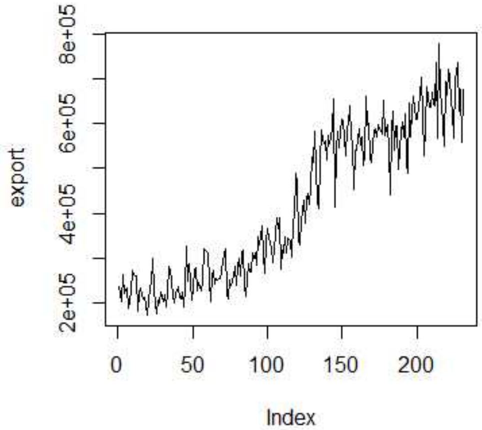
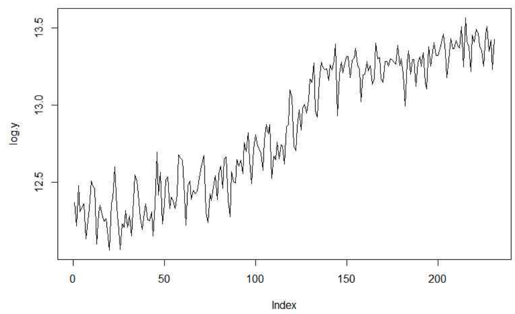
HW5 수출액 예측 및 VAR(1) 시뮬레이션

A. 첨부파일은 1998.1-2003.12까지의 우리나라 식료 및 직접소비재 월별 수출자료이다. 이 자료에 대해 ARMA 모델링을 하고 향후 1년치 까지를 예측하라.

export.data=read.csv(“export.csv”, head=T); attach(export.data);

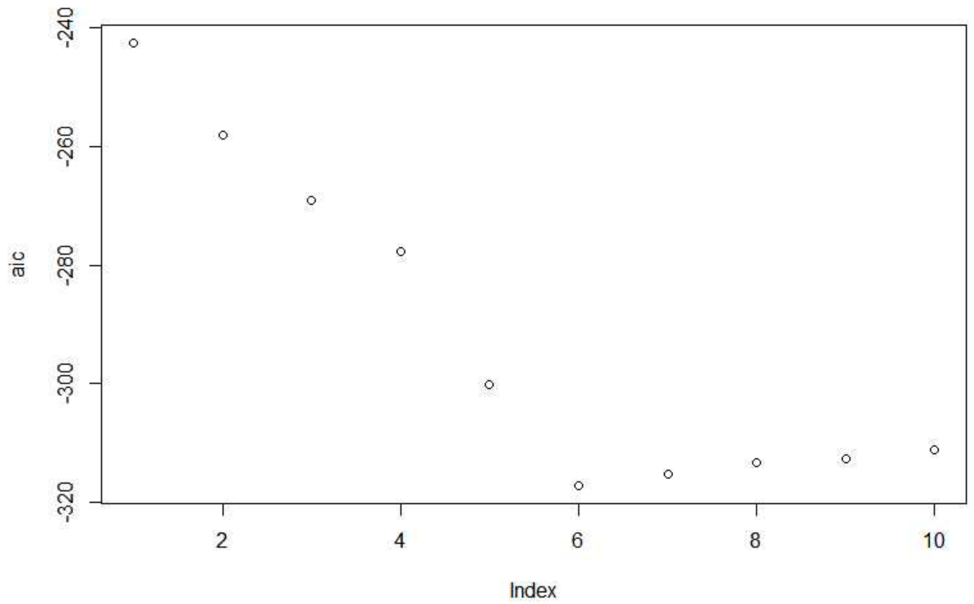
1. 원계열과 로그계열 시도표를 비교하여 ARIMA 모델링시 로그변환 하는 것이 그렇지 않은 것 보다 더 좋은 이유를 설명하시오. 이후의 질문은 모두 로그 계열에 대한 것이다.

	로그계열시도표 y=export.data$export log.y=log(y) plot(log.y,type = 'l') 로그변환은 분산 안정화에 도움이 되기 때문에 로그 변환을 하는 것이 원계열보다 좋다. 

2. 로그계열 시도표를 보고 차분을 할 필요가 있는지를 설명하시오?

로그계열 시도표를 보면 증가 추세가 있다.
비정상이다. 1차 차분을 하여 정상화 과정을 거치는 것이 바람직하다.

3. 단위근 검정을 하기 위해 로그계열의 AIC order 를 구하면 얼마인가?



AR(6)일 때 AIC 가 최소가 된다.
p=6이 AR 차수로 적합하다.
따라서 AIC order = 6 이다.

(p = 6)

단위근 검정결과 ADF 값과 이의 p-값은? (추세가 있음에 유의). 단위근 검정결과를 가지고 차분이 필요한 이유를 설명하시오.

