Non-Stationary and Seasonal Time Series (Chap 6)

6.1 ARIMA Models

DEFINITION: $\{X_t\}$ is an **ARIMA(p,d,q) process** if

$$Y_t := (1 - B)^d X_t$$

is a causal ARMA(p,q) process.

Remarks :

1. $\{X_t\}$ satisfies the difference equation

$$\phi^*(B)\, X_t = (1 - B)^d\, \phi(B)\, X_t = \theta(B)\, Z_t$$

1

where $\phi^*(z)$ has a zero of order d at z=1.

2. If $d > 0$, a polynomial of degree d can be added to $\{X_t\}$ without violating the difference equation

$$\phi(B)(1 - B)^d X_t = \theta(B)\, Z_t\,.$$

Ex 6.1.1. (An ARIMA(1,1,0) process).

$$(1 - \phi B)\,(1 - B)\, X_t = Z_t\,,\ \{Z_t\} \sim \mathrm{WN}(0,\sigma^2)\,.$$

Write,

$$X_t = X_0 + \sum_{j=0}^{t} Y_j\,,$$

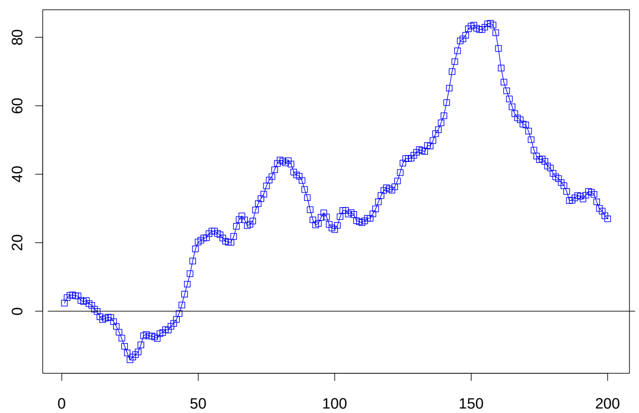
where $\quad Y_t = (1 - B)\, X_t = \sum_{j=0}^{\infty} \psi_j\, Z_{t-j}\,.$

2

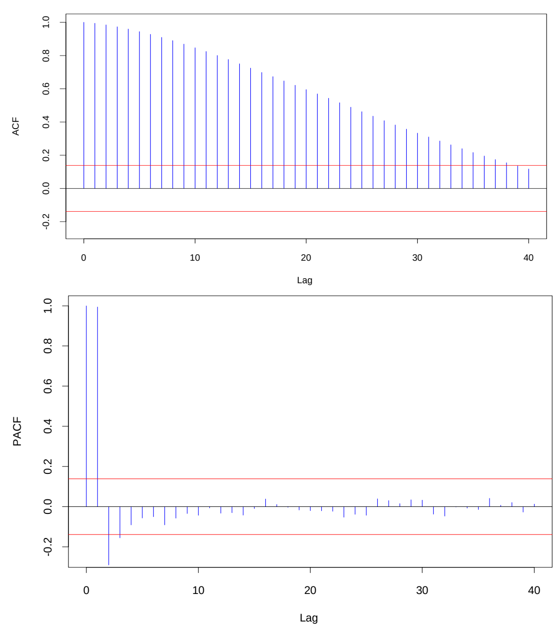
Figure 6.1 (200 observations from ARIMA(1,1,0)

$$(1-.8B)(1-B)X_t = Z_t, \quad \{Z_t\} \sim \mathrm{WN}(0,1)$$

with $X_t = 0$.)



Figs 6.2 & 6.3 (Sample ACF and PACF of data in Fig 6.1)



MLE AR(1) Model for $Y_t = (1 - B)\, X_t$:

$$(1 - .808B)\,(1 - B)\, X_t = Z_t\,, \quad \{Z_t\} \sim \mathrm{WN}(0, .978),$$

(Fitted model is close to true: $\phi = .8$, $\sigma^2 = 1$.)

MLE AR(2) model :

$$(1 - 1.808B - .801B^2)\, X_t = Z_t\,,$$

$$(1 - .825B)\,(1 - .983B)\, X_t = Z_t\,,$$

$$\{Z_t\} \sim \mathrm{WN}(0, .970).$$

(Fitted model is nearly non-stationary with a root of the AR polynomial near the unit circle.)

5

6.2 Identification Techniques

(a) Preliminary Transformations

Goal : Transform the data (if necessary) to achieve a more plausible realization of a stationary ts.

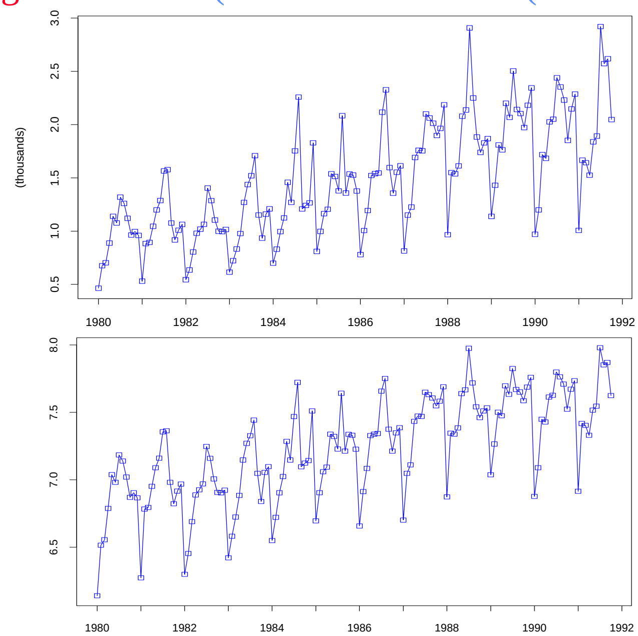
Box-Cox Transformation. Useful for

- transforming skewed data to symmetric data.
- stabilizing the variance.

$$f_\lambda(U_t) = \begin{cases} \lambda^{-1}(U_t^\lambda - 1), & U_t \geq 0, \lambda > 0, \\ \ln(U_t), & U_t > 0, \lambda = 0, \end{cases}$$

6

Figs 1.1 & 1.17 (Red wine data & ln (red wine).)



Elimination of trend and seasonality.

(i) classical decomposition-- trend, seasonal and random components.

