Chapter 3

The ARIMA Process and Testing for Unit Roots

3.1 The ARIMA Process

The ARIMA process is defined as follows:

Definition 3.1 (The ARIMA \((p,d,q)\) process). Let \(d\) be a non-negative integer. The process \(\{X_t, \ t \in \mathbb{Z}\}\) is said to be an ARIMA \((p,d,q)\) process if \(\nabla^d X_t\) is a causal ARMA \((p,q)\) process.

A causal ARIMA \((p,d,q)\) process \(\{X_t\}\) satisfies:

\[
\phi(B)X_t = \phi^*(B)(1-B)^dX_t = \theta(B)Z_t, \quad \{Z_t\} \sim \text{WN}(0,\sigma^2), \tag{3.1}
\]

where \(\phi^*(z) \neq 0\) for all \(|z| \leq 1\). The process \(Y_t := \nabla^d X_t = (I - B)^d X_t\) satisfies:

\[
\phi^*(B)Y_t = \theta(B)Z_t.
\]

Example 3.1 (Random Walk).

Consider the simple random walk process:

\[X_t = X_{t-1} + Z_t \quad \{Z_t\} \sim \text{WN}(0,\sigma^2) \quad 0 < \sigma^2 < +\infty.\]

This is not a stationary process; \(\mathbf{V}(X_t) = t\sigma^2 \xrightarrow{t \to +\infty} +\infty\); the central limit theorem gives that \(X_t \xrightarrow{t \to +\infty} (d) N(0,\sigma^2)\). A stationary process may be obtained from \(X\) by differencing; let

\[Y_t = \nabla X_t = X_t - X_{t-1} = (I - B)X_t.\]

then \(Y_t\) is a stationary process;

\[Y_t = Z_t \sim \text{WN}(0,\sigma^2).\]

It follows that the random walk \(\{X_t : t \in \mathbb{Z}_+\}\) is an ARIMA\((0,1,0)\) process. \(\square\)
It is clear that, for $d \geq 1$ there are no stationary solutions of Equation (3.1). Furthermore, neither the mean nor ACVF of $\{X_t\}$ are determined by (3.1), since any process $X_t + Y_t$, where $Y_t$ disappears by differencing $d$ times, satisfies equation (3.1). For example, if $Y$ is a random variable, then $\nabla(X_t + Y) = \nabla X_t$.

For $|\phi| < 1$, the process

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

is a causal AR(1) process and is stationary, while for $\phi = 1$, the process is not stationary, but is an ARIMA(0,1,0) process.

Recall that a causal AR(1) process has autocorrelation function

$$\rho(h) = \phi^{|h|}, \quad |\phi| < 1.$$ 

and hence, for any $h$,

$$\lim_{|\phi| \uparrow 1} |\rho(h)| = 1.$$ 

Similarly it holds for any ARMA process that its ACVF decreases slowly if some of the roots of $\phi(z) = 0$ are near the unit circle. From a sample of finite length, it is difficult to distinguish between an ARIMA($p,1,q$) process and an ARMA($p+1,q$) where $\phi(z)$ has a root near the unit circle. An estimated ACVF that decreases slowly indicates that differencing may be advisable.

Suppose that $\{X_t\}$ is a causal and invertible ARMA($p,q$) process:

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

where $\theta(z) \neq 0$ for all $|z| \leq 1$ and $\phi(z)$ has no roots in the unit circle. Then

$$\phi(B)\nabla X_t = \phi(B)(1 - B)X_t = \theta(B)(1 - B)Z_t, \quad \{Z_t\} \sim \text{WN}(0, \sigma^2),$$

from which it follows that $\nabla X_t$ is a causal, but non-invertible ARMA($p,q+1$) process. A unit root in the moving average polynomial indicates that $X_t$ has been overdifferenced.

### 3.1.1 The Dickey Fuller Test

For given time series data, there are tests available to indicate whether or not there are unit roots present. One common test is the Dickey Fuller test, introduced by Dickey and Fuller [3](1979), which has been refined to produce the Augmented Dickey Fuller Test (abbreviated to ADF).

**Dickey Fuller** Consider the AR(1) model:

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

Subtracting $X_{t-1}$ from both sides gives:
\[ \nabla X_t = (\phi - 1) X_{t-1} + Z_t \Rightarrow \nabla X_t = \beta X_{t-1} + Z_t \quad \{Z_t\} \sim \text{WN}(0, \sigma^2). \]

The Dickey Fuller test simply takes a linear regression of \(\{\nabla X_t\}\) against \(X_{t-1}\) and estimates the parameter \(\beta\) in the model, with error bounds. The test may also include a constant, and a deterministic drift; using linear regression, assuming \(\{Z_t\} \sim \text{IID}N(0, \sigma^2)\), one tests whether the parameter \(\beta\) is significant in either

\[ \nabla X_t = \alpha + \beta X_{t-1} + Z_t \]

or

\[ \nabla X_t = \alpha_0 + \alpha_1 t + \beta X_{t-1} + Z_t. \]

While standard multiple linear regression techniques may be used, the approach by Dickey and Fuller represents a refinement where the estimates are made in a different way and the distribution of the test statistic \(DF_r := \frac{\hat{\beta}}{\text{sd}(\hat{\beta})}\) turns out not to be exactly \(t\) distributed. The distribution is known as the Dickey–Fuller distribution.

The tests have low statistical power; they cannot distinguish between a true unit-root (\(\beta = 0\)) and near unit-root (\(\beta\) close to zero). This is called the ‘near observation equivalence’ problem.

**The Augmented Dickey Fuller Test**  The testing procedure for the ADF test is the same as for the Dickey–Fuller test but it is applied to the model

\[ \nabla X_t = \alpha_0 + \alpha_1 t + \beta X_{t-1} + \delta_1 \nabla X_{t-1} + \cdots + \delta_{p-1} \nabla X_{t-p+1} + Z_t \quad \{Z_t\} \sim \text{IID}N(0, \sigma^2) \]

The lag length \(p\) is determined when applying the test, using standard model building techniques from multiple linear regression analysis. The unit root test is then carried out under the null hypothesis \(\beta = 0\) against the alternative hypothesis \(\beta < 0\). The test statistic

\[ DF_r = \frac{\hat{\beta}}{\text{sd}(\hat{\beta})} \]

is computed it can be compared to the relevant critical value for the Dickey–Fuller test.

**Testing for unit roots using R**  The following gives a demonstration of a unit root test. Consider the log series of U.S. quarterly GDP from 1947.I to 2008.IV. The file is found in q-gdp4708.txt in the course directory. The data is plotted in Figure 3.1.