검정결과 ADF 값: -2.5556 p-값: 0.3425 차분이 필요한 이유: 비정상 계열에 대해서 ARMA 추정할 경우 추정계수가 bias되는 경향이 있기 때문이다. H0 : Yt가 단위근에 의한 비정상적 H1 : Yt가 결정적 추세와 정상적인 오차 p-값이 0.05 이상이면 H0는 기각될 수 없으며 1차 차분이 계열을 정상화하는데 필요하다. p = 0.3425로 0.05 이상이므로 유의수준 5% 하	코드 > library(fUnitRoots) > adfTest(log.y, type = "ct", lags = 5) # p=6이 AR 차수로 적합하므로 lags=5로 지정한다. Title: Augmented Dickey-Fuller Test Test Results:

에서 H0를 기각할 수 없다.

```
PARAMETER:
  Lag Order: 5
STATISTIC:
  Dickey-Fuller: -2.5556
P VALUE:
  0.3425
```

4. 로그계열에 대해 ARIMA(p,1,q) 모형을 추정하려고 한다. 차수 p, q 의 BIC order 는 얼마 인가? (단 p 는 9 이하, q 는 4 이하로 한다.) (p,q) = (6,2)

bic

bic order = (6,2)

코드

```
> aic=matrix(rep(0, 9*4), 9,4); bic=matrix(rep(0, 9*4),
9,4);
> for (p in 1:9) {for (q in 1:4)
+   {aic[p,q] = Arima(log.y, order=c(p-1,0,q-1))$aic
+    bic[p,q] = Arima(log.y, order=c(p-1,0,q-1))$bic}}
> bic
           [,1]       [,2]       [,3]       [,4]
 [1,]  276.9746   73.86107  -22.48592  -92.63801
 [2,] -232.1903 -280.02128 -287.13431 -281.75822
 [3,] -244.3815 -235.22786 -280.60231 -276.25009
 [4,] -251.7686 -283.52988 -297.33511 -276.42784
 [5,] -256.9668 -282.63474 -238.57321 -270.87973
 [6,] -276.0417 -284.64770 -272.63456 -326.25079
 [7,] -289.6487 -284.28786 -327.88034 -290.73569
 [8,] -284.3035 -278.87794 -282.00871 -278.95994
 [9,] -279.0005 -275.71506 -279.89127 -312.21360
> min(bic)
[1] -327.8803
```

5. 추정된 모형의 방정식은?

```
Y = log(export)
> arima.fit =Arima(Y, order = c(6,1,2))
> arima.fit
Series: Y
ARIMA(6,1,2)

Coefficients:
         ar1      ar2      ar3      ar4      ar5     ar6      ma1     ma2
      0.2823  -0.7445  -0.3625  -0.3481  -0.1165  0.0069  -0.9744  0.9999
s.e.  0.0674   0.0692   0.0826   0.0830   0.0701  0.0676   0.0187  0.0295

sigma^2 = 0.011:  log likelihood = 191.56
AIC=-365.12   AICc=-364.3   BIC=-334.18
```

$$\Delta \ln y_t = 0.2823 \cdot \Delta \ln y_{t-1} - 0.7445 \cdot \Delta \ln y_{t-2} - 0.3625 \cdot \Delta \ln y_{t-3} - 0.3481 \cdot \Delta \ln y_{t-4} - 0.1165 \cdot \Delta \ln y_{t-5} + 0.0069 \cdot \Delta \ln y_{t-6} + a_t - 0.9744\, a_{t-1} + 0.9999 \cdot a_{t-2}$$

6. 미래 1년 (12 기간)까지 로그계열의 예측치, 95% 예측구간을 구하고 그림을 그려라.