(ii) differencing

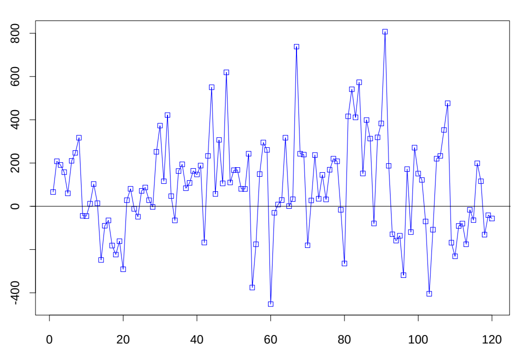


Fig 6.11 (Deseasonalized and linearly detrended $\ln(U_t)$.)

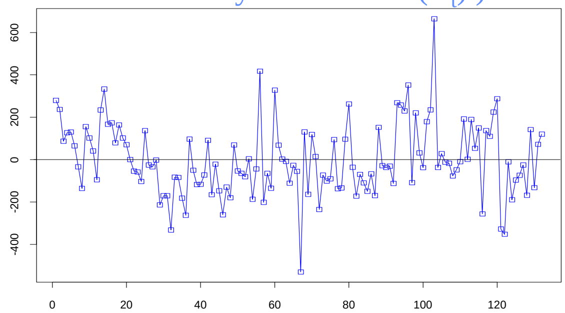
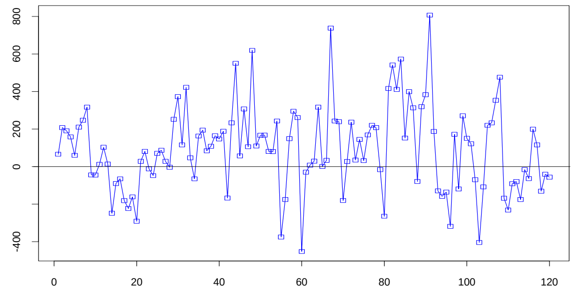


Fig 6.12

$(\ln(U_t) - \ln(U_{t-12}))$



(b) Identification and Estimation

(i) Transform data (if necessary) to make a stationary looking realization. Examine sample ACF and PACF to get an idea for p and q.

(ii) Use option Preliminary Estimation of ITSM to estimate preliminary models. (Use min AICC for automatic order selection of AR models.)

(iii) Find maximum likelihood models for the candidate models in (ii) using the option ARMA estimation.

(iv) Find the minimum AICC model among the fitted models in (iii).

(v) Examine estimated coefficients and standard errors of coefficients to see if any should be set to 0. (Fit subset models using the option, Constrain optimized coefficients, if necessary.)

(vi) Check final candidate models for goodness of fit. (Tests performed in the File and analyze residuals selection after maximizing the likelihood.)

Ex 6.2.1 (The red wine data.) Let

$$X_t = \ln(U_t) - \ln(U_{t-12}) - .0631$$

Fig 6.13

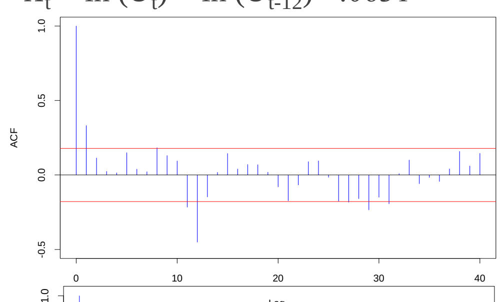
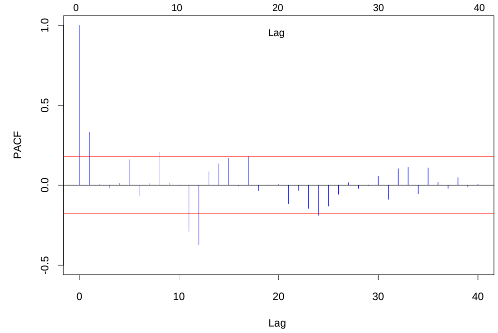


Fig 6.14



Modelling the red wine data: Sample ACF and PACF suggest AR(12) or MA(13).

Preliminary AR model: The automatic AR model fitting option of Preliminary estimation with Burg's algorithm selects the AR(12)

$$(1-.245B-.069B^2-.012B^3-.021B^4-.200B^5+.025B^6+.004B^7$$
$$-.133B^8+.010B^9-.095B^{10}+.118B^{11}+.384B^{12})\, X_t = Z_t\,,$$
$$\{Z_t\} \sim \mathrm{WN}(0,.0135),\quad \mathrm{AICC} = -158.77.$$

(MLE gives similar model with AICC = -158.87.)

Subset AR Model:

$$(1-.261B-.217B^5-.140B^8+.389B^{12})\, X_t = Z_t\,,$$
$$\{Z_t\} \sim \mathrm{WN}(0,.0140),\quad \mathrm{AICC} = -173.07.$$

13

Preliminary MA model: The best fitting preliminary MA model (using the innovations algorithm) is the MA(13)

$$X_t = (1+.270B+.190B^2+.087B^3+.025B^4+.258B^5+.131B^6+.100B^7$$
$$+.168B^8+.048B^9+.213B^{10}+.035B^{11}-.494B^{12}-.167B^{13})\, Z_t\,,$$
$$\{Z_t\} \sim \mathrm{WN}(0,.0121),\quad \mathrm{AICC} = -167.17.$$

(Std errors suggest zero coefficients at lags 3, 4, 6, 7, 9, 11. MLE gives similar model with AICC = -158.87.)

Subset MA Model: (non-invertible model)

$$X_t = (1+.247B+.205B^2+.229B^5+.275B^8+.265B^{10}-.608B^{12}$$
$$-.263B^{13})\, Z_t\,,\quad \{Z_t\} \sim \mathrm{WN}(0,.0110),\quad \mathrm{AICC} = -181.52$$

Remark: Subset MA model appears to be best fit.

14

6.3 Unit Roots in Time Series Models

AR unit root test

Consider the simple AR(1) model

$$X_t - \mu = \phi_1 (X_{t-1} - \mu) + Z_t, \quad \{Z_t\} \sim WN(0,\sigma^2)$$

Wish to test

H_0 : $\phi_1 = 1$ (unit root $\Longrightarrow$ $Y_t \sim I(1)$)

H_1 : $|\phi_1| < 1$ (no unit root $\Longrightarrow$ $Y_t \sim I(0)$)

Note: AR(1) model can be written as

$$\nabla X_t = \phi_0^* + (\phi_1 - 1) X_{t-1} + Z_t = \phi_0^* + \phi_1^* X_{t-1} + Z_t,$$

where $\phi_1^* = (\phi_1 - 1)$. So H_0 is equivalent to $\phi_1^* = 0$.