The following indicates that the unit root test cannot be rejected. The test used is the Dickey-Fuller test.

```r
> install.packages("fUnitRoots")
> library(fUnitRoots)
> q.gdp4708 <- read.table("~/data/q-gdp4708.txt", header=T, quote="\"")
> View(q.gdp4708)
```

53
Figure 3.1: US quarterly GDP 1947 - 2008

\[ \text{da} \leftarrow \text{q.gdp4708} \]
\[ \text{gdp} = \log(\text{da}[4]) \]
\[ \text{m1} = \text{ar(diff(gdp), method = 'mle')} \]
\[ \text{m1}$\text{order} \]
[1] 10
\[ \text{adfTest(gdp, lags=10, type=c("c"))} \]

Title:
Augmented Dickey-Fuller Test

Test Results:
PARAMETER:
   Lag Order: 10
STATISTIC:
   Dickey-Fuller: -1.6109
   P VALUE:
   0.4569

With a \( p \) value of 0.4569, the null hypothesis (of no unit root) is not rejected; in fact, it looks likely that this data set has a unit root. In fact, the \textit{auto.arima} function to find the best fitting ARIMA model suggests that the unit root is of order 2.

\[ \text{library("forecast")} \]
\[ \text{auto.arima(q.gdp4708$gdp)} \]
Series: q.gdp4708$gdp
ARIMA(0,2,1)
Coefficients:

\[ ma1 = -0.6438 \]

\[ \text{s.e.} \quad 0.0685 \]

\[ \sigma^2 \text{ estimated as } 1361: \log \text{ likelihood} = -1236.89 \]

\[ \text{AIC} = 2477.79 \quad \text{AICc} = 2477.84 \quad \text{BIC} = 2484.8 \]

3.2 SARIMA Processes

Seasonal series are characterised by a strong serial correlation at the seasonal lag and multiples thereof. Seasonal ARIMA models allow for randomness in the seasonal pattern from one cycle to the next.

**Definition 3.2** (The SARIMA \((p,d,q) \times (P,D,Q)_s\) Process). A process \( \{X_t\} \) is said to be a Seasonal ARIMA \((p,d,q) \times (P,D,Q)\) process with period \(s\) if the differenced process

\[ Y_t := (1 - B)^d (1 - B^s)^D X_t \]

is a causal ARMA process,

\[ \phi(B) \Phi(B^s) Y_t = \theta(B) \Theta(B^s) Z_t \quad \{Z_t\} \sim WN(0, \sigma^2) \]

where

\[ \phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p \]

\[ \Phi(z) = 1 - \Phi_1 z^s - \ldots - \Phi_P z^{Ps} \]

\[ \theta(z) = 1 + \theta_1 z + \ldots + \theta_q z^q \]

\[ \Theta(z) = 1 + \Theta_1 z^s + \ldots + \Theta_Q z^{Qs}. \]

Note that the process \( \{Y_t\} \) is causal if and only if both \( \phi(z) \neq 0 \) and \( \Phi(z) \neq 0 \) for all \(|z| \leq 1\).
Chapter 4

Estimation of the mean, autocovariance and autocorrelation

Let \( \{X_t\} \) be a stationary time series with mean \( \mu \), autocovariance function \( \gamma(\cdot) \), and, when it exists, spectral density \( f(\cdot) \). Now consider the problem of estimating the mean \( \mu \), ACVF \( \gamma(\cdot) \) and ACF \( \rho(\cdot) = \frac{\gamma(\cdot)}{\gamma(0)} \) from observations of \( X_1, X_2, \ldots, X_n \).

4.1 Asymptotic Normality

For a large class of strictly linear time series, estimators of the mean \( \mu \), ACVF \( \gamma(\cdot) \) and ACF \( \rho(\cdot) \) will satisfy a central limit theorem. With this in view, asymptotic normality is defined.

**Definition 4.1 (Asymptotic Normality).** Let \( Y_1, Y_2, \ldots \) be a sequence of random variables. They are said to be asymptotically normal, written \( Y_n \sim AN(\mu_n, \sigma^2_n) \) if and only if \( \mu_n = \mathbb{E}[Y_n] \) and \( \sigma^2_n = \text{V}(Y_n) \) for each \( n \) and

\[
\lim_{n \to \infty} \mathbb{P} \left( \frac{Y_n - \mu_n}{\sigma_n} \leq x \right) = \Phi(x),
\]

where \( \Phi(x) = \mathbb{P}(Z \leq x) \), \( Z \sim N(0, 1) \).

Let \( Y_1, Y_2, \ldots \) be a sequence of random \( k \)-vectors. The sequence is said to be asymptotically normal, written \( Y_n \sim AN(\mu_n, \Sigma_n) \) if and only if \( \mu_n = \mathbb{E}[Y_n] \) and \( \Sigma_{n,i,j} = \text{C}(Y_{n,i}, Y_{n,j}) \) (\( \Sigma_n \) is the covariance matrix of \( Y_n \) for each \( n \)) and

\[
\Lambda'Y_n \sim AN(\Lambda'\mu_n, \Lambda'\Sigma_n\Lambda) \quad \forall \Lambda \in \mathbb{R}^k.
\]

4.2 Estimation of \( \mu \)

The estimator \( \bar{X}_n := \frac{1}{n} \sum_{j=1}^{n} X_j \) is the natural unbiased estimate of \( \mu \).

**Theorem 4.2.** Let \( \{X_t\} \) be a stationary time series with mean \( \mu \) and autocovariance function \( \gamma(\cdot) \) which satisfies \( \lim_{h \to +\infty} |\gamma(h)| = 0 \). Then
Suppose $\sum_{h=-\infty}^{\infty} |\gamma(h)| < +\infty$ and let $f$ denote the spectral density. Then

$$nV(X_n) \xrightarrow{n \to +\infty} \sum_{h=-\infty}^{\infty} \gamma(h) = 2\pi f(0)$$

**Proof** For the first statement,

$$nV(X_n) = \frac{1}{n} \sum_{i,j=1}^{n} C(X_i, X_j) = \sum_{|h|<n} \left(1 - \frac{|h|}{n}\right) \gamma(h) \leq \sum_{|h|<n} |\gamma(h)|.$$  

Since $|\gamma(h)| \xrightarrow{|h| \to +\infty} 0$, it follows that

$$\frac{1}{n} \sum_{|h|<n} \left(1 - \frac{|h|}{n}\right) \gamma(h) \xrightarrow{n \to +\infty} 0,$$

from which it follows that $V(X_n) \xrightarrow{n \to +\infty} 0$.

