Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

Løsningsforslag - Øving 7, 2005

Exercise 6.11

a)

The correlation structure related to the seasonal component indicates an AR(1) model, while the correlation structure between the seasonal increments is given by an MA(1) model. Indeed, if Y_t is given by the following ARMA model,

$$(1-\Phi B^{12})Y_t = (1+\theta B)Z_t,$$

where $Z_t \sim WN(0,\sigma^2)$, then the correlation structure of Y_t is found to be

$$\rho_Y(12j) = \Phi^{|j|}, \quad j = 0, \pm 1, \pm 2, \ldots$$

$$\rho_Y(12j \pm 1) = \theta\,\Phi^{|j|}, \quad j = 0, \pm 1, \pm 2, \ldots$$

$$\rho_Y(h) = 0, \quad \text{otherwise.}$$

This leads to a SARIMA$(0,1,1)\times(1,1,0)_{12}$ with $\Phi = 0.8$ and $\theta = 0.4$. To find the white noise variance, we use the equation

$$\mathrm{E}[(Y_t - \Phi\,Y_{t-12})(Y_t - \Phi\,Y_{t-12})] = \mathrm{E}[(Z_t + \theta\,Z_{t-1})(Z_t + \theta\,Z_{t-1})],$$

which gives the result

$$\gamma_Y(0) - 2\Phi\gamma_Y(12) + \Phi^2\gamma_Y(0) = \gamma_Y(0) - \Phi^2\gamma_Y(0) = \sigma^2(1+\theta^2).$$

Hence

$$\sigma^2 = \gamma_Y(0)\frac{1-\Phi^2}{1+\theta^2}$$

which leads to $\sigma^2 = 7.76$.

b)

Under the common assumptions made in Section 6.4 in the textbook, we may write

$$P_n X_{n+1} = X_n + X_{n-11} - X_{n-12} + P_n Y_{n+1}$$

where we have used that $P_n X_{n+h-j} = X_{n+h-j}$ for $j \geq h$. For large n it follows from equation (3.3.12) in the textbook, which applies to Y_t, that

$$P_n Y_{n+1} = \Phi\, Y_{n-11} + \theta_{n1}(Y_n - \hat{Y}_n)$$

That is,

$$P_n X_{n+1} = X_n + X_{n-11} - X_{n-12} + \Phi\, Y_{n-11} + \theta_{n1}(Y_n - \hat{Y}_n)$$

Similarly, for $P_n X_{n+12}$ we find that

$$P_n X_{n+12} = P_n X_{n+11} + X_n - X_{n-1} + P_n Y_{n+12}$$

Again, from equation (3.3.12) it follows that $P_n Y_{n+h} = \Phi\, Y_{n+h-12}$ for $h \geq 2$. Hence we find that

$$P_n X_{n+12} = X_n - X_{n-1} + P_n X_{n+11} + \Phi Y_n$$

Similarly,

$$P_n X_{n+11} = X_{n-1} - X_{n-2} + P_n X_{n+10} + \Phi Y_{n-1}$$

etc.

$$P_n X_{n+2} = X_{n-10} - X_{n-11} + P_n X_{n+1} + \Phi Y_{n-10}$$

Collecting the results:

$$\begin{aligned} P_n X_{n+12} &= X_n - X_{n-11} + P_n X_{n+1} + \Phi(Y_n + \ldots + Y_{n-10}) \\ &= 2X_n - X_{n-12} + \Phi(Y_n + \ldots + Y_{n-11}) + \theta_{n1}(Y_n - \hat{Y}_n) \end{aligned}$$

c)

For large n we may use equation (6.5.13) in the textbook to calculate the mean squared errors of the predictors in b). We need to calculate the expansion coefficients ψ_j of the relation

$$\sum_{j=0}^{\infty} \psi_j\, z^j = \frac{1 + \theta z}{(1-z)(1-\Phi z^{12})(1-z^{12})}$$

We obtain

$$\begin{aligned}\sum_{j=0}^{\infty} \psi_j \, z^j &= (1 + \theta z)(1 + z + z^2 + \ldots)(1 + \Phi z^{12} + \ldots)(1 + z^{12} + \ldots) \\ &= (1 + (1 + \theta)z + (1 + \theta)z^2 + \ldots)(1 + (1 + \Phi)z^{12} + \ldots)\end{aligned}$$

