Exercise 5.1

According to the large sample distribution of Yule-Walker estimators, we have that

$$\begin{bmatrix} \hat{\phi}_1 - \phi_1 \\ \hat{\phi}_2 - \phi_2 \end{bmatrix} \xrightarrow{d} N\big(0, \frac{\sigma^2}{n}\Gamma_2^{-1}\big)$$

where

$$\hat{\Gamma}_2 = \begin{bmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) \end{bmatrix} = \begin{bmatrix} 1382.2 & 1114.4 \\ 1114.4 & 1382.2 \end{bmatrix}$$

This gives

$$\hat{\Gamma}_2^{-1} = \begin{bmatrix} 0.002067 & -0.001667 \\ -0.001667 & 0.002067 \end{bmatrix}$$

The estimates $\hat{\phi}_1$ and $\hat{\phi}_2$ are obtained from the equation

$$\begin{bmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \end{bmatrix} = \hat{\Gamma}_2^{-1} \begin{bmatrix} \hat{\gamma}(1) \\ \hat{\gamma}(2) \end{bmatrix} = \begin{bmatrix} 1.318 \\ -0.634 \end{bmatrix}$$

while $\hat{\sigma}^2$ is given by the equation

$$\hat{\sigma}^2 = \hat{\gamma}(0) - [1.318, -0.634] \begin{bmatrix} 1114.4 \\ 591.73 \end{bmatrix} = 1382.2 - 1093.0 = 289.2$$

Estimated standard deviation of $\hat{\phi}_1$:

$$\frac{\hat{\sigma}}{\sqrt{n}}\sqrt{\big(\hat{\Gamma}_2^{-1}\big)_{11}} = 0.0773$$

Estimated standard deviation of $\hat{\phi}_2$:

$$\frac{\hat{\sigma}}{\sqrt{n}}\sqrt{\big(\hat{\Gamma}_2^{-1}\big)_{22}} = 0.0773$$

Approximate 95% confidence bounds

$$\phi_1 \;:\; 1.318 \pm 1.96 \times 0.0773 = 1.166, 1.470$$

$$\phi_2 \;:\; -0.634 \pm 1.96 \times 0.0773 = -0.786, -0.482$$

Exercise 5.2

(a) Using the DL algorithm it is obtained that

$$\hat{\phi}_{11} = \hat{\rho}(1) = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = 0.80625$$

$$\hat{v}_1 = \hat{\gamma}(0)\big(1 - \hat{\rho}(1)^2\big) = 483.72$$

$$\hat{\phi}_{22} = \frac{\hat{\gamma}(2) - \hat{\phi}_{11}\,\hat{\gamma}(1)}{\hat{v}_1} = -0.63412$$

$$\hat{\phi}_{21} = \hat{\phi}_{11}(1 - \hat{\phi}_{22}) = 1.3175$$

$$\hat{v}_2 = \hat{v}_1\big(1 - \hat{\phi}_{22}^2\big) = 289.21$$

$$\hat{\phi}_{33} = \frac{\hat{\gamma}(3) - \hat{\phi}_{21}\,\hat{\gamma}(2) - \hat{\phi}_{22}\,\hat{\gamma}(1)}{\hat{v}_2} = 0.08047$$

$$\hat{\phi}_{32} = \hat{\phi}_{22} - \hat{\phi}_{33}\hat{\phi}_{21} = -0.7401$$

$$\hat{\phi}_{31} = \hat{\phi}_{21} - \hat{\phi}_{33}\hat{\phi}_{22} = 1.3685$$

$$\hat{v}_3 = \hat{v}_2\big(1 - \hat{\phi}_{33}^2\big) = 287.34$$

(b) If the data are from an AR(2)-process, $\phi_{33} = 0$ and $\hat{\phi}_{33}$ is an observation from $N(0, 1/100)$. It is seen that $|\hat{\phi}_{33}| = 0.08047 < 1.96/10$. Hence, we would not reject the zero hypothesis $H_0: \; \phi_{33} = 0$ at the significance level 0.05.