For the second statement, suppose that $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$. Then it follows firstly, that for any $N$

$$\sum_{|h|\leq N} \left(1 - \frac{|h|}{n}\right) \gamma(h) \xrightarrow{n \to +\infty} \sum_{|h|\leq N} \gamma(h),$$

and secondly that

$$\left| \sum_{|h|\geq N+1} \left(1 - \frac{|h|}{n}\right) \gamma(h) \right| \leq \sum_{|h|\geq N+1} |\gamma(h)| \xrightarrow{N \to +\infty} 0,$$

from which it follows that

$$nV(X_n) \xrightarrow{n \to +\infty} \sum_{h=-\infty}^{\infty} \gamma(h) = 2\pi f(0).$$

**Theorem 4.3** (Central Limit Theorem for Strictly Stationary $m$-Dependent Sequences). Let $\{X_t\}$ be a strictly stationary $m$-dependent sequence of random variables. That is, $X_t \perp X_s$ for all $s$ such that $|t-s| > m$. Let $\mu = 0$ and let $\gamma$ denote the autocovariance function. Let $v_m = \gamma(0) + 2 \sum_{j=1}^{m} \gamma(j)$ and suppose that $v_m \neq 0$. Then

1. $\lim_{n \to +\infty} nV(X_n) = v_m$ and
2. $X_n \sim AN(0, \frac{v_m}{n}).$
Proof The first statement has been dealt with in the previous theorem. For the second statement, for each integer \( k \) such that \( k > 2m \), let

\[
Y_{nk} = \frac{1}{n^{1/2}} \left\{ (X_1 + \ldots + X_{k-m}) + \ldots + (X_{(r-1)k+1} + \ldots + X_{rk-m}) \right\}
\]

where \( r = \left\lfloor \frac{n}{k} \right\rfloor \). It follows that \( n^{1/2}Y_{nk} \) is the sum of \( r \) i.i.d. random variables, each with mean zero and variance:

\[
R_{k-m} = \text{Var}(X_1 + \ldots + X_{k-m})
\]

\[
= \sum_{i=1}^{k-m} \sum_{j=1}^{k-m} \text{C}(X_i, X_j)
\]

\[
= (k - m)\gamma(0) + 2 \sum_{i=1}^{k-m} \sum_{j=i+1}^{k-m} \gamma(j - i)
\]

\[
= (k - m)\gamma(0) + 2 \sum_{j=1}^{k-m} \gamma(j) \sum_{i=1}^{k-m-j} 1
\]

\[
= \sum_{|j| < k-m} (k - m - |j|)\gamma(j).
\]

It follows from the Central Limit Theorem that

\[
\frac{(n^{1/2}Y_{nk})}{(\lfloor \frac{n}{k} \rfloor R_{k-m})^{1/2}} \xrightarrow{d} N(0, 1)
\]

which may be expressed as

\[
Y_{nk} \xrightarrow{n \to +\infty} N(0, \frac{1}{k} R_{k-m}) \xrightarrow{k \to +\infty} N(0, v_m).
\]

It remains to show that

\[
\lim_{k \to +\infty} \limsup_{n \to +\infty} \mathbb{P}\left( \left| n^{1/2}X_n - Y_{nk} \right| > \epsilon \right) = 0 \quad \forall \epsilon > 0.
\]

To establish this,

\[
(n^{1/2}X_n - Y_{nk}) = \frac{1}{n^{1/2}} \sum_{j=1}^{r-1} (X_{jk-m+1} + \ldots + X_{jk}) + \frac{1}{n^{1/2}} (X_{rk-m+1} + \ldots + X_n).
\]

The terms are independent. It follows that

\[
\text{Var}(n^{1/2}X_n - Y_{nk}) = \frac{1}{n} \left( \left(\lfloor \frac{n}{k} \rfloor \right) - 1 \right) R_m + R_{h(n)}.
\]

\[
h(n) = n - k \left\lfloor \frac{n}{k} \right\rfloor + m \quad 0 \leq h(n) \leq k + m.
\]

It follows that

\[
\limsup_{n \to +\infty} \text{Var}(n^{1/2}X_n - Y_{nk}) = \frac{1}{k} R_m
\]

59
and the result follows by Chebyshev.

While this establishes asymptotic normality for the mean of an MA(q) process for any q, it does not establish asymptotic normality for general ARMA processes; even the AR(1) process requires an infinite number of terms. For AR(1),

\[ X_t = \sum_{j=0}^{\infty} \theta_j Z_{t-j} \quad \{Z_t\} \sim \text{IID}(0, \sigma^2). \]

The following result shows that for any linear time series driven by IID noise, the sample average is asymptotically normal.

**Theorem 4.4.** Let \( \{X_t\} \) be a strictly linear time series defined by

\[ X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j} \quad \{Z_t\} \sim \text{IID}(0, \sigma^2) \]

where \( \{\psi_j\} \) satisfy \( \sum_{j=-\infty}^{\infty} |\psi_j| < \infty \) and \( \sum_{j=-\infty}^{\infty} \psi_j \neq 0 \). The infinite sum is taken in the following sense: let

\[ X_{tm} = \mu + \sum_{j=-m}^{m} \psi_j Z_{t-j} \quad \{Z_t\} \sim \text{IID}(0, \sigma^2) \]

then \( \{X_t\} \) is the stationary time series where for any \( N \geq 1 \),

\[ \{X_{-N}, \ldots, X_N\} \overset{(d)}{=} \lim_{m \to +\infty} \{X_{m,-N}, \ldots, X_{m,N}\}. \]

Then

\[ \frac{\sqrt{n}(\bar{X}_n - \mu)}{v} \sim AN(0, 1) \]

where

\[ v = \sum_{h=-\infty}^{\infty} \gamma(h) = \sigma^2 \left( \sum_{j=-\infty}^{\infty} \psi_j \right)^2 \]

and \( \gamma \) is the ACVF of \( \{X_t\} \).

**Proof** (Omitted)

### 4.3 Estimation of \( \gamma(\cdot) \)

The natural estimates of \( \gamma \) and \( \rho \) are:

\[ \hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (X_t - \bar{X}_n)(X_{t+h} - \bar{X}_n), \quad 0 \leq h \leq n - 1 \]

and

\[ \hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \]
respectively. The estimators are biased, but nevertheless are the estimators used. They are asymptotically unbiased. This normalisation is chosen to ensure that the matrix

\[
\hat{\Gamma}_h = \begin{pmatrix}
\hat{\gamma}(0) & \ldots & \hat{\gamma}(h) \\
\vdots & & \vdots \\
\hat{\gamma}(h) & \ldots & \hat{\gamma}(0)
\end{pmatrix}
\]

is non-negative definite.