It follows that $\psi_0 = 1$ and $\psi_1 = \ldots = \psi_{11} = 1 + \theta$. Hence

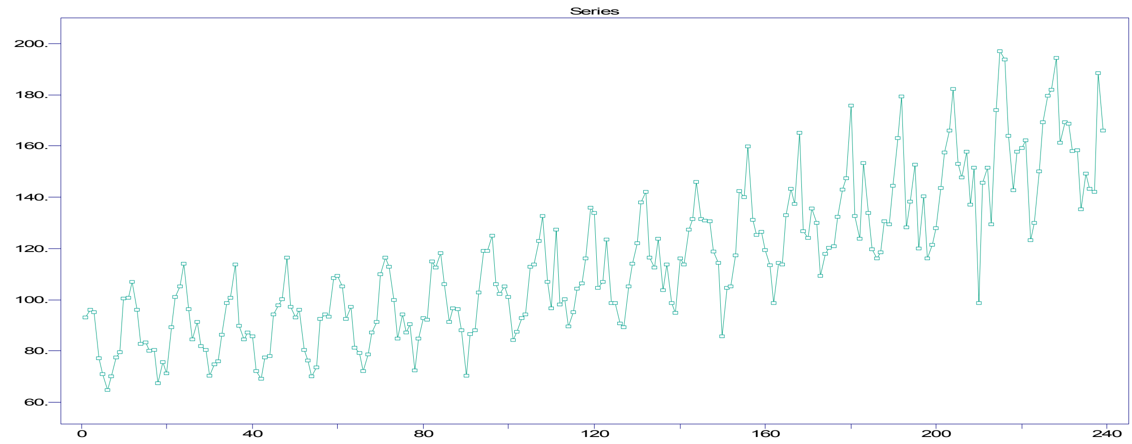
$$\sigma_n(1)^2 = \psi_0^2 \, \sigma^2 = \sigma^2$$

and

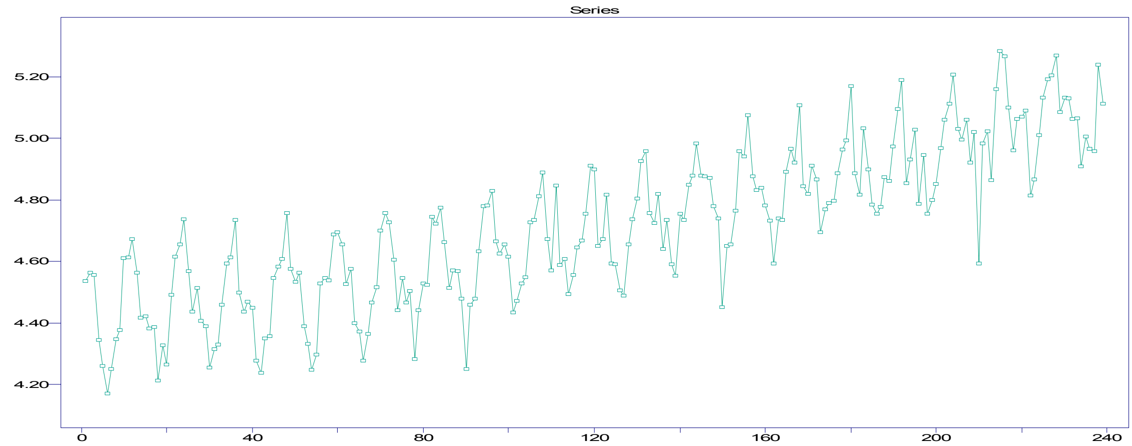
$$\sigma_n(12)^2 = \sum_{j=0}^{11} \psi_j^2 \, \sigma^2 = \sigma^2(1 + 11(1 + \theta)^2)$$

