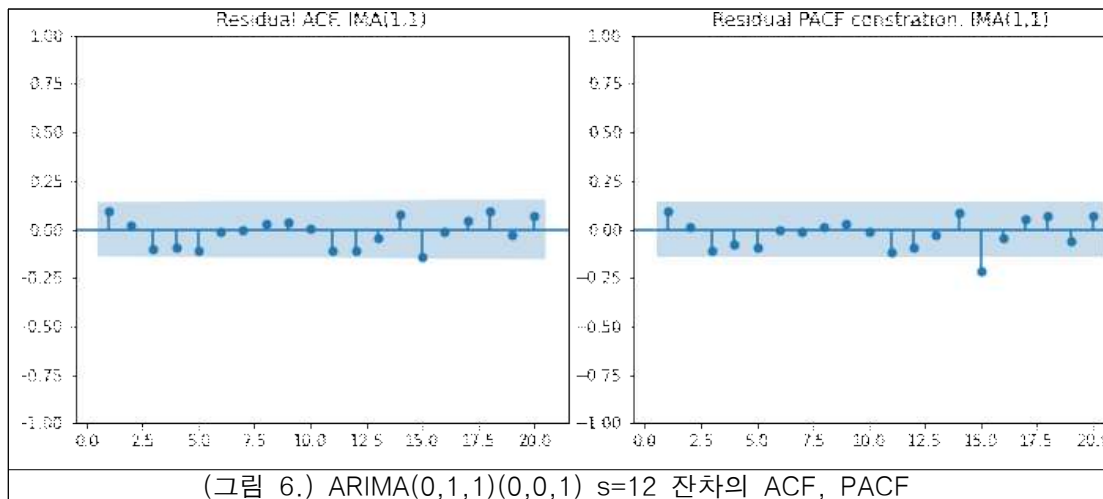


그림 6.에서 이전에 관측된 7주기의 계절성이 삭제된 것을 볼 수 있다.

- Final Model

Model	MSE	# of parameters
IMA(1,1)	0.1007	1
ARIMA(0,1,1)(0,0,1) s=12	0.0984	2

(표 5.) 최종 모형 비교

표 5.에서 IMA 모형이 MSE가 조금 더 높지만, 유의한 차이는 아니다. 따라서, 모수 절약의 법칙에 따라 IMA(1,1)을 선택한다.

데이터 2

12171676 컴퓨터공학과 이종범

주제 :

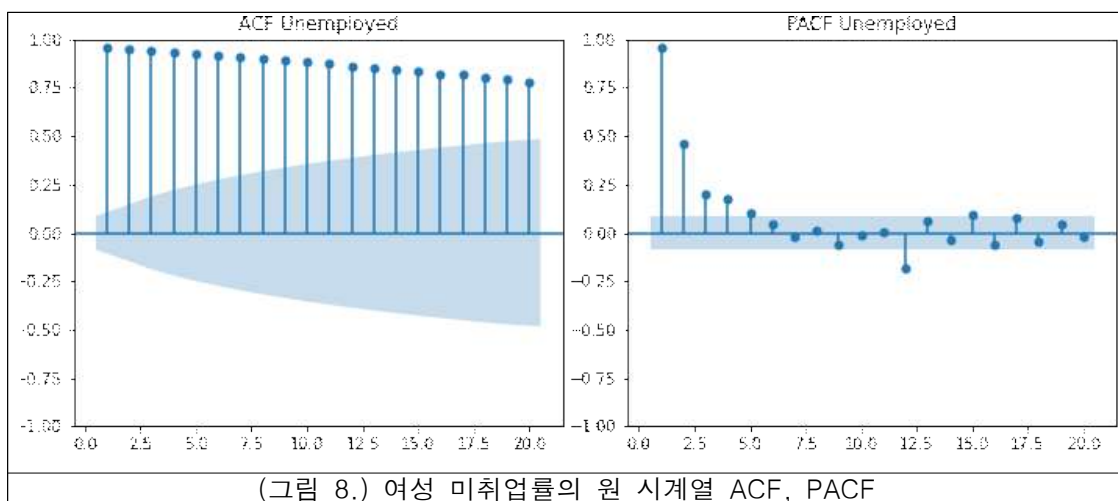
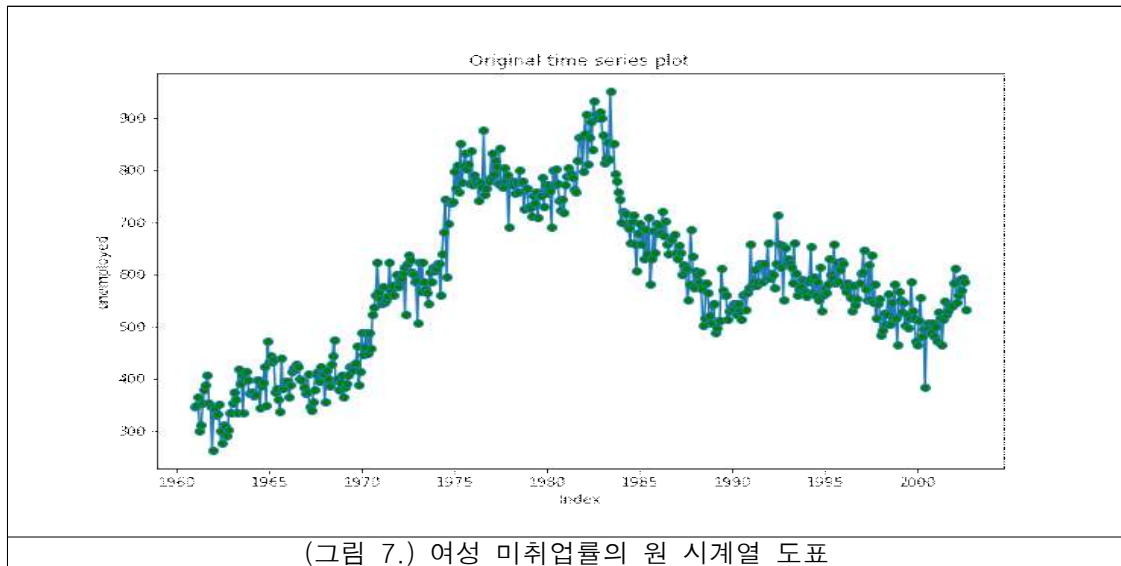
Monthly series of unemployed females between ages 16 and 19 in US from January 1961 to October 2002. (n = 500)