In the sequel, let \( \gamma = (\gamma(0), \gamma(1), \ldots, \gamma(h))^t \), with similar notation for estimators of \( \gamma \). That is, \( \hat{\gamma} = (\hat{\gamma}(0), \ldots, \hat{\gamma}(h))^t \) and \( \gamma^* = (\gamma^*(0), \ldots, \gamma^*(h))^t \) where the estimator \( \gamma^*(p) \) is used for a stationary time series where it is known that \( \mu = 0 \) and is defined as \( \gamma^*(p) = \frac{1}{n} \sum_{t=1}^n X_t X_{t+p} \). The following results show that the distribution of \( \hat{\gamma} \) is asymptotically normal. The first deals with the covariance structure of \( \gamma^* \), the following move onto expressing the problem in such a way that the central limit may be used.

**Theorem 4.5.** Let \( \{X_t\} \) be a strictly linear time series with mean 0;

\[
X_t = \sum_{j=\infty}^{\infty} \psi_j Z_{t-j} \quad \{Z_t\} \sim IID(0, \sigma^2)
\]

satisfying \( \sum_{j=-\infty}^{\infty} |\psi_j| < \infty \) and \( E[Z_t^4] = \eta \sigma^4 < \infty \). Let

\[
\gamma^*(h) = \frac{1}{n} \sum_{t=1}^n X_t X_{t+h} \quad h = 0, 1, 2, \ldots
\]

Then

\[
\lim_{n \to +\infty} nC(\gamma^*(p), \gamma^*(q)) = (\eta - 3) \gamma(p) \gamma(q) + \sum_{k=-\infty}^{\infty} (\gamma(k) \gamma(k - p + q) + \gamma(k + q) \gamma(k - p))
\]

**Note** If \( \{Z_t, t \in \mathbb{Z}\} \) is Gaussian, then \( \eta = 3 \).

**Proof** Omitted. \( \square \)

The covariance structure of \( \gamma^* \) has now been established; the following results show asymptotic normality. Firstly, Theorem 4.6 proves the result for a MA(2m + 1) process where m is finite; Theorem 4.8 extends it to strictly linear processes.

**Theorem 4.6.** Let \( \{X_t\} \) be the moving average process

\[
X_t = \sum_{j=-m}^m \psi_j Z_{t-j} \quad \{Z_t\} \sim IID(0, \sigma^2),
\]

where \( E[Z_t^4] = \eta \sigma^2 < +\infty \). Let \( \gamma \) be the autocovariance function of \( X \). Let \( \gamma^* = (\gamma^*(0), \ldots, \gamma^*(h))^t \) and \( \gamma = (\gamma(0), \ldots, \gamma(h))^t \). Then
\[ \gamma^* \sim AN(\gamma, n^{-1}V) \]

where \( V = (v_{pq})_{p,q=0,\ldots,h} \) is the covariance matrix with entries

\[
v_{pq} = \sum_{k=-\infty}^{\infty} T_k = (\eta - 3)\gamma(p)\gamma(q) + \sum_{k=-\infty}^{\infty} (\gamma(k)\gamma(k-p+q) + \gamma(k+q)\gamma(k-p)).
\]

**Proof** This follows directly from the central limit theorem (Theorem 4.4); consider the random \( h+1 \) vectors

\[ Y_t = (X_tX_t, X_tX_{t+1}, \ldots, X_tX_{t+h})^t \]

then \( Y_t \) is a strictly stationary \((2m+h)\) dependent sequence and

\[ \frac{1}{n} \sum_{t=1}^{n} Y_t = (\gamma^*(0), \ldots, \gamma^*(h)). \]

For any linear combination \( \lambda^t \gamma^* \) such that \( \lambda^t V \lambda > 0 \), it follows that \( \{\lambda^t Y_t\} \) satisfies the hypotheses of Theorem 4.4 and hence

\[ \frac{\sqrt{n}(\lambda^t \gamma^* - \lambda^t \gamma)}{\sqrt{\lambda^t V \lambda}} \rightarrow_{(d)} N(0,1) \]

from which the result follows.

Lemma 4.7. Let \( \{X_n, n = 1, 2, \ldots\} \) and \( Y_{nj}, j = 1, 2, \ldots; n = 1, 2, \ldots \) be random \( k \)-vectors such that

1. \( Y_{nj} \rightarrow Y_j \) for each \( j = 1, 2, \ldots \)
2. \( Y_j \rightarrow Y \) as \( j \rightarrow +\infty \) and
3. \( \lim_{j \rightarrow +\infty} \lim_{n \rightarrow +\infty} \mathbb{P} (|X_n - Y_{nj}| > \epsilon) = 0 \) for every \( \epsilon > 0 \).

Then

\[ X_n \rightarrow Y \quad n \rightarrow +\infty. \]

**Proof** Clear from the definitions.

Theorem 4.8. The result of Theorem 4.6 remains true for a process

\[ X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j} \quad \{Z_j\} \sim IID(0, \sigma^2) \]

where \( \sum_{j=-\infty}^{\infty} |\psi_j| < +\infty \) and \( \mathbb{E}[Z_t^4] = \eta \sigma^2 < +\infty. \)
Proof. The proof follows directly by applying Theorem 4.6 to the process

\[ X_{tm} = \sum_{j=-m}^{m} \psi_{j} Z_{t-j}, \quad \{Z_{j}\} \sim IID(0, \sigma^{2}). \]

Let

\[ \gamma_{m}^{*}(p) = \frac{1}{n} \sum_{i=1}^{n} X_{tm} X_{(i+p)m}, \]

then

\[ n^{1/2}(\gamma_{m}^{*} - \gamma_{m}) \to \Sigma_{m} \]

where \( \gamma_{m} \) is the autocovariance function of \( \{X_{tm}\} \) and the vector notation is as in the previous theorem. Then \( \Sigma_{m} \sim N(0, V_{m}) \), where \( V_{m} \) is the covariance matrix from Theorem 4.6. As \( m \to +\infty \), \( V_{m} \to V \).

The proof now follows by Lemma 4.7, provided it can be shown that

\[ \lim_{m \to +\infty} \limsup_{n \to +\infty} \mathbb{P} \left( \left| n^{1/2} \gamma_{m}^{*}(p) - \gamma_{m}(p) - \gamma^{*}(p) + \gamma(p) \right| > \epsilon \right) = 0 \]

for \( p = 0, 1, \ldots, h \). This follows by Chebyshev; the probability is bounded by

\[ \frac{n}{\epsilon^{2}} V \left( \gamma_{m}^{*}(p) - \gamma^{*}(p) \right) = \frac{1}{\epsilon^{2}} \left( n V(\gamma_{m}^{*}(p)) + n V(\gamma^{*}(p)) - 2 n C(\gamma_{m}^{*}(p), \gamma^{*}(p)) \right). \]

Firstly,

\[ \lim_{m \to +\infty} \lim_{n \to +\infty} n V(\gamma_{m}^{*}(p)) = \lim_{n \to +\infty} V(\gamma^{*}(p)) = v_{pp} \]

\[ \lim_{m \to +\infty} \lim_{n \to +\infty} n C(\gamma_{m}^{*}(p), \gamma^{*}(p)) = v_{pp} \]

from which the result follows.

\[ \square \]

Proposition 4.9. Let \( \{X_{i}\} \) be a moving average process satisfying

\[ X_{i} = \sum_{j=-\infty}^{\infty} \psi_{j} Z_{i-j}, \quad \{Z_{i}\} \sim IID(0, \sigma^{2}) \]

where \( \sum_{j=-\infty}^{\infty} |\psi_{j}| < +\infty \) and \( \mathbb{E}[Z_{i}^{4}] = \eta \sigma^{4} < +\infty \). Let \( \gamma \) be the autocovariance function of \( \{X_{i}\} \).