Exercise 6.9

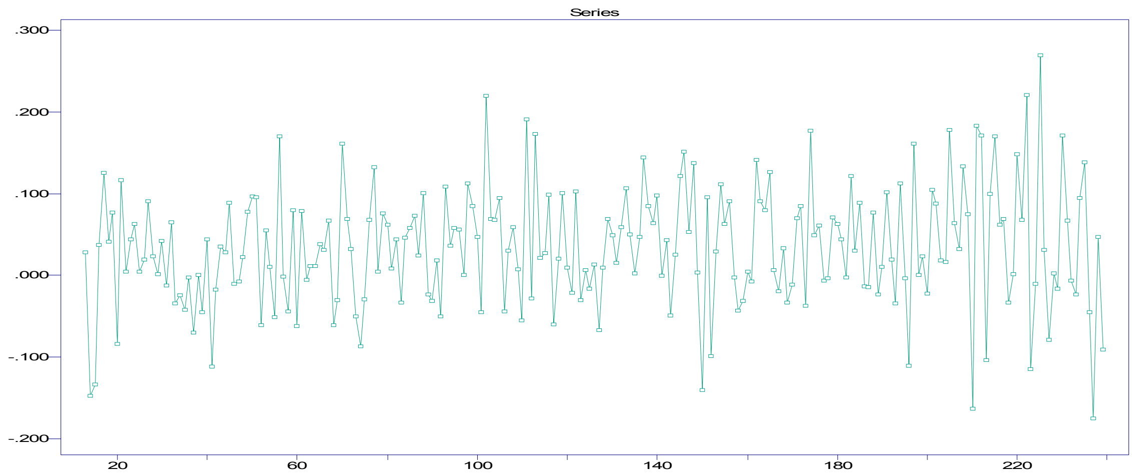
(a)



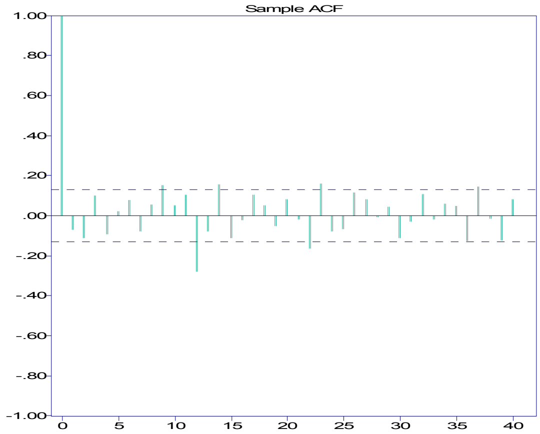
Apply a Box-Cox log-transformation to stabilize the increase in variability with level.

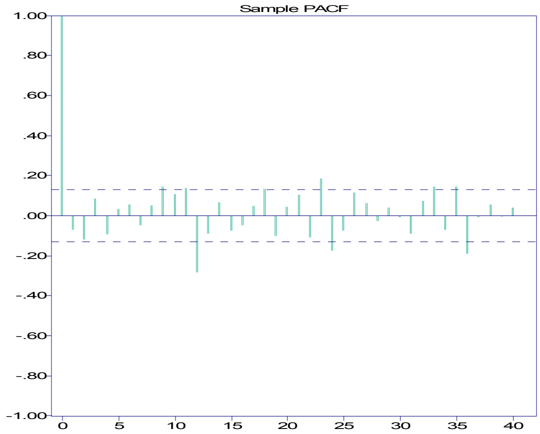


Apply ($1-B^{12}$) to eliminate seasonal component.



Subtract the mean value, since we will be fitting a zero-mean model.





The sample ACF suggests an MA(12), or perhaps an MA(14), process, while the sample PACF suggests an AR(12) process. However, the evidence would seem to slightly favour an MA process. Using the autofit option confirms that the best fit is obtained by an MA(14) model for p,q < 16.

==

ITSM::(Maximum likelihood estimates)

==

Method: Maximum Likelihood

ARMA Model:

X(t) = Z(t) - .01666 Z(t-1) - .02190 Z(t-2) + .1429 Z(t-3)
- .07141 Z(t-4) + .07244 Z(t-5) + .1022 Z(t-6) + .01836 Z(t-7)
+ .07015 Z(t-8) + .2698 Z(t-9) + .01332 Z(t-10) + .1669 Z(t-11)
- .6871 Z(t-12) - .03055 Z(t-13) + .2184 Z(t-14)

WN Variance = .003183

MA Coefficients

-.016656	-.021898	.142897	-.071407
.072443	.102231	.018356	.070150
.269795	.013323	.166938	-.687113
-.030550	.218354		

Standard Error of MA Coefficients

.064771	.064748	.045985	.045627
.045864	.042497	.042782	.042782
.042497	.045864	.045627	.045985
.064748	.064771		

(Residual SS)/N = .00318334

AICC = -.609626E+03

BIC = -.613495E+03

-2Log(Likelihood) = -.641901E+03

Accuracy parameter = .100000E-08

Number of iterations = 1

Number of function evaluations = 400335

Uncertain minimum.

