# 2020학년도 (1)학기 과제물(온라인제출용)

교과목명 : 예측방법론 학 번 : 201935-950005

성 명: 김동영

연 락 처 : 010-2600-0622

o 과제유형 :(공통) 형

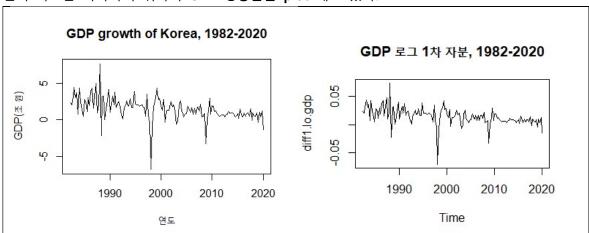
o 과 제 명 : 기말(과제/주관식)

- 1. GDP 계절조정계열(분기, 실질)의 전기대비 성장률을 바탕으로 다음 물음에 답하시오.(30점)
  - (1) ARIMA모형으로 추정하시오.
  - (2) 추정된 ARIMA모형의 잔차가 임의적인지 점검하시오.
  - (3) 향후 4분기(1년)을 예측하시오.

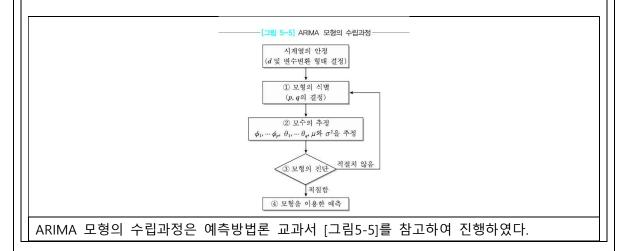
### 문제 1번

GDP 계절조정계열(분기, 실질)을 KOSIS에서 2020년 1분기까지 다운로드 하였다. (교수님이 카페에 올리신 자료와 최근 2019년 근방의 숫자가 약간 다른점이 있으나 2020까지 최근 업데이트 된 KOSIS를 이용하였다. 추후 교수님자료와 혹자가 다운받은 자료를 auto.arima 추정하였을때 다른값이 나온점은 흥미로웠다.)

먼저 자료를 파악하기 위하여 GDP 성장률을 plot 해 보았다.



성장률 그래프와 비교해 보기 위해 GDP 로그1 차 차분도 함께 그려보았다. 계절조정계열을 가공한것이기 때문에 두 그래프 모두 계절요인이 없고 추세가 없어진 것을 알 수 있다. 한국의 GDP성장률이 지속적으로 감소한다는 사실을 인지하고 보니 그래프가 조금씩 감소하는 약간의 추세는 있음을 알 수 있다. 차이점은 성장률이 Y축 값이 더 크게 부각되는 것을 볼수 있다(왼쪽은 범위가 -6~6정도 인 반면, 오른쪽은 -1~1정도 인것으로 보임).



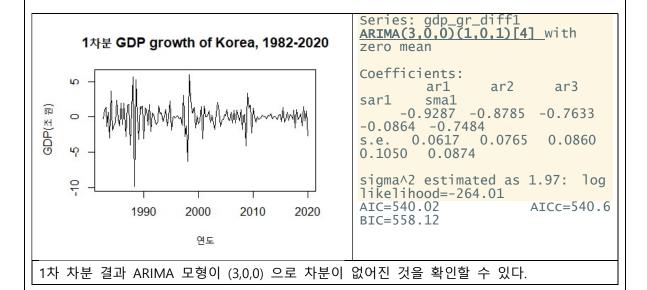
GDP성장률의 대체적인 ARIMA 모형 추정 위하여 auto.arima function을 이용한 결과는 아래와 같다.