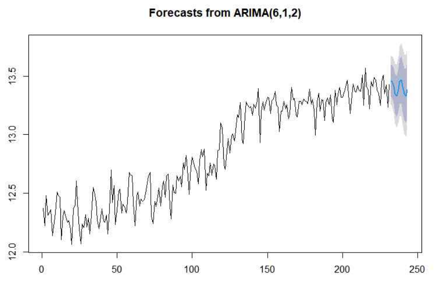
예측값

```
> arima.fit =Arima(Y, order = c(6,1,2))
> arima.hat=forecast(arima.fit, h=12)
> arima.hat
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
232	13.46016	13.32517	13.59515	13.25371	13.66661
233	13.45592	13.31451	13.59733	13.23965	13.67219
234	13.41629	13.26656	13.56602	13.18730	13.64528
235	13.35146	13.18555	13.51736	13.09773	13.60518
236	13.32882	13.15270	13.50493	13.05948	13.59816
237	13.38403	13.20009	13.56797	13.10272	13.66534
238	13.45449	13.26044	13.64854	13.15772	13.75126
239	13.46864	13.26430	13.67298	13.15613	13.78115
240	13.41533	13.20177	13.62889	13.08872	13.74194
241	13.34717	13.12523	13.56911	13.00774	13.68660
242	13.33137	13.10193	13.56080	12.98048	13.68226
243	13.38422	13.14765	13.62078	13.02242	13.74602

Forecast 는 예측치이고
Lo 95는 95% 예측구간이다.

그림

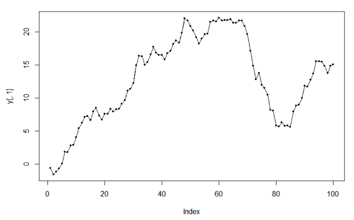
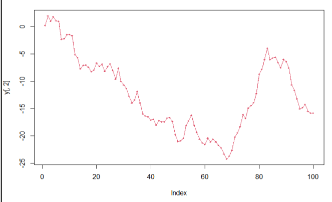


plot(arima.hat)

B. 다음 2차원 VAR 모형을 고려하자. $Y_t = AY_{t-1} + e_t$, t = 1, ..., n, n = 100

$$Y_t = (Y_{1t}, Y_{2t})', \quad A = \begin{pmatrix} 1.2 & 0.2 \\ -0.4 & 0.6 \end{pmatrix}, \quad e_t \sim iid\ N(0, I_2), \quad Y_1 = e_1$$

1. set.seed(1) 한다음 $Y_t = (Y_{1t}, Y_{2t})'$, t = 1, ..., n 발생시켜 시도표를 그려라.

y1 시도표 	y2 시도표 	코드 plot(y[,1], type = "b", pch = 19, cex = 0.6) lines(y[,1]) plot(y[,2], type = "b", pch = 18, cex = 0.6, col = 2) lines(y[,2], col = 2)

2. Y_1, Y_2 가 I(0) 계열인지 또는 I(1) 계열인지를 행렬 A 의 eigenvalue로부터 결정하시오.

A 의 eigenvalue > eigen(A) eigen() decomposition $values [1] **1.0 0.8** $vectors [,1] [,2] [1,] 0.7071068 -0.4472136 [2,] -0.7071068 0.8944272 A 의 eigenvalue는 1.0과 0.8이다.	결론 eigenvalue이 절대값이 1 보다 작으면 정상계열이다. 이 경우 하나가 절대값이 1이기 때문에 Y1과 Y2 모두 I(1) 계열이다. 정상계열이 아니다.

3. Y_t 에 대해 VEC 모형을 쓰시오. 이때 행렬 Π 의 rank 는 얼마인가? 이로부터 Y_1, Y_2 가 공적분관계임을 설명하시오. (rank = 1)

```
> summary(rank.result)

######################
# Johansen-Procedure #
######################

Test type: maximal eigenvalue statistic (lambda max) , with linear trend

Eigenvalues (lambda):
[1] 0.33680438 0.06355571
```

```
Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 1 |  6.44  6.50  8.18 11.65
r = 0  | 40.25 12.91 14.90 19.19

Eigenvectors, normalised to first column:
(These are the cointegration relations)

         y1.l2     y2.l2
y1.l2 1.000000 1.0000000
y2.l2 1.053659 0.3896221

Weights W:
(This is the loading matrix)

          y1.l2       y2.l2
y1.d  0.2095887 -0.04808254
y2.d -0.3438920 -0.03446522
```

```
             test 10pct  5pct  1pct
2) r <= 1 |  6.44  6.50  8.18 11.65
1) r = 0  | 40.25 12.91 14.90 19.19
```

1) 40.25 이 critical value 중 제일 큰 값인 19.19보다 크므로 H0를 유의수준 1%에서 기각한다. 따라서 r>0 이다.
2) 6.44는 6.50 보다 작으므로 유의수준 10%에서 H0를 기각 못한다. 따라서 r<=1이다.

따라서 rank = 1 이다.

Y1과 Y2가 비슷한걸로 보아 Y1과 Y2가 공적분관계라는 것을 알 수 있다.

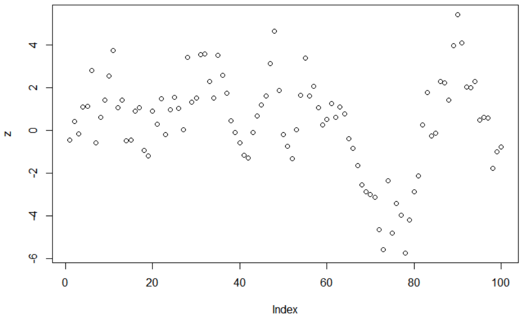
4. 공적분 벡터 β 를 구하시오 (1 0.9998379)
$\Pi = \alpha\beta$ 롤 표현할 때 α 는? $(\ 0.2142 - 0.4055\ )'$

```
# vec(0) 모형 추정
> vecm.fit = VECM(Data, lag = 0, r = 1, estim = "ML", include = "none")
> summary(vecm.fit)
```

#############
###Model VECM
#############
Full sample size: 100   End sample size: 99
Number of variables: 2   Number of estimated slope parameters 2
AIC -22.56111 BIC -14.77575 SSR 171.4906
Cointegrating vector (estimated by ML):
   y1   y2
r1 1 0.9998379

              ECT
Equation y1 0.2142(0.0412)***
Equation y2 -0.4055(0.0425)***

5. $z_t = \beta Y_t$ 의 시도표를 그려보고 이것이 I(0) 또는 I(1) 인지를 설명하시오.

<table>
<tr>
<td>z 의 시도표
z=1*y[,1]+0.9998379*y[,2]
plot(z)
</td>
<td>설명
I(0) 이다.
y1과 y1가 비슷하다.</td>
</tr>
</table>