최종 분석결과 :

- 모형 : $ARIMA(0,1,1)(0,0,1)_s=12$
- 파라미터 : $\theta_1 = -0.5744, \theta_{s1} = -0.2315$, *no constant*

3. Model identification

- 원 시계열 자료를 활용한 해석



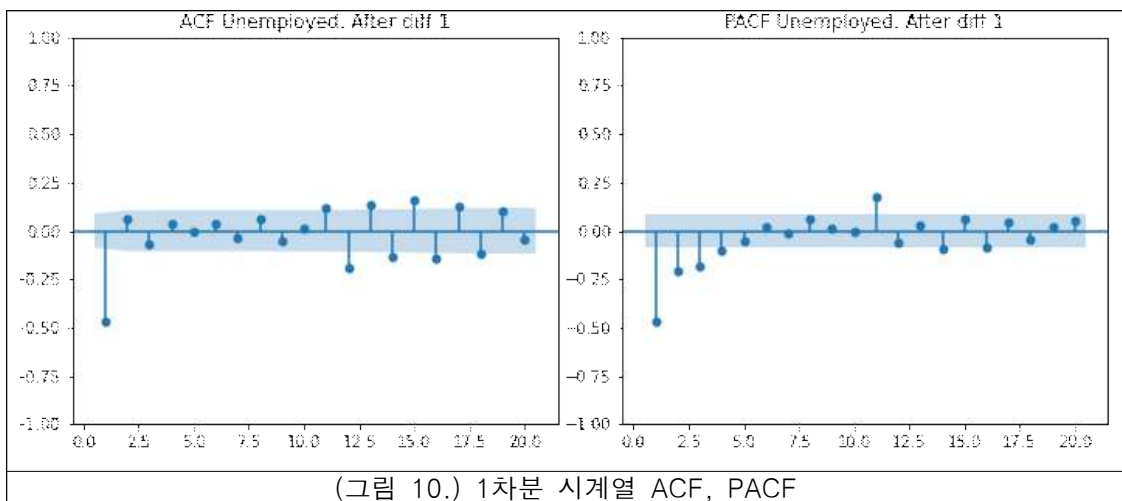
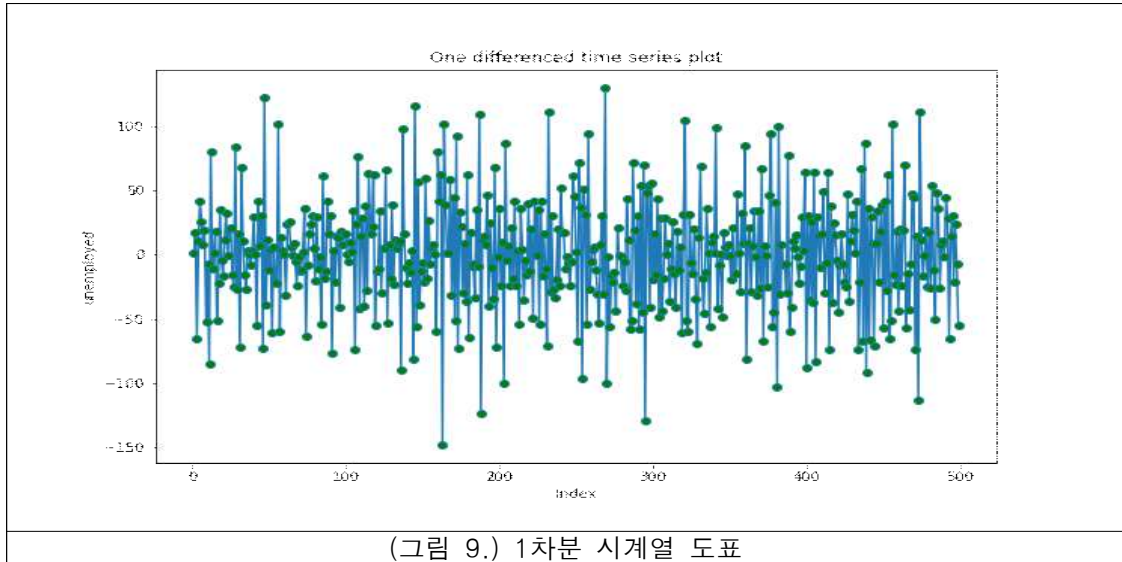
해석 :