Series: gdp\_gr ARIMA(3,1,0)(1,0,1)[4] Coefficients: ar1 ar2 ar3 sar1 sma1 -0.9287 -0.8785 -0.7633 -0.0864 -0.7484 s.e. 0.0617 0.0765 0.0860 0.1050 0.0874 sigma^2 estimated as 1.97: log likelihood=-264.01 AIC=540.02 AICc=540.6 BIC=558.12

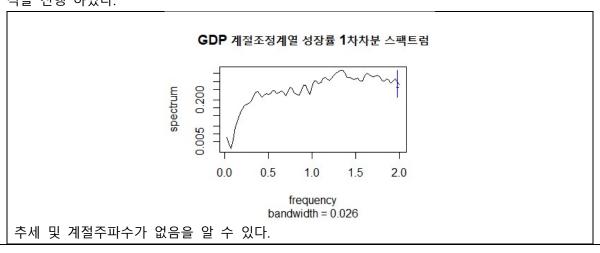
ARIMA(3,1,0)(1,0,1)[4] 로 1차 차분이 필요해 보인다. 이유는 성장률의 감소일것으로 추정된다. 추가로 계절변동요인은 이미 계절조정계열의 성장률 이기 때문에 무시하기로 하였다.

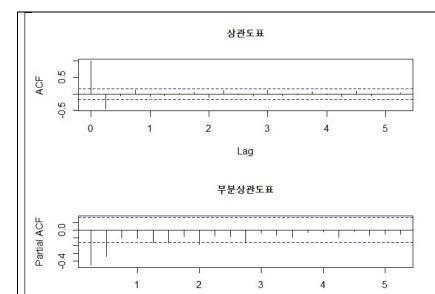
시계열의 안정 (d 및 변수변환 형태 결정)

시계열의 안정을 위하여 1차 차분을 하여 다시 auto.arima를 해 보았다.



성장률을 1차 차분한 모형을 기준으로 스펙트럼, 상관도표, 부분상관도표, adf test, box test 분석을 진행 하였다.





상관도표, 부분상관도표 모두 2차이후 작은값을 보인다. 상관도표와 부분상관도표 만으로는 ARMA 모형추정의 한계가 있기 때문에 참고만 하였다.

Augmented Dickey-Fuller Test

data: gdp\_gr\_diff1
Dickey-Fuller = -8.0104, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary

Lag

p-value = 0.01로 단위근이 없음을 확인하였다. 추세가 없는 상태에서 당연한 결과라고 할 수 있다.

Box-Ljung test

data: gdp\_gr\_diff1

X-squared = 38.689, df = 8, p-value = 5.614e-06

륭-박스 검정결과를 통하여 p-value가 작으므로 8차까지 자기상관계수가 0이라는 귀무가설을 기각한다. GDP성장률의 1차차분이 백색잡음은 아님을 알수있다.

① 모형의 식별 (p, q의 결정)

현재까지 분석결과로는 모형의 식별단계에서 (3,1,0)으로 가정해 본다.

모수의 추정을 정확하게 하기 위하여 여러가지 모형을 아래와 같이 가정해 보았다.

번호	ARIMA 모형	근거
모형1	ARIMA(3,1,0)	위의 여러가지 모형식별결과
모형2	ARIMA(3,1,0)(1,0,1)	최초 auto.arima 결과를 그대로 이용해보기로 하였다.
모형3	ARIMA(3,1,1)	과대적합진단을 위한 모델
모형4	ARIMA(3,0,0)	GDP 성장률에 추세가 없다고 가정한 모델



모형진단 결과는 아래와 같다.