15

Now let $\hat{\phi}_1^*$ be the OLS estimator of $\phi*$ found by regressing ∇X_t on 1 and X_{t-1}. The estimated standard error is

$$SE(\hat{\phi}_1^*) = S(\sum_{t=2}^{n}(X_{t-1} - \overline{X})^2)^{1/2}, \qquad S^2 = \sum_{t=2}^{n}(\nabla X_t - \hat{\phi}_0^* - \hat{\phi}_1^* X_{t-1})^2 / (n-3).$$

Dickey-Fuller test statistic:

$$\tau_u = \hat{\phi}_1^* / SE(\hat{\phi}_1^*)$$

The .01, .05, and .10 quantiles of limiting distribution are -3.43, -2.86, and -2.57, respectively.

This procedure can be extended to the case when Y_t follows an AR(p) process. For example, suppose

$$X_t - \mu = \phi_1 (X_{t-1} - \mu) + \ldots + \phi_p (X_{t-p} - \mu) + Z_t, \quad \{Z_t\} \sim WN(0,\sigma^2)$$

which can be rewritten as

$$\nabla X_t = \phi_0^* + \phi_1^* X_{t-1} + \phi_2^* \nabla X_{t-1} + \ldots + \phi_p^* \nabla X_{t-p+1} + Z_t$$

16

Now if the model

$$\nabla X_t = \phi_0^* + \phi_1^* X_{t-1} + \phi_2^* \nabla X_{t-1} + \ldots + \phi_p^* \nabla X_{t-p+1} + Z_t$$

has a unit root, then $\phi_1^*=0$ and the differenced series $\{\nabla X_t\}$ follows an AR(p-1). As in the AR(1) case, ϕ_1^* can be estimated as the coefficient of X_{t-1} in the OLS regression of ∇X_t onto 1, X_{t-1}, ∇X_{t-1}, . . ., ∇X_{t-p+1}.

Augmented Dickey - Fuller test statistic:

$$\hat{\tau}_u = \hat{\phi}_1^* / SE(\hat{\phi}_1^*)$$

Has the same limit distribution as the DF test statistic.

17

Example 6.3.1 (Data in Ex 6.1.1)

PACF suggests an AR(2) or possibly an AR(3). We try the latter. Regressing ∇X_t on 1, X_{t-1}, ∇X_{t-1}, and ∇X_{t-2} for t= 2, …, 200 using OLS, we obtain

$$\nabla X_t = \underset{(.1135)}{.1503} - \underset{(.0028)}{.0041} X_{t-1} + \underset{(.0707)}{.9335} \nabla X_{t-1} - \underset{(.0708)}{.1548} \nabla X_{t-2} + Z_t$$

where $\{Z_t\} \sim$ WN(0,.9639). The augmented Dickey - Fuller test statistic is

$$\hat{\tau}_u = \hat{\phi}_1^* / SE(\hat{\phi}_1^*) = -.0041/.0028 = -1.464$$

Since -1.464 > -2.57, the unit root hypothesis is not rejected at α = .10. On the other hand, if we mistakenly used a t-distr with 193 df, the p-value is .074. The t-distribution can be used for the other coefficients. Based on the t-test, the intercept appears to be 0.

18

Example 6.3.1 (cont)

Repeating the analysis w/o intercept, we obtain the model

$$\nabla X_t = -.0012\, X_{t-1} + .9395\, \nabla X_{t-1} - .1585\, \nabla X_{t-2} + Z_t$$
$$\quad (.0018) \qquad (.0707) \qquad (.0709)$$

with a test statistic of

$$\hat{\tau} = \hat{\phi}_1^* / SE(\hat{\phi}_1^*) = -.0012/.0018 = -.667$$

The .01, .05, .10 cutoff values of the corresponding test statistic w/o an intercept are now -2.58, -1.95, and -1.62. Since -.667 is larger than -1.62, we clearly do not reject the null (unit root).

19

Example 6.1.1 in B&D

```
   Variable      Coeff        Std Error       T-Stat      Signif
*******************************************************************
*************
1.   Constant 0.150279326  0.113477590      1.32431   0.18696677
2.   X{1}    -0.004113559  0.002839488     -1.44870   0.14904485
3.   DX{1}    0.933549178  0.070748894     13.19525   0.00000000
4.   DX{2}   -0.154797403  0.070790303     -2.18670   0.02996668

Dickey-Fuller tau statistic      -1.44870
```

Do not reject H_0.

```
Dickey-Fuller Unit Root Test, Series X
Regression Run From 4 to 200
Observations 198
With intercept with 2 lags
T-test statistic      -1.44870
Critical values: 1%= -3.465 5%= -2.876 10%= -2.574
```

20

MA unit roots (B&D 6.3.2)

If $\{X_t\}$ is a causal-invertible ARMA(p,q) process,

$$\phi(B)X_t = \theta(B)Z_t, \qquad \{Z_t\} \sim \mathrm{WN}(0,\sigma^2)$$

then the differenced process ∇X_t is an ARMA(p,q+1) with MA polynomial $\theta(z)(1-z)$. That is, the MA polynomial now has a unit root. Consequently, testing for a unit root in the MA polynomial is equivalent to testing that the time series has been overdifferenced.

As a second application, it is possible to distinguish between the two competing models:

$$\nabla X_t = a + V_t$$

$$X_t = c_0 + c_1 t + W_t$$

where $\{V_t\}$ and $\{W_t\}$ are invertible ARMA processes. The differenced series in first case has no unit root while in the second case, differenced series has one unit root.

MA unit roots. Consider the simple case of an MA(1),

$$X_t = Z_t + \theta Z_{t-1}, \qquad \{Z_t\} \sim \mathrm{WN}(0,\sigma^2)$$

H_0 : $\theta = -1$ (unit root)

H_1 : $\theta > -1$ (no unit root)

Davis and Dunsmuir (1996) derived test based on MLE.