Exercise 5.5

10 series of 200 observations of an MA(1) process with$\theta = \theta_1 = 0.6$ and white noise variance equal to 1.0, are generated by using for example ITSM. For the moment estimate ofθ, the following equation is used

$$\hat{\theta} = \left(1 - \sqrt{1 - 4\hat{\rho}(1)^2}\right) / \left(2\hat{\rho}(1)\right)$$

which is valid for $\hat{\rho}(1) \leq 0.5$. The moment estimates are conveniently calculated by pasting the values for $\hat{\rho}(1)$ into f.ex. Excel.

The Innovations and MLE estimates are obtained directly from ITSM for each data series. The sample mean and sample variance for each type of estimator can also be calculated by using f.ex. Excel.

Obtained results for the seed numbers chosen.

Simulation	rho(1)hat	thetahat	Moments	Innov. Alg.	MLE
1	0.3901	0.479966	0.48	0.592	0.6686
2	0.4001	0.500208	0.5002	0.5909	0.6135
3	0.4611	0.665026	0.665	0.5649	0.557
4	0.4569	0.649853	0.6499	0.6567	0.61
5	0.5489	#NUM!	1	0.755	0.6771
6	0.464	0.676099	0.6761	0.548	0.4859
7	0.4274	0.562754	0.5628	0.5284	0.5521
8	0.4238	0.553757	0.5538	0.6318	0.6378
9	0.3825	0.46532	0.4653	0.5324	0.5741
10	0.3611	0.426912	0.4269	0.5103	0.5581
	Sample variance =		0.027606	0.005475	0.003502
	Sample mean =		0.598	0.59104	0.59342
	Asymptotic theory:		0.023704	0.005	0.0032

The MLE is clearly best in terms of sample variance, The moment estimator is worst. The empirical means and variances of the estimators are in good agreement with the asymptotic theory.

Exercise 5.9

We can write the joint Gaussian density of the first n observations $X_1, \ldots, X_n$ for $n > p$ as

$$f(x_1,\ldots,x_n) = f(x_1,\ldots,x_p)\cdot f_{X_{p+1}|X_t,t\le p}(x_{p+1}|\,x_t,\,t\le p)\cdot\ldots\cdot f_{X_n|X_t,t\le n-1}(x_n|\,x_t,\,t\le n-1)$$

The first factor is, by equation (5.2.1) in the textbook

$$\frac{1}{(\sigma\sqrt{2\pi})^p\sqrt{\det G_p}}\ \exp\left\{-\frac{1}{2\sigma^2}\Big(\mathbf{x}_p' G_p^{-1}\mathbf{x}_p\Big)\right\}$$

where $\mathbf{x}_p = (x_1,\ldots,x_p)'$, $G_p = \sigma^{-2}\Gamma_p = \sigma^{-2}E[\mathbf{X}_p'\mathbf{X}_p]$. The remaining factors are

$$\frac{1}{\sigma\sqrt{2\pi}}\ \exp\left\{-\frac{1}{2\sigma^2}\Big(x_t-\phi_1 x_{t-1}-\ldots-\phi_p x_{t-p}\Big)^2\right\}\,,\quad t = p+1,\ldots,n$$

since, conditional on $X_s = x_s,\ s<t,\ X_t$ is normally distributed with mean value $\phi_1 x_{t-1} + \ldots + \phi_p x_{t-p}$ and variance σ^2. Multiplying these factors together gives the required result.