ㅁ칭	74 71
모형	결과
ARIMA(3,1,0)	Call: $arima(x = gdp\_gr, order = c(3, 1, 0))$
	Coefficients:
	sigma^2 estimated as 2.264: log likelihood = -276.21, aic = 560.42
	Training set error measures:  ME RMSE MAE MPE MAPE
	MASE Training set -0.03551289 1.499643 0.95942 -29.22999 133.4487 0.8112962 ACF1 Training set -0.01516938
ARIMA(3,1,0)(1,0,1)	Call: arima(x = gdp_gr, order = c(3, 1, 0), seasonal = list(order = c(1, 0, 1)))
	Coefficients:     ar1 ar2 ar3 sar1 sma1     -0.9287 -0.8785 -0.7633 -0.0864 -0.7484 s.e. 0.0617 0.0765 0.0860 0.1050 0.0874
	sigma $^2$ estimated as 1.904: log likelihood = -264.01, aic = 540.02
	Training set error measures:  ME RMSE MAE MPE MAPE MASE
	Training set -0.2088347 1.375415 0.8547859 -53.41842 135.4803 0.7228165 ACF1
	Training set 0.05764642
ARIMA(3,1,1)	$arima(x = gdp\_gr, order = c(3, 1, 1))$
	Coefficients:     ar1 ar2 ar3 ma1     0.1019 0.0203 0.0743 -0.9448 s.e. 0.0848 0.0842 0.0836 0.0230
	sigma $^2$ estimated as 1.918: log likelihood = -264.36, aic = $538.72$
	Training set error measures:  ME RMSE MAE MPE MAPE MASE
	Training set -0.1929803 1.380299 0.869904 -54.66371 139.0628 0.7356005  ACF1 Training set -0.01464549
A D IA 4A (2, 0, 0)	Call:
ARIMA(3,0,0)	$arima(x = gdp\_gr, order = c(3, 0, 0))$

```
Coefficients:
    ar1 ar2 ar3 intercept
    0.2049 0.1132 0.1700 1.4865
s.e. 0.0805 0.0817 0.0805 0.2213

sigma^2 estimated as 1.997: log likelihood = -268.35,
    aic = 546.7

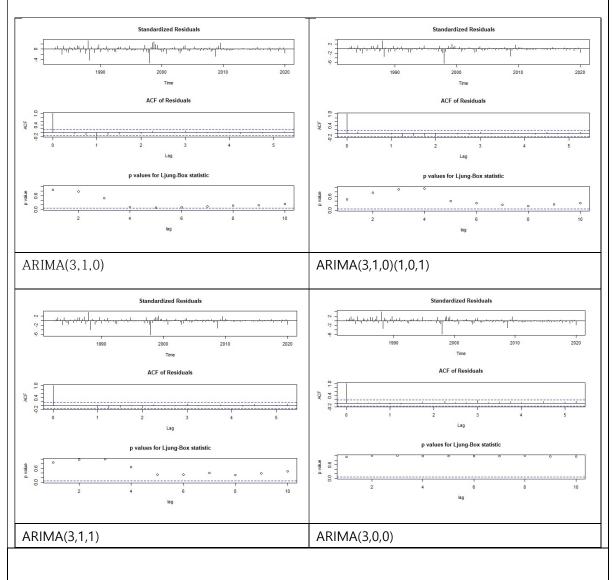
Training set error measures:
    ME RMSE MAE MPE MAPE

MASE
Training set -0.006977581 1.413146 0.9189023 -28.98198
120.6714 0.777034
    ACF1
Training set 0.005035202
```

AIC값을 비교한 결과

ARIMA(3,1,1) 모형이 가장 적절한 것으로 판단하였다.

## 문제2번 잔차분석



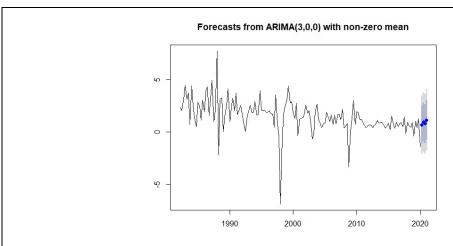
## Tsdiag function으로 잔차분석한 결과

- 표준화된 잔차는 네가지 모델 모두 특별한 패턴을 보이지 않았다.
- 잔차의 상관도표는 모두 자기상관계수가 점선 내에 있음을 알 수있다.
- 륭-박스 검정의 유의확률은 유의수준인 점선보다 (3,1,1), (3,0,0)에서 확연히 크게 나타 났다. 하지만 <u>ARIMA(3,0,0) 모델에서 특정패턴을 보이지 않아 가장 적절한 것으로 판단</u> 된다.