Reject H_0 if $\hat{\theta} > -1 + c_\alpha / n$ where c_α is the $(1-\alpha)$ quantile of the limit distribution of $n(\hat{\theta}+1)$.

($c_{.01}$= 11.93, $c_{.05}$= 6.80, $c_{.1}$= 4.90). For n =100, cutoff value for α =.05 is -1 + 6.80/100= -.932

Likelihood ratio test: Reject H_0 if

$$\lambda_n := -2\ln\left(\frac{L(-1, S(-1)/n)}{L(\hat{\theta}, \hat{\sigma}^2)}\right) > c_{LR,\alpha}$$

($c_{LR,.01}$= 4.41, $c_{LR,.05}$ = 1.94, $c_{LR,.1}$ = 1.00)

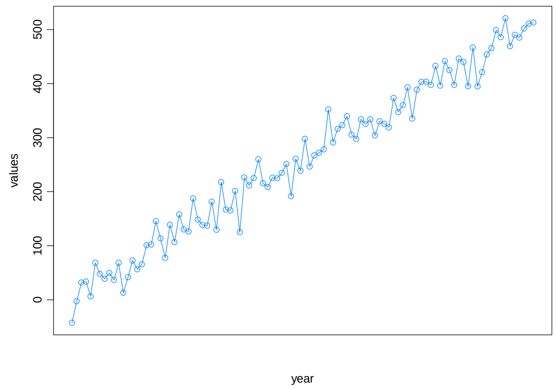
Remarks. 1. Must use MLE in the two tests described above.

2. If one also estimates a mean term, then one can reject H_0 at level α = .045 if

$$\hat{\theta} > -1.$$

This is quite striking!!

Example (trend + noise).



Results (using the program RATS)

```
open data maUnitRoot.dat
calendar 1900:1
all 2000:1
data(format=free,org=columns) / maexample
boxjenk(ma=1,diffs=1, constant,maxl) maexample

Annual Data From 189:01 To 287:01
Usable Observations      99       Degrees of Freedom    97
Centered R**2       0.971370      R Bar **2   0.971075
Uncentered R**2     0.993158      T x R**2      98.323
Mean of Dependent Variable        261.77833334
Std Error of Dependent Variable 147.44751847
Standard Error of Estimate        25.07688503
Sum of Squared Residuals         60998.465791
Log Likelihood                     -460.74002
Durbin-Watson Statistic              2.211689
Q(24-1)                             37.817041
Significance Level of Q            0.02664675
Variable       Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.   CONSTANT     5.126101     0.092936     55.15704  0.00000000
2.   MA{1}       -1.000001  1945.595505 -5.13982e-04  0.99959096
```

Do not reject H_0.

Overshorts Data

```
open data oshorts.dat
calendar 1900:1
all 1957:1
data(format=free,org=columns) / oshorts

boxjenk(ma=1,diffs=1, constant,maxl) oshorts

   Usable Observations     56      Degrees of Freedom    54
Centered R**2     -0.036851      R Bar **2  -0.056052
Uncentered R**2  -0.027585      T x R**2      -1.545
Mean of Dependent Variable       -5.50000000
Std Error of Dependent Variable  58.44126811
Standard Error of Estimate       60.05682450
Sum of Squared Residuals        194768.39710
Log Likelihood                    -309.80010
Durbin-Watson Statistic             2.901067
Q(14-1)                            31.407250
Significance Level of Q           0.00293852

   Variable     Coeff       Std Error      T-Stat     Signif
*******************************************************************
1.  CONSTANT  -0.10465      0.48179     -0.21722  0.82885596
2.  MA{1}     -1.00000  18352.34609 -5.44890e-05  0.99995672
```

Do not reject H_0.

Note if one mean corrects the data first, then the results are

```
open data oshorts.dat
calendar 1900:1
all 1957:1
data(format=free,org=columns) / oshorts

boxjenk(ma=1,diffs=1, demean,maxl) oshorts

Dependent Variable OSHORTS
Annual Data From 1714:01 To 1769:01
Usable Observations     56      Degrees of Freedom    55
Centered R**2     -0.196407      R Bar **2  -0.196407
Uncentered R**2  -0.185714      T x R**2     -10.400
Mean of Dependent Variable       -5.50000000
Std Error of Dependent Variable  58.44126811
Standard Error of Estimate       63.92327992
Sum of Squared Residuals        224740.21436
Log Likelihood                    -312.63169
Durbin-Watson Statistic             2.775851
Q(14-1)                            29.886183
Significance Level of Q           0.00489079

   Variable     Coeff       Std Error      T-Stat     Signif
********************************************************************
1.  MA{1}    -0.903114819  0.069149611    -13.06030  0.00000000
```

Cutoff value = -1 + 6.8/56 = -.879 implies we still reject H_0.

6.4 Forecasting ARIMA Models

$$(1 - B)^d X_t = Y_t, \quad t = 1, 2, \ldots,$$

where $\{Y_t\}$ is a causal ARMA(p, q) process, and that the random vector $(X_{1-d}, \ldots, X_0)$ is uncorrelated with Y_t, $t > 0$. The difference equations can be rewritten in the form

$$X_t = Y_t - \sum_{j=1}^{d} \binom{d}{j} (-1)^j X_{t-j}, \quad t = 1, 2, \ldots. \tag{6.4.1}$$

It is convenient, by relabeling the time axis if necessary, to assume that we observe $X_{1-d}, X_{2-d}, \ldots, X_n$. (The observed values of $\{Y_t\}$ are then $Y_1, \ldots, Y_n$.) As usual, we shall use P_n to denote best linear prediction in terms of the observations up to time n (in this case $1, X_{1-d}, \ldots, X_n$ or equivalently $1, X_{1-d}, \ldots, X_0, Y_1, \ldots, Y_n$).