Constrained optimization would seem a likely alternative for this case.

==
ITSM::(Maximum likelihood estimates)
==

Method: Maximum Likelihood

ARMA Model:
X(t) = Z(t) + .0000 Z(t-1) + .0000 Z(t-2) + .1238 Z(t-3)
 + .0000 Z(t-4) + .0000 Z(t-5) + .1654 Z(t-6) + .0000 Z(t-7)
 + .0000 Z(t-8) + .3228 Z(t-9) + .0000 Z(t-10) + .2346 Z(t-11)
 - .7181 Z(t-12) + .0000 Z(t-13) + .2321 Z(t-14)

WN Variance = .002988

MA Coefficients

.000000	.000000	.123764	.000000
.000000	.165379	.000000	.000000
.322751	.000000	.234625	-.718141
.000000	.232084		

Standard Error of MA Coefficients

.000000	.000000	.066166	.000000
.000000	.071322	.000000	.000000
.069721	.000000	.057810	.078165
.000000	.065422		

(Residual SS)/N = .00298788

AICC = -.624464E+03
BIC = -.648953E+03

-2Log(Likelihood) = -.638975E+03

Accuracy parameter = .000190000

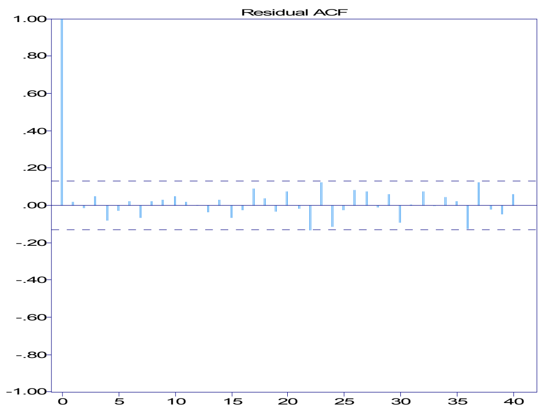
Number of iterations = 18

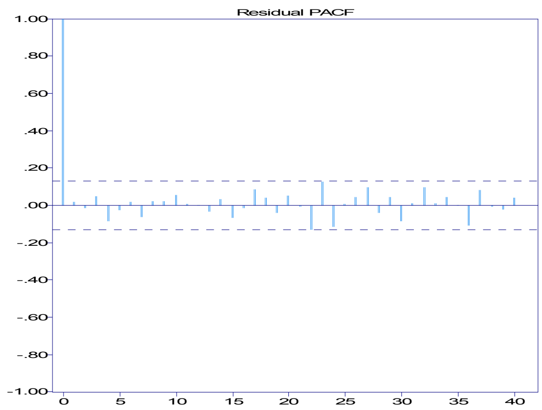
Number of function evaluations = 190

Optimization stopped within accuracy level.

The constrained model is seen to give a better AICC value, and is therefore preferred.

The sample ACF/PACF of the residuals:





The ACF and PACF give a reasonable indication of white noise residuals.

==

ITSM::(Tests of randomness on residuals)

==

Ljung - Box statistic = 24.225 Chi-Square (20), p-value = .23271

McLeod - Li statistic = 37.843 Chi-Square (26), p-value = .06262

Turning points = .15600E+03~AN(.15000E+03,sd = 6.3272), p-value = .34298

Diff sign points = .11600E+03~AN(.11300E+03,sd = 4.3589), p-value = .49130

Rank test statistic = .14105E+05~AN(.12826E+05,sd = .57188E+03), p-value = .02526

Jarque-Bera test statistic (for normality) = 2.8487 Chi-Square (2), p-value = .24066

Order of Min AICC YW Model for Residuals = 0

The Ljung-Box statistic, which is perhaps the most widely used test for white noise residuals, strongly indicates that the residual time series is a white noise.