#### 문제3번

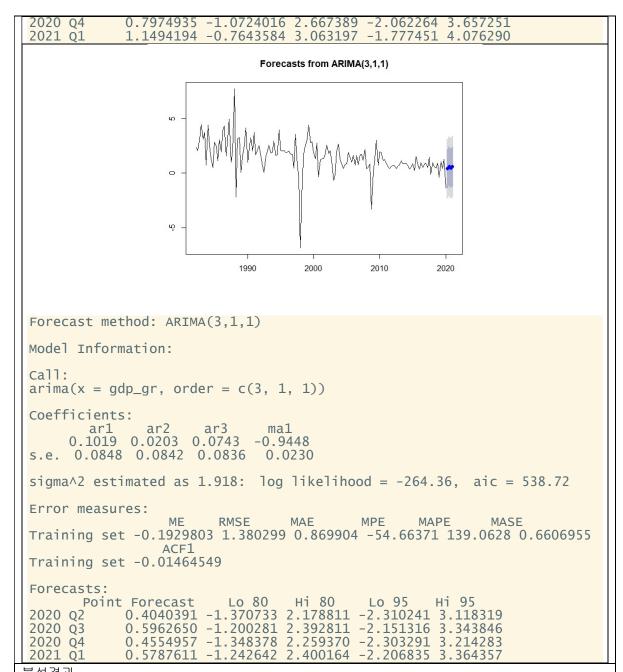
2020 Q3

잔차분석 결과를 통해 최종 모델을 ARIMA(3,0,0)으로 지정하고 향후 1년을 예측하였다.



```
Forecast method: ARIMA(3,0,0) with non-zero mean
Model Information:
arima(x = gdp_gr, order = c(3, 0, 0))
Coefficients:
ar1 ar2 ar3 intercept
0.2049 0.1132 0.1700 1.486
s.e. 0.0805 0.0817 0.0805 0.221
                                1.4865
                                 0.2213
sigma^2 estimated as 1.997: log likelihood = -268.35, log likelihood = -268.35
Error measures:
                     ME
                           RMSE
                                      MAE
                                                MPE
                                                        MAPE
Training set -0.006977581 1.413146 0.9189023 -28.98198 120.6714
0.6979099
                  ACF1
Training set 0.005035202
Forecasts:
      Point Forecast
                            Lo 80
                                     ні 80
                                               Lo 95
                                                         ні 95
             0.6867617 -1.1242574 2.497781 -2.082953 3.456476
2020 Q2
```

0.9569199 -0.8917412 2.805581 -1.870363 3.784203



#### 분석결과

#### ARIMA(3,0,0)

		Poin	t Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2020	Q2		<mark>0.6867617</mark>	-1.12426	2.497781	-2.08295	3.456476
2020	Q3		<mark>0.9569199</mark>	-0.89174	2.805581	-1.87036	3.784203
2020	Q4		<mark>0.7974935</mark>	-1.0724	2.667389	-2.06226	3.657251
2021	Q1		<mark>1.1494194</mark>	-0.76436	3.063197	-1.77745	4.07629

#### ARIMA(3,1,1)0|

2020	Q2	C	<mark>.4040391</mark>	-1.37073	2.178811	-2.31024	3.118319
2020	Q3		<mark>0.596265</mark>	-1.20028	2.392811	-2.15132	3.343846
2020	Q4	C	<mark>.4554957</mark>	-1.34838	2.25937	-2.30329	3.214283
2021	Q1	C	<mark>).5787611</mark>	-1.24264	2.400164	-2.20684	3.364357

## 로 첫번째 모델이 더 GDP를 긍정적으로 예측한다.

아직 KOSIS에 2020년 2분기 GDP자료가 업데이트되지 않아 어떤예측이 정확한지 알수는 없으나 추후 확인해보는 것도 개인적으로 수행 할 계획이다.

- 🖹 🧁 경제활동별, 지출항목별 규모

## 추가분석 사항

KDI의 경제전망 보고서 (2019년 하반기)를 참고하면

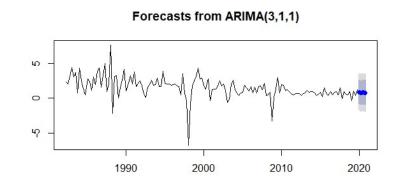
#### 2. 2020년 국내경제 전망

- 우리 경제는 2020년에 내수와 수출의 개선이 제한적인 수준에 머물면서 2019년 (2.0%)보다 소폭 높은 2.3% 내외의 성장률을 기록할 전망
  - 민간소비는 소비심리가 개선되겠으나, 국내총소득이 낮은 증가세를 보임에 따라 미약한 회복새를 보임 전망
  - 설비투자는 반도체 수요의 회복과 함께 기저효과의 영향도 더해지면서 양호한 증가 세를 나타낼 전망
  - 건설투자는 건축부문의 감소세를 토목부문이 일부 상쇄하면서 부진이 완화될 전망
  - ◎ 수출은 신흥국의 투자수요 확대가 상품수출의 증가로 이어지면서 점차 개선될 전망