Exercise 5.11

The reduced likelihood is

$$l_2(\phi) = \ln\frac{S(\phi)}{2} + \frac{1}{2}\left(\ln r_0 + \ln r_1\right)$$

where from Exercise 5.9

$$S(\phi) = \frac{x_1^2}{r_0} + \frac{(x_2-\phi x_1)^2}{r_1} = x_1^2(1-\phi^2) + (x_2-\phi x_1)^2 = x_1^2 + x_2^2 - 2\phi x_1 x_2$$

since $r_0 = (1-\phi^2)^{-1}$ and $r_1 = 1$. Therefore

$$l_2(\phi) = -\frac{1}{2}\ \ln(1-\phi^2) + \ln\big(\frac{1}{2}(x_1^2+x_2^2-2\phi x_1 x_2)\big)$$

and

$$\frac{\partial l_2(\phi)}{\partial\phi} = -\frac{1}{2}\,\big(\frac{-2\phi}{1-\phi^2}\big) + \frac{-2x_1x_2}{x_1^2+x_2^2-2\phi x_1x_2} = 0$$

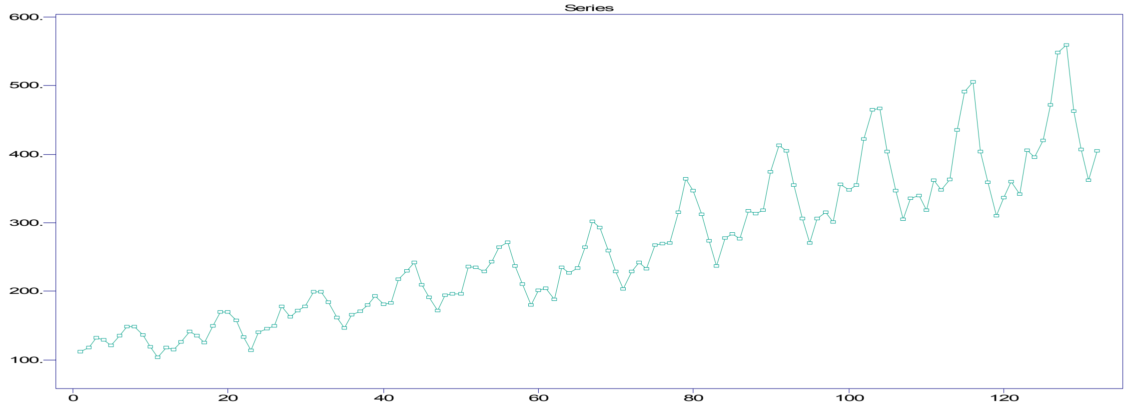
for $\phi = \frac{2x_1x_2}{x_1^2+x_2^2}$. Hence

$$\hat{\phi} = \frac{2x_1x_2}{x_1^2+x_2^2}$$

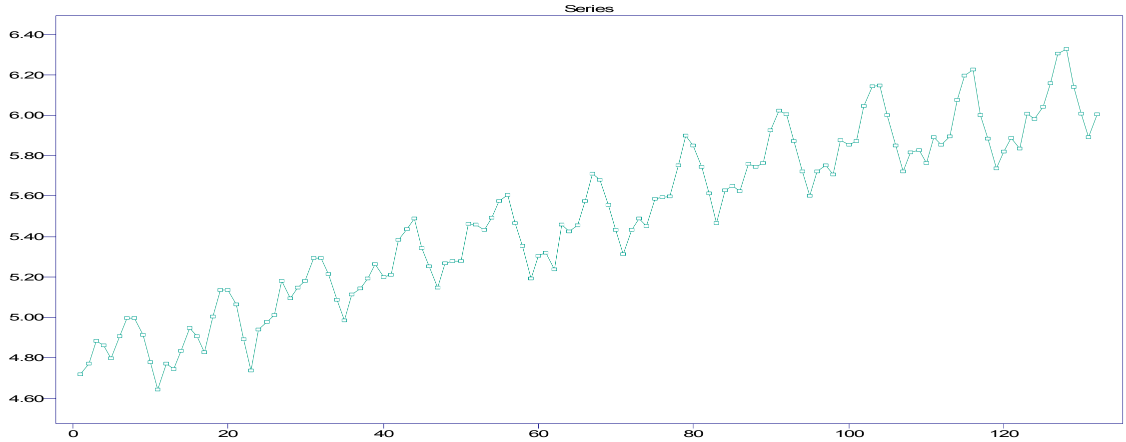
$$\hat{\sigma}^2 = \frac{1}{2}\,S(\hat{\phi}) = \frac{1}{2}\frac{(x_1^2-x_2^2)^2}{x_1^2+x_2^2}$$