Then for any non negative integer \( h \)

\[ \hat{\gamma} \sim AN \left( \gamma, \frac{1}{n} V \right) \]

where \( V \) is the covariance matrix with entries

\[ v_{pq} = (\eta - 3)\gamma(p)\gamma(q) + \sum_{k=-\infty}^{\infty} (\gamma(k)\gamma(k-p+q) + \gamma(k+q)\gamma(k-p)). \]

Proof. (Direct computation, which is omitted) \[ \square \]
4.4 Estimating the Autocorrelation Function $\rho(.)$

For the ACF $\rho(.)$, $\eta$ disappears.

**Theorem 4.10.** Let $\rho = (\rho(1), \ldots, \rho(h))^t$ and $\hat{\rho} = (\hat{\rho}(1), \ldots, \hat{\rho}(h))^t$. Let $\{X_t\}$ satisfy:

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j Z_j \quad \{Z_t\} \sim IID(0, \sigma^2)$$

where $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ and $E[Z_t^4] < \infty$, then

$$\hat{\rho} \sim AN(\rho, n^{-1}W)$$

where $W = (w_{ij})_{i,j=1,\ldots,h}$ is the covariance matrix whose entries are given by:

$$w_{ij} = \sum_{k=-\infty}^{\infty} \{\rho(k+i)\rho(k+j) + \rho(k-i)\rho(k+j) + 2\rho(i)\rho(j)\rho^2(k) - 2\rho(i)\rho(k)\rho(k+j) - 2\rho(j)\rho(k)\rho(k+i)\}.$$  \hspace{1cm} (4.1)

**Proof** This follows from the Delta Method (see the course Statistics): details omitted. \hfill \Box

The expression (4.1) is called Bartlett’s formula. It may be re-arranged to obtain the more convenient form:

$$w_{ij} = \sum_{k=1}^{\infty} \{\rho(k+i) + \rho(k-i) - 2\rho(i)\rho(k)\} \times \{\rho(k+j) + \rho(k-j) - 2\rho(j)\rho(k)\}.$$  \hspace{1cm} (4.2)

The assumption $E[Z_t^4] < \infty$ is relaxed at the expense of a slightly stronger assumption on the sequence $\{\psi_j\}$.

**Theorem 4.11.** If $\{X_t\}$ is a strictly linear time series where $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ and $\sum_{j=-\infty}^{\infty} \psi_j^2 |j| < \infty$, then

$$\hat{\rho} \sim AN(\rho, n^{-1}W)$$

where $W$ is given by the previous theorem.

**Proof** Omitted. \hfill \Box

Using similar techniques, the asymptotic correlations between the estimators can be established;

$$\lim_{n \to +\infty} \text{Corr} (\hat{\gamma}(i), \hat{\gamma}(j)) = \frac{w_{ij}}{\sqrt{w_{ii}w_{jj}}} \quad \lim_{n \to +\infty} \text{Corr} (\hat{\rho}(i), \hat{\rho}(j))) = \frac{w_{ij}}{\sqrt{w_{ii}w_{jj}}}.$$

\hfill \Box

64
Tutorial 3

Unit Roots, ARIMA, SARIMA

1. The data for this exercise is found in the file d-sp55008.txt. Start with:

   ```
   > d.sp55008 <- read.table("~/data/d-sp55008.txt", header=T, quote="\"")
   > View(d.sp55008)
   > library(fUnitRoots)
   > da=d.sp55008
   > sp5=log(da[,7])
   ```

   and test whether the time series of column sp5 has a unit root.

2. This exercise is connected with the distribution of the estimator of the coefficient \( \phi \) for an AR(1) process, when the process does not have a unit root, but where \( \phi \) is close to 1. Try the following:

   ```
   > #simulate distributions of AR(1) samples
   > T <- 100 # sample size
   > N <- 1000 # repetitions
   > # process \( X(t) = 0 + 0.95 X(t-1) + Z(t) \)
   > rho <- array(,N)
   > for (n in 1:N){y <- arima.sim(n=T, list(ar=0.95));
   rho[n] <-ar.ols(y,aic=FALSE,order.max=1)$ar}
   > plot(density(rho),lwd=2)
   > lines(c(0.95,0.95),c(-1,11),lty=2)
   ```

   This gives a sample of size of 1000 estimated coefficients \( \hat{\phi} \) for an AR(1) process where \( \phi = 0.95 \) is the true value. The graph shows the estimated density function for the estimator \( \hat{\phi} \). For what proportion of samples is \( \hat{\phi} \geq 1 \)?

   Try the Augmented Dickey Fuller test on the samples generated. For what proportion of samples is the hypothesis of a unit root rejected?

   Try the following and, interpret the results.

   ```
   > #consider how the estimate is affected by the standard deviation
   > rho2 <- array(,N)
   > for (n in 1:N){y <- arima.sim(n=T,list(ar=0.95),n.start=1,start.innov=runnorm(1,sd=30));
   rho2[n] <- ar.ols(y,aic=FALSE, order.max = 1)$ar}
   > plot(density(rho2),col="red",lwd=2)
   > lines(density(rho),col="blue",lwd=2)
   > lines(c(0.95,0.95),c(-1,26),lty=2)
   ```
3. In this exercise, we examine the dependence of the AR(1) process on the initial condition. The vector \( e \) is the vector of innovations or noise; it is a random sample from a \( N(0, 1) \) distribution. To this is added (at the beginning of the column) an observation from an independent identically distributed normal to produce \( e_1 \) and an observation from a normal with \( \sigma = 50 \) to produce \( e_2 \). The difference between \( y_1 \) and \( y_2 \) is therefore simply the initial condition; the only difference is the first innovation.