전반적으로 긍정적인 전망은 아니지만 2019년보다는 개선된 성장률 2.3%를 보일 것으로 전망하였다. 아직 코로나바이러스로 인한 경제 영향을 예측하지 못한 상황이었기 때문에 이러한 예상이 나왔다고 판단된다. 그래서 GDP 자료에서 2020년 1분기 자료를 빼고 ARIMA모형을 추정하여 2020년 1Q를 예상하였을 때 어느정도 예상이 가능하였는지 분석해 보았다.

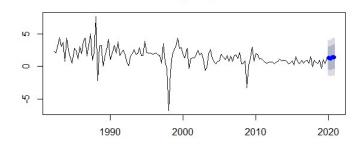
		Α	В	С
	1	년도	국내총생산(시 장가격, GDP)	전기대비성 장률
	148	2018 3/4	452561,1	0,45845045
	149	2019	456769,7	0,92995178
li	150	2019 1/4	455081	-0,3697049
	151	2019 1/2	459813,4	1,03990279
	152	2019 3/4	461705,9	0,41158
	153	2020	467494,9	1,25382847
	154	2020 1/4	460970,3	-1,3956516
	155			Company of the first

하이라이트 된 2020년 1Q를 제외하고 ARIMA모형을 다시 분석하였다.



		Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2020	Q1	0.8351516	<mark>-0.93014</mark>	2.600439	<mark>-1.86462</mark>	3.534926
2020	Q2	0.7563426	-1.03176	2.544449	-1.97833	3.491015
2020	Q3	0.8060605	-0.98898	2.601101	-1.93922	3.551338
2020	Q4	0.7772056	-1.03584	2.590247	-1.9956	3.550014

#### Forecasts from ARIMA(3,0,0) with non-zero mean



		Point Forecast	Lo	08 c	Hi 80	Lo 95		Hi 95	
2020	Q1	1.267991		<mark>-0.52792</mark>	3.063907		<mark>-1.47863</mark>	4.014608	
2020	Q2	1.25333		-0.58014	3.086798		-1.55072	4.057377	
2020	Q3	1.39376		-0.45865	3.246169		-1.43926	4.226775	
2020	Q4	1.423476		-0.47073	3.317683		-1.47346	4.320416	

마지막으로 선택된 ARIMA(3,0,0)의 예측결과를 보면 LO95 범위에 -1.47863으로 실제값인

-1.395보다 더 작으므로 point forecast 값인 1.267 예측치보다는 낮지만 예측범위를 완전 벗어 난다고 보긴 힘들다. 이를 통해 코로나사태가 한국경제에 ¹Black swan급의 영향을 주었다고 보 기는 힘들다고 판단된다. 또한, 추후 지속적으로 대한민국 경제에 영향을 줄 수도 있기 때문에 계속해서 상황을 주시해야 할 것이다.