Note that if $|x_1| = |x_2|$, then $\hat{\phi} = \mathrm{sgn}(x_1x_2)$ and $\hat{\sigma}^2 = 0$

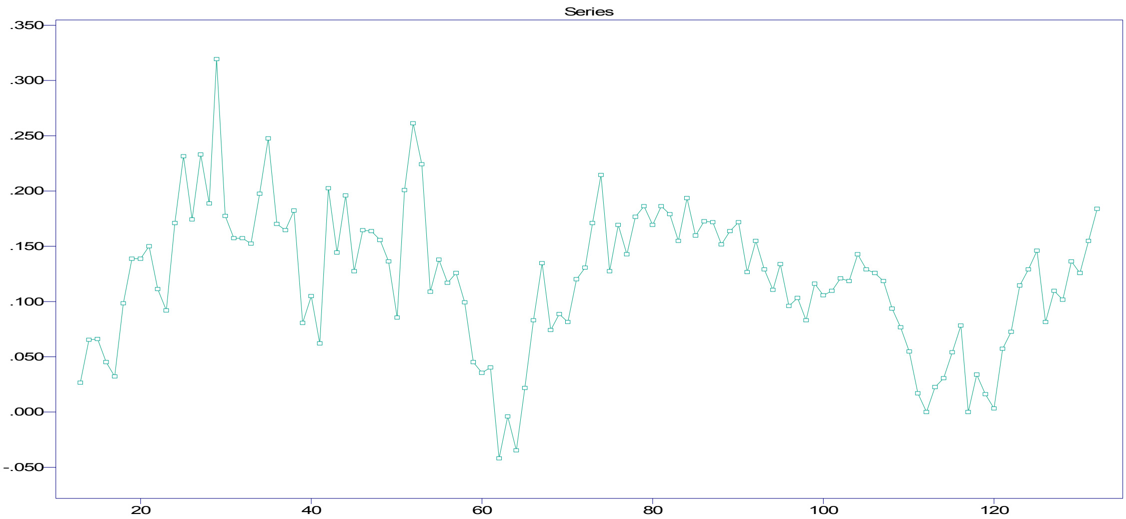
Exercise 6.7
(a)



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



Apply $(1- B^{12})$ to eliminate seasonal component. (Not clear whether (1-B) should then be applied, but it's better not to difference more than necessary, so we try without.



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

Looking at the sample ACF suggests an MA(12) process, while the sample PACF suggests an AR(13) process.

```
====================================
ITSM::(Preliminary estimates)
====================================

Method: Burg

AR(13)  Model:

X(t) = .5755 X(t-1) + .2621 X(t-2) - .1051 X(t-3) + .05330 X(t-4)
       + .09710 X(t-5) + .04217 X(t-6) - .08817 X(t-7) + .05339 X(t-8)
       + .1122 X(t-9) - .09645 X(t-10) - .1238 X(t-11) - .3537 X(t-12)
       + .3036 X(t-13)
     + Z(t)

WN Variance = .001258

AR Coefficients
       .575521      .262059     -.105093      .053296
       .097104      .042166     -.088171      .053387
       .112158     -.096446     -.123826     -.353748
       .303608


Ratio of AR coeff. to 1.96 * (standard error)
      3.456789     1.409346     -.553431      .280714
       .513497      .222763     -.466236      .282044
       .593106     -.507990     -.652084    -1.902449
      1.823582


WN variance estimate (Burg): .00127933

AICC = -.424881E+03


 ====================================
ITSM::(Preliminary estimates)
====================================

Method: Innovations

MA(12)  Model:

X(t) = Z(t) + .6077 Z(t-1) + .6004 Z(t-2) + .4408 Z(t-3)
```