\( y_1[-1] \) removes the first observation and \( y_1[-T] \) removes the last. Since the process is constructed from the equation

\[
X_t = 0.95X_{t-1} + Z_t
\]

where \( Z_t \) are the innovations, regressing \( y_1[-1] \) against \( y_1[-T] \) should result in an estimated regression line approximately \( y = 0.95x \). The plots show the estimated regression line and the line \( y = 0.95x \).

\[
\begin{verbatim}
> #AR(1) with different standard deviations
> e<-rnorm(99,sd=1)
> e1<-c(rnorm(1,sd=1),e)
> e2<-c(rnorm(1,sd=50),e)
> y1<-arima.sim(n=T,list(ar=0.95),innov=e1)
> y2 <- arima.sim(n=T, list(ar=0.95),innov=e2)
> plot(cbind(y1,y2))
> reg1<-lm(y1[-1]~y1[-T])
> plot(y1[-T],y1[-1],xlim=range(y1),ylim=range(y1))
> abline(reg1,lty=2)
> lines(range(y1),0.95*range(y1),lty=1)
> reg2<-lm(y2[-1]~y2[-T])
> plot(y2[-T],y2[-1],xlim=range(y2),ylim=range(y2))
> abline(reg2,lty=2)
> lines(range(y2),0.95*range(y2),lty=1)
\end{verbatim}
\]

Show your results and draw conclusions.

4. **Deterministic Seasonal Behaviour**

**Control** Send me a brief summary of your conclusions (noble@mimuw.edu.pl)

The data set `m-deciles08.txt` contains the monthly simple returns of the CRSP Decile 1 Index from January 1970 to December 2008, for 468 observations. First, make time series plot. Is there any apparent seasonality? What is the sample ACF? What are the significant lags?

A seasonal ARMA model for a process with lags at 1, and \( 12k \) for \( k \geq 1 \), would be:

\[
(1 - \phi_1 B)(1 - \phi_{12} B^{12})X_t = (1 - \theta_{12} B^{12})Z_t \quad \{Z_t\} \sim WN(0, \sigma^2)
\]

66
What are the parameter estimates $\phi_1$, $\phi_{12}$ and $\theta_{12}$ for such a model?

The seasonal behaviour might be deterministic. Try to create a dummy variable

$$J = \begin{cases} 
1 & \text{January} \\
0 & \text{otherwise}
\end{cases}$$

and make a regression

$$X_t = \beta_0 + \beta_1 J_t + \epsilon_t.$$ 

What is the fitted model? Does it fit the data well? What do you conclude about the ‘January effect?’
Exercises

(These should be considered at home)

Unit Roots, ARIMA, SARIMA

1. Let \( \{X_t\} \) be an ARIMA\((p,d,q)\) process satisfying the difference equations
   \[
   \phi(B)(1-B)^d X_t = \theta(B)Z_t \quad \{Z_t\} \sim \text{WN}(0,\sigma^2).
   \]
   Show that these difference equations are also satisfied by the process
   \[
   W_t = X_t + A_0 + A_1 t + \ldots + A_{d-1} t^{d-1}
   \]
   where \(A_0, \ldots, A_{d-1}\) are arbitrary random variables.

2. Determine the ACVF of the process with spectral density
   \[
   f(\lambda) = \frac{1}{\pi^2} \frac{\pi - |\lambda|}{|\lambda|}, \quad \lambda \in [-\pi,\pi]
   \]
   3. Determine the power transfer function (Definition 2.13) of the time invariant linear filter (Definition 2.12) with coefficients
   \[
   c_0 = 1, \quad c_1 = -2\alpha, \quad c_2 = 1 \quad \text{and} \quad c_j = 0 \quad \text{for} \quad j \neq 0, 1, 2.
   \]
   If you want to use the filter to suppress sinusoidal oscillations with period 6, which values of \(\alpha\) should you use?

Estimation of Mean, ACVF, ACF

1. Let \( \{X_t\} \) be a stationary process with mean zero and an absolutely summable ACVF \(\gamma(h)\) such that \(\sum_{h=-\infty}^{\infty} |\gamma(h)| < +\infty\) and \(\sum_{h=-\infty}^{\infty} \gamma(h) = 0\). Show that \(nV(\overline{X}_n) \xrightarrow{n \to +\infty} 0\).

2. Suppose that \(\{X_t\} \sim IID(\mu,\sigma^2)\) where \(0 < \sigma^2 < +\infty\). If \(\overline{X}_n = \frac{1}{n}(X_1 + \ldots + X_n)\) has a probability density function \(f(x)\) which is continuous and positive at \(x = 0\), show that \(E\left[\frac{1}{\overline{X}_n}\right] = +\infty\).
   What is the limiting distribution of \(\frac{1}{\overline{X}_n}\) when \(\mu = 0\)?

3. Let \(\{X_t\}\) be a causal AR(1) process with mean \(\mu\); that is \(Y_t := X_t - \mu\) satisfies
   \[
   Y_t - \phi Y_{t-1} = Z_t \quad \{Z_t\} \sim IID(0,\sigma^2) \quad |\phi| < 1.
   \]
   Show that \(\overline{X}_n \sim AN\left(\mu, \frac{\sigma^2}{(1-\phi)^2n}\right)\).
   In a sample of size \(n = 100\) from an AR(1) process with \(\phi = 0.6\) and \(\sigma^2 = 2\), the sample average is \(\overline{x}_n = 0.271\). Construct an approximate 95% confidence interval for \(\mu\) and determine whether or not 0 is in the interval.

4. Let \(\{X_t\}\) be a stationary process with mean \(\mu\). Show that among unbiased linear estimators of \(\mu\), the one with the least variance is given by
   \[
   \hat{\mu}_n = (1^t \Gamma_n^{-1} 1)^{-1} 1^t \Gamma_n^{-1} \overline{X}_n
   \]
   where \(\Gamma_n\) is the covariance matrix of \(\overline{X} = (X_1, \ldots, X_n)^t\) and \(1 = (1, \ldots, 1)^t\). Show also that \(V(\hat{\mu}_n) = (1^t \Gamma_n 1)^{-1}\).
5. Show that for any series \( \{x_1, \ldots, x_n\} \), the sample autocovariances satisfy \( \sum_{|h| < n} \hat{\gamma}(h) = 0 \).

6. Let \( \{X_t\} \) be an AR(1) process with mean \( \mu \). That is:

\[
(X_t - \mu) = \phi(X_{t-1} - \mu) + Z_t \quad \{Z_t\} \sim IID(0, \sigma^2)
\]

where \(|\phi| < 1\). Find constants \( a_n > 0 \) and \( b_n \) such that \( \exp \{ \bar{X}_n \} \sim AN(b_n, a_n) \).
Short Answers

Unit Roots, ARIMA, SARIMA

1. Straightforward: \((I - B)^q t^q = t^q - (t - 1)^q = -\sum_{j=0}^{q-1} \binom{q}{j} t^j (-1)^{q-j}\) which is a polynomial of degree \(q - 1\), while \((I - B)c = c - c = 0\). It follows that

\[
(I - B)^d \left(A_0 + A_1 t + \ldots + A_d t^{d-1}\right) = (I - B)^d A_0 + A_1 (I - B)^d t + \ldots + A_d (I - B)^d t^{d-1}
\]

and (clearly) \((I - B)^d A_0 = 0, (I - B)^d t^q = 0\) for \(q = 1, \ldots, d - 1\), from which the result follows.