(b)

The 95% confidence bounds are calculated for each estimated coefficient as follows:

θ_3: $0.1238 \pm 1.96 \cdot 0.0662 = -0.0059, 0.2536$

θ_6: $0.1654 \pm 1.96 \cdot 0.0713 = 0.0257, 0.3051$

θ_9: $0.3228 \pm 1.96 \cdot 0.0697 = 0.1862, 0.4594$

θ_{11}: $0.2346 \pm 1.96 \cdot 0.0578 = 0.1213, 0.3479$

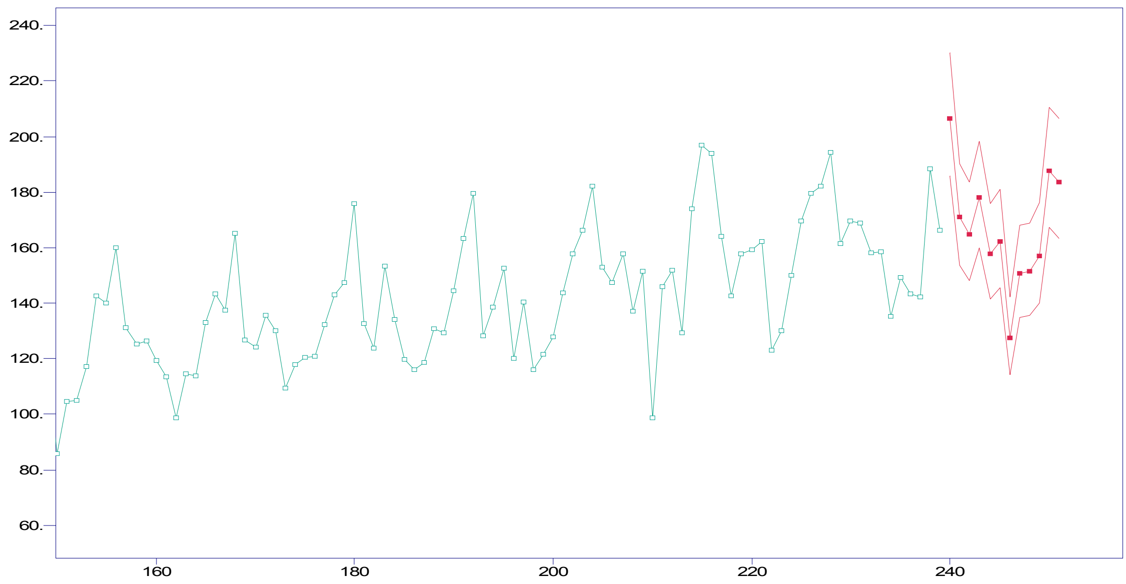
θ_{12}: $-0.7181 \pm 1.96 \cdot 0.0782 = 0.1008, 0.3274$

θ_{14}: $0.2321 \pm 1.96 \cdot 0.0654 = -0.8463, -0.5899$

(c)

The whiteness of residuals was discussed under point (a).

(d)



Plot (excerpt) of the original time series with 12 forecast values and 95% confidence bounds.

(e)

====================

ITSM::(ARMA Forecast)

====================

Approximate 95% Prediction Bounds

Step	Prediction	Lower	Upper
1	.20667E+03	.18567E+03	.23004E+03
2	.17094E+03	.15357E+03	.19027E+03
3	.16487E+03	.14812E+03	.18352E+03
4	.17811E+03	.15988E+03	.19841E+03
5	.15784E+03	.14169E+03	.17584E+03
6	.16234E+03	.14573E+03	.18085E+03
7	.12740E+03	.11420E+03	.14213E+03
8	.15059E+03	.13498E+03	.16800E+03
9	.15137E+03	.13568E+03	.16887E+03
10	.15709E+03	.14006E+03	.17619E+03
11	.18783E+03	.16747E+03	.21066E+03
12	.18362E+03	.16327E+03	.20650E+03