+ .3443 Z(t-4) + .3930 Z(t-5) + .4010 Z(t-6) + .3737 Z(t-7)
+ .2877 Z(t-8) + .4015 Z(t-9) + .3140 Z(t-10) + .2456 Z(t-11)
- .2776 Z(t-12)

WN Variance = .001221

MA Coefficients

.607663	.600420	.440781	.344328
.393026	.400975	.373714	.287661
.401482	.314030	.245646	-.277590

Ratio of MA coeff. to 1.96 * (standard error)

3.396234	2.867790	1.873116	1.387390
1.536964	1.511921	1.360238	1.017326
1.396908	1.060022	.814666	-.910980

WN variance estimate (Innovations): .00116874

AICC = -.429279E+03

The best value of the AICC is obtained for the MA(12), so we proceed with that.

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

Method: Maximum Likelihood

MA(12) Model:

X(t) = Z(t) + .5933 Z(t-1) + .6646 Z(t-2) + .4825 Z(t-3)
+ .3040 Z(t-4) + .4000 Z(t-5) + .4004 Z(t-6) + .3744 Z(t-7)
+ .2962 Z(t-8) + .3346 Z(t-9) + .3245 Z(t-10) + .3975 Z(t-11)
- .2381 Z(t-12)

WN Variance = .001133

MA Coefficients

.593338	.664645	.482456	.304049
.400000	.400416	.374408	.296165
.334570	.324540	.397504	-.238073

Standard Error of MA Coefficients

.089093	.139933	.128916	.116267
.111234	.131150	.128830	.107477

.112888 .126111 .138832 .132335

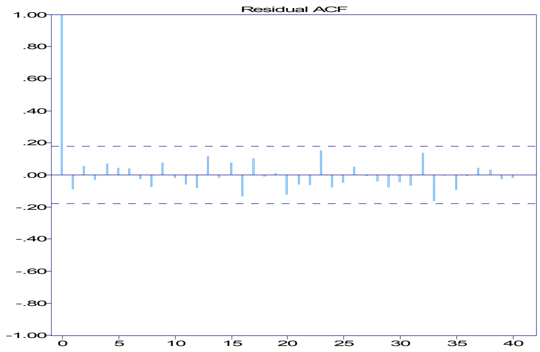
AICC = -.434962E+03

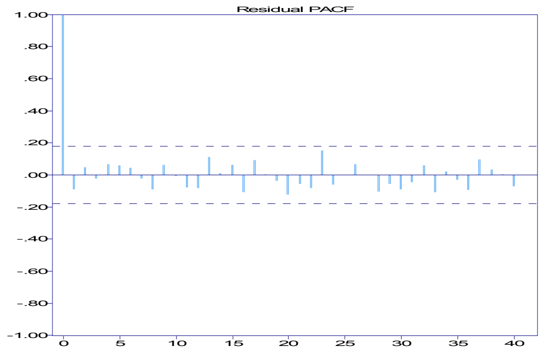
In view of the possibility that we could have applied (1-B) to the data, we may now try an ARMA(1,12) model. Using MLE, successively setting small coefficients to zero (constrained optimization), we get

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

Method: Maximum Likelihood


ARMA(1,12)  Model:

X(t) = .8745 X(t-1)
    + Z(t) - .3041 Z(t-1) + .0000 Z(t-2) - .2253 Z(t-3)
    - .2367 Z(t-4) + .0000 Z(t-5) + .0000 Z(t-6) + .0000 Z(t-7)
    + .0000 Z(t-8) + .0000 Z(t-9) + .0000 Z(t-10) + .0000 Z(t-11)
    - .6715 Z(t-12)

WN Variance = .001038

AR Coefficients
      .874466

Standard Error of AR Coefficients
      .050206

MA Coefficients
     -.304059    .000000    -.225261    -.236657
      .000000    .000000     .000000     .000000
      .000000    .000000     .000000    -.671455