<sup>1</sup> 도저히 일어날 것 같지 않은 일이 일어나는 것을 얘기하는 것으로, 월가 투자전문가인 나심 니콜라스 탈레브가 그의 저서 '검은 백조(The black swan)'를 통해 서브프라임 모기지 사태를 예언하면서 두루 쓰이게 됐다

```
코드:
library(tseries)
library(forecast)
setwd("D:/00. 방통대 통계/예측방법론/기말고사")
gdp <- read.csv("GDP_1982-2020.csv", header = TRUE)
gdp <- ts(gdp[,2]/1000, start = 1982, frequency = 4)
gdp_gr \leftarrow ts((gdp-lag(gdp,-1))/lag(gdp,-1)*100,
            start=c(1982,2), frequency = 4)
auto.arima(gdp_gr)
gdp_gr_diff1 <- diff(gdp_gr, 1)</pre>
auto.arima(gdp_gr_diff1)
###### gdp growth
plot(gdp_gr, ylab="GDP(조 원)", xlab="연도", col="black", main="GDP growth of Korea, 1982-
2020")
plot(gdp_gr_diff1, ylab="GDP(조 원)", xlab="연도", col="black", main="1차분 GDP growth of
Korea, 1982-2020")
##### gdp growth traits
spectrum(gdp_gr_diff1, spans=c(3,3), main="GDP 계절조정계열 성장률 1차차분 스팩트럼") #
계절인자 없음
par(mfrow=c(2,1))
```

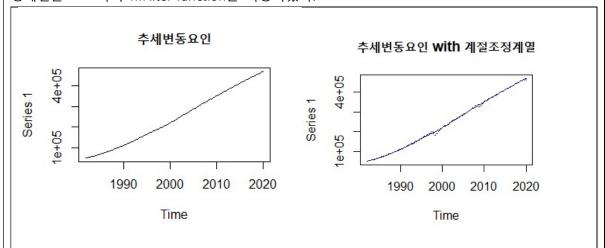
```
acf(gdp_gr, main="상관도표") #arima 판단?
pacf(gdp_gr, main="부분상관도표") #arima 판단?
acf(gdp_gr_diff1, main="상관도표") #arima 판단?
pacf(gdp_gr_diff1, main="부분상관도표") #arima 판단?
adf.test(gdp_gr) # 단위근 없음
adf.test(gdp_gr_diff1) # 단위근 없음
Box.test(gdp_gr, lag=8, type="Ljung") #8차까지 자기상관계수가 0이라는 귀무가설
Box.test(gdp_gr_diff1, lag=8, type="Ljung") #8차까지 자기상관계수가 0이라는 귀무가설
##### arima 모형추정
gdp_gr_fit1 = arima(gdp_gr, order=c(3,1,0))
gdp_gr_fit2 = arima(gdp_gr, order=c(3,1,0), seasonal = list(order = c(1,0,1)))
gdp_gr_fit3 = arima(gdp_gr, order=c(3,1,1))
gdp_gr_fit4 = arima(gdp_gr, order=c(3,0,0))
summary(gdp_gr_fit1)
summary(gdp_gr_fit2)
summary(gdp_gr_fit3)
summary(gdp_gr_fit4)
##### arima 모형의 진단
tsdiag(gdp_gr_fit1) ### 잔차분석
tsdiag(gdp_gr_fit2) ### 잔차분석
tsdiag(gdp_gr_fit3) ### 잔차분석
tsdiag(gdp_gr_fit4) ### 잔차분석
```

```
#### 향후 4분기 예측
plot(forecast(gdp_gr_fit1, h=4))
plot(forecast(gdp_gr_fit2, h=4))
plot(forecast(gdp_gr_fit3, h=4))
plot(forecast(gdp_gr_fit4, h=4))
summary(forecast(gdp_gr_fit4, h=4))
summary(forecast(gdp_gr_fit4, h=4))
```

- 2. GDP 계절조정계열(분기, 실질)을 바탕으로 다음 물음에 답하시오.(20점)
- (1) GDP의 추세변동을 HP필터를 이용하여 구하시오.
- (2) GDP의 순환변동치를 구하여 그래프로 그리고, 그 의미를 정리하시오.

#### 문제1번

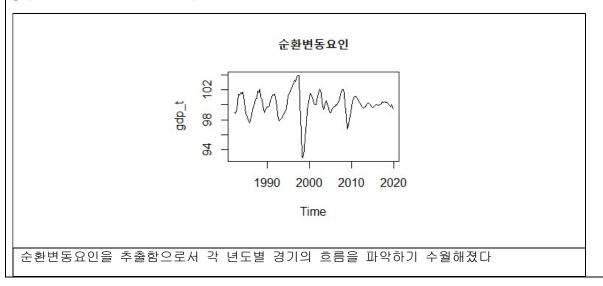
추세변동요인 추출법 중에 호드릭-프레스콧 필터를 사용하여 추세변동을 구하였다. GDP 계절조 정계열을 로그하여 mFilter function을 사용하였다.