Our goal is to compute the best linear predictors $P_n X_{n+h}$. This can be done by applying the operator P_n to each side of (6.4.1) (with $t = n+h$) and using the linearity of P_n to obtain

$$P_n X_{n+h} = P_n Y_{n+h} - \sum_{j=1}^{d} \binom{d}{j} (-1)^j P_n X_{n+h-j}. \tag{6.4.2}$$

27

To find the mean squared error of prediction it is convenient to express $P_n Y_{n+h}$ in terms of $\{X_j\}$. For $n \geq 0$ we denote the one-step predictors by $\hat{Y}_{n+1} = P_n Y_{n+1}$ and $\hat{X}_{n+1} = P_n X_{n+1}$. Then from (6.4.1) and (6.4.2) we have

$$X_{n+1} - \hat{X}_{n+1} = Y_{n+1} - \hat{Y}_{n+1}, \quad n \geq 1,$$

and hence from (3.3.12), if $n > m = \max(p, q)$ and $h \geq 1$, we can write

$$P_n Y_{n+h} = \sum_{i=1}^{p} \phi_i P_n Y_{n+h-i} + \sum_{j=h}^{q} \theta_{n+h-1,j} \left(X_{n+h-j} - \hat{X}_{n+h-j} \right). \tag{6.4.3}$$

Setting $\phi^*(z) = (1 - z)^d \phi(z) = 1 - \phi_1^* z - \cdots - \phi_{p+d}^* z^{p+d}$, we find from (6.4.2) and (6.4.3) that

$$P_n X_{n+h} = \sum_{j=1}^{p+d} \phi_j^* P_n X_{n+h-j} + \sum_{j=h}^{q} \theta_{n+h-1,j} \left(X_{n+h-j} - \hat{X}_{n+h-j} \right), \tag{6.4.4}$$

28

6.4.1 The Forecast Function

Inspection of equation (6.4.4) shows that for fixed $n > m = \max(p, q)$, the h-step predictors

$$g(h) := P_n X_{n+h},$$

satisfy the homogeneous linear difference equations

$$g(h) - \phi_1^* g(h-1) - \cdots - \phi_{p+d}^* g(h-p-d) = 0, \quad h > q, \tag{6.4.7}$$

where $\phi_1^*, \ldots, \phi_{p+d}^*$ are the coefficients of $z, \ldots, z^{p+d}$ in

$$\phi^*(z) = (1-z)^d \phi(z).$$

The solution of (6.4.7) is well known from the theory of linear difference equations (see TSTM, Section 3.6). If we assume that the zeros of $\phi(z)$ (denoted by $\xi_1, \ldots, \xi_p$) are all distinct, then the solution is

$$g(h) = a_0 + a_1 h + \cdots + a_d h^{d-1} + b_1 \xi_1^{-h} + \cdots + b_p \xi_p^{-h}, \quad h > q - p - d, \tag{6.4.8}$$

where the coefficients $a_1, \ldots, a_d$ and $b_1, \ldots, b_p$ can be determined from the $p + d$ equations obtained by equating the right-hand side of (6.4.8) for $q - p - d < h \leq q$ with the corresponding value of $g(h)$ computed numerically (for $h \leq 0$, $P_n X_{n+h} = X_{n+h}$, and for $1 \leq h \leq q$, $P_n X_{n+h}$ can be computed from (6.4.4) as already described). Once the constants a_i and b_i have been evaluated, the algebraic expression (6.4.8) gives the predictors for all $h > q - p - d$. In the case $q = 0$, the values of $g(h)$ in the equations for $a_0, \ldots, a_d, b_1, \ldots, b_p$ are simply the *observed* values $g(h) = X_{n+h}$, $-p - d \leq h \leq 0$, and the expression (6.4.6) for the mean squared error is exact.

6.5 Seasonal ARIMA models

If d and D are nonnegative integers, then $\{X_t\}$ is a **seasonal ARIMA$(p,d,q) \times (P,D,Q)_s$ process with period** s if the differenced series $Y_t = (1-B)^d(1-B^s)^D X_t$ is a causal ARMA process defined by

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)Z_t, \qquad \{Z_t\} \sim \mathrm{WN}\left(0,\sigma^2\right), \tag{6.5.1}$$

where $\phi(z) = 1-\phi_1 z - \cdots - \phi_p z^p$, $\Phi(z) = 1-\Phi_1 z - \cdots - \Phi_P z^P$, $\theta(z) = 1+\theta_1 z + \cdots + \theta_q z^q$, and $\Theta(z) = 1+\Theta_1 z + \cdots + \Theta_Q z^Q$.

Notes: Typically D is rarely more than 1. Also P and Q are often quite small (1 or 2)

Suppose we have r years of monthly data, which we tabulate as follows:

Year/Month	1	2	...	12
1	Y_1	Y_2	...	Y_{12}
2	Y_{13}	Y_{14}	...	Y_{24}
3	Y_{25}	Y_{26}	...	Y_{36}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	
r	$Y_{1+12(r-1)}$	$Y_{2+12(r-1)}$	...	$Y_{12+12(r-1)}$

Each column is a (between year) time series and is assumed to follow the same ARMA(P,Q) model, i.e., for each j,

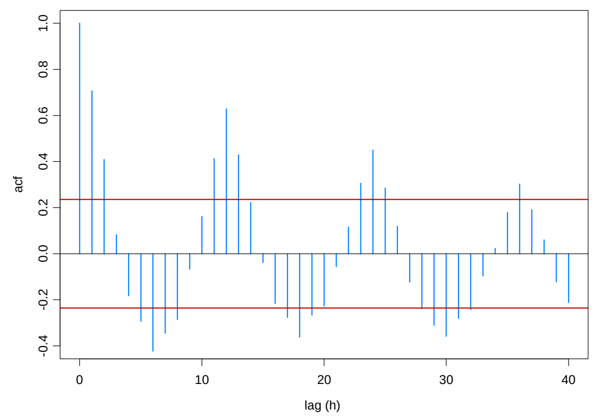
$$\begin{aligned} Y_{j+12t} = {} & \Phi_1 Y_{j+12(t-1)} + \cdots + \Phi_P Y_{j+12(t-P)} + U_{j+12t} \\ & + \Theta_1 U_{j+12(t-1)} + \cdots + \Theta_Q U_{j+12(t-Q)}, \end{aligned}$$

where

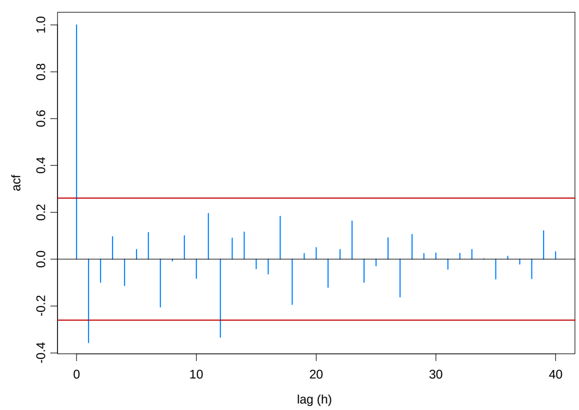
$$\{U_{j+12t}, t = \ldots,-1,0,1,\ldots\} \sim \mathrm{WN}\left(0,\sigma_U^2\right).$$

The ARMA(p,q) is used to model the within year dependence.