Standard Error of MA Coefficients
      .112903    .000000     .101972     .106371
      .000000    .000000     .000000     .000000
      .000000    .000000     .000000     .100523

AICC = -.449503E+03
```

The AICC value obtained for the ARMA(1,12) model is better than for the MA(12) model above, so the ARMA model is the preferred one (based on the AICC criterion).

ACF/PACF of residuals





The ACF and PACF give good indication of white noise residuals.

==

ITSM::(Tests of randomness on residuals)

==

Ljung - Box statistic = 20.704 Chi-Square (20), p-value = .41472

McLeod - Li statistic = 21.199 Chi-Square (25), p-value = .68144

Turning points = 87.000~AN(78.667,sd = 4.5838), p-value = .06906

Diff sign points = 58.000~AN(59.500,sd = 3.1754), p-value = .63666

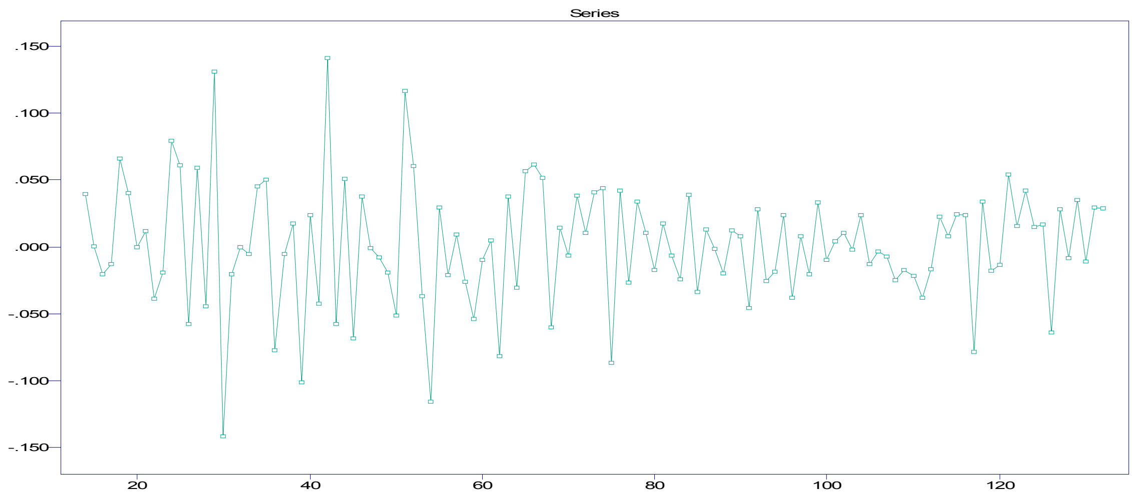
Rank test statistic = .33900E+04~AN(.35700E+04,sd = .22044E+03), p-value = .41418

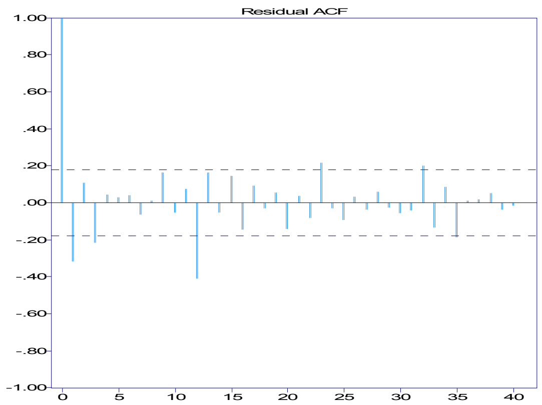
Jarque-Bera test statistic (for normality) = 8.7697 Chi-Square (2), p-value = .01246

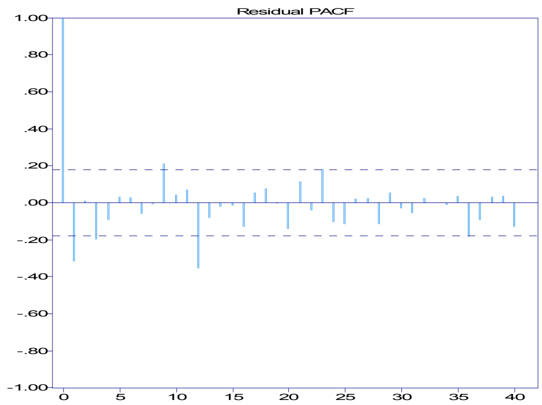
Order of Min AICC YW Model for Residuals = 0

These tests support a claim of IID noise.

An alternative approach is to apply $(1-B^{12})(1-B)$ to log(data):