오른쪽 그래프를 통해 어떻게 HP 필터가 계절조정계열중 추세변동요인을 나타내는지 눈으로 확인할 수 있다.

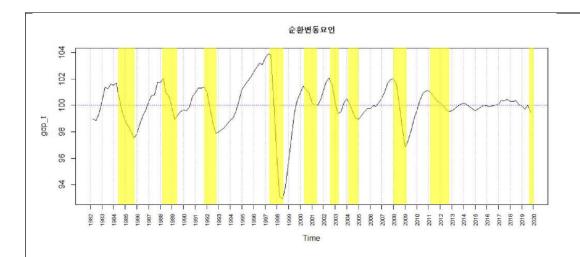
#### 문제2번

로그변환된 계절조정 GDP를 3기 중심화 평균하여 추세,순환변동계열을 생성하여 변동요인 gdpsam으로 지정하였다.<sup>2</sup> Gdpsam을 추세변동계열로 나누어 추세변동계열을 구한다.



.

<sup>&</sup>lt;sup>2</sup> 예측방법론 교과서 184.p 참고



순환변동요인 그래프의 X축에 각 년도를 입력하고 점선을 그어 교과서 180p에 나오는 경기동행종합지수와 같이 만들어 보았다. 2013년 이후 100주변에서 계속 머물고 있다가 최근 코로나바이러스 영향으로 하강국면에 들어선것으로 보여진다. 유가하락, 미중무역갈등 등 많은 변수들이 있어 과거의 자료들을 가지고 도출한 ARIMA모형만으로 예측을 한다면 한계가 있을 것으로판단하고 있다.

```
코드:
library(mFilter)
setwd("D:/00. 방통대 통계/예측방법론/기말고사")
gdp <- read.csv("GDP_1982-2020.csv", header = TRUE)
gdp <- ts(gdp[,2], start = 1982, frequency = 4)
#gdp_gr2 <- ts(gdp_gr[,3], start = c(1982,2), frequency = 4)
#gdp_gr <- ts((gdp-lag(gdp,-1))/lag(gdp,-1)*100,start=c(1982,2), frequency = 4)
log.gdp <- log(gdp)
log.gdp.hp = mFilter(log.gdp, filter="HP")
gdp_t = exp(log.gdp.hp$trend)
gdpsam = exp((log.gdp+lag(log.gdp, -1)+lag(log.gdp, 1))/3)
gdp_c = gdpsam/gdp_t*100
```

 $gdp_c2 = gdp/gdp_t*100$ 

```
gdp_i = gdp/gdpsam*100
plot(gdp_t, main="추세변동요인")
plot(gdp_t, main="추세변동요인 with 계절조정계열")
lines(gdp, lty=3, col = "blue")
plot(gdp_i, main="불규칙 변동요인")
plot(gdp_c, main="순환변동요인")

plot(gdp_c, main="순환변동요인")

axis(1, seq(1982,2020,1), las = 2, cex.axis=0.7)
abline(v=seq(1982,2020,1), lty=3, col="gray")
abline(h=100,lty=3, col="blue")
```

- 3. 국내외 예측기관의 우리나라 금년 경제예측의 사례를 2개 이상 비교 정리하고 의견을 정리하시오.(20점)
  - 2019년 하반기에 발행된 2020년 경제예측보고서와 2020년 상반기에 발행된 2020년 경제예측보고서를 국내기관: KDI 한국개발연구원 기준으로 비교해 보았다.
  - 2019 하반기의 2020년 예측보고서와 2020년 상반기의 2020년 예측보고서를 비교하는 이유는 유가하락 및 코로나사태를 통해 어떻게 예측이 바뀌었는지 비교하기 위함이다.