- (1) 그림 7.을 통해, 시간에 따라 분산이 달라져 비정상 시계열로 추정
- (2) 그림 8.의 ACF가 매우 천천히 감소하는 특징을 보인다. 비정상 시계열 추정
- (3) 그림 8.의 PACF Lag5 이후 sinusoidal decay 관측
- (4) 그림 8.의 PACF에서 S=12에서 상승하는 특징을 보인다.

s=12 Seasonality 추정

-> 차분을 수행하여, 시계열 정상화 이후에 모형 제시

– 차분 시계열 자료를 활용한 해석



해석 :

(1) 그림 9.을 통해, 충분히 등분산으로 생각할 수 있고, 시간에 따른 평균, 분산 차이가 없어 보인다.

(2) 그림 10. ACF에서 Cut off after lag1 특성이 보인다.

-> IMA(1,1)의 특성

(3) 그림 10. PACF에서 lag 1 이후 sinusoidal decay의 특성을 보인다.

-> IMA(1,1)의 특성

(4) 그림 10. ACF, PACF 모두 주기 12마다 상승하는 것이 관측된다.

-> s=12의 계절성

-> IMA(1,1), ARIMA(0,1,1)(0,0,1) s=12 모형 시도

4. Parameter estimation diagnostics

- IMA(1,1)

파라미터 추정

	coeff	P_value
θ_1	-0.6007	0.000...
MSE (residuals)	1324.66	0.000...

(표 6.) IMA(1,1) 모형의 파라미터

Integration이 존재하여, 상수항 제거 후 추정.(소프트웨어 권장사항)

표 6.에서 IMA(1,1) 파라미터는 $\theta_1 = -0.6007$, no constant임을 확인.

잔차 진단

Lag	12	24	36	48
Chi-Square	22.29	52.99	65.45	80.71
P-value	0.0344	0.0006	0.002	0.002

(표 7.) IMA(1,1) 모형의 Box-Pierce Chi-square statistics

표 7.에서 모형의 잔차가 귀무가설을 기각하므로, 잔차는 랜덤이라 할 수 없다.