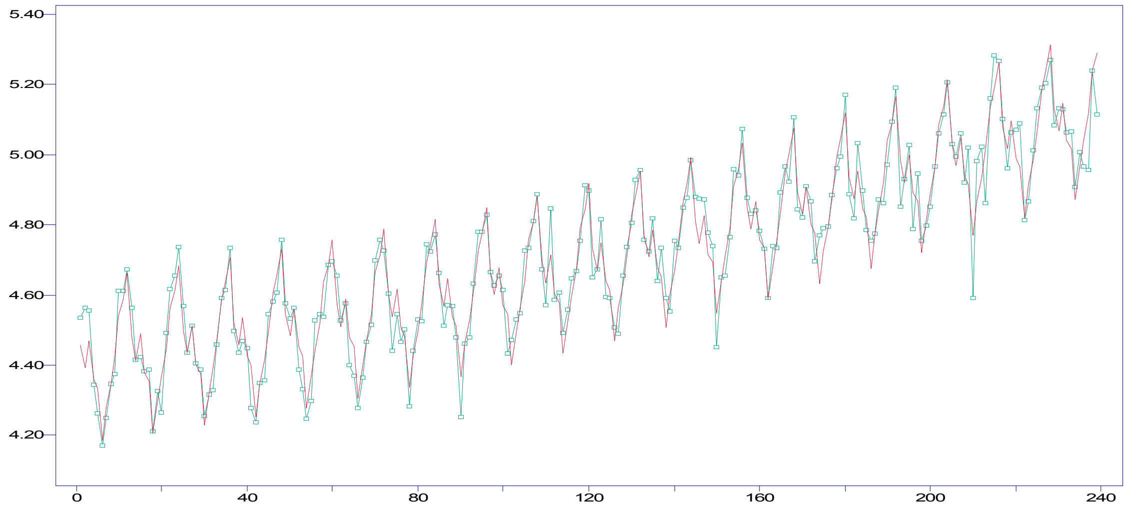
(f) Observed forecast errors

Observed values	Forecast values	Errors
199.2	206.67	7.47
182.7	170.94	-11.76
145.2	164.87	19.67
182.1	178.11	-3.99
158.7	157.84	-0.86
141.6	162.34	20.74
132.6	127.40	-5.2
139.6	150.59	10.99
147.0	151.37	4.37
166.6	157.09	-9.51
157.0	187.83	30.83
180.4	183.62	3.22

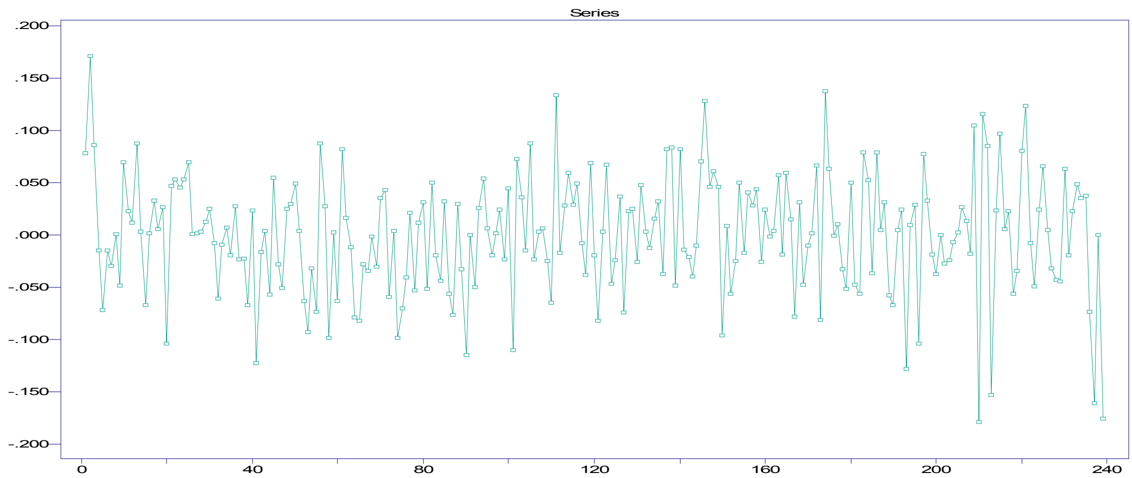
The last observed value is 180.4, which is within the 95% prediction bounds:
164.69, 208.72.

Exercise 6.10

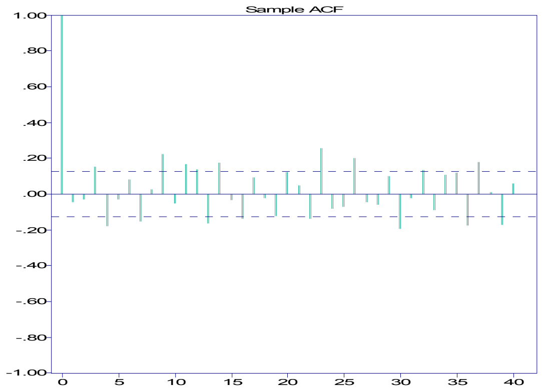
The classical decomposition method applied to the log-data.