Example 6.5.4 (Monthly Accidental Deaths)



ACF of the deaths data after differencing at lags d=1, and D=1.



Plot suggests q=1, Q=1. Model for the differences is then

$$Y_t = 28.831 + (1+\theta_1 B)(1+\Theta_1 B^{12})\, Z_t$$

After differencing the data at lags 1 and 12 and subtracting the mean in ITSM, select Model>Specify. In the dialog box, enter a model of order 13 with θ_1=-.3, θ_{12}=-.3, and θ_{13}=.09. This corresponds to the model

$$Y_t = (1-.3B)(1-.3B^{12})\, Z_t.$$

Then choose Model>Estimation>max likelihood and click on the button Constrain Optimization. Specification the number of multiplicative relations (one in this case) and define the relationship by entering 1, 12, 13 to indicate

$$\theta_1\, \theta_{12} = \theta_{13}$$

Click OK and return to the maximum likelihood dialog box for optimization. The fitted model is

$$\nabla\nabla_{12} X_t = 28.831 + (1-.478B)(1-.591B^{12})\, Z_t, \quad \{Z_t\}\sim \mathrm{WN}(0{,}942506)$$

35

6.6 Regression with ARMA Errors

Model :

$$Y_t = x_{t1}\beta_1 + \cdots + x_{tk}\beta_k + W_t\,, \quad t=1, \ldots, n,$$

or in matrix notation,

$$\mathbf{Y} = \mathrm{X}\,\beta + \mathbf{W},$$

where

$\mathbf{Y} = (Y_1, \ldots, Y_n)'$ is the response vector,

X is the n x k design matrix of explanatory variables,

$\beta = (\beta_1, \ldots, \beta_k)'$ is the parameter vector and

$\mathbf{W} = (W_1, \ldots, W_n)'$ are observations from the ARMA process

$$\phi(B)\, W_t = \theta(B)\, Z_t\,, \quad \{Z_t\} \sim \mathrm{WN}(0,\sigma^2)$$

36

Examples of explanatory variables:

(i) quadratic trend: $x_{t1} = 1,\ x_{t2} = t,\ x_{t3} = t^2.$

(ii) sinusoidal trend: $x_{t1} = 1,\ x_{t2} = \cos(\omega t),\ x_{t3} = \sin(\omega t).$

(iii) leading indicator

Ordinary Least Squares (OLS) Estimation:

Minimize $\sum_{t=1}^{n} (Y_t - \beta' \mathbf{x}_t)^2$, wrt β.

Soln is : $\hat{\beta}_{OLS} = (X'X)^{-1}X'\mathbf{Y}$

Properties of the OLS estimator $\hat{\beta}_{OLS} = (X'X)^{-1}X'\mathbf{Y}$.

(i) $E\,\hat{\beta}_{OLS} = \beta$ (unbiased)

(ii) $\mathrm{Cov}(\hat{\beta}_{OLS}) = (X'X)^{-1}X'\Gamma_n X(X'X)^{-1}$,

$\Gamma_n = E(\mathbf{W}\mathbf{W}')$.

Generalized Least Squares (GLS) Estimation:

Minimize $(\mathbf{Y} - X\beta)'\,\Gamma_n^{-1}(\mathbf{Y} - X\beta)$, wrt β.

Soln is : $\hat{\beta}_{GLS} = (X'\,\Gamma_n^{-1}\,X)^{-1}X'\,\Gamma_n^{-1}\,\mathbf{Y}$

Properties of the GLS estimator $\hat{\beta}_{GLS}$:

(i) $E\,\hat{\beta}_{GLS} = \beta$ (unbiased)

(ii) $\mathrm{Cov}(\hat{\beta}_{GLS}) = (X'\Gamma_n^{-1}X')^{-1}$.

(iii) $\hat{\beta}_{GLS}$ is the best linear unbiased estimator of β

(iv) $\mathrm{Var}(\mathbf{c}'\,\hat{\beta}_{GLS}) \leq \mathrm{Var}(\mathbf{c}'\,\hat{\beta}_{OLS})$ for all vectors **c.**

39

Maximum likelihood estimation:

$L(\beta, \phi, \theta, \sigma^2) = (2\pi)^{-n/2}|\Gamma_n|^{-1/2}\exp\{-.5\,(\mathbf{Y}-X\beta)'\,\Gamma_n^{-1}(\mathbf{Y}-X\beta)\,\}$.

For fixed ϕ, θ, σ^2, likelihood is maximized by taking

$\hat{\beta} = \hat{\beta}_{GLS}(\phi, \theta)$.

Iterative computation of ML estimates:

Step 0. Set $\hat{\beta}_0 = \hat{\beta}_{OLS}$, put $\hat{W}_t = Y_t - \hat{\beta}_0'\mathbf{x}_t$, t=1, ..., n.

Let $\hat{\phi}_0$, $\hat{\theta}_0$ be MLE based on fitting ARMA to $\hat{W}_t$.

Step i. Calculate $\hat{\beta}_i = \hat{\beta}_{GLS}(\hat{\phi}_{i-1}, \hat{\theta}_{i-1})$, put

$\hat{W}_t = Y_t - \hat{\beta}_i'\mathbf{x}_t$ and let $\hat{\phi}_i$, $\hat{\theta}_i$ be MLE based on fitting ARMA to $\hat{W}_t$.