2. \[
\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda = \frac{1}{\pi^2} \left( 2\pi^2 K_0(h) + \int_{-\pi}^{\pi} \lambda e^{ih\lambda} d\lambda - \int_{0}^{\pi} \lambda e^{ih\lambda} d\lambda \right)
\]

\[
= 2K_0(h) - \frac{2}{\pi^2} \int_{0}^{\pi} \lambda \cos(h\lambda) d\lambda
\]

\[
u = \lambda, \quad u' = 1, \quad v' = \cos(h\lambda), \quad v = \frac{1}{h} \sin(h\lambda)
\]

\[
\gamma(0) = 2
\]

For \(h \neq 0\),

\[
\gamma(h) = \frac{2}{\pi^2} \left\{ \frac{\lambda}{h} \sin(h\lambda) \right\} + \frac{1}{h} \int_{0}^{\pi} \sin(h\lambda) d\lambda
\]

\[
= \frac{2}{\pi^2 h^2} \left( 1 - \cos(h\pi) \right)
\]

\[
= \begin{cases} 
\frac{4}{\pi^2 h^2} & h \text{ odd} \\
0 & h \text{ even}
\end{cases}
\]

3. Filtered process: \(Y_t = X_t - 2\alpha X_{t-1} + X_{t-2}\) so try to filter \(\sin(\frac{2\pi}{6} t)\): \(\alpha\) satisfies:

\[
0 = \sin(\frac{2\pi}{6} t) - 2\alpha \sin(\frac{2\pi}{6} (t - 1)) + \sin(\frac{2\pi}{6} (t - 2)) = 0
\]

Use \(\sin(a - b) = \sin(a) \cos(b) - \cos(a) \sin(b)\) to get:

\[
0 = \sin(\frac{2\pi}{6} t) - 2\alpha \sin(\frac{2\pi}{6}) \cos(\frac{2\pi}{6} t) + 2\alpha \cos(\frac{2\pi}{6} t) \sin(\frac{2\pi}{6}) + \sin(\frac{2\pi}{6} t) \cos(\frac{4\pi}{6}) - \cos(\frac{2\pi}{6} t) \sin(\frac{4\pi}{6})
\]

\[
= \sin(\frac{\pi}{3} t) - \alpha \sin(\frac{\pi}{3} t) + \sqrt{3} \alpha \cos(\frac{\pi}{3} t) - \frac{1}{2} \sin(\frac{\pi}{3} t) - \sqrt{3} \cos(\frac{\pi}{3} t)
\]

\[
= \left( \frac{1}{2} - \alpha \right) \sin(\frac{\pi}{3} t) - \sqrt{3} \left( \frac{1}{2} - \alpha \right) \cos(\frac{\pi}{3} t).
\]

It follows that \(\alpha = \frac{1}{2}\).
Estimation of Mean, ACVF, ACF

1. 

\[ nV(X_n) = \frac{1}{n} \sum_{j=1}^{n} \sum_{k=1}^{n} C(X_j, X_k) = \frac{1}{n} \sum_{j=1}^{n} V(X_j) + \frac{2}{n} \sum_{j=1}^{n} \sum_{k=j+1}^{n} C(X_j, X_k) \]

\[ = \gamma(0) + \frac{2}{n} \sum_{j=1}^{n} \sum_{k=1}^{n-j} \gamma(k) = \gamma(0) + \frac{2}{n} \sum_{j=1}^{n} (n-j) \gamma(j) \]

\[ = \sum_{j=-n}^{n} \left( 1 - \frac{|j|}{n} \right) \gamma(j). \]

Absolute summability gives that for any \( \epsilon > 0 \) there exists an \( N(\epsilon) < +\infty \) such that

\[ \sup_n \left| \sum_{N+1 \leq |j| \leq (N+1) \wedge n} \left( 1 - \frac{|j|}{n} \right) \gamma(j) \right| < +\epsilon, \]

from which, for \( n > N(\epsilon) \),

\[ |nV(X_n)| = |nV(X_n) - \sum_{j=-\infty}^{\infty} \gamma(j)| \leq \epsilon + \sup_n \left| \sum_{j=-N}^{N} \frac{|j|}{n} \gamma(j) \right| \xrightarrow{n \to +\infty} \epsilon. \]

It follows that, for any \( \epsilon > 0 \), \( \lim_{n \to +\infty} |nV(X_n)| < \epsilon \) and hence \( \lim_{n \to +\infty} |nV(X_n)| < \epsilon = 0. \)

2. The first part clear: \( E \left[ \frac{1}{|Y|} \right] = \int_{|y|} f(y)dy = +\infty \) if \( f \) is continuous and \( f(0) \neq 0. \)

For the second part, use: \( X_n \sim AN(0, \frac{\sigma^2}{n}) \). Let \( v = \frac{\sigma^2}{n} \). Then, using \( Y = \frac{1}{X_n} \), it follows that for any \( 0 < y < +\infty, \)

\[ \lim_{n \to +\infty} \mathbb{P}(-y \leq Y \leq y) = \mathbb{P}(|X_n| \geq \frac{1}{y}) = \lim_{n \to +\infty} 2 \int_{1/y}^{\infty} \frac{\sqrt{n}}{\sqrt{2\pi} \sigma} \exp \left\{ -\frac{nx^2}{2\sigma^2} \right\} dx \]

\[ = \lim_{n \to +\infty} 2 \int_{\sqrt{1/y}}^{\infty} \frac{1}{\sqrt{2\pi} \sigma} \exp \left\{ -\frac{x^2}{2\sigma^2} \right\} dx = 0. \]

It follows that for any borel measurable \( A \subset \mathbb{R}, \mathbb{P}(Y \in A) \xrightarrow{n \to +\infty} 0. \)

3. The question is equivalent to showing that \( Y_n \sim AN \left( 0, \frac{\sigma^2}{\phi} \right) \). Asymptotic normality follows from Theorem 4.4: the hypotheses are that \( \sum_{j=-\infty}^{\infty} |\psi_j| < +\infty \) and \( \{Z_i\} \sim IID(0, \sigma^2) \) for \( \sigma^2 < +\infty \). For the process given, \( Y_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j} \) so that \( \sum |\psi_j| = \sum_{j=0}^{\infty} |\phi|^j = \frac{1}{1-|\phi|} < +\infty \); causal implies \( |\phi| < 1 \) where the inequality is strict. Therefore the hypotheses of the theorem are satisfied and hence the distribution of \( X_n \) is asymptotically normal. Clearly \( E[X_n] = \mu; \) it remains to compute the asymptotic variance.