KDI 경제전밍	-														
2019년 하반기 경제전망 보고서							2020년 상변	반기 경제	∥전망	보고서					
i l	2018 <sup>p</sup>		2019°		2020		Ī	2018 <sup>p</sup>	2019 <sup>p</sup>	2020		2021			
	연간	상반기º	하반기	연간	상반기	하반기	연간		연간	연간이	상반기	하반기	연간	연간	
국내총생산	2.7	1.9	2.0	2,0	2.4	2.2	2,3	국내총생산	2,7	2.0	-0.2	0.5	0.2	3.9	
(계절조정 전기대비)		(0.3)	(0.7)		(0.4)	(0.6)									
● 내수와 수출의 개선이 제한적인 수준으로							● 코	로나19으	영향	으로 민	l간소비	l와 수	출이 급	3	
머물면서 2020년 대한민국의 성장률을 폭으로 위축되어 2020년 성장률을 0.2										0.2%	<u>2</u>				

- 2.3%로 예측
- 실업률은 완만한 경제성장세와 정부일자리 정책의 영향으로 2020년 3.5%로 떨어질 것 으로 전망
- 경상수지는 수출입 모두 확대되면서 금년과 유사한 흑자폭 유지 전망
- 원유는 2020년 배럴당 60달러 내외를 예측

- 로 예측
- 실업률은 가파른 경기 위축에도 불구하고 경제활동참가율이 큰 폭으로 하락하면서 2020년 3.9%로 소폭 상승 전망
- ▶ 경상수지는 수출물량이 축소 되겠으나 교역 조건이 개선되면서 흑자폭 유지
- 원유는 45%정도 하락한 35달러 내외 2021 년에는 40달러 내외 전망

아직 1년중 절반밖에 지나가지 않았지만, 2020년에는 예측하기 힘든 사건들이 연속적으로 발생하며 경제에 도 많은 파급효과를 가져온 년도라도 할 수 있다. 먼저 미중 무역분쟁이 잠잠해 지는듯 싶었다가 미국, 사 우디, 러시아 등 산유국의 분쟁으로 연초 유가가 급락하였고 이로 인해 경기가 급격히 위축되는 등 수요 측과 공급 측 요인이 모두 작용하여 저물가 기조가 지속되었다.³ 추가로 코로나바이러스가 전세계를 강타 하였고 항공업계는 많은 업체가 도산할 정도로 직격탄을 맞았다. 필자도 해외에서 거주하며 한국 및 각국 을 빈번하게 오가던 생활에 익숙했던 것이 10년전 일처럼 지금은 한국뿐만 아니라 어느 나라도 마음 놓고 갈수 없는 세상이 되어버렸다. 예측보고서의 모든 예측이 경제에 부정적으로 숫자들이 변하였고 미래를 예 측하는 것이 의미가 있는지 의심이 갈 정도로 예측불가능한 사건들이 발생하면서 경제 예측이 어려워 진 다고 할 수 있다. 하지만 이러한 예측은 특정한 가정을 하고있고 이것이 예측 분석의 핵심이라고 생각한 다. IMF world economy outlook 보고서에도 가장 먼저 Assumptions and conventions 이라고 하는 각종 고 정요인을 설명하는 부분으로 시작하는 것을 보면 얼마나 그 가정이 예측 분석에 중요한지 알 수 있다. 이 러한 assumption이 없다면 그 어떤 것도 예측할 수 없을 것이며 학문이 발달하고 데이터가 발달할수록 이 러한 assumption을 더욱 견고하고 유연하게 하여 예측의 정확성을 높이는 것이 예측 분석의 키포인트라고 생각한다. 그러므로 코로나 사태와 같은 black swan급의 문제 발생하였을 때, 예측이 틀렸고 해당 보고서가 의미가 없다고 생각하기 보단, 어떤 예측이 어떻게 변하였는지를 파악하는 것이 먼저라고 생각한다. 기회 가 된다면 추후에는 순환변동요인이 있을 때마다 각종 예측보고서의 assumption 파트가 어떻게 변하는지 확인하는 과제를 수행한다면 경제를 이해하는데 더욱 도움이 될 것이라고 판단한다.

<sup>&</sup>lt;sup>3</sup> KDI 2020 상반기 경제전망 보고서 3부 2장 세계경제 동향