40

Ex 6.6.1 (The Overshort data). Model is

$$Y_t = \beta + W_t\,, \quad W_t = Z_t + \theta\, Z_{t-1}, \quad \{Z_t\} \sim \mathrm{WN}(0,\sigma^2)$$

OLS estimate: $\hat{\beta}_{OLS} = \overline{Y}_{57} = -4.035.$

Ignoring the dependence in the data, the variance of the estimate is

$$\hat{\gamma}(0)/57 = 59.92.$$

Fitted MA(1) Model: $W_t = Z_t - .818\, Z_{t-1},$

$$\{Z_t\} \sim \mathrm{WN}(0,2040.77)$$

Assuming this is the true model, we have

$$\mathrm{Var}(\hat{\beta}_{OLS}) = 2.214 << 59.92.$$

GLS estimate (under the fitted model):

$$\hat{\beta}_{GLS} = -4.75, \quad \mathrm{Var}(\hat{\beta}_{GLS}) = 1.408.$$

95% Confidence Interval for β.

$$-4.75 + (1.96)(1.408)^{.5} = (-7.07,-2.43)$$

(C.I. suggests storage tank is leaking--same conclusion would not have been reached w/o taking into account dependence in data.)

Iterative Estimates:

Iteration	$\hat{\beta}_i$	$\hat{\theta}_i$
0	-4.035	-.818
1	-4.745	-.847
2	-4.779	-.849
3	-4.782	-.849

Ex 6.6.2 (The Lake Data). Model is

$$Y_t = \beta_0 + \beta_1 t + W_t, \quad W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + Z_t,$$

Assuming the fitted AR(2) model,

$$W_t = 1.008 W_{t-1} - .295 W_{t-2} + Z_t, \ \{Z_t\} \sim \mathrm{WN}(0,.451)$$

is the true model,

$$\hat{\beta}_{OLS} = (10.202, -.0242)' \qquad \hat{\beta}_{GLS} = (10.091, -.0216)'$$

with estimated covariance matrices,

$$\begin{bmatrix} .222 & -.003 \\ -.003 & .00007 \end{bmatrix} \text{ and } \begin{bmatrix} .214 & -.003 \\ -.003 & .00006 \end{bmatrix}$$

43

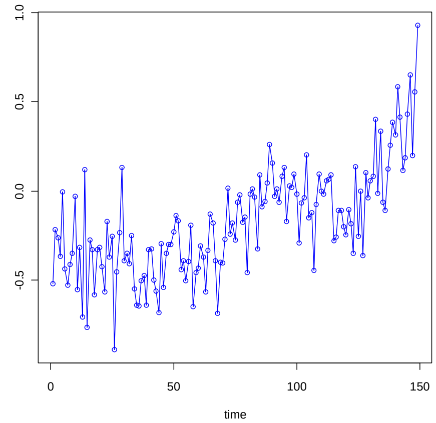
Assuming the data are independent, the estimated covariance matrix for the OLS estimate is

$$\begin{bmatrix} .07203 & -.001 \\ -.001 & .00002 \end{bmatrix}$$

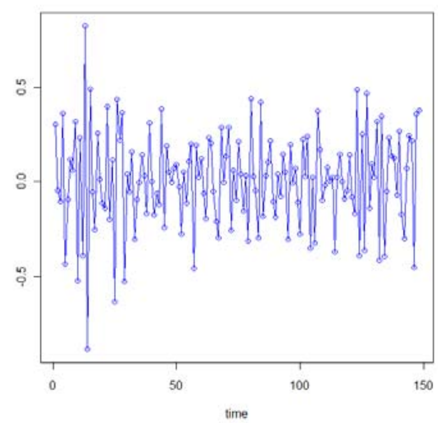
The estimated variances are nearly 3 times smaller when dependence in the data is ignored.

44

Example (Temperature data: average temperature in northern hemisphere, 1850-1999)



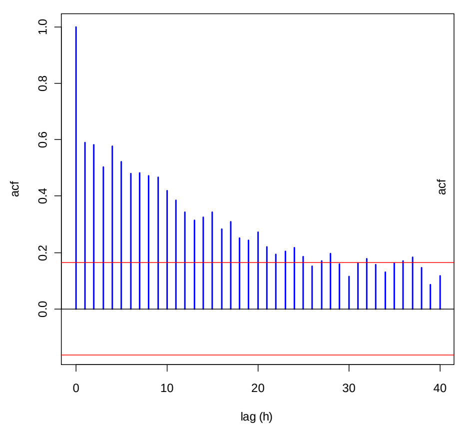
Temperature



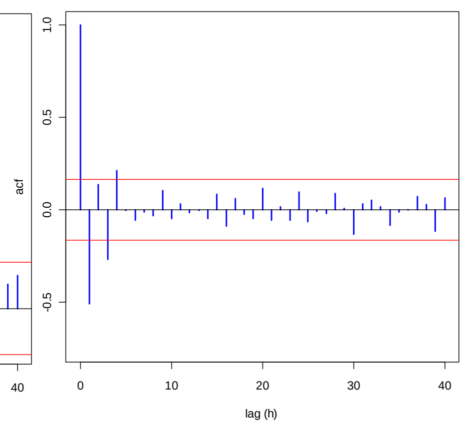
Differenced temp

45

Example (Temperature data: average temperature in northern hemisphere, 1850-1999)



ACF Temperature



ACF Differenced temp

46

1209 temperature proxies: Choose the principal components with the 10 largest eigenvalues. The matrix “covariate” is the 149 by 1209 matrix of covariates (mean corrected). Here is the R-code to extract the principal components.

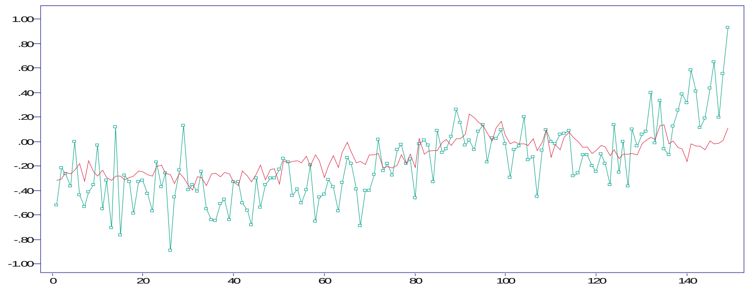
```
eig <- eigen(cov(covariate))
covariate <- as.matrix(covariate)
pp <- covariate%*%eig$vector
junk=lm(temperature~pp[,1:10])
```

Coeff	Value	Std Error
0	-.17545638	.01795629
1	.00001608	.00000225
2	.00000529	.00000462
3	.00003906	.00000705
4	-.00003211	.00000991
5	-.00002932	.00001061
6	.00003181	.00001085
7	-.00004351	.00001195
8	.00003576	.00001241
9	-.00000029	.00001339
10	-.00000266	.00001442