The sample ACF suggests an MA(12) model, while the sample PACF indicates an AR(12) model. Again, after preliminary estimation, the MA(12) model seems to be the better one. However, repeated (constrained) MLE gives an AICC value larger than above. Hence we prefer the first model.

==

ITSM::(Maximum likelihood estimates)

==

Method: Maximum Likelihood

ARMA Model:

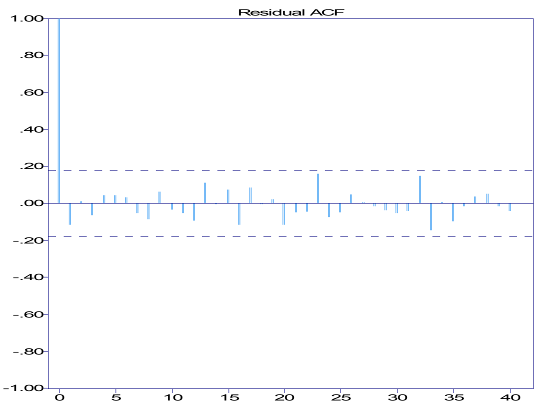
X(t) = Z(t) - .3366 Z(t-1) + .0000 Z(t-2) - .2439 Z(t-3)
- .2142 Z(t-4) + .0000 Z(t-5) + .0000 Z(t-6) + .0000 Z(t-7)
+ .0000 Z(t-8) + .0000 Z(t-9) + .0000 Z(t-10) + .0000 Z(t-11)
- .6404 Z(t-12)

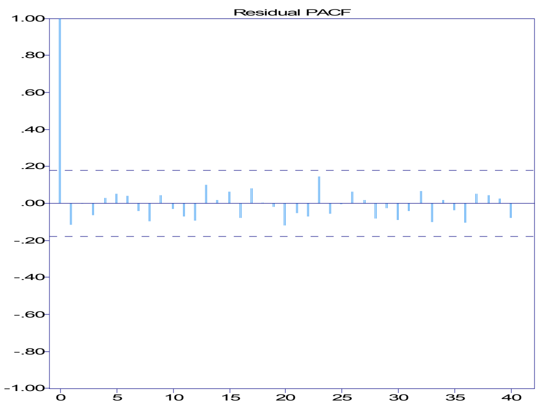
WN Variance = .001126

MA Coefficients

-.336593	.000000	-.243946	-.214239
.000000	.000000	.000000	.000000
.000000	.000000	.000000	-.640373

AICC = -.439855E+03





The ACF and PACF give good indication of white noise residuals.

==

ITSM::(Tests of randomness on residuals)

==

Ljung - Box statistic = 18.773 Chi-Square (20), p-value = .53662

McLeod - Li statistic = 29.020 Chi-Square (24), p-value = .21938

Turning points = 86.000~AN(78.000,sd = 4.5644), p-value = .07965

Diff sign points = 57.000~AN(59.000,sd = 3.1623), p-value = .52709

Rank test statistic = .34150E+04~AN(.35105E+04,sd = .21770E+03), p-value = .66089

Jarque-Bera test statistic (for normality) = 2.9459 Chi-Square (2), p-value = .22925

Order of Min AICC YW Model for Residuals = 0

These tests support a claim of IID noise.

(b)