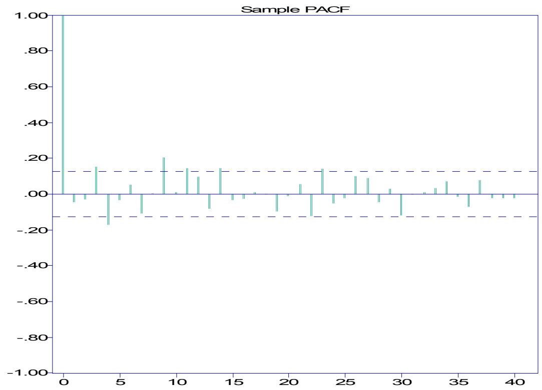


The 'stationary' time series



The sample ACF/PACF





The sample ACF/PACF do not convey a clear picture of an ARMA model to choose. Using the autofit option leads to an ARMA(2,2) model for p,q < 16.

==
ITSM::(Maximum likelihood estimates)
==

Method: Maximum Likelihood

ARMA Model:
X(t) = - 1.148 X(t-1) - .9991 X(t-2)
    + Z(t) + 1.129 Z(t-1) + .9896 Z(t-2)

WN Variance = .002684

AR Coefficients
    -1.147526   -.999137

Standard Error of AR Coefficients
    .002831   .002829

MA Coefficients
    1.129353   .989600

Standard Error of MA Coefficients
    .009794   .009793

(Residual SS)/N = .00268405

AICC = -.721957E+03
BIC = -.722607E+03

-2Log(Likelihood) = -.732215E+03

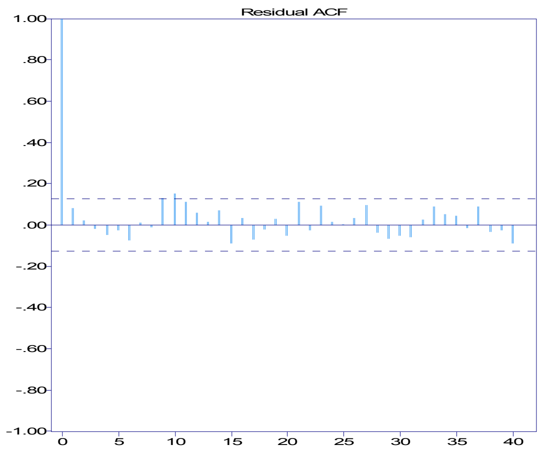
Accuracy parameter = .100000E-08

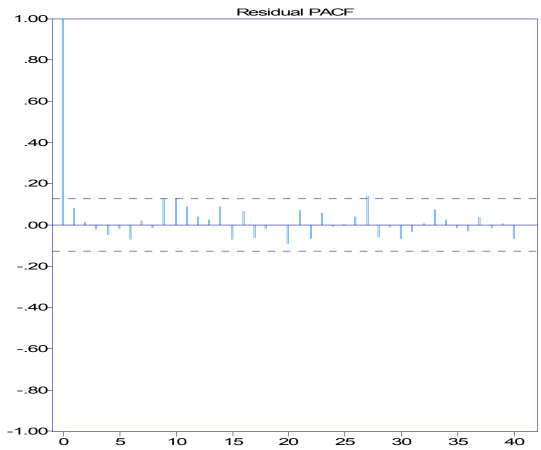
Number of iterations = 2

Number of function evaluations = 219772

Uncertain minimum.

The ACF/PACF of residual time series:





==

ITSM::(Tests of randomness on residuals)

==

Ljung - Box statistic = 30.295 Chi-Square (20), p-value = .06522

McLeod - Li statistic = 17.099 Chi-Square (24), p-value = .84440

Turning points = .15700E+03~AN(.15800E+03,sd = 6.4936), p-value = .87761

Diff sign points = .11000E+03~AN(.11900E+03,sd = 4.4721), p-value = .04417

Rank test statistic = .14580E+05~AN(.14221E+05,sd = .61772E+03), p-value = .56058

Jarque-Bera test statistic (for normality) = 3.0790 Chi-Square (2), p-value = .21448

Order of Min AICC YW Model for Residuals = 0

The AICC is lower for the present model than the one obtained in Exercise 6.9, but the residual statistics are not so good.