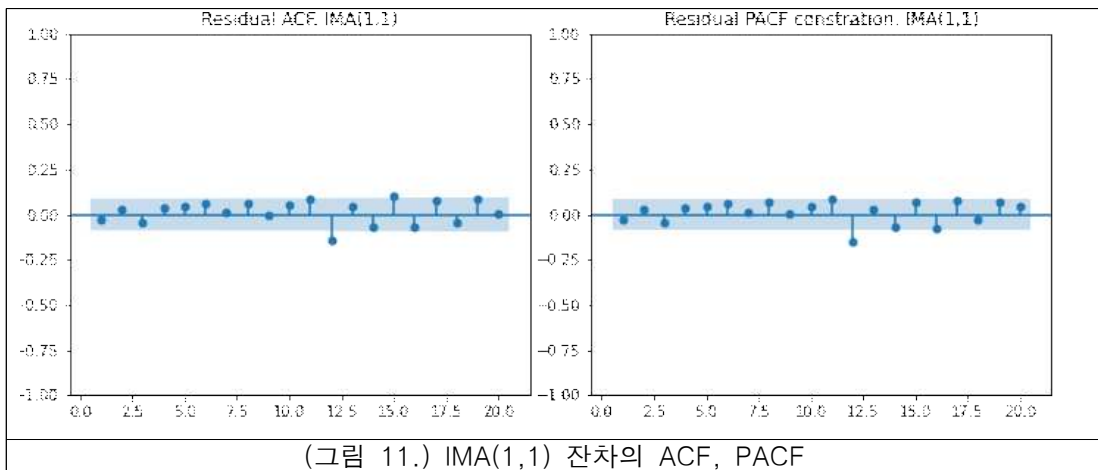


그림 11.를 통해 검정 결과와 달리, 잔차가 random인 것처럼 추정된다. 하지만, 여전히 12주기마다 상승하는 계절 패턴이 보인다.

- ARIMA(0,1,1)(0,0,1) s=12

파라미터 추정

	coeff	P_value
θ_1	-0.5774	0.000
θ_{s1}	-0.2315	0.000
MSE (residuals)	1280.2005	0.000..

(표 8.) ARIMA(0,1,1)(0,0,1) s=12 모형의 파라미터

표 8.에서 IMA보다 MSE도 적게 나왔으며, 검정 결과 모두 기각함을 볼 수 있다.

잔차 진단

Lag	12	24	36	48
Chi-Square	13.4	34.48	47.2	60.8
P-value	0.3407	0.076	0.1	0.1

(표 9.) ARIMA(0,1,1)(0,0,1) s=12 모형의 Box-Pierce Chi-square statistics

표 9.에서 모형의 잔차가 귀무가설을 채택하므로, 잔차는 랜덤이라 할 수 있다.

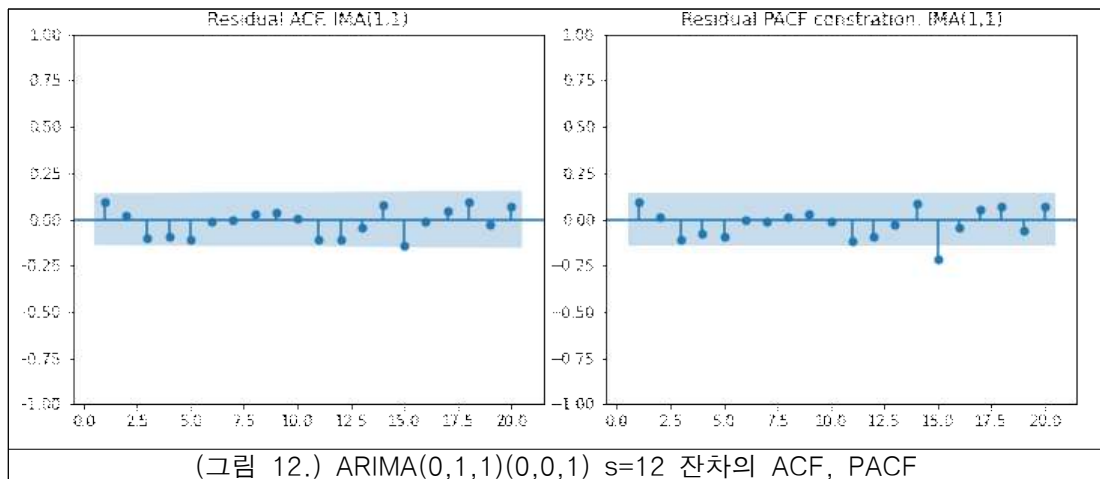


그림 12.에서 이전에 관측된 12주기의 계절성이 삭제된 것을 볼 수 있다.

- Final Model

Model	MSE	# of parameters
IMA(1,1)	1324.6617	1
ARIMA(0,1,1)(0,0,1) s=12	1280.2005	2

(표 10.) 최종 모형 비교

표 10.에서 계절 ARIMA 모형이 MSE가 더 낮으며, 잔차도 랜덤을 만족하므로 더 적합한 모형이라고 할 수 있다.