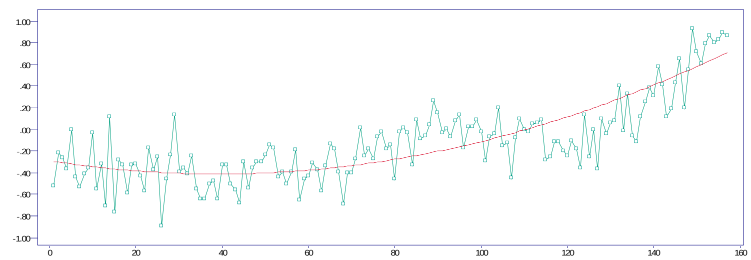
MLE: After fitting a high order AR model, updating the regression estimates and then choosing Auto (ARMA fitting), we get an ARMA(3,3) model and final GLS estimates:

Coeff	Value	Std Error
0	-.06882710	.15552852
1	.00000780	.00000284
2	.00000613	.00000478
3	.00001633	.00000787
4	-.00001410	.00001429
5	-.00001720	.00001803
6	.00003063	.00001756
7	-.00003393	.00001314
8	.00002733	.00001810
9	.00000066	.00001796
10	.00002262	.00001311

Plot of regression function



with quadratic



49

Section 10.2 and Example 6.6.3 Application to Intervention Analysis

During the period for which a time series is observed, it is sometimes the case that a change occurs (e.g. tax laws, reporting methods, etc.) which affects the level of the series.

Simple Intervention Model :

$$Y_t = wX_t + W_t\,, \quad t=1, \ldots, n,$$

where X_t is a deterministic function of time.

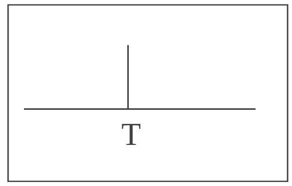
3.16 Application to Intervention Analysis

50

Examples of intervention models:

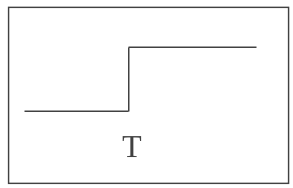
Pulse at fixed time T: momentary effect on the level of the series.

$$X_t = \begin{cases} 1, & \text{if } t = T, \\ 0, & \text{if } t \neq T. \end{cases}$$



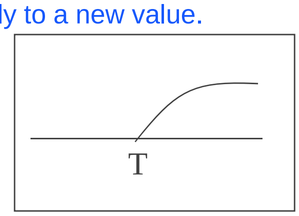
Step change at fixed time T: change in level occurs after a time T.

$$X_t = \begin{cases} 1, & \text{if } t \geq T, \\ 0, & \text{if } t < T. \end{cases}$$



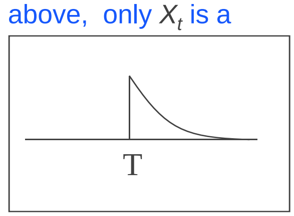
Delayed response: level changes gradually to a new value.

$$Y_t = w(X_t + v X_{t-1} + v^2 X_{t-2} + v^3 X_{t-3} + \ldots) + W_t ,$$



where X_t is the step function described in previous example.

Decayed response: immediate change in level which then decays back to the original value. Model same as above, only X_t is a pulse at time t=T.



Ex 6.6.3 (Seat belt legislation (SBL.TSM).

This data is Y_t, t=1,…,120, representing the number of monthly deaths and serious injuries for 10 year beginning in January 1975. Seatbelt legislation was introduced in February 1983 in the hope of reducing the mean number of "monthly deaths and serious injuries." We consider the regression,

$$Y_t = a + b\, f(t) + W_t\ ,\ t = 1, \ldots, 120,$$

where f(t)=0 for t=1,…,98, and f(t)=1 for t=99,…,120.

- Open the file sbl.tsm and select Regression > Specify, click on the Include auxiliary variables imported from file box and click on the browse button. Open the file called sblin.tsm and indicate that you have only 1 regressor. Make sure you **check** the intercept and **uncheck** polynomial buttons before clicking OK on the dialog box.

- Now click on the GLS icon and inspect the plots of the current data set and its ACF/PACF.
- These plots clearly suggest a strong seasonal component with period 12. Difference the data and consider

$$X_t = Y_t - Y_{t-12},$$
$$= b\, g(t) + N_t\ ,\ t = 13, \ldots, 120,$$

where g(t)=1 for t=99,…,110, and 0 otherwise, $N_t = W_t - W_{t-12}$ is modeled as an ARMA process.

- Open the file sbld.tsm and select Regression > Specify, click on the Include auxiliary variables imported from file box and click on the browse button. Open the file called sbldin.tsm and indicate that you have only 1 regressor.
- Press the blue GLS button

- Sample ACF/PACF of residuals suggests MA(13) or AR(13). Use autofit with AR and MA orders up to 13.

- Update gls estimates (click on blue GLS button). Press the MLE button and repeat. Final model is

$$X_t = -328.45\, g(t) + N_t\ ,$$

where N_t is an MA(13) process.

55

Example (Car deaths and injuries (CDID.TSM).

This data is $Y_t = (1\text{-}B^4)D_t - 305.2$, where $\{D_t\}$ is the number of car drivers killed or seriously injured (quarterly) in Great Britain in 1969-1984. There is a potential intervention at the 16th value of the data corresponding to the imposition of the Arab oil embargo in November 1973. Model this data as follows:

- Start a univariate project by reading in the data cdid.tsm

- Read in the intervention function using the option (Regression > Specify), click on the Include auxiliary variables imported from file box and click on the browse button. Open the file called cdin.tsm and indicate that you have only 1 regressor. Make sure you **uncheck** the intercept and polynomial buttons before clicking OK on the dialog box.

56

- Now click on the GLS icon and inspect the plots of the current data set and its ACF/PACF.
- These plots suggest modeling the noise as an MA(4). Click on the Aut button (click OK on the warning message).
- Re-estimate the MA(4) model by clicking on the MLE icon (ignore the warning again).
- Re-estimate the regression coefficients by clicking on the MLE icon and repeat until the parameter estimates no longer change. Inspect the final residuals for whiteness.
- Is the regression coefficient negative? How does one interpret the result.

57