\[ \lim_{n \to +\infty} nV(Y_n) = \sum_{j=-\infty}^{\infty} \gamma(j) = \frac{\sigma^2}{1-\phi^2} \left( 1 + 2 \sum_{j=1}^{\infty} \phi^j \right) = \frac{\sigma^2}{1-\phi^2} \left( 1 + \frac{2\phi}{1-\phi} \right) = \frac{\sigma^2}{(1-\phi)^2} \]
For the confidence interval, recall that a symmetric interval estimator is defined as two functions \( \mu_- \) and \( \mu_+ \) such that \( P_\mu(\mu_-(X) \geq \mu) = \frac{\alpha}{2} \) and \( P_\mu(\mu_+(X) \leq \mu) = \frac{\alpha}{2} \) and the symmetric confidence interval of confidence level \( 1 - \alpha \) or significance \( \alpha \) is: \( \mu \in (\mu_-(\bar{X}), \mu_+(\bar{X})) \), where \( \bar{X} \) is the vector of observed values. Here

\[
\mu \in \left( \bar{x}_n \pm \frac{\sigma}{(1 - \phi)\sqrt{n}} \times 1.96 \right) = (0.271 \pm \frac{\sqrt{2}}{0.4 \times 10} \times 1.96) = (-0.422, 0.963)
\]

0 is in the interval and is therefore a plausible value.

4. Linear, unbiased means \( \hat{\mu} = a_0 + \sum_{j=1}^{n} a_j X_j \), where, for all \( \mu \), \( a_0 + \mu \sum_{j=1}^{n} a_j = \mu \), hence \( a_0 = \mu(1 - \sum_{j=1}^{n} a_j) \) hence (since this identity holds for all \( \mu \)), \( a_0 = 0 \) and \( \sum_{j=1}^{n} a_j = 1 \). It follows that \( a_n = 1 - \sum_{j=1}^{n-1} a_j \), where \( a_1, \ldots, a_{n-1} \) are free variables. Then, to find minimum, take derivatives: let \( \bar{a} = (a_1, \ldots, a_n)^t \), then

\[
V(\hat{\mu}) = \bar{a}^t \bar{\Gamma} \bar{a}
\]

\[
\frac{\partial}{\partial a_j} V(\hat{\mu}) = 2\bar{a}^t (\Gamma_{j,j} - \sum_{j=n}^{1} a_j) = 0 \quad j = 1, \ldots, n-1
\]

From this,

\[
\bar{a}^t \Gamma_{1,1} = \bar{a}^t \Gamma_{2,2} = \ldots = \bar{a}^t \Gamma_{n,n} = c
\]

for some constant \( c \), from which

\[
\bar{a}^t \Gamma = c \bar{1}^t
\]

for a constant \( c \) to be determined, where \( \bar{1} \) is the \( n \)-vector with each entry 1. Hence

\[
\bar{a}^t = c \bar{1}^t \bar{\Gamma}^{-1}
\]

Since \( 1 = \sum_{j=1}^{n} a_j = \bar{a}^t \bar{1} \), it follows that \( 1 = \bar{a}^t \bar{1} = c \bar{1}^t \bar{\Gamma}^{-1} \bar{1} \) and hence

\[
\hat{\mu} = (\bar{1}^t \bar{\Gamma}^{-1} \bar{1})^{-1} \bar{1}^t \bar{\Gamma}^{-1} \bar{X}
\]

as required. The variance is:

\[
V(\hat{\mu}) = (\bar{1}^t \bar{\Gamma}^{-1} \bar{1})^{-1} \bar{1}^t \bar{\Gamma}^{-1} \bar{\Gamma}^{-1} \bar{1} (\bar{1}^t \bar{\Gamma}^{-1} \bar{1})^{-1} = (\bar{1}^t \bar{\Gamma}^{-1} \bar{1})^{-1}
\]

5.

\[
n\hat{\gamma}(h) = \sum_{j=1}^{n} (x_j - \bar{x})(x_{j+h} - \bar{x}) = \sum_{j=1}^{n} x_j x_{j+h} + (n - |h|)\bar{x}^2 - \bar{x} \sum_{j=1}^{n} (x_j + x_{j+h})
\]

72
\[
\sum_{h=-(n-1)}^{n-1} \gamma(h) = \gamma(0) + 2 \sum_{h=1}^{n-1} \gamma(h)
\]

\[
= \frac{n}{n} x_j^2 - n\bar{x}^2 + 2 \sum_{h=1}^{n-1} \sum_{j=1}^{n-h} x_j x_{j+h} + 2 \sum_{h=1}^{n-1} (n-h)\bar{x}^2 - 2\bar{x} \sum_{h=1}^{n-1} \sum_{j=1}^{n-h} (x_j + x_{j+h})
\]

\[
= \sum_{j=1}^{n} x_j^2 + \sum_{j=1}^{n-1} x_j \sum_{j=1}^{n-j} x_{j+h} + \sum_{j=2}^{n} x_j \sum_{h=1}^{j-1} x_{j-h} + \bar{x}^2 (-n + n(n-1))
\]

\[
-2\bar{x} \left( \sum_{j=1}^{n} (n-j)x_j + \sum_{j=2}^{n} (j-1)x_j \right)
\]

\[
= n^2\bar{x}^2 + \bar{x}^2(n^2 - 2n) - 2n(n-1)\bar{x}^2
\]

\[
= 0
\]

6. Firstly, \( \bar{X}_n \sim \text{AN}(\mu, \frac{\sigma^2}{(1-\phi)^2}) \). Now use delta method:

\[
f(\bar{X}_n) = f(\mu) + (\bar{X}_n - \mu)f'(x)* |x* - \mu| < |\bar{X}_n - \mu|.
\]

Then

\[
f(\bar{X}_n) \sim \text{AN}(f(\mu), f'(\mu)^2\text{V}(\bar{X}_n))
\]

so that

\[
\exp \{ \bar{X}_n \} \sim \text{AN}(e^\mu, e^{2\mu}\text{V}(\bar{X}_n)) = \text{AN} \left( e^\mu, \frac{1}{n} e^{2\mu} \frac{\sigma^2}{(1-\phi)^2} \right)
\]

hence \( b_n = e^\mu \) and \( a_n = \frac{1}{n} e^{2\mu} \frac{\sigma^2}{(1-\phi)^2} \).