(b)

The 95% confidence bounds are calculated for each estimated coefficient as follows:

ϕ_1: $- 1.1475 \pm 1.96 \cdot 0.0028 = - 1.1530, - 1.1420$

ϕ_2: $- 0.9991 \pm 1.96 \cdot 0.0028 = - 1.0046, - 0.9936$

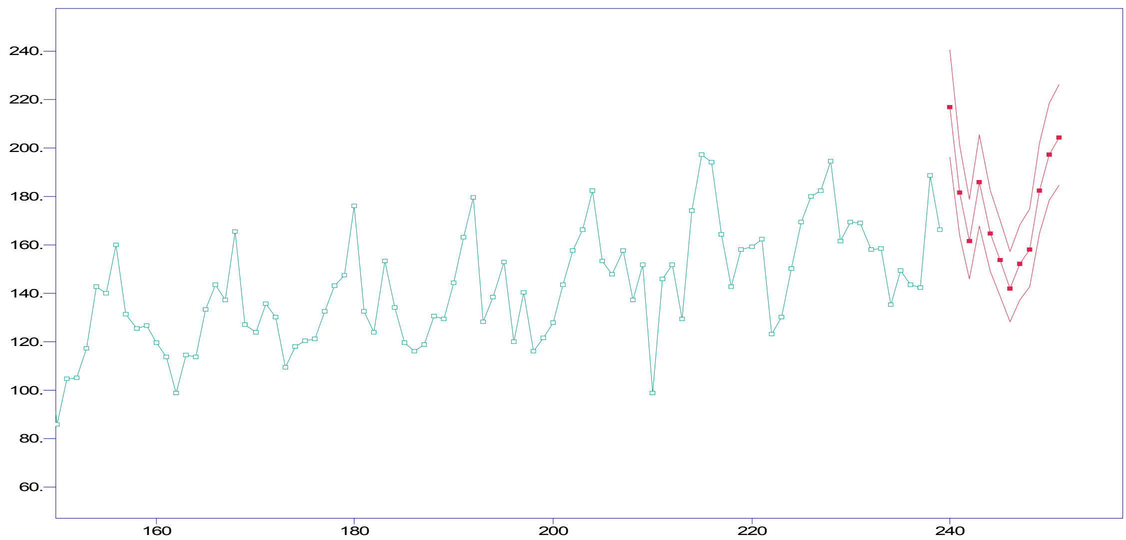
θ_1: $1.1293 \pm 1.96 \cdot 0.0098 = 1.1101, 1.1485$

θ_2: $0.9896 \pm 1.96 \cdot 0.0098 = 0.9704, 1.0088$

(c)

The whiteness of residuals was discussed under point (a).

(d)



Plot (excerpt) of the original time series with 12 forecast values and 95% confidence bounds.

(e)

====================

ITSM::(ARMA Forecast)

====================

Approximate 95% Prediction Bounds

Step	Prediction	Lower	Upper
1	.21689E+03	.19592E+03	.24010E+03
2	.18143E+03	.16388E+03	.20085E+03
3	.16157E+03	.14595E+03	.17887E+03
4	.18559E+03	.16764E+03	.20545E+03
5	.16476E+03	.14883E+03	.18240E+03
6	.15378E+03	.13891E+03	.17025E+03
7	.14193E+03	.12821E+03	.15713E+03
8	.15187E+03	.13718E+03	.16813E+03
9	.15794E+03	.14266E+03	.17486E+03

10	.18239E+03	.16475E+03	.20193E+03
11	.19722E+03	.17814E+03	.21835E+03
12	.20424E+03	.18447E+03	.22612E+03

(f)

Observed values	Forecast values	Errors
199.2	216.89	17.69
182.7	181.43	-1.27
145.2	161.57	16.37
182.1	185.59	3.49
158.7	164.76	6.06
141.6	153.78	12.18
132.6	141.93	9.33
139.6	151.87	12.27
147.0	157.94	10.94
166.6	182.39	15.79
157.0	197.22	40.22
180.4	204.24	23.84

The last observed value is 180.4, which is not within the 95% prediction bounds: 184.47, 226.12.

Comparing the forecast errors of this table with the corresponding one in Exercise 6.9, it is seen that while the errors are comparable for the shorter forecasts, the errors in Exercise 6.9 are significantly lower for the longer forecasts. The present model also leads to a systematic overprediction (except for one value).