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

ϕ_1: 0.875 ± 1.96 · 0.0502 = 0.777, 0.973

θ_1: -0.304 ± 1.96 · 0.1130 = -0.525, -0.083
θ_3: -0.225 ± 1.96 · 0.1020 = -0.436, -0.025

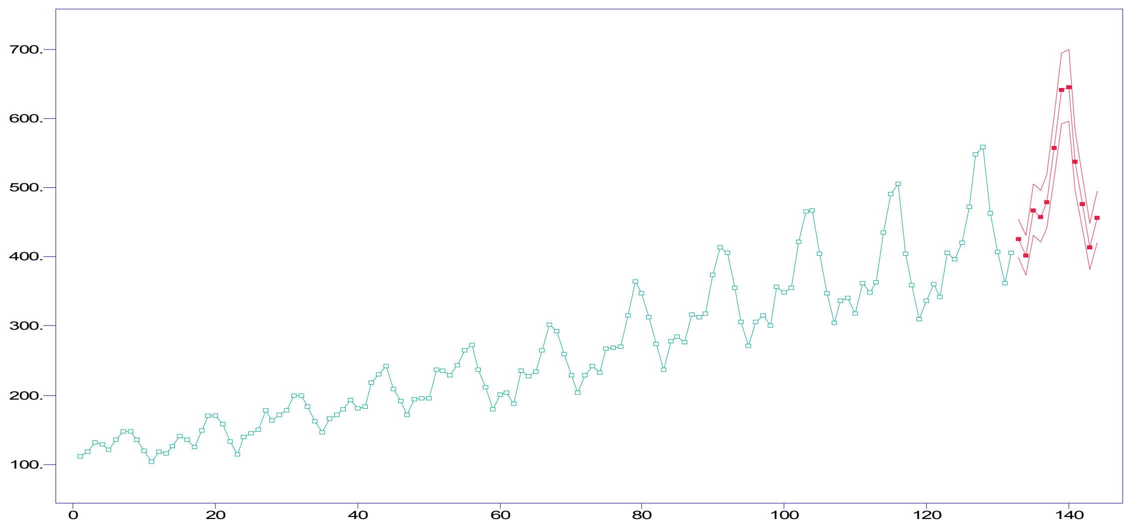
θ_4: -0.237 ± 1.96 · 0.1065 = -0.445, -0.027

θ_{12}: -0.671 ± 1.96 · 0.0502 = -0.868, -0.474

(c)

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

(d)



Graph of original data with forecasts and 95% prediction bounds.

(e)

====================
ITSM::(ARMA Forecast)
====================

Approximate 95 Percent Prediction Bounds

Step	Prediction	Lower	Upper
1	.42550E+03	.39946E+03	.45323E+03
2	.40116E+03	.37302E+03	.43141E+03
3	.46696E+03	.43139E+03	.50547E+03
4	.45737E+03	.42205E+03	.49564E+03
5	.47881E+03	.44181E+03	.51892E+03

6	.55671E+03	.51366E+03	.60337E+03
7	.64091E+03	.59133E+03	.69466E+03
8	.64491E+03	.59500E+03	.69901E+03
9	.53726E+03	.49566E+03	.58234E+03
10	.47652E+03	.43962E+03	.51651E+03
11	.41389E+03	.38183E+03	.44863E+03
12	.45591E+03	.42060E+03	.49419E+03

(f) Observed forecast errors

Observed values	Forecast values	Errors
417	425.50	8.50
391	401.16	10.16
419	466.96	47.96
461	457.37	3.63
472	478.81	6.81
535	556.71	21.71
622	640.91	18.91
606	644.91	38.91
508	537.26	29.26
461	476.52	15.52
390	413.89	23.89
432	455.91	23.91

The last observed value is 432, which is within the 95% prediction bounds: 420.6